The Horton-Strahler number of Galton-Watson trees with possibly infinite variance

Robin Khanfir

THE HORTON-STRAHLER NUMBER OF GALTON-WATSON TREES WITH POSSIBLY INFINITE VARIANCE

Robin KHANFIR*

July 10, 2023

Abstract

The Horton-Strahler number, also known as the register function, provides a tool for quantifying the branching complexity of a rooted tree. We consider the Horton-Strahler number of critical Galton-Watson trees conditioned to have size n and whose offspring distribution is in the domain of attraction of an α-stable law with $\alpha \in [1,2]$. We give tail estimates and when $\alpha \neq 1$, we prove that it grows as $\frac{1}{\alpha}\log_{\alpha/(\alpha-1)} n$ in probability. This extends the result in Brandenberger, Devroye & Reddad [6] dealing with the finite variance case for which $\alpha = 2$. We also characterize the cases where $\alpha = 1$, namely the spectrally positive Cauchy regime, which exhibits more complex behaviors.

Keywords Horton-Strahler number · Register function · Galton-Watson trees · tail estimates · stable laws

Mathematics Subject Classification 60C05 · 60J80 · 60F05 · 05C05 · 60E07

1 Introduction

The Horton-Strahler number of a finite rooted tree is an integer that quantifies its branching complexity. One possible formal definition is given recursively as follows.

Definition 1.1 Let t be a finite rooted tree. Its *Horton-Strahler number* $\mathcal{S}(t)$ is defined as follows.

(a) If t reduces to a single node, then $\mathcal{S}(t) = 0$.

(b) Otherwise, $\mathcal{S}(t)$ is the maximum of the Horton-Strahler numbers of the subtrees $t_1, ..., t_k$ that are attached to the root, plus one if that maximum is not uniquely achieved. Namely,

$$\mathcal{S}(t) = \max_{1\leq i\leq k} \mathcal{S}(t_i) + \mathbf{1}_{\left\{\#\operatorname{argmax}_{1\leq i\leq k}\mathcal{S}(t_i)\geq 2\right\}}.$$

□

Alternatively, $\mathcal{S}(t)$ is also the height of the largest perfect binary tree that can be embedded into t (see Section 3.1 for more details). In this article, we provide estimates on the Horton-Strahler number of critical Galton-Watson trees conditioned to be large. Before discussing our results precisely, let us provide a brief history with general references on related topics.

Background. The Horton-Strahler number was introduced independently by the two hydrogeologists Horton [21] and Strahler [33] to obtain quantitative empirical laws about river systems, that

*Sorbonne Université, Campus Pierre et Marie Curie, LPSM, Case courrier 158, 4, place Jussieu, 75252 Paris Cedex 05 France. Email: *robin.khanfir@sorbonne-universite.fr*

are represented by trees whose leaves are springs and whose root corresponds to the outlet of the basin. Many key physical characteristics of stream networks have been since modeled with the help of this number: see e.g. Peckham [30], Fac-Beneda [15], Chavan & Srinivas [8] and Bamufleh et al. [3]. The Horton-Strahler number appears independently in other scientific fields (anatomy, botany, molecular biology, physics, social network analysis, etc). In computer science, it is used to optimize the amount of memory or time needed to manipulate data structures. It is sometimes called the *register function* in this context because the minimum number of registers needed to evaluate an arithmetic expression A is equal to the Horton-Strahler number of the syntax tree of A. We refer to Viennot [35] for a survey on those various applications.

The Horton-Strahler number is encountered in many areas of mathematics: see for instance Esparza, Luttenberger & Schlund [14] for connections with mathematical logic, formal language theory, algebra, combinatorics, topology, approximation theory, and more. In the probability area, let us mention Kovchegov & Zaliapin [27], which considers the Horton-Strahler number through the prism of pruning operations on trees. Here, we rather focus on probabilistic works that discuss the Horton-Strahler number of uniform samples of standard families of combinatorial trees. Flajolet, Raoult & Vuillemin [17] and Kemp [22] consider the Horton-Strahler number of a uniform random ordered rooted binary tree T_n with n leaves (a uniform n-*Catalan tree*) and they prove that

$$\mathbb{E}[\mathcal{S}(T_n)] = \log_4 n + D(\log_4 n) + o(1)$$

as $n \to \infty$, where $\log_b x = \ln x / \ln b$ stands for the logarithm of x to the base b, and D is a 1-periodic continuous function. Here, all the random variables that we consider are defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$ whose expectation is denoted by $\mathbb{E}$. In particular, $\mathcal{S}(T_n)$ is subject to deterministic oscillations. Moreover, Devroye & Kruszewski [9] proved that $\mathcal{S}(T_n)$ is highly concentrated around its expected value via exponential tail estimates. These results were extended to k-ary trees by Drmota & Prodinger [10].

More recently, Brandenberger, Devroye & Reddad [6] showed that the Horton-Strahler number of a critical Galton-Watson tree with finite variance offspring distribution conditioned to have n vertices always grows as $\log_4 n$ in probability, which extends all the results that have been previously obtained on first-order behavior. In a companion paper [24], we go further and we study, among other things, the fluctuations and deterministic oscillations of the Horton-Strahler number of large Catalan trees.

Framework and main results. Let us give a precise overview of the present article which provides tail estimates and characterizes the first-order behavior of the Horton-Strahler number of critical Galton-Watson trees conditioned to have size n and whose offspring distribution is in the domain of attraction of an α-stable law with $\alpha \in [1, 2]$. This framework extends the finite variance case, where $\alpha = 2$, and it includes the so-called spectrally positive Cauchy (or 1-stable) laws. To that end, let us introduce our basic notations and assumptions.

Throughout this work, $\mu = (\mu(k))_{k\in\mathbb{N}}$ stands for a probability measure on the set $\mathbb{N}$ of nonnegative integers. We shall always assume μ to be *non-trivial* and *critical*, namely

$$\mu(0) > 0 \quad \text{and} \quad \sum_{k\in\mathbb{N}} k\mu(k) = 1. \tag{1}$$

We view it as the *critical offspring distribution* of a (rooted and ordered) Galton-Watson tree denoted by τ, which is then almost surely finite. See Section 2.1 for a formal definition. Several results are expressed in terms of the following functions

$$\forall s \in [0,1], \quad \varphi(s) = \sum_{k\in\mathbb{N}} s^k \mu(k) \quad \text{and} \quad \psi(s) = \varphi(1-s) - (1-s), \tag{2}$$

that are strictly convex thanks to (1).

Our first contribution consists in four propositions that connect $\mathcal{S}(\tau)$ to other simple characteristics of τ and that hold under the sole assumption that μ is non-trivial and critical. More precisely,

– Proposition 3.5 provides an upper bound for the size $\#\tau$ of τ when $\mathcal{S}(\tau)$ is small,
– Proposition 3.6 provides a lower bound for the height $|\tau|$ of τ when $\mathcal{S}(\tau)$ is large,
– Proposition 3.7 provides a lower bound for $\mathcal{S}(\tau)$ when the height $|\tau|$ of τ is large ,
– Proposition 3.9 provides a lower bound for $\mathcal{S}(\tau)$ when the maximal out-degree of τ is large.

We will use these propositions to derive the order of magnitude of $\mathcal{S}(\tau)$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$. These estimates are fairly general and they are sufficiently accurate to be interesting in their own right.

Our second goal is to estimate the tail of the Horton-Strahler number of Galton-Watson trees whose offspring distribution has possibly infinite variance. We work under the assumption that

(3) μ *belongs to the domain of attraction of a stable law.*

We denote the scaling index of (the type of) the limiting law by α: since μ has a finite mean, we get $\alpha \in [1, 2]$ and since μ is supported on $[0, \infty)$, we only deal with spectrally positive stable laws: namely, their *skewness parameter* β is equal to 1 and the support of their Lévy measure is included in $(0, \infty)$. By standard results, see e.g. [5, Theorem 8.3.1], Assumption (3) is equivalent to the existence of a function $L : [0, \infty) \mapsto (0, \infty)$ that is *slowly varying* and such that

$$\mu\left([n, \infty)\right) \sim n^{-\alpha} L(n) \text{ if } \alpha \in [1, 2) \quad \text{and} \quad \sum_{k=0}^{n} k^2 \mu(k) - 1 \sim 2L(n) \text{ if } \alpha = 2. \tag{4}$$

Note that $\sum_{k=0}^{n} k^2\mu(k) - 1$ is nondecreasing and ultimately positive since μ satisfies (1) (we refer to Proposition 2.9 for more details). We first discuss our results in the cases where $\alpha \in (1, 2]$, and then we consider the cases where $\alpha = 1$ that feature more complicated behaviors.

The cases where $\alpha \in (1, 2]$**.** In these cases, our results are simply expressed in terms of the index α only. First of all, we prove in Proposition 3.12 that if μ is in the domain of attraction of an α-stable law with $\alpha \in (1, 2]$, then the tail of $\mathcal{S}(\tau)$ follows the universal exponential decay

$$-\log_{\frac{\alpha}{\alpha-1}} \mathbb{P}(\mathcal{S}(\tau) \geq n) \sim n,$$

where $\log_b x = \ln x / \ln b$ stands for the logarithm of x to the base base b. This extends the work of Brandenberger, Devroye & Reddad [6] who proved the $\alpha = 2$ case under the assumption that μ has a finite variance.

We next discuss the behaviour of $\mathcal{S}(\tau)$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ and to that end, we assume that

(5) μ *is aperiodic*

(namely that μ is not supported by a proper additive subgroup of $\mathbb{Z}$) because this implies that $\mathbb{P}(\#\tau = n) > 0$ for all n that are large enough. Then, we prove the following.

Theorem 1.2 *Assume that μ is critical and aperiodic and that it belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. Then, the following convergence holds in probability.*

$$\frac{\alpha \mathcal{S}(\tau)}{\log_{\frac{\alpha}{\alpha-1}} n} \text{ under } \mathbb{P}(\,\cdot\,|\,\#\tau = n) \longrightarrow 1. \tag{6}$$

This extends the work of Brandenberger, Devroye & Reddad [6] who proved the $\alpha = 2$ case under the assumption that μ has a finite variance.

Our proof of Theorem 1.2 relies on results in Duquesne [11] on the height of α-stable trees and in Kortchemski [25] on the asymptotic behavior of the positive excursion of the random walk $(W_n)_{n\in\mathbb{N}}$. These results are recalled precisely in Section 2.2, Proposition 2.11. Let us mention that (6) holds true when μ is not aperiodic by restricting to the integers such that $\mathbb{P}(\#\tau = n) > 0$.

The 1-stable cases. In these more complex cases, we need to specify a converging sequence of rescaled centered sums of μ-distributed independent random variables. More precisely, we denote by $(W_n)_{n\in\mathbb{N}}$ a (left-continuous) random walk starting at $W_0 = 0$, whose jump law is given by

$$\mathbb{P}(W_1 = k) = \mu(k+1), \quad k \in \{-1\} \cup \mathbb{N}. \tag{7}$$

Note that $\mathbb{E}[W_1] = 0$. Then μ belongs to the domain of attraction of a 1-stable law if and only if there exists a $(0,\infty)$-valued sequence $(a_n)_{n\in\mathbb{N}}$ tending to ∞ such that

$$\frac{W_n + b_n}{a_n} \xrightarrow[n\to\infty]{\text{(law)}} X, \text{ where } b_n = n\mathbb{E}[W_1 \mathbf{1}_{\{|W_1|>a_n\}}] \text{ and } \mathbb{E}[e^{-\lambda X}] = e^{\lambda \ln \lambda} \tag{8}$$

for all $n \in \mathbb{N}$ and $\lambda \in (0,\infty)$. Let us mention here that $a_n \sim nL(a_n)$ and that $b_n/a_n \to \infty$, necessarily. The law of X is a *spectrally positive Cauchy (or 1-stable) law*. Its Fourier transform is given by $\mathbb{E}[\exp(iuX)] = \exp(-\frac{\pi}{2}|u| - iu\ln|u|)$, $u \in \mathbb{R}$. We refer to Proposition 2.12 for details.

The asymptotic behavior of $\mathcal{S}(\tau)$ in the 1-stable case is expressed in terms of the sequence $(b_n)_{n\in\mathbb{N}}$ and the following function Υ that is derived from ψ in (2) as follows.

$$\forall s \in (0,1), \quad \Upsilon(s) = \int_s^1 \frac{\mathrm{d}r}{r\ln\Lambda(r)}, \quad \text{where} \quad \Lambda(s) = \frac{s\psi'(s)}{s\psi'(s) - \psi(s)}. \tag{9}$$

We refer to Section 3.3 for more details on the definition of Υ. Proposition 3.12 asserts that if μ is in the domain of attraction of a 1-stable law, then the tail of $\mathcal{S}(\tau)$ satisfies

$$\Upsilon\Big(\mathbb{P}\left(\mathcal{S}(\tau) \geq n\right)\Big) \sim n. \tag{10}$$

In contrast to the cases where $\alpha \in (1,2]$ for which we can show that $\Upsilon(s) \sim_{0^+} \log_{\frac{\alpha}{\alpha-1}} 1/s$, the asymptotic behavior of Υ when $\alpha = 1$ depends on the slowly varying function L appearing in (4): see Proposition 3.11 for a precise statement. As discussed in Examples 3.13, the following holds.

- (**a**) If $L(n) \sim (\ln n)^{-1-\kappa}$ with $\kappa \in (0,\infty)$, then $\Upsilon(s) \sim_{0^+} \frac{\ln 1/s}{\ln\ln 1/s}$ and $-\ln\mathbb{P}(\mathcal{S}(\tau)\geq n) \sim n\ln n$.
- (**b**) If $L(n) \sim \exp(-(\ln n)^\kappa)$ with $\kappa \in (0,1)$, then $\Upsilon(s) \sim_{0^+} \frac{1}{1-\kappa}\frac{\ln 1/s}{\ln\ln 1/s}$ and $-\ln\mathbb{P}(\mathcal{S}(\tau)\geq n) \sim (1-\kappa)n\ln n$.
- (**c**) If $L(n) \sim \exp(-\ln n/\ln\ln n)$ then $\Upsilon(s) \sim_{0^+} \frac{\ln 1/s}{\ln\ln\ln 1/s}$ and $-\ln\mathbb{P}(\mathcal{S}(\tau)\geq n) \sim n\ln\ln n$.

When the Galton-Watson tree τ is conditioned to be large, the size of its Horton-Strahler number is of order $\Upsilon(\frac{1}{b_n})$ as proved by the following theorem that first handles the case where τ is conditioned to have at least n vertices.

Theorem 1.3 *Assume that μ is critical and that it belongs to the domain of attraction of a 1-stable law. Then, the following convergence holds in probability.*

$$\frac{\mathcal{S}(\tau)}{\Upsilon(\frac{1}{b_n})} \text{ under } \mathbb{P}(\,\cdot\,|\,\#\tau \geq n) \longrightarrow 1, \tag{11}$$

where Υ is given by (9) and $(b_n)_{n\in\mathbb{N}}$ by (8).

Our proof of Theorem 1.3 and of Theorem 1.4 below, rely on several results of Kortchemski & Richier [26] and of Berger [4] that specify the asymptotic behavior of positive excursion of the random walk $(W_n)_{n\in\mathbb{N}}$. They are recalled precisely in Section 2.2, Lemma 2.13 and Propositions 2.15 and 2.16.

As discussed by Kortchemski & Richier [26] and Berger [4], it is not clear how one could control τ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ by assuming only that $\mu([n,\infty)) \sim L(n)/n$ as in (4). Here, we work under the stronger assumption that $\mu(n) \sim L(n)/n^2$, which implies the previous one and also implies that μ is aperiodic.

Theorem 1.4 *Assume that μ is critical and that there is a slowly varying function L such that*

$$\mu(n) \sim \frac{L(n)}{n^2}. \tag{12}$$

Then, the following convergence holds in probability.

$$\frac{\mathcal{S}(\tau)}{\Upsilon(\frac{1}{b_n})} \text{ under } \mathbb{P}(\,\cdot\,|\,\#\tau = n) \longrightarrow 1, \tag{13}$$

where Υ is given by (9) and $(b_n)_{n\in\mathbb{N}}$ by (8).

As already mentioned, when $\alpha = 1$, the rescaling sequence $\Upsilon(\frac{1}{b_n})$ depends on the slowly varying function L appearing in (4): as discussed in Examples 4.2, the following holds true.

- $(\mathbf{a})$ If $L(n)\sim(\ln n)^{-1-\kappa}$ with $\kappa\in(0,\infty)$, then $\Upsilon(\frac{1}{b_n})\sim\frac{\ln n}{\ln\ln n}$.
- $(\mathbf{b})$ If $L(n)\sim\exp(-(\ln n)^{\kappa})$ with $\kappa\in(0,1)$, then $\Upsilon(\frac{1}{b_n})\sim\frac{1}{1-\kappa}\frac{\ln n}{\ln\ln n}$.
- $(\mathbf{c})$ If $L(n)\sim\exp(-\ln n/\ln\ln n)$, then $\Upsilon(\frac{1}{b_n})\sim\frac{\ln n}{\ln\ln\ln n}$.

Organisation of paper. In Section 2, we properly set our framework and we recall from previous works the tools that we use later on in the paper: Section 2.1 is devoted to Galton-Watson trees and Section 2.2 to known limit theorems for random walks and Galton-Watson trees. In Section 3, we study the distribution of the Horton-Strahler number of Galton-Watson trees. In Section 3.1, we first establish new technical results on Horton-Strahler numbers (especially in Lemmas 3.1 and 3.2). Section 3.2 focuses on proving Propositions 3.5, 3.6, 3.7, and 3.9 that link the Horton-Strahler number to size, height, and maximal out-degree of Galton-Watson trees. Section 3.3 is devoted to the tail asymptotics of Galton-Watson trees whose offspring distribution belongs to the domain of attraction of a stable law: in particular, we prove Proposition 3.12, that is one of the main results of the paper and we discuss Examples 3.13 $(\mathbf{a})$, $(\mathbf{b})$, and $(\mathbf{c})$. Section 4 is devoted to the proof of Theorems 1.2, 1.3 and 1.4 and we discuss Examples 4.2 $(\mathbf{a})$, $(\mathbf{b})$, and $(\mathbf{c})$, in the end of Section 4.3.

2 Framework and tools

In this section, we recall a set of well-known results that are used in the rest of the article and in the proofs of Theorems 1.2, 1.3 and 1.4. With the exception of Lemma 2.13, this section contains no new result.

