



**HAL**  
open science

# Height of weighted recursive trees with sub-polynomially growing total weight

Michel Pain, Delphin Sénizergues

► **To cite this version:**

Michel Pain, Delphin Sénizergues. Height of weighted recursive trees with sub-polynomially growing total weight. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 2024, 60 (3), 10.1214/23-AIHP1379 . hal-03720465

**HAL Id: hal-03720465**

**<https://hal.science/hal-03720465v1>**

Submitted on 13 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Height of weighted recursive trees with sub-polynomially growing total weight

Michel Pain<sup>1,a</sup>  and Delphin Sénizergues<sup>2,b</sup> 

<sup>1</sup>*Institut de Mathématiques de Toulouse (UMR 5219), Université de Toulouse, Toulouse, France; CNRS*, <sup>a</sup>[michel.pain@math.univ-toulouse.fr](mailto:michel.pain@math.univ-toulouse.fr)

<sup>2</sup>*Modal'X, UPL, Université Paris Nanterre, 92000 Nanterre, France, and FP2M, CNRS FR 2036*, <sup>b</sup>[dsenizer@parisnanterre.fr](mailto:dsenizer@parisnanterre.fr)

---

**Abstract.** Weighted recursive trees are built by adding successively vertices with predetermined weights to a tree: each new vertex is attached to a parent chosen at random with probability proportional to its weight. In the case where the total weight of the tree at step  $n$  grows polynomially in  $n$ , we obtained in [20] an asymptotic expansion for the height of the tree, which falls into the universality class of the maximum of branching random walks. In this paper, we consider the case of a total weight growing sub-polynomially in  $n$  and obtain asymptotics for the height of the tree in several regimes, showing that universality is broken and that the model exhibits new behavior.

**Résumé.** Les arbres récursifs pondérés sont construits en ajoutant successivement des sommets aux poids prédéterminés: chaque nouveau sommet est attaché à un parent choisi au hasard avec une probabilité proportionnelle à son poids. Dans le cas où le poids total des  $n$  premiers sommets croît polynomialement en  $n$ , nous avons obtenu dans [20] un développement asymptotique pour la hauteur de l'arbre, qui tombe dans la classe d'universalité du maximum d'une marche aléatoire branchante. Dans le présent article, nous considérons le cas où le poids total croît de manière sous-polynomiale en  $n$  et nous décrivons le comportement asymptotique de la hauteur de l'arbre dans plusieurs régimes. Nous décrivons dans ce cas de nouveaux comportements qui sortent de la classe d'universalité mentionnée au-dessus.

*MSC2020 subject classifications:* Primary 60J80; secondary 05C05 ; 60G70

*Keywords:* random trees, weighted recursive trees, asymptotic height

---

## 1. Introduction

Weighted recursive trees (WRTs) are an example of a model of growing random trees that has attracted a lot of attention in the recent years, see for example [8, 12–16, 18, 20, 21]. This model is parametrized by a sequence  $(w_n)_{n \geq 1}$  of non-negative numbers that we call *weights* and that account for some inhomogeneity between vertices in the growth dynamic of the tree. We are interested in the impact of these weights on the behavior of the height of the obtained trees. In our previous paper [20], we considered weight sequences for which the height of the trees was comparable to other well-known models of growing trees and fell in the same universality class as the maximal displacement for branching random walks. We refer the reader to the introduction of [20], particularly Section 1.4, for a discussion containing a review of the literature concerning the asymptotic height of related models of growing random trees, links with branching processes and a review of the literature about WRTs. In the current paper, we work with weight sequences that take us out of this universality class, in a regime where correlations due to the tree structure are stronger and have a more significant impact on the height of the tree.

### 1.1. Presentation of the model

*Definition of WRTs.* Let us define the model of weighted recursive trees, first introduced in [3] by Borovkov and Vatutin. For any sequence of non-negative real numbers  $(w_n)_{n \geq 1}$  with  $w_1 > 0$ , we define the distribution  $\text{WRT}((w_n)_{n \geq 1})$  on sequences of growing rooted trees, which is called the *weighted recursive tree with weights*  $(w_n)_{n \geq 1}$ . We construct a sequence of rooted trees  $(T_n)_{n \geq 1}$  starting from  $T_1$  containing only one root-vertex  $u_1$  by letting it evolve in the following manner: the tree  $T_{n+1}$  is obtained from  $T_n$  by adding a vertex  $u_{n+1}$  with label  $n + 1$ . The parent of this new vertex is

chosen to be the vertex with label  $K_{n+1}$ , where

$$\forall k \in \{1, \dots, n\}, \quad \mathbb{P}(K_{n+1} = k \mid \mathbf{T}_n) = \frac{w_k}{W_n},$$

denoting, for each  $n \geq 1$ ,  $W_n := \sum_{i=1}^n w_i$  the sum of the  $n$  first weights. A key quantity is the sequence  $(a_n)_{n \geq 1}$  defined as

$$(1.1) \quad a_n := \sum_{i=1}^n \frac{w_i}{W_i}, \quad \text{for any } n \geq 1.$$

Its role can be explained by the following fact (see Remark 2.2): the height of  $\mathbf{u}_{n+1}$  has the same distribution as  $\sum_{i=1}^n B_i$ , where the  $B_i$ 's are independent Bernoulli r.v. with parameter  $w_i/W_i$ . In particular, the height of  $\mathbf{u}_{n+1}$  has mean  $a_n$  and its variance is  $\sum_{i=1}^n \frac{w_i}{W_i} (1 - \frac{w_i}{W_i})$ , which behaves like  $a_n$  as  $n \rightarrow \infty$  if  $w_n/W_n \rightarrow 0$ . Hence, if moreover  $a_n \rightarrow \infty$ , the height of  $\mathbf{u}_{n+1}$ , and therefore the height of a vertex chosen in  $\mathbf{T}_n$  proportionally to its weight, is close to  $a_n$  with fluctuations of the order of  $\sqrt{a_n}$ .

*Objective of the paper.* In [20], in the case where the total weight  $W_n$  grows polynomially, that is  $W_n = \lambda \cdot n^\gamma + O(n^{\gamma-\varepsilon})$  for some  $\lambda, \gamma, \varepsilon > 0$ , we proved (under some weak additional assumption) the following asymptotic expansion for the height of  $\mathbf{T}_n$ ,

$$(1.2) \quad \text{ht}(\mathbf{T}_n) = c_1(\gamma) \log n - c_2(\gamma) \log \log n + O_{\mathbb{P}}(1),$$

as  $n \rightarrow \infty$ , where  $c_1(\gamma), c_2(\gamma)$  are positive constants depending only on  $\gamma$ , and the term  $O_{\mathbb{P}}(1)$  denotes a tight sequence of random variables. This extended a previous result [21] by one of the authors, which described the first order for the height. Moreover, because of the precise value of  $c_2(\gamma)$ , this showed that, in this regime, the height of  $\mathbf{T}_n$  falls into the universality class of the maximum of branching random walks. By heuristics presented in Section 1.4, we believe that the regime where this universal behavior holds is exactly the one where  $W_n = n^{\gamma+o(1)}$  (or equivalently  $a_n \sim \gamma \log n$ ).

In this paper, we show that new behaviors, outside of this universality class, appear when the total weight  $W_n$  grows sub-polynomially or converges to a finite limit. This corresponds to the case where  $a_n$  grows sub-logarithmically or converges. To state our results, we distinguish two main cases. In the first main case, we consider sequences of weights such that  $a_n$  varies as powers of  $\log n$ : either  $a_n$  grows like  $(\log n)^p$  for some  $p \in (0, 1)$ , or  $a_n$  converges to some finite limit at speed  $(\log n)^{-q}$  for some  $q > 0$ . The case of a slower divergence or convergence is also included and the result in this first main case turns out to be unified. In the second main case, we investigate the case where  $a_n$  converges at a speed  $\exp(-\log^\beta n)$  for some  $\beta > 0$ . This leads to three sub-cases with different asymptotics for the height of  $\mathbf{T}_n$ . Before proving these results, we first establish general criteria for obtaining some upper and lower bounds (Proposition 3.2 and Proposition 4.3). These bounds are valid in great generality and could also be used to obtain the asymptotic expansion for  $\text{ht}(\mathbf{T}_n)$  in regimes that are not treated in this paper.

Note that in this paper, we only state results on the height  $\text{ht}(\mathbf{T}_n)$  of the tree, whereas in [20] we also expressed an expansion similar to (1.2) for the diameter  $\text{diam}(\mathbf{T}_n)$ , which can be obtained from the latter by just multiplying all the terms by 2. The same would hold in the setting studied in this paper by the same argument, see Remark 4.4.

*Slowly varying function.* In all cases mentioned above, we allow corrective factors in the behavior of  $a_n$  and those are expressed using a slowly varying function satisfying some regularity conditions. For a function  $L: [1, \infty) \rightarrow (0, \infty)$ , we consider the following family of assumptions  $(\text{SV}_k)$  for  $k \geq 0$ :

$(\text{SV}_0)$   $L$  is slowly varying, that is  $L$  is measurable and, for any  $\lambda > 0$ ,  $L(\lambda x)/L(x) \rightarrow 1$  as  $x \rightarrow \infty$ ;

or, for some integer  $k \geq 1$ ,

$(\text{SV}_k)$   $L$  is  $k$  times differentiable and, for any  $1 \leq i \leq k$ ,  $x^i L^{(i)}(x)/L(x) \rightarrow 0$  as  $x \rightarrow \infty$ .

Note that  $(\text{SV}_1)$  implies  $(\text{SV}_0)$ , see e.g. Lemma B.1. Moreover,  $(\text{SV}_k)$  is the typical behavior that one would want for a  $k$  times differentiable slowly varying function:  $(\text{SV}_k)$  is a restriction to the  $k$  first derivatives of the definition of a smoothly varying function of index 0, see e.g. [1, Equation (1.8.1')].

*Notation.* Some of our assumptions and results are expressed using the Landau big- $O$  and small- $o$  notation: we write  $x_n = O(y_n)$  if there exists a constant  $C$  such that  $|x_n| \leq C|y_n|$  for all  $n \geq 1$ ; we write  $x_n = o(y_n)$  if for every  $\varepsilon > 0$  there exists  $N$  such that for all  $n \geq N$  we have  $|x_n| \leq \varepsilon|y_n|$ . Moreover, we write  $x_n \sim y_n$  if  $x_n = y_n(1 + o(1))$ . Throughout the paper, we also denote  $\mathbb{N} := \{0, 1, 2, \dots\}$  the set of non-negative integers, and use the notation  $\llbracket a, b \rrbracket := [a, b] \cap \mathbb{Z}$  and  $\llbracket a, b \rrbracket := [a, b] \cap \mathbb{Z}$  for integer intervals. We denote by  $C$  a positive constant that can change from line to line.

*Relation between  $a_n$  and  $W_n$ .* We express our assumptions on the sequence of weights  $(w_i)_{i \geq 1}$  in terms of the asymptotic behavior of  $a_n$ , defined in (1.1), and a control on the quantity  $\sum_{i=n}^{\infty} w_i^2/W_i^2$ . In fact, we prove in Section C the following relation between  $a_n$  and  $W_n$ : if  $\sum_{i=1}^{\infty} w_i^2/W_i^2 < \infty$ , then there exists a constant  $K$  such that, as  $n \rightarrow \infty$ ,

$$(1.3) \quad a_n = \log W_n + K + O\left(\sum_{i=n+1}^{\infty} \left(\frac{w_i}{W_i}\right)^2\right).$$

Hence, it is easy to reinterpret our assumptions in term of the asymptotic behavior of  $W_n$  instead. Note that the last display does not assume that  $a_n$  (or equivalently  $W_n$ ) tends to infinity with  $n$ , and is also useful when those quantities are bounded.

## 1.2. Variance varying like powers of $\log n$

We will first state results about the case where  $a_n$  can be written as

$$(1.4) \quad a_n = \int_1^{\log n} x^{-\alpha} L(x) dx + o\left((\log n)^{-1-\alpha} (\log \log n)^2 L(\log n)\right),$$

where  $\alpha > 0$  and  $L$  is a positive function on  $[1, \infty)$  satisfying (SV<sub>2</sub>). By Lemma B.3, such a function  $L$  exists if, for some function  $J: [1, \infty) \rightarrow (0, \infty)$  satisfying (SV<sub>3</sub>), we have

- $a_n = (\log n)^{1-\alpha} J(\log n) + o\left((\log n)^{-1-\alpha} (\log \log n)^2 J(\log n)\right)$  if  $\alpha \in (0, 1)$ ;
- $a_{\infty} - a_n = (\log n)^{1-\alpha} J(\log n) + o\left((\log n)^{-1-\alpha} (\log \log n)^2 J(\log n)\right)$  if  $\alpha > 1$ , where  $a_{\infty} = \lim_{n \rightarrow \infty} a_n$ .

In the case  $\alpha = 1$ , no such general criterion exists and  $a_n$  can either converge or go to infinity. We also assume that

$$(1.5) \quad \sum_{i \geq n} \frac{w_i^2}{W_i^2} = O\left(\frac{1}{n}\right).$$

This assumption guarantees some regularity for the weight sequence  $(w_n)_{n \geq 1}$  and ensures for example that we cannot be in a degenerate case where most of the weight of the tree is concentrated on a very sparse subset of vertices. For comparison, if  $w_n$  was regular in  $n$ , that is  $\frac{w_n}{W_n} \approx \frac{d}{dn} a_n \approx n^{-1} (\log n)^{-\alpha} L(\log n)$ , then we would have  $\sum_{i \geq n} w_i^2/W_i^2 = O(n^{-1} (\log n)^{-2\alpha} L^2(\log n))$ .

**Theorem 1.1.** *Assume (1.4) for  $\alpha > 0$  and some function  $L: [1, \infty) \rightarrow (0, \infty)$  satisfying (SV<sub>2</sub>). Also assume (1.5). Then, almost surely, as  $n \rightarrow \infty$ , we have*

$$\begin{aligned} \text{ht}(\mathbb{T}_n) = \frac{\log n}{\alpha \log \log n} + \frac{\log n}{(\alpha \log \log n)^2} & \left( \sum_{k \geq 0} \left( \frac{\log L(\log n)}{\alpha \log \log n} \right)^k [\log L(\log n) + (k+1) \log \log \log n] \right. \\ & \left. + 1 + \alpha + \log \alpha + o(1) \right). \end{aligned}$$

*Remark 1.2.* Keeping only the term  $k = 0$  in the expansion of  $\text{ht}(\mathbb{T}_n)$  in Theorem 1.1 above, we get

$$\begin{aligned} \text{ht}(\mathbb{T}_n) = \frac{\log n}{\alpha \log \log n} + \frac{\log n}{(\alpha \log \log n)^2} & \left( \log L(\log n) + \log \log \log n + O\left(\frac{(\log L(\log n))^2}{\log \log n}\right) \right. \\ & \left. + 1 + \alpha + \log \alpha + o(1) \right). \end{aligned}$$

The second order term for  $\text{ht}(\mathbb{T}_n)$  can be the one involving  $\log L(\log n)$  or the one involving  $\log \log \log n$  (or none of those two if for example  $L(x) = (\log \log x)^{-1}$ ). Depending on  $L$ , the terms in the series in the expansion of  $\text{ht}(\mathbb{T}_n)$  can be included in the  $o(1)$  for  $k$  large enough, but not necessarily (take for example  $L(x) = \exp\left(\frac{\log x}{\log \log \log x}\right)$ ).

*Remark 1.3.* Recall  $\text{ht}(u_{n+1})$  has the same distribution as  $\sum_{i=1}^n B_i$ , where the  $B_i$ 's are independent Bernoulli r.v. with parameter  $w_i/W_i$  (see Remark 2.2). In order to see the influence of the correlations due to the tree structure, it is mean-

ingful to compare the height of  $T_n$  with the maximum of  $n$  independent variables with the same distribution as  $\sum_{i=1}^n B_i$ <sup>1</sup>. In Section A.1, we explain why this maximum always provides an upper bound for  $\text{ht}(T_n)$ , we compare the expansion of this upper bound with that of the actual behavior of  $\text{ht}(T_n)$  and identify the first term for which they differ. In the framework of Theorem 1.1, an interesting transition occurs:

- If  $\alpha < 1$ , the first terms in the expansion in the i.i.d. case are the same as for  $\text{ht}(T_n)$  up to the term of order  $\log n / (\log \log n)^2$  for which the coefficient differs.
- If  $\alpha = 1$ , the first term is still the same but the difference between the expansions grows faster than  $\log n / (\log \log n)^2$ .
- If  $\alpha > 1$ , the first term differs by a multiplicative constant.

We decided to stop the expansion at the order  $\log n / (\log \log n)^2$ , because it was enough to see the difference with the case of i.i.d. variables in all cases. However, our method (and in particular Proposition 3.2 and 4.3) could give the expansion up to the order  $\log n / (\log \log n)^k$  for any  $k \geq 1$ , if we assume a small enough error term in (1.4) and that  $L$  is sufficiently smooth (see Section 5.1 for details).

*Examples.* Let us apply the above result to specific sequences of weights. The computations needed to ensure that those weight sequences indeed satisfy the assumptions of Theorem 1.1 are available in Section D of the appendix.

- **For**  $w_n = \frac{\lambda(1-\alpha)}{n(\log n)^\alpha} \cdot \exp(\lambda(\log n)^{1-\alpha})$  **with**  $\alpha \in (0, 1)$  **and**  $\lambda > 0$ , we can check that (1.5) is satisfied and that (1.4) holds with a function  $L$  such that  $L(x) = \lambda(1 - \alpha)$  for  $x$  large enough. Applying Theorem 1.1 then yields

$$\text{ht}(T_n) = \frac{\log n}{\alpha \log \log n} + \frac{\log n}{(\alpha \log \log n)^2} (\log \log \log n + 1 + \alpha + \log(\lambda\alpha(1 - \alpha))) + o(1).$$

This example of behavior for the weight sequence is similar to that of the (random) sequences appearing in [2, 18] that are constructed using “memory kernel”  $\mu_1$ . See [18, Section 1.1] for more details.

- **For**  $w_n = \frac{1}{n}$ , we can check that (1.5) is satisfied and (1.4) holds with  $\alpha = 1$  and a function  $L$  such that  $L(x) = 1 + o(1)$  as  $x \rightarrow \infty$ . Applying Theorem 1.1 yields

$$\text{ht}(T_n) = \frac{\log n}{\log \log n} + \frac{\log n}{(\log \log n)^2} (\log \log \log n + 2 + o(1)).$$

- **For**  $w_n = \frac{1}{n(\log n)^\alpha}$  **with**  $\alpha > 1$ , we can check that (1.5) is satisfied and (1.4) holds with a function  $L$  such that  $L(x) = \frac{1}{W_\infty} + o(1)$  as  $x \rightarrow \infty$ , where  $W_\infty := \lim_{n \rightarrow \infty} W_n$ . Using Theorem 1.1 we get

$$\text{ht}(T_n) = \frac{\log n}{\alpha \log \log n} + \frac{\log n}{(\alpha \log \log n)^2} (\log \log \log n + 1 + \alpha + \log \alpha - \log W_\infty + o(1)).$$

Note that in this case, changing the value of a finite number of weights in the sequence would affect the asymptotic behavior of the height at the third order because of the presence of the term containing  $W_\infty$ .

The sequence of weights could also be chosen randomly. For example, let  $w_n := X_n \tilde{w}_n$ , where  $(X_n)_{n \geq 1}$  is a sequence of i.i.d. non-negative random variables and  $(\tilde{w}_n)_{n \geq 1}$  is one of the sequences in the examples above. Then, if  $\mathbb{E}[X_1^2] < \infty$ , one can check that almost surely (1.4) and (1.5) are satisfied with the same function  $L$  as above and therefore the conclusion on  $\text{ht}(T_n)$  still holds a.s. In particular, in the third example above, the third order term containing  $\log W_\infty$  becomes random.

### 1.3. Quickly converging variance

Here, we assume that  $a_n$  converges to some finite limit  $a_\infty$  and that

$$(1.6) \quad a_\infty - a_n = \exp\left(-(\alpha - 1) \log^\beta n\right) \cdot J\left(\exp\left(\log^\beta n\right)\right) \cdot \left(1 + o\left((\log n)^{(4\beta - 2) \wedge 0}\right)\right),$$

<sup>1</sup>It can seem crude to take  $n$  copies of  $\sum_{i=1}^n B_i$  instead of independent variables  $Z_1, \dots, Z_n$  where  $Z_k \stackrel{(d)}{=} \sum_{i=1}^k B_i$ . But note that  $\max_{1 \leq k \leq n} Z_k$  is stochastically larger than the maximum of  $n/2$  independent copies of  $\sum_{i=1}^{n/2} B_i$  and replacing  $n$  by  $n/2$  in all expansions considered does not change anything.

where  $\alpha > 1$ ,  $\beta > 0$  and  $J$  is a positive function on  $[1, \infty)$  satisfying  $(\text{SV}_2)$  if  $\beta < \frac{1}{2}$  and  $(\text{SV}_0)$  otherwise. The assumption required for  $\sum_{i=n}^{\infty} w_i^2/W_i^2$  depends on  $\beta$ . When  $\beta < 1$  we use the same assumption (1.5) as in the previous cases. If  $\beta = 1$ , we will further assume

$$(1.7) \quad \forall \varepsilon > 0: \quad \sum_{i \geq n} \frac{w_i^2}{W_i^2} = O\left(\frac{1}{n^{2\alpha-1-\varepsilon}}\right).$$

Compared to (1.5), this last display assumes a faster convergence to 0 for  $\sum_{i \geq n} w_i^2/W_i^2$ , but note that in this regime weights are also decreasing faster: if  $w_n$  was regular in  $n$ , we would have  $\sum_{i \geq n} w_i^2/W_i^2 = O(n^{1-2\alpha} J^2(n))$ . Finally, if  $\beta > 1$  we assume

$$(1.8) \quad \exists \varepsilon > 0: \quad \sum_{i \geq n} \frac{w_i^2}{W_i^2} = O\left(\frac{1}{n^\varepsilon} \exp\left(-2(\alpha-1)\log^\beta n\right) \cdot J^2\left(\exp\left(\log^\beta n\right)\right)\right).$$

Here, if  $w_n$  was regular in  $n$ , we would have a similar bound with  $1/n^\varepsilon$  replaced by  $(\log n)^{\beta-1}/n$ . Also note that, for some  $J$  slowly varying,  $J(\exp(\log^\beta n))$  can grow faster than any polynomial.

