Electronic Journal of Probability

Electron. J. Probab. **22** (2017), no. 21, 1–25.
ISSN: 1083-6489 DOI: 10.1214/17-EJP31

Asymptotics of heights in random trees constructed by aggregation

Bénédicte Haas*

Abstract

To each sequence (a_n) of positive real numbers we associate a growing sequence (T_n) of continuous trees built recursively by gluing at step n a segment of length a_n on a uniform point of the pre–existing tree, starting from a segment T_1 of length a_1. Previous works [5, 10] on that model focus on the influence of (a_n) on the compactness and Hausdorff dimension of the limiting tree. Here we consider the cases where the sequence (a_n) is regularly varying with a non–negative index, so that the sequence (T_n) explodes. We determine the asymptotics of the height of T_n and of the subtrees of T_n spanned by the root and ℓ points picked uniformly at random and independently in T_n, for all $\ell \in \mathbb{N}$.

Keywords: random trees; line–breaking; uniform recursive trees.
AMS MSC 2010: 05C05; 60J05.
Submitted to EJP on June 23, 2016, final version accepted on January 26, 2017.
Supersedes arXiv:1606.06536.

1 Introduction

A well–known construction of the Brownian continuum random tree presented by Aldous in the first of his series of papers [2, 3, 4] works as follows. Consider a Poisson point process on $\mathbb{R}_+$ with intensity $t\mathrm{d}t$ and "break" the half–line $\mathbb{R}_+$ at each point of the process. This gives an ordered sequence of closed segments with random lengths. Take the first segment and glue on it the second segment at a point chosen uniformly at random (i.e. according to the normalized length measure). Then consider the continuous tree formed by these two first segments and glue on it the third segment at a point chosen uniformly at random. And so on. This gluing procedure, called the *line–breaking construction* by Aldous [2], gives in the limit a version of the Brownian CRT.

We are interested in a generalization of this construction, starting from any sequence of positive terms

$$(a_n, n \geq 1).$$

*Université Paris 13, Sorbonne Paris Cité LAGA, CNRS (UMR 7539) 93430 Villetaneuse, France. E-mail: haas@math.univ-paris13.fr

For each n, we let b_n denote a closed segment of length a_n. The construction process then holds as above: we start with $T_1 := \mathsf{b}_1$ and then recursively glue the segment b_n on a point chosen uniformly on T_{n-1}, for all $n \geq 1$. The trees T_n are viewed as metric spaces, once endowed with their length metrics, which will be noted d in all cases. This yields in the limit a random real tree obtained as the completion of the increasing union of the trees T_n,

$$\mathcal{T} := \overline{\cup_{n\geq 1} T_n}$$

that may be infinite. We let d denote its metric as well, and decide to *root* this tree at one of the two extremities of b_1.

This model has been recently studied by Curien and Haas [10] and Amini et al. [5]. The paper [10] gives necessary conditions and sufficient conditions on the sequence (a_n) for $\mathcal{T}$ to be compact (equivalently bounded) and studies its Hausdorff dimension. Typically, if

$$a_n \leq n^{\alpha+\circ(1)} \quad \text{and} \quad a_1 + \ldots + a_n = n^{1+\alpha+\circ(1)} \quad \text{for some } \alpha < 0,$$

then almost surely the tree $\mathcal{T}$ is compact and its set of leaves has Hausdorff dimension $1/|\alpha|$, which ensures that the tree itself has Hausdorff dimension $\max(1, 1/|\alpha|)$. This, as an example, retrieves the compactness of the Brownian CRT and that its Hausdorff dimension is 2. On the other hand, the tree $\mathcal{T}$ is almost surely unbounded as soon as the sequence (a_n) does not converge to 0. The issue of finding an exact condition on (a_n) for $\mathcal{T}$ to be bounded is still open. However, Amini et al. [5] obtained an exact condition for $\mathcal{T}$ to be bounded, provided that (a_n) is *non–increasing*. In that case, almost surely,

$$\mathcal{T} \text{ is bounded if and only if } \sum_{i\geq 1} i^{-1} a_i < \infty.$$

Concerning related works, Ross and Wen [19] have recently shown that when the $a_n, n \geq 1$ are random lengths obtained by breaking the half–line with a Poisson point process with intensity $t^{\ell}\mathrm{d}t$ for some positive integer ℓ, the corresponding trees arise as scaling limits of some growing combinatorial trees constructed inhomogeneously, thus generalizing the convergence of trees constructed via Rémy's algorithm towards the Brownian CRT. There are also related constructions with different gluing rules. Sénizergues [20] studies a generalization of the above (a_n)–model where the segments are replaced by d–dimensional independent random metric measured spaces ($d \in (0,\infty)$) and the gluing rules depend both on the diameters and the measures of the metric spaces. He shows an unexpected and intriguing Hausdorff dimension. In other directions, Addario–Berry, Broutin and Goldschmidt [1] provide a line–breaking construction of the continuum limit of critical random graphs, extending Aldous' line–breaking construction to random real trees with vertex identifications; while Goldschmidt and Haas [14] propose a construction of the stable Lévy trees introduced by Duquesne, Le Gall and Le Jan [13, 15] that generalizes Aldous' line–breaking construction to this class of trees (except in the Brownian case, the stable Lévy trees are not binary and the gluing procedure is then slightly more complex). See also [8], [12] and [18] for other related models.

The aim of the present paper is to examine the cases where the (a_n)–model obviously leads to an unbounded tree and we will almost always assume that

$$\text{the sequence } (a_n) \text{ is regularly varying with index } \alpha \geq 0.$$

We recall that this means that for all $c > 0$,

$$\frac{a_{\lfloor cn \rfloor}}{a_n} \underset{n\to\infty}{\longrightarrow} c^{\alpha},$$

the prototype example being the sequence (n^{α}). We refer to Bingham et al. [7] for background on that topic. Our goal is to understand how the tree T_n then grows as $n \to \infty$. To that end, we will study the asymptotic behavior of the height of a typical point of T_n and of the height of T_n. We will see that in general these heights do not grow at the same rate. We will also complete the study of the height of a typical point first by providing a functional convergence, and second by studying the behavior of the subtrees of T_n spanned by the root and ℓ points picked uniformly at random and independently in T_n, for all $\ell \in \mathbb{N}$.

Height of a typical point and height of T_n. We are interested in the asymptotic behavior of the following quantities:

- D_n: height of a typical point, i.e. given T_n, we pick $X_n \in T_n$ uniformly at random in T_n and let
$$D_n = d(X_n, \mathsf{root})$$
be its distance to the root;
- the height of the tree:
$$H_n = \max_{v \in T_n} d(v, \mathsf{root}).$$

In the particular case where all the lengths a_n are identical, the sequence (T_n) can be coupled with a growing sequence of uniform recursive trees with i.i.d. uniform $(0,1)$ lengths on their edges. This is explained in Section 5. The asymptotic behavior of the height of a uniform vertex and the height of a random recursive tree *without* edge lengths (i.e. endowed with the graph distance) are well–known, [11, 16, 17]. From this and the strong law of large number, we immediately get the asymptotic of D_n. The behavior of H_n is less obvious. However, Broutin and Devroye [9] develop the material to study the height of random recursive trees with i.i.d. edge lengths, using the underlying branching structure and large deviation techniques. From this, we will deduce that:

Theorem 1.1. *If $a_n = 1$ for all $n \geq 1$,*
$$\frac{D_n}{\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{2} \qquad \text{and} \qquad \frac{H_n}{\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{e^{\beta^*}}{2\beta^*},$$
where β^* is the unique solution in $(0,\infty)$ to the equation $2(e^{\beta}-1) = \beta e^{\beta}$. Approximately, $\beta^* \sim 1.594$ and $e^{\beta^*}/2\beta^* \sim 1.544$.

This will be carried out in Section 5. Our main contribution concerns the cases where the index of regular variation α is strictly positive. In that case we introduce a random variable $\xi_{(\alpha)}$ characterized by its Laplace transform $\mathbb{E}[\exp(\lambda \xi_{(\alpha)})] = \exp(\phi_{(\alpha)}(\lambda))$, $\lambda \in \mathbb{R}$ where
$$\phi_{(\alpha)}(\lambda) = \frac{\alpha+1}{\alpha}\int_0^1 \big(\exp(\lambda u) - 1\big)\frac{1-u}{u}\mathrm{d}u = \frac{\alpha+1}{\alpha}\sum_{k\geq 1}\frac{\lambda^k}{(k+1)!k}. \tag{1.1}$$

The Lévy–Khintchine formula ensures that $\xi_{(\alpha)}$ is infinitely divisible. Note also that $\xi_{(\alpha)}$ is stochastically decreasing in α. Our main result is:

Theorem 1.2. *Assume that (a_n) is regularly varying with index $\alpha > 0$. Then,*

(i)
$$\frac{D_n}{a_n} \xrightarrow[n\to\infty]{(\mathrm{d})} \xi_{(\alpha)}$$

(ii)
$$\frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \xrightarrow[n\to\infty]{\mathrm{a.s.}} 1.$$

More precisely, in (i), $\mathbb{E}[\exp(\lambda a_n^{-1} D_n)]$ *converges to* $\mathbb{E}[\exp(\lambda \xi_{(\alpha)})]$ *for all* $\lambda \in \mathbb{R}$*, which in particular implies the convergence of all positive moments.*

The proof of (i) is undertaken in Section 2.2 and relies on the powerful observation from [10] that D_n can be written as the sum of i.i.d. random variables. The proof of (ii), and in particular of the lower bound, is more intricate. It relies on the *second moment method* and requires to get the joint distribution of the paths from the root to two points marked independently, uniformly in the tree T_n (established in Section 3.1) as well as precise deviations bounds for the convergence (i) (established in Section 2.2). The core of the proof of (ii) is undertaken in Section 4.

The two previous statements on the asymptotic behavior of D_n can actually be grouped together and slightly generalized as follows:

Proposition 1.3. *Assume that* (a_n) *is regularly varying with index* $\alpha \geq 0$*. Then,*

$$\frac{D_n}{\sum_{i=1}^n i^{-1} a_i} \underset{n\to\infty}{\longrightarrow} \begin{cases} \alpha \xi_{(\alpha)} & \text{if } \alpha > 0 \quad (\textit{convergence in distribution}) \\ \dfrac{1}{2} & \text{if } \alpha = 0 \text{ and } \sum_{i=1}^{\infty} i^{-1} a_i = \infty \quad (\textit{convergence in probability}) \\ D_\infty & \text{if } \alpha = 0 \text{ and } \sum_{i=1}^{\infty} i^{-1} a_i < \infty \quad (\textit{convergence in distribution}) \end{cases}$$

where D_∞ *denotes a positive random variable with finite expectation.*

This will be explained in the remark around (2.8) in Section 2.2.

Height of the n–th leaf and height of a uniform leaf. In the recursive construction of (T_n), we can label the leaves $L_1, L_2, \ldots$ by order of appearance, so that the leaf L_n belongs to the segment b_n. We then let $L_{n,\star}$ denote a leaf chosen uniformly at random amongst the n leaves of T_n. Theorem 1.2 (i) implies that when (a_n) varies regularly with index $\alpha > 0$,

$$\frac{d(L_n, \mathsf{root})}{a_n} \xrightarrow[n\to\infty]{(\mathrm{d})} 1 + \xi_{(\alpha)} \quad \text{ and } \quad \frac{d(L_{n,\star}, \mathsf{root})}{a_n} \xrightarrow[n\to\infty]{(\mathrm{d})} (1 + \xi_{(\alpha)}) U^\alpha, \tag{1.2}$$

where U is uniform on $(0, 1)$, independent of $\xi_{(\alpha)}$. The first convergence is simply due to the fact that the distance $d(L_n, \mathsf{root})$ is distributed as $a_n + D_{n-1}$, since the segment b_n is inserted on a uniform point of T_{n-1}. The second convergence is explained in Section 2.2. When $a_n = 1$ for all n, $d(L_n, \mathsf{root})$ and $d(L_{n,\star}, \mathsf{root})$ both divided by $\ln(n)$ converge to $1/2$, almost surely and in probability respectively (see Section 5).