2.1 Galton-Watson trees

Rooted ordered trees. We recall Ulam's formalism on rooted ordered trees. Let $\mathbb{N}^* = \{1, 2, 3, ...\}$ be the set of positive integers and let $\mathbb{U}$ be the following set of finite words

$$\mathbb{U} = \bigcup_{n\in\mathbb{N}} (\mathbb{N}^*)^n$$

with the convention $(\mathbb{N}^*)^0 = \{\varnothing\}$. The set of words $\mathbb{U}$ is totally ordered by the *lexicographic order* denoted by $\leq$. Let $u = (u_1, ..., u_n) \in \mathbb{U}$ and $v = (v_1, ..., v_m) \in \mathbb{U}$, we write $u * v = (u_1, ..., u_n, v_1, ..., v_m) \in \mathbb{U}$ for the *concatenation* of u and v. We denote by $|u| = n$ the *height* of u, and if $n \geq 1$ then we denote by $\overleftarrow{u} = (u_1, ..., u_{n-1})$ the *parent* of u. We also say that u is a *child* of v when $\overleftarrow{u} = v$. The *genealogical order* $\preceq$ is a partial order on $\mathbb{U}$ defined by $u \preceq v \iff \exists u' \in \mathbb{U},\ v = u * u'$. When $u \preceq v$, we will say that u is an ancestor of v. When $u \preceq v$ but $u \neq v$, we write $u \prec v$. Finally, we write $u \wedge v \in \mathbb{U}$ for the *most recent common ancestor* of u and v, that is their common ancestor with maximal height.

Definition 2.1 We say that a subset t of $\mathbb{U}$ is a *tree* when the following is satisfied:

(a) $\varnothing \in t$,

(b) for all $u \in t$, if $u \neq \varnothing$ then $\overleftarrow{u} \in t$,

(c) For all $u \in t$, there exists an integer $k_u(t) \in \mathbb{N}$ such that $u*(i) \in t \iff 1 \leq i \leq k_u(t)$.

We denote by $\mathbb{T}$ the space of all trees. □

Let $t \in \mathbb{T}$. The size of t is simply the (possibly infinite) number $\#t$ of its vertices and we say that t is finite if $\#t < \infty$. As a graph, the *edges* of t are given by the unordered pairs $\{u, \overleftarrow{u}\}$ for all $u \in t\backslash\{\varnothing\}$. Therefore the *degree* of $u \in t$ is $k_u(t) + 1$ if $u \neq \varnothing$ and $k_\varnothing(t)$ otherwise. Namely, $k_u(t)$ is the *out-degree of* u (alternatively, if one views t as a family tree whose ancestor is $\varnothing$, then $k_u(t)$ stands for the number of children of u). We use the following notations for the *height of* t and its *maximal out-degree.*

$$|t| = \max_{u \in t} |u| \quad \text{and} \quad \Delta(t) = \max_{u \in t} k_u(t). \tag{14}$$

We also denote the subtree stemming from $u \in t$ and the tree pruned at u respectively by

$$\theta_u t = \{v \in \mathbb{U} \,:\, u * v \in t\} \quad \text{and} \quad \mathrm{Cut}_u\, t = t\backslash\{v \in t \,:\, u \prec v\} \tag{15}$$

Observe that $\theta_u t$ and $\mathrm{Cut}_u\, t$ both belong to $\mathbb{T}$.

Galton-Watson trees and the Many-To-One Principle. Let us equip the set of trees $\mathbb{T}$ with the sigma-field $\mathscr{F}(\mathbb{T})$ generated by the sets $\{t \in \mathbb{T} : u \in t\}$, where u ranges in $\mathbb{U}$. Formally, a *random tree* is a function $\tau : \Omega \to \mathbb{T}$ that is $(\mathscr{F}, \mathscr{F}(\mathbb{T}))$-measurable.

Definition 2.2 Let $\mu = (\mu(k))_{k \in \mathbb{N}}$ be a probability measure on $\mathbb{N}$. A *Galton-Watson tree with offspring distribution* μ (a $\mathsf{GW}(\mu)$*-tree*, for short) is a random tree τ that satisfies the following.

(a) $k_\varnothing(\tau)$ has law μ.

(b) For all $k \in \mathbb{N}^*$ such that $\mu(k) > 0$, the subtrees $\theta_{(1)}\tau, \ldots, \theta_{(k)}\tau$ under $\mathbb{P}(\,\cdot \,|\, k_\varnothing(\tau) = k)$ are independent with the same law as τ under $\mathbb{P}$. □

It is well-known that a $\mathsf{GW}(\mu)$-tree τ is almost surely finite if and only if its offspring distribution is subcritical or critical and non-trivial, namely if (1) holds, and in that case for all finite tree $t \in \mathbb{T}$,

$$\mathbb{P}(\tau = t) = \prod_{u \in t} \mu\big(k_u(t)\big). \tag{16}$$

As observed by Kesten [23], a critical $\mathsf{GW}(\mu)$-tree conditioned to be large locally converges in law to a tree τ_∞ with a single infinite line of descent and whose law can be informally described as follows: all individuals of τ_∞ reproduce independently, the individuals of the infinite line of descent reproduce according to the μ-size-biased distribution $(k\mu(k))_{k \in \mathbb{N}}$ whereas the others reproduce according to μ. More precisely, we introduce the following.

Definition 2.3 Let μ be a non-trivial critical offspring distribution. A *size-biased* $\mathsf{GW}(\mu)$*-tree* is a random tree τ_∞ that satisfies the following.

(a) For all $n \in \mathbb{N}$, there is a unique $u \in \tau_\infty$ such that $|u| = n$ and $\#(\theta_u \tau_\infty) = \infty$. We denote this vertex by U_n. Note that $U_0 = \varnothing$ and that $\overleftarrow{U_{n+1}} = U_n$. Hence, there exists a $\mathbb{N}^*$-valued sequence of random variables $(J_n)_{n \in \mathbb{N}^*}$ such that U_n is the word $(J_1, \ldots, J_n)$ for all $n \in \mathbb{N}^*$.

(b) The random variables $(k_{U_n}(\tau_\infty), J_{n+1})$, for $n \in \mathbb{N}$, are independent and distributed as follows:

$$\forall j, k \in \mathbb{N}^*, \quad \mathbb{P}(J_{n+1} = j \,;\, k_{U_n}(\tau_\infty) = k) = \mathbf{1}_{\{j \leq k\}} \mu(k). \tag{17}$$

(c) Conditionally given $(k_{U_n}(\tau_\infty), J_{n+1})_{n\in\mathbb{N}}$, the finite subtrees stemming from the infinite line of descent, which are the $\theta_{U_n*(j)}\tau_\infty$ for $j \in \{1,\dots,k_{U_n}(\tau_\infty)\}\backslash\{J_{n+1}\}$ and $n\in\mathbb{N}$, are independent $\mathsf{GW}(\mu)$-trees. □

Remark 2.4 Note that the individual U_{n+1} on the infinite line of descent has $J_{n+1}-1$ siblings strictly on the left hand side and $k_{U_n}(\tau_\infty)-J_{n+1}$ siblings strictly on the right hand side. Their joint law is given by the following bivariate generating function:

$$\forall s,r\in[0,1],\quad \mathbb{E}\big[r^{J_{n+1}-1}s^{k_{U_n}(\tau_\infty)-J_{n+1}}\big]=\frac{\varphi(r)-\varphi(s)}{r-s}, \tag{18}$$

where φ is given by (2) and where the quotient is equal to $\varphi'(r)$ when $r=s$. □

As mentioned above, size-biased trees are the local limits of critical Galton-Watson trees conditioned to be large and therefore appear in many results concerning the asymptotic behavior of branching processes: we refer to Lyons, Pemantle & Peres [29], Aldous & Pitman [2] and Abraham & Delmas [1] for general results in this vein. One key tool involving size-biased $\mathsf{GW}(\mu)$-trees is the so-called *Many-To-One Principle*, which is part of folklore (see e.g. Duquesne [12, Equation (24)] for a proof) and which we use in our article in the following form.

Proposition 2.5 (Many-To-One Principle) *Let τ be a $\mathsf{GW}(\mu)$-tree and let τ_∞ be a size-biased $\mathsf{GW}(\mu)$-tree. We keep the notations of Definition 2.3. Then, for all $n\in\mathbb{N}$ and for all bounded functions $G_1:\mathbb{T}\times\mathbb{U}\longrightarrow\mathbb{R}$ and $G_2:\mathbb{T}\longrightarrow\mathbb{R}$, it holds that*

$$\mathbb{E}\left[\sum_{u\in\tau}\mathbf{1}_{\{|u|=n\}}G_1(\mathrm{Cut}_u\,\tau,u)\;G_2(\theta_u\tau)\right]=\mathbb{E}\big[G_1(\mathrm{Cut}_{U_n}\,\tau_\infty,U_n)\big]\,\mathbb{E}\big[G_2(\tau)\big].$$

Lukasiewicz path associated with a tree. We recall here a key combinatorial tool to study Galton-Watson trees via random walks.

Definition 2.6 Let $t\in\mathbb{T}$ be *finite* and let $u(t)=(u_j(t))_{0\le j<\#t}$ be the sequence of its vertices listed in increasing lexicographic order: $u_0(t)=\varnothing<u_1(t)<\ldots<u_j(t)<u_{j+1}(t)<\ldots<u_{\#t-1}(t)$. The sequence $u(t)$ is often called the *depth-first exploration* of t. We then define a $\mathbb{Z}$-valued path $W(t)=(W_j(t))_{0\le j\le\#t}$ by setting $W_0(t)=0$ and

$$W_{j+1}(t)=W_j(t)+k_{u_j(t)}(t)-1 \tag{19}$$

for all $0\le j<\#t$, that is the *Lukasiewicz path* of t. □

In probability, Lukasiewicz paths originate from queuing systems theory to study the waiting line of a single server subject to the Last-In-First-Out policy and they have been used by Le Gall & Le Jan [18] to define Lévy trees. In the following lemma, we recall that Lukasiewicz paths are adapted processes that completely encode finite trees and that are particularly well-suited to study Galton-Watson trees (see e.g. Le Gall [19, Proposition 1.1] and [19, Corollary 1.6] for more details).

Proposition 2.7 *Let $t\in\mathbb{T}$ be finite. Let τ be a $\mathsf{GW}(\mu)$-tree where μ is non-trivial and critical. Then the following holds true.*

(i) *Lukasiewicz paths provide a one-to-one correspondence between the set of finite trees and the set of finite nonnegative excursions of left-continuous walks, which is defined by*

$$\bigcup_{n\in\mathbb{N}^*}\big\{(w_j)_{0\le j\le n}\in\mathbb{Z}^n\;:\;w_0=0,\;w_n=-1,\;w_j\ge 0\text{ and }w_{j+1}-w_j\ge -1\text{ for all }0\le j<n\big\}.$$

(*ii*) *For all $m \in \mathbb{N}$, we denote by $R_m t$ the tree t restricted to its $m+1$ first vertices (with respect to the lexicographic order). Namely,*

$$R_m t = \{u_j(t) \ : \ 0 \leq j \leq \min(m, \#t - 1)\} \tag{20}$$

where $(u_j(t))$ stands for the vertices of t listed in lexicographic order. Then, $R_m t$ is a measurable function of $\big(W_j(t) \ ; \ 0 \leq j \leq \min(m, \#t)\big)$.

(*iii*) *Let $(W_n)_{n\geq 0}$ be a $\mathbb{Z}$-valued random walk whose jump distribution is given by (7). We set $\mathtt{H}_1 = \inf\{j \in \mathbb{N} \ : \ W_j = -1\}$, which is an a.s. finite stopping time since $\mathbb{E}[W_1] = 0$. Then,*

$$\big(W_j(\tau)\big)_{0\leq j\leq \#\tau} \overset{(law)}{=} \big(W_j\big)_{0\leq j\leq \mathtt{H}_1}.$$

In particular, $\#\tau$ and $\mathtt{H}_1$ have the same law.

(*iv*) *More generally, for all $p \in \mathbb{N}$, we set*

$$\mathtt{H}_0 = 0 \quad \textit{and} \quad \mathtt{H}_p = \inf\{j \in \mathbb{N} \ : \ W_j = -p\} \ . \tag{21}$$

Then there is an i.i.d. sequence $(\tau_p)_{p\in\mathbb{N}}$ of $\mathsf{GW}(\mu)$-trees such that

$$\forall p \in \mathbb{N}, \quad \big(p + W_{\mathtt{H}_p + j}\big)_{0\leq j\leq \mathtt{H}_{p+1}-\mathtt{H}_p} = \big(W_j(\tau_p)\big)_{0\leq j\leq \#\tau_p} \ . \tag{22}$$

2.2 Limit theorems.

In this section we recall – mostly from Bingham, Goldies & Teugels [5] and Feller [16] – limit theorems for sums of i.i.d. random variables belonging to the domain of attraction of stable laws. We also recall useful limit theorems for Galton-Watson trees and random walks from Berger [4], Duquesne [11], Kortchemski [25], and Kortchemski & Richier [26].

Regularly and slowly varying functions. Recall that a measurable and locally bounded function $l : (0,\infty) \to (0,\infty)$ is *slowly varying* at infinity (resp. at 0^+) if $l(cx) \sim l(x)$ as $x \to \infty$ (resp. as $x \to 0^+$) for all $c \in (0,\infty)$. Also recall that $f : (0,\infty) \to (0,\infty)$ is *regularly varying of index* $\alpha \in \mathbb{R}$ at ∞ (resp. at 0^+) if there exists a slowly varying function l at ∞ (resp. at 0^+) such that $f(x) = x^\alpha l(x)$. Below we gather in a single proposition several well-known results on slowly and regularly varying functions that are used in this article.

Proposition 2.8 *Let l be a slowly varying function at ∞. Then the following holds true.*

(*i*) (Potter's bound) *For all $\varepsilon \in (0,\infty)$ and all $c \in (1,\infty)$, there exists $x_0 \in (0,\infty)$ such that for all $x \in [x_0,\infty)$ and all $\lambda \in [1,\infty)$, it holds*

$$\frac{1}{c}\lambda^{-\varepsilon} \leq \frac{l(x\lambda)}{l(x)} \leq c\lambda^{\varepsilon}.$$

Therefore, $\ln l(x) = o(\ln x)$ and if f is regularly varying with index $\rho \in \mathbb{R}\backslash\{0\}$, then $\ln f(x) \sim \rho \ln x$. Moreover, if $x_n \sim y_n \to \infty$ then $l(x_n) \sim l(y_n)$ and $f(x_n) \sim f(y_n)$.

(*ii*) (Karamata's Abelian Theorem for Tails) *Let $\rho \in (0,\infty)$. Then $\int^\infty y^{-1-\rho} l(y)\,\mathrm{d}y < \infty$ and, as $x \to \infty$,*

$$\int_x^\infty y^{-1-\rho} l(y)\,\mathrm{d}y \ \sim \ \tfrac{1}{\rho} x^{-\rho} l(x) \quad \textit{and} \quad \int_1^x y^{\rho-1} l(y)\,\mathrm{d}y \ \sim \ \tfrac{1}{\rho} x^{\rho} l(x).$$

(*iii*) *Suppose that $\int^\infty y^{-1} l(y)\,\mathrm{d}y < \infty$ and set $\overline{l}(x) = \int_x^\infty y^{-1} l(y)\,\mathrm{d}y$. Then $\overline{l}$ is slowly varying at ∞ and $\lim_{x\to\infty} \overline{l}(x)/l(x) = \infty$.*

(iv) (Monotone Density Theorem) *Let $u:[0,\infty)\to\mathbb{R}$ be a locally Lebesgue integrable function and set $U(x)=\int_0^x u(y)\,\mathrm{d}y$ for all $x\in\mathbb{R}$. Assume that there are $c,\rho\in(0,\infty)$ such that $U(x)\sim cx^\rho l(x)$ as $x\to\infty$ and furthermore assume that u is ultimately monotone. Then $u(x)\sim c\rho x^{\rho-1}l(x)$ as $x\to\infty$.*

(v) (Karamata's Abelian Theorem for Laplace Transform) *Let $U:[0,\infty)\to[0,\infty)$ be a measurable function such that $\widehat{U}(\lambda):=\lambda\int_0^\infty e^{-\lambda x}U(x)\,\mathrm{d}x<\infty$ for all $\lambda\in(0,\infty)$. Assume that there are $c\in(0,\infty)$ and $\rho\in(-1,\infty)$ such that $U(x)\sim\frac{c}{\Gamma(1+\rho)}x^\rho l(x)$ as $x\to\infty$, then $\widehat{U}(\lambda)\sim c\lambda^{-\rho}\,l(1/\lambda)$ as $\lambda\to 0^+$.*

Proof. For (i), see e.g. [5, Theorem 1.5.6]. For (ii), see e.g. [5, Theorem 1.5.8 and Proposition 1.5.10]. For (iii), see e.g. [5, Theorem 1.5.9b]. For (iv), see e.g. [5, Theorem 1.7.2]. For (v), see e.g. [5, Theorem 1.7.6]. ■

Limit theorems: the cases where $\alpha\in(1,2]$. We next recall equivalent formulations of the property for a probability measure μ on $\mathbb{N}$ to belong to the domain of attraction of a stable law of index $\alpha\in(1,2]$ (we handle 1-stable laws separately).

Proposition 2.9 *Let μ be a probability measure on $\mathbb{N}$ that satisfies (1). Recall from (2) the definition of ψ. Let $(W_n)_{n\in\mathbb{N}}$ be a random walk whose jumps distribution is specified in (7). Let $\alpha\in(1,2]$. Then, the following assertions are equivalent.*

(a) *μ belongs to the domain of attraction of an α-stable law.*

(b) *There exists $L:(0,\infty)\to(0,\infty)$ that varies slowly at ∞ such that if $\alpha\in(1,2)$ then $\mu([n,\infty))\sim n^{-\alpha}L(n)$, and if $\alpha=2$ then $\sum_{0\le k\le n}k^2\mu(k)-1\sim 2L(n)$ which is ultimately positive by (1).*

(c) *If $\alpha\in(1,2)$ then $\psi(s)\sim_{0^+}\frac{\alpha-1}{\Gamma(2-\alpha)}s^\alpha L(1/s)$, and if $\alpha=2$ then $\psi(s)\sim_{0^+}s^2L(1/s)$.*

(d) *There exists a $(0,\infty)$-valued sequence $(a_n)_{n\in\mathbb{N}}$ tending to ∞ such that $\frac{1}{a_n}W_n$ converges in law to the spectrally positive α-stable random variable X_α whose Laplace exponent is given for all $\lambda\in[0,\infty)$ by $\ln\mathbb{E}[\exp(-\lambda X_\alpha)]=\frac{\alpha-1}{\Gamma(2-\alpha)}\lambda^\alpha$ if $\alpha\in(1,2)$ and by λ^2 if $\alpha=2$.*

Moreover, if one of the four equivalent assumptions from above holds true, then $a_n^\alpha\sim nL(a_n)$ and $a_n\sim n^{1/\alpha}L^*(n)$ where the function $L^*:x\in(0,\infty)\mapsto x^{-1/\alpha}\inf\{y\in(0,\infty):y^\alpha/L(y)>x\}$ is slowly varying at ∞.