**Theorem 1.4.** *Assume that (1.6) holds for some  $\beta > 0$ , some  $\alpha > 1$  and some function  $J: [1, \infty) \rightarrow (0, \infty)$  satisfying  $(\text{SV}_2)$  if  $\beta < \frac{1}{2}$  and  $(\text{SV}_0)$  otherwise. Then, almost surely, as  $n \rightarrow \infty$ ,*

1. *if  $\beta \in (0, 1)$  and (1.5) holds, we have*

$$\text{ht}(\mathbb{T}_n) \sim \frac{(\log n)^{1-\beta}}{(\alpha-1)(1-\beta)};$$

2. *if  $\beta = 1$  and (1.7) holds, we have*

$$\text{ht}(\mathbb{T}_n) \sim \frac{\log \log n}{\log \alpha};$$

3. *if  $\beta \in (1, \infty)$  and (1.8) holds, we have*

$$\text{ht}(\mathbb{T}_n) \sim \frac{\log \log \log n}{\log \beta}.$$

*Remark 1.5.* As in Remark 1.3, we can compare  $\text{ht}(\mathbb{T}_n)$  with the maximum of  $n$  independent copies of  $\sum_{i=1}^{n-1} B_i$ . In the framework of Theorem 1.4,  $\text{ht}(\mathbb{T}_n)$  is always much smaller than this maximum a.s., see Section A.1.

*Remark 1.6.* We see that in the setting of Theorem 1.4, the faster the weight sequence converges, the slower the height of the tree grows to infinity. Note that unless the weight sequence is such that  $w_i = 0$  for all  $i$  large enough, the height of  $\mathbb{T}_n$  almost surely tends to infinity with  $n$ , irrespective of the behavior of  $a_n$ .

*Example.* Consider the weight sequence  $w_n = n^{-\alpha}$  for  $\alpha > 1$ . Then we can apply Theorem 1.4 with  $\beta = 1$ , see Section D in the appendix where we check that the assumptions are satisfied, to get that almost surely as  $n \rightarrow \infty$  we have

$$\text{ht}(\mathbb{T}_n) \sim \frac{\log \log n}{\log \alpha}.$$

#### 1.4. Link with time-inhomogeneous branching random walk

As mentioned at the beginning of the introduction, the WRT can be compared to a BRW and we make here this analogy more precise. First note that, as explained in [20, Section 1.4], the WRT can be linked rigorously to a particular BRW with types, which does not satisfy the branching property and therefore is not tractable (or at least not easier to study than the WRT itself). Hence, we believe it is not a good strategy to try to prove results on the WRT through a direct comparison with a BRW. The goal here is different: we present a *non-rigorous* analogy, but which provided useful insights for the proofs.

A branching random walk on the real line is a discrete-time process defined as follows. Initially there is one individual at position 0, which forms the 0th generation. For any  $t \in \mathbb{N}$ , each individual of the  $t$ th generation has a random number of children who jump independently of each other from their parent's location according to some fixed jump distribution and these children form the  $(t+1)$ th generation. In the following, we also consider the case of time-inhomogeneous BRW, where the jump distribution can depend on the generation number (but not the reproduction law).

We denote by  $V_t$  the set of particles in generation  $t$  and for a particle  $u \in V_t$  we denote  $X_u$  its position. We introduce the measure

$$\mu_t^{\text{BRW}} := \frac{1}{|V_t|} \sum_{u \in V_t} \delta_{X_u},$$

the empirical measure of positions of particles in generation  $t$ , so that in particular, the maximal displacement at time  $t$  corresponds to the supremum of the support of  $\mu_t^{\text{BRW}}$ . A property that will be important in the comparison with the WRT is the following: Sample a random variable  $Z_{t+1}$  with distribution  $\mu_{t+1}^{\text{BRW}}$ . Then, conditional on  $\mu_t^{\text{BRW}}$ , the conditional distribution of  $Z_{t+1}$  can be described by the following equality in law

$$(1.9) \quad (Z_{t+1} | \mu_t^{\text{BRW}}) \stackrel{(d)}{=} Z_t + \Delta_{t+1},$$

where the two terms on the right-hand-side are independent and  $Z_t \sim \mu_t^{\text{BRW}}$  and  $\Delta_{t+1}$  has the jump distribution corresponding to the transition from time  $t$  to  $t+1$ .

Now, for a given  $n$ , we would like to compare the height of the tree  $T_n$  with the maximal displacement at time  $t$  in a BRW, with mean number of children  $e$ , say. With this choice, the total number of particles at time  $t$  in the BRW behaves roughly as  $e^t$ , so it is natural to restrict ourselves to times of the form  $n = \lfloor e^t \rfloor$  (in what follows we omit the floor notation so as to keep expressions simple). Now, we consider the measure

$$\mu_t^{\text{WRT}} = \frac{1}{W_{e^t}} \sum_{i=1}^{e^t} w_i \delta_{\text{ht}(u_i)}$$

which is the (weighted) empirical distribution of heights of vertices in the tree  $T_{e^t}$ . If we sample a random variable  $H_{t+1}$  with distribution  $\mu_{t+1}^{\text{WRT}}$  then we obtain the following equality in distribution<sup>2</sup>

$$(1.10) \quad (H_{t+1} | \mu_t^{\text{WRT}}) \stackrel{(d)}{=} H_t + \sum_{i=e^t+1}^{e^{t+1}} B_i,$$

where all the terms on the right-hand-side are independent with  $H_t \sim \mu_t^{\text{WRT}}$  and the  $B_i$ 's have respective distribution  $\text{Bernoulli}(w_i/W_i)$ . From the similarity between (1.9) and (1.10) it is natural for our comparison to fix the jump distribution between time  $t$  and time  $t+1$  to be close to that of

$$(1.11) \quad \sum_{i=e^t+1}^{e^{t+1}} B_i \stackrel{(d)}{\simeq} \text{Poisson} \left( \sum_{i=e^t+1}^{e^{t+1}} \frac{w_i}{W_i} \right) = \text{Poisson}(a_{e^{t+1}} - a_{e^t}),$$

at least for  $t$  large enough so that the Poisson approximation can be justified.

To summarize the analogy: the (weighted) empirical distribution of heights for vertices in  $T_n$  can be compared with empirical distribution of the positions of particles at time  $t = \log n$  in a BRW with mean number of children  $e$  which jumps between time  $s$  and  $s+1$  with distribution  $\text{Poisson}(a_{e^{s+1}} - a_{e^s})$ . In particular, the height of  $T_n$  should behave like the maximal displacement in the BRW. One can expect this comparison to be more precise than the one with independent copies of the  $\sum_{i=1}^n B_i$ , because it relies not only on a many-to-one formula, but also on the hierarchical construction of the model. Indeed, it is precise enough to predict all the expansions in Theorem 1.1 and Theorem 1.4.

In the case treated in [20] where  $W_n$  grows polynomially, we have  $a_n \simeq \gamma \log n$  and therefore the jumps approximately have distribution  $\text{Poisson}(\gamma)$ , which does not depend on  $s$  in this case. This is a realistic approximation: the asymptotic expansion of  $\text{ht}(T_n)$  obtained in [20], see (1.2), is exactly the same as the one of the maximal position at time  $t = \log n$  in a BRW with mean number of children  $e$  and jump distribution  $\text{Poisson}(\gamma)$ .

In the framework of Section 1.2 with  $L = 1$ , it follows from (1.11) that jumps at time  $s$  in the BRW have approximately a  $\text{Poisson}(s^{-\alpha})$  distribution, so the BRW becomes time-inhomogeneous with decreasing variance. To our knowledge, this time-inhomogeneous BRW does not fit into the framework studied so far in the literature, that we review briefly below.

A model of time-inhomogeneous BRW has been introduced by Bovier and Kurkova [6] and is defined as follows. Fix a profile of variance  $\sigma^2: [0, 1] \rightarrow (0, \infty)$ . Then, for each given horizon  $t$ , consider the binary BRW where jumps at time  $s$

<sup>2</sup>The term involving a sum of independent Bernoulli r.v. arises from applying the many-to-one formula (Lemma 2.1) to the tree  $T_{e^{t+1}}^{(e^t)}$  where the  $e^t$  first vertices are merged together (see definition in Section 2.1), conditionally on the tree  $T_{e^t}$ . The fact that the first term has distribution  $\mu_t^{\text{WRT}}$  is obtained by considering the distribution of  $T_{e^{t+1}}$  conditional on  $T_{e^t}$  and  $T_{e^{t+1}}^{(e^t)}$ .



have distribution  $\mathcal{N}(0, \sigma^2(s/t))$ . They showed that the first order for the maximal position at time  $t$  is linear in  $t$  with an explicit speed in terms of  $\sigma^2$ . In particular, if  $\sigma^2$  is decreasing, the speed is strictly smaller than the one of the maximum of  $2^t$  independent random walks of length  $t$  with jumps  $\mathcal{N}(0, \sigma^2(s/t))$  at time  $s$ . This is therefore called the strongly correlated case, since the branching structure has an effect even on the first order of the maximum. In comparison, if  $\sigma^2$  is non-decreasing, the speed is the same as for independent random walk (this is the weakly correlated case).

In the case of a decreasing variance profile, the next terms in the expansion of the maximum have been studied in [11, 17] in a continuous-time setting and in [19] for non-Gaussian jumps and also time-inhomogeneous reproduction law. Fang and Zeitouni [11] showed that the second term is negative and of order  $t^{1/3}$  and Mallein [19] identified it precisely. This is much larger than the logarithmic correction appearing for the maximum of classical BRW. Maillard and Zeitouni [17] proved tightness of the maximum after centering by an additional logarithmic term. Tightness around the median had been obtained previously by Fang [9] in a general setting. Other results have been proved in the case where  $\sigma^2$  takes only two values [4, 10] and in the weakly correlated case [5].

Now we come back to our framework with a BRW with  $\text{Poisson}(s^{-\alpha})$  jumps. For any fixed horizon  $t$ , one can rewrite the law of the jump at time  $s$  as  $\text{Poisson}(t^{-\alpha} \cdot (s/t)^{-\alpha})$ . Hence, two effects can be expected here: a first one, coming from the inhomogeneity in the variance profile that is given by the function  $u \in [0, 1] \mapsto u^{-\alpha}$  and a second one, coming from the overall reduction of the variance by a factor  $t^{-\alpha}$ . If the jumps had law  $\mathcal{N}(0, t^{-\alpha} \cdot (s/t)^{-\alpha})$ , a scaling argument would allow us to compare with the previously mentioned case and the maximum should be of order  $t^{-\alpha/2} \cdot t$  for  $\alpha < 2^3$ . In the case of  $\text{Poisson}(s^{-\alpha})$  jumps and mean number of children  $e$ , Theorem 1.1 suggests that the first order of the maximum should be  $t/(\alpha \log t)$ . The reason for this different behavior is that the total variance of a particle's position at time  $t$  is much smaller than  $t$ , while the total number of particles is still roughly  $e^t$ , so typically the maximal particle is reaching a position much larger than its variance, and therefore behaves in a very large deviations regime for which there is no universality<sup>4</sup>. In particular, even if we are in the strongly correlated case, the difference with the maximum of independent random walks does not necessarily appear at the first order, see Remark 1.3.

### 1.5. Organization of the paper

The paper is organized as follows: In Section 2 we introduce some notation and recall two results that hold in all generality for WRTs, a “many-to-one” and a “many-to-two” lemma, which allow us to respectively re-express first and second moment estimates on some key quantities as probabilities of events involving an inhomogeneous random walk with Bernoulli increments. Then, we use the many-to-one lemma to get some first moment estimate in Section 3, which leads to a general upper bound on the height of the tree, stated in Proposition 3.2. In Section 4, we then use the many-to-two lemma to get some second moment estimate, which yields the general lower bound on the height of the tree contained in Proposition 4.3. Finally in Section 5, we prove Theorem 1.1 and Theorem 1.4 by applying Proposition 3.2 and Proposition 4.3 to the particular behavior of the weight sequences that we consider. The strategy for the proof is presented in Section 3.1, Section 4.1 and Section 5.1.

In the appendix, Section A.1 considers crude upper bounds on  $\text{ht}(T_n)$  that one can make using the many-to-one lemma in a naive way and compares those to the actual behavior of  $\text{ht}(T_n)$  obtained in Theorem 1.1 and Theorem 1.4. Section A.2 presents a simple method to obtain lower bounds, at least when the sequence of weights is regular enough, which provides the correct first order in the regimes considered in this paper. In Section B, we prove a few technical results concerning slowly varying functions. Section C contains the proof of (1.3) which is helpful for checking that the assumption of our results are satisfied for particular weight sequences. Section D contains some computations ensuring that we can apply our results to the examples presented in the introduction.

## 2. Definitions and many-to-few lemmas

In this section we recall some definitions and some notation concerning WRTs and state two useful lemmas: the many-to-one lemma (Lemma 2.1) and the many-to-two lemma (Lemma 2.3).

<sup>3</sup>In the literature, the function  $\sigma$  is not allowed to diverge at 0, but as long as it is integrable the result for the 1st order of the maximum should hold.

<sup>4</sup>In a homogeneous BRW at time  $t$ , there are roughly  $e^t$  particles with positions given by the value  $S_t$  at time  $t$  of a homogeneous random walk, for which one has large deviation estimates  $\mathbb{P}(S_t \geq at) \simeq \exp(-I(a)t)$  for  $a$  greater than the mean jump. If one neglects the dependence between these random walks, this explains that the maximum is of order  $t$  in that case (more precisely the maximum is roughly  $at$  where  $a$  is such that  $I(a) = 1$ ). In the inhomogeneous case with jumps of variance  $s^{-\alpha}$  at time  $s$  with  $\alpha \in (0, 1)$ , the variance of  $S_t$  would be  $\sum_{s=1}^t s^{-\alpha} \simeq (1-\alpha)^{-1} t^{1-\alpha}$ . So the maximum of  $\exp(t^{1-\alpha})$  independent copies of  $S_t$  would be governed by large deviation events. But, we are interested here in the behavior of  $e^t$  copies, so very large deviation events (where  $S_t \gg t^{1-\alpha}$ ) will occur. In this regime there is no universality for the tail estimates for  $S_t$ .



## 2.1. Some definitions

*Labels and ancestors of a vertices.* For any  $u \in T_n$ , we write  $\text{lab}(u)$  for the label of vertex  $u$  in the tree  $T_n$ , which is an integer between 1 and  $n$ . For any  $k \leq n$  we write  $u(k)$  for the most recent ancestor of  $u$  that has label smaller or equal to  $k$ . For any  $u, v \in T_n$ , we denote  $u \wedge v$  the most recent common ancestor of  $u$  and  $v$  in the tree  $T_n$ .

*An auxiliary tree.* For any integer  $N \geq 1$ , we construct a new tree  $T_n^{(N)}$  from  $T_n$ : we first remove all vertices with labels 2 through  $N$  and then attach all of them and all of their children to the root. Note that  $T_n^{(N)}$  has distribution  $\text{WRT}((w_i^{(N)})_{i \geq 1})$ , where the sequence of weights  $(w_i^{(N)})_{i \geq 1}$  is related to the sequence  $(w_i)_{i \geq 1}$  as follows:

$$(2.1) \quad w_i^{(N)} = \begin{cases} W_N & \text{if } i = 1, \\ 0 & \text{if } 2 \leq i \leq N, \\ w_i & \text{if } i \geq N + 1. \end{cases}$$

In other words, the sequence  $(w_i^{(N)})_{i \geq 1}$  is obtained from  $(w_i)_{i \geq 1}$  by transferring all the weight of vertices labelled 2 through  $N$  to the root and leaving the rest unchanged. We also write  $W_n^{(N)} := \sum_{i=1}^n w_i^{(N)}$ .

## 2.2. Many-to-few lemmas

In this section we state the many-to-few lemmas, which allow to compute first and second moments of functionals of the heights along the lineage of a vertex chosen in  $T_n$  according to its weight. They have been proved in [20], using a specific construction of  $T_n$  coupled with two distinguished vertices introduced by Mailler and Uribe Bravo [18, Proposition 9].

**Lemma 2.1** (Many-to-one). *For any function  $F : \mathbb{N}^n \rightarrow \mathbb{R}$  we have*

$$\mathbb{E} \left[ \sum_{i=1}^n \frac{w_i}{W_n} \cdot F(\text{ht}(u_i(1)), \text{ht}(u_i(2)), \dots, \text{ht}(u_i(n))) \right] = \mathbb{E} [F(H_1, H_2, \dots, H_n)],$$

where, we defined for  $j \geq 1$ ,  $H_j = \sum_{i=2}^j B_i$  with  $(B_i)_{i \geq 2}$  a sequence of independent random variables such that  $B_i$  has distribution  $\text{Bernoulli}(w_i/W_i)$ .

**Proof.** This is Lemma 2.3 in [20] with  $\theta = 0$  (note that  $\theta$  has a specific non zero value in [20], but all results in Section 2 of [20] stay true for any  $\theta \in \mathbb{R}$ , as mentioned in Remark 2.2 there).  $\square$

*Remark 2.2.* One consequence of the many-to-one lemma, mentioned in the introduction, is that  $\text{ht}(u_{n+1})$  has the same distribution as  $H_n + 1$ . Indeed,  $u_{n+1}$  is attached to a vertex  $u_i$ , for  $1 \leq i \leq n$ , chosen proportionally to its weight  $w_i$  and in that case  $\text{ht}(u_{n+1}) = \text{ht}(u_i(n)) + 1$ . Therefore, for any test function  $f : \mathbb{N} \rightarrow \mathbb{R}$ , we have

$$\mathbb{E} [f(\text{ht}(u_{n+1}))] = \mathbb{E} \left[ \sum_{i=1}^n \frac{w_i}{W_n} \cdot f(\text{ht}(u_i(n)) + 1) \right] = \mathbb{E} [f(H_n + 1)],$$

where we used the many-to-one lemma in the last equality.

**Lemma 2.3** (Many-to-two). *For any  $\ell \geq 1$ , we introduce a sequence  $(B_i^\ell, \bar{B}_i^\ell)_{i \geq 2}$  of independent couples of random variables such that*

- if  $i < \ell$ ,  $B_i^\ell$  has distribution  $\text{Bernoulli}(w_i/W_i)$  and  $\bar{B}_i^\ell = B_i^\ell$ ;
- if  $i = \ell$ ,  $B_i^\ell = \bar{B}_i^\ell = 1$ ;
- if  $i > \ell$ ,  $B_i^\ell$  and  $\bar{B}_i^\ell$  have the same distribution as two independent  $\text{Bernoulli}(w_i/W_i)$  random variables conditioned not to be both equal to 1.

Then, for any  $n \geq 1$  and any functions  $F : \mathbb{N}^n \rightarrow \mathbb{R}$  and  $f : \llbracket 1, n \rrbracket \rightarrow \mathbb{R}$ , we have

$$\begin{aligned} & \mathbb{E} \left[ \sum_{1 \leq i, j \leq n} \frac{w_i w_j}{W_n^2} \cdot f(\text{lab}(u_i \wedge u_j)) \cdot F(\text{ht}(u_i(1)), \dots, \text{ht}(u_i(n))) \cdot F(\text{ht}(u_j(1)), \dots, \text{ht}(u_j(n))) \right] \\ &= \sum_{\ell=1}^n \frac{w_\ell^2}{W_\ell^2} \cdot \left( \prod_{i=\ell+1}^n \left( 1 - \frac{w_i^2}{W_i^2} \right) \right) \cdot f(\ell) \cdot \mathbb{E} [F(H_1^\ell, \dots, H_n^\ell) \cdot F(\bar{H}_1^\ell, \dots, \bar{H}_n^\ell)], \end{aligned}$$

where  $H_j^\ell = \sum_{i=2}^j B_i^\ell$  and  $\overline{H}_j^\ell = \sum_{i=2}^j \overline{B}_i^\ell$  for  $j \geq 1$ .

**Proof.** This is Lemma 2.4 in [20] with  $\theta = 0$ . In the reference, for  $i > \ell$ , the distribution of  $(B_i^\ell, \overline{B}_i^\ell)$  is described as follows, with  $q_i := w_i/W_i$ : the random variable  $B_i^\ell$  has distribution Bernoulli( $q_i/(1+q_i)$ ) and, given  $B_i^\ell$ , the other one  $\overline{B}_i^\ell$  has distribution Bernoulli( $q_i \mathbb{1}_{\{B_i^\ell=0\}}$ ). One can check that this is equivalent to the description given in the statement of the lemma.  $\square$

### 3. Upper bound for the height

In this section, we provide a criterion for obtaining an upper bound for the height of  $T_n$  for general weight sequences  $(w_i)_{i \geq 1}$ . This is the content of Proposition 3.2.

#### 3.1. Strategy

We work with an increasing sequence of integers  $(i_r)_{r \geq 0}$  with  $i_0 = 1$ . We define an associated nondecreasing sequence of integers  $(t_n)_{n \geq 1}$  as follows: for  $n \geq 1$ ,  $t_n$  is the unique integer  $r \geq 0$  such that  $n \in \llbracket i_{r-1} + 1, i_r \rrbracket$ , with the convention  $i_{-1} = 0$ . In particular,  $t_{i_r} = r$  for any  $r \geq 0$ .

This sequence  $(i_r)_{r \geq 0}$  has to be thought as a time change such that  $\text{ht}(T_{i_r}) \simeq r$ . Then, proving that  $\text{ht}(T_{i_r}) \leq r$  for  $r$  large enough implies that  $\text{ht}(T_n) \leq t_n$  for  $n$  large enough (at least when  $n = i_{t_n}$ , but one can fill the gaps easily by monotonicity). The sequence  $(i_r)_{r \geq 0}$  will be chosen explicitly for each specific case in Section 5. It has to be chosen so that  $\text{ht}(T_{i_r})$  is sufficiently smaller than  $r$  to allow the argument to work, but as close to  $r$  as possible to get the best upper bound possible.