Functional convergence. The convergence of the height of a typical point can actually be improved into a functional convergence when the index of regular variation is strictly positive. As above, let X_n be a point picked uniformly in T_n and for each positive integer $k \leq n$, let $X_n(k)$ denote its projection onto T_k. Let then

$$D_n(k) := d\left(X_n(k), \mathsf{root}\right), \quad 1 \leq k \leq n$$

be the non–decreasing sequence of distances of these branch–points to the root. If a climber decides to climb from the root to the typical point X_n at speed 1, $D_n(k)$ is the time he will spend in T_k. The proof of Theorem 1.2 (i) can be adapted to get the behavior as $n \to \infty$ of the sequence $(D_n(k), 1 \leq k \leq n)$. To do so, introduce for $\alpha > 0$ the càdlàg Markov process with independent, positive increments defined by

$$\xi_{(\alpha)}(t) := \sum_{t_i \leq t} v_i, \quad t \geq 0, \tag{1.3}$$

where (t_i, v_i) is a Poisson point process with intensity $(\alpha+1)t^{-\alpha-1}\mathbb{1}_{\{v\leq t^\alpha\}}\mathrm{d}t\mathrm{d}v$ on $(0,\infty)^2$. (We note that $\xi_{(\alpha)}(1)$ is distributed as the r.v. $\xi_{(\alpha)}$ defined via (1.1).) This process is α–self-similar, in the sense that for all $a>0$,

$$\big(\xi_{(\alpha)}(at), t\geq 0\big) \overset{(\mathrm{d})}{=} \big(a^\alpha \xi_{(\alpha)}(t), t\geq 0\big).$$

Proposition 1.4. *If (a_n) is regularly varying with index $\alpha>0$,*

$$\left(\frac{D_n\left(\lfloor nt \rfloor\right)}{a_n}, 0\leq t\leq 1\right) \xrightarrow[n\to\infty]{(\mathrm{d})} \big(\xi_{(\alpha)}(t), 0\leq t\leq 1\big)$$

for the Skorokhod topology on $D([0,1],\mathbb{R}_+)$, the set of càdlàg functions from $[0,1]$ to $\mathbb{R}_+$.

This is proved in Section 2.3.

Gromov–Prohorov–type convergence. Last, fix ℓ a positive integer, and given T_n, let $X_n^{(1)},\dots,X_n^{(\ell)}$ be ℓ points picked independently and uniformly at random in T_n. Our goal is to describe the asymptotic behavior of $T_n(\ell)$, the subtree of T_n spanned by these ℓ marked points and the root. To that end, for all $1\leq i,j\leq \ell$, we denote by $\mathsf{B}_n^{(i,j)}$ the point in $T_n(\ell)$ at which the paths from the root to $X_n^{(i)}$ and from the root to $X_n^{(j)}$ separate, with the convention that $\mathsf{B}_n^{(i,j)} = X_n^{(i)}$ when $X_n^{(i)}$ belongs to the path from the root to $X_n^{(j)}$. For regularly varying sequences of lengths (a_n), the tree $T_n(\ell)$ appropriately rescaled converges to a “star–tree” with ℓ branches with random i.i.d. lengths. More precisely:

Proposition 1.5. (i) *Assume that (a_n) is regularly varying with index $\alpha>0$. Then,*

$$\left(\left(\frac{d\big(X_n^{(i)},\mathsf{root}\big)}{a_n}, 1\leq i\leq \ell\right), \frac{\max_{1\leq i\neq j\leq \ell} d\big(\mathsf{B}_n^{(i,j)},\mathsf{root}\big)}{a_n}\right) \xrightarrow[n\to\infty]{(\mathrm{d})} \left(\left(\xi_{(\alpha)}^{(i)}, 1\leq i\leq \ell\right), 0\right)$$

where $\xi_{(\alpha)}^{(1)},\dots,\xi_{(\alpha)}^{(\ell)}$ are i.i.d. with distribution (1.1).

(ii) *Assume that (a_n) is regularly varying with index 0 and that $\sum_{i=1}^\infty i^{-1}a_i=\infty$. Then,*

$$\left(\left(\frac{d\big(X_n^{(i)},\mathsf{root}\big)}{\sum_{i=1}^n i^{-1}a_i}, 1\leq i\leq \ell\right), \frac{\max_{1\leq i\neq j\leq \ell} d\big(\mathsf{B}_n^{(i,j)},\mathsf{root}\big)}{a_n}\right) \xrightarrow[n\to\infty]{\mathbb{P}} \left(\left(\frac{1}{2},\dots,\frac{1}{2}\right), 0\right).$$

Notation. Throughout the paper, we use the notation

$$A_n := \sum_{i=1}^n a_i, \quad \text{for all } n\in\mathbb{N}.$$

2 Height of a typical point

Fix n, and given T_n, let X_n be a point picked uniformly on T_n. The goal of this section is to establish different results on the distribution of the distance of this marked point to the root, mainly when the sequence (a_n) is regularly varying with a strictly positive index. Our approach entirely relies on the fact that this distance can be written as the sum of independent, non–negative random variables. More precisely, as noticed in [10], the distances $D_n(k)$ to the root of the projections of X_n onto $T_k, k\leq n$ can jointly be written in the following form:

$$D_n(k) = \sum_{i=1}^k a_i V_i \mathbb{1}_{\left\{U_i\leq \frac{a_i}{A_i}\right\}}, \quad \forall k\leq n, \tag{2.1}$$

where $U_i, V_i, 1 \le i \le n$ are all uniformly distributed on $(0,1)$ and independent. In particular, the distance D_n of X_n to the root writes

$$D_n = \sum_{i=1}^{n} a_i V_i \mathbb{1}_{\left\{U_i \le \frac{a_i}{A_i}\right\}}. \tag{2.2}$$

To see this, we roughly proceed as follows. Consider the projection $X_n(n-1)$ of X_n onto T_{n-1}. By construction, it is uniformly distributed on T_{n-1} given T_{n-1}, and then:

- either $X_n \in T_{n-1}$ and $X_n(n-1) = X_n$, which occurs with probability A_{n-1}/A_n,
- or $X_n \in T_n \backslash T_{n-1}$ and $d(X_n, X_n(n-1)) = a_n V_n$ with V_n uniform on $(0,1)$ and independent of T_{n-1}, which occurs with probability a_n/A_n.

Iterating this argument gives (2.1). An obvious consequence is that

$$\mathbb{E}\left[D_n\right] = \frac{1}{2} \cdot \sum_{i=1}^{n} \frac{a_i^2}{A_i}. \tag{2.3}$$

The rest of this section is organized as follows. In Section 2.1 we start by recalling some classical bounds for regularly varying sequences that will be used throughout the paper. The first part of Section 2.2 concerns the asymptotic behavior of the height D_n, with the proofs of Theorem 1.2 (i) and its corollaries (1.2), as well as Proposition 1.3. The second part of Section 2.2 is devoted to the implementation of bounds (Lemma 2.3) that will be crucial for the proof of Theorem 1.2 (ii) on the behavior of the height of T_n, proof that will be undertaken in Section 4. Last, Section 2.3 contains the proof of Proposition 1.4.

2.1 Bounds for regularly varying sequences

Assume that (a_n) is regularly varying with index $\alpha \ge 0$. We recall some classical bounds that will be useful at different places in the paper.

Fix $\varepsilon > 0$. From [7, Theorem 1.5.6 and Theorem 1.5.11], there exists an integer i_ε such that for all $n \ge i \ge i_\varepsilon$,

$$(1-\varepsilon)\left(\frac{i}{n}\right)^{\alpha+\varepsilon} \le \frac{a_i}{a_n} \le (1+\varepsilon)\left(\frac{i}{n}\right)^{\alpha-\varepsilon} \tag{2.4}$$

and

$$\frac{(1-\varepsilon)(\alpha+1)}{i} \le \frac{a_i}{A_i} \le \frac{(1+\varepsilon)(\alpha+1)}{i} \tag{2.5}$$

([7, Theorem 1.5.6] and [7, Theorem 1.5.11] are stated for regularly varying functions, but can be used for regularly varying sequences, using that $f(x) := a_{\lfloor x \rfloor}$ is a varying regularly function).

Moreover, still by [7, Theorem 1.5.11],

$$\frac{a_n}{\sum_{i=1}^{n} i^{-1} a_i} \underset{n\to\infty}{\longrightarrow} \alpha. \tag{2.6}$$

2.2 One dimensional convergence and deviations

For $\alpha > 0$, recall the definition of the random variable $\xi_{(\alpha)}$ defined via its Laplace transform

$$\mathbb{E}[\exp(\lambda \xi_{(\alpha)})] = \exp(\phi_{(\alpha)}(\lambda)), \quad \lambda \in \mathbb{R}$$

with $\phi_{(\alpha)}$ given by (1.1). With the expression (2.2), it is easy to find the asymptotic behavior of (D_n) by computing its Laplace transform and then get Theorem 1.2 (i). We more precisely have:

Lemma 2.1. *Assume that* (a_n) *is regularly varying with index* $\alpha > 0$*. Then,*

(i) *For all* $\lambda \in \mathbb{R}$

$$\mathbb{E}\left[\exp\left(\lambda \frac{D_n}{a_n}\right)\right] \underset{n\to\infty}{\longrightarrow} \exp(\phi_{(\alpha)}(\lambda)).$$

(ii) *For all* $c > 1$*, there exists* n_c *such that for all* $n \geq n_c$ *and all* $\lambda \geq 0$*,*

$$\mathbb{E}\left[\exp\left(\lambda \frac{D_n}{a_n}\right)\right] \leq \exp\left(c(1+\alpha^{-1})\lambda \exp(c\lambda)\right).$$

Proof. (i) For all $\lambda \neq 0$, we get from (2.2) that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda \frac{D_n}{a_n}\right)\right] &= \prod_{i=1}^{n} \mathbb{E}\left[\exp\left(\lambda \frac{a_i}{a_n} V \mathbb{1}_{\left\{U \leq \frac{a_i}{A_i}\right\}}\right)\right] \\ &= \prod_{i=1}^{n}\left(1 - \frac{a_i}{A_i} + \frac{a_i}{A_i}\frac{\left(\exp\left(\lambda\frac{a_i}{a_n}\right)-1\right)}{\lambda\frac{a_i}{a_n}}\right)\end{aligned}$$

where U, V are uniform on $(0,1)$ and independent (if $a_i = 0$ for some i we use the convention $(\exp(0)-1)/0 = 1$). Now assume that $\lambda > 0$ (the following lines hold similarly for $\lambda < 0$ by adapting the bounds). Using (2.4), (2.5) together with the fact that $\ln(1+x) \sim x$ as $x \to 0$ and that $x \mapsto x^{-1}(\exp(x)-1)$ is increasing on $(0,\infty)$ and converges to 1 as $x \to 0$, leads to the existence of an integer j_ε such that for $n \geq j_\varepsilon$

$$\begin{aligned} & c_\varepsilon(n) + (1-\varepsilon)^2(\alpha+1)\sum_{i=j_\varepsilon}^{n}\frac{1}{i}\left(\left(\frac{\exp\left(\lambda(1-\varepsilon)\left(\frac{i}{n}\right)^{\alpha+\varepsilon}\right)-1}{\lambda(1-\varepsilon)\left(\frac{i}{n}\right)^{\alpha+\varepsilon}}\right)-1\right) \\ \leq\ & \ln\left(\mathbb{E}\left[\exp\left(\lambda\frac{D_n}{a_n}\right)\right]\right) \\ \leq\ & c_\varepsilon(n) + (1+\varepsilon)^2(\alpha+1)\sum_{i=j_\varepsilon}^{n}\frac{1}{i}\left(\left(\frac{\exp\left(\lambda(1+\varepsilon)\left(\frac{i}{n}\right)^{\alpha-\varepsilon}\right)-1}{\lambda(1+\varepsilon)\left(\frac{i}{n}\right)^{\alpha-\varepsilon}}\right)-1\right), \end{aligned} \tag{2.7}$$

where

$$c_\varepsilon(n) := \sum_{i=1}^{j_\varepsilon-1}\ln\left(1-\frac{a_i}{A_i}+\frac{a_i}{A_i}\frac{\left(\exp\left(\lambda\frac{a_i}{a_n}\right)-1\right)}{\lambda\frac{a_i}{a_n}}\right) \underset{n\to\infty}{\longrightarrow} 0$$

since $a_n \to \infty$. Writing $\frac{1}{i} = \frac{1}{n} \times \frac{n}{i}$, we recognize Riemann sums in the lower and upper bounds, which, letting first $n \uparrow \infty$ and then $\varepsilon \downarrow 0$ gives

$$\ln\left(\mathbb{E}\left[\exp\left(\lambda\frac{D_n}{a_n}\right)\right]\right) \underset{n\to\infty}{\longrightarrow} (\alpha+1)\int_0^1 \frac{1}{x}\left(\left(\frac{\exp(\lambda x^\alpha)-1}{\lambda x^\alpha}\right)-1\right)\mathrm{d}x =: \phi_{(\alpha)}(\lambda).$$

It is easy to see with the change of variables $y = x^\alpha$ in the integral and then the power series expansion of the exponential function that this expression of $\phi_{(\alpha)}(\lambda)$ indeed corresponds to (1.1).