Proof. For $(a)\Leftrightarrow(b)$, see e.g. [5, Theorem 8.3.1]. The equivalence $(a)\Leftrightarrow(d)$ follows from the definition of the domain of attraction of a stable law: the limiting law is necessarily spectrally positive since μ is supported by $\mathbb{N}$ and among the spectrally positive α-stable types, it is always possible to choose a centered one as $\alpha\in(1,2]$ (see e.g [5, Section 8.3] and [16, Chapter XVII.5]).

For $(b)\Leftrightarrow(c)$ see e.g. [5, Theorem 8.1.6]. More precisely, recall from (2) the definitions of φ and ψ. Then, for all $\lambda\in(0,\infty)$, set

$$f_1(\lambda):=\varphi(e^{-\lambda})-1+\lambda=\psi(1-e^{-\lambda})+\tfrac{1}{2}\lambda^2+\mathrm{O}(\lambda^3). \tag{23}$$

If $\alpha\in(1,2)$ then [5, Theorem 8.1.6] asserts that (b) is equivalent to $f_1(\lambda)\sim_{0^+}\frac{\alpha-1}{\Gamma(2-\alpha)}\lambda^\alpha L(1/\lambda)$, which implies that $(b)\Leftrightarrow(c)$ in these cases (with Proposition 2.8 (i)). If $\alpha=2$ then [5, Theorem 8.1.6] asserts that (b) is equivalent to $f_1(\lambda)\sim_{0^+}\lambda^2(\frac{1}{2}+L(1/\lambda))$ and (23) implies that $(b)\Leftrightarrow(c)$.

Let us prove the last point of the proposition. We assume that $(a$-$d)$ hold true. By Grimvall [20, Theorem 2.1], $\lim_{n\to\infty}\mathbb{E}[\exp(-\frac{\lambda}{a_n}W_n)]=\mathbb{E}[e^{-\lambda X_\alpha}]$ for all $\lambda\in(0,\infty)$. Observe that $\mathbb{E}[\exp(-\lambda W_n)]=(e^\lambda(f_1(\lambda)+1-\lambda))^n$. If $\alpha\in(1,2)$, we easily get $nf_1(\lambda/a_n)\sim\frac{\alpha-1}{\Gamma(2-\alpha)}\lambda^\alpha$ as $n\to\infty$ for all $\lambda\in(0,\infty)$ and thus $a_n^\alpha\sim nL(a_n)$. If $\alpha=2$, we get

$$\tfrac{n\lambda}{a_n}+n\ln\left[1-\tfrac{\lambda}{a_n}\left(1-\tfrac{\lambda}{a_n}\left(\tfrac{1}{2}+L(a_n)\right)(1+o(1))\right)\right]\sim\tfrac{nL(a_n)}{a_n^2}\lambda^2,$$

which implies that $a_n^2 \sim nL(a_n)$. In both cases, we observe that $a_n^\alpha / L(a_n) \sim n$ and we use [5, Theorem 1.5.12] to complete the proof of the proposition. ■

We next recall standard results on the size of Galton-Watson trees, which are expressed in terms of random walks thanks to Proposition 2.7.

Proposition 2.10 *Let μ be a probability measure on $\mathbb{N}$ that satisfies (1). Let $(W_n)_{n\in\mathbb{N}}$ be a random walk whose jumps distribution is specified by (7). Recall from (21) the definition of the stopping times $(\mathtt{H}_p)_{p\in\mathbb{N}}$. Let τ be a $\mathsf{GW}(\mu)$-tree. Then the following holds true.*

- (i) *$(\mathtt{H}_p)_{p\in\mathbb{N}}$ is a random walk with positive jumps.*
- (ii) (Kemperman) $\mathbb{P}(\mathtt{H}_p = n) = \frac{p}{n}\mathbb{P}(W_n = -p)$ *for all $n, p \in \mathbb{N}^*$.*
- (iii) *Suppose that μ satisfies Proposition 2.9 $(a$-$d)$ and is aperiodic. Then, $\mathbb{P}(\#\tau = n) \sim \frac{c_\alpha}{na_n}$ and $\mathbb{P}(\#\tau \geq n) \sim \frac{\alpha c_\alpha}{a_n}$ where c_α is the value at 0 of the (continuous version) of the density of X_α, and where (a_n) and X_α are as in Proposition 2.9 (d).*
- (iv) *Suppose that μ satisfies Proposition 2.9 $(a$-$d)$. Recall that $|\tau|$ stands for the height of τ. Then,*

$$\frac{\mathbb{P}(|\tau| \geq n)}{\psi\left(\mathbb{P}(|\tau| \geq n)\right)} \sim (\alpha - 1)n.$$

Proof. Note that (i) is an immediate consequence of the Markov property and the left-continuity of W. For (ii), see e.g. [5, Theorem 8.9.15]. Let us prove (iii). By (ii) and by Proposition 2.7 (iii), we see that $\mathbb{P}(\#\tau = n) = \frac{1}{n}\mathbb{P}(W_n = -1)$. We next use Gnedenko's local limit Theorem (see e.g. [5, Theorem 8.4.1]) to get $\lim_{n\to\infty} |a_n\mathbb{P}(W_n = -1) - c_\alpha| = 0$ and thus $\mathbb{P}(\#\tau = n) \sim \frac{c_\alpha}{na_n}$. Since (a_n) varies regularly with index $1/\alpha$ by Proposition 2.9, we get $\mathbb{P}(\#\tau \geq n) \sim \frac{\alpha c_\alpha}{a_n}$ by Karamata's Abelian Theorem for Tails (see Proposition 2.8 (ii)). For (iv), see Slack [32, Lemma 2]. ■

We next recall two limit theorems that are used to prove Theorem 1.2. One follows from the convergence of rescaled $\mathsf{GW}(\mu)$-trees to stable trees due to Duquesne [11]. The other is the uniform integrability of the density of the law of (roughly speaking) the Lukasiewicz path of a $\mathsf{GW}(\mu)$-tree τ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ with respect to $\mathbb{P}(\,\cdot\,|\,\#\tau \geq n)$ that has been proved in Kortchemski [25].

Proposition 2.11 *Assume that μ satisfies (1), that it is aperiodic, and that it belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. More precisely, we assume that Proposition 2.9 (d) holds true. Let τ be a $\mathsf{GW}(\mu)$-tree. Recall from (14) that $|\tau|$ stands for the maximal height of τ, and that $W(\tau)$ stands for its Lukasiewicz path as in Definition 2.6. Then, the following holds true.*

- (i) *There exists a random variable $M \in (0, \infty)$ such that $\frac{a_n}{n}|\tau|$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ converges in law to M as $n \to \infty$.*
- (ii) *Let $r \in (0, 1)$. Then for all large enough $n \in \mathbb{N}$, there is a function $D_n^{(r)} : \mathbb{N} \to [0, \infty)$ such that*

$$\mathbb{E}\left[f\left(W_{\min(\lfloor nr\rfloor,\cdot)}(\tau)\right) \,\middle|\, \#\tau = n\right] = \mathbb{E}\left[f\left(W_{\min(\lfloor nr\rfloor,\cdot)}(\tau)\right) D_n^{(r)}\left(W_{\lfloor rn\rfloor}(\tau)\right) \,\middle|\, \#\tau \geq n\right] \tag{24}$$

for all bounded function $f : \mathbb{N}^{\mathbb{N}} \to [0, \infty)$. Moreover, these functions satisfy

$$\lim_{c\to\infty} \limsup_{n\to\infty} \mathbb{E}\left[\mathbf{1}_{\{D_n^{(r)}(W_{\lfloor rn\rfloor}(\tau))\geq c\}} D_n^{(r)}\left(W_{\lfloor rn\rfloor}(\tau)\right) \,\middle|\, \#\tau \geq n\right] = 0. \tag{25}$$

Proof. The point (i) is a consequence of the convergence of rescaled Galton Watson trees to the α-stable tree: see Duquesne [11, Theorem 3.1]. Here, M is the height of the normalized α-stable tree that is a $(0, \infty)$-valued random variable. See Kortchemski [25, Equation (1) and Lemma 2] for (24). For (25), [25, Equation (12)] shows that $D_n^{(r)}(a_n\,\cdot\,)$ uniformly converges on all compact intervals of $(0, \infty)$ towards a continous function. Moreover, the laws of $a_n^{-1}W_{\lfloor rn\rfloor}(\tau)$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ are tight in $(0, \infty)$ (see [25, Equation (17)]), which completes the proof of (25) by (24). ■

Limit theorems: the 1-stable case. We now consider the spectrally positive 1-stable law, which features more complicated behaviors.

Proposition 2.12 *Let μ be a probability measure on $\mathbb{N}$ that satisfies (1). Recall from (2) the definition of ψ. Let $(W_n)_{n\in\mathbb{N}}$ be a random walk whose jumps distribution is given by (7).*

(i) The following assertions are equivalent.

- *(a) μ belongs to the domain of attraction of a 1-stable law.*
- *(b) There exists $L:(0,\infty)\to(0,\infty)$ that varies slowly at ∞ such that $\int^\infty y^{-1}L(y)\,\mathrm{d}y<\infty$ and such that $\mu([n,\infty))\sim n^{-1}L(n)$.*
- *(c) There exists a $(0,\infty)$-valued sequence $(a_n)_{n\in\mathbb{N}}$ tending to ∞ such that $\frac{1}{a_n}(W_n+b_n)$ converges in law to X, where $b_n=n\mathbb{E}[W_1\mathbf{1}_{\{|W_1|>a_n\}}]$ and where X is the spectrally positive 1-stable random variable whose Laplace exponent is given for all $\lambda\in(0,\infty)$ by $\ln\mathbb{E}[\exp(-\lambda X)]=\lambda\ln\lambda$.*

(ii) Assume that $(a$-$c)$ hold true. Then, $\sum_{k\geq n}k\mu(k)\sim\ell(n)$ where ℓ is the slowly varying function defined by

$$\forall x\in(0,\infty),\quad \ell(x)=\int_x^\infty\frac{L(y)}{y}\,\mathrm{d}y. \tag{26}$$

This implies that $\psi(s)\sim_{0^+}s\ell(1/s)$. Moreover, $L(x)=o(\ell(x))$ as $x\to\infty$.

(iii) Assume that $(a$-$c)$ hold true. Then, $a_n\sim nL(a_n)$ and $a_n\sim nL^*(n)$ where the function $L^*:x\in(0,\infty)\mapsto x^{-1}\inf\{y\in(0,\infty):y/L(y)>x\}$ varies slowly at ∞. Moreover, $b_n\sim n\ell(a_n)$ where ℓ is given by (26) and therefore $a_n=o(b_n)$.

Proof. The equivalences $(a)\Leftrightarrow(b)$ and $(a)\Leftrightarrow(c)$ are proved as in Proposition 2.9 (for the form of b_n, see e.g. [16, Chapter IX.8, Equation (8.15)]). To prove (ii), first observe that $\sum_{k\geq n}k\mu(k)=\sum_{j\geq1}\mu([\max(n,j),\infty))=n\mu([n,\infty))+\sum_{j>n}\mu([j,\infty))$. Then note that $n\mu([n,\infty))\sim L(n)$ and that $\sum_{j>n}\mu([j,\infty))\sim\ell(n)$ where ℓ is defined by (26). By Proposition 2.8 (iii), $L(n)=o(\ell(n))$ and we get $\sum_{k\geq n}k\mu(k)\sim\ell(n)$. By [5, Theorem 8.1.6], this is equivalent to $f_1(\lambda)\sim_{0^+}\lambda\ell(1/\lambda)$ where f_1 is given by (23). This implies $\psi(s)\sim_{0^+}s\ell(1/s)$ by Proposition 2.8 (i), which completes the proof of (ii). Then, we prove $a_n\sim nL(a_n)$ and $a_n\sim nL^*(n)$ as in Proposition 2.9. Next observe that $b_n=n\sum_{k>a_n}k\mu(k+1)\sim n\ell(a_n)$. Thus $a_n/b_n\sim L(a_n)/\ell(a_n)\to0$ by (ii). This completes the proof of (iii). ■

Suppose that Proposition 2.12 (c) holds true. Since $a_n=o(b_n)$, it holds $b_n^{-1}W_{\lfloor ns\rfloor}\to-s$ in law, and thus in probability, for all $s\in[0,\infty)$. Standard arguments (or a stronger result such as Skorokhod [31, Theorem 2.7]) entail the following convergence

$$\left(\tfrac{1}{b_n}W_{\lfloor ns\rfloor}\,;\,s\in[0,\infty)\right)\longrightarrow\left(-s\,;\,s\in[0,\infty)\right) \tag{27}$$

in probability for the topology of uniform convergence on all compact intervals. Recall from (21) the definition of the stopping times $(\mathtt{H}_p)_{p\in\mathbb{N}}$. Then, (27) implies for all $x\in[0,\infty)$ that

$$\tfrac{1}{n}\mathtt{H}_{\lfloor b_nx\rfloor}\longrightarrow x \tag{28}$$

in probability. Namely the law of $\mathtt{H}_1$ is *relatively stable* (see e.g. [5, Section 8.8 § 1]). Since the total size of a $\mathsf{GW}(\mu)$-tree τ has the same distribution as $\mathtt{H}_1$ (by Proposition 2.7 (iii)), the law of $\#\tau$ is thus relatively stable. By use of Berger [4, Theorem 2.4 and Lemma 4.3] and Kortchemski & Richier [26, Proposition 12], we get the following.

Lemma 2.13 *Let μ be a probability measure on $\mathbb{N}$ that satisfies (1) and that belongs to the domain of attraction of a 1-stable law. Let τ be a $\mathsf{GW}(\mu)$-tree. Then,*

$$\mathbb{P}(\#\tau\geq n)\sim\frac{L(b_n)}{b_n\ell(b_n)}, \tag{29}$$

where (b_n), L and ℓ are as in Proposition 2.12. If the more restrictive assumption (12) holds, then

$$\mathbb{P}(\#\tau=n)\sim\frac{1}{n}\mathbb{P}(\#\tau\geq n)\sim\frac{L(b_n)}{b_n^2}. \tag{30}$$

Proof. Since $b_n/n \sim \ell(a_n) \to 0$ as $n \to \infty$, (28) implies that there is a sequence (c_n) tending to ∞ such that $\frac{1}{c_n}\mathtt{H_n} \to 1$ in probability: namely, the law of $\mathtt{H}_1$ is relatively stable. Then, [5, Theorem 8.8.1] asserts the existence of a function l that slowly varies at ∞ such that $\sum_{0\le k\le n}\mathbb{P}(\mathtt{H}_1 \ge k) \sim l(n)$ and $1-\mathbb{E}[e^{-\lambda \mathtt{H}_1}] \sim_{0^+} \lambda\, l(1/\lambda)$. Thus, (28) easily entails that $\ell(a_n)l(n) \to 1$. By Berger [4, Lemma 4.3], we get $\ell(a_n) \sim \ell(b_n)$ and thus, $\sum_{0\le k\le n}\mathbb{P}(\#\tau \ge k) = \sum_{0\le k\le n}\mathbb{P}(\mathtt{H}_1 \ge k) \sim 1/\ell(b_n)$. By Kortchemski & Richier [26, Proposition 12], we get (29) because within the notations of [26], we necessarily get $\Lambda(n) \sim 1/\ell(b_n)$. Moreover, Berger [4, Theorem 2.4] asserts that if (12) holds then $\mathbb{P}(W_n = -1) \sim nL(b_n)/b_n^2$, so (30) follows from Kemperman's identity (Proposition 2.10 (ii)). ■

Remark 2.14 Although the difficult part of (29) is the very content of Kortchemski & Richier [26, Proposition 12], the relatively explicit form of the right member of (29) seems novel under the sole assumption that μ belongs to the domain of attraction of a 1-stable law. □

We next recall two limit theorems on the maximal out-degree of a $\mathsf{GW}(\mu)$-tree when μ belongs to the domain of attraction of a 1-stable law. They are part of more general results due to Kortchemski & Richier [26].

Proposition 2.15 *Let τ be a $\mathsf{GW}(\mu)$-tree with offspring distribution μ that satisfies (1) and that belongs to the domain of attraction of a 1-stable law. Recall from (14) that the notation $\Delta(\tau)$ stands for the maximal out-degree of τ. Then, the following holds true.*

(i) The following convergence holds in distribution on $[0,\infty)$:

$$\frac{1}{b_n}\Delta(\tau) \ \text{ under } \mathbb{P}(\,\cdot\,|\,\#\tau \ge n) \ \longrightarrow J, \tag{31}$$

where the law of J is given by $\mathbb{P}(J \ge x) = 1/x$ for all $x \in [1,\infty)$.

(ii) Under the more restrictive assumption (12), the following convergence holds in probability:

$$\frac{1}{b_n}\Delta(\tau) \ \text{ under } \mathbb{P}(\,\cdot\,|\,\#\tau = n) \ \longrightarrow 1. \tag{32}$$

Proof. For (i), see [26, Theorem 6]. For (ii), see [26, Theorem 1]. ■

Recall that $W(\tau)$ stands for the Lukasiewicz path of τ, as in Definition 2.6. We conclude this section by recalling a result from Kortchemski & Richier [26] that shows that the law of $W(\tau)$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ is closed in variation distance to the law of the Vervaat transform of the path $(W_0, W_1, \ldots, W_{n-1}, -1)$ under $\mathbb{P}$. More precisely, let $(W_n)_{n\in\mathbb{N}}$ be as in Proposition 2.12. For all $n \in \mathbb{N}^*$, we introduce the following notations.

$$I_n = -\min_{0\le j\le n-1} W_j \quad \text{and} \quad \sigma_n = \inf\{0 \le k \le n-1 \ : \ W_k = -I_n\} \tag{33}$$

and

$$Z_j^{(n)} = \begin{cases} W_{\sigma_n+j} + I_n & \text{if } \ 0 \le j < n-\sigma_n, \\ I_n - 1 + W_{j-(n-\sigma_n)} & \text{if } \ n \ge j \ge n-\sigma_n. \end{cases} \tag{34}$$

Namely, $Z^{(n)}$ is constructed by reading the increments of $(W_0, W_1, ..., W_{n-1}, -1)$ from left to right in cyclic order by starting at time σ_n: this is a kind of *Vervaat transform* of $(W_0, W_1, ..., W_{n-1}, -1)$ (see Vervaat [34] for more details). We shall use Kortchemski & Richier [26, Theorem 21], which we re-state by convenience into the following proposition.

Proposition 2.16 *Assume that μ satisfies (1) and (12). We keep the above notation. Then,*

$$\sup_{A\in\mathscr{B}(\mathbb{R}^{n+1})} \Big|\, \mathbb{P}\big(W(\tau) \in A \,\big|\, \#\tau = n\big) - \mathbb{P}\big(Z^{(n)} \in A\big) \,\Big| \longrightarrow 0,$$

where $\mathscr{B}(\mathbb{R}^{n+1})$ stands for the Borel sigma-field of $\mathbb{R}^{n+1}$.