The upper bound relies on a first moment calculation on the number of high vertices. However, we need first to introduce a barrier controlling the height along the ancestral line of these high vertices, otherwise this first moment calculation does not give the desired upper bound, see Section A.1. This type of barrier argument is classical in the BRW literature and is used in the time-inhomogeneous setting, see e.g. [17, 19]. A reasonable barrier here consists in enforcing that the ancestor  $u_m(i_r)$  of a high vertex  $u_m$  has a height at most  $r$ , since we expect that  $\text{ht}(T_{i_r}) \leq r$ . Therefore, we introduce the following quantity, for  $n \geq 1$ ,

$$(3.1) \quad Q_n := \sum_{m=1}^n \frac{w_m}{W_n} \mathbb{1}_{\text{ht}(u_m)=t_n} \mathbb{1}_{\forall r \in \llbracket 0, t_n \rrbracket, \text{ht}(u_m(i_r)) \leq r},$$

which is the *weighted* number of vertices of height  $t_n$  in  $T_n$ , whose ancestral line respected the barrier constraint. Working with this weighted version is more convenient when applying the many-to-few lemmas.

The first moment of  $Q_n$  is estimated in Lemma 3.1: by the many-to-one lemma, it equals

$$(3.2) \quad \mathbb{P}(H_n = t_n \quad \text{and} \quad \forall r \in \llbracket 0, t_n \rrbracket, H_{i_r} \leq r).$$

The walk  $(H_k)_{k \geq 1}$  on this event typically stays close to the barrier (at first order). This behavior becomes typical after the following time-inhomogeneous change of measure: the distribution of  $B_i$  for  $i_{r-1} < i \leq i_r$  is biased by  $e^{\theta_r B_i}$ , where  $\theta_r$  is chosen such that  $H_{i_r} - H_{i_{r-1}}$  has mean 1 under the new measure. Actually, we do not need to specify precisely the choice of this sequence  $(\theta_r)_{r \geq 1}$  in this section, but in Section 5, we always choose  $\theta_r = -\log(a_{i_r} - a_{i_{r-1}})$  for  $r$  large enough, which ensures that  $H_{i_r} - H_{i_{r-1}}$  has approximately distribution Poisson(1). It is helpful to keep this in mind while reading the proofs.

Finally, a criterion for the upper bound is established in Proposition 3.2. Heuristically, the mean (non-weighted) number of vertices of height  $t_n$  in  $T_n$  satisfying the barrier constraint is roughly  $n\mathbb{E}[Q_n]$  and the criterion says that, if  $i_t \mathbb{E}[Q_{i_t}]$  is summable in  $t$ , then  $\text{ht}(T_n) \leq t_n$  for large  $n$ .

#### 3.2. First moment estimate

**Lemma 3.1.** *Let  $(\theta_r)_{r \geq 1}$  be a non-decreasing sequence of non-negative numbers and*

$$(3.3) \quad p_i := \frac{e^{\theta_{t_i}} \frac{w_i}{W_i}}{1 + (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i}}, \quad i \geq 2.$$

Let  $n \geq 1$ . Let  $(Y_i)_{i \geq 2}$  be a sequence of independent Bernoulli( $p_i$ ) random variables. We set  $S_0 := 0$  and, for  $r \geq 1$ ,

$$(3.4) \quad X_r := \sum_{i=i_{r-1}+1}^{i_r \wedge n} Y_i \quad \text{and} \quad S_r := \sum_{s=1}^r (X_s - 1).$$

Then, we have

$$(3.5) \quad \begin{aligned} \mathbb{E}[Q_n] &= \left( \prod_{i=2}^n \left( 1 + (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \right) \right) \exp \left( - \sum_{r=1}^{t_n} \theta_r \right) \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r \right) \mathbb{1}_{S_{t_n}=0, \forall r < t_n, S_r \leq 0} \right] \\ &\leq \exp \left( \sum_{i=2}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^{t_n} \theta_r \right). \end{aligned}$$

The calculation relies on a time-inhomogeneous change of measure similar to [19, Lemma 3.3] but here we do not need to estimate precisely the expectation appearing in the middle part of (3.5), since we are not aiming at the same level of precision.

**Proof.** It follows from the many-to-one lemma (Lemma 2.1) that

$$\mathbb{E}[Q_n] = \mathbb{P} \left( \sum_{i=2}^n B_i = t_n, \forall r < t_n, \sum_{i=2}^{i_r} B_i \leq r \right),$$

where the  $B_i$ 's are independent Bernoulli r.v. with parameter  $w_i/W_i$ . Now note that  $Y_i$  has the distribution of  $B_i$  biased by  $e^{\theta_{t_i} B_i}$ , that is

$$\forall \phi: \{0, 1\} \rightarrow \mathbb{R}, \quad \mathbb{E}[\phi(Y_i)] = \frac{\mathbb{E}[\phi(B_i) e^{\theta_{t_i} B_i}]}{\mathbb{E}[e^{\theta_{t_i} B_i}]}.$$

Therefore, we have

$$\mathbb{E}[Q_n] = \mathbb{E} \left[ \exp \left( \sum_{i=2}^n \theta_{t_i} B_i \right) \right] \mathbb{E} \left[ \exp \left( - \sum_{i=2}^n \theta_{t_i} Y_i \right) \mathbb{1}_{\sum_{i=1}^n Y_i = t_n, \forall r < t_n, \sum_{i=1}^{i_r} Y_i \leq r} \right].$$

Then, we can rewrite

$$\sum_{i=2}^n \theta_{t_i} Y_i = \sum_{r=1}^{t_n} \theta_r \sum_{i=i_{r-1}+1}^{i_r \wedge n} Y_i = \sum_{r=1}^{t_n} \theta_r + \sum_{r=1}^{t_n} \theta_r (X_r - 1) = \sum_{r=1}^{t_n} \theta_r + \sum_{r=1}^{t_n-1} (\theta_r - \theta_{r+1}) S_r + \theta_{t_n} S_{t_n},$$

using a summation by part and the fact that  $S_0 = 0$ . Hence, using that we are on the event  $S_{t_n} = 0$ , we get the equality in the lemma. Bounding the expectation by 1 because  $(\theta_r)_{r \geq 1}$  is non-decreasing, and using the inequality  $1 + x \leq e^x$ , the inequality follows.  $\square$

### 3.3. General criterion for the upper bound

**Proposition 3.2.** *Suppose that there exists a sequence  $(\theta_r)_{r \geq 1}$  such that, for some  $r_0 \geq 1$ ,  $(\theta_r)_{r \geq r_0}$  is a non-decreasing sequence of non-negative numbers and such that the sequence*

$$(3.6) \quad i_t \cdot \exp \left( \sum_{i=2}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r \right)$$

*is summable in  $t \geq 1$ . Then,  $\sup_{n \geq 1} (\text{ht}(\mathbf{u}_n) - t_n) < \infty$  almost surely.*

**Proof.** For  $n \geq 1$ , first note that  $\text{ht}(\mathbf{u}_{n+1}) = \text{ht}(\mathbf{u}_{n+1}(n)) + 1$ , since by definition the vertex  $\mathbf{u}_{n+1}$  is directly attached to the vertex  $\mathbf{u}_{n+1}(n)$ . We write

$$(3.7) \quad \{\exists n \geq 1, \text{ht}(\mathbf{u}_{n+1}) > t_{n+1}\} = \bigcup_{n \geq 1} \{\text{ht}(\mathbf{u}_{n+1}) = t_{n+1} + 1 \text{ and } \forall i \leq n, \text{ht}(\mathbf{u}_{n+1}(i)) \leq t_i\},$$

where we used that if  $\text{ht}(\mathbf{u}_{n+1}(n)) \leq t_n$  then  $\text{ht}(\mathbf{u}_{n+1}) \leq t_n + 1 \leq t_{n+1} + 1$ . Given  $\mathbf{T}_n$ , vertex  $\mathbf{u}_{n+1}$  is attached to  $\mathbf{u}_m$  with probability  $w_m/W_n$  for  $1 \leq m \leq n$ , and in that case  $\text{ht}(\mathbf{u}_{n+1}) = \text{ht}(\mathbf{u}_m) + 1$ : therefore, we have

$$(3.8) \quad \mathbb{P}(\text{ht}(\mathbf{u}_{n+1}) = t_{n+1} + 1 \text{ and } \forall i \leq n, \text{ht}(\mathbf{u}_{n+1}(i)) \leq t_i) = \mathbb{E} \left[ \sum_{m=1}^n \frac{w_m}{W_n} \mathbb{1}_{\text{ht}(\mathbf{u}_m) = t_{n+1}} \mathbb{1}_{\forall i \leq n, \text{ht}(\mathbf{u}_m(i)) \leq t_i} \right].$$

If  $n = i_t$ , then  $t_{n+1} = t_n + 1$  and so the last expectation is zero, because the indicator functions both require  $\text{ht}(\mathbf{u}_m) = t_{n+1}$  and  $\text{ht}(\mathbf{u}_m) = \text{ht}(\mathbf{u}_m(n)) \leq t_n$ . If  $n \neq i_t$ , then  $t_{n+1} = t_n$  and, noting that

$$\{\text{ht}(\mathbf{u}_m) = t_n \text{ and } \forall i \leq n, \text{ht}(\mathbf{u}_m(i)) \leq t_i\} = \{\text{ht}(\mathbf{u}_m) = t_n \text{ and } \forall r < t_n, \text{ht}(\mathbf{u}_m(i_r)) \leq r\},$$

it shows that the quantities displayed in (3.8) are equal to  $\mathbb{E}[Q_n]$ . Coming back to (3.7) and using a union-bound, we hence get

$$(3.9) \quad \mathbb{P} \left( \sup_{n \geq 1} (\text{ht}(\mathbf{u}_n) - t_n) > 0 \right) \leq \sum_{n \geq 1} \mathbb{E}[Q_n].$$

Now we fix some large  $N \geq 1$  and consider the trees  $(\mathbf{T}_n^{(N)})_{n \geq 1}$  introduced in Section 2.1. Let  $Q_n^{(N)}$  be defined as in (3.1) but for the tree  $\mathbf{T}_n^{(N)}$ . For all  $n \geq 1$ , we have  $\text{ht}(\mathbf{u}_n) \leq \text{ht}(\mathbf{u}_n^{(N)}) + N$ , so

$$(3.10) \quad \mathbb{P} \left( \sup_{n \geq 1} (\text{ht}(\mathbf{u}_n) - t_n) > N \right) \leq \mathbb{P} \left( \sup_{n \geq 1} (\text{ht}(\mathbf{u}_n^{(N)}) - t_n) > 0 \right) \leq \sum_{n \geq 1} \mathbb{E}[Q_n^{(N)}]$$

where the second inequality is obtained by applying (3.9) to the sequence  $(\mathbf{T}_n^{(N)})_{n \geq 1}$  which is a weighted recursive tree with weight sequence  $(w_n^{(N)})_{n \geq 1}$ . Now, going back to its definition in (3.1), we can check that  $Q_n^{(N)} = 0$  for any  $2 \leq n \leq N$ . Indeed all the terms in the sum vanish: we have  $t_n = 1$ , so the term corresponding to  $m = 1$  is zero because of the indicator  $\mathbb{1}_{\{\text{ht}(\mathbf{u}_1^{(N)}) = 1\}}$ ; the other ones vanish because  $w_m^{(N)} = 0$  for  $2 \leq m \leq n$ . On the other hand, letting  $\tilde{\theta}_r := \theta_r \mathbb{1}_{r \geq r_0}$ , so that  $(\tilde{\theta}_r)_{r \geq 0}$  is a non-decreasing sequence of non-negative numbers, we apply Lemma 3.1 to get

$$\mathbb{E}[Q_n^{(N)}] \leq \exp \left( \sum_{i=2}^n (e^{\tilde{\theta}_{t_i}} - 1) \frac{w_i^{(N)}}{W_i^{(N)}} - \sum_{r=1}^{t_n} \tilde{\theta}_r \right) \leq \exp \left( \sum_{i=i_{r_0-1}+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=r_0}^{t_n} \theta_r \right),$$

using that, for all  $i \geq 2$ , we have  $w_i^{(N)}/W_i^{(N)} \leq w_i/W_i$ . In the end, using (3.10) and the discussion that follows we get that

$$\begin{aligned} \mathbb{P} \left( \sup_{n \geq 1} (\text{ht}(\mathbf{u}_n) - t_n) > N \right) &\leq \sum_{n \geq N+1} \mathbb{E}[Q_n^{(N)}] \leq \sum_{n \geq N+1} \exp \left( \sum_{i=i_{r_0-1}+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=r_0}^{t_n} \theta_r \right) \\ &\leq \sum_{t \geq t_{N+1}-1} i_t \cdot \exp \left( \sum_{i=i_{r_0-1}+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=r_0}^t \theta_r \right), \end{aligned}$$

by grouping terms. The summand in the last sum differs from (3.6) only by a finite factor independent of  $t$  and therefore is summable by assumption of the lemma. Hence, this last sum tends to 0 as  $N \rightarrow \infty$ , because  $t_{N+1} \rightarrow \infty$ . This finishes the proof.  $\square$

#### 4. Lower bound for the height

In this section, we provide a criterion for obtaining a lower bound for the height of  $\mathbf{T}_n$  for general weight sequences  $(w_i)_{i \geq 1}$ . This is the content of Proposition 4.3. We additionally prove Lemma 4.5, which provides us with some estimates on one of the quantities that appear in the assumptions of Proposition 4.3.

#### 4.1. Strategy

As for the upper bound, we work with an increasing sequence of integers  $(i_r)_{r \geq 0}$  with  $i_0 = 1$ , which provides a time change such that  $\text{ht}(\mathbb{T}_{i_r}) \simeq r$ . Again, the sequence  $(i_r)_{r \geq 0}$  will be chosen explicitly for each specific case in Section 5. The main difference for the lower bound is that it has to be chosen so that  $\text{ht}(\mathbb{T}_{i_r})$  is slightly larger than  $r$  (rather than slightly smaller for the upper bound). This ensures that there are typically a large number of vertices at height  $r$  in  $\mathbb{T}_{i_r}$  and that we can expect some concentration in a first and second moment argument.

As in Section 3, we consider the sequence of integers  $(t_n)_{n \geq 1}$  such that for any  $n \geq 1$  we have  $n \in \llbracket i_{t_n-1} + 1, i_{t_n} \rrbracket$ , and the quantity  $Q_n$  defined in (3.1). We first establish a lower bound for  $\mathbb{E}[Q_n]$  in Lemma 4.1 and an upper bound for  $\mathbb{E}[Q_n^2]$  in Lemma 4.2. These bounds show that  $\mathbb{E}[Q_n^2]$  and  $\mathbb{E}[Q_n]^2$  are of the same order. This implies that  $\mathbb{P}(Q_n > 0) \geq c > 0$  for  $n$  large enough, which imply in particular that  $\mathbb{P}(\text{ht}(\mathbb{T}_n) \geq t_n) \geq c$ . However, as for the case of polynomially growing  $W_n$  treated in [20], we emphasize that we cannot proceed here as for the BRW and use the branching property at some large fixed time to conclude that  $\text{ht}(\mathbb{T}_n) \geq t_n - C$  with high probability. Indeed, subtrees rooted at the first vertices of the WRT do not behave independently.

Instead, we use the same method as in [20]: we apply this moment calculation to the auxiliary tree  $\mathbb{T}_n^{(N)}$ , whose height provides a lower bound for  $\text{ht}(\mathbb{T}_n)$ . For large  $N$ , the benefit is the following: in  $\mathbb{T}_n^{(N)}$ , two particles contributing to  $Q_n$  chosen independently have w.h.p. the root as most recent common ancestor. This results from two facts: (i) the barrier prevents the tree from having too large groups of close cousins contributing to  $Q_n$ , so that the most recent common ancestor has to be close to the root, (ii) in  $\mathbb{T}_n^{(N)}$ , the root has a much larger weight than each of its close descendants, so it is unlikely that one of this descendant has two children with progeny contributing to  $Q_n$  (this mimics the branching property argument in a rigorous way). This guarantees that  $\mathbb{E}[Q_n^2]$  and  $\mathbb{E}[Q_n]^2$  are now asymptotically equivalent and therefore  $\mathbb{P}(Q_n > 0) \rightarrow 1$ .

The main differences with [20] are the following: (i) the second moment calculation requires a time-inhomogeneous change of measure as in Section 3, (ii) here we do not want to estimate precisely the remaining expectations involving the random walk  $S_n$  (as the one appearing for the first moment), (iii) because of the latter,  $N$  has to grow appropriately with  $n$ , whereas it was fixed in [20].

#### 4.2. Moments calculation

The following lower bound for  $\mathbb{E}[Q_n]$  is a consequence of Lemma 3.1. The sum  $\sum_{i=2}^n e^{2\theta t_i} \frac{w_i^2}{W_i^2}$  has to be thought as an error term. For now, we do not aim at estimating the expectation involving the walk  $S$  appearing in the statement: although it plays a negligible role, it has to be handled carefully.

**Lemma 4.1.** *Let  $(\theta_r)_{r \geq 1}$  be a non-decreasing sequence of non-negative numbers. Let  $n \geq 1$  and  $(S_r)_{r \geq 0}$  defined as in (3.4). Then, we have*

$$\mathbb{E}[Q_n] \geq \exp\left(\sum_{i=2}^n (e^{\theta t_i} - 1) \frac{w_i}{W_i} - \sum_{i=2}^n \frac{e^{2\theta t_i}}{2} \frac{w_i^2}{W_i^2} - \sum_{r=1}^{t_n} \theta_r\right) \mathbb{E}\left[\exp\left(\sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r\right) \mathbb{1}_{S_{t_n}=0, \forall r < t_n, S_r \leq 0}\right].$$

**Proof.** This follows directly from the equality in Lemma 3.1, together with the inequality  $1 + x \geq \exp(x - \frac{x^2}{2})$  for  $x \geq 0$ .  $\square$

We now prove an upper bound for  $\mathbb{E}[Q_n^2]$ . This relies on the many-to-two lemma and a time-inhomogeneous change of measure. We keep aside the term  $\ell = 1$  given by the many-to-two lemma, because it will be the main term once we work with the auxiliary tree  $\mathbb{T}_n^{(N)}$ . In that case, the first term in the bound below is roughly  $\mathbb{E}[Q_n]^2$  and the second one is an error term. Note in the proof that we manage to bound the term  $\ell = 1$  in terms of  $\mathbb{E}[Q_n]^2$  using Lemma 4.1, but without estimating the expectation involving the walk  $S$ .

**Lemma 4.2.** *Let  $(\theta_r)_{r \geq 1}$  be a non-decreasing sequence of non-negative numbers. Recall the definition of  $p_i$  in (3.3). Let  $n \geq 1$  such that  $n = i_{t_n}$ . Then, we have*

$$\begin{aligned} \mathbb{E}[Q_n^2] &\leq \left(\prod_{i=2}^n (1 - p_i^2)^{-1}\right) \cdot \exp\left(\sum_{i=2}^n e^{2\theta t_i} \frac{w_i^2}{W_i^2}\right) \cdot \mathbb{E}[Q_n]^2 \\ &\quad + \exp\left(\sum_{i=2}^n (e^{\theta t_i} - 1) \frac{w_i}{W_i} - \sum_{r=1}^{t_n} \theta_r\right) \sum_{\ell=2}^n \frac{w_\ell^2}{W_\ell^2} \cdot \exp\left(\theta_{t_\ell} + \sum_{i=\ell+1}^n (e^{\theta t_i} - 1) \frac{w_i}{W_i} - \sum_{r=t_\ell+1}^{t_n} \theta_r\right). \end{aligned}$$

**Proof.** Applying the many-to-two lemma (Lemma 2.3), we get

$$(4.1) \quad \mathbb{E} [Q_n^2] = \sum_{\ell=1}^n \frac{w_\ell^2}{W_\ell^2} \cdot \left( \prod_{i=\ell+1}^n \left( 1 - \frac{w_i^2}{W_i^2} \right) \right) \cdot \mathbb{E} \left[ \mathbb{1}_{H_n^\ell=t_n} \mathbb{1}_{\forall r < t_n, H_{i_r}^\ell \leq r} \cdot \mathbb{1}_{\overline{H}_n^\ell=t_n} \mathbb{1}_{\forall r < t_n, \overline{H}_{i_r}^\ell \leq r} \right],$$

where  $H^\ell$  and  $\overline{H}^\ell$  are the walks associated to the jumps  $B_i^\ell$  and  $\overline{B}_i^\ell$ , jointly defined in Lemma 2.3. Let the  $(Y_i^\ell, \overline{Y}_i^\ell)$ 's for  $i \geq 2$  be independent couples of Bernoulli r.v., with a distribution defined by

$$\forall \phi: \{0, 1\}^2 \rightarrow \mathbb{R}, \quad \mathbb{E} [\phi(Y_i^\ell, \overline{Y}_i^\ell)] = \frac{\mathbb{E} [\phi(B_i^\ell, \overline{B}_i^\ell) \exp(\theta_{t_i}(B_i^\ell + \overline{B}_i^\ell \mathbb{1}_{i>\ell}))]}{\mathbb{E} [\exp(\theta_{t_i}(B_i^\ell + \overline{B}_i^\ell \mathbb{1}_{i>\ell}))]}.$$

Recalling the definition of  $p_i$  in (3.3), one can check that

- if  $i < \ell$ , then  $Y_i^\ell$  has distribution Bernoulli( $p_i$ ) and  $\overline{Y}_i^\ell = Y_i^\ell$ ;
- if  $i = \ell$ , then  $Y_i^\ell = \overline{Y}_i^\ell = 1$ ;
- if  $i > \ell$ , then  $Y_i^\ell$  and  $\overline{Y}_i^\ell$  have the same distribution as two independent Bernoulli( $p_i$ ) r.v. conditioned not to be both equal to 1.