(ii). Fix $c > 1$. Using the upper bound (2.7) and the fact that

$$\frac{\exp(x)-1}{x} - 1 \leq x\exp(x) \quad \text{for all } x > 0$$

we see that for all $0 < \eta < \alpha$ and then for all n large enough

$$\ln\left(\mathbb{E}\left[\exp\left(\lambda\frac{D_n}{a_n}\right)\right]\right) \leq c_\eta(n) + (1+\eta)^3(\alpha+1)\lambda\exp(\lambda(1+\eta))\sum_{i=j_\eta}^{n}\frac{1}{i}\left(\frac{i}{n}\right)^{\alpha-\eta}.$$

Note also, using $\ln(1+x) \le x$, that

$$c_\eta(n) \le \sum_{i=1}^{j_\eta - 1} \frac{a_i}{A_i} \lambda \frac{a_i}{a_n} \exp\left(\lambda \frac{a_i}{a_n}\right)$$

which, clearly, is smaller than $\eta\lambda\exp(\eta\lambda)$ for n large enough and all $\lambda \ge 0$. Gathering this together, we get that for all n large enough (depending on η) and all $\lambda \ge 0$,

$$\ln\left(\mathbb{E}\left[\exp\left(\lambda \frac{D_n}{a_n}\right)\right]\right) \le \left(\eta + (1+\eta)^4\right) \frac{1+\alpha}{\alpha-\eta} \lambda \exp\left(\lambda(1+\eta)\right).$$

Taking η small enough so that $\left(\eta + (1+\eta)^4\right)\alpha \le c(\alpha-\eta)$ gives the expected upper bound. □

Remark (height of a uniform leaf). We keep the notation of the introduction and let $L_{n,\star}$ denote a leaf chosen uniformly at random amongst the n leaves of T_n. Then the previous result implies that when (a_n) is regularly varying with index $\alpha > 0$,

$$\frac{d(L_{n,\star}, \mathsf{root})}{a_n} \xrightarrow[n\to\infty]{\text{(d)}} (1+\xi_{(\alpha)})U^\alpha$$

with U uniformly distributed on $(0,1)$ and independent of $\xi_{(\alpha)}$. To see this, one could use that the distribution of $(1+\xi_{(\alpha)})U^\alpha$ is characterized by its positive moments (since it has exponential moments, since $\xi_{(\alpha)}$ has), together with the fact that for each $p \ge 0$, the p–th moment $\mathbb{E}[(d(L_{n,\star}, \mathsf{root})/a_n)^p]$ converges to $\mathbb{E}[((1+\xi_{(\alpha)})U^\alpha)^p]$. To prove this last convergence, note that

$$\mathbb{E}\left[\left(\frac{d(L_{n,\star}, \mathsf{root})}{a_n}\right)^p\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\left(\frac{d(L_i, \mathsf{root})}{a_i}\right)^p\right]\left(\frac{a_i}{a_n}\right)^p.$$

Since $d(L_i, \mathsf{root}) - a_i$ is uniformly distributed on T_{i-1} (by construction) we know from the previous lemma that, divided by a_i, it converges in distribution to $\xi_{(\alpha)}$, and that more precisely there is convergence of all positive and exponential moments. Together with (2.4), this leads to the convergence of $\mathbb{E}[(d(L_{n,\star}, \mathsf{root})/a_n)^p]$ to $\mathbb{E}[(1+\xi_{(\alpha)})^p]/(\alpha p+1)$, as expected.

Remark (other sequences (a_n)). It is easy to adapt Part (i) of the proof to get that for a general sequence (a_n) of positive terms such that $(\sum_{i=1}^n A_i^{-1}a_i^2)^{-1}\max_{1\le i\le n} a_i \to 0$ as $n \to \infty$,

$$\frac{D_n}{\sum_{i=1}^n A_i^{-1}a_i^2} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{2}. \tag{2.8}$$

It is easy to check that the above condition on (a_n) holds if (a_n) is regularly varying with index 0 and $\sum_{i=1}^\infty i^{-1}a_i = \infty$ (recall (2.5),(2.6)), leading in that case to

$$\frac{D_n}{\sum_{i=1}^n i^{-1}a_i} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{2}.$$

In particular this recovers the first part of Theorem 1.1. To illustrate with other 0–regularly varying sequences, consider $a_n = (\ln(n))^\gamma, \gamma \in \mathbb{R}$. Then:

$$\begin{cases} \dfrac{D_n}{(\ln(n))^{\gamma+1}} \xrightarrow[n\to\infty]{\mathbb{P}} \dfrac{1}{2(\gamma+1)} & \text{when } \gamma > -1 \\ \dfrac{D_n}{\ln(\ln(n))} \xrightarrow[n\to\infty]{\mathbb{P}} \dfrac{1}{2} & \text{when } \gamma = -1 \\ D_n \xrightarrow[n\to\infty]{\text{a.s.}} D_\infty & \text{when } \gamma < -1, \end{cases}$$

where the last line is due to the fact that (D_n) is stochastically increasing (by (2.2)) and that $\lim_n \mathbb{E}[D_n]$ is finite when $\gamma < -1$ (by (2.3),(2.5)), which implies that (D_n) converges in distribution to a r.v. D_∞ with finite expectation. Note that this last argument actually holds for any sequence (a_n) such that $\sum_{i=1}^{\infty} A_i^{-1} a_i^2 < \infty$ (which is equivalent to $\sum_{i=1}^{\infty} i^{-1} a_i < \infty$ when (a_n) is regularly varying, necessarily with index 0). All these remarks lead to Proposition 1.3, using again (2.6) when $\alpha > 0$.

We come back to the case where $\alpha > 0$ and note the following behavior of the maximum of n i.i.d. copies of $\xi_{(\alpha)}$.

Proposition 2.2. *Let $\xi_{(\alpha,1)}, \ldots, \xi_{(\alpha,n)}$ be i.i.d. copies of $\xi_{(\alpha)}$. Then,*

$$\frac{\max\left\{\xi_{(\alpha,1)}, \ldots, \xi_{(\alpha,n)}\right\} \times \ln(\ln(n))}{\ln(n)} \underset{n\to\infty}{\overset{\mathbb{P}}{\longrightarrow}} 1.$$

Proof. From (1.1), we know that the random variable $\xi_{(\alpha)}$ is infinitely divisible and the support of its Lévy measure is $[0,1]$. By [7, Theorem 8.2.3], this implies that

$$\exp\left(\lambda x \ln(x)\right) \mathbb{P}\left(\xi_{(\alpha)} > x\right) \underset{x\to\infty}{\longrightarrow} 0 \quad \text{when } \lambda < 1$$

and

$$\exp\left(\lambda x \ln(x)\right) \mathbb{P}\left(\xi_{(\alpha)} > x\right) \underset{x\to\infty}{\longrightarrow} \infty \quad \text{when } \lambda > 1.$$

Besides, the independence of the $\xi_{(\alpha,i)}, 1 \le i \le n$ leads to

$$\ln\left(\mathbb{P}\left(\max\left\{\xi_{(\alpha,1)}, \ldots, \xi_{(\alpha,n)}\right\} \le u\frac{\ln(n)}{\ln(\ln(n))}\right)\right) \underset{n\to\infty}{\sim} -n\mathbb{P}\left(\xi_{(\alpha)} > u\frac{\ln(n)}{\ln(\ln(n))}\right),$$

for all $u > 0$. With the above estimates, it is straightforward that the right–hand side converges to 0 when $u > 1$ and to $-\infty$ when $u < 1$. □

We will not directly use this result later in the paper, but this may be seen as a hint that the height H_n may be asymptotically proportional to $n^\alpha \ln(n)/\ln(\ln(n))$. To prove this rigorously, we will actually use the following estimates.

Lemma 2.3. *Assume that (a_n) is regularly varying with index $\alpha > 0$ and fix $\gamma > 0$.*

(i) *Then for all $\gamma' < \gamma$,*

$$n^{\gamma'} \mathbb{P}\left(\frac{D_n}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \underset{n\to\infty}{\longrightarrow} 0$$

whereas for all $\gamma' > \gamma$,

$$n^{\gamma'} \mathbb{P}\left(\frac{D_n}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \underset{n\to\infty}{\longrightarrow} \infty.$$

(ii) *Fix $c \in (0,1)$. Then for all $\gamma' < \gamma$*

$$n^{\gamma'} \mathbb{P}\left(\frac{D_n - D_n(\lfloor nc \rfloor)}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \underset{n\to\infty}{\longrightarrow} 0$$

whereas for all $\gamma' > \gamma$,

$$n^{\gamma'} \mathbb{P}\left(\frac{D_n - D_n(\lfloor nc \rfloor)}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \underset{n\to\infty}{\longrightarrow} \infty.$$

Proof. Of course, since $D_n - D_n(\lfloor nc \rfloor) \leq D_n$, we only need to prove the convergence to 0 in (i) for $\gamma' < \gamma$ and the convergence to ∞ in (ii) for $\gamma' > \gamma$.

(i) Let $\gamma' < \gamma$ and take a, d such that $a > \gamma'/\gamma, d > 1$ and $ad < 1$. From the upper bound of Lemma 2.1 (ii), we see that for n large enough

$$\begin{aligned} & n^{\gamma'} \mathbb{P}\left(\frac{D_n}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \\ = \quad & n^{\gamma'} \mathbb{P}\left(a \ln(\ln(n)) \frac{D_n}{a_n} > a\gamma \ln(n)\right) \\ \leq \quad & \exp(\gamma' \ln(n)) \cdot \mathbb{E}\left[\exp\left(a \ln(\ln(n)) \frac{D_n}{a_n}\right)\right] \cdot \exp(-a\gamma \ln(n)) \\ \underset{\text{Lemma2.1(ii)}}{\leq} \quad & \exp\Big(\ln(n) \times (\gamma' - a\gamma) + (1 + \alpha^{-1}) ad \ln(\ln(n)) (\ln(n))^{ad}\Big) \end{aligned}$$

and this converges to 0 since $ad < 1$ and $a\gamma > \gamma'$.

(ii) Let $\gamma' > \gamma$. We will (stochastically) compare the random variable $a_n^{-1}(D_n - D_n(\lfloor nc \rfloor))$ with a binomial $\mathrm{Bin}(\lfloor an \rfloor, b/n)$ distribution, with appropriate $a, b > 0$. A simple application of Stirling's formula will then lead to the expected result. Recall from (2.2) and (2.1) that

$$\frac{D_n - D_n(\lfloor nc \rfloor)}{a_n} = \sum_{i=\lfloor nc \rfloor + 1}^{n} \frac{a_i}{a_n} V_i \mathbb{1}_{\left\{U_i \leq \frac{a_i}{A_i}\right\}},$$

with $U_i, V_i, i \geq 1$ i.i.d. uniform on $(0,1)$. Then note from (2.4) and (2.5) that for all $\varepsilon, d \in (0,1)$

$$\sum_{i=\lfloor dn \rfloor + 1}^{n} \frac{a_i}{a_n} V_i \mathbb{1}_{\left\{U_i \leq \frac{a_i}{A_i}\right\}} \geq (1-\varepsilon) d^{\alpha+\varepsilon} \sum_{\lfloor dn \rfloor + 1}^{n} V_i \mathbb{1}_{\left\{U_i \leq \frac{\alpha+1}{2n}\right\}}$$

provided that n is large enough. Now take $\varepsilon \in (0,1)$ small enough and $d \in (c,1)$ large enough so that $\gamma < (1-\varepsilon)^2 d^{\alpha+\varepsilon} \gamma'$. Setting $N_{n,\varepsilon,d} := \sum_{\lfloor dn \rfloor + 1}^{n} \mathbb{1}_{\{V_i \geq 1-\varepsilon\}}$, we have,

$$\begin{aligned} & \mathbb{P}\left(\frac{D_n - D_n(\lfloor nc \rfloor)}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \\ \geq \quad & \mathbb{P}\left((1-\varepsilon) d^{\alpha+\varepsilon} \sum_{\lfloor dn \rfloor + 1}^{n} V_i \mathbb{1}_{\left\{U_i \leq \frac{\alpha+1}{2n}\right\}} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \\ \underset{(U_i) \text{ indep. } (V_i)}{\geq} \quad & \mathbb{P}\left((1-\varepsilon)^2 d^{\alpha+\varepsilon} \sum_{i=1}^{\lfloor \varepsilon(1-d)n/2 \rfloor} \mathbb{1}_{\left\{U_i \leq \frac{\alpha+1}{2n}\right\}} > \gamma \frac{\ln(n)}{\ln(\ln(n))}, N_{n,\varepsilon,d} \geq \frac{\varepsilon(1-d)n}{2}\right) \\ \geq \quad & \mathbb{P}\left(\mathrm{Bin}\left(\left\lfloor \frac{\varepsilon(1-d)n}{2} \right\rfloor, \frac{\alpha+1}{2n}\right) > \frac{\gamma}{(1-\varepsilon)^2 d^{\alpha+\varepsilon}} \frac{\ln(n)}{\ln(\ln(n))}\right) \\ & - \mathbb{P}\left(\mathrm{Bin}\left(n - \lfloor dn \rfloor, \varepsilon\right) < \frac{\varepsilon(1-d)n}{2}\right). \end{aligned} \tag{2.9}$$