Proof. See Kortchemski & Richier [26, Theorem 21] and note $\mathbb{P}(I_n > 1) \to 1$ by e.g. (27). ■

3 Distribution of the Horton-Strahler number of Galton-Watson trees

3.1 Alternative definitions of the Horton-Strahler number and basic results

In this section, we prove basic results on the Horton-Strahler number: firstly, we provide alternative definitions, secondly we state a key upper bound in Lemma 3.1, and finally we prove in Lemma 3.2 a recursive equation satisfied by the tail of the Horton-Strahler of $\mathsf{GW}(\mu)$-trees that is the starting point of the analysis of their asymptotic behavior.

Let us first recall alternative definitions of the Horton-Strahler number. The first one uses *Horton pruning* of a finite tree $t \in \mathbb{T}$ that is defined as follows: *remove the leaves of* t *and merge each line into one edge (a* line *in* t *is a maximal sequence of vertices* $v_0, \dots, v_n \in t$ *such that* $k_{v_1}(t) = \dots = k_{v_{n-1}}(t) = 1$ *and* $v_j = \overleftarrow{v_{j+1}}$ *for all* $0 \le j < n$*). The resulting tree is called the Horton-pruned tree, which we denote here by* $\mathtt{Prun}(t)$. Then, $\mathcal{S}(t)$ is the minimal number of Horton prunings that are necessary to obtain $\{\varnothing\}$ from t. Namely,

$$\mathcal{S}(t) = \min\big\{n \in \mathbb{N} \,:\, \mathtt{Prun}_n(t) = \{\varnothing\}\big\}, \tag{35}$$

where $\mathtt{Prun}_n$ stands for the n-th iteration of $\mathtt{Prun}$ for all $n \in \mathbb{N}^*$ and where $\mathtt{Prun}_0$ stands for the operation that merges each line into one edge. We refer to Kovchegov & Zaliapin [27, Section 2.3] for a proof and more details.

Another useful definition uses embeddings of perfect binary trees. More precisely, let $t, t' \in \mathbb{T}$ be finite. Then, $\phi : t \to t'$ is an *embedding if it is injective and if* $\phi(u \wedge v) = \phi(u) \wedge \phi(v)$ *for all* $u, v \in t$. For all $n \in \mathbb{N}$, we denote by $T_{2,n} = \bigcup_{0 \le k \le n} \{1,2\}^k$ the n-perfect binary tree, with the convention that $\{1,2\}^0 = \{\varnothing\}$. Then, for all finite tree $t \in \mathbb{T}$,

$$\mathcal{S}(t) = \max\big\{n \in \mathbb{N} \,:\, \exists \phi : T_{2,n} \to t \text{ embedding}\big\}. \tag{36}$$

This result seems to be 'part of the folklore'. Let us however provide a short proof.

Proof of (36). We reason by induction on the height of t. Note that (36) obviously holds true if $|t| = 0$. Now assume $k := k_\varnothing(t) \ge 1$ and let $\phi : T_{2,n+1} \to t$ be an embedding, we separate the cases according to the positions of $\phi(\varnothing), \phi(1), \phi(2)$.
If $(i) \preceq \phi(\varnothing)$ with $1 \le i \le k$, then we have $(i) \preceq \phi(u)$ for all $u \in T_{2,n+1}$ by definition, and we check that setting $\phi(u) = (i) * \phi_i(u)$ defines an embedding $\phi_i : T_{2,n+1} \to \theta_{(i)}t$. Conversely, an embedding $\phi_i : T_{2,n+1} \to \theta_{(i)}t$ induces an embedding $\phi : T_{2,n+1} \to t$ such that $(i) \preceq \phi(\varnothing)$. Thus,

$$\max\big\{n \in \mathbb{N} \,:\, \exists \phi : T_{2,n} \to t \text{ embedding}, (i) \preceq \phi(\varnothing)\big\} = \mathcal{S}(\theta_{(i)}t).$$

Otherwise, $\varnothing = \phi(\varnothing) = \phi(1) \wedge \phi(2)$ so we have distinct $1 \le i, j \le k$ such that $(i) \preceq \phi(1)$ and $(j) \preceq \phi(2)$. Similarly as before, we see that setting $\phi((1)*u) = (i)*\phi_i(u)$ and $\phi((2)*u) = (j)*\phi_j(u)$ respectively defines two embeddings $\phi_i : T_{2,n} \to \theta_{(i)}t$ and $\phi_j : T_{2,n} \to \theta_{(j)}t$. Conversely, two embeddings $\phi_i : T_{2,n} \to \theta_{(i)}t$ and $\phi_j : T_{2,n} \to \theta_{(j)}t$ induces an embedding $\phi : T_{2,n+1} \to t$ such that $\phi(\varnothing) = \varnothing$, $(i) \preceq \phi(1)$, and $(j) \preceq \phi(2)$. Thus,

$$\max\big\{n \in \mathbb{N} \,:\, \exists \phi : T_{2,n} \to t \text{ embedding}, (i) \preceq \phi(1), (j) \preceq \phi(2)\big\} = 1 + \min\big(\mathcal{S}(\theta_{(i)}t), \mathcal{S}(\theta_{(j)}t)\big).$$

Taking the maximum over $\phi(\varnothing), \phi(1), \phi(2)$ and recalling Definition 1.1 concludes the proof. ■

The definition (36) immediately implies the following. Let $t, t' \in \mathbb{T}$ be finite.

$$\textit{If there is an embedding } \phi : t \to t', \textit{ then } \quad \mathcal{S}(t) \le \mathcal{S}(t'). \tag{37}$$

We next use (36) and (37) to get an upper bound of $\mathcal{S}(t)$ in terms of $\mathcal{S}(R_m t)$, where $R_m t$ is defined in (20) (recall that it is the tree consisting in the first $m+1$ vertices of t in lexicographic order) and of $R_m t^\star$, where $t^\star$ is the *mirror image of* t that is formally defined as follows.

Let $t \in \mathbb{T}$ and $u \in t\backslash\{\varnothing\}$ be the word $(j_1, \ldots, j_n)$. We set $u_{|0} = \varnothing$ and $u_{|p} = (j_1, \ldots, j_p)$ for all $1 \le p \le n$. Then, the mirror image of u is the word $u^\star = (j_1^\star, \ldots, j_n^\star)$ where

$$j_p^\star := k_{u_{|p-1}}(t) - j_p + 1, \quad 1 \le p \le n.$$

We also set $\varnothing^\star = \varnothing$. Then $t^\star = \{u^\star \,:\, u \in t\}$ and it is easy to show that $t^\star \in \mathbb{T}$. Observe that $u \mapsto u^\star$ is a bijective embedding so $\mathcal{S}(t^\star) = \mathcal{S}(t)$ by (37). We stress that the word $u^\star$ depends on the tree t on which u is observed. Nevertheless, this notation should not lead to confusion here because the underlying tree will always be clear according to context. Since $k_{u^\star}(t^\star) = k_u(t)$, (16) implies that if τ is a $\mathsf{GW}(\mu)$-tree whose offspring distribution μ satisfies (1), then so is $\tau^\star$. Furthermore, since $\#\tau^\star = \#\tau$, if $n \in \mathbb{N}^*$ is such that $\mathbb{P}(\#\tau = n) > 0$, then we easily check that

$$\text{under } \mathbb{P}(\,\cdot\,|\,\#\tau = n), \quad \tau^\star \overset{(\text{law})}{=} \tau. \tag{38}$$

The following lemma plays a key role in the proofs of Theorems 1.2 and 1.4.

Lemma 3.1 *Let $t \in \mathbb{T}$ be finite. Then the following holds true.*

(i) For all $u \in t$, we set $t_{\le u} = \{v \in t \,:\, v \le u\}$. Then,

$$\mathcal{S}(t) \le 1 + \max\Big(\mathcal{S}(t_{\le u})\,,\, \max\{\mathcal{S}(\theta_v t) \,:\, v \in t,\, \overleftarrow{v} \preceq u \text{ and } v > u\}\Big). \tag{39}$$

(ii) Let $m \in \mathbb{N}$ be such that $2m \ge \#t + |t|$. Then,

$$\mathcal{S}(t) \le 1 + \max\big(\mathcal{S}(R_m t), \mathcal{S}(R_m t^\star)\big). \tag{40}$$

(iii) If τ is a random finite tree such that $\tau^\star$ has the same law as τ, then for all $n, m \in \mathbb{N}^*$, we get

$$\mathbb{P}(\mathcal{S}(\tau) \ge n) \le 2\mathbb{P}\big(\mathcal{S}(R_m\tau) \ge n - 1\big) + \mathbb{P}(\#\tau + |\tau| > 2m). \tag{41}$$

Proof. We first prove (i). If $\mathcal{S}(t) = 0$, then (39) is obviously true. We next suppose that $\mathcal{S}(t) = n \ge 1$, and to simplify notation, we set

$$[\![\varnothing, u]\!] = \{v \in t \,:\, v \preceq u\} \quad \text{and} \quad B = \{v \in t \,:\, \overleftarrow{v} \preceq u \text{ and } v > u\}.$$

By (36), there is an embedding $\phi : T_{2,n} \to t$. Observe that $\phi(1)$ and $\phi(2)$ cannot both belong to $[\![\varnothing, u]\!]$, otherwise it would imply $\phi(1) \wedge \phi(2) = \phi(\varnothing) \in \{\phi(1), \phi(2)\}$ (by definition of embeddings since $\varnothing = (1) \wedge (2)$), which contradicts the injectivity of ϕ. Thus, there is $j \in \{1, 2\}$ such that either $\phi(j) \in t_{\le u}\backslash[\![\varnothing, u]\!]$ or $\phi(j) \in t\backslash t_{\ge u}$. In the first case, by definition of embeddings, $\phi(j) \preceq \phi((j) * v)$ for all $v \in T_{2,n-1}$ and so $\phi((j) * T_{2,n-1}) \subset t_{\le u}$. Then, (36) entails $n - 1 \le \mathcal{S}(t_{\le u})$. Suppose next that $\phi(j) \in t\backslash t_{\ge u}$. Since $t\backslash t_{\ge u}$ is the disjoint union of the $v * (\theta_v t)$ for $v \in B$, there exists $v \in B$ such that $\phi((j) * T_{2,n-1}) \subset \phi(j) * (\theta_{\phi(j)} t) \subset v * (\theta_v t)$. Then, (36) entails $n - 1 \le \max_{v \in B} \mathcal{S}(\theta_v t)$, which completes the proof of (39).

Let us now prove (ii). To that end, we denote by $\overline{u}$ the $\le$-minimal leaf of $\theta_u t$. We also set $t_{\ge u} = \{v \in t : v \ge \overline{u}\} \cup [\![\varnothing, \overline{u}]\!]$. By definition, $t = t_{\le u} \cup t_{\ge u}$ and $[\![\varnothing, u]\!] = t_{\le u} \cap t_{\ge u}$. Then, note that for all $v \in B$, $v * (\theta_v t) \subset t_{\ge u}$. Moreover, note that $(t_{\ge u})^\star = t^\star_{\le \overline{u}^\star}$. Therefore, (37) and (i) imply

$$\mathcal{S}(t) \le 1 + \max\big(\mathcal{S}(t_{\le u}), \mathcal{S}(t_{\ge u})\big) = 1 + \max\big(\mathcal{S}(t_{\le u}), \mathcal{S}(t^\star_{\le \overline{u}^\star})\big). \tag{42}$$

Next set $m + 1 = \#t_{\le u} = \#\{v \in t : v \le u\}$ and $m' + 1 = \#t_{\ge u} = \#\{v \in t^\star : v \le \overline{u}^\star\}$. Observe that $R_m t = t_{\le u}$ and $R_{m'} t^\star = t^\star_{\le \overline{u}^\star}$. Moreover, $\#t = \#t_{\le u} + \#t_{\ge u} - \#[\![\varnothing, u]\!] = m + m' + 1 - |u|$. Thus, $m + m' < \#t + |t|$. If $2m \ge \#t + |t|$, then $m' < \frac{1}{2}(\#t + |t|) \le m$ and $R_{m'} t^\star \subset R_m t^\star$. By (37), we get $\mathcal{S}(R_{m'} t^\star) \le \mathcal{S}(R_m t^\star)$ and we obtain (40) by (42).

Inequality (41) is an easy consequence of (40): we leave the details to the reader. ■

We next prove the main equation that is satisfied by the tail distribution of the Horton-Strahler number of a Galton-Watson tree.

Lemma 3.2 *Let τ be a $\mathsf{GW}(\mu)$-tree whose offspring distribution μ satisfies (1). Recall φ and ψ from (2). For all $n \in \mathbb{N}$, we set $q_n = \mathbb{P}(\mathcal{S}(\tau) > n)$. Then,*

$$1 - q_0 = \frac{\mu(0)}{1-\mu(1)} \quad \text{and} \quad 1 - q_{n+1} = \varphi(1-q_n) + (q_n - q_{n+1})\varphi'(1-q_n), \quad n \in \mathbb{N}. \tag{43}$$

This equation can be rewritten in terms of ψ as $q_n - q_{n+1} = \psi(q_n)/\psi'(q_n)$.

Proof. By Definition 1.1, $\mathcal{S}(\tau)=0$ if and only if $k_\varnothing(\tau)=0$ or $(k_\varnothing(\tau)=1; \mathcal{S}(\theta_{(1)}\tau)=0)$. Thus, by Definition 2.2, we get $\mathbb{P}(\mathcal{S}(\tau)=0)=\mu(0)+\mu(1)\mathbb{P}(\mathcal{S}(\tau)=0)$, which gives the first equality in (43).

Let us prove the recursive relation in (43). Let $n \in \mathbb{N}$. By Definition 1.1, $\mathcal{S}(\tau) \le n+1$ if and only if $\mathcal{S}(\theta_u\tau) \le n$ for all children u of $\varnothing$ in τ (if any) with the possible exception of one child v, which may satisfy $\mathcal{S}(\theta_v\tau) = n+1$. More precisely,

$$\mathbf{1}_{\{\mathcal{S}(\tau)\le n+1\}} = \prod_{1\le j\le k_\varnothing(\tau)} \mathbf{1}_{\{\mathcal{S}(\theta_{(j)}\tau)\le n\}} + \sum_{1\le j\le k_\varnothing(\tau)} \mathbf{1}_{\{\mathcal{S}(\theta_{(j)}\tau)=n+1\}} \prod_{\substack{1\le i\le k_\varnothing(\tau)\\ j\neq i}} \mathbf{1}_{\{\mathcal{S}(\theta_{(i)}\tau)\le n\}}. \tag{44}$$

Taking the expectation yields (43) by Definition 2.2 and since φ is the generating function of μ. ■

Remark 3.3 Although it seems difficult to solve (43) explicitly in general, it can be done actually for the so-called *α-stable offspring distribution* μ_α, $\alpha \in (1,2]$, whose generating function is

$$\forall s \in [0,1], \quad \varphi_\alpha(s) = s + \tfrac{1}{\alpha}(1-s)^\alpha.$$

Namely, if τ_α is a $\mathsf{GW}(\mu_\alpha)$-tree then $\mathcal{S}(\tau_\alpha)$ is a geometric random variable with parameter $\frac{1}{\alpha}$, i.e.

$$\mathbb{P}(\mathcal{S}(\tau_\alpha) = n) = \tfrac{1}{\alpha}\big(1-\tfrac{1}{\alpha}\big)^n, \quad n \in \mathbb{N}.$$

This is explicitly proved in Kovchegov & Zaliapin [28, Lemma 10] and earlier for the $\alpha=2$ case, see Burd, Waymire & Winn [7, Proposition 2.5]. See also Duquesne & Winkel [13] who show that α-stable offspring distributions are the only laws that are invariant under any hereditary pruning. □

We end this section with a lemma that lists basic properties of ψ, that are useful to analyse (43).

Lemma 3.4 *Let μ be a probability on $\mathbb{N}$ that satisfies (1). Recall from (2) the definition of ψ. Then, the following holds true.*

- (i) *ψ is nonnegative, increasing, strictly convex, and analytic on $(0,1]$. Moreover, ψ' is increasing, concave, and $\psi(0) = \psi'(0) = 0$.*
- (ii) *For all $s \in (0,1]$, $\frac{1}{2}\psi'(s) \le \frac{\psi(s)}{s} \le \psi'(s)$.*

Proof. The point (i) is elementary. The upper bound in (ii) is a consequence of the convexity of ψ and to obtain the lower bound we apply *Hermite-Hadamard inequality* that asserts that for all convex functions $f : [a,b] \to \mathbb{R}$, it holds

$$f\big(\tfrac{1}{2}(a+b)\big) \le \frac{1}{b-a}\int_a^b f(t)\,\mathrm{d}t \le \tfrac{1}{2}\big(f(a)+f(b)\big). \tag{45}$$

We apply (45) to ψ' after observing $\frac{1}{s}\psi(s) = \frac{1}{s}\int_0^s \psi'(s)\mathrm{d}s$. ■

3.2 Tail estimates of joint laws

In this section, we provide four estimates of the tail of the joint distribution of the Horton-Strahler number of a Galton-Watson tree with either its size, its height, or its maximal out-degree. We fix an offspring distribution μ that satisfies (1) and τ stands for a $\mathsf{GW}(\mu)$-tree. Recall from (2) the definition of φ and ψ. Observe that $\psi'(1)=1-\varphi'(0)=1-\mu(1)>0$ by (1). It is convenient to write

$$q_n = \mathbb{P}(\mathcal{S}(\tau) > n), \quad n \in \{-1\} \cup \mathbb{N}, \tag{46}$$

where $q_{-1} = 1$ obviously.