We set  $S_0^\ell := 0$ ,  $\overline{S}_0^\ell := 0$  and, for  $r \geq 1$ ,

$$X_r^\ell := \sum_{i=i_{r-1}+1}^{i_r} Y_i^\ell, \quad S_r^\ell := \sum_{s=1}^r (X_s^\ell - 1), \quad \overline{X}_r^\ell := \sum_{i=i_{r-1}+1}^{i_r} \overline{Y}_i^\ell \quad \text{and} \quad \overline{S}_r^\ell := \sum_{s=1}^r (\overline{X}_s^\ell - 1).$$

Then, we use this change of measure to write

$$\begin{aligned} & \mathbb{E} \left[ \mathbb{1}_{H_n^\ell=t_n} \mathbb{1}_{\forall r < t_n, H_{i_r}^\ell \leq r} \cdot \mathbb{1}_{\overline{H}_n^\ell=t_n} \mathbb{1}_{\forall r < t_n, \overline{H}_{i_r}^\ell \leq r} \right] \\ &= \mathbb{E} \left[ \exp \left( \sum_{i=2}^n \theta_{t_i} (B_i^\ell + \overline{B}_i^\ell \mathbb{1}_{i>\ell}) \right) \right] \mathbb{E} \left[ \exp \left( - \sum_{i=2}^n \theta_{t_i} (Y_i^\ell + \overline{Y}_i^\ell \mathbb{1}_{i>\ell}) \right) \mathbb{1}_{S_{t_n}^\ell=0} \mathbb{1}_{\forall r < t_n, S_r^\ell \leq 0} \cdot \mathbb{1}_{\overline{S}_{t_n}^\ell=0} \mathbb{1}_{\forall r < t_n, \overline{S}_r^\ell \leq 0} \right]. \end{aligned}$$

For the first expectation, using the definition of  $(B_i^\ell, \overline{B}_i^\ell)$ , we have

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \sum_{i=2}^n \theta_{t_i} (B_i^\ell + \overline{B}_i^\ell \mathbb{1}_{i>\ell}) \right) \right] &= \prod_{i=2}^{\ell-1} \left( 1 + (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \right) \cdot e^{\theta_{t_\ell} \mathbb{1}_{\ell \geq 2}} \cdot \prod_{i=\ell+1}^n \left( 1 + (e^{\theta_{t_i}} - 1) \frac{2 \frac{w_i}{W_i}}{1 + \frac{w_i}{W_i}} \right) \\ &\leq \exp \left( \sum_{i=2}^{\ell-1} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} + \theta_{t_\ell} \mathbb{1}_{\ell \geq 2} + \sum_{i=\ell+1}^n 2(e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \right) \\ &\leq \exp \left( \sum_{i=2}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} + \theta_{t_\ell} \mathbb{1}_{\ell \geq 2} + \sum_{i=\ell+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \right). \end{aligned}$$

On the other hand, for the second expectation, we rewrite

$$\sum_{i=2}^n \theta_{t_i} Y_i^\ell = \sum_{r=1}^{t_n} \theta_r + \sum_{r=1}^{t_n} \theta_r (X_r^\ell - 1) = \sum_{r=1}^{t_n} \theta_r + \sum_{r=1}^{t_n-1} (\theta_r - \theta_{r+1}) S_r^\ell + \theta_{t_n} S_{t_n}^\ell - \theta_1 S_0^\ell,$$

with a summation by part, and similarly (noting that  $\ell + 1 \leq i_{t_\ell} + 1$  and using here that  $n = i_{t_n}$ )

$$\begin{aligned} \sum_{i=\ell+1}^n \theta_{t_i} \overline{Y}_i^\ell &= \sum_{i=\ell+1}^{i_{t_\ell}} \theta_{t_i} \overline{Y}_i^\ell + \sum_{i=i_{t_\ell}+1}^n \theta_{t_i} \overline{Y}_i^\ell = \sum_{i=\ell+1}^{i_{t_\ell}} \theta_{t_i} \overline{Y}_i^\ell + \sum_{r=t_\ell+1}^{t_n} \theta_r + \sum_{r=t_\ell+1}^{t_n} \theta_r (\overline{X}_r^\ell - 1) \\ &= \sum_{i=\ell+1}^{i_{t_\ell}+1} \theta_{t_i} \overline{Y}_i^\ell + \sum_{r=t_\ell+1}^{t_n} \theta_r + \sum_{r=t_\ell+1}^{t_n-1} (\theta_r - \theta_{r+1}) \overline{S}_r^\ell + \theta_{t_n} \overline{S}_{t_n}^\ell - \theta_{t_\ell+1} \overline{S}_{t_\ell}^\ell, \end{aligned}$$



Using the bounds  $\sum_{i=\ell+1}^{i_{t_\ell}+1} \theta_{t_i} \bar{Y}_i^\ell \geq 0$  and  $-\theta_{t_\ell+1} \bar{S}_{t_\ell}^\ell \geq 0$  (on the event  $S_{t_\ell}^\ell \leq 0$ ), and recalling that on the event of interest we have  $S_{t_n}^\ell = \bar{S}_{t_n}^\ell = 0$  we get

$$\begin{aligned} & \mathbb{E} \left[ \exp \left( - \sum_{i=2}^n \theta_{t_i} (Y_i^\ell + \bar{Y}_i^\ell \mathbb{1}_{i>\ell}) \right) \mathbb{1}_{S_{t_n}^\ell=0} \mathbb{1}_{\forall r < t_n, S_r^\ell \leq 0} \cdot \mathbb{1}_{\bar{S}_{t_n}^\ell=0} \mathbb{1}_{\forall r < t_n, \bar{S}_r^\ell \leq 0} \right] \\ & \leq \exp \left( - \sum_{r=1}^{t_n} \theta_r - \sum_{r=t_\ell+1}^{t_n} \theta_r \right) \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r^\ell + \sum_{r=t_\ell+1}^{t_n-1} (\theta_{r+1} - \theta_r) \bar{S}_r^\ell \right) \mathbb{1}_{S_{t_n}^\ell = \bar{S}_{t_n}^\ell = 0, \forall r < t_n, S_r^\ell \leq 0, \bar{S}_r^\ell \leq 0} \right]. \end{aligned}$$

Coming back to (4.1) and using the bound  $\prod_{i=\ell+1}^n (1 - w_i^2/W_i^2) \leq 1$ , we proved

$$(4.2) \quad \begin{aligned} \mathbb{E} [Q_n^2] & \leq \sum_{\ell=1}^n \frac{w_\ell^2}{W_\ell^2} \cdot \exp \left( \sum_{i=2}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} + \theta_{t_\ell} \mathbb{1}_{\ell \geq 2} + \sum_{i=\ell+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^{t_n} \theta_r - \sum_{r=t_\ell+1}^{t_n} \theta_r \right) \\ & \cdot \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r^\ell + \sum_{r=t_\ell+1}^{t_n-1} (\theta_{r+1} - \theta_r) \bar{S}_r^\ell \right) \mathbb{1}_{S_{t_n}^\ell = \bar{S}_{t_n}^\ell = 0, \forall r < t_n, S_r^\ell \leq 0, \bar{S}_r^\ell \leq 0} \right]. \end{aligned}$$

We now focus on the term  $\ell = 1$  (so that  $t_\ell = 0$ ). Recall the definition of  $(Y_i)_{i \geq 2}$  and  $(S_r)_{r \geq 0}$  in Lemma 3.1 and let  $(\bar{Y}_i)_{i \geq 2}$  and  $(\bar{S}_r)_{r \geq 0}$  be independent copies of them. Then, by definition of the  $(Y_i^1, \bar{Y}_i^1)$ 's, the expectation on the RHS of (4.2) equals (we use here that  $n = i_{t_n}$  so that the indicator function appearing in (3.4) does not play a role)

$$\begin{aligned} & \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r + \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) \bar{S}_r \right) \mathbb{1}_{S_{t_n} = \bar{S}_{t_n} = 0, \forall r < t_n, S_r \leq 0, \bar{S}_r \leq 0} \mid \forall i \in \llbracket 2, n \rrbracket, (Y_i, \bar{Y}_i) \neq (1, 1) \right] \\ & \leq \frac{\mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r + \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) \bar{S}_r \right) \mathbb{1}_{S_{t_n} = \bar{S}_{t_n} = 0, \forall r < t_n, S_r \leq 0, \bar{S}_r \leq 0} \right]}{\mathbb{P}(\forall i \in \llbracket 2, n \rrbracket, (Y_i, \bar{Y}_i) \neq (1, 1))} \\ & = \left( \prod_{i=2}^n (1 - p_i^2)^{-1} \right) \cdot \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r \right) \mathbb{1}_{S_{t_n} = 0, \forall r < t_n, S_r \leq 0} \right]^2. \end{aligned}$$

Therefore, the  $\ell = 1$  term on the RHS of (4.2) is at most

$$\begin{aligned} & \exp \left( \sum_{i=2}^n 2(e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^{t_n} 2\theta_r \right) \cdot \left( \prod_{i=2}^n (1 - p_i^2)^{-1} \right) \cdot \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n-1} (\theta_{r+1} - \theta_r) S_r \right) \mathbb{1}_{S_{t_n} = 0, \forall r < t_n, S_r \leq 0} \right]^2 \\ & \leq \left( \prod_{i=2}^n (1 - p_i^2)^{-1} \right) \cdot \exp \left( \sum_{i=2}^n e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2} \right) \cdot \mathbb{E} [Q_n]^2, \end{aligned}$$

by Lemma 4.1. For the  $\ell \geq 2$  terms on the RHS of (4.2), we simply bound the expectation by 1, using that  $(\theta_k)_{k \geq 1}$  is non-decreasing. This proves the result.  $\square$

### 4.3. General criterion for the lower bound

The proposition below is the main result of this section: it provides a general lower bound in terms of some quantities which we will have to control in each case.

**Proposition 4.3.** *Assume that there exist*

- a non-decreasing sequence  $(T(n))_{n \geq 2}$  of non-negative integers such that  $T(n) < t_n$  for any  $n \geq 2$ , and  $T(n) \rightarrow \infty$  as  $n \rightarrow \infty$ ;
- a sequence  $(\theta_r)_{r \geq 1}$  such that, for some  $r_0 \geq 1$ ,  $(\theta_r)_{r \geq r_0}$  is a non-decreasing sequence of non-negative numbers.

Let  $(S_r)_{r \geq 0}$  be defined as in (3.4). Define, for  $n \geq 2$ ,

$$E(n) := \mathbb{E} \left[ \exp \left( \sum_{r=1}^{t_n - T(n) - 1} (\theta_{T(n)+r+1} - \theta_{T(n)+r}) (S_{T(n)+r} - S_{T(n)}) \right) \mathbb{1}_{S_{t_n} = S_{T(n)}, \forall r < t_n - T(n), S_{T(n)+r} \leq S_{T(n)}} \right]$$

$$\delta_1(n) := \sum_{i \geq i_{T(n)}+1} e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2}$$

$$\delta_2(n) := \sum_{\ell \geq i_{T(n)}+1} e^{2\theta_{t_\ell}} \frac{w_\ell^2}{W_\ell^2} \cdot \exp \left( \sum_{r=T(n)+1}^{t_\ell-1} \theta_r - \sum_{i=i_{T(n)}+1}^{i_{t_\ell}-1} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \right).$$

If the following series are convergent

$$(4.3) \quad \sum_{t \geq 1} \delta_1(i_t) < \infty \quad \text{and} \quad \sum_{t \geq 1} \frac{\delta_2(i_t)}{E(i_t)^2} < \infty$$

then almost surely, for  $n$  large enough,  $\text{ht}(\mathbb{T}_n) \geq t_n - T(n) - 1$ .

Before diving into the proof, we provide some insight concerning the quantities appearing above. First,  $E(n)$  is the expectation involving the walk  $S$  appearing in Lemma 4.1, but for the auxiliary tree  $\mathbb{T}_n^{(N)}$  with  $N = i_{T(n)}$ . It always plays a negligible role, but we need a lower bound for it, which is the content of Lemma 4.5 established in Section 4.4. Then,  $\delta_1(n)$  and  $\delta_2(n)$  are small as a consequence of our assumption on  $\sum_{i=n}^{\infty} w_i^2/W_i^2$ . Moreover, for  $\delta_2(n)$ , we need the exponential term to be not too large. This amounts to requiring that the quantity

$$i_t \cdot \exp \left( \sum_{i=2}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r \right)$$

diverges fast enough. This has to be compared with the criterion for the upper bound (Proposition 3.2), where this quantity had to be summable. Recall that this quantity is approximately  $i_t \mathbb{E}[Q_{i_t}]$ , which is roughly the number of particles contributing to  $Q_{i_t}$ . This number needs to be large so that we have a good concentration in the first and second moment argument.

**Proof of Proposition 4.3.** We write  $T = T(n)$  and  $N = N(n) = i_T$ . We consider the trees  $(\mathbb{T}_n^{(N)})_{n \geq 1}$  introduced in Section 2.1. We also define  $i_0^{(N)} := 1$  and

$$i_r^{(N)} := i_{T+r}, \quad \theta_r^{(N)} := \theta_{T+r}, \quad \text{for } r \geq 1.$$

We define  $(t_i^{(N)})_{i \geq 1}$  as before, but associated to the sequence  $(i_r^{(N)})_{r \geq 1}$ . One can check that

$$(4.4) \quad t_i^{(N)} = \begin{cases} t_i - T, & \text{if } i \geq N + 1, \\ 1, & \text{if } 2 \leq i \leq N, \\ 0, & \text{if } i = 1. \end{cases}$$

We also define  $Q_n^{(N)}$  as in (3.1) but for the tree  $(\mathbb{T}_n^{(N)})_{n \geq 1}$  and the sequences  $(i_r^{(N)})_{r \geq 0}$  and  $(t_i^{(N)})_{i \geq 1}$ . By construction, we have  $\text{ht}(\mathbb{T}_n) \geq \text{ht}(\mathbb{T}_n^{(N)})$  and therefore,

$$(4.5) \quad \mathbb{P}(\text{ht}(\mathbb{T}_n) \geq t_n - T) \geq \mathbb{P}(\text{ht}(\mathbb{T}_n^{(N)}) \geq t_n^{(N)}) \geq \mathbb{P}(Q_n^{(N)} > 0) \geq \frac{\mathbb{E}[Q_n^{(N)}]^2}{\mathbb{E}[(Q_n^{(N)})^2]},$$

where the last inequality follows from the Cauchy–Schwarz inequality.

First assume that  $n = i_{t_n}$ . We apply Lemma 4.1 and Lemma 4.2 to  $Q_n^{(N)}$ , using that  $(\mathbb{T}_n^{(N)})_{n \geq 1}$  is a weighted recursive tree with weight sequence  $(w_n^{(N)})_{n \geq 1}$ . Recall the definition of  $p_i$  and  $(S_r)_{r \geq 0}$  in (3.3) and (3.4) and define similarly  $p_i^{(N)}$  and  $(S_r^{(N)})_{r \geq 0}$  but for the sequences  $(w_n^{(N)})_{n \geq 1}$ ,  $(i_r^{(N)})_{r \geq 0}$  and  $(\theta_r^{(N)})_{r \geq 1}$ . Note that, for  $2 \leq i \leq N$ ,  $w_i^{(N)} = 0$  and, for  $i \geq N + 1$ ,  $\theta_{t_i^{(N)}} = \theta_{t_i}$  and  $w_i^{(N)}/W_i^{(N)} = w_i/W_i$ , so we get

$$p_i^{(N)} = \begin{cases} p_i, & \text{if } i \geq N + 1, \\ 0, & \text{if } 2 \leq i \leq N, \end{cases} \quad \text{and} \quad (S_r^{(N)})_{r \geq 0} \stackrel{(d)}{=} (S_{T(n)+r} - S_{T(n)})_{r \geq 0}.$$

We consider  $n$  large enough in such a way that  $T \geq r_0$ , so that  $(\theta_r^{(N)})_{r \geq 1}$  is a non-decreasing sequence of non-negative numbers. Therefore, Lemma 4.1 implies

$$\begin{aligned} \mathbb{E} [Q_n^{(N)}] &\geq \exp \left( \sum_{i=N+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{i=N+1}^n \frac{e^{2\theta_{t_i}} w_i^2}{2 W_i^2} - \sum_{r=T+1}^{t_n} \theta_r \right) \cdot E(n) \\ &\geq \exp \left( \sum_{i=N+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=T+1}^{t_n} \theta_r \right) \cdot \frac{E(n)}{2}, \end{aligned}$$

for  $n$  large enough, using that  $\sum_{i=N+1}^n e^{2\theta_{t_i}} w_i^2 / W_i^2 \leq \delta_1(n)$  and  $\delta_1(n) \rightarrow 0$  as a consequence of (4.3). On the other hand, Lemma 4.2 yields

$$\begin{aligned} &\mathbb{E} \left[ (Q_n^{(N)})^2 \right] \\ &\leq \left( \prod_{i=N+1}^n (1 - p_i^2)^{-1} \right) \cdot \exp \left( \sum_{i=N+1}^n e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2} \right) \cdot \mathbb{E} [Q_n^{(N)}]^2 \\ &\quad + \exp \left( \sum_{i=N+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=T+1}^{t_n} \theta_r \right) \cdot \sum_{\ell=N+1}^n \frac{w_\ell^2}{W_\ell^2} \cdot \exp \left( \theta_{t_\ell} + \sum_{i=\ell+1}^n (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=t_\ell+1}^{t_n} \theta_r \right). \end{aligned}$$

The fact that  $\delta_1(n) \rightarrow 0$  ensures that  $e^{\theta_{t_i}} \frac{w_i}{W_i} \rightarrow 0$  as  $i \rightarrow \infty$  and therefore  $p_i \sim e^{\theta_{t_i}} \frac{w_i}{W_i}$ . Hence, for  $n$  large enough (and hence  $N$  large enough), we have

$$\left( \prod_{i=N+1}^n (1 - p_i^2)^{-1} \right) \cdot \exp \left( \sum_{i=N+1}^n e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2} \right) \leq \exp \left( 3 \sum_{k=N+1}^n e^{2\theta_{t_k}} \frac{w_k^2}{W_k^2} \right) \leq \exp(3\delta_1(n)) \leq 1 + 4\delta_1(n),$$

for  $n$  large enough, using again that  $\delta_1(n) \rightarrow 0$ . Combining what precedes, we get, for  $n$  large enough,

$$\begin{aligned} \frac{\mathbb{E} \left[ (Q_n^{(N)})^2 \right]}{\mathbb{E} [Q_n^{(N)}]^2} &\leq 1 + 4\delta_1(n) + \frac{4}{E(n)^2} \sum_{\ell=N+1}^n \frac{w_\ell^2}{W_\ell^2} \cdot \exp \left( \theta_{t_\ell} - \sum_{i=N+1}^{\ell} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} + \sum_{r=T+1}^{t_\ell} \theta_r \right) \\ &\leq 1 + 4\delta_1(n) + \frac{4}{E(n)^2} \cdot \delta_2(n), \end{aligned}$$

bounding  $\sum_{i=i_{t_\ell-1}+1}^{\ell} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \geq 0$ . Coming back to (4.5), this proves, for  $n$  large enough satisfying  $n = i_{t_n}$ ,

$$\mathbb{P}(\text{ht}(\mathbb{T}_n) \geq t_n - T(n)) \geq \left( 1 + 4\delta_1(n) + \frac{4\delta_2(n)}{E(n)^2} \right)^{-1}.$$

By assumption (4.3) and the Borel–Cantelli lemma, it follows that almost surely, for  $n$  large enough satisfying  $n = i_{t_n}$ , we have  $\text{ht}(\mathbb{T}_n) \geq t_n - T(n)$ .

We now want to get a lower bound for  $\text{ht}(\mathbb{T}_m)$  when  $m \in \llbracket i_{t_n} + 1, i_{t_n+1} - 1 \rrbracket$  for some  $n = i_{t_n}$ . In particular,  $t_m = t_n + 1$ . Since  $n \leq m$ , we have  $\text{ht}(\mathbb{T}_m) \geq \text{ht}(\mathbb{T}_n)$  and  $T(m) \geq T(n)$ . Therefore,

$$\text{ht}(\mathbb{T}_m) - (t_m - T(m) - 1) \geq \text{ht}(\mathbb{T}_n) - (t_n - T(n)).$$

The result follows.  $\square$

*Remark 4.4.* Under the assumptions of Proposition 4.3, one can also conclude that almost surely, for  $n$  large enough,  $\text{diam}(\mathbb{T}_n) \geq 2(t_n - T(n) - 1)$ , by following the same proof as the one of [20, Theorem 1.3]. Indeed, the key point in this argument is that, in the second moment of  $Q_n^{(N)}$ , the term  $\ell = 1$  is the dominant one: this means that two vertices chosen independently according to their weight in  $Q_n^{(N)}$  typically have the root as most recent common ancestor. Noting that the upper bound  $\text{diam}(\mathbb{T}_n) \geq 2 \text{ht}(\mathbb{T}_n)$  is direct, this implies that Theorem 1.1 and Theorem 1.4 also hold for  $\text{diam}(\mathbb{T}_n)$  by multiplying all the terms in the expansion by 2.

#### 4.4. Lower bound for the term $E(n)$

We conclude this section by stating and proving Lemma 4.5 below which provides a lower bound on the term  $E(n)$  that appears in the assumptions of Proposition 4.3.

**Lemma 4.5.** *Let  $(\theta_r)_{r \geq 1}$  be a non-decreasing sequence of non-negative numbers such that, for  $t$  large enough, we have  $e^{\theta_t}(a_t - a_{t-1}) = 1$ . Recall the definition of  $E(n)$  from Proposition 4.3. Assume that  $T(n) \rightarrow \infty$  and  $t_n - T(n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Also assume that*

$$(4.6) \quad \sum_{i \geq 1} e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2} < \infty.$$

Then, there exists a constant  $C > 0$ , such that for all  $n \geq 1$  satisfying  $n = i_{t_n}$  we have

1.  $E(n) \geq \exp(-C(t_n - T(n)))$ ;
2.  $E(n) \geq \exp\left(-(\theta_{t_n} - \theta_{T(n)}) \cdot (t_n - T(n))^{1/2} - C \log(t_n - T(n))\right) - C \sum_{i=i_{T(n)+1}}^n e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2}$ .

Note that it need not be the case that the second estimate is better than the first one as it depends on the rate of growth of  $\theta_t$ .

**Proof.** The two lower bounds are obtained in the same fashion, by identifying an event for the walk  $(S_{T(n)+r} - S_{T(n)})_{r \geq 0}$  for which we can give a lower bound for the exponential term in the expectation. Recall the definition of  $p_i$  from (3.3). It is easy to check that we have

$$(4.7) \quad p_i = e^{\theta_{t_i}} \frac{w_i}{W_i} + O\left(e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2}\right) \quad \text{and} \quad p_i^2 = O\left(e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2}\right),$$

using in particular that  $e^{\theta_{t_i}} \frac{w_i}{W_i} \rightarrow 0$  as a consequence of assumption (4.6).