One the one hand, the theory of large deviations for the binomial distribution gives

$$\mathbb{P}\left(\mathrm{Bin}\left(n - \lfloor dn \rfloor, \varepsilon\right) < \varepsilon(1-d)n/2\right) \leq \exp(-hn),$$

with $h > 0$. On the other hand, a simple application of Stirling's formula implies that

$$\mathbb{P}\left(\mathrm{Bin}\left(\left\lfloor \frac{\varepsilon(1-d)n}{2} \right\rfloor, \frac{\alpha+1}{2n}\right) > \frac{\gamma}{(1-\varepsilon)^2 d^{\alpha+\varepsilon}} \frac{\ln(n)}{\ln(\ln(n))}\right) \geq n^{-\frac{\gamma}{(1-\varepsilon)^2 d^{\alpha+\varepsilon}} + o(1)} \tag{2.10}$$

(this is well–known, a proof is given below). Together with the lower bound (2.9), these two facts indeed lead to

$$n^{\gamma'}\mathbb{P}\left(\frac{D_n - D_n(\lfloor nc\rfloor)}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}\right) \underset{n\to\infty}{\longrightarrow} \infty$$

since $\gamma < (1-\varepsilon)^2 d^{\alpha+\varepsilon}\gamma'$. We finish with a quick proof of (2.10). More generally, let $a, b, x > 0$. Then

$$\mathbb{P}\left(\mathrm{Bin}\left(\lfloor an\rfloor, \frac{b}{n}\right) > \frac{x\ln(n)}{\ln(\ln(n))}\right) \geq \binom{\lfloor an\rfloor}{\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1}\left(\frac{b}{n}\right)^{\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1}\left(1-\frac{b}{n}\right)^{\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1}.$$

Using Stirling's formula, the binomial term rewrites

$$\binom{\lfloor an\rfloor}{\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1} = \exp\left(\left(\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1\right)\left(\ln(an) - \ln\left(\frac{x\ln(n)}{\ln(\ln(n))}\right)+1+\circ(1)\right)\right).$$

Hence,

$$\begin{aligned}
&\mathbb{P}\left(\mathrm{Bin}\left(\lfloor an\rfloor, \frac{b}{n}\right) > \frac{x\ln(n)}{\ln(\ln(n))}\right)\\
\geq\ & \exp\left(\left(\left\lfloor\frac{x\ln(n)}{\ln(\ln(n))}\right\rfloor+1\right)\left(\ln(an) - \ln\left(\frac{x\ln(n)}{\ln(\ln(n))}\right)+1+\ln\left(\frac{b}{n}\right)+\circ(1)\right)\right)\\
=\ & \exp\big(-x\ln(n)(1+\circ(1))\big).
\end{aligned}$$

□

2.3 Functional convergence

In this section we prove Proposition 1.4. To lighten notation, we let for all $n\in\mathbb{N}$

$$\xi_n(t) := \frac{D_n(\lfloor nt\rfloor)}{a_n} = \frac{\sum_{i=1}^{\lfloor nt\rfloor} a_i V_i \mathbb{1}_{\{U_i\leq a_i/A_i\}}}{a_n}, \qquad 0\leq t\leq 1,$$

where $U_i, V_i, 1\leq i\leq n$ are i.i.d. uniform on $(0,1)$ (recall the construction (2.1)). Our goal is to prove that the process (ξ_n) converges to the process $\xi_{(\alpha)}$ defined by (1.3) for the Skorokhod topology on $D([0,1],\mathbb{R}_+)$. We start by proving the finite–dimensional convergence, relying on manipulations done in Section 2.2. Then we use Aldous' tightness criterion to conclude that the convergence holds with respect to the topology of Skorokhod.

Finite–dimensional convergence. The processes $\xi_n, n\geq 1$ and $\xi_{(\alpha)}$ all have independent increments, by construction. It remains to prove that

$$\xi_n(t)-\xi_n(s) \underset{n\to\infty}{\overset{(\mathrm{d})}{\longrightarrow}} \xi_{(\alpha)}(t)-\xi_{(\alpha)}(s)$$

for all $0\leq s\leq t\leq 1$. From the proof of Lemma 2.1 (i), we immediately get that for all $\lambda\geq 0$

$$\begin{aligned}
\mathbb{E}\left[\exp\big(\lambda(\xi_n(t)-\xi_n(s))\big)\right] &= \prod_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}\mathbb{E}\left[\exp\left(\lambda\frac{a_i}{a_n}V\mathbb{1}_{\left\{U\leq\frac{a_i}{A_i}\right\}}\right)\right]\\
&\underset{n\to\infty}{\to} \exp\left((\alpha+1)\int_s^t\frac{1}{x}\left(\left(\frac{\exp(\lambda x^\alpha)-1}{\lambda x^\alpha}\right)-1\right)\mathrm{d}x\right).
\end{aligned}$$

On the other hand, Campbell's theorem applied to the Poisson point process (t_i, v_i) on $(0,\infty)^2$ with intensity $(\alpha+1)t^{-\alpha-1}\mathbb{1}_{\{v\leq t^\alpha\}}\mathrm{d}t\mathrm{d}v$ implies that for all $\lambda\geq 0$

$$\mathbb{E}\left[\exp\big(\lambda(\xi_{(\alpha)}(t)-\xi_\alpha(s)\big)\right] = \exp\left((\alpha+1)\int_s^t\int_0^{x^\alpha}(\exp(\lambda v)-1)\,\mathrm{d}v\frac{\mathrm{d}x}{x^{1+\alpha}}\right)$$

which indeed coincides with the above limit of $\mathbb{E}\left[\exp\left(\lambda\big(\xi_n(t)-\xi_n(s)\big)\right)\right]$.

Tightness. We use Aldous' tightness criterion ([6, Theorem 16.10]) that ensures that (ξ_n) is tight with respect to the Skorokhod topology on $D([0,1],\mathbb{R}_+)$ if:

- $\lim_{c\to\infty}\limsup_{n\to\infty}\mathbb{P}(\sup_{t\in[0,1]}\xi_n(t)>c)=0$
- and for all $\varepsilon>0$

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{\tau\in\mathsf{S}_n}\sup_{0\le\theta\le\delta}\mathbb{P}\big(\left|\xi_n\left((\tau+\theta)\wedge 1\right)-\xi_n\left(\tau\right)\right|>\varepsilon\big)=0 \tag{2.11}$$

  where S_n is the set of stopping times with respect to the filtration generated by the process ξ_n.

The first point is obvious, since the processes ξ_n are non–decreasing and we already know that $\xi_n(1)$ converges in distribution. For the second point, note that if $\tau\in\mathsf{S}_n$, then $\lfloor n\tau\rfloor$ is a stopping time with respect to the filtration generated by the process $(D_n(k),0\le k\le n)$. Hence

$$\begin{aligned}
&\sup_{0\le\theta\le\delta}\mathbb{P}\big(\left|\xi_n\left((\tau+\theta)\wedge 1\right)-\xi_n\left(\tau\right)\right|>\varepsilon\big)\\
=\;&\mathbb{P}\big(\xi_n((\tau+\delta)\wedge 1)-\xi_n(\tau)>\varepsilon\big)\\
\le\;&\sum_{k=0}^{n}\mathbb{P}(\lfloor n\tau\rfloor=k)\mathbb{P}\left(\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor}\frac{a_i}{a_n}V_i\mathbb{1}_{\{U_i\le\frac{a_i}{A_i}\}}>\varepsilon\right).
\end{aligned}$$

We may assume that $\varepsilon<\alpha$. Then, using $\mathbb{P}(X>\varepsilon)\le\varepsilon^{-1}\mathbb{E}[X]$ for any non–negative r.v. X, we get

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor)}\frac{a_i}{a_n}V^i\mathbb{1}_{\{U_i\le\frac{a_i}{A_i}\}}>\varepsilon\right)&\le\frac{1}{2\varepsilon}\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor}\frac{a_i^2}{a_nA_i}\\
&\underset{\text{by (2.4),(2.5), for } n\ge n_\varepsilon \text{ and all } k\le n}{\le}\frac{C_{\alpha,\varepsilon}}{n^{\alpha-\varepsilon}}\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor}i^{\alpha-\varepsilon-1}\\
&\underset{\text{for } n\ge n_\varepsilon \text{ and all } k\le n}{\le}C_{\alpha,\varepsilon}\max(\delta^{\alpha-\varepsilon},\delta)
\end{aligned}$$

where $C_{\alpha,\varepsilon}$ depends only on α,ε. To get the last line we have used that either $\alpha-\varepsilon-1\ge 0$ and then (since $k+1\le 2n$)

$$\frac{1}{n^{\alpha-\varepsilon}}\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor}i^{\alpha-\varepsilon-1}\le\frac{((2+\delta)n)^{\alpha-\varepsilon-1}n\delta}{n^{\alpha-\varepsilon}}=(2+\delta)^{\alpha-\varepsilon-1}\delta.$$

Or $\alpha-\varepsilon-1<0$ and then

$$\frac{1}{n^{\alpha-\varepsilon}}\sum_{i=k+1}^{k+1+\lfloor n\delta\rfloor}i^{\alpha-\varepsilon-1}\le\frac{\min\big((k+1+n\delta)^{\alpha-\varepsilon},(k+1)^{\alpha-\varepsilon-1}n\delta\big)}{(\alpha-\varepsilon)n^{\alpha-\varepsilon}}\le\frac{(2\delta)^{\alpha-\varepsilon}}{\alpha-\varepsilon}$$

where the last inequality is obtained by considering the first term in the minimum when $k+1\le n\delta$ and the second term when $k+1>n\delta$.

In conclusion, we have proved that for all n large enough and all stopping times $\tau\in\mathsf{S}_n$,

$$\sup_{0\le\theta\le\delta}\mathbb{P}\big(\left|\xi_n\left((\tau+\theta)\wedge 1\right)-\xi_n\left(\tau\right)\right|>\varepsilon\big)\le C_{\alpha,\varepsilon}\max(\delta^{\alpha-\varepsilon},\delta).$$

which gives (2.11).

3 Multiple marking

In order to prove Theorem 1.2 (ii), we need the joint distribution of the paths from the root to two points marked independently, uniformly in the tree T_n. This is studied in Section 3.1. Then in Section 3.2, we turn to ℓ marked points and the proof of Proposition 1.5.

3.1 Marking two points

The result of this section is available for any sequence (a_n) of positive terms.

Given T_n, let $X_n^{(1)}, X_n^{(2)}$ denote two points taken independently and uniformly in T_n, and $D_n^{(1)}, D_n^{(2)}$ their respective distances to the root. For all $1 \leq k \leq n$, let also $D_n^{(1)}(k)$ (resp. $D_n^{(2)}(k)$) denote the distance to the root of the projection of $X_n^{(1)}$ (resp. $X_n^{(2)}$) onto $T_k \subset T_n$. Our goal is to describe the joint distribution of the paths $\left(\left(D_n^{(1)}(k), D_n^{(2)}(k)\right), 1 \leq k \leq n\right)$ – we recall that the marginals are given by (2.1). To that end, we introduce a sequence $\left(B^{(i,1)}, B^{(i,2)}\right), i \geq 1$ of independent pairs of random variables defined by:

$$\left\{\begin{array}{lcc} \mathbb{P}\left((B^{(i,1)}, B^{(i,2)}) = (1,1)\right) & = & 0 \\ \mathbb{P}\left((B^{(i,1)}, B^{(i,2)}) = (1,0)\right) & = & \frac{a_i}{A_i + a_i} \\ \mathbb{P}\left((B^{(i,1)}, B^{(i,2)}) = (0,1)\right) & = & \frac{a_i}{A_i + a_i} \\ \mathbb{P}\left((B^{(i,1)}, B^{(i,2)}) = (0,0)\right) & = & \frac{A_{i-1}}{A_i + a_i}. \end{array}\right. \tag{3.1}$$

Note the two following facts (which will be useful later on):
• $B^{(i,1)}$ (resp. $B^{(i,2)}$) is stochastically smaller than a Bernoulli r.v. with success parameter a_i/A_i
• the distribution of $B^{(i,1)}$ given that $B^{(i,2)} = 0$ (resp. $B^{(i,2)}$ given that $B^{(i,1)} = 0$) is a Bernoulli r.v. with success parameter a_i/A_i.