Proposition 3.5 *Let τ be a $\mathsf{GW}(\mu)$-tree where μ satisfies (1). Then, for all $n \in \mathbb{N}$,*

$$\mathbb{E}\big[\#\tau \mathbf{1}_{\{\mathcal{S}(\tau)=0\}}\big] = \frac{\mu(0)}{(1-\mu(1))^2} \quad \textit{and} \quad \psi'(q_{n-1})\mathbb{E}\big[\#\tau \mathbf{1}_{\{\mathcal{S}(\tau)\le n\}}\big] \le 2. \tag{47}$$

Proof. Although $\mathbb{E}[\#\tau]=\infty$, let us first prove that $e_n := \mathbb{E}[\#\tau \mathbf{1}_{\{\mathcal{S}(\tau)\le n\}}] < \infty$. By Definition 1.1, if $\mathcal{S}(\tau) \le n$ then $\mathcal{S}(\theta_u\tau) \le n$ for all children u of $\varnothing$ in τ (if any) and $\mathcal{S}(\theta_u\tau) = n$ for at most one child. This implies that for all $m, n \in \mathbb{N}$,

$$\begin{aligned} \min(m, \#\tau)\mathbf{1}_{\{\mathcal{S}(\tau)\le n\}} \le 1 + \sum_{i=1}^{k_\varnothing(\tau)} \min\big(m, \#\theta_{(i)}\tau\big)\mathbf{1}_{\{\mathcal{S}(\theta_{(i)}\tau)\le n-1\}} \\ + \sum_{i=1}^{k_\varnothing(\tau)} \min\big(m, \#\theta_{(i)}\tau\big)\mathbf{1}_{\{\mathcal{S}(\theta_{(i)}\tau)\le n\}} \prod_{\substack{1\le j\le k_\varnothing(\tau)\\ j\ne i}} \mathbf{1}_{\{\mathcal{S}(\theta_{(j)}\tau)\le n-1\}} \end{aligned} \tag{48}$$

which makes sense even when $n = 0$: the first sum in the right-hand side of (48) being null. To simplify notations, we set $e_n(m) = \mathbb{E}[\min(m, \#\tau)\mathbf{1}_{\{\mathcal{S}(\tau)\le n\}}]$ for all integers $m \ge 0$ and $n \ge -1$ (with $e_{-1}(m)=0$). Taking the expectation in (48) gives $e_n(m) \le 1+e_{n-1}(m)+e_n(m)\varphi'(1-q_{n-1})$. It easily implies that $e_n \le (1 + e_{n-1})/(1-\varphi'(1-q_{n-1}))$ because $\lim_{m\to\infty} e_n(m) = e_n$. This recursively entails $e_n < \infty$ for all integer $n \ge -1$.

To simplify notations, we set $k_\varnothing = k_\varnothing(\tau)$, $\tau_j = \theta_{(j)}\tau$ and $\mathcal{S}_j = \mathcal{S}(\theta_{(j)}\tau)$ for all $1 \le j \le k_\varnothing(\tau)$. We first explicitly compute e_0 by observing that $\#\tau\mathbf{1}_{\{\mathcal{S}(\tau)=0\}} = \mathbf{1}_{\{k_\varnothing=0\}} + \mathbf{1}_{\{k_\varnothing=1\}}(1+\tau_1)\mathbf{1}_{\{\mathcal{S}_1=0\}}$. Taking the expectation entails $e_0 = \mu(0) + \mu(1)(1-q_0+e_0)$. By (43), this becomes

$$e_0 = \mathbb{E}[\#\tau \mathbf{1}_{\{\mathcal{S}(\tau)=0\}}] = \frac{1-q_0}{1-\mu(1)} = \frac{\mu(0)}{(1-\mu(1))^2}. \tag{49}$$

Next, by using the fact that $\#\tau = 1 + \sum_{1\le j\le k_\varnothing} \#\tau_j$ and by the decomposition (44), we get

$$\begin{aligned} \#\tau\, \mathbf{1}_{\{\mathcal{S}(\tau)\le n+1\}} = \mathbf{1}_{\{\mathcal{S}(\tau)\le n+1\}} + \sum_{j=1}^{k_\varnothing} \#\tau_j\, \mathbf{1}_{\{\mathcal{S}_j\le n\}} \prod_{\substack{1\le i\le k_\varnothing\\ j\ne i}} \mathbf{1}_{\{\mathcal{S}_i\le n\}} \\ + \sum_{j=1}^{k_\varnothing} \#\tau_j \mathbf{1}_{\{\mathcal{S}_j=n+1\}} \prod_{\substack{1\le i\le k_\varnothing\\ j\ne i}} \mathbf{1}_{\{\mathcal{S}_i\le n\}} + \sum_{\substack{1\le i,j\le k_\varnothing\\ j\ne i}} \#\tau_j \mathbf{1}_{\{\mathcal{S}_j\le n\}}\mathbf{1}_{\{\mathcal{S}_i=n+1\}} \prod_{\substack{1\le l\le k_\varnothing\\ l\ne i,j}} \mathbf{1}_{\{\mathcal{S}_l\le n\}}. \end{aligned}$$

Taking the expectation term-by-term, we get

$$e_{n+1} = 1 - q_{n+1} + e_n\varphi'(1-q_n) + (e_{n+1}-e_n)\varphi'(1-q_n) + e_n(q_n-q_{n+1})\varphi''(1-q_n)$$

for all $n \in \mathbb{N}$. Recall from (2) that $\psi(s)=\varphi(1-s)-1+s$. By Lemma 3.2, we find

$$e_{n+1}\psi'(q_n) = 1 - q_{n+1} + e_n(q_n-q_{n+1})\psi''(q_n) = 1 - q_{n+1} + e_n\frac{\psi''(q_n)}{\psi'(q_n)}\int_0^{q_n}\psi'(s)\,\mathrm{d}s \tag{50}$$

since $\psi(0) = 0$ by Lemma 3.4 (i). Still from Lemma 3.4, we know that ψ' is concave, so we get $\psi'(s) \le \psi'(q_n) - (q_n - s)\psi''(q_n)$. Thus,

$$\frac{\psi''(q_n)}{\psi'(q_n)} \int_0^{q_n} \psi'(s)\,\mathrm{d}s \le q_n\psi''(q_n) - \frac{(q_n\psi''(q_n))^2}{2\psi'(q_n)} = \tfrac{1}{2}\psi'(q_n)\big(1-(1-x)^2\big),$$

where $x = \psi''(q_n)/(\psi'(q_n)/q_n)$ belongs to $[0,1]$ since ψ' is concave. Thus, we get

$$x_{n+1} := e_{n+1}\psi'(q_n) \le 1 + \tfrac{1}{2}e_n\psi'(q_n) \le 1 + \tfrac{1}{2}e_n\psi'(q_{n-1}) = 1 + \tfrac{1}{2}x_n$$

since (q_n) is decreasing and ψ' is increasing. This entails $x_n \le 2 - 2^{-n}(1-x_0)$, which leads to (47) because $x_0 = \psi'(1)e_0 = (1-\mu(1))e_0 = 1 - q_0 < 1$ by (49). ■

For all $t \in \mathbb{T}$, we set

$$Z(t) = \max\{|u| \ : \ u \in t \text{ such that } \mathcal{S}(\theta_u t) = \mathcal{S}(t)\}.$$

Note that $Z(t) \le |t|$ where recall from (14) that $|t|$ stands for the height of t.

Proposition 3.6 *Let τ be a $\mathsf{GW}(\mu)$-tree where μ satisfies (1). Recall ψ from (2) and q_n from (46). Recall from (14) that $|\tau|$ stands for the height of τ. Then, $\mathbb{P}(Z(\tau) \ge m \,\big|\, \mathcal{S}(\tau) = n) = (1-\psi'(q_{n-1}))^m$ for all $m, n \in \mathbb{N}$. This implies that for all $\lambda \in (0,\infty)$,*

$$\limsup_{n\to\infty} \mathbb{P}\big(\psi'(q_{n-1})|\tau| \le \lambda \,\big|\, \mathcal{S}(\tau) = n\big) \le 1 - e^{-\lambda}. \tag{51}$$

Proof. By Definition 1.1, for all integers $n \ge 0$ and $m \ge 1$, it holds

$$\mathbf{1}_{\{Z(\tau)\ge m\,;\,\mathcal{S}(\tau)=n\}} = \sum_{i=1}^{k_\varnothing(\tau)} \mathbf{1}_{\{Z(\theta_{(i)}\tau)\ge m-1\,;\,\mathcal{S}(\theta_{(i)}\tau)=n\}} \prod_{\substack{1\le j\le k_\varnothing(\tau)\\ j\ne i}} \mathbf{1}_{\{\mathcal{S}(\theta_{(j)}\tau)\le n-1\}}$$

By taking the expectation, we get

$$\mathbb{P}\big(Z(\tau) \ge m\,;\, \mathcal{S}(\tau) = n\big) = \mathbb{P}\big(Z(\tau) \ge m-1\,;\, \mathcal{S}(\tau) = n\big)\,\varphi'(1-q_{n-1}),$$

which implies the first desired equality as $\varphi'(1-q_{n-1}) = 1-\psi'(q_{n-1})$. Note that $\psi'(q_{n-1}) \to \psi'(0) = 0$. Thus, $\lim_{n\to\infty} \mathbb{P}(Z(\tau) \ge \lambda/\psi'(q_{n-1}) \,|\, \mathcal{S}(\tau) = n) = e^{-\lambda}$ which implies (51) since $Z(\tau) \le |\tau|$. ■

In [6], Brandenberger, Devroye & Reddad study the Horton-Strahler number of a Galton-Watson tree conditioned to have exactly n vertices under the assumption that the variance of the offspring distribution is finite. To that end, they use a (little more than) local convergence of the conditioned Galton-Watson tree towards the corresponding size-biased tree. We adapt and extend this idea in a more general context using only the Many-To-One Principle to get the following.

Proposition 3.7 *Let τ be a $\mathsf{GW}(\mu)$-tree where μ satisfies (1). Recall ψ from (2) and q_n from (46). For all integers $n, m \in \mathbb{N}$ such that $2\psi'(\mathbb{P}(|\tau| \ge \lfloor n/2 \rfloor)) \le \psi'(q_m)$, the following inequality holds:*

$$\mathbb{P}\big(\mathcal{S}(\tau) \le m \,\big|\, |\tau| \ge n\big) \le \exp\big(-\tfrac{1}{8}\, n\psi'(q_m)\big). \tag{52}$$

Proof. For all $n \in \mathbb{N}$, we denote by φ_n the n-iterate of φ with the convention $\varphi_0 = \mathrm{Id}$. It is classical that $\mathbb{P}(|\tau| < n) = \varphi_n(0)$. Then, observe the following: the $\le$-smallest vertex of τ at height n is the only vertex $u \in \tau$ such that $|u| = n$ and such that for all $v \in \tau$ with $v < u$ and $\overleftarrow{v} \prec u$, we have

$|\theta_v\tau| + |v| < n$. Moreover for all $p \in \mathbb{N}$, (37) implies that if $\mathcal{S}(\tau) \leq m$ then $\mathcal{S}(\theta_v\tau) \leq m$ for all $v \in \tau$. Therefore,

$$\mathbf{1}_{\{|\tau|\geq n\,;\,\mathcal{S}(\tau)\leq m\}} \leq \sum_{u\in\tau} \mathbf{1}_{\{|u|=n\}} \prod_{\substack{w\in\tau, i\geq 1\\ w*(i)\preceq u}} \Big(\prod_{j=1}^{i-1} \mathbf{1}_{\{|\theta_{w*(j)}\tau|+|w*(j)|<n\}}\Big)\Big(\prod_{j=i+1}^{k_w(\tau)} \mathbf{1}_{\{\mathcal{S}(\theta_{w*(j)}\tau)\leq m\}}\Big). \tag{53}$$

Here we adopt the following convention: a product over an empty set of indices is taken equal to 1. Recall Definition 2.3 of the size-biased GW(μ)-tree τ_∞ and recall from (18) in Remark 2.4 the joint law of the number of left/right siblings of individuals on the infinite line of descent. We now use the Many-To-One Principle (Proposition 2.5) after taking the expectation in (53) to get

$$\mathbb{P}\big(|\tau| \geq n\,;\,\mathcal{S}(\tau) \leq m\big) \leq \prod_{p=0}^{n-1} \frac{\varphi(1-q_m) - \varphi(\varphi_p(0))}{1-q_m-\varphi_p(0)}. \tag{54}$$

We now use convexity properties of φ and ψ given in Lemma 3.4 to get an upper bound of the right-hand side of (54). First observe that the convexity of φ implies that for all real numbers $s, r \in [0,1]$ such that $s \leq r$, we have $(\varphi(r)-\varphi(s))/(r-s) \leq (1-\varphi(s))/(1-s)$. Therefore,

$$\prod_{p=0}^{\lfloor n/2\rfloor-1} \frac{\varphi(1-q_m) - \varphi(\varphi_p(0))}{1-q_m-\varphi_p(0)} \leq \prod_{p=0}^{\lfloor n/2\rfloor-1} \frac{1-\varphi(\varphi_p(0))}{1-\varphi_p(0)} = 1-\varphi_{\lfloor n/2\rfloor}(0)\,. \tag{55}$$

To get an upper bound of $(\varphi(1-q_m)-\varphi(\varphi_p(0)))/(1-q_m-\varphi_p(0))$ when $p \geq \lfloor n/2\rfloor$, we use the following: let $s, r \in [0,1]$ and suppose that $2\psi'(s) \leq \psi'(r)$, then

$$\begin{aligned}
\frac{\varphi(1-r)-\varphi(1-s)}{s-r} &= \frac{1-\varphi(1-s)}{s} + \frac{\varphi(1-s)-1}{s} + \frac{\varphi(1-r)-\varphi(1-s)}{s-r}\\
&= \frac{1-\varphi(1-s)}{s} + \frac{\psi(s)}{s} - \frac{\psi(r)-\psi(s)}{r-s}\\
&= \frac{1-\varphi(1-s)}{s} + \frac{\psi(s)}{s} - \frac{1}{r-s}\int_s^r \psi'(x)\,\mathrm{d}x\\
&\leq \frac{1-\varphi(1-s)}{s} + \psi'(s) - \tfrac{1}{2}\big(\psi'(r)+\psi'(s)\big)\\
&\leq \frac{1-\varphi(1-s)}{s} - \tfrac{1}{4}\psi'(r) \leq \frac{1-\varphi(1-s)}{s}\big(1-\tfrac{1}{4}\psi'(r)\big).
\end{aligned} \tag{56}$$

Here, we have used Hermite-Hadamard inequality (45) for concave functions, and the two convexity inequalities $\frac{1}{s}\psi(s) \leq \psi'(s)$ and $\frac{1}{s}(1-\varphi(1-s)) \leq \varphi'(1) = 1$.

Assume that $m, n \in \mathbb{N}$ satisfy $2\psi'(1-\varphi_{\lfloor n/2\rfloor}(0)) \leq \psi'(q_m)$. Then, for all $p \geq \lfloor n/2\rfloor$, we have $2\psi'(1-\varphi_p(0)) \leq \psi'(q_m)$. Applying (56) successively with $s = 1-\varphi_p(0)$ and $r = q_m$ gets us

$$\begin{aligned}
\mathbb{P}(|\tau| \geq n\,;\,\mathcal{S}(\tau) \leq m) &\leq \big(1-\varphi_{\lfloor n/2\rfloor}(0)\big)\big(1-\tfrac{1}{4}\psi'(q_m)\big)^{n/2} \prod_{p=\lfloor n/2\rfloor}^{n-1} \frac{1-\varphi_{p+1}(0)}{1-\varphi_p(0)}\\
&\leq \big(1-\tfrac{1}{4}\psi'(q_m)\big)^{n/2}\mathbb{P}(|\tau| \geq n),
\end{aligned}$$

by (54) and (55). This easily entails (52) since $\ln(1-x) \leq -x$ for all $x \in [0,1)$. ■

Although Proposition 3.7 holds under the sole assumption that μ satisfies (1), its application requires knowing the behavior of the tail of the height $|\tau|$ of the GW(μ)-tree τ. When μ belongs to the domain of attraction of a stable law of index $\alpha \in (1,2]$, Proposition 2.10 (iv) provides such information, and then Proposition 3.7 entails the following more convenient result.

Corollary 3.8 *Let τ be a $\mathsf{GW}(\mu)$-tree where μ satisfies (1) and belongs to the domain of attraction of a stable law of index $\alpha \in (1,2]$. Recall ψ from (2) and q_n from (46). There exists a constant $C_\mu \in (0,\infty)$ that only depends on μ such that for all integers $n, m \in \mathbb{N}$, it holds*

$$\mathbb{P}\big(\mathcal{S}(\tau) \leq m \,\big|\, |\tau| \geq n\big) \leq C_\mu \exp\big(-\tfrac{1}{8}\, n\psi'(q_m)\big). \tag{57}$$

Proof. To simplify the notations, we set $s_n = \mathbb{P}(|\tau| \geq \lfloor n/2 \rfloor)$ for all $n \in \mathbb{N}$. By Proposition 2.10 (iv), the constant $2c_\mu := \sup_{n\in\mathbb{N}}(n+1)\frac{\psi(s_n)}{s_n}$ is finite and positive. Then, we get from Lemma 3.4 (ii) that $2\psi'(s_n) \leq 8c_\mu/(n+1)$ for all $n \in \mathbb{N}$, by definition of c_μ. Let us set $C_\mu = e^{c_\mu} \in (0,\infty)$ and let $n, m \in \mathbb{N}$. If $\psi'(q_m) \geq 8c_\mu/(n+1)$, then $2\psi'(s_n) \leq \psi'(q_m)$ so Proposition 3.7 yields (57) because $1 \leq C_\mu$. Otherwise, $\psi'(q_m) < 8c_\mu/(n+1)$ and $C_\mu \exp(-\frac{1}{8}n\psi'(q_m)) \geq 1$ by choice of C_μ. Thus, (57) clearly holds in that case, which completes the proof. ■

Our last estimate relies on the same idea as Proposition 3.7: if the maximal out-degree of a Galton-Watson tree is large, then the tree contains several independent disjoint copies of itself.