We start by proving 1. Noting that on the event  $\{\forall r \leq t_n - T(n), S_{T(n)+r} = S_{T(n)}\}$  the exponential term in  $E(n)$  equals 1, we get

$$(4.8) \quad E(n) \geq \mathbb{P}(\forall r \leq t_n - T(n), S_{T(n)+r} = S_{T(n)}) = \prod_{r=1}^{t_n - T(n)} \mathbb{P}(X_{T(n)+r} = 1).$$

Then, for any  $t \geq 1$ , recalling that  $X_t = \sum_{i=i_{t-1}+1}^{i_t} Y_i$ , where the  $Y_i$ 's are independent Bernoulli( $p_i$ ) r.v., we have

$$(4.9) \quad \mathbb{P}(X_t = 1) = \sum_{i=i_{t-1}+1}^{i_t} p_i \prod_{\substack{j=i_{t-1}+1 \\ j \neq i}}^{i_t} (1 - p_j) \geq \left( \sum_{i=i_{t-1}+1}^{i_t} p_i \right) \cdot \left( \prod_{i=i_{t-1}+1}^{i_t} (1 - p_i) \right).$$

Using (4.7) and then (4.6), we get that as  $t \rightarrow \infty$

$$\sum_{i=i_{t-1}+1}^{i_t} p_i = \sum_{i=i_{t-1}+1}^{i_t} e^{\theta_{t_i}} \frac{w_i}{W_i} + \sum_{i=i_{t-1}+1}^{i_t} O\left(e^{2\theta_{t_i}} \frac{w_i^2}{W_i^2}\right) = e^{\theta_t}(a_t - a_{t-1}) + o(1) = 1 + o(1),$$

using that for  $t$  large enough  $e^{\theta_t}(a_t - a_{t-1}) = 1$ . Moreover, using the fact that  $p_i \rightarrow 0$ , the Taylor expansion  $\log(1 - x) \underset{x \rightarrow 0}{=} -x + O(x^2)$  and the previous display we get

$$\left( \prod_{i=i_{t-1}+1}^{i_t} (1 - p_i) \right) = \exp\left(-\left( \sum_{i=i_{t-1}+1}^{i_t} p_i \right) + O\left( \sum_{i=i_{t-1}+1}^{i_t} p_i^2 \right)\right) = \exp(-1 + o(1)).$$

Coming back to (4.9), this shows that  $\mathbb{P}(X_t = 1)$  is bounded below by some positive value for  $t$  sufficiently large. Plugging this back into (4.8), we get part 1 of the lemma.

In order to prove 2, we first compare  $(S_r)$  with a time-homogeneous random walk  $(\widehat{S}_r)$  that has Poisson(1) – 1 increments. Applying [20, Lemma A.1] to our case, we can couple  $(S_r)$  and  $(\widehat{S}_r)$  in such a way that

$$\begin{aligned} \mathbb{P}\left(\exists r \in \llbracket 0, t_n - T(n) \rrbracket, S_{T(n)+r} - S_{T(n)} \neq \widehat{S}_r\right) &\leq 2 \sum_{t=T(n)+1}^{t_n} \left| \sum_{i=i_{t-1}+1}^{i_t} p_i - 1 \right| + 2 \sum_{i=i_{T(n)}+1}^n p_i^2 \\ &\leq 2 \sum_{s=T(n)+1}^{t_n} |e^{\theta_s}(a_{i_s} - a_{i_{s-1}}) - 1| + C \sum_{i=i_{T(n)}+1}^n e^{2\theta_i} \frac{w_i^2}{W_i^2}, \end{aligned}$$

where the first sum on the right-hand side is zero for  $n$  large enough. Then, setting  $r_n = t_n - T(n)$  for brevity, it is enough to prove that

$$(4.10) \quad \mathbb{E} \left[ \exp \left( \sum_{r=1}^{r_n-1} (\theta_{T(n)+r+1} - \theta_{T(n)+r}) \widehat{S}_r \right) \mathbb{1}_{\widehat{S}_{r_n}=0, \forall r < r_n, \widehat{S}_r \leq 0} \right] \geq \exp \left( -(\theta_{t_n} - \theta_{T(n)}) \cdot r_n^{1/2} - C \log r_n \right).$$

For this, we will prove that

$$(4.11) \quad \liminf_{n \rightarrow \infty} r_n^{3/2} \cdot \mathbb{P} \left( \widehat{S}_{r_n} = 0 \text{ and } \forall r < r_n, -r_n^{1/2} \leq \widehat{S}_r \leq 0 \right) > 0,$$

which implies (4.10) by restricting ourselves to this event and bounding  $\widehat{S}_r \geq -r_n^{1/2}$  inside the exponential. We now prove (4.11). First, according to [7, Equation (4.6)] (note that the probability below equals  $q_{r_n}^+(0, 0)$  for the walk  $-\widehat{S}$  with their notation), we have, for some constant  $C_1 > 0$ , as  $n \rightarrow \infty$ ,

$$(4.12) \quad \mathbb{P} \left( \widehat{S}_{r_n} = 0 \text{ and } \forall r < r_n, \widehat{S}_r \leq 0 \right) \sim \frac{C_1}{r_n} \cdot \mathbb{P} \left( \widehat{S}_{r_n} = 0 \right) \sim \frac{C_1}{\sqrt{2\pi} \cdot r_n^{3/2}}.$$

Now, let  $(\widehat{\mathbf{S}}_u^{(n)})_{u \in [0,1]}$  denotes the process defined by linear interpolation of the points  $(r/r_n, \widehat{S}_r/r_n^{1/2})$  for  $0 \leq r \leq r_n$ . Then, [7, Corollary 2.5] proves that the process  $-\widehat{\mathbf{S}}^{(n)}$ , conditionally on the event  $\{\widehat{S}_{r_n} = 0 \text{ and } \forall r < r_n, -r_n^{1/2} \leq \widehat{S}_r \leq 0\}$ , converges in distribution in  $(C([0, 1], \mathbb{R}), \|\cdot\|_\infty)$  toward the normalized Brownian excursion  $(\mathbf{e}_u)_{u \in [0,1]}$ . In particular,

$$(4.13) \quad \mathbb{P} \left( \forall r < r_n, \widehat{S}_r \geq -r_n^{1/2} \mid \widehat{S}_{r_n} = 0 \text{ and } \forall r < r_n, \widehat{S}_r \leq 0 \right) \xrightarrow{n \rightarrow \infty} \mathbb{P} \left( \max_{u \in [0,1]} \mathbf{e}_u \leq 1 \right) > 0,$$

using that the map  $f \in C([0, 1], \mathbb{R}) \mapsto \max_{[0,1]} f$  is continuous. Combining (4.12) and (4.13) shows (4.11) and this concludes the proof.  $\square$

## 5. Application to different regimes for the weight sequence

In this section, we apply the general results of Section 3 and Section 4 to specific regimes for the weight sequence. We first present the general ideas and then study the different regimes separately.

### 5.1. Strategy

The proofs of Theorem 1.1 and Theorem 1.4 follow from applying Proposition 3.2 and Proposition 4.3 to well-chosen sequences  $(i_t)_{t \geq 0}$  and  $(\theta_t)_{t \geq 1}$ . As discussed before,  $(i_t)_{t \geq 0}$  should be chosen such that  $\text{ht}(T_{i_t})$  is either slightly less (upper bound) or slightly more (lower bound) than  $t$ . For the reasons explained in Section 3.1, we will always choose  $\theta_t = -\log(a_{i_t} - a_{i_{t-1}})$  for  $t$  large enough. Then, the proof is mainly computational. We get an asymptotic expansion for  $\theta_t$  and use it to check the assumptions of Proposition 3.2 and Proposition 4.3. In particular, this expansion has to be precise enough to ensure that the sequence  $(\theta_t)$  is eventually non-decreasing.

Choosing  $(i_t)_{t \geq 0}$  is not difficult if one already knows the expansion of  $\text{ht}(T_n)$  that one tries to prove. However, it may not be immediate to find a good conjecture for the terms in this expansion in the first place so we explain here how we guessed the behavior of  $\text{ht}(T_n)$  in the different cases that we studied. For Theorem 1.1, when  $\alpha \leq 1$ , we started with the crude upper and lower bounds presented in Section A, which match at the first order. This first order implies the choice  $i_t = \exp(\alpha t \log t(1 + o(1)))$ , but this is already enough to deduce the expansion (5.9) for  $\theta_t$ . Then, recall we want (3.6)

to be summable for the upper bound and sufficiently quickly divergent for the lower bound, but to get precise bounds we want the first orders to cancel out. Hence the precise definition of  $i_t$  in (5.1) and (5.7) includes exactly the right terms to cancel those coming from the exponential in (3.6). When  $\alpha > 1$ , the crude upper bound does not give the correct first order, but we simply used the same definition for  $i_t$ . Also note that, as claimed in Remark 1.3, one could get the next orders in  $\text{ht}(\mathbb{T}_n)$ . Indeed, this precise definition of  $i_t$  implies a more precise expansion for  $\theta_t$ , so plugging it in (3.6), one can find what the next orders in the definition of  $i_t$  should be, and so on.

For Theorem 1.4, we trusted the first order given by the crude lower bound given in Section A.2 to find the definition of  $i_t$ . Indeed, it is rather intuitive that the method used for this crude lower bound should work better when  $w_n$  tends faster to 0. Hence, since it is yielding the right first order for Theorem 1.1, it had to be also correct for Theorem 1.4.

## 5.2. Variance varying like powers of $\log n$

Our aim in this section is to prove Theorem 1.1. We work under assumption (1.4) and (1.5). We first define and study a sequence  $\ell_t$ , which is then used to set  $i_t \simeq \exp(\ell_t)$  later.

**Lemma 5.1.** *Let  $\eta \in \mathbb{R}$ . For  $t \geq 2$ , we define*

$$(5.1) \quad \ell_t := \alpha t \log t - (1 - \alpha)t \log \log t - t \log L(\alpha t \log t) - t[1 + \alpha + (1 - \alpha) \log \alpha + \eta].$$

Then, we have the following asymptotics, as  $t \rightarrow \infty$ ,

1.  $\ell_t \sim \alpha t \log t$ ;
2.  $\ell_t - \ell_{t-1} \sim \alpha \log t$ ;
3.  $\log \ell_t - \log \ell_{t-1} \sim \frac{1}{t}$ ;
4.  $\log(\ell_{t+1} - \ell_t) - \log(\ell_t - \ell_{t-1}) = o\left(\frac{1}{t}\right)$ .

**Proof.** Part 1 follows directly from the fact that  $\log L(x) = o(\log x)$  (see Lemma B.11).

For Part 2, we prove a more precise expansion that will be useful for Part 4. We introduce  $\tilde{\ell}_t := \ell_t + t \log L(\alpha t \log t)$ . One can check through explicit calculation that

$$(5.2) \quad \tilde{\ell}_t - \tilde{\ell}_{t-1} = \alpha \log t - (1 - \alpha) \log \log t - [1 + (1 - \alpha) \log \alpha + \eta] - \frac{1 - \alpha}{\log t} + O\left(\frac{1}{t}\right),$$

so it remains to deal with the part involving  $L$ . We decompose

$$(5.3) \quad \begin{aligned} & t \log L(\alpha t \log t) - (t-1) \log L(\alpha(t-1) \log(t-1)) \\ &= \log L(\alpha t \log t) + (t-1) [\log L(\alpha t \log t) - \log L(\alpha(t-1) \log(t-1))]. \end{aligned}$$

Then, using Lemma B.13, the function  $\tilde{L}(x) = \exp(xL'(x)/L(x))$  satisfies (SV<sub>1</sub>) and we have

$$(5.4) \quad \begin{aligned} & \log L(\alpha t \log t) - \log L(\alpha(t-1) \log(t-1)) \\ &= \log \tilde{L}(\alpha t \log t) \cdot \log\left(\frac{t \log t}{(t-1) \log(t-1)}\right) + o\left(\log^2\left(\frac{t \log t}{(t-1) \log(t-1)}\right)\right) \\ &= \log \tilde{L}(\alpha t \log t) \cdot \left[\frac{1}{t} + \frac{1}{t \log t} + O\left(\frac{1}{t^2}\right)\right] + o\left(\frac{1}{t^2}\right). \end{aligned}$$

Hence, combining this with (5.2) and (5.3), we proved

$$(5.5) \quad \begin{aligned} \ell_t - \ell_{t-1} &= \alpha \log t - (1 - \alpha) \log \log t - \log L(\alpha t \log t) - \log \tilde{L}(\alpha t \log t) \cdot \left[1 + \frac{1}{\log t} + O\left(\frac{1}{t}\right)\right] \\ &\quad - [1 + (1 - \alpha) \log \alpha + \eta] - \frac{1 - \alpha}{\log t} + O\left(\frac{1}{t}\right). \end{aligned}$$

Using  $\log L(x) = o(\log x)$  and  $\log \tilde{L}(x) = o(\log x)$ , this proves  $\ell_t - \ell_{t-1} \sim \alpha \log t$ .

For Part 3, we simply have

$$\log \ell_t - \log \ell_{t-1} = \log\left(1 - \frac{\ell_t - \ell_{t-1}}{\ell_{t-1}}\right) \sim \frac{\ell_t - \ell_{t-1}}{\ell_{t-1}} \sim \frac{1}{t},$$



using Parts 1 and 2.

For Part 4, we introduce  $f(t) := \alpha \log t - (1 - \alpha) \log \log t - [1 + (1 - \alpha) \log \alpha + \eta] - \frac{1 - \alpha}{\log t}$  and  $g(t) = \log L(\alpha t \log t) - \log \tilde{L}(\alpha t \log t) \cdot [1 + \frac{1}{\log t}]$ , so that (5.5) yields

$$\log(\ell_t - \ell_{t-1}) = \log \left( f(t) + g(t) + O \left( \frac{1 + \log \tilde{L}(\alpha t \log t)}{t} \right) \right) = \log f(t) + \log \left( 1 + \frac{g(t)}{f(t)} + o \left( \frac{1}{t} \right) \right),$$

using that  $f(t) \sim \alpha \log t$  and  $\log \tilde{L}(\alpha t \log t) = o(\log t)$ . Note that  $g(t)/f(t) \rightarrow 0$  so

$$\log \left( 1 + \frac{g(t)}{f(t)} + o \left( \frac{1}{t} \right) \right) = \log \left( \left( 1 + \frac{g(t)}{f(t)} \right) \left( 1 + o \left( \frac{1}{t} \right) \right) \right) = \log \left( 1 + \frac{g(t)}{f(t)} \right) + o \left( \frac{1}{t} \right).$$

Proceeding similarly, we also have

$$\log \left( 1 + \frac{g(t+1)}{f(t+1)} \right) = \log \left( 1 + \frac{g(t)}{f(t)} \right) + O \left( \frac{g(t+1)}{f(t+1)} - \frac{g(t)}{f(t)} \right).$$

combining what precedes yields

$$(5.6) \quad \log(\ell_{t+1} - \ell_t) - \log(\ell_t - \ell_{t-1}) = \log \frac{f(t+1)}{f(t)} + O \left( \frac{g(t+1)}{f(t+1)} - \frac{g(t)}{f(t)} \right) + o \left( \frac{1}{t} \right).$$

On the one hand, we have through explicit calculation

$$\log \frac{f(t+1)}{f(t)} = \log \left( 1 + \frac{f(t+1) - f(t)}{f(t)} \right) = O \left( \frac{1}{t \log t} \right).$$

On the other hand, it follows from Lemma B.12 that  $\log L(\alpha(t+1) \log(t+1)) - \log L(\alpha t \log t) = o(1/t)$  and the same holds for  $\tilde{L}$  (because  $\tilde{L}$  satisfies (SV<sub>1</sub>)). Combining these facts, we get

$$\frac{g(t+1)}{f(t+1)} - \frac{g(t)}{f(t)} = \frac{g(t+1) - g(t)}{f(t+1)} + g(t) \frac{f(t) - f(t+1)}{f(t)f(t+1)} = o \left( \frac{1}{t} \right).$$

Coming back to (5.6), this proves Part 4. □

As a consequence of Parts 1 and 2 of Lemma 5.1, there exists  $s_0$  such that,

$$(5.7) \quad i_t := \begin{cases} t+1, & \text{if } 0 \leq t < s_0, \\ \lfloor \exp(\ell_t) \rfloor, & \text{if } t \geq s_0, \end{cases}$$

defines an increasing sequence of integers. Indeed, we can choose  $s_0$  large enough such that, for any  $t \geq s_0$ ,  $\lfloor \exp(\ell_t) \rfloor \geq t$  and  $\ell_{t+1} - \ell_t \geq 1$ . Moreover, we will assume that  $s_0$  is chosen large enough such that, for  $t \geq s_0$ , we have  $a_{i_t} - a_{i_{t-1}} > 0$ . The fact that this last condition can be met is a consequence of the proof of the next lemma (see (5.11)).

**Lemma 5.2.** *Let  $\eta \in \mathbb{R}$ . Let  $i_t$  be defined as in (5.7) for some  $s_0$  and*

$$(5.8) \quad \theta_t := \begin{cases} 0, & \text{if } 1 \leq t \leq s_0, \\ -\log(a_{i_t} - a_{i_{t-1}}), & \text{if } t > s_0. \end{cases}$$

*Then, there exists  $r_0 \geq s_0$  such that  $(\theta_t)_{t \geq r_0}$  is a non-decreasing sequence of non-negative numbers. Moreover, as  $t \rightarrow \infty$ ,*

$$(5.9) \quad \theta_t = \alpha \log t - (1 - \alpha) \log \log t - (1 - \alpha) \log \alpha - \log L(\alpha t \log t) + o(1).$$

**Proof.** Consider  $t > s_0$ , so that  $i_t = \lfloor \exp(\ell_t) \rfloor$  and  $i_{t-1} = \lfloor \exp(\ell_{t-1}) \rfloor$ . Recall also from Lemma 5.1 that  $\ell_t \sim \alpha t \log t$  and  $\ell_{t+1} - \ell_t \sim \alpha \log t$ . It follows from (1.4) that

$$(5.10) \quad \begin{aligned} a_{i_t} - a_{i_{t-1}} &= \int_{\log i_{t-1}}^{\log i_t} x^{-\alpha} L(x) dx + o \left( (\log i_t)^{-1-\alpha} (\log \log i_t)^2 L(\log i_t) \right) \\ &= \int_{\ell_{t-1}}^{\ell_t} x^{-\alpha} L(x) dx + o \left( \frac{1}{t} (\ell_t - \ell_{t-1}) \ell_t^{-\alpha} L(\ell_t) \right), \end{aligned}$$

using that  $\log i_t - \ell_t = O(1/i_t)$  to change the endpoints of the integral. Then, Lemma B.2 yields

$$(5.11) \quad \begin{aligned} a_{i_t} - a_{i_{t-1}} &= \ell_t^{-\alpha} (\ell_t - \ell_{t-1}) L(\ell_t) \left( 1 + \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t} (1 + o(1)) + o\left(\frac{1}{t}\right) \right) \\ &= \ell_t^{-\alpha} (\ell_t - \ell_{t-1}) L(\ell_t) \left( 1 + \frac{\alpha}{2t} + o\left(\frac{1}{t}\right) \right). \end{aligned}$$

It follows that

$$\theta_t = \alpha \log \ell_t - \log(\ell_t - \ell_{t-1}) - \log L(\ell_t) - \frac{\alpha}{2t} + o\left(\frac{1}{t}\right).$$

It follows from Lemma B.12 that  $\log L(\ell_t) = \log L(\alpha t \log t) + o(1)$  and so the expansion (5.9) follows from Lemma 5.1.1-2. Hence,  $\theta_t \sim \alpha \log t$  and so  $\theta_t \geq 0$  for  $t$  large enough. Moreover, applying again Lemma B.12, we have  $\log L(\ell_{t+1}) - \log L(\ell_t) = o(1/t)$ , so using Lemma 5.1.3-4, we get

$$\theta_{t+1} - \theta_t = \frac{\alpha}{t} + o\left(\frac{1}{t}\right),$$

which shows the sequence  $(\theta_t)$  is eventually non-decreasing.  $\square$

We now estimate the asymptotic behavior of the quantities appearing in the assumptions of Proposition 3.2 and Proposition 4.3.

**Lemma 5.3.** *Let  $\eta \in \mathbb{R}$ . Let  $i_t$  be defined as in (5.7) and  $\theta_t$  as in (5.8). Then, as  $t \rightarrow \infty$ ,*

$$(5.12) \quad i_t \cdot \exp\left(\sum_{i=2}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r\right) = \exp(-\eta t + o(t)).$$

Moreover, as  $s \rightarrow \infty$ , uniformly in  $t \geq s$ ,

$$(5.13) \quad \frac{i_t}{i_s} \cdot \exp\left(\sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=s+1}^t \theta_r\right) = \exp(-\eta(t-s) + o(t-s)).$$

**Proof.** We focus on the proof of (5.13). Then, (5.12) follows from the same lines (or it can be seen as a consequence of (5.13) by letting  $s \rightarrow \infty$  slowly enough in comparison with  $t$ ). The case  $s = t$  is obvious so we assume  $t \geq s + 1$ , so that  $o(1)$  terms can be included in  $o(t-s)$ . In the following,  $o(\dots)$  terms are meant to hold as  $s \rightarrow \infty$  uniformly in  $t \geq s + 1$ . We first estimate  $\sum_{r=s+1}^t \theta_r$ , by summing the expansion of  $\theta_r$  in (5.9). Using standard integral comparison, we have

$$\begin{aligned} \sum_{r=s+1}^t \log r &= t \log t - t - s \log s + s + o(t-s), \\ \sum_{r=s+1}^t \log \log r &= t \log \log t - s \log \log s + o(t-s). \end{aligned}$$

For the term involving  $L$  we use a summation by part to get

$$\begin{aligned} &\sum_{r=s+1}^t \log L(\alpha r \log r) \\ &= \sum_{r=s+2}^{t-1} r [\log L(\alpha r \log r) - \log L(\alpha(r+1) \log(r+1))] + t \log L(\alpha t \log t) - s \log L(\alpha(s+1) \log(s+1)) \\ &= t \log L(\alpha t \log t) - s \log L(\alpha s \log s) + o(t-s), \end{aligned}$$

where we used that  $\log L(\alpha r \log r) - \log L(\alpha(r+1) \log(r+1)) = o(1/r)$  by Lemma B.12, to bound the sum and to replace  $\log L(\alpha(s+1) \log(s+1))$  by  $\log L(\alpha s \log s)$ . Combining this yields

$$\begin{aligned} \sum_{r=s+1}^t \theta_r &= \alpha t \log t - (1-\alpha)t \log \log t - t \log L(\alpha t \log t) - t[\alpha + (1-\alpha) \log \alpha] \\ &\quad - (\alpha s \log s - (1-\alpha)s \log \log s - s \log L(\alpha s \log s) - s[\alpha + (1-\alpha) \log \alpha]) + o(t-s). \end{aligned}$$

On the other hand, using the definition of  $a_n$  and then recalling that  $e^{\theta_r}(a_{i_r} - a_{i_{r-1}}) = 1$ , we have

$$\sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} = \sum_{r=s+1}^t (e^{\theta_r} - 1)(a_{i_r} - a_{i_{r-1}}) = \sum_{r=s+1}^t (1 + o(1)) = t - s + o(t-s).$$

Combining this with  $i_t/i_s = \exp(\ell_t - \ell_s + o(1))$  and the definition of  $\ell_t$  in (5.1) yields (5.13).  $\square$

**Proof of Theorem 1.1.** Let  $i_t$  and  $\theta_t$  defined as in (5.7) and (5.8) for some fixed  $\eta \in \mathbb{R}$ . Recall that the sequence of integers  $(t_n)_{n \geq 1}$  is defined at the beginning of Section 3.1. One can check that, as  $n \rightarrow \infty$ ,

$$t_n = \frac{\log n}{\alpha \log \log n} \left( 1 + \frac{\log L(\log n) + \log \log \log n}{\alpha \log \log n - \log L(\log n) - \log \log \log n} + \frac{1 + \alpha + \log \alpha + \eta}{\alpha \log \log n} + o\left(\frac{1}{\log \log n}\right) \right).$$

*Upper bound.* We take  $\eta > 0$  so that it follows from (5.12) that (3.6) is satisfied. Hence, we can apply Proposition 3.2 to get  $\sup_{n \geq 1} (\text{ht}(u_n) - t_n) < \infty$  almost surely. Letting  $\eta \rightarrow 0$  yields the upper bound.