Lemma 3.1. *Let $U_i, V_i, V_i^{(1)}, V_i^{(2)}, i \geq 1$ be independent r.v. uniformly distributed on $(0,1)$, all independent of a sequence $((B^{(i,1)}, B^{(i,2)}), i \geq 1)$ of independent pairs of Bernoulli r.v. distributed as (3.1). Then for all $n \geq 1$ and all bounded continuous functions $f : \mathbb{R}^{2\times n} \to \mathbb{R}$,*

$$\begin{aligned} & \mathbb{E}\left[f\left(\left(D_n^{(1)}(k), D_n^{(2)}(k)\right), 1 \leq k \leq n\right)\right] \qquad (3.2) \\ = \ & \sum_{\kappa=1}^{n} \left(\frac{a_\kappa}{A_\kappa}\right)^2 \left(\prod_{i=\kappa+1}^{n} \left(1 - \left(\frac{a_i}{A_i}\right)^2\right)\right) \times \mathbb{E}\left[f\left(\left(\Delta_\kappa^{(1)}(k), \Delta_\kappa^{(2)}(k)\right), 1 \leq k \leq n\right)\right] \end{aligned}$$

where for $j = 1, 2$,

$$\Delta_\kappa^{(j)}(k) = \sum_{i=1}^{(\kappa-1)\wedge k} a_i V_i \mathbb{1}_{\left\{U_i \leq \frac{a_i}{A_i}\right\}} + a_\kappa V_\kappa^{(j)} \mathbb{1}_{\{k \geq \kappa\}} + \sum_{i=\kappa+1}^{k} a_i V_i^{(j)} B^{(i,j)}. \tag{3.3}$$

This lemma implies in particular that the distribution of the splitting index $S_n(2)$ of the two paths linking respectively $X_n^{(1)}$ and $X_n^{(2)}$ to the root, i.e.

$$S_n(2) := \inf\left\{1 \leq k \leq n : p_k(X_n^{(1)}) \neq p_k(X_n^{(2)})\right\},$$

where $p_k(X_n^{(i)}), i = 1, 2$ denotes the projection of $X_n^{(i)}$ onto T_k, is given by

$$\mathbb{P}\left(S_n(2) = \kappa\right) = \left(\frac{a_\kappa}{A_\kappa}\right)^2 \prod_{i=\kappa+1}^{n} \left(1 - \left(\frac{a_i}{A_i}\right)^2\right), \quad 1 \leq \kappa \leq n \tag{3.4}$$

(which is indeed a probability distribution!). Moreover, given $S_n(2)=\kappa$, the dependence of the two paths above the index $\kappa+1$ is only driven by pairs of random variables $\left(B^{(i,1)},B^{(i,2)}\right), i\geq \kappa+1$, as described in (3.3).

Proof. We proceed by induction on $n\geq 1$. For $n=1$, the formula of the lemma reduces to

$$\mathbb{E}\left[f\big(D_1^{(1)},D_1^{(2)}\big)\right]=\mathbb{E}\left[f\big(a_1V_1^{(1)},a_1V_1^{(2)}\big)\right]$$

which is obviously true since the two marked points are independently and uniformly distributed on a segment of length a_1. Consider now an integer $n\geq 2$ and assume that the formula of the lemma holds for $n-1$. When marking $X_n^{(1)},X_n^{(2)}$, four disjoint situations may arise:

- with probability $(a_n/A_n)^2$, the two marked points are on the branch b_n. Conditionally on this event, $D_n^{(1)}(k)=D_n^{(2)}(k), 1\leq k\leq n-1$ which corresponds to the path to the root of a point uniformly distributed on T_{n-1}, which is distributed as

$$\sum_{i=1}^{k} a_iV_i\mathbb{1}_{\left\{U_i\leq \frac{a_i}{A_i}\right\}},\quad 1\leq k\leq n-1.$$

  Moreover $D_n^{(1)}(n)-D_n^{(1)}(n-1)$ and $D_n^{(2)}(n)-D_n^{(1)}(n-1)$ are independent, independent of the path $(D_n^{(1)}(k),k\leq n-1)$, and uniformly distributed on b_n, which has length a_n. All this leads to the term $\kappa=n$ in the sum (3.2).
- with probability $A_{n-1}a_n/A_n^2$, $X_n^{(1)}\in T_{n-1}$ and $X_n^{(2)}\in\mathsf{b}_n$. Conditionally on this event, $D_n^{(1)}(k),1\leq k\leq n-1$ and $D_n^{(2)}(k),1\leq k\leq n-1$ correspond to the respective paths to the root of two points marked independently, uniformly in T_{n-1}. Their joint distribution is therefore given by the induction hypothesis. Moreover $D_n^{(1)}(n)=D_n^{(1)}(n-1)$ and $D_n^{(2)}(n)-D_n^{2}(n-1)$ is independent of the paths $(D_n^{(1)}(k),D_n^{(2)}(k)),1\leq k\leq n-1$ and is uniformly distributed on b_n. To sum up, setting for $\kappa\leq n-1$ $\overline{\Delta}_\kappa^{(1)}(n):=\Delta_\kappa^{(1)}(n-1)$, $\overline{\Delta}_\kappa^{(2)}(n):=\Delta_\kappa^{(2)}(n-1)+a_nV_n^{(2)}$ and $\overline{\Delta}_\kappa^{(1)}(k):=\Delta_\kappa^{(1)}(k)$, $\overline{\Delta}_\kappa^{(2)}(k):=\Delta_\kappa^{(2)}(k)$ for $k\leq n-1$, we have:

$$\begin{aligned}
&\mathbb{E}\left[f\left(\big(D_n^{(1)}(k),D_n^{(2)}(k)\big),1\leq k\leq n\right)\mathbb{1}_{\left\{X_n^{(1)}\in T_{n-1},X_n^{(2)}\in\mathsf{b}_n\right\}}\right]\\
=\;&\frac{A_{n-1}a_n}{A_n^2}\times\sum_{\kappa=1}^{n-1}\left(\frac{a_\kappa}{A_\kappa}\right)^2\left(\prod_{i=\kappa+1}^{n-1}\left(1-\left(\frac{a_i}{A_i}\right)^2\right)\right)\\
&\qquad\times\mathbb{E}\left[f\left(\big(\overline{\Delta}_\kappa^{(1)}(k),\overline{\Delta}_\kappa^{(2)}(k)\big),1\leq k\leq n\right)\right]\\
=\;&\sum_{\kappa=1}^{n-1}\left(\frac{a_\kappa}{A_\kappa}\right)^2\left(\prod_{i=\kappa+1}^{n}\left(1-\left(\frac{a_i}{A_i}\right)^2\right)\right)\\
&\qquad\times\mathbb{E}\left[f\left(\big(\Delta_\kappa^{(1)}(k),\Delta_\kappa^{(2)}(k)\big),1\leq k\leq n\right)\mathbb{1}_{\left\{B^{(n,1)}=0,B^{(n,2)}=1\right\}}\right],
\end{aligned}$$

  where we have used for the second equality that

$$\frac{A_{n-1}a_n}{A_n^2}=\left(1-\left(\frac{a_n}{A_n}\right)^2\right)\times\mathbb{P}\big(B^{(n,1)}=0,B^{(n,2)}=1\big).$$

- with probability $A_{n-1}a_n/A_n^2$, $X_n^{(2)}\in T_{n-1}$ and $X_n^{(1)}\in\mathsf{b}_n$, which is symmetric to the previous case.

- with probability $(A_{n-1}/A_n)^2$ the two marked points are in T_{n-1}. Conditionally on this event, $D_n^{(1)}(n) = D_n^{(1)}(n-1)$, $D_n^{(2)}(n) = D_n^{(2)}(n-1)$ and $D_n^{(1)}(k), 1 \le k \le n-1$ and $D_n^{(2)}(k), 1 \le k \le n-1$ correspond to the paths to the root of two points marked independently, uniformly in T_{n-1}. Their joint distribution is therefore given by the induction hypothesis, and setting for $\kappa \le n-1$, $\overline{\Delta}_\kappa^{(1)}(n) := \Delta_\kappa^{(1)}(n-1)$, $\overline{\Delta}_\kappa^{(2)}(n) := \Delta_\kappa^{(2)}(n-1)$ and $\overline{\Delta}_\kappa^{(1)}(k) := \Delta_\kappa^{(1)}(k)$, $\overline{\Delta}_\kappa^{(2)}(k) := \Delta_\kappa^{(2)}(k)$ for $k \le n-1$, we have:

$$\begin{aligned}
&\mathbb{E}\left[f\left(\left(D_n^{(1)}(k), D_n^{(2)}(k)\right), 1 \le k \le n\right) \mathbb{1}_{\left\{X_n^{(1)} \in T_{n-1}, X_n^{(2)} \in T_{n-1}\right\}}\right] \\
= \quad & \frac{A_{n-1}^2}{A_n^2} \times \sum_{\kappa=1}^{n-1} \left(\frac{a_\kappa}{A_\kappa}\right)^2 \left(\prod_{i=\kappa+1}^{n-1} \left(1 - \left(\frac{a_i}{A_i}\right)^2\right)\right) \\
&\qquad\qquad \times \mathbb{E}\left[f\left(\left(\overline{\Delta}_\kappa^{(1)}(k), \overline{\Delta}_\kappa^{(2)}(k)\right), 1 \le k \le n\right)\right] \\
= \quad & \sum_{\kappa=1}^{n-1} \left(\frac{a_\kappa}{A_\kappa}\right)^2 \left(\prod_{i=\kappa+1}^{n} \left(1 - \left(\frac{a_i}{A_i}\right)^2\right)\right) \\
&\qquad \times \mathbb{E}\left[f\left(\left(\Delta_\kappa^{(1)}(k), \Delta_\kappa^{(2)}(k)\right), 1 \le k \le n\right) \mathbb{1}_{\left\{B^{(n,1)}=0, B^{(n,2)}=0\right\}}\right],
\end{aligned}$$

where we have used for the second equality that

$$\frac{A_{n-1}^2}{A_n^2} = \left(1 - \left(\frac{a_n}{A_n}\right)^2\right) \times \mathbb{P}\big(B^{(n,1)} = 0, B^{(n,2)} = 0\big).$$

Gathering these four situations finally leads to the formula of the lemma for n. □

3.2 Marking ℓ points and behavior of $T_n(\ell)$

The goal of this section is to prove Proposition 1.5. We start with a few notation. For each n, given T_n, let $X_n^{(1)}, \dots, X_n^{(\ell)}$ be ℓ points picked independently and uniformly in T_n. Let then $D_n^{(1)}, \dots, D_n^{(\ell)}$ be their respective distances to the root, and for all $1 \le k \le n$, $D_n^{(1)}(k), \dots, D_n^{(\ell)}(k)$ be the respective distances to the root of the projections of $X_n^{(1)}, \dots, X_n^{(\ell)}$ onto $T_k \subset T_n$.

In the tree $T_n(\ell)$, the subtree of T_n spanned from the root and $X_n^{(1)}, \dots, X_n^{(\ell)}$, we let, using the notation of the introduction,

$$\mathsf{B}_n(\ell) := \mathsf{B}_n^{(i_0,j_0)} \quad \text{if} \quad d(\mathsf{B}_n^{(i_0,j_0)}, \mathsf{root}) = \max_{1 \le i \ne j \le \ell} (d(\mathsf{B}_n^{(i,j)}, \mathsf{root}))$$

be the point amongst the $\mathsf{B}_n^{(i,j)}, 1 \le i \ne j \le \ell$ the farthest from the root (note that it is well–defined a.s.). We may and will also see $\mathsf{B}_n(\ell)$ as a point of T_n.