Proposition 3.9 *Let τ be a $\mathsf{GW}(\mu)$-tree where μ satisfies (1). Recall ψ from (2) and q_n from (46). Recall from (14) that $\Delta(\tau)$ stands for the maximal out-degree of τ. Then, for all integers $n, m \geq 1$ such that $\mathbb{P}(\Delta(\tau) \geq n) > 0$, the following inequality holds true:*

$$\mathbb{P}\big(\mathcal{S}(\tau) \leq m \,\big|\, \Delta(\tau) \geq n\big) \leq e^{-nq_m}. \tag{58}$$

Moreover, for all $\lambda \in (0,\infty)$, we also have

$$\mathbb{E}\big[e^{-\lambda \#\tau} \,\big|\, \Delta(\tau) \geq n\big] \leq \mathbb{E}\big[e^{-\lambda\#\tau}\big]^n. \tag{59}$$

Proof. On the event $\{\mathcal{S}(\tau) \leq m\,;\, \Delta(\tau) \geq n\}$, we decompose τ along the ancestral line of the $\leq$-first vertex u such that $k_u(\tau) \geq n$ and, by (37), we see $\mathcal{S}(\theta_{u*(j)}\tau) \leq m$ for all $1 \leq j \leq n$. Hence,

$$\mathbf{1}_{\{\mathcal{S}(\tau)\leq m\,;\,\Delta(\tau)\geq n\}} \leq \sum_{u\in\tau} \mathbf{1}_{\{k_u(\tau)\geq n\}} \Big(\prod_{v\in\tau:v<u} \mathbf{1}_{\{k_v(\tau)<n\}}\Big)\Big(\prod_{j=1}^{n} \mathbf{1}_{\{\mathcal{S}(\theta_{u*(j)}\tau)\leq m\}}\Big).$$

We take the expectation and apply the Many-To-One Principle (Proposition 2.5) 'forwards and backwards' to get

$$\begin{aligned}\mathbb{P}\left(\mathcal{S}(\tau) \leq m\,;\, \Delta(\tau) \geq n\right) &\leq \mathbb{E}\Big[\sum_{u\in\tau} \mathbf{1}_{\{k_u(\tau)\geq n\}} \prod_{v\in\tau:v<u} \mathbf{1}_{\{k_v(\tau)<n\}}\Big]\mathbb{P}\big(\mathcal{S}(\tau)\leq m\big)^n \\ &= \mathbb{P}\big(\Delta(\tau)\geq n\big)\big(1-q_m\big)^n,\end{aligned}$$

which entails (58) since $\ln(1-x) \leq -x$ for all $x \in [0,1)$. To prove (59), we observe that

$$e^{-\lambda\#\tau}\mathbf{1}_{\{\Delta(\tau)\geq n\}} \leq \sum_{u\in\tau} \mathbf{1}_{\{k_u(\tau)\geq n\}} \Big(\prod_{v\in\tau:v<u} \mathbf{1}_{\{k_v(\tau)<n\}}\Big)\Big(\prod_{j=1}^{n} e^{-\lambda\#\theta_{u*(j)}\tau}\Big).$$

Then, as in the previous argument, we take the expectation and apply the Many-To-One Principle (Proposition 2.5) 'forwards and backwards' to get the desired result. ■

3.3 Tail estimates.

In this section, we prove results on the tail of the Horton-Strahler number of a $\mathsf{GW}(\mu)$-tree. These results are expressed in terms of the functions Λ and Υ that are given by

$$\forall s \in (0,1), \quad \Upsilon(s) = \int_s^1 \frac{\mathrm{d}r}{r\ln\Lambda(r)}, \quad \text{where} \quad \Lambda(s) = \frac{s\psi'(s)}{s\psi'(s)-\psi(s)}, \tag{9}$$

and where ψ is defined in (2). Their basic properties are listed in the following lemma.

Lemma 3.10 *Let μ be a probability on $\mathbb{N}$ that satisfies (1). Then, the functions Λ and Υ in (9) are well-defined, $\Lambda \geq 2$, and Υ is continuous, positive, and decreasing on $(0,1)$.*

Proof. The inequality $\frac{1}{s}\psi(s) \geq \frac{1}{2}\psi'(s)$ from Lemma 3.4 (ii) entails that $\Lambda(s) \geq 2$ for all $s \in (0,1)$, and (iii) follows immediately. ■

Under the additional assumption that μ belongs to the domain of attraction of a stable law, the following lemma shows that Λ varies slowly at 0^+.

Proposition 3.11 *Let us assume that μ satisfies (1) and that it belongs to the domain of attraction of a stable law of index $\alpha \in [1,2]$.*

(i) If $\alpha \in (1,2]$, then $\lim_{s\to 0^+} \Lambda(s) = \frac{\alpha}{\alpha-1}$ and $\Upsilon(s) \sim_{0^+} \log_{\frac{\alpha}{\alpha-1}} 1/s$.

(ii) Suppose that $\alpha = 1$. Let L be such that $\mu([n,\infty)) \sim n^{-1}L(n)$ and ℓ be the slowly varying function given by $\ell(x) = \int_x^\infty y^{-1}L(y)\,\mathrm{d}y$ for all $x \in (0,\infty)$ (see Proposition 2.12). Then,

$$\Lambda(s) \sim_{0^+} \frac{\ell(1/s)}{L(1/s)} \quad \text{and} \quad \lim_{s\to 0^+} \Lambda(s) = \infty.$$

Proof. Let us prove (i) first. By Proposition 2.9, ψ is regularly varying of index α at 0^+. Then, recall that ψ' is increasing, and since $\psi(s) = \int_0^s \psi'(x)\,\mathrm{d}x$, the Monotone Density Theorem (recalled in Proposition 2.8 (iv)) implies that $s\psi'(s) \sim_{0^+} \alpha\psi(s)$. This entails $\lim_{s\to 0^+} \Lambda(s) = \frac{\alpha}{\alpha-1}$ and thus $\Upsilon(s) \sim_{0^+} \log_{\frac{\alpha}{\alpha-1}} 1/s$. This completes the proof of (i).

Let us assume that $\alpha = 1$ to prove (ii). To simplify, we denote by ξ a random variable whose distribution is μ. Thus, $x\mathbb{P}(\xi > x) \sim_\infty L(x)$. Proposition 2.12 (ii) asserts that $\mathbb{E}[\xi \mathbf{1}_{\{\xi>x\}}] \sim_\infty \ell(x)$ and $\psi(s) \sim_{0^+} s\ell(1/s)$. The Monotone Density Theorem again asserts that $s\psi'(s) \sim_{0^+} s\ell(1/s)$. We next consider $s\psi'(s) - \psi(s)$, $s \in (0,1)$ being fixed. To simplify notations, we set $\lambda = -\ln(1-s)$ and we first observe that

$$\begin{aligned} s\psi'(s) - \psi(s) &= \mathbb{E}\left[1 - e^{-\lambda\xi} - \tfrac{s\xi}{1-s}e^{-\lambda\xi}\right] = \mathbb{E}\left[\int_0^\xi \left(\lambda e^{-\lambda x} - \tfrac{s}{1-s}e^{-\lambda x} + \tfrac{sx\lambda}{1-s}e^{-\lambda x}\right)\mathrm{d}x\right], \\ \text{(by Fubini)} &= \tfrac{s\lambda}{1-s}\int_0^\infty \mathbb{P}(\xi > x) x e^{-\lambda x}\,\mathrm{d}x \;-\; \left(\tfrac{s}{1-s} - \lambda\right)\int_0^\infty \mathbb{P}(\xi > x) e^{-\lambda x}\,\mathrm{d}x. \end{aligned}$$

We estimate the second term of the right-hand side by writing

$$\left(\tfrac{s}{1-s} + \ln(1-s)\right)\int_0^\infty \mathbb{P}(\xi > x) e^{-\lambda x}\,\mathrm{d}x \;\sim_{0^+}\; \tfrac{1}{2}s^2\mathbb{E}[\xi] = \tfrac{1}{2}s^2. \tag{60}$$

Next, Karamata's Abelian Theorem for Laplace transform (as recalled in Proposition 2.8 (v)) asserts

$$\lambda \int_0^\infty x\mathbb{P}(\xi > x)\, e^{-\lambda x}\,\mathrm{d}x \;\sim_{0^+}\; L(1/\lambda). \tag{61}$$

Since $\lambda = -\ln(1-s)$, (60) and (61) entail that $s\psi'(s) - \psi(s) \sim_{0^+} sL(1/s)$ thanks to Potter's bound (see Proposition 2.8 (i)), which implies the desired estimate for Λ since $s\psi'(s) \sim_{0^+} s\ell(1/s)$. Finally, $\lim_{s\to\infty} \Lambda(s) = \infty$ comes from an application of Proposition 2.8 (iii). ■

Proposition 3.12 *Let us assume that μ satisfies (1) and that it belongs to the domain of attraction of a stable law of index $\alpha \in [1,2]$. Recall the definition of Υ from (9). Then,*

$$\Upsilon\big(\mathbb{P}(\mathcal{S}(\tau) > n)\big) \sim n. \tag{62}$$

In particular, if $\alpha \in (1,2]$, we get $-\ln \mathbb{P}(\mathcal{S}(\tau) > n) \sim n \ln \frac{\alpha}{\alpha-1}$.

Proof. Recall from (46) that $q_n = \mathbb{P}(\mathcal{S}(\tau) > n)$ and note that $q_n \to 0$. Lemma 3.2 asserts that $q_n - q_{n+1} = \psi(q_n)/\psi'(q_n)$, namely $q_{n+1}^{-1} = q_n^{-1}\Lambda(q_n)$. To simplify, we set $Q_n = -\ln q_n$. Therefore,

$$Q_{n+1} = Q_n + \ln \Lambda(e^{-Q_n}), \quad n \in \mathbb{N}. \tag{63}$$

By Proposition 3.11, Λ varies slowly at 0^+. Thus, Potter's bound (recalled in Proposition 2.8 (i)) applies to $l(y) = \Lambda(1/y)$: for any $\varepsilon \in (0, \frac{1}{10})$, there is $n_\varepsilon \in \mathbb{N}$ such that for all $n \geq n_\varepsilon$ and for all $x \in [Q_n, Q_{n+1}]$,

$$\begin{aligned} \left|\ln \Lambda(e^{-x}) - \ln \Lambda(e^{-Q_n})\right| &= \left|\ln\left(\frac{\Lambda(e^{-(x-Q_n)}e^{-Q_n})}{\Lambda(e^{-Q_n})}\right)\right| \leq \varepsilon + \varepsilon(x - Q_n) \\ &\leq \varepsilon + \varepsilon \ln \Lambda(e^{-Q_n}). \end{aligned}$$

by (63) (we apply Proposition 2.8 (i) with $l(y) = \Lambda(1/y)$, $c = e^\varepsilon$ and $\lambda = e^{x-Q_n}$). This implies

$$\begin{aligned} \left|\frac{1}{\ln \Lambda(e^{-x})} - \frac{1}{\ln \Lambda(e^{-Q_n})}\right| &\leq \frac{\varepsilon}{\ln \Lambda(e^{-Q_n})} \cdot \frac{1 + \ln \Lambda(e^{-Q_n})}{\ln \Lambda(e^{-x})} \\ &\leq \frac{\varepsilon}{\ln \Lambda(e^{-Q_n})} \cdot \frac{1 + \ln \Lambda(e^{-Q_n})}{(1-\varepsilon)\ln \Lambda(e^{-Q_n}) - \varepsilon} \\ &\leq \frac{\varepsilon}{\ln \Lambda(e^{-Q_n})} \cdot \frac{(\ln 2)^{-1} + 1}{1 - \varepsilon - \varepsilon(\ln 2)^{-1}} \leq \frac{100\varepsilon}{\ln \Lambda(e^{-Q_n})} \end{aligned}$$

because $\Lambda \geq 2$, by Lemma 3.10, and since $\varepsilon \in (0, \frac{1}{10})$. Therefore, by (63), we get

$$\int_{Q_n}^{Q_{n+1}} \left|\frac{1}{\ln \Lambda(e^{-x})} - \frac{1}{\ln \Lambda(e^{-Q_n})}\right| \mathrm{d}x \leq 100\varepsilon \cdot \frac{Q_{n+1} - Q_n}{\ln \Lambda(e^{-Q_n})} = 100\varepsilon.$$

This implies that for all $\varepsilon \in (0, \frac{1}{10})$ and for all $n \geq n_\varepsilon$,

$$\left|\sum_{n_\varepsilon \leq k < n} \frac{Q_{k+1} - Q_k}{\ln \Lambda(e^{-Q_k})} - \int_{Q_{n_\varepsilon}}^{Q_n} \frac{\mathrm{d}x}{\ln \Lambda(e^{-x})}\right| = \left|n - n_\varepsilon - \int_{Q_{n_\varepsilon}}^{Q_n} \frac{\mathrm{d}x}{\ln \Lambda(e^{-x})}\right| \leq 100\varepsilon(n - n_\varepsilon).$$

Hence, $\lim_{n\to\infty} \frac{1}{n} \int_{Q_0}^{Q_n} \frac{\mathrm{d}x}{\ln \Lambda(e^{-x})} = 1$, which proves (62) after the change of variable $r = e^{-x}$. ■

In the 1-stable cases, we provide below three examples of slowly varying functions L that may govern the tail of μ via the estimate $n\mu([n, \infty)) \sim L(n)$. We use the notations of Proposition 3.11.

Example 3.13 **(a)** Let $\kappa \in (0, \infty)$. We consider the case where $L(x) = (\ln x)^{-1-\kappa}$ for $x \in (0, \infty)$. Indeed, this function varies slowly at ∞ and is such that $\int^\infty y^{-1}L(y)\,\mathrm{d}y < \infty$. Clearly, we have $\ell(x) = \frac{1}{\kappa}(\ln x)^{-\kappa} = \frac{1}{\kappa}L(x)\ln x$. Therefore, Proposition 3.11 (ii) yields

$$\Lambda(s) \sim_{0^+} \tfrac{1}{\kappa} \ln 1/s \quad \text{and} \quad \Upsilon(s) \sim_{0^+} \frac{\ln 1/s}{\ln \ln 1/s}.$$

If one sets $x_n = -\ln \mathbb{P}(\mathcal{S}(\tau) > n)$, Proposition 3.12 asserts that $x_n / \ln x_n \sim n$, which implies that

$$-\ln \mathbb{P}(\mathcal{S}(\tau) > n) \sim n \ln n.$$

(b) Let $\kappa \in (0, 1)$. We next consider the case where $L(x) = \exp(-(\ln x)^\kappa)$ for $x \in (0, \infty)$. Again, this function varies slowly at ∞ and verifies $\int^\infty y^{-1}L(y)\,\mathrm{d}y < \infty$. An integration by parts gives

$$\ell(x) = \tfrac{1}{\kappa}(\ln x)^{1-\kappa}L(x) + \tfrac{1-\kappa}{\kappa} \int_x^\infty \frac{L(y)}{y(\ln y)^\kappa}\,\mathrm{d}y \sim_\infty \tfrac{1}{\kappa}(\ln x)^{1-\kappa}L(x),$$

and Proposition 3.11 (ii) thus implies that

$$\Lambda(s) \sim_{0^+} \tfrac{1}{\kappa}(\ln 1/s)^{1-\kappa} \quad \text{and} \quad \Upsilon(s) \sim_{0^+} \tfrac{1}{1-\kappa} \cdot \frac{\ln 1/s}{\ln\ln 1/s}.$$

If one sets $x_n = -\ln \mathbb{P}(\mathcal{S}(\tau) > n)$, Proposition 3.12 asserts that $x_n/\ln x_n \sim (1-\kappa)n$, which yields

$$-\ln \mathbb{P}(\mathcal{S}(\tau) > n) \sim (1-\kappa)\, n \ln n.$$

(c) We finally consider the case where $L(x) = \exp(-\ln x/\ln\ln x)$ for all $x \in (e^e, \infty)$. This function still slowly varies and verifies $\int^\infty y^{-1}L(y)\,\mathrm{d}y < \infty$. We make the change of variable $z = \ln y/\ln\ln y$ and then an integration by parts to compute

$$\ell(x) \sim_\infty \int_{\frac{\ln x}{\ln\ln x}}^\infty e^{-z}\ln(z)\,\mathrm{d}z \sim_\infty \Big[-e^{-z}\ln z\Big]_{\frac{\ln x}{\ln\ln x}}^\infty \sim_\infty L(x)\ln\ln x.$$

Therefore, Proposition 3.11 (ii) implies that

$$\Lambda(s) \sim_{0^+} \ln\ln 1/s \quad \text{and} \quad \Upsilon(s) \sim_{0^+} \frac{\ln 1/s}{\ln\ln\ln 1/s}.$$

If one sets $x_n = -\ln \mathbb{P}(\mathcal{S}(\tau) > n)$, Proposition 3.12 asserts that $x_n/\ln\ln x_n \sim n$, which entails

$$-\ln \mathbb{P}(\mathcal{S}(\tau) > n) \sim n \ln\ln n.$$

□

We use further the following estimates.

Proposition 3.14 *Let us assume that μ satisfies (1) and that it belongs to the domain of attraction of a stable law of index $\alpha \in [1,2]$. Then, for all $\lambda \in (0,\infty)$ and all $\kappa \in \mathbb{R}$, the following holds true:*

$$\ln\ln 1/s \underset{s\to 0^+}{=} o\left(\Upsilon(s)\right), \quad \Upsilon(\lambda s \ln^\kappa 1/s) \sim_{0^+} \Upsilon(s), \quad \textit{and} \quad \Upsilon\left(s\Lambda(s)^\kappa\right) \sim_{0^+} \Upsilon(s).$$

In particular, Υ is slowly varying at 0^+.

Proof. Recall that Λ is slowly varying at 0^+ from Proposition 3.11 so $\ln\Lambda(s) = o(\ln 1/s)$ by Proposition 2.8 (i). This allows us to write

$$\ln\ln 1/s \sim_{0^+} \int_s^{1/2} \frac{\mathrm{d}r}{r\ln 1/r} \underset{s\to 0^+}{=} o\big(\Upsilon(s)\big).$$

Next, without loss of generality, we may assume that $\lambda \in [1,\infty)$, so that $1 \le \lambda \ln^{|\kappa|} 1/s$ when s is small enough. Recall from Lemma 3.10 that $\Lambda \ge 2$. Thus,

$$\left|\Upsilon(\lambda s\ln^\kappa 1/s) - \Upsilon(s)\right| \le \int_{\frac{s}{\lambda}\ln^{-|\kappa|} 1/s}^{\lambda s \ln^{|\kappa|} 1/s} \frac{\mathrm{d}r}{r\ln\Lambda(r)} \le \frac{1}{\ln 2}\ln\left(\lambda^2 \ln^{2|\kappa|} 1/s\right) \underset{s\to 0^+}{=} o\big(\Upsilon(s)\big),$$

which is the second estimate. Then, observe that $\lim_{s\to 0^+} s\Lambda(s)^{|\kappa|} = 0$ because Λ slowly varies at 0^+. Another application of Proposition 2.8 (i), together with $\Lambda \ge 2$, entails that $\ln\Lambda(r) \ge \frac{1}{2}\ln\Lambda(s)$ for all small enough s and for all $r \in (0,1)$ such that $s\Lambda(s)^{-|\kappa|} \le r \le s\Lambda(s)^{|\kappa|}$. Therefore,

$$\left|\Upsilon(s\Lambda(s)^\kappa) - \Upsilon(s)\right| \le \int_{s\Lambda(s)^{-|\kappa|}}^{s\Lambda(s)^{|\kappa|}} \frac{\mathrm{d}r}{r\ln\Lambda(r)} \le \frac{2}{\ln\Lambda(s)}\ln\left(\Lambda(s)^{2|\kappa|}\right) = 4|\kappa| \underset{s\to 0^+}{=} o\big(\Upsilon(s)\big).$$

which completes the proof. ■

4 Proofs of Theorems 1.2, 1.3 and 1.4

4.1 Proof of Theorem 1.2

In all this section, we assume that the critical offspring distribution μ belongs to the domain of attraction of an α-stable law with $\alpha \in (1,2]$, and more precisely, we assume that Proposition 2.9 $(a\text{-}d)$ hold. Let us set $\gamma = \ln \frac{\alpha}{\alpha-1}$. Recall from Propositions 3.11 (i) and 3.12 that

$$\gamma \Upsilon(s) \sim_{0+} \ln 1/s \quad \text{and} \quad -\ln \mathbb{P}(\mathcal{S}(\tau) > n) \sim \gamma n. \tag{64}$$

The proof of Theorem 1.2 is separated into two parts: a lower bound and an upper bound.