*Lower bound.* We fix some  $\eta < 0$  and we aim at checking that (4.3) is satisfied to apply Proposition 4.3 with  $T(n) := (\log n)^{4/5}$ . Using assumption (1.5), we have

$$\delta_1(n) = \sum_{r \geq T(n)+1} e^{2\theta_r} \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} \leq \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot O\left(\frac{1}{i_{r-1}}\right) = \sum_{r \geq T(n)+1} \exp(-\alpha r \log r(1+o(1))),$$

using  $i_{r-1} = \exp(\alpha r \log r(1+o(1)))$  and  $\theta_r \sim \alpha \log r$ . It follows that

$$\sum_{t \geq 1} \delta_1(i_t) = \sum_{t \geq 1} O(\exp(-\alpha T(i_t))) = \sum_{t \geq 1} O\left(\exp\left(-\alpha t^{4/5}\right)\right) < \infty,$$

proving the first part of (4.3). We now aim at checking the second part of (4.3). We have  $T(i_t) \sim (\alpha t \log t)^{4/5}$  and, recalling  $\theta_t \sim \alpha \log t$ ,  $\theta_{T(i_t)} \sim \frac{4}{5} \alpha \log t$ . Hence, we get, for any  $C > 0$ ,

$$(\theta_t - \theta_{T(i_t)}) \cdot (t - T(i_t))^{3/4} - C \log(t - T(i_t)) \sim \frac{1}{5} \alpha t^{3/4} \log t$$

and it follows from the second part of Lemma 4.5 combined with (1.5) that

$$(5.14) \quad E(i_t) \geq \exp\left(-\frac{1}{5} \alpha t^{3/4} \log t(1+o(1))\right) - \frac{C}{i_{T(i_t)}} \geq \exp\left(-\frac{1}{5} \alpha t^{3/4} \log t(1+o(1))\right),$$

because  $1/i_{T(i_t)} = \exp(-\alpha(\alpha t \log t)^{4/5} \cdot \frac{4}{5} \log t \cdot (1+o(1)))$ . On the other hand, using (1.5) and (5.13), we have

$$\begin{aligned} \delta_2(n) &= \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot \exp\left(\sum_{s=T(n)+1}^{r-1} \theta_s - \sum_{i=i_{T(n)}+1}^{i_{r-1}} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i}\right) \cdot \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} \\ &\leq \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot \frac{i_{r-1}}{i_{T(n)}} \exp(\eta(r-1-T(n)) + o(r-1-T(n))) \cdot O\left(\frac{1}{i_{r-1}}\right) \\ (5.15) \quad &= \frac{1}{i_{T(n)}} \sum_{s \geq 0} \exp(2\theta_{s+T(n)+1} + \eta s + o(s) + O(1)). \end{aligned}$$

Then, by (5.9), for  $n$  large enough, we have  $\theta_{s+T(n)+1} \leq 2\alpha \log(s+T(n)+1) \leq 2\alpha(\log(s+1) + \log T(n))$ . Summing over  $s$ , this yields  $\delta_2(n) = O(T(n)^{2\alpha}/i_{T(n)})$ . Combining this with (5.14), we get

$$\begin{aligned} \sum_{t \geq 1} \frac{\delta_2(i_t)}{E(i_t)^2} &= \sum_{t \geq 1} \frac{T(i_t)^{2\alpha}}{i_{T(i_t)}} \exp\left(-\frac{1}{5}\alpha t^{3/4} \log t(1+o(1))\right) \\ &= \sum_{t \geq 1} \exp\left(-\alpha(\alpha t \log t)^{4/5} \cdot \frac{4}{5} \log t \cdot (1+o(1))\right) < \infty, \end{aligned}$$

using previous estimates to note that  $1/i_t$  is the dominating factor. This proves the second part of (4.3). Hence, we can apply Proposition 4.3 to get that almost surely, for  $n$  large enough,  $\text{ht}(T_n) \geq t_n - T(n) - 1$ . Letting  $\eta \rightarrow 0$  yields the lower bound.  $\square$

### 5.3. Quickly converging sequences

The goal of this section is to prove Theorem 1.4. We work under assumption (1.6) and one of the assumptions (1.5), (1.7) or (1.8), depending on the value of  $\beta$ . We consider the cases  $\beta \in (0, 1)$ ,  $\beta = 1$  and  $\beta \in (1, \infty)$  in separate subsections. All those subsections are organized in the same fashion as the previous one.

#### 5.3.1. Case $\beta \in (0, 1)$

We start by introducing a sequence  $(\ell_t)$  meant to satisfy  $\ell_t \simeq \exp(\log^\beta i_t)$ , which is the quantity appearing in assumption (1.6) when  $n = i_t$ .

**Lemma 5.4.** *Let  $\kappa > 0$ . For  $t \geq 1$ , we define*

$$(5.16) \quad \ell_t := \exp\left(\kappa \cdot t^{\frac{\beta}{1-\beta}}\right).$$

Then, as  $t \rightarrow \infty$ ,

1.  $\log \ell_t - \log \ell_{t-1} = \frac{\kappa\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}} \cdot (1 + O(\frac{1}{t}))$ ;
2. If  $\beta \in (0, \frac{1}{2})$ , then  $\ell_t - \ell_{t-1} = \ell_t \cdot \frac{\kappa\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}} \cdot \left(1 - \frac{\kappa\beta}{2(1-\beta)} t^{\frac{2\beta-1}{1-\beta}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)\right)$ .

**Proof.** We start by noting that

$$t^{\frac{\beta}{1-\beta}} - (t-1)^{\frac{\beta}{1-\beta}} = t^{\frac{\beta}{1-\beta}} \cdot \left(1 - \left(1 - \frac{1}{t}\right)^{\frac{\beta}{1-\beta}}\right) = \frac{\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}} \cdot \left(1 + O\left(\frac{1}{t}\right)\right).$$

Plugging the last display in the expression for  $\ell_t$  proves Part 1.

Now assume that  $\beta \in (0, \frac{1}{2})$ . Then,  $\frac{2\beta-1}{1-\beta} < 0$ , so it follows from 1 that  $\log \ell_{t-1} - \log \ell_t \rightarrow 0$  or equivalently  $(\ell_t - \ell_{t-1})/\ell_t \rightarrow 0$ . Therefore, the following expansion holds

$$\begin{aligned} \frac{\ell_t - \ell_{t-1}}{\ell_t} &= 1 - \exp(\log \ell_{t-1} - \log \ell_t) \\ &= -(\log \ell_{t-1} - \log \ell_t) - \frac{1}{2}(\log \ell_{t-1} - \log \ell_t)^2 + o\left((\log \ell_{t-1} - \log \ell_t)^2\right) \\ &= \frac{\kappa\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}} \cdot \left(1 - \frac{\kappa\beta}{2(1-\beta)} t^{\frac{2\beta-1}{1-\beta}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)\right), \end{aligned}$$

using 1 and that  $O(t^{-1}) = o\left(t^{\frac{2\beta-1}{1-\beta}}\right)$ . This proves 2.  $\square$

We now define

$$(5.17) \quad i_t := \begin{cases} t+1, & \text{if } 0 \leq t < s_0, \\ \left\lfloor \exp\left((\log \ell_t)^{\frac{1}{\beta}}\right) \right\rfloor = \left\lfloor \exp\left(\kappa^{\frac{1}{\beta}} t^{\frac{1}{1-\beta}}\right) \right\rfloor, & \text{if } t \geq s_0, \end{cases}$$

by choosing some  $s_0$  large enough so that  $(i_t)_{t \geq 1}$  is increasing. Moreover, we will assume that  $s_0$  is chosen large enough such that, for  $t \geq s_0$ , we have  $a_{i_t} - a_{i_{t-1}} > 0$ . As in the previous case, the fact that this last condition can be met is a consequence of the proof of the next lemma.

**Lemma 5.5.** *Let  $\kappa > 0$ . Let  $i_t$  be defined as in (5.17) and*

$$(5.18) \quad \theta_t := \begin{cases} 0, & \text{if } 1 \leq t \leq s_0, \\ -\log(a_{i_t} - a_{i_{t-1}}), & \text{if } t > s_0. \end{cases}$$

*Then, there exists  $r_0 \geq s_0$  such that  $(\theta_t)_{t \geq r_0}$  is a non-decreasing sequence of non-negative numbers. Moreover, as  $t \rightarrow \infty$ ,*

$$(5.19) \quad \theta_t \sim (\alpha - 1)\kappa \cdot t^{\frac{\beta}{1-\beta}}$$

**Proof.** First note that the  $o(\dots)$  term appearing in assumption (1.6) for  $a_n$ , when taking  $n = i_{t-1}$ , equals

$$o\left(\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot (\log i_{t-1})^{(4\beta-2) \wedge 0}\right) = o\left(\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot \left(t^{2 \cdot \frac{2\beta-1}{1-\beta}} \wedge 1\right)\right),$$

and the one for  $n = i_t$  can be included in this one. Moreover, note that the derivative of  $x \mapsto \exp(\log^\beta x)$  is smaller than 1 for  $x \geq 1$  so that from the definition of  $\ell_t$  and  $i_t$  we have  $|\exp(\log^\beta(i_t)) - \ell_t| \leq 1$ . Hence, we get

$$(5.20) \quad a_{i_t} - a_{i_{t-1}} = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) - \ell_t^{1-\alpha} \cdot J(\ell_t) + o\left(\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot \left(t^{2 \cdot \frac{2\beta-1}{1-\beta}} \wedge 1\right)\right)$$

We now distinguish the three cases  $\beta \in (0, \frac{1}{2})$ ,  $\beta = \frac{1}{2}$  and  $\beta \in (\frac{1}{2}, 1)$ .

*Case  $\beta \in (0, \frac{1}{2})$ .* In this case, we first use Lemma B.3 and the assumption that  $J$  satisfies (SV<sub>2</sub>) to deduce that there exists a function  $L: [1, \infty) \rightarrow (0, \infty)$  satisfying (SV<sub>1</sub>) such that  $L(x) \sim (\alpha - 1)J(x)$  and, for  $x$  large enough,  $x^{1-\alpha}J(x) = \int_x^\infty y^{-\alpha}L(y)dy$ , so that we can rewrite

$$\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) - \ell_t^{1-\alpha} \cdot J(\ell_t) = \int_{\ell_{t-1}}^{\ell_t} y^{-\alpha}L(y)dy.$$

Now, since  $\frac{2\beta-1}{1-\beta} < 0$  so it follows from Lemma 5.41 that  $\frac{\ell_t}{\ell_{t-1}} \rightarrow 1$  and we can apply Lemma B.2 and plug the result back into (5.20) to get

$$\begin{aligned} a_{i_t} - a_{i_{t-1}} &= \ell_t^{-\alpha}(\ell_t - \ell_{t-1})L(\ell_t) \left(1 + \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t}(1 + o(1))\right) + o\left(\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot t^{2 \cdot \frac{2\beta-1}{1-\beta}}\right) \\ &= \ell_t^{1-\alpha} \cdot L(\ell_t) \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t} \cdot \left(1 + \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)\right), \end{aligned}$$

where we used that  $\frac{\ell_t - \ell_{t-1}}{\ell_t} \sim \frac{\kappa\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}}$  by Lemma 5.42, and the fact that  $L(x) \sim (\alpha - 1)J(x)$  to factor in the error term. This ensures that

$$\begin{aligned} \theta_t - \theta_{t-1} &= (\alpha - 1)(\log \ell_t - \log \ell_{t-1}) - (\log L(\ell_t) - \log L(\ell_{t-1})) - \log\left(\frac{\ell_t - \ell_{t-1}}{\ell_t} \cdot \frac{\ell_{t-1}}{\ell_{t-1} - \ell_{t-2}}\right) \\ &\quad - \left(\log\left(1 + \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)\right) - \log\left(1 + \frac{\alpha}{2} \cdot \frac{\ell_{t-1} - \ell_{t-2}}{\ell_{t-1}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)\right)\right). \end{aligned}$$

The first term is directly handled by Lemma 5.41. For the second term, by Lemma B.12, we have

$$\log(L(\ell_t)) - \log(L(\ell_{t-1})) = o(\log(\ell_t) - \log(\ell_{t-1})) = o\left(t^{\frac{2\beta-1}{1-\beta}}\right),$$

using again Lemma 5.41 in the last equality. For the third term, Lemma 5.42 yields

$$\log\left(\frac{\ell_t - \ell_{t-1}}{\ell_t} \cdot \frac{\ell_{t-1}}{\ell_{t-1} - \ell_{t-2}}\right) = \log\left(\left(\frac{t}{t-1}\right)^{\frac{2\beta-1}{1-\beta}} \cdot \frac{1 - \frac{\kappa\beta}{2(1-\beta)}t^{\frac{2\beta-1}{1-\beta}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)}{1 - \frac{\kappa\beta}{2(1-\beta)}(t-1)^{\frac{2\beta-1}{1-\beta}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right)}\right) = o\left(t^{\frac{2\beta-1}{1-\beta}}\right),$$

noting that  $-1 < \frac{2\beta-1}{1-\beta} < 0$ . In order to control the last term, we use the expansion  $\log(1+x) = x + O(x^2)$  and the fact that the expansion in Lemma 5.42 gives

$$\frac{\ell_t - \ell_{t-1}}{\ell_t} - \frac{\ell_{t-1} - \ell_{t-2}}{\ell_{t-1}} = o\left(t^{\frac{2\beta-1}{1-\beta}}\right) \quad \text{and} \quad \left(\frac{\ell_t - \ell_{t-1}}{\ell_t}\right)^2 = o\left(t^{\frac{2\beta-1}{1-\beta}}\right).$$

Putting everything together we get

$$\theta_t - \theta_{t-1} = (\alpha - 1) \frac{\kappa\beta}{1-\beta} \cdot t^{\frac{2\beta-1}{1-\beta}} + o\left(t^{\frac{2\beta-1}{1-\beta}}\right),$$

which is eventually positive when  $t$  is large, so the sequence  $(\theta_t)_{t \geq s_0}$  is eventually non-decreasing as claimed. The claim on the asymptotic expansion of  $\theta_t$  might be obtained for example by summing the expansion that we have for the increments.

Case  $\beta = \frac{1}{2}$ . Recalling  $\ell_t = e^\kappa \ell_{t-1}$ , we write

$$\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) - \ell_t^{1-\alpha} \cdot J(\ell_t) = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot \left(1 - e^{(1-\alpha)\kappa} \frac{J(\ell_t)}{J(\ell_{t-1})}\right).$$

Since  $J$  satisfies (SV<sub>0</sub>) and  $\ell_t = e^\kappa \ell_{t-1}$  and  $\ell_t \rightarrow \infty$  as  $t \rightarrow \infty$  we have  $J(\ell_t)/J(\ell_{t-1}) = 1 + o(1)$ , so coming back to (5.20) we get

$$a_{i_t} - a_{i_{t-1}} = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot \left(1 - e^{(1-\alpha)\kappa}\right) \cdot (1 + o(1)).$$

Hence, it follows that

$$\theta_t = (\alpha - 1) \log \ell_{t-1} - \log J(\ell_{t-1}) - \log \left(1 - e^{(1-\alpha)\kappa}\right) + o(1).$$

This proves the asymptotics of  $\theta_t$  in (5.19). It remains to prove that  $(\theta_t)_{t \geq 1}$  is eventually non-decreasing: using again  $\log J(\ell_t) - \log J(\ell_{t-1}) = o(1)$ , we have  $\theta_{t+1} - \theta_t = (\alpha - 1)\kappa + o(1)$ , which is positive for  $t$  large enough.

Case  $\beta \in (\frac{1}{2}, 1)$ . In this case we can write

$$\ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) - \ell_t^{1-\alpha} \cdot J(\ell_t) = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot (1 + o(1)),$$

using that  $\ell_{t-1} = o(\ell_t)$  and  $\log J(x) = o(\log x)$  by Lemma B.11. This yields, plugging this back into (5.20),

$$\theta_t = -\log(a_{i_t} - a_{i_{t-1}}) = (\alpha - 1) \log \ell_{t-1} - \log J(\ell_{t-1}) + o(1).$$

This proves the asymptotics (5.19). The fact that  $(\theta_t)_{t \geq 1}$  is eventually non-decreasing follows again from  $\log J(\ell_t) - \log J(\ell_{t-1}) = o(\log \ell_t - \log \ell_{t-1})$  by Lemma B.11, together with the fact that  $\log \ell_t - \log \ell_{t-1} \rightarrow \infty$  by Lemma 5.41.  $\square$

**Lemma 5.6.** *Let  $\kappa > 0$  be such that  $\eta := \kappa^{\frac{1}{\beta}} - (\alpha - 1)(1 - \beta)\kappa$  is non-zero. Let  $i_t$  be defined as in (5.17) and  $\theta_t$  as in (5.18). Then, as  $t \rightarrow \infty$ ,*

$$(5.21) \quad i_t \cdot \exp\left(\sum_{i=2}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r\right) = \exp\left(\eta \cdot t^{\frac{1}{1-\beta}} \cdot (1 + o(1))\right).$$

Moreover, as  $s \rightarrow \infty$ , uniformly in  $t \geq s$ ,

$$(5.22) \quad \frac{i_t}{i_s} \cdot \exp\left(\sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} - \sum_{r=s+1}^t \theta_r\right) = \exp\left(\eta \cdot \left(t^{\frac{1}{1-\beta}} - s^{\frac{1}{1-\beta}}\right) \cdot (1 + o(1))\right).$$

**Proof.** We focus on the proof of (5.22), because (5.21) follows along the same lines. The case  $s = t$  is obvious so we assume  $t \geq s + 1$ . The following asymptotics are meant to hold as  $s \rightarrow \infty$ , uniformly in  $t \geq s + 1$ . Summing the asymptotic equivalent in (5.19), we get

$$\sum_{r=s+1}^t \theta_r \sim (\alpha - 1)\kappa \cdot \int_s^t x^{\frac{\beta}{1-\beta}} dx \sim (\alpha - 1)(1 - \beta)\kappa \cdot \left(t^{\frac{1}{1-\beta}} - s^{\frac{1}{1-\beta}}\right).$$



On the other hand, it follows from  $e^{\theta_r}(a_{i_r} - a_{i_{r-1}}) = 1$  that  $\sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} \sim t - s = o(t^{\frac{1}{1-\beta}} - s^{\frac{1}{1-\beta}})$ . Then, using the definition of  $i_t$ , we get (5.22).  $\square$

**Proof of Theorem 1.41.** Let  $i_t$  and  $\theta_t$  defined as in (5.17) and (5.18) for some fixed  $\kappa > 0$ . Recall that the sequence of integers  $(t_n)_{n \geq 1}$  is defined at the beginning of Section 3.1. It follows from the definition of  $i_t$  that, as  $n \rightarrow \infty$ ,

$$t_n = \kappa^{-\frac{1-\beta}{\beta}} (\log n)^{1-\beta}.$$

Moreover, we can check that the quantity  $\eta = \kappa^{\frac{1}{\beta}} - (\alpha - 1)(1 - \beta)\kappa$  appearing in Lemma 5.6 is of the same sign as  $\kappa - (\alpha - 1)^{\frac{\beta}{1-\beta}} (1 - \beta)^{\frac{\beta}{1-\beta}}$ .

*Upper bound.* We choose  $\kappa < (\alpha - 1)^{\frac{\beta}{1-\beta}} (1 - \beta)^{\frac{\beta}{1-\beta}}$  so that  $\eta < 0$  and therefore (3.6) follows from (5.21). Hence, we can apply Proposition 3.2 to get  $\text{ht}(\mathbb{T}_n) \leq t_n + O(1)$  almost surely. Letting  $\kappa \rightarrow (\alpha - 1)^{\frac{\beta}{1-\beta}} (1 - \beta)^{\frac{\beta}{1-\beta}}$  yields the upper bound.