We will need the following random variables. For all $i \ge 1$, let $\left(B^{(i,1)}, \dots, B^{(i,\ell)}\right)$ be an exchangeable ℓ–uplet with distribution

$$\left\{\begin{array}{lcl}
\mathbb{P}\left((B^{(i,1)}, \dots, B^{(i,\ell)}) = (u_1, \dots, u_\ell)\right) & = & 0 \quad \text{for all } (u_i) \in \{0,1\}^\ell \text{ with at least two } 1 \\
\mathbb{P}\left((B^{(i,1)}, \dots, B^{(i,\ell)}) = (1, 0, \dots, 0)\right) & = & \frac{a_i}{A_{i-1} + \ell a_i} \\
\mathbb{P}\left((B^{(i,1)}, \dots, B^{(i,\ell)}) = (0, 0 \dots, 0)\right) & = & \frac{A_{i-1}}{A_{i-1} + \ell a_i}.
\end{array}\right. \tag{3.5}$$

In order to study the asymptotic behavior of $(T_n(\ell))$, we set up the following lemma, which is similar to Lemma 3.1, although less explicit.

Lemma 3.2. *For all $k \in \mathbb{N}$ and all $n \in \mathbb{N}$, $n > k$, the distribution of*

$$\left(D_n^{(1)} - D_{k+1}^{(1)}, \dots, D_n^{(\ell)} - D_{k+1}^{(\ell)}\right) \text{ given that } \mathsf{B}_n(\ell) \in T_k$$

is the same as that of

$$\left(\sum_{i=k+2}^{n} a_i V_i^{(1)} B^{(i,1)}, \dots, \sum_{i=k+2}^{n} a_i V_i^{(\ell)} B^{(i,\ell)} \right),$$

where the random variables $V_i^{(j)}, i \geq 1, 1 \leq j \leq \ell$ are i.i.d. uniform on $(0,1)$, the ℓ–uplets $\left(B^{(i,1)}, \dots, B^{(i,\ell)}\right)$ are distributed via (3.5), $\forall i \geq 1$, independently of each other and independently of $(V_i^{(j)}, i \geq 1, 1 \leq j \leq \ell)$.

Proof. The proof is similar to that of Lemma 3.1 and holds by induction on $n > k$. We sketch it briefly. For $n = k+1$ the statement is obvious since both $\ell-$uplets are then equal to $(0, \dots, 0)$. Assume now that the statement holds for some $n > k$. Then observe what happens for $n+1$: given that $\mathsf{B}_{n+1}(\ell) \in T_k$, two situations may occur:

- either none of the marked points belongs to the segment b_{n+1}. This occurs with a probability proportional to $(A_n)^\ell$ and then

$$\left(D_{n+1}^{(1)} - D_{k+1}^{(1)}, \dots, D_{n+1}^{(\ell)} - D_{k+1}^{(\ell)}\right) \text{ given that } \mathsf{B}_{n+1}(\ell) \in T_k$$

  is distributed as

$$\left(D_n^{(1)} - D_{k+1}^{(1)}, \dots, D_n^{(\ell)} - D_{k+1}^{(\ell)}\right) \text{ given that } \mathsf{B}_n(\ell) \in T_k.$$

- or a unique marked point belongs to the segment b_{n+1}. The probability that $X_{n+1}^{(1)}$ belongs to b_{n+1} (and not the other $\ell - 1$ marked points) is proportional to $a_{n+1}(A_n)^{\ell-1}$ and in that case,

$$\left(D_{n+1}^{(1)} - D_{k+1}^{(1)}, \dots, D_{n+1}^{(\ell)} - D_{k+1}^{(\ell)}\right) \text{ given that } \mathsf{B}_{n+1}(\ell) \in T_k$$

  is distributed as

$$\left(D_n^{(1)} + a_{n+1}V - D_{k+1}^{(1)}, \dots, D_n^{(\ell)} - D_{k+1}^{(\ell)}\right) \text{ given that } \mathsf{B}_n(\ell) \in T_k,$$

  where V is uniform on $(0,1)$ and independent of $D_n^{(i)} - D_{k+1}^{(i)}, 1 \leq i \leq \ell, \mathsf{B}_n(\ell)$.

This leads to the statement for $n+1$. $\square$

Proof of Proposition 1.5. Throughout this proof it is assumed that (a_n) is regularly varying with index $\alpha > 0$ (the proof is identical under the assumptions (ii) of Proposition 1.5). With the notation of this section, our goal is to prove that

$$\left(\frac{D_n^{(i)}}{a_n}, 1 \leq i \leq \ell, \frac{d(\mathsf{B}_n(\ell), \mathsf{root})}{a_n} \right) \xrightarrow[n\to\infty]{(\mathrm{d})} \left(\left(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell \right), 0 \right)$$

where $\xi_{(\alpha)}^{(1)}, \dots, \xi_{(\alpha)}^{(\ell)}$ are i.i.d. with distribution (1.1). We first claim that

$$\frac{d(\mathsf{B}_n(\ell), \mathsf{root})}{a_n} \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$

since $a_n \to \infty$ and $d(\mathsf{B}_n(\ell), \mathsf{root}) \leq \sum_{1\leq i\neq j\leq \ell} d(\mathsf{B}_n^{(i,j)}, \mathsf{root})$, which is stochastically bounded since the splitting index $S_n(2)$ of the paths of two marked points converges in distribution, by (3.4). By Slutsky's Theorem, it remains to prove that

$$\left(\frac{D_n^{(i)}}{a_n}, 1 \leq i \leq \ell\right) \xrightarrow[n\to\infty]{(\mathrm{d})} \left(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell\right).$$

We start by observing that for all $k \geq 1$,

$$\left(\frac{D_n^{(i)} - D_{k+1}^{(i)}}{a_n}, 1 \leq i \leq \ell\right) \quad \text{given that } \mathsf{B}_n(\ell) \in T_k \quad \xrightarrow[n\to\infty]{(\mathrm{d})} \left(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell\right),$$

which obviously leads to (since $a_n \to \infty$)

$$\left(\frac{D_n^{(i)}}{a_n}, 1 \leq i \leq \ell\right) \quad \text{given that } \mathsf{B}_n(\ell) \in T_k \quad \xrightarrow[n\to\infty]{(\mathrm{d})} \left(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell\right).$$

The above observation relies on the following consequence of Lemma 3.2: for all $(\lambda_i)_{1\leq i\leq \ell} \in \mathbb{R}^\ell$ and all $n > k$,

$$\begin{aligned}
&\ln\left(\mathbb{E}\left[\exp\left(\sum_{i=1}^{\ell} \lambda_i \frac{D_n^{(i)} - D_{k+1}^{(i)}}{a_n}\right)\right] |\mathsf{B}_n(\ell) \in T_k\right) \\
= \quad &\sum_{j=k+2}^{n} \ln\left(1 + \frac{a_j}{A_{j-1} + \ell a_j} \sum_{i=1}^{\ell}\left(\frac{\exp(\lambda_i \frac{a_j}{a_n}) - 1}{\lambda_i \frac{a_j}{a_n}} - 1\right)\right)
\end{aligned}$$

(with the usual convention $x^{-1}(\exp(x) - 1) = 1$ when $x = 0$). A slight modification of the proof of Lemma 2.1 implies that this logarithm converges to $\sum_{i=1}^{\ell} \phi_{(\alpha)}(\lambda_i)$, which then leads to the expected convergences in distribution. The end of the proof is then easy. Let $\mathbf{V}_n = (a_n^{-1} D_n^{(i)}, 1 \leq i \leq \ell)$ and $f : \mathbb{R}^\ell \to \mathbb{R}$ be a continuous, bounded function. Fix $\varepsilon > 0$. There exists $k_\varepsilon \in \mathbb{N}$ such that $\mathbb{P}(\mathsf{B}_n(\ell) \notin T_{k_\varepsilon}) \leq \varepsilon$ for all n, since, as already mentioned, the splitting index of the paths of two marked points converges in distribution, by (3.4). Then, writing

$$\mathbb{E}\left[f(\mathbf{V}_n)\right] = \mathbb{E}\left[f(\mathbf{V}_n)|B_n(\ell) \in T_{k_\varepsilon}\right] \mathbb{P}(\mathsf{B}_n(\ell) \in T_{k_\varepsilon}) + \mathbb{E}\left[f(\mathbf{V}_n)\mathbb{1}_{\{\mathsf{B}_n(\ell)\notin T_{k_\varepsilon}\}}\right]$$

we get that

$$\begin{aligned}
&\mathbb{E}\left[f\big(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell\big)\right](1 - \varepsilon) - \sup_{x\in\mathbb{R}^\ell} |f(x)|\varepsilon \\
\leq \quad &\liminf_{n\to\infty} \mathbb{E}\left[f(\mathbf{V}_n)\right] \\
\leq \quad &\limsup_{n\to\infty} \mathbb{E}\left[f(\mathbf{V}_n)\right] \leq \mathbb{E}\left[f\big(\xi_{(\alpha)}^{(i)}, 1 \leq i \leq \ell\big)\right] + \sup_{x\in\mathbb{R}^\ell} |f(x)|\varepsilon.
\end{aligned}$$

Letting $\varepsilon \to 0$ gives the result. □

4 Height of T_n when $\alpha > 0$

Throughout this section we assume that (a_n) is regularly varying with index $\alpha > 0$. Our goal is to prove that

$$\frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \xrightarrow[n\to\infty]{\mathrm{a.s.}} 1$$

(Theorem 1.2 (ii)). We split the proof into two parts, starting with the fact that

$$\limsup_{n\to\infty} \frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \leq 1 \quad \text{a.s.}, \tag{4.1}$$

which is an easy consequence of Borel–Cantelli's lemma and Lemma 2.3 (i). We will then show that

$$\liminf_{n\to\infty} \frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \geq 1 \quad \text{a.s.}, \tag{4.2}$$

using the second moment method and, again, Borel–Cantelli's lemma. To carry this out, we will use Lemma 3.1 on the two marked points, as well as the estimates of Lemma 2.3 (ii).

4.1 Proof of the limsup (4.1)

In the infinite tree $\cup_{n\geq 1} T_n$, label the leaves by order of appearance: for each $i \geq 1$, the leaf L_i is the one that belongs to the branch b_i. Then consider for $i \geq 2$ the projection of L_i onto T_{i-1} and denote by $\overline{D}_{i-1}$ the distance of this projection to the root, which is distributed as D_{i-1}. Let $\overline{D}_0 = 0$ and note that

$$H_n = \max_{1\leq i\leq n} \{d(L_i, \mathsf{root})\} = \max_{1\leq i\leq n} \{\overline{D}_{i-1} + a_i\}.$$

Now, let $c_1 > c_2 > 1$. By Lemma 2.3 (i),

$$\sum_{i\geq 1} \mathbb{P}\left(\overline{D}_{i-1} \geq c_2 \frac{a_{i-1}\ln(i-1)}{\ln(\ln(i-1))}\right) < \infty.$$

Hence by Borel–Cantelli's lemma, almost surely

$$\overline{D}_{i-1} < c_2 \frac{a_{i-1}\ln(i-1)}{\ln(\ln(i-1))}$$

for all i large enough. This leads, together with the fact that (a_i) is regularly varying – see in particular (2.4) – to the almost sure existence of a (random) i_0 such that

$$\overline{D}_{i-1} + a_i < c_1 \frac{a_n \ln(n)}{\ln(\ln(n))}$$

for all $n \geq i \geq i_0$. Hence,

$$\limsup_{n\to\infty} \frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \leq c_1 \quad \text{a.s.}$$

This holds for all $c_1 > 1$, hence (4.1).

4.2 Proof of the liminf (4.2)

Let $X_n^{(i)}, 1 \leq i \leq n$ be n points marked independently and uniformly in T_n. Then let $D_n^{(i)}, 1 \leq i \leq n$ denote their respective distances to the root, and for all $k < n$, $D_n^{(i)}(k), 1 \leq i \leq n$ denote the distances to the root of their respective projections onto T_k. Of course, $H_n \geq \max_{1\leq i\leq n} D_n^{(i)}$ and it is sufficient to prove the liminf for this maximum of *dependent* random variables. To that end, we first establish the following lemma, using the second moment method.

Lemma 4.1. *For all $c \in (0,1)$ and all $\gamma < 1$,*

$$\mathbb{P}\left(\frac{\max_{1\leq i\leq n}\left(D_n^{(i)} - D_n^{(i)}(\lfloor nc\rfloor)\right)}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \leq n^{\gamma-1+\circ(1)}. \tag{4.3}$$

Since H_n is larger than the maximum involved in this probability, this immediately implies that

$$\mathbb{P}\left(\frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} \leq \gamma\right) \underset{n\to\infty}{\longrightarrow} 0 \quad \text{for all } \gamma < 1.$$

This is however not sufficient since we want an *almost sure* bound for the liminf (4.2). We will turn to this conclusion later on. We first prove the lemma.