Lower bound. We first prove for all $\varepsilon \in (0,1)$ that

$$\mathbb{P}\big(\alpha\gamma \mathcal{S}(\tau) \le (1-\varepsilon)\ln n \;\big|\; \#\tau = n\big) \xrightarrow[n\to\infty]{} 0. \tag{65}$$

The idea is to apply Proposition 3.7, or rather Corollary 3.8, and to switch the conditioning on the size of τ with the conditioning on the height of τ. To that end, we use Proposition 2.11 (i) to find

$$\limsup_{\eta\to 0} \limsup_{n\to\infty} \mathbb{P}\big(\,|\tau| \le \eta \tfrac{n}{a_n} \;\big|\; \#\tau = n\big) = 0.$$

Therefore, to prove (65), we only need to prove for all $\eta \in (0,1)$ that

$$\limsup_{n\to\infty} \mathbb{P}\big(\,|\tau| > \eta \tfrac{n}{a_n}\,;\, \alpha\gamma \mathcal{S}(\tau) \le (1-\varepsilon)\ln n \;\big|\; \#\tau = n\big) = 0.$$

We roughly bound the conditional probability by the ratio of the probabilities as follows.

$$\mathbb{P}\big(\,|\tau| > \eta \tfrac{n}{a_n}\,;\, \alpha\gamma \mathcal{S}(\tau) \le (1-\varepsilon)\ln n \;\big|\; \#\tau = n\big) \le \frac{\mathbb{P}\big(\,\alpha\gamma \mathcal{S}(\tau) \le (1-\varepsilon)\ln n \;\big|\; |\tau| > \eta \tfrac{n}{a_n}\big)}{\mathbb{P}(\#\tau = n)}. \tag{66}$$

Since $a_n \sim n^{\frac{1}{\alpha}+o(1)}$ by Proposition 2.9 and Potter's bound (see Proposition 2.8 (i)), Proposition 2.10 (iii) entails the following estimate for the denominator of the right-hand side of (66):

$$\mathbb{P}(\#\tau = n) \sim \frac{c_\alpha}{n a_n} \sim n^{-1-\frac{1}{\alpha}+o(1)}. \tag{67}$$

To bound the right-hand side of (66), we want to apply Corollary 3.8 and to that end, we first control $\psi'\big(\mathbb{P}\big(\mathcal{S}(\tau) > \frac{1-\varepsilon}{\alpha\gamma}\ln n\big)\big)$. By Potter's bound and Proposition 2.9 (c), we get $\ln \psi(s) \sim_{0+} \alpha \ln s$. Lemma 3.4 (ii) then yields $\ln \psi'(s) \sim_{0+} (\alpha-1)\ln s$. Together with (64), this implies that

$$\psi'\big(\mathbb{P}\big(\mathcal{S}(\tau) > \tfrac{1-\varepsilon}{\alpha\gamma}\ln n\big)\big) \sim n^{-(1-\varepsilon)\frac{\alpha-1}{\alpha}+o(1)}. \tag{68}$$

We eventually apply Corollary 3.8, and then (67) and (68), to get that

$$\mathbb{P}\big(\mathcal{S}(\tau) \le \tfrac{1-\varepsilon}{\alpha\gamma}\ln n \;\big|\; |\tau| > \tfrac{\eta n}{a_n}\big) \le C_\mu \exp\Big(-\tfrac{\eta n}{8a_n}\,\psi'\big(\mathbb{P}\big(\mathcal{S}(\tau) > \tfrac{1-\varepsilon}{\alpha\gamma}\ln n\big)\big)\Big) \le \exp\big(-n^{\varepsilon\frac{\alpha-1}{2\alpha}}\big)$$

for all sufficiently large n. The previous upper bound combined with (66) and (67) implies (65).

Upper bound. We want to prove for all $\varepsilon \in (0,1)$ that

$$\mathbb{P}\big(\alpha\gamma \mathcal{S}(\tau) \ge (1+2\varepsilon)\ln n \;\big|\; \#\tau = n\big) \xrightarrow[n\to\infty]{} 0. \tag{69}$$

To that end, we first show

$$\mathbb{P}\big(\alpha\gamma \mathcal{S}(\tau) \ge (1+\varepsilon)\ln n \;\big|\; \#\tau \ge n\big) \xrightarrow[n\to\infty]{} 0. \tag{70}$$

Indeed, we use the following rough bound

$$\mathbb{P}\big(\alpha\gamma\mathcal{S}(\tau)\geq(1+\varepsilon)\ln n\,\big|\,\#\tau\geq n\big)\leq\frac{\mathbb{P}\big(\alpha\gamma\mathcal{S}(\tau)\geq(1+\varepsilon)\ln n\big)}{\mathbb{P}(\#\tau\geq n)}.$$

Then, by Propositions 2.10 (iii) and 2.9 on the one hand, and by (64) on the other hand, we observe

$$\mathbb{P}(\#\tau\geq n)\sim\frac{c_\alpha}{a_n}\sim n^{-\frac{1}{\alpha}+o(1)}\quad\text{and}\quad\mathbb{P}\big(\alpha\gamma\mathcal{S}(\tau)\geq(1+\varepsilon)\ln n\big)\sim n^{-(1+\varepsilon)\frac{1}{\alpha}+o(1)},$$

which implies (70). □

We then recall from (38) that under $\mathbb{P}(\,\cdot\,|\,\#\tau=n)$, τ and its mirror image have the same law. Therefore, (41) in Lemma 3.1 applies with $m=\lfloor\frac{7}{8}n\rfloor$ and one gets, for all sufficiently large n,

$$\mathbb{P}\big(\mathcal{S}(\tau)\geq\tfrac{1+2\varepsilon}{\alpha\gamma}\ln n\,\big|\,\#\tau=n\big)\leq2\mathbb{P}\big(\mathcal{S}(R_{\lfloor\frac{7}{8}n\rfloor}\tau)\geq\tfrac{1+\varepsilon}{\alpha\gamma}\ln n\,\big|\,\#\tau=n\big)+\mathbb{P}\big(\,|\tau|>\tfrac{1}{2}n\,\big|\,\#\tau=n\big).$$

By Proposition 2.11 (i), $\mathbb{P}\big(\,|\tau|>\frac{1}{2}n\,\big|\,\#\tau=n\big)\to0$ when $n\to\infty$. To control the first term of the right-hand side of the previous inequality, we use Proposition 2.11 (ii): to that end, recall from Proposition 2.7 (ii) that $R_{\lfloor\frac{7}{8}n\rfloor}\tau$ is a measurable function of the Lukasiewicz path $(W_k(\tau))_{0\leq k\leq\lfloor\frac{7}{8}n\rfloor}$. Therefore, for all $n\in\mathbb{N}$ and $c\in(0,\infty)$,

$$\begin{aligned}
x_n\;&:=\;\mathbb{P}\big(\,\mathcal{S}(R_{\lfloor\frac{7}{8}n\rfloor}\tau)\geq\tfrac{1+\varepsilon}{\alpha\gamma}\ln n\,\big|\,\#\tau=n\big)\\
&=\;\mathbb{E}\Big[\mathbf{1}_{\left\{\mathcal{S}(R_{\lfloor\frac{7}{8}n\rfloor}\tau)\geq\frac{1+\varepsilon}{\alpha\gamma}\ln n\right\}}D_n^{(7/8)}\big(W_{\lfloor\frac{7}{8}n\rfloor}(\tau)\big)\,\Big|\,\#\tau\geq n\Big]\\
&\leq\;\mathbb{E}\Big[\mathbf{1}_{\{\mathcal{S}(\tau)\geq\frac{1+\varepsilon}{\alpha\gamma}\ln n\}}D_n^{(7/8)}\big(W_{\lfloor\frac{7}{8}n\rfloor}(\tau)\big)\,\Big|\,\#\tau\geq n\Big]\\
&\leq\;c\,\mathbb{P}\big(\mathcal{S}(\tau)\geq\tfrac{1+\varepsilon}{\alpha\gamma}\ln n\,\big|\,\#\tau\geq n\big)+\mathbb{E}\Big[\mathbf{1}_{\left\{D_n^{(7/8)}(W_{\lfloor\frac{7}{8}n\rfloor}(\tau))\geq c\right\}}D_n^{(7/8)}\big(W_{\lfloor\frac{7}{8}n\rfloor}(\tau)\big)\,\Big|\,\#\tau\geq n\Big]
\end{aligned}$$

by (37). We make $n\to\infty$ then $c\to\infty$, and thus, by (70) and (25) in Proposition 2.11, we get $\limsup_{n\to\infty}x_n=0$. This implies (69) and readily completes the proof of Theorem 1.2. ■

4.2 Proof of Theorem 1.3

In all this section, we assume that the critical offspring distribution μ belongs to the domain of attraction of a 1-stable law, and more precisely, we assume that Proposition 2.12 $(a$-$c)$ hold. It is convenient to set

$$\forall x\in[0,\infty),\quad q_x=\mathbb{P}\big(\mathcal{S}(\tau)\geq x\big).$$

Recall from (14) that $\Delta(\tau)$ stands for the maximal out-degree of τ, which is the relevant quantity to consider in the 1-stable cases. Note that $\mathbb{P}(\Delta(\tau)\geq n)\geq\mu([n,\infty))>0$ for all $n\in\mathbb{N}$. We first prove the following estimates.

Lemma 4.1 *Under the same assumption and notations as in Theorem 1.3, the following holds true for all $\varepsilon\in(0,1)$ and all $\kappa\in(0,\infty)$.*

(i) *There is $n_0\in\mathbb{N}$ that depends on ε and κ such that for all integers $n\geq n_0$,*

$$b_nq_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}\leq(\ln n)^{-\kappa}\quad\text{and}\quad b_nq_{(1-\varepsilon)\Upsilon(\frac{1}{b_n})}\geq(\ln n)^{\kappa}.$$

(ii) *Recall from (9) the definition of Λ. Then, $\lim_{n\to\infty}b_n\Lambda(\frac{1}{b_n})q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}=0$.*

(iii) *It holds $\lim_{n\to\infty}n^\kappa\,\mathbb{P}\big(\mathcal{S}(\tau)\leq(1-\varepsilon)\Upsilon\big(\frac{1}{b_n}\big)\,\big|\,\Delta(\tau)>\frac{1}{2}b_n\big)=0$.*

Proof. Let us prove (i). By Proposition 3.12 then Proposition 3.14, we have

$$\Upsilon\big(q_{(1\pm\varepsilon)\Upsilon(\frac{1}{b_n})}\big) \sim (1\pm\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \sim (1\pm\varepsilon)\Upsilon\big(\tfrac{(\ln b_n)^{\mp\kappa}}{b_n}\big).$$

Then, for all sufficiently large n, $\Upsilon\big(q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}\big) > \Upsilon\big(\frac{(\ln b_n)^{-\kappa}}{b_n}\big)$ and $\Upsilon\big(q_{(1-\varepsilon)\Upsilon(\frac{1}{b_n})}\big) < \Upsilon\big(\frac{(\ln b_n)^{\kappa}}{b_n}\big)$, which implies that

$$b_n q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})} \le (\ln b_n)^{-\kappa} \quad \text{and} \quad b_n q_{(1-\varepsilon)\Upsilon(\frac{1}{b_n})} \ge (\ln b_n)^{\kappa}$$

because Υ decreases (see Lemma 3.10). By Proposition 2.12 (iii), b_n varies regularly with index 1 and by Proposition 2.8 (i), $\ln b_n \sim \ln n$, which implies the desired result.

We use similar arguments to prove (ii): by Proposition 3.12 then Proposition 3.14, we get

$$\Upsilon\big(q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}\big) \sim (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \sim (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\Lambda\big(\tfrac{1}{b_n}\big)^{-1-\kappa}\big),$$

so $\Upsilon\big(q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}\big) > \Upsilon\big(\frac{1}{b_n}\Lambda\big(\frac{1}{b_n}\big)^{-1-\kappa}\big)$ for all sufficiently large n, which implies that

$$b_n\Lambda\big(\tfrac{1}{b_n}\big) q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})} \le \Lambda\big(\tfrac{1}{b_n}\big)^{-\kappa}.$$

This implies (ii) since $\lim_{x\to\infty}\Lambda(1/x)=\infty$ by Proposition 3.11 (ii).

To prove (iii), we first use (i) with e.g. $\kappa=2$. Then, Proposition 3.9 implies that

$$n^{\kappa}\mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \Delta(\tau) > \tfrac{1}{2}b_n\big) \le n^{\kappa}\exp\big(-\tfrac{1}{2}b_n q_{(1-\varepsilon)\Upsilon(\frac{1}{b_n})}\big) \le n^{\kappa}\exp\big(-\tfrac{1}{2}(\ln n)^2\big)$$

for all sufficiently large n, which entails the desired result. ■

The proof of Theorem 1.3 is cut into two parts: firstly an upper bound and secondly a lower bound.

Upper bound. We first prove for all $\varepsilon \in (0,1)$ that

$$\mathbb{P}\big(\mathcal{S}(\tau) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \#\tau \ge n\big) \xrightarrow[n\to\infty]{} 0. \tag{71}$$

To that end, we use a direct upper bound, then we combine Lemma 2.13 with Proposition 3.11 (ii):

$$\mathbb{P}\big(\mathcal{S}(\tau) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \#\tau \ge n\big) \le \frac{q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}}{\mathbb{P}(\#\tau \ge n)} \sim \Lambda\big(\tfrac{1}{b_n}\big)\, b_n\, q_{(1+\varepsilon)\Upsilon(\frac{1}{b_n})}.$$

Next, we use Lemma 4.1 (ii) to conclude. ■

Lower bound. We next prove for all $\varepsilon \in (0,1)$ that

$$\mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \#\tau \ge n\big) \xrightarrow[n\to\infty]{} 0. \tag{72}$$

To that end, we first prove that

$$\lim_{n\to\infty} \mathbb{P}\big(\#\tau < n \,\big|\, \Delta(\tau) \ge 2b_n\big) \xrightarrow[n\to\infty]{} 0. \tag{73}$$

Indeed, by (59) in Proposition 3.9 combined with Markov inequality, for all $\lambda \in (0,\infty)$, it holds

$$\mathbb{P}\big(\#\tau < n \,\big|\, \Delta(\tau) \ge 2b_n\big) \le e^{\lambda}\mathbb{E}\big[e^{-\frac{\lambda}{n}\#\tau}\big]^{\lfloor 2b_n\rfloor}.$$

Recall from Proposition 2.7 (iv) that $\mathbb{E}\big[\exp\big(-\frac{\lambda}{n}\#\tau\big)\big]^{\lfloor 2b_n\rfloor} = \mathbb{E}\big[\exp\big(-\frac{\lambda}{n}\mathtt{H}_{\lfloor 2b_n\rfloor}\big)\big]$. The convergence (28) with $x=2$ then implies that $\lim_{n\to\infty}\mathbb{E}\big[e^{-\frac{\lambda}{n}\#\tau}\big]^{\lfloor 2b_n\rfloor} = e^{-2\lambda}$. Thus,

$$\limsup_{n\to\infty}\mathbb{P}\big(\#\tau < n \,\big|\, \Delta(\tau)\geq 2b_n\big) \leq e^{-\lambda} \xrightarrow[\lambda\to\infty]{} 0$$

which entails (73). □

We next prove that for all $c\in(0,\infty)$,

$$\limsup_{n\to\infty}\frac{\mathbb{P}\big(\Delta(\tau)\geq cb_n\big)}{\mathbb{P}(\#\tau\geq n)} \leq \frac{4}{c}. \tag{74}$$

Indeed, we first prove the case where $c=2$. Observe that

$$u(n) := \frac{\mathbb{P}\big(\Delta(\tau)\geq 2b_n\big)}{\mathbb{P}(\#\tau\geq n)} = \mathbb{P}\big(\Delta(\tau)\geq 2b_n \,\big|\, \#\tau\geq n\big) + u(n)\mathbb{P}\big(\#\tau<n\,\big|\,\Delta(\tau)\geq 2b_n\big).$$

By (73), for all sufficiently large n, we get $\mathbb{P}\big(\#\tau<n\,\big|\,\Delta(\tau)\geq 2b_n\big)\leq\frac{1}{2}$. It therefore implies $u(n)\leq 2\mathbb{P}\big(\Delta(\tau)\geq 2b_n\,\big|\,\#\tau\geq n\big)\to 2\mathbb{P}(J\geq 2)=1$ by (31) in Proposition 2.15. This proves (74) when $c=2$. Then, recall from Proposition 2.12 (iii) that (b_n) is 1-regularly varying. Then for all sufficiently large n, we get $cb_n\geq 2b_{\lfloor\frac{1}{4}cn\rfloor}$ and thus

$$\frac{\mathbb{P}\big(\Delta(\tau)\geq cb_n\big)}{\mathbb{P}(\#\tau\geq n)} \leq \frac{\mathbb{P}\big(\Delta(\tau)\geq 2b_{\lfloor\frac{1}{4}cn\rfloor}\big)}{\mathbb{P}(\#\tau\geq n)} = u(\lfloor\tfrac{1}{4}cn\rfloor)\frac{\mathbb{P}(\#\tau\geq\lfloor\frac{1}{4}cn\rfloor)}{\mathbb{P}(\#\tau\geq n)}.$$

This implies the desired result since we know from Lemma 2.13 that the sequence $n\mapsto\mathbb{P}(\#\tau\geq n)$ varies regularly with exponent -1, and since $\limsup_{n\to\infty}u(\lfloor\frac{1}{4}cn\rfloor)\leq 1$ as proved above. □

We complete the proof of (72) as follows.

$$\begin{aligned} v_n &:= \mathbb{P}\big(\mathcal{S}(\tau)\leq(1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\,\big|\,\#\tau\geq n\big)\\ &\leq \frac{\mathbb{P}\big(\mathcal{S}(\tau)\leq(1-\varepsilon)\Upsilon\big(\frac{1}{b_n}\big)\,;\,\Delta(\tau)>\frac{1}{2}b_n\big)}{\mathbb{P}(\#\tau\geq n)} + \mathbb{P}\big(\Delta(\tau)\leq\tfrac{1}{2}b_n\,\big|\,\#\tau\geq n\big)\\ &\leq \frac{\mathbb{P}\big(\Delta(\tau)>\frac{1}{2}b_n\big)}{\mathbb{P}(\#\tau\geq n)}\mathbb{P}\big(\mathcal{S}(\tau)\leq(1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\,\big|\,\Delta(\tau)>\tfrac{1}{2}b_n\big) + \mathbb{P}\big(\Delta(\tau)\leq\tfrac{1}{2}b_n\,\big|\,\#\tau\geq n\big). \end{aligned}$$

This first term of the above right-hand side converges to 0 by (74) with $c=1/2$ and by Lemma 4.1 (iii). Moreover, (31) in Proposition 2.15 asserts that $\lim_{n\to\infty}\mathbb{P}\big(\Delta(\tau)\leq\frac{1}{2}b_n\,\big|\,\#\tau\geq n\big)=0$. This completes the proof of (72), and of Theorem 1.3. ■

4.3 Proof of Theorem 1.4

Throughout this section, we assume there exists a function L that varies slowly at ∞ such that $\mu(n)\sim n^{-2}L(n)$. This implies that $\mu([n,\infty))\sim n^{-1}L(n)$ so, by Proposition 2.12, μ belongs to the domain of attraction of a 1-stable law. The proof of Theorem 1.4 is separated into two parts: firstly a lower bound and secondly an upper bound.