*Lower bound.* We fix some  $\kappa > (\alpha - 1)^{\frac{\beta}{1-\beta}} (1 - \beta)^{\frac{\beta}{1-\beta}}$ , so that  $\eta > 0$ , and we aim at checking that (4.3) is satisfied to apply Proposition 4.3 with  $T(n) := (\log n)^{(1-\beta)^{3/2}}$ . Using assumption (1.5), we have

$$\delta_1(n) = \sum_{r \geq T(n)+1} e^{2\theta_r} \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} \leq \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot O\left(\frac{1}{i_{r-1}}\right) = \sum_{r \geq T(n)+1} \exp\left(-\kappa^{\frac{1}{\beta}} r^{\frac{1}{1-\beta}} (1 + o(1))\right),$$

using (5.17) and (5.19). It follows that

$$\sum_{t \geq 1} \delta_1(i_t) = \sum_{t \geq 1} O\left(\exp\left(-\kappa^{\frac{1}{\beta}} T(i_t)^{\frac{1}{1-\beta}}\right)\right) = \sum_{t \geq 1} O\left(\exp\left(-\kappa^{(1+(1-\beta)^{1/2})/\beta} \cdot t^{(1-\beta)^{-1/2}}\right)\right) < \infty,$$

proving the first part of (4.3). We now aim at checking the second part of (4.3). It follows from the first part of Lemma 4.5 that  $E(i_t) \geq \exp(-Ct)$ . For  $\delta_2(n)$  we proceed as in (5.15): using (1.5) and (5.22), we have

$$\begin{aligned} \delta_2(n) &\leq \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot \frac{i_{r-1}}{i_{T(n)}} \exp\left(-\eta \cdot \left(r^{\frac{1}{1-\beta}} - (T(n) + 1)^{\frac{1}{1-\beta}}\right) \cdot (1 + o(1))\right) \cdot O\left(\frac{1}{i_{r-1}}\right) \\ &\leq \frac{O(1)}{i_{T(n)}} \sum_{s \geq 0} \exp\left(4(\alpha - 1)\kappa \cdot (s + T(n) + 1)^{\frac{\beta}{1-\beta}} - \frac{\eta}{2} \cdot \left((s + T(n) + 1)^{\frac{1}{1-\beta}} - (T(n) + 1)^{\frac{1}{1-\beta}}\right)\right), \end{aligned}$$

for  $n$  large enough, using (5.19) to bound  $\theta_r$ . Then, for any real  $a, b > 0$  and integer  $M \geq 1$ , we have

$$\begin{aligned} &\sum_{s \geq 0} \exp\left(a(s + M)^{\frac{\beta}{1-\beta}} - b \cdot \left((s + M)^{\frac{1}{1-\beta}} - M^{\frac{1}{1-\beta}}\right)\right) \\ &\leq \sum_{s=0}^{M-1} \exp\left(a(2M)^{\frac{\beta}{1-\beta}}\right) + \sum_{s \geq M} \exp\left(a(2s)^{\frac{\beta}{1-\beta}} - b \cdot s^{\frac{1}{1-\beta}}\right) \leq M \exp\left(a(2M)^{\frac{\beta}{1-\beta}}\right) + C(a, b), \end{aligned}$$

bounding the second sum by the same sum starting at  $s = 0$ . It follows that

$$\delta_2(n) \leq \frac{1}{i_{T(n)}} \exp\left(O\left(T(n)^{\frac{\beta}{1-\beta}}\right)\right) = \exp\left(O\left(T(n)^{\frac{\beta}{1-\beta}}\right) - \kappa^{\frac{1}{\beta}} T(n)^{\frac{1}{1-\beta}}\right).$$

Recalling that  $E(i_t) \geq \exp(-Ct)$  and  $T(i_t) = (\log i_t)^{(1-\beta)^{3/2}} \sim \kappa^{(1-\beta)^{3/2}/\beta} \cdot t^{(1-\beta)^{1/2}}$ , the second part of (4.3) follows. Hence, we can apply Proposition 4.3 to get that almost surely, for  $n$  large enough,  $\text{ht}(\mathbb{T}_n) \geq t_n - T(n) - 1$ . Letting  $\kappa \rightarrow (\alpha - 1)^{\frac{\beta}{1-\beta}} (1 - \beta)^{\frac{\beta}{1-\beta}}$  yields the lower bound.  $\square$

5.3.2. Case  $\beta = 1$ 

Let  $\kappa > 1$ . For  $t \geq 1$ , we define

$$(5.23) \quad \ell_t := \exp(\kappa^t)$$

and

$$(5.24) \quad i_t := \begin{cases} t+1, & \text{if } 0 \leq t < s_0, \\ \lfloor \ell_t \rfloor = \lfloor \exp(\kappa^t) \rfloor, & \text{if } t \geq s_0, \end{cases}$$

by choosing some  $s_0$  large enough so that  $(i_t)_{t \geq 1}$  is increasing.

**Lemma 5.7.** *Let  $\kappa > 1$ . Let  $i_t$  be defined as in (5.24) and*

$$(5.25) \quad \theta_t := \begin{cases} 0, & \text{if } 1 \leq t \leq s_0, \\ -\log(a_{i_t} - a_{i_{t-1}}), & \text{if } t > s_0. \end{cases}$$

*Then, there exists  $r_0 \geq s_0$  such that  $(\theta_t)_{t \geq r_0}$  is a non-decreasing sequence of non-negative numbers. Moreover, as  $t \rightarrow \infty$ ,*

$$(5.26) \quad \theta_t \sim (\alpha - 1)\kappa^{t-1}.$$

**Proof.** The proof is similar to the case  $\beta \in (\frac{1}{2}, 1)$ . We start by writing

$$\ell_t^{1-\alpha} \cdot J(\ell_{t-1}) - \ell_t^{1-\alpha} \cdot J(\ell_t) = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot (1 + o(1)),$$

using that  $\log J(x) = o(\log x)$  by Lemma B.11. This allows us to write, using assumption (1.6),  $a_{i_t} - a_{i_{t-1}} = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot (1 + o(1))$ . This entails that

$$\theta_t = -\log(a_{i_t} - a_{i_{t-1}}) = (\alpha - 1)\kappa^{t-1} - \log J(\exp(\kappa^{t-1})) + o(1) \sim (\alpha - 1)\kappa^{t-1}.$$

This asymptotic expansion already ensures that the sequence  $\theta_t$  is eventually non-decreasing so the lemma is proved.  $\square$

**Lemma 5.8.** *Let  $\kappa > 1$  such that  $\kappa \neq \alpha$ . Let  $i_t$  be defined as in (5.24) and  $\theta_t$  as in (5.25). Then, as  $t \rightarrow \infty$ ,*

$$(5.27) \quad i_t \cdot \exp\left(\sum_{i=2}^{i_t} (e^{\theta_{i_t}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r\right) = \exp\left(\frac{\kappa - \alpha}{\kappa - 1} \cdot \kappa^t \cdot (1 + o(1))\right).$$

*Moreover, as  $s \rightarrow \infty$ , uniformly in  $t \geq s$ ,*

$$(5.28) \quad \exp\left(\sum_{r=s+1}^t \theta_r - \sum_{i=i_s+1}^{i_t} (e^{\theta_{i_t}} - 1) \frac{w_i}{W_i}\right) = \exp\left(\frac{\alpha - 1}{\kappa - 1} \cdot (\kappa^t - \kappa^s) \cdot (1 + o(1))\right).$$

**Proof.** This follows from the fact that

$$\sum_{r=1}^t \theta_r = \sum_{r=1}^t (\alpha - 1)\kappa^{r-1}(1 + o(1)) = \frac{\alpha - 1}{\kappa - 1} \cdot \kappa^t \cdot (1 + o(1))$$

and similarly for  $t \geq s + 1$ , (the case  $t = s$  can be treated aside),

$$\sum_{r=s+1}^t \theta_r = \sum_{r=s+1}^t (\alpha - 1)\kappa^{r-1}(1 + o(1)) = \frac{\alpha - 1}{\kappa - 1} \cdot (\kappa^t - \kappa^s) \cdot (1 + o(1)),$$

the definition of  $i_t$  in (5.24), and the fact that  $\sum_{i=i_s+1}^{i_t} (e^{\theta_{i_t}} - 1) \frac{w_i}{W_i} = t - s + o(t - s) = o(\kappa^t - \kappa^s)$ .  $\square$

**Proof of Theorem 1.42.** Let  $i_t$  and  $\theta_t$  defined as in (5.24) and (5.25) for some fixed  $\kappa > 1$ . It follows from the definition of  $i_t$  that

$$t_n = \frac{\log \log n}{\log \kappa} + O(1).$$

*Upper bound.* We choose  $\kappa < \alpha$  and then (3.6) follows from (5.27). Hence, we can apply Proposition 3.2 to get  $\text{ht}(T_n) \leq t_n + O(1)$  almost surely. Letting  $\kappa \rightarrow \alpha$  yields the upper bound.

*Lower bound.* We fix some  $\kappa > \alpha$  and we aim at checking that (4.3) is satisfied to apply Proposition 4.3 with  $T(n) := (\log \log n)^{1/2}$ . Using (1.7), for any  $\varepsilon > 0$ , we have

$$\delta_1(n) = \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot O\left(\frac{1}{i_{r-1}^{2\alpha-1-\varepsilon}}\right) = \sum_{r \geq T(n)+1} O\left(\exp(-(1-\varepsilon)\kappa^{r-1}(1+o(1)))\right),$$

using (5.26). Choosing  $\varepsilon = 1/4$ , this yields  $\delta_1(i_t) = O(\exp(-\kappa^{T(i_t)}/2))$ , which is summable in  $t$  since  $T(i_t) \sim \sqrt{t \log \kappa}$ . This proves the first part of (4.3). For the second part, we first have  $E(i_t) \geq \exp(-Ct)$  by the first part of Lemma 4.5. Then, using (1.7) as for  $\delta_1(n)$  and (5.28), we have, for any  $\varepsilon > 0$ ,

$$\begin{aligned} \delta_2(n) &= \sum_{r \geq T(n)+1} \exp\left(\sum_{s=T(n)+1}^{r-1} \theta_s - \sum_{i=i_{T(n)+1}}^{i_{r-1}} (e^{\theta_{i_t}} - 1) \frac{w_i}{W_i}\right) \cdot e^{2\theta_r} \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} \\ &= \sum_{r \geq T(n)+1} \exp\left(\frac{\alpha-1}{\kappa-1} \cdot (\kappa^{r-1} - \kappa^{T(n)})(1+o(1)) - (1-\varepsilon)\kappa^{r-1}(1+o(1))\right) \\ &\leq \sum_{r \geq T(n)+1} \exp\left(-\left(\frac{\kappa-\alpha}{\kappa-1} - \varepsilon\right) \cdot \kappa^{r-1}(1+o(1))\right). \end{aligned}$$

Hence, choosing  $\varepsilon = \frac{1}{3} \cdot \frac{\kappa-\alpha}{\kappa-1}$ , we get  $\delta_2(n) = O(\exp(-\varepsilon \kappa^{T(n)}))$ . It follows easily that  $\sum_{t \geq 1} \delta_2(i_t)/E(i_t) < \infty$ , which proves the second part of (4.3). Applying Proposition 4.3 and letting  $\kappa \rightarrow \alpha$ , we get the lower bound.  $\square$

### 5.3.3. Case $\beta > 1$

Let  $\kappa > 1$ . For  $t \geq 1$ , we define

$$(5.29) \quad \ell_t := \exp(\exp(\beta \kappa^t))$$

and

$$(5.30) \quad i_t := \begin{cases} t+1, & \text{if } 0 \leq t < s_0, \\ \lfloor \exp(\log^{1/\beta} \ell_t) \rfloor = \lfloor \exp(\exp(\kappa^t)) \rfloor, & \text{if } t \geq s_0, \end{cases}$$

by choosing some  $s_0$  large enough so that  $(i_t)_{t \geq 1}$  is increasing.

**Lemma 5.9.** *Let  $\kappa > 1$ . Let  $i_t$  be defined as in (5.30) and*

$$(5.31) \quad \theta_t := \begin{cases} 0, & \text{if } 1 \leq t \leq s_0, \\ -\log(a_{i_t} - a_{i_{t-1}}), & \text{if } t > s_0. \end{cases}$$

*Then, there exists  $r_0 \geq s_0$  such that  $(\theta_t)_{t \geq r_0}$  is a non-decreasing sequence of non-negative numbers. Moreover, as  $t \rightarrow \infty$ ,*

$$(5.32) \quad \theta_t = (\alpha - 1) \log \ell_{t-1} - \log J(\ell_{t-1}) + O(1).$$

**Proof.** Proceeding as in the proof of Lemma 5.7, we get  $a_{i_t} - a_{i_{t-1}} = \ell_{t-1}^{1-\alpha} \cdot J(\ell_{t-1}) \cdot (1+o(1))$ . The result follows by taking the logarithm. In particular,  $\theta_t \sim (\alpha - 1) \exp(\beta \kappa^{t-1})$ , so it is eventually non-negative and non-decreasing.  $\square$

**Lemma 5.10.** *Let  $\kappa > 1$ . Let  $i_t$  be defined as in (5.30) and  $\theta_t$  as in (5.31). Then, as  $t \rightarrow \infty$ ,*

$$(5.33) \quad i_t \cdot \exp\left(\sum_{i=2}^{i_t} (e^{\theta_{i_t}} - 1) \frac{w_i}{W_i} - \sum_{r=1}^t \theta_r\right) = \exp\left(e^{\kappa^t} - (\alpha - 1)e^{\beta \kappa^{t-1}}(1+o(1))\right).$$

Moreover, as  $s \rightarrow \infty$ , uniformly in  $t \geq s$ ,

$$(5.34) \quad \exp\left(\sum_{r=s+1}^t \theta_r - \sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i}\right) = \exp\left((\alpha - 1)e^{\beta\kappa^{t-1}}(1 + o(1))\mathbb{1}_{t \geq s+1}\right).$$

**Proof.** Recall that, as for the other cases, we have  $\sum_{i=i_s+1}^{i_t} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i} = O(t)$  for any  $s \leq t$ . Then, the two claims are obtained by summing the asymptotic equivalent  $\theta_r \sim (\alpha - 1) \exp(\beta\kappa^{r-1})$ .  $\square$

**Proof of Theorem 1.43.** Let  $i_t$  and  $\theta_t$  defined as in (5.30) and (5.31) for some fixed  $\kappa > 1$ . It follows from the definition of  $i_t$  that

$$t_n = \frac{\log \log \log n}{\log \kappa} + O(1).$$

*Upper bound:* We choose  $\kappa < \beta$  so that  $e^{\kappa^t} = o((\alpha - 1)e^{\beta\kappa^{t-1}})$ . Therefore, (3.6) follows from (5.34) with  $s = 1$ . Applying Proposition 3.2 and then letting  $\kappa \rightarrow \beta$  yields the upper bound.

*Lower bound:* We fix some  $\kappa > \beta$  and we aim at checking that (4.3) is satisfied to apply Proposition 4.3 with  $T(n) := \log \log \log \log n$ . Using assumption (1.8), there is  $\varepsilon > 0$  such that, as  $r \rightarrow \infty$ ,

$$(5.35) \quad \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} = O\left(\frac{1}{i_{r-1}^\varepsilon} \ell_{r-1}^{-2(\alpha-1)} J^2(\ell_{r-1})\right) = O\left(\frac{1}{i_{r-1}^\varepsilon} e^{-2\theta_r}\right),$$

using (5.32). Applying this, we get

$$\delta_1(n) = \sum_{r \geq T(n)+1} O\left(\frac{1}{i_{r-1}^\varepsilon}\right) = O\left(\sum_{r \geq T(n)+1} \exp(-\varepsilon e^{\kappa^r})\right) = O\left(\exp(-\varepsilon e^{\kappa^{T(n)}})\right).$$

Since  $T(i_t) \sim \log t$ , it follows that  $\sum_{t \geq 1} \delta_1(i_t) < \infty$ , which proves the first part of (4.3). We now aim at checking the second part of (4.3). It follows from the first part of Lemma 4.5 that  $E(i_t) \geq \exp(-Ct)$ . For  $\delta_2(n)$ , using (5.34) and (5.35), we get

$$\begin{aligned} \delta_2(n) &= \sum_{r \geq T(n)+1} e^{2\theta_r} \cdot \exp\left(\sum_{s=T(n)+1}^{r-1} \theta_s - \sum_{i=i_{T(n)}+1}^{i_{r-1}} (e^{\theta_{t_i}} - 1) \frac{w_i}{W_i}\right) \cdot \sum_{i=i_{r-1}+1}^{i_r} \frac{w_i^2}{W_i^2} \\ &= \sum_{r \geq T(n)+1} \exp\left((\alpha - 1)e^{\beta\kappa^{r-2}}(1 + o(1))\mathbb{1}_{t \geq s+1}\right) \cdot O\left(\frac{1}{i_{r-1}^\varepsilon}\right) \\ &= \sum_{r \geq T(n)+1} \exp\left(-\varepsilon e^{\kappa^{r-1}}(1 + o(1))\right), \end{aligned}$$

using that  $e^{\beta\kappa^{r-2}} = o(e^{\kappa^{r-1}})$ , since  $\kappa > \beta$ . It is then easy to check that  $\sum_{t \geq 1} \delta_2(i_t)/E(i_t) < \infty$ , proving the second part of (4.3). Applying Proposition 4.3 and letting  $\kappa \rightarrow \beta$ , we get the lower bound.  $\square$

## Appendix A: Crude bounds for the height

### A.1. Crude upper bound

In this section, we use the many-to-one lemma in a naive way (without a barrier) in order to get some crude upper bound on the height of  $T_n$ . As explained below, this amounts to bounding  $\text{ht}(T_n)$  by the maximum of  $n$  independent copies of  $\sum_{i=1}^n B_i$ , where the  $B_i$ 's are independent Bernoulli( $w_i/W_i$ ) random variables. We then compare the obtained upper bound with the actual behavior of  $\text{ht}(T_n)$  in the regimes that are studied in Theorem 1.1 and Theorem 1.4.

*Crude upper bound approach.* We first work under no restriction on the weight sequence  $(w_i)$ . Applying the result contained in Remark 2.2 to the function  $f(y) = \mathbb{1}_{\{y \geq x\}}$ , we get

$$\mathbb{P}(\text{ht}(\mathbf{u}_n) \geq x) = \mathbb{P}(H_{n-1} + 1 \geq x),$$

where we recall that we defined for  $k \geq 1$ ,  $H_k = \sum_{i=2}^k B_i$  with  $(B_i)_{i \geq 1}$  a sequence of independent random variables such that  $B_i$  has distribution Bernoulli( $w_i/W_i$ ). Using a union-bound on all the vertices yields

$$(A.1) \quad \mathbb{P}(\text{ht}(\mathbf{T}_n) \geq x) \leq \sum_{i=1}^n \mathbb{P}(\text{ht}(\mathbf{u}_i) \geq x) = \sum_{i=1}^n \mathbb{P}(H_{i-1} + 1 \geq x) \leq n \cdot \mathbb{P}(H_n + 1 \geq x),$$

where for the last inequality we used that  $(H_k)_{k \geq 1}$  is non-decreasing. Note that the right-hand side of (A.1) is small if and only if  $x$  is larger than the maximum of  $n$  independent copies of  $H_n + 1 = \sum_{i=1}^n B_i$ . Even if the inequalities in (A.1) are not optimal, there are not so far from the truth in the regimes considered in this paper (see Footnote 1).

We now have to estimate the maximum of  $n$  independent copies of  $\sum_{i=1}^{n-1} B_i$ , that we denote by  $M_n$ , in the different regimes considered in Theorem 1.1 and Theorem 1.4. This is done via standard techniques and we omit the calculations, even if they can be tedious in some of the cases.

*Variance varying like a power of  $\log n$ .* We consider here the framework of Theorem 1.1. First assume that  $\alpha \in (0, 1)$ . In that case one can check that

$$M_n = \frac{\log n}{\alpha \log \log n} + \frac{\log n}{(\alpha \log \log n)^2} \left( \sum_{k \geq 0} \left( \frac{\log L(\log n)}{\alpha \log \log n} \right)^k (\log L(\log n) + (k+1) \log \log \log n) + 1 + \log \alpha - \log(1 - \alpha) + o(1) \right).$$

The single difference with the expansion of  $\text{ht}(\mathbf{T}_n)$  in Theorem 1.1 appears in the coefficient of the term of order  $\log n / (\log \log n)^2$ , which is  $1 + \log \alpha - \log(1 - \alpha)$  instead of  $1 + \log \alpha - \alpha$ .

Now assume that  $\alpha = 1$  and  $a_n$  diverges. Assume also that  $a_{\log n} = o(a_n)$  as  $n \rightarrow \infty$ <sup>5</sup>. Let  $J(x) := \int_1^x \frac{L(u)}{u} du$ , so that  $a_n \sim J(\log n)$ . One can check that  $L(x) = o(J(x))$  as  $x \rightarrow \infty$  and it follows that  $J$  satisfies (SV<sub>1</sub>). Then, the following expansion holds.

$$M_n = \frac{\log n}{\log \log n} + \frac{\log n}{(\log \log n)^2} \left( \sum_{k \geq 0} \left( \frac{\log J(\log n)}{\log \log n} \right)^k (\log J(\log n) + (k+1) \log \log \log n) + 1 + o(1) \right).$$

Therefore, the first order is the same for  $M_n$  and  $\text{ht}(\mathbf{T}_n)$ , but the order of magnitude of the difference is greater than  $\log n / (\log \log n)^2$  because  $\log J(x) - \log L(x) \rightarrow \infty$ .

Assume again  $\alpha = 1$ , but now with  $a_n$  converging to some finite limit  $a_\infty$ . Assume also that  $a_\infty - a_n = o(a_\infty - a_{\log n})$  (the same comment as in Footnote 5 holds). Let  $I(x) := \int_x^\infty \frac{L(u)}{u} du$ , one can check as for  $J$  that  $L(x) = o(I(x))$  as  $x \rightarrow \infty$  and therefore  $I$  satisfies (SV<sub>1</sub>). In this case, we get

$$M_n = \frac{\log n}{\log \log n} + \frac{\log n}{(\log \log n)^2} \left( \log I(\log \log n) + \log \log \log n + 1 + o(1) \right).$$

Again, the difference with  $\text{ht}(\mathbf{T}_n)$  is not at the first order but is greater than  $\log n / (\log \log n)^2$  because  $\log I(\log x) \geq \log I(x)$  and  $\log I(x) - \log L(x) \rightarrow \infty$ .

Finally, in the case  $\alpha > 1$ , we can check that

$$(A.2) \quad M_n \sim \frac{\log n}{\log \log n},$$

which differs from the first order of  $\text{ht}(\mathbf{T}_n)$  by a multiplicative constant.

<sup>5</sup>We do not discuss the opposite case here, for which it is tedious to get an expansion for  $M_n$ . However, it is easier to check that the claim in Remark 1.3 still holds.