Proof of Lemma 4.1. We start with standard arguments, in order to use the second moment method. Fix $\gamma \in (0,1)$ and introduce

$$A_n^{(i)} := \left\{\frac{D_n^{(i)} - D_n^{(i)}(\lfloor nc\rfloor)}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}\right\}, \quad 1 \leq i \leq n,$$

and

$$S_n := \sum_{i=1}^{n} \mathbb{1}_{A_n^{(i)}}.$$

Since the sequence $(D_n^{(i)} - D_n^{(i)}(\lfloor nc\rfloor), 1 \leq i \leq n)$ is exchangeable, we have:

$$\mathbb{E}\left[S_n\right] = n\mathbb{P}\big(A_n^{(1)}\big)$$

and

$$\mathrm{Var}\left(S_n\right) = n\mathbb{P}\big(A_n^{(1)}\big) + n(n-1)\mathbb{P}\big(A_n^{(1)} \cap A_n^{(2)}\big) - \big(n\mathbb{P}\big(A_n^{(1)}\big)\big)^2.$$

Note that with this notation, (4.3) rewrites $\mathbb{P}\left(S_n = 0\right) \leq n^{\gamma-1+\circ(1)}$. To prove this upper bound, we use the second moment method:

$$\begin{aligned}\mathbb{P}\left(S_n = 0\right) &\leq \frac{\mathrm{Var}\left(S_n\right)}{\left(\mathbb{E}[S_n]\right)^2} \\ &\leq \frac{1}{n\mathbb{P}\big(A_n^{(1)}\big)} + \frac{\mathbb{P}\big(A_n^{(1)} \cap A_n^{(2)}\big)}{\big(\mathbb{P}\big(A_n^{(1)}\big)\big)^2} - 1.\end{aligned}$$

By Lemma 2.3 (ii), we know that $n\mathbb{P}\big(A_n^{(1)}\big) = n^{1-\gamma+\circ(1)}$. It remains to show that

$$\frac{\mathbb{P}\big(A_n^{(1)} \cap A_n^{(2)}\big)}{\big(\mathbb{P}\big(A_n^{(1)}\big)\big)^2} \leq 1 + n^{\gamma-1+\circ(1)}.$$

To that end, recall the notation and statement of Lemma 3.1:

$$\mathbb{P}\big(A_n^{(1)} \cap A_n^{(2)}\big) = \sum_{\kappa=1}^{n} p_\kappa \mathbb{P}\left(\frac{\Delta_\kappa^{(j)}(n) - \Delta_\kappa^{(j)}(\lfloor nc\rfloor)}{a_n} > \gamma \frac{\ln(n)}{\ln(\ln(n))}, j = 1,2\right)$$

where $p_\kappa := \left(\frac{a_\kappa}{A_\kappa}\right)^2 \left(\prod_{i=\kappa+1}^{n}\left(1 - \left(\frac{a_i}{A_i}\right)^2\right)\right)$ for $1 \leq \kappa \leq n$. We split this sum into two parts:

(i) First, using the notation of Section 3.1 and the remarks just before Lemma 3.1, we see that

$$\begin{aligned}
&\sum_{\kappa=1}^{\lfloor nc\rfloor} p_\kappa \mathbb{P}\left(\frac{\Delta_\kappa^{(j)}(n)-\Delta_\kappa^{(j)}(\lfloor nc\rfloor)}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}, j=1,2\right)\\
=\ & \sum_{\kappa=1}^{\lfloor nc\rfloor} p_\kappa \mathbb{E}\Bigg[\mathbb{1}_{\left\{\frac{\Delta_\kappa^{(1)}(n)-\Delta_\kappa^{(1)}(\lfloor nc\rfloor)}{a_n}>\gamma\frac{\ln(n)}{\ln(\ln(n))}\right\}}\\
&\qquad\qquad \times \mathbb{P}\left(\frac{\sum_{i=\lfloor nc\rfloor+1}^n a_i V_i^{(2)} B^{(i,2)}}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))} \;\Big|\; B^{(i,1)}, V_i^{(1)}, 1\le i\le n\right)\Bigg]\\
\le\ & \mathbb{P}\big(A_n^{(1)}\big)\sum_{\kappa=1}^{\lfloor nc\rfloor} p_\kappa \mathbb{P}\left(\frac{\Delta_\kappa^{(1)}(n)-\Delta_\kappa^{(1)}(\lfloor nc\rfloor)}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}\right)\\
\le\ & \mathbb{P}\big(A_n^{(1)}\big)^2.
\end{aligned}$$

The first inequality is due to the fact that the sum $\sum_{i=\lfloor nc\rfloor+1}^n a_i V_i^{(2)} B^{(i,2)}$ given $B^{(i,1)}, V_i^{(1)}$, $1\le i\le n$ is stochastically smaller than $D_n^{(1)} - D_n^{(1)}(\lfloor nc\rfloor)$ since the distribution of $B^{(i,2)}$ conditional on $B^{(i,1)}=0$ is a Bernoulli r.v. with success parameter a_i/A_i, and moreover $B^{(i,2)}=0$ a.s. when $B^{(i,1)}=1$. The second inequality follows immediately from Lemma 3.1.

(ii) Second,

$$\begin{aligned}
&\sum_{\kappa=\lfloor nc\rfloor+1}^{n} p_\kappa \mathbb{P}\left(\frac{\Delta_\kappa^{(j)}(n)-\Delta_\kappa^{(j)}(\lfloor nc\rfloor)}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}, j=1,2\right)\\
\le\ & \sum_{\kappa=\lfloor nc\rfloor+1}^{n} p_\kappa \mathbb{P}\left(\frac{\sum_{i=\kappa+1}^n a_i V_i^{(1)} B^{(i,1)} + a_\kappa V_\kappa^{(1)} + \sum_{i=\lfloor nc\rfloor+1}^{\kappa-1} a_i V_i \mathbb{1}_{\{U_i\le \frac{a_i}{A_i}\}}}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}\right)\\
\le\ & n^{-\gamma+\circ(1)}\sum_{\kappa=\lfloor nc\rfloor+1}^{n} p_\kappa = n^{-\gamma-1+\circ(1)}.
\end{aligned}$$

Indeed, note that

$$\sum_{i=\kappa+1}^n a_i V_i^{(1)} B^{(i,1)} + a_\kappa V_\kappa^{(1)} + \sum_{i=\lfloor nc\rfloor+1}^{\kappa-1} a_i V_i \mathbb{1}_{\{U_i\le a_i/A_i\}}$$

is stochastically dominated by $a_\kappa + D_n^{(1)} - D_n^{(1)}(\lfloor nc\rfloor)$ since $B^{(i,1)}$ is dominated by a Bernoulli r.v. with success parameter a_i/A_i, for all i. So by Lemma 2.3 (ii) and the fact that $a_\kappa \le 2a_n$ uniformly in $\kappa\in\{\lfloor nc\rfloor,\ldots,n\}$ for n large enough (see (2.4)), we get that

$$\mathbb{P}\left(\frac{\sum_{i=\kappa+1}^n a_i V_i^{(1)} B^{(i,1)} + a_\kappa V_\kappa^{(1)} + \sum_{i=\lfloor nc\rfloor+1}^{\kappa} a_i V_i \mathbb{1}_{\{U_i\le a_i/A_i\}}}{a_n} > \gamma\frac{\ln(n)}{\ln(\ln(n))}\right) \le n^{-\gamma+\circ(1)}$$

with a $\circ(1)$ independent of $\kappa\in\{\lfloor nc\rfloor,\ldots,n\}$. Moreover, by (2.5),

$$\sum_{\kappa=\lfloor nc\rfloor+1}^{n} p_\kappa \le \sum_{\kappa=\lfloor nc\rfloor+1}^{n}\left(\frac{a_\kappa}{A_\kappa}\right)^2 = n^{-1+\circ(1)}.$$

Finally, gathering the two upper bounds established in (i) and (ii) and using again that $\mathbb{P}(A_n^{(1)}) = n^{-\gamma+\circ(1)}$, we have proved that

$$\frac{\mathbb{P}\big(A_n^{(1)}\cap A_n^{(2)}\big)}{\big(\mathbb{P}\big(A_n^{(1)}\big)\big)^2} \le 1 + \frac{n^{-\gamma-1+\circ(1)}}{n^{-2\gamma+\circ(1)}} = 1 + n^{\gamma-1+\circ(1)}$$

as wanted. □

It remains to deduce (4.2) from Lemma 4.1. To that end, fix $\gamma \in (0,1)$. A first consequence of Lemma 4.1 is that

$$\mathbb{P}\left(\frac{H_n}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \leq n^{\gamma-1+\circ(1)}. \tag{4.4}$$

Now let $c \in (0,1)$ and note that

$$H_n \geq \max\left(H_{\lfloor nc \rfloor}, \max_{1 \leq i \leq n}\left(D_n^{(i)} - D_n^{(i)}(\lfloor nc \rfloor)\right)\right)$$

with $H_{\lfloor nc \rfloor}$ and $\max_{1 \leq i \leq n}\left(D_n^{(i)} - D_n^{(i)}(\lfloor nc \rfloor)\right)$ *independent*. Hence,

$$\begin{aligned}\mathbb{P}\left(\frac{H_n}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) &\leq \mathbb{P}\left(\frac{H_{\lfloor nc \rfloor}}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \\ &\quad \times \mathbb{P}\left(\frac{\max_{1 \leq i \leq n}\left(D_n^{(i)} - D_n^{(i)}(\lfloor nc \rfloor)\right)}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \\ &\leq n^{\gamma c^{-\alpha} - 1 + \circ(1)} \cdot n^{\gamma - 1 + \circ(1)}\end{aligned}$$

by (4.4) applied to $\lfloor nc \rfloor$ instead of n (together with the regular variation assumption on (a_n)) and Lemma 4.1. Next, fix an integer k such that $(1-\gamma)k > 1$. Iterating the previous argument, we get that

$$\mathbb{P}\left(\frac{H_n}{a_n} \leq \gamma \frac{\ln(n)}{\ln(\ln(n))}\right) \leq n^{\gamma \sum_{j=0}^{k-1} c^{-\alpha j} - k + \circ(1)}.$$

We now choose $c \in (0,1)$ sufficiently close to 1 so that $\gamma \sum_{j=0}^{k-1} c^{-\alpha j} - k < -1$ and conclude with Borel–Cantelli lemma that almost surely

$$\frac{H_n \cdot \ln(\ln(n))}{a_n \ln(n)} > \gamma \quad \text{for all } n \text{ large enough.}$$

This holds for all $\gamma < 1$. Hence (4.2).

5 The case $a_n = 1$

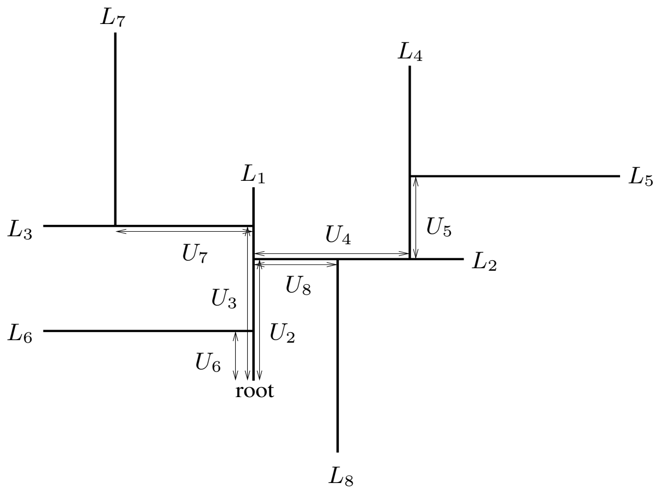
The goal of this section is to prove Theorem 1.1. To that end we start by associating to a sequence (T_n) built recursively from a sequence (a_n) of positive lengths (with no constraints on the a_ns for the moment) a sequence of graph–theoretic trees (R_n) that codes its genealogy as follows:

- R_1 is the tree composed by a unique vertex, labeled $\textcircled{1}$
- if in T_n the branch b_n is glued on the branch b_i, $i < n$, then R_n is obtained from R_{n-1} by grafting a new vertex, labeled $\textcircled{n}$, to the vertex $\textcircled{i}$.