Lower bound. We prove for all $\varepsilon\in(0,1)$ that

$$\mathbb{P}\big(\mathcal{S}(\tau)\leq(1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\,\big|\,\#\tau=n\big)\xrightarrow[n\to\infty]{}0. \tag{75}$$

Indeed, observe that

$$\begin{aligned} v_n &:= \mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \#\tau = n\big) \\ &\le \frac{\mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\,;\, \Delta(\tau) > \frac{1}{2}b_n\big)}{\mathbb{P}(\#\tau = n)} + \mathbb{P}\big(\Delta(\tau) \le \tfrac{1}{2}b_n \,\big|\, \#\tau = n\big) \\ &\le \frac{\mathbb{P}\big(\Delta(\tau) > \frac{1}{2}b_n\big)}{n\mathbb{P}(\#\tau = n)}\, n\,\mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \Delta(\tau) > \tfrac{1}{2}b_n\big) + \mathbb{P}\big(\Delta(\tau) \le \tfrac{1}{2}b_n \,\big|\, \#\tau = n\big). \end{aligned}$$

Then recall from (30) in Lemma 2.13 that $n\mathbb{P}(\#\tau = n) \sim \mathbb{P}(\#\tau \ge n)$. Thus, by (74), we get

$$\limsup_{n\to\infty} \frac{\mathbb{P}\big(\Delta(\tau) > \frac{1}{2}b_n\big)}{n\mathbb{P}(\#\tau = n)} \le 8.$$

By Lemma 4.1 (iii), $\lim_{n\to\infty} n\mathbb{P}\big(\mathcal{S}(\tau) \le (1-\varepsilon)\Upsilon\big(\frac{1}{b_n}\big) \,\big|\, \Delta(\tau) > \frac{1}{2}b_n\big) = 0$. Then, (32) in Proposition 2.15 applies and asserts that $\lim_{n\to\infty} \mathbb{P}\big(\Delta(\tau) \le \frac{1}{2}b_n \,\big|\, \#\tau = n\big) = 0$, which finally implies (75). ■

Upper bound. We finally show for all $\varepsilon \in (0,1)$ that

$$\mathbb{P}\big(\mathcal{S}(\tau) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) \,\big|\, \#\tau = n\big) \xrightarrow[n\to\infty]{} 0. \tag{76}$$

To prove (76), we use the result of Kortchemski & Richier [26, Theorem 21] that is recalled in Proposition 2.16 and that shows that the law of the Lukasiewicz path $(W_j(\tau))_{0\le j\le n}$ of τ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$ is close in total variation distance to the law of $(Z_j^{(n)})_{0\le j\le n}$ as defined in (34).

First, let us briefly recall the definition of $Z^{(n)}$: let $(W_n)_{n\in\mathbb{N}}$ be a left-continuous random walk starting at 0 and whose jump distribution is given by (7). We recall from (33) the two notations $I_n = -\min_{0\le j\le n-1} W_j$ and $\sigma_n = \inf\{0 \le k \le n-1 : W_k = -I_n\}$. Then,

$$Z_j^{(n)} = W_{\sigma_n + j} + I_n \text{ if } 0 \le j < n - \sigma_n \quad \text{and} \quad Z_j^{(n)} = I_n - 1 + W_{j-(n-\sigma_n)} \text{ if } n \ge j \ge n - \sigma_n.$$

Next, we interpret Proposition 2.16 in terms of trees, i.e. we view $Z^{(n)}$ as the Lukasiewicz path of a random tree $\tau^{(n)}$. Namely, observe that $Z_0^{(n)} = 0$, $Z_n^{(n)} = -1$, and the other values of $Z^{(n)}$ are nonnegative. Plus, maybe except at the cutting time $n-\sigma_n-1$ when $Z^{(n)}_{n-\sigma_n} - Z^{(n)}_{n-\sigma_n-1} = -1 - W_{n-1}$, the jumps of $Z^{(n)}$ are larger or equal to -1. Consequently, if $W_{n-1} \le 0$ then Proposition 2.7 (i) applies and $Z^{(n)}$ is the Lukasiewicz path associated with a tree $\tau^{(n)}$. If $W_{n-1} > 0$ then we take $\tau^{(n)}$ equal to the *star-tree with* $n-1$ *leaves*, which we denote by $\bigstar_n$ (or any tree with n vertices).

$$\text{If } W_{n-1} \le 0 \text{ then } (W_j(\tau^{(n)}))_{0\le j\le n} = (Z_j^{(n)})_{0\le j\le n} \quad \text{and} \quad \text{if } W_{n-1} > 0 \text{ then } \tau^{(n)} = \bigstar_n. \tag{77}$$

Moreover, observe that Proposition 2.12 (c) easily implies that $\mathbb{P}\big(\tau^{(n)} = \bigstar_n\big) = \mathbb{P}(W_{n-1} > 0) \to 0$. Thus, thanks to Proposition 2.16, in order to prove (76), we only need to show

$$\mathbb{P}\big(\tau^{(n)} \ne \bigstar_n\,;\, \mathcal{S}(\tau^{(n)}) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\big) \xrightarrow[n\to\infty]{} 0. \tag{78}$$

To that end, we discuss a decomposition of $\tau^{(n)}$ at the 'cutting time' $n-\sigma_n-1$ and we recall from (22) in Proposition 2.7 the existence of an i.d.d. sequence $(\tau_p)_{p\in\mathbb{N}}$ of $\mathsf{GW}(\mu)$-trees given by

$$\forall p \in \mathbb{N}, \quad (p + W_{\mathtt{H}_p + j})_{0\le j\le \mathtt{H}_{p+1} - \mathtt{H}_p} = \big(W_j(\tau_p)\big)_{0\le j\le \#\tau_p},$$

where $\mathtt{H}_p = \inf\{j \in \mathbb{N} : W_j = -p\}$. Let us denote by $u_0^{(n)} = \varnothing < u_1^{(n)} < \ldots < u_{n-1}^{(n)}$ the vertices of $\tau^{(n)}$ listed in lexicogaphic order. From the definition of $Z^{(n)}$ and thanks to Proposition 2.7 (ii), we first derive that if $W_{n-1} \le 0$ then

$$R_{n-\sigma_n-1}\big(\tau^{(n)}\big) = R_{n-\sigma_n-1}\big(\tau_{I_n}\big) \quad \text{and} \quad k_{u^{(n)}_{n-\sigma_n-1}}(\tau^{(n)}) = -W_{n-1}, \tag{79}$$

where we recall from (20) in Proposition 2.7 that $R_m(t) = R_m t$ stands for the tree consisting of the first $m+1$ vertices of t taken in lexicographic order. We now look at the subtrees that are grafted either at $u^{(n)}_{n-\sigma_n-1}$ or to the right of its ancestral line: namely, the subtrees that are grafted at the following set of vertices

$$B = \big\{v \in \tau^{(n)} \ : \ \overleftarrow{v} \preceq u^{(n)}_{n-\sigma_n-1} \text{ and } v > u^{(n)}_{n-\sigma_n-1}\big\}.$$

If we denote by $v(0) < v(1) < \ldots < v(\#B-1)$ the vertices of B listed in lexicographic order, then by definition of $Z^{(n)}$ and by definition of the trees τ_p as recalled above, we get that if $W_{n-1} \le 0$ then

$$\theta_{v(p)}\tau^{(n)} = \tau_p, \quad 0 \le p \le I_n - 1 = \#B - 1. \tag{80}$$

To prove (78), we then use (39) in Lemma 3.1 (i) that asserts on the event $\{W_{n-1} \le 0\}$ that

$$\mathcal{S}\big(\tau^{(n)}\big) \le 1 + \max\big(\mathcal{S}\big(R_{n-\sigma_n-1}\big(\tau^{(n)}\big)\big), \max_{v\in B} \mathcal{S}\big(\theta_v\tau^{(n)}\big)\big).$$

Then, the identities (79) and (80), together with the monotony property (37) of $\mathcal{S}$, yield that

$$\mathcal{S}\big(\tau^{(n)}\big) \le 1 + \max_{0\le p\le I_n} \mathcal{S}\big(\tau_p\big).$$

Therefore, we obtain the following:

$$\begin{aligned} x_n &:= \mathbb{P}\Big(\tau^{(n)} \neq \bigstar_n \, ; \, \mathcal{S}(\tau^{(n)}) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big)\Big) \\ &\le \mathbb{P}\Big(\max_{0\le p\le \lfloor 2b_n\rfloor} \mathcal{S}\big(\tau_p\big) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) - 1\Big) + \mathbb{P}\big(I_n > 2b_n\big) \\ &\le (2b_n+1)\,\mathbb{P}\Big(\mathcal{S}(\tau) \ge (1+\varepsilon)\Upsilon\big(\tfrac{1}{b_n}\big) - 1\Big) + \mathbb{P}\big(\tfrac{1}{n}\mathtt{H}_{\lfloor 2b_n\rfloor} \le 1\big), \end{aligned}$$

since the τ_p are $\mathsf{GW}(\mu)$-trees. Then, recall from (28) that $\frac{1}{n}\mathtt{H}_{\lfloor 2b_n\rfloor} \to 2$ in probability, which implies that $\mathbb{P}\big(\frac{1}{n}\mathtt{H}_{\lfloor 2b_n\rfloor} \le 1\big) \to 0$. Then, Lemma 4.1 (i) asserts that $b_n\mathbb{P}\big(\mathcal{S}(\tau) \ge (1+\varepsilon)\Upsilon\big(\frac{1}{b_n}\big) - 1\big) \to 0$, which finally implies (78) and which completes the proof of Theorem 1.4. ■

We conclude this article by providing an asymptotic equivalent of the rescaling sequence $\Upsilon\big(\frac{1}{b_n}\big)$, that gives the order of $\mathcal{S}(\tau)$ under $\mathbb{P}(\,\cdot\,|\,\#\tau = n)$, in the three cases considered in Example 3.13.

Example 4.2 We recall from Proposition 2.12 (iii) that b_n varies regularly with index 1, which implies $\ln b_n \sim \ln n$ by Proposition 2.8 (i).

(a) Let $\kappa \in (0,\infty)$. We consider the case where $L(x) = (\ln x)^{-1-\kappa}$ for all $x \in (0,\infty)$. Recall from Example 3.13 that $\Upsilon(s) \sim_{0^+} (\ln\frac{1}{s})/(\ln\ln\frac{1}{s})$. Therefore, we easily get

$$\Upsilon\big(\tfrac{1}{b_n}\big) \sim \frac{\ln n}{\ln\ln n}.$$

(b) Let $\kappa \in (0,1)$. We next consider the case where $L(x) = \exp(-(\ln x)^\kappa)$ for all $x \in (0,\infty)$. Recall from Example 3.13 that $\Upsilon(s) \sim_{0^+} \frac{1}{1-\kappa}(\ln\frac{1}{s})/(\ln\ln\frac{1}{s})$. Therefore, we easily get

$$\Upsilon\big(\tfrac{1}{b_n}\big) \sim \frac{1}{1-\kappa}\cdot\frac{\ln n}{\ln\ln n}.$$

(c) Let us finally consider the case where $L(x) = \exp(-\ln x/\ln\ln x)$ for all $x \in (e^e,\infty)$. Recall from Example 3.13 that $\Upsilon(s) \sim_{0^+} (\ln\frac{1}{s})/(\ln\ln\ln\frac{1}{s})$. Therefore, we easily get

$$\Upsilon\big(\tfrac{1}{b_n}\big) \sim \frac{\ln n}{\ln\ln\ln n}.$$

□

Acknowledgements I am deeply indebted to my Ph.D. advisor Thomas Duquesne for introducing me to the Horton-Strahler number, for suggesting me several problems about it, and for many insightful discussions. I am especially grateful for his help to prove Proposition 3.11 and to improve the quality of this paper. I warmly thank Guillaume Boutoille et Yoan Tardy for a stimulating discussion about the correct order of magnitude that should appear in the Cauchy regime. Many thanks are due to Quentin Berger for some feedback about the proof of the estimate (29). I also thank Igor Kortchemski for some precisions about his work.

References

[1] Romain Abraham and Jean-François Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electronic Journal of Probability*, 19:1 – 19, 2014.

[2] David Aldous and Jim Pitman. Tree-valued Markov chains derived from Galton-Watson processes. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 34(5):637 – 686, 1998.

[3] Sameer Bamufleh, Abdullah Al-Wagdany, Amro Elfeki, and Anis Chaabani. Developing a geomorphological instantaneous unit hydrograph (GIUH) using equivalent Horton-Strahler ratios for flash flood predictions in arid regions. *Geomatics, Natural Hazards and Risk*, 11(1):1697 – 1723, 2020.

[4] Quentin Berger. Notes on Random Walks in the Cauchy Domain of Attraction. *Probability Theory and Related Fields*, 175:1 – 44, 2019.

[5] Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels. *Regular Variation*. Number 27 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1989.

[6] Anna Brandenberger, Luc Devroye, and Tommy Reddad. The Horton–Strahler number of conditioned Galton–Watson trees. *Electronic Journal of Probability*, 26:1 – 29, 2021.

[7] Gregory A. Burd, Edward C. Waymire, and Ronald D. Winn. A Self-Similar Invariance of Critical Binary Galton-Watson Trees. *Bernoulli*, 6(1):1 – 21, 2000.

[8] Sagar R. Chavan and Venkata V. Srinivas. Effect of DEM source on equivalent Horton–Strahler ratio based GIUH for catchments in two Indian river basins. *Journal of Hydrology*, 528:463 – 489, 2015.

[9] Luc Devroye and Paul Kruszewski. A note on the Horton-Strahler number for random trees. *Information Processing Letters*, 56(2):95 – 99, 1995.

[10] Michael Drmota and Helmut Prodinger. The Register Function for T-Ary Trees. *ACM Transactions on Algorithms*, 2(3):318 – 334, 2006.

[11] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *The Annals of Probability*, 31(2):996 – 1027, 2003.

[12] Thomas Duquesne. An elementary proof of Hawkes's conjecture on Galton-Watson trees. *Electronic Communications in Probability*, 14:151 – 164, 2009.

[13] Thomas Duquesne and Matthias Winkel. Hereditary tree growth and Lévy forests. *Stochastic Processes and their Applications*, 129(10):3690 – 3747, 2019.

[14] Javier Esparza, Michael Luttenberger, and Maximilian Schlund. History of Strahler Numbers — with a Preface. International Conference on Language and Automata Theory and Applications, 2014, 2016. Available online at https://archive.model.in.tum.de/um/bibdb/esparza/latarevised16.pdf.

[15] Joanna Fac-Beneda. Fractal structure of the Kashubian hydrographic system. *Journal of Hydrology*, 488:48 – 54, 2013.

[16] William Feller. *An Introduction to Probability Theory and Its Applications. Vol. II.* . Second. John Wiley & Sons Inc., New York, 1971.

[17] Philippe Flajolet, Jean-Claude Raoult, and Jean E. Vuillemin. The number of registers required for evaluating arithmetic expressions. *Theoretical Computer Science*, 9(1):99 – 125, 1979.

[18] Jean-Francois Le Gall and Yves Le Jan. Branching processes in levy processes: The exploration process. *The Annals of Probability*, 26(1):213 – 252, 1998.

[19] Jean-François Le Gall. Random trees and applications. *Probability Surveys*, 2:245 – 311, 2005.

[20] Anders Grimvall. On the Convergence of Sequences of Branching Processes. *The Annals of Probability*, 2(6):1027 – 1045, 1974.

[21] Robert E. Horton. Erosional development of streams and their drainage basins ; hydrophysical approach to quantitative morphology. *GSA Bulletin*, 56(3):275 – 370, 1945.

[22] Rainer Kemp. The average number of registers needed to evaluate a binary tree optimally. *Acta Informatica*, 11:363 – 372, 1979.

[23] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 22(4):425 – 487, 1986.

[24] Robin Khanfir. Fluctuations of the Horton-Strahler number of stable Galton-Watson trees. *Work in preparation*.

[25] Igor Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton–Watson trees. In Catherine Donati-Martin, Antoine Lejay, and Alain Rouault, editors, *Séminaire de Probabilités XLV*, Lecture Notes in Mathematics, pages 537 – 558. Springer International Publishing, Heidelberg, 2013.

[26] Igor Kortchemski and Loïc Richier. Condensation in critical Cauchy Bienaymé-Galton–Watson trees. *The Annals of Applied Probability*, 29(3):1837 – 1877, 2019.

[27] Yevgeniy Kovchegov and Ilya Zaliapin. Random self-similar trees: A mathematical theory of Horton laws. *Probability Surveys*, 17:1 – 213, 2020.

[28] Yevgeniy Kovchegov and Ilya Zaliapin. Invariance and attraction properties of Galton–Watson trees. *Bernoulli*, 27(3):1789 – 1823, 2021.

[29] Russell Lyons, Robin Pemantle, and Yuval Peres. Conceptual Proofs of L Log L Criteria for Mean Behavior of Branching Processes. *The Annals of Probability*, 23(3):1125 – 1138, 1995.

[30] Scott D. Peckham. New Results for Self-Similar Trees with Applications to River Networks. *Water Resources Research*, 31(4):1023 – 1029, 1995.

[31] Anatoliy V. Skorokhod. Limit Theorems for Stochastic Processes with Independent Increments. *Theory of Probability & Its Applications*, 2(2):138 – 171, 1957.

[32] R. S. Slack. A branching process with mean one and possibly infinite variance. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 9:139 – 145, 1968.

[33] Arthur N. Strahler. Hypsometric (area-altitude) analysis of erosional topography. *GSA Bulletin*, 63(11):1117 – 1142, 1952.

[34] Wim Vervaat. A Relation between Brownian Bridge and Brownian Excursion. *The Annals of Probability*, 7(1):143 – 149, 1979.

[35] Xavier Viennot. Trees. In *Mots, mélanges offert à M.P. Schützenberger*. Hermès, Paris, 1990. Available online at http://www.xaviernviennot.org/xavier/.