The vertex $\textcircled{1}$ is considered as the root of $R_n, \forall n \geq 1$. This sequence of genealogical trees has been used by [5] to study the boundedness of $\overline{\cup_{n \geq 1} T_n}$.

From now on it is assumed that $a_n = 1$ for all $n \geq 1$. In that case, for all n, R_n is obtained by grafting the new vertex $\textcircled{n}$ to one vertex chosen uniformly at random amongst the $n-1$ vertices of R_{n-1}. Hence R_n is a *uniform recursive tree* with n leaves. Let d_{R_n} denote the graph distance on R_n. It is well–known that

$$\frac{d_{R_n}(\textcircled{n}, \textcircled{1})}{\ln(n)} \xrightarrow[n\to\infty]{\text{a.s.}} 1 \qquad \text{and} \qquad \max_{1 \leq i \leq n} \frac{d_{R_n}(\textcircled{i}, \textcircled{1})}{\ln(n)} \xrightarrow{\text{a.s.}} e, \tag{5.1}$$



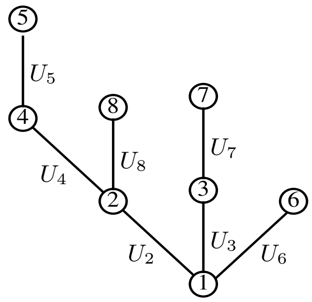


Figure 1: On the left, a version of the tree T_8. On the right, the associated genealogical tree with edge–lengths $\mathcal{R}_8$. The discrete (graph–theoretic) tree R_8 is obtained from $\mathcal{R}_8$ by forgetting the uniform lengths $U_i, 2 \leq i \leq 8$.

see [11, 16, 17]. Next we add lengths to the edges of the trees $R_n, n \geq 2$. By construction, there exists a sequence of i.i.d. uniform r.v. $U_i, i \geq 1$ such that in $\cup_{n \geq 1} T_n$,

$$d(L_i, \mathsf{root}) = \sum_{j=1}^{k} U_{i_j} + U_i + 1 \quad \text{if} \quad \mathsf{b}_1 \to \mathsf{b}_{i_1} \to \ldots \to \mathsf{b}_{i_k} \to \mathsf{b}_i$$

where the sequence $\mathsf{b}_1 \to \mathsf{b}_{i_1} \to \ldots \to \mathsf{b}_{i_k} \to \mathsf{b}_i$ represents the segments involved in the path from the root to L_i (recall from the introduction that the leaves are labelled by order of insertion). For all n and all $2 \leq i \leq n$, we decide to allocate the length U_i to the edge in R_n between the vertex $\textcircled{i}$ and its parent. We denote by $\mathcal{R}_n$ this new tree with edge–lengths and by $d_{\mathcal{R}_n}$ the corresponding metric, so that finally,

$$d(L_i, \mathsf{root}) = d_{\mathcal{R}_n}(\textcircled{i}, \textcircled{1}) + 1, \quad \text{for all leaves } L_i \in T_n. \tag{5.2}$$

See Figure 1 for an illustration.

Height of a typical vertex in T_n, height of leaf L_n, height of a uniform leaf of T_n. The strong law of large numbers and the convergence on the left of (5.1) then clearly yield that $d_{\mathcal{R}_n}(\textcircled{n}, \textcircled{1}) / \ln(n)$ converges a.s. to $1/2$. This in turn yields that $d(L_n, \mathsf{root}) / \ln(n)$ converges a.s. to $1/2$ and that

$$\frac{D_n}{\ln(n)} \xrightarrow[n \to \infty]{\mathbb{P}} \frac{1}{2}$$

since b_{n+1} is inserted on a uniform point of T_n. (More precisely, if we note, for each n, $\overline{D}_n$ the distance to the root of the insertion point of b_{n+1} on T_n, we obtain versions of the D_ns that converge almost surely: $\overline{D}_n / \ln(n) \to 1/2$ a.s.).

Moreover, from the (a.s.) convergence of $d(L_n, \mathsf{root}) / \ln(n)$ to $1/2$, it is easy to get the convergence in probability of $d(L_{n,\star}, \mathsf{root}) / \ln(n)$ to $1/2$, where $L_{n,\star}$ is a uniform leaf of T_n. We let the reader adapt the proof seen in Section 2.2 for regularly varying sequences (a_n) with a strictly positive index.

Height of T_n. From (5.2) it is clear that the height H_n of T_n has the same asymptotic behavior as the height of $\mathcal{R}_n$. Using results by Broutin and Devroye [9] on the asymptotic behavior of heights of certain trees with edge–lengths, we obtain:

Proposition 5.1. *As $n \to \infty$,*

$$\max_{1\le i\le n} \frac{d_{\mathcal{R}_n}(\text{ⓘ}, \text{①})}{\ln(n)} \xrightarrow{\mathbb{P}} \frac{e^{\beta^*}}{2\beta^*},$$

where β^* is the unique solution in $(0,\infty)$ to the equation $2(e^\beta - 1) = \beta e^\beta$.

Proof. We use several remarks or techniques of [9] and invite the reader to refer to this paper for details. First, according to the paragraph following Theorem 3 in [9], the random recursive trees $(\mathcal{R}_n)$ can be coupled with random binary trees with edge–lengths so as to fit the framework of [9, Theorem 1] on the asymptotic of heights of binary trees with edge–lengths. From this theorem, we then know that

$$\max_{1\le i\le n} \frac{d_{\mathcal{R}_n}(\text{ⓘ}, \text{①})}{\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} c$$

where c is defined a few lines below. Let us first introduce some notation.

Let E denote an exponential r.v. with parameter 1 and Z a real–valued r.v. with distribution $(\delta_0(\mathrm{d}x) + \mathbb{1}_{[0,1]}(x)\mathrm{d}x)/2$, where δ_0 denotes the Dirac measure at 0 and $\mathrm{d}x$ the Lebesgue measure on $\mathbb{R}$. Note that $\mathbb{E}[E] = 1$ and $\mathbb{E}[Z] = 1/4$. Moreover,

$$\Lambda_Z(t) := \ln\left(\mathbb{E}\left[e^{tZ}\right]\right) = \ln\left(1 + \frac{e^t - 1}{t}\right) - \ln(2), \quad \text{for } t \neq 0$$

and $\Lambda_Z(0) = 0$. The corresponding Fenchel–Legendre transform $\Lambda_Z^*(t) := \sup_{\lambda\in\mathbb{R}}\{\lambda t - \Lambda_Z(\lambda)\}$ is then given by

$$\Lambda_Z^*(t) = t\lambda(t) - \ln(h(\lambda(t))) + \ln(2) \quad \text{for } 0 < t < 1,$$

and $\Lambda_Z^*(t) = +\infty$ for $t \notin (0,1)$, where $h(u) = 1 + (e^u - 1)/u$, for $u \in \mathbb{R}$ ($h(0) = 2$), and for $t \in (0,1)$, $\lambda(t)$ is defined by

$$t = \frac{h'(\lambda(t))}{h(\lambda(t))}$$

(the function $u \in \mathbb{R} \mapsto h'(u)/h(u) \in (0,1)$ – with the convention $h'(0)/h(0) = 1/4$ – is bijective, increasing). For the r.v. E, we more simply have

$$\Lambda_E^*(t) = t - 1 - \ln(t) \quad \text{for } 0 < t < 1$$

and $\Lambda_E^*(t) = +\infty$ for $t \notin (0,1)$. According to [9, Theorem 1], the limit c introduced above is defined as the unique maximum of α/ρ along the curve

$$\begin{aligned} &\left\{(\alpha,\rho) : \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \ln(2), 0 < \rho < 1, \frac{1}{4} \le \alpha < 1\right\} \quad (5.3)\\ = \quad &\left\{(\alpha,\rho) : \alpha\lambda(\alpha) - \ln(h(\lambda(\alpha))) + \rho - 1 - \ln(\rho) = 0), 0 < \rho < 1, \frac{1}{4} \le \alpha < 1\right\} \end{aligned}$$

(according to [9, Lemma 1], this curve is increasing and concave).

It remains to determine this maximum. We reason like Broutin and Devroye at the end of their proof of [9, Theorem 3]. The slope of the curve is

$$\frac{\mathrm{d}\rho}{\mathrm{d}\alpha} = \frac{\lambda(\alpha)}{\frac{1}{\rho} - 1}$$

and on the other hand, at the maximum

$$\frac{\mathrm{d}\rho}{\mathrm{d}\alpha} = \frac{\rho}{\alpha}.$$

Hence, at the maximum

$$\alpha_{\max}\lambda(\alpha_{\max}) = 1 - \rho_{\max}.$$

Plugging this in (5.3) gives $\rho_{\max} = 1/h(\lambda(\alpha_{\max}))$, which gives in turn

$$\alpha_{\max}\lambda(\alpha_{\max}) = 1 - \frac{1}{h(\lambda(\alpha_{\max}))}.$$

Setting $\beta_{\max} = \lambda(\alpha_{\max}) \Leftrightarrow \alpha_{\max} = h'(\beta_{\max})/h(\beta_{\max})$, this is equivalent to

$$\frac{h'(\beta_{\max})}{h(\beta_{\max})}\beta_{\max} = 1 - \frac{1}{h(\beta_{\max})}.$$

Simple manipulations then give

$$2(e^{\beta_{\max}} - 1) = \beta_{\max}e^{\beta_{\max}},$$

which then leads to

$$c = \frac{\alpha_{\max}}{\rho_{\max}} = \frac{1}{2}\frac{e^{\beta_{\max}}}{\beta_{\max}}.$$

□

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *Critical random graphs: limiting constructions and distributional properties*, Electron. J. Probab. **15** (2010), no. 25, 741–775. MR-2650781

[2] D. Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR-1085326

[3] D. Aldous, *The continuum random tree. II. An overview*, Stochastic analysis (Durham, 1990), London Math. Soc. Lecture Note Ser., vol. 167, Cambridge Univ. Press, Cambridge, 1991, pp. 23–70. MR-1166406

[4] D. Aldous, *The continuum random tree III*, Ann. Probab. **21** (1993), no. 1, 248–289.

[5] O. Amini, L. Devroye, S. Griffiths, and N. Olver, *Explosion and linear transit times in infinite trees*, To appear in Probab. Theory Related Fields. Preprint – arXiv:1411.4426.

[6] P. Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999. MR-1700749

[7] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1989. MR-1015093

[8] K. A. Borovkov and V. A. Vatutin, *On the asymptotic behaviour of random recursive trees in random environments*, Adv. in Appl. Probab. **38** (2006), no. 4, 1047–1070. MR-2285693

[9] N. Broutin and L. Devroye, *Large deviations for the weighted height of an extended class of trees*, Algorithmica **46** (2006), no. 3-4, 271–297. MR-2291957

[10] N. Curien and B. Haas, *Random trees constructed by aggregation*, To appear in the Annales de l'Institut Fourier. Preprint – arXiv:1411.4255.

[11] L. Devroye, *Applications of the theory of records in the study of random trees*, Acta Inform. **26** (1988), no. 1-2, 123–130. MR-969872

[12] L. Devroye, O. Fawzi, and N. Fraiman, *Depth properties of scaled attachment random recursive trees*, Random Structures Algorithms **41** (2012), no. 1, 66–98. MR-2943427

[13] T. Duquesne and J. F. Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque (2002), no. 281, vi+147. MR-MR1954248

[14] C. Goldschmidt and B. Haas, *A line-breaking construction of the stable trees*, Electron. J. Probab. **20** (2015), no. 16, 1–24.

[15] J. F. Le Gall and Y. Le Jan, *Branching processes in Lévy processes: the exploration process*, Ann. Probab. **26** (1998), no. 1, 213–252. MR-1617047

[16] H. M. Mahmoud, *Limiting distributions for path lengths in recursive trees*, Probab. Engrg. Inform. Sci. **5** (1991), no. 1, 53–59. MR-1183165

[17] B. Pittel, *Note on the heights of random recursive trees and random m-ary search trees*, Random Structures Algorithms **5** (1994), no. 2, 337–347. MR-1262983

[18] F. Rembart and M. Winkel, *Recursive construction of continuum random trees*, Preprint – arXiv:1607.05323.

[19] N. Ross and Y. Wen, *Scaling limits for some random trees constructed inhomogeneously*, Preprint – arXiv:1611.01306.

[20] D. Sénizergues, *Random gluing of metric spaces*, In preparation.

Acknowledgments. I warmly thank Louigi Addario–Berry for pointing out the reference [9]. This work is partially supported by the ANR GRAAL ANR–14–CE25–0014.