*Quickly converging variance.* We consider here the framework of Theorem 1.1. If  $\beta > 1$ , we get the same behavior as in (A.2). If  $\beta = 1$ , we have

$$M_n \sim \frac{\log n}{\alpha \log \log n},$$

where  $\alpha > 1$  is the constant appearing in assumption (1.6). If  $\beta > 1$ , we get

$$M_n \sim \frac{\log n}{(\alpha - 1)(\log \log n)^\beta}.$$

In all these cases, the upper bound is not even the right order of magnitude for  $\text{ht}(\mathbb{T}_n)$  since we have  $\text{ht}(\mathbb{T}_n) = o(M_n)$  almost surely.

## A.2. Crude lower bound

In this section, we present a simple method for proving lower bounds for the height of  $\mathbb{T}_n$ , which surprisingly gives a correct first order in the regimes studied in this paper. Assume for simplicity that the sequence  $(w_i)_{i \geq 1}$  is non-increasing and that all weights  $w_i$  are non-zero.

The strategy is the following. We build a path in the tree  $\mathbb{T}_n$  starting from the root, by then considering its first child and then the first child of the latter, and so on. More formally, we define a random sequence of indices  $(I_r)_{r \geq 1}$  recursively as follows: let  $I_0 = 1$  and, for any  $r \geq 0$ ,  $I_{r+1}$  is the index of the first vertex attached to  $u_{I_r}$ . Note that this sequence is a.s. well-defined: any vertex  $u_i$  has infinitely many children a.s. by Borel–Cantelli lemma and the fact that  $\sum_{k > i} w_i/W_k = \infty$  in the regime where  $W_k$  grows sub-polynomially. Finding a lower bound for the length of this path provides a lower bound for  $\text{ht}(\mathbb{T}_n)$ .

To this end, we note that, for any  $r \geq 1$  and  $t \geq 0$ ,

$$\mathbb{P}(I_{r+1} - I_r > t \mid I_r) = \prod_{i=I_r+1}^{I_r+t} \left(1 - \frac{w_{I_r}}{W_i}\right) \leq \exp\left(-\sum_{i=I_r+1}^{I_r+t} \frac{w_{I_r}}{W_i}\right).$$

Then, for any increasing sequence  $(i_r)_{r \geq 0}$  such that  $i_0 = 1$ , we have, noting that  $I_0 = 1 = i_0$ ,

$$\begin{aligned} \mathbb{P}(\exists r \geq 0 : I_r > i_r) &\leq \sum_{r \geq 0} \mathbb{P}(I_{r+1} > i_{r+1}, I_r \leq i_r) = \sum_{r \geq 0} \mathbb{E}[\mathbb{P}(I_{r+1} - I_r > i_{r+1} - I_r \mid I_r) \mathbb{1}_{I_r \leq i_r}] \\ (A.3) \quad &\leq \sum_{r \geq 0} \mathbb{E}\left[\exp\left(-\sum_{i=I_r+1}^{i_{r+1}} \frac{w_{I_r}}{W_i}\right) \mathbb{1}_{I_r \leq i_r}\right] \leq \sum_{r \geq 0} \exp\left(-\sum_{i=i_r+1}^{i_{r+1}} \frac{w_{i_r}}{W_i}\right), \end{aligned}$$

using that  $(w_i)_{i \geq 1}$  is non-increasing. Hence, if we find a sequence  $(i_r)_{r \geq 0}$  such that the right-hand side of (A.3) is smaller than  $\varepsilon$ , then we can deduce that  $\mathbb{P}(\text{ht}(\mathbb{T}_n) \geq t_n - 1) \geq \mathbb{P}(I_{t_n-1} \leq n) \geq 1 - \varepsilon$ , where  $(t_n)_{n \geq 1}$  is defined in such a way that  $i_{t_n-1} < n \leq i_{t_n}$  for all  $n \geq 1$ .

We now illustrate this method on with example: assume  $w_i = i^{-\alpha}$  for some  $\alpha > 1$  (this corresponds to the case  $\beta = 1$  in Theorem 1.4). Then, using the bound  $W_i \geq W_1 = 1$ , we have

$$\sum_{i=i_r+1}^{i_{r+1}} \frac{w_{i_r}}{W_i} \geq i_r^{-\alpha} \cdot (i_{r+1} - i_r).$$

In order for this to be large, we want  $i_r^\alpha = o(i_{r+1})$ . This leads naturally to the choice  $i_r = \exp(\kappa^r)$  with  $\kappa > \alpha$  for which the series on the right-hand side of (A.3) is convergent. But, to get it as small as we want, we rather take  $i_r = \delta \cdot \exp(\kappa^r)$  for small enough  $\delta > 0$ . This proves a lower bound with  $t_n = (\log \log n) / \log \kappa + O(1)$ , which yields the lower bound in Theorem 1.4 by letting  $\kappa \rightarrow \alpha$  (at least in probability, but one can strengthen it to an almost sure result).

This method, despite being very crude, gives the correct first order for the lower bound in all cases of Theorem 1.1 and Theorem 1.4, at least in the case of non-increasing  $(w_i)_{i \geq 1}$ . One can weaken this assumption, but then it does not work to choose the first child of each vertex along the path, one has to choose the first child with sufficient weight. Pushing this method to get the next orders in the framework of Theorem 1.1, one sees that the lower bound obtained does not match the truth at the term of order  $\frac{\log n}{\log \log n} \log \log \log n$ . Moreover, in the case of polynomially-growing  $W_n$  treated in [20], this method would not give the correct first order.



## Appendix B: Slowly varying functions

In this short section, we state and prove a few results concerning slowly varying functions that we need in Section 5 to prove Theorem 1.1 and Theorem 1.4.

**Lemma B.1.** *The following holds.*

1. Assume  $L: [1, \infty) \rightarrow (0, \infty)$  satisfies (SV<sub>0</sub>). Then, as  $x \rightarrow \infty$ ,  $\log L(x) = o(\log x)$  and, for any  $\lambda > 1$ , uniformly in  $y > \lambda x$ ,

$$(B.1) \quad \log L(y) - \log L(x) = o\left(\log\left(\frac{y}{x}\right)\right).$$

2. Assume  $L: [1, \infty) \rightarrow (0, \infty)$  satisfies (SV<sub>1</sub>). Then, as  $x \rightarrow \infty$ , uniformly in  $y > x$ ,

$$(B.2) \quad \log L(y) - \log L(x) = o\left(\log\left(\frac{y}{x}\right)\right).$$

3. Assume  $L: [1, \infty) \rightarrow (0, \infty)$  satisfies (SV<sub>2</sub>). Let  $\tilde{L}(x) := \exp(xL'(x)/L(x))$ . Then,  $\tilde{L}$  satisfies (SV<sub>1</sub>) and, uniformly in  $y > x$ ,

$$(B.3) \quad \log L(y) - \log L(x) = \log \tilde{L}(y) \cdot \log\left(\frac{y}{x}\right) + o\left(\log^2\left(\frac{y}{x}\right)\right),$$

**Proof.** *Part 1.* This follows easily from Karamata's Representation Theorem [1, Theorem 1.3.1], which shows that, for some  $A$  large enough,  $L$  can be written as

$$L(x) = \exp\left(f(x) + \int_A^x \frac{g(u)}{u} du\right), \quad x \geq A,$$

where  $f$  and  $g$  are measurable functions such that  $f(x)$  has a finite limit as  $x \rightarrow \infty$  and  $g(x) \rightarrow 0$ .

*Part 2.* When  $L$  satisfies (SV<sub>1</sub>), the representation above holds with  $f = 0$  and  $g(x) = xL'(x)/L(x)$ . In particular, we can write, using  $xL'(x)/L(x) \rightarrow 0$ ,

$$(B.4) \quad \log L(y) - \log L(x) = \int_x^y \frac{1}{u} \cdot \frac{uL'(u)}{L(u)} du = o\left(\int_x^y \frac{1}{u} du\right) = o\left(\log\left(\frac{y}{x}\right)\right),$$

which proves (B.2).

*Part 3.* It follows from a derivative calculation that  $\tilde{L}$  satisfies (SV<sub>1</sub>). Moreover, we have

$$\begin{aligned} \log L(y) - \log L(x) &= \int_x^y \frac{1}{u} \log \tilde{L}(u) du = \log \tilde{L}(y) \cdot \log\left(\frac{y}{x}\right) + \int_x^y \frac{1}{u} \cdot (\log \tilde{L}(u) - \log \tilde{L}(y)) du \\ &= \log \tilde{L}(y) \cdot \log\left(\frac{y}{x}\right) + \int_x^y \frac{1}{u} \cdot o\left(\log\left(\frac{y}{u}\right)\right) du, \end{aligned}$$

using (B.2) but for  $\tilde{L}$ . This yields (B.3). □

**Lemma B.2.** *Let  $\alpha > 0$ . Assume  $L: [1, \infty) \rightarrow (0, \infty)$  satisfies (SV<sub>1</sub>). Let  $(\ell_t)_{t \geq 1}$  be an increasing sequence such that  $\ell_t \rightarrow \infty$  and  $\ell_{t-1}/\ell_t \rightarrow 1$ . Then, as  $t \rightarrow \infty$ ,*

$$(B.5) \quad \int_{\ell_{t-1}}^{\ell_t} x^{-\alpha} L(x) dx = \ell_t^{-\alpha} (\ell_t - \ell_{t-1}) L(\ell_t) \left(1 + \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t} (1 + o(1))\right).$$

**Proof.** We decompose

$$(B.6) \quad \int_{\ell_{t-1}}^{\ell_t} x^{-\alpha} L(x) dx = (\ell_t - \ell_{t-1}) \ell_t^{-\alpha} L(\ell_t) + \int_{\ell_{t-1}}^{\ell_t} (x^{-\alpha} - \ell_t^{-\alpha}) L(x) dx + \ell_t^{-\alpha} \int_{\ell_{t-1}}^{\ell_t} (L(x) - L(\ell_t)) dx.$$

We first deal with the third term on the right-hand side of (B.6). Using the same representation for  $L$  as in (B.4) and the fact that  $xL'(x)/L(x) \rightarrow 0$ , we have, uniformly in  $x \in [\ell_{t-1}, \ell_t]$ ,

$$(B.7) \quad \frac{L(x) - L(\ell_t)}{L(\ell_t)} = \exp\left(-\int_x^{\ell_t} \frac{1}{y} \cdot \frac{yL'(y)}{L(y)} dy\right) - 1 = \exp\left(o\left(\log \frac{\ell_t}{\ell_{t-1}}\right)\right) - 1 = o\left(\frac{\ell_t - \ell_{t-1}}{\ell_t}\right),$$

using that  $(\ell_t - \ell_{t-1})/\ell_t \rightarrow 0$  as a consequence of the assumptions of the lemma. Therefore, we get

$$\ell_t^{-\alpha} \int_{\ell_{t-1}}^{\ell_t} (L(x) - L(\ell_t)) dx = o\left(\ell_t^{-\alpha} \frac{\ell_t - \ell_{t-1}}{\ell_t} L(\ell_t) \int_{\ell_{t-1}}^{\ell_t} dx\right) = o\left(\ell_t^{-\alpha} \frac{(\ell_t - \ell_{t-1})^2}{\ell_t} L(\ell_t)\right),$$

proving that this term can be included in the  $o(1)$  term on the right-hand side of (B.5). We now deal with the second term on the right-hand side of (B.6). By (B.7), we have  $L(x) = L(\ell_t)(1 + o(1))$  uniformly in  $x \in [\ell_{t-1}, \ell_t]$ . Therefore, with the change of variable  $x = \ell_t - u$ , we have

$$\begin{aligned} \int_{\ell_{t-1}}^{\ell_t} (x^{-\alpha} - \ell_t^{-\alpha})L(x) dx &= \ell_t^{-\alpha} L(\ell_t)(1 + o(1)) \int_0^{\ell_t - \ell_{t-1}} \left(\left(1 - \frac{u}{\ell_t}\right)^{-\alpha} - 1\right) du \\ &\sim \ell_t^{-\alpha} L(\ell_t) \int_0^{\ell_t - \ell_{t-1}} \frac{\alpha u}{\ell_t} du \\ &\sim (\ell_t - \ell_{t-1}) \ell_t^{-\alpha} L(\ell_t) \cdot \frac{\alpha}{2} \cdot \frac{\ell_t - \ell_{t-1}}{\ell_t}. \end{aligned}$$

Coming back to (B.6), this concludes the proof.  $\square$

**Lemma B.3.** Assume  $J: [1, \infty) \rightarrow (0, \infty)$  satisfies (SV $_k$ ) for some  $k \geq 1$ .

1. For any  $\alpha \in (0, 1)$ , there exists a function  $L: [1, \infty) \rightarrow (0, \infty)$  satisfying (SV $_{k-1}$ ) such that  $L(x) \sim (1 - \alpha)J(x)$  and, for  $x$  large enough,

$$(B.8) \quad x^{1-\alpha} J(x) = \int_1^x y^{-\alpha} L(y) dy.$$

2. For any  $\alpha > 1$ , there exists a function  $L: [1, \infty) \rightarrow (0, \infty)$  satisfying (SV $_{k-1}$ ) such that  $L(x) \sim (\alpha - 1)J(x)$  and, for  $x$  large enough,

$$(B.9) \quad x^{1-\alpha} J(x) = \int_x^\infty y^{-\alpha} L(y) dy.$$

**Proof.** Let  $\alpha \in \mathbb{R}$ . Let  $L_0(x) := (1 - \alpha)J(x) + xJ'(x)$ , so that  $x^{-\alpha}L_0(x) = \frac{d}{dx}(x^{1-\alpha}J(x))$ . Then, if  $\alpha \neq 1$ , it follows from (SV $_1$ ) for  $J$  that  $L_0(x) \sim (1 - \alpha)J(x)$ . Using this fact and (SV $_k$ ) for  $J$ , we get, for any  $i \in \llbracket 1, k-1 \rrbracket$ ,

$$(B.10) \quad \frac{x^i L_0^{(i)}(x)}{L_0(x)} = (1 - \alpha) \frac{x^i J^{(i)}(x)}{L_0(x)} + \frac{x^{i+1} J^{(i+1)}(x)}{L_0(x)} \xrightarrow{x \rightarrow \infty} 0.$$

*Part 1.* Assume that  $\alpha \in (0, 1)$ . We write

$$x^{1-\alpha} J(x) = J(1) + \int_1^x y^{-\alpha} L_0(y) dy,$$

Since  $L_0(x) \sim (1 - \alpha)J(x)$ , there exists  $x_0$  such that for  $x \geq x_0$ ,  $L_0(x) > 0$ . We define  $L(x) = L_0(x)$  for  $x \geq x_0$ . This ensures that  $L$  satisfies (SV $_{k-1}$ ) by (B.10). For  $x < x_0$ , we can choose  $L(x)$  such that  $L$  is  $k-1$  times differentiable positive and

$$J(1) + \int_1^{x_0} y^{-\alpha} L_0(y) dy = \int_1^{x_0} y^{-\alpha} L(y) dy.$$

Then (B.8) holds for any  $x \geq x_0$ .

*Part 2.* Assume that  $\alpha > 1$ . Then, there exists  $x_0$  such that for  $x \geq x_0$ , we have  $L_0(x) < 0$ . We choose  $L(x) = -L_0(x)$  for  $x \geq x_0$ . Moreover, we have

$$x^{1-\alpha} J(x) = - \int_x^\infty y^{-\alpha} L_0(y) dy,$$

so (B.9) holds for any  $x \geq x_0$ .  $\square$

### Appendix C: Converting assumptions on $a_n$ to $W_n$

**Proof of (1.3).** We denote  $\delta(x) := \log(1-x) + x$  and then we write

$$\frac{W_1}{W_n} = \prod_{i=2}^n \frac{W_{i-1}}{W_i} = \exp\left(\sum_{i=2}^n \log\left(1 - \frac{w_i}{W_i}\right)\right) = \exp\left(-\sum_{i=2}^n \frac{w_i}{W_i} + \sum_{i=2}^n \delta\left(\frac{w_i}{W_i}\right)\right).$$

This yields

$$a_n = 1 + \sum_{i=2}^n \frac{w_i}{W_i} = 1 + \log\left(\frac{W_n}{W_1}\right) - \sum_{i=2}^\infty \delta\left(\frac{w_i}{W_i}\right) + \sum_{i=n+1}^\infty \delta\left(\frac{w_i}{W_i}\right) = \log(W_n) + K + O\left(\sum_{i=n+1}^\infty \frac{w_i^2}{W_i^2}\right),$$

as  $n \rightarrow \infty$ , since  $\delta(x) = O(x^2)$  as  $x \rightarrow 0$ .  $\square$

### Appendix D: Application to specific weight sequences

This section contains the computations needed to verify that we can indeed apply Theorem 1.1 and Theorem 1.4 to the examples that we presented in Section 1.2 and Section 1.3 in the introduction. Before diving into computations, let us note that if  $a_n$  satisfies

$$a_n = \int_1^{\log n} x^{-\alpha} L(x) dx + C + o((\log n)^{-1-\alpha} (\log \log n)^2 L(\log n)).$$

for some  $\alpha > 0$ , some constant  $C$ , and some function  $L$  that satisfies (SV<sub>2</sub>), we can change the values of  $L$  on a finite interval without changing its regularity in such a way that assumption (1.4) holds. Hence, in the examples, we only check without loss of generality that the above display holds for some constant  $C$ . To do that we additionally use (1.3), which allows us to only check the asymptotics of  $W_n$ . The main difficulty is that, in order to get a small enough error term, we sometimes need to include corrective terms in the definition of  $L$ .

**Case  $w_n = \frac{\lambda(1-\alpha)}{n(\log n)^\alpha} \cdot \exp(\lambda(\log n)^{1-\alpha})$  with  $\alpha \in (0, 1)$  and  $\lambda > 0$ .** First, using a sum-integral comparison we get that  $W_n = \exp(\lambda(\log n)^{1-\alpha}) + O(1)$ . It is easy to check from there that  $\sum_{i \geq n} w_i^2 / W_i^2 = O(n^{-1}(\log n)^{-2\alpha})$  so that (1.5) holds and also that

$$\log(W_n) = \lambda(\log n)^{1-\alpha} + O(\exp(-\lambda(\log n)^{1-\alpha})) = \int_1^{\log n} x^{-\alpha} \lambda(1-\alpha) dx + O((\log n)^{-1-\alpha}).$$

Using the last display and (1.3), we get that  $a_n$  satisfies (1.4) with a function  $L$  chosen such that  $L$  is twice differentiable and  $L(x) = \lambda(1-\alpha)$  for  $x$  sufficiently large. This ensures that we can apply Theorem 1.1.

**Case  $w_n = \frac{1}{n}$ .** We have  $W_n = \sum_{i=1}^n \frac{1}{i} = \log n + \gamma + O\left(\frac{1}{n}\right)$ , where  $\gamma$  is the Euler–Mascheroni constant. We can easily check that  $\sum_{i \geq n} w_i^2 / W_i^2 = O(n^{-1}(\log n)^{-2})$ , which ensures that (1.5) holds. We also get

$$\begin{aligned} \log W_n &= \log\left(\log n + \gamma + O\left(\frac{1}{n}\right)\right) = \log \log n + \frac{\gamma}{\log n} + O\left(\frac{1}{(\log n)^2}\right) \\ &= \int_1^{\log n} \frac{1}{x} \left(1 - \frac{\gamma}{x}\right) dx + \gamma + O\left(\frac{1}{(\log n)^2}\right). \end{aligned}$$

Now we use (1.3) and the last display to ensure that  $a_n$  satisfies (1.4) with a function  $L$  chosen such that  $L$  is twice differentiable and  $L(x) = 1 - \frac{\gamma}{x}$  for  $x$  sufficiently large. This ensures that we can apply Theorem 1.1.

**Case**  $w_n = \frac{1}{n(\log n)^\alpha}$  **with**  $\alpha > 1$ . In that case,  $W_n = W_\infty - \frac{(\log n)^{1-\alpha}}{\alpha-1} + O\left(\frac{1}{n}\right)$ . We can first easily check that  $\sum_{i \geq n} w_i^2/W_i^2 = O(n^{-1}(\log n)^{-2\alpha})$ , which ensures that (1.5) holds. Also for any  $k \geq 1$  we have

$$\begin{aligned} \log W_n &= \log(W_\infty - (W_\infty - W_n)) \\ &= \log W_\infty + \log\left(1 - \frac{(\log n)^{1-\alpha}}{W_\infty(\alpha-1)} + O\left(\frac{1}{n}\right)\right) \\ &= \log W_\infty - \sum_{i=1}^k \frac{1}{i} \frac{(\log n)^{i(1-\alpha)}}{W_\infty^i(\alpha-1)^i} + O\left((\log n)^{(k+1)(1-\alpha)}\right) \\ &= \log W_\infty - \int_{\log n}^{\infty} x^{-\alpha} \left( \sum_{i=1}^k \frac{x^{(i-1)(1-\alpha)}}{W_\infty^i(\alpha-1)^{i-1}} \right) dx + O\left((\log n)^{(k+1)(1-\alpha)}\right) \\ &= \int_1^{\log n} x^{-\alpha} \left( \sum_{i=0}^{k-1} \frac{x^{i(1-\alpha)}}{W_\infty^{i+1}(\alpha-1)^i} \right) dx + C + O\left((\log n)^{(k+1)(1-\alpha)}\right). \end{aligned}$$

Now choosing  $k$  so that  $(k+1)(1-\alpha) \leq -1-\alpha$ , and using (1.3) again we get that  $a_n$  satisfies (1.4) with a function  $L$  chosen such that  $L$  is twice differentiable and  $L(x) = \sum_{i=0}^{k-1} \frac{x^{i(1-\alpha)}}{W_\infty^{i+1}(\alpha-1)^i}$  for  $x$  sufficiently large. This ensures that we can apply Theorem 1.1.

**Case**  $w_n = n^{-\alpha}$  **for**  $\alpha > 1$ . It is easy to check that we then have  $\sum_{i \geq n} w_i^2/W_i^2 = O(n^{1-2\alpha})$  so that (1.7) holds. We have  $W_n = W_\infty - \sum_{i=n+1}^{\infty} n^{-\alpha} = W_\infty - \frac{n^{1-\alpha}}{\alpha-1} + o(n^{1-\alpha})$  so that using (1.3) we get

$$a_\infty - a_n = \log W_\infty - \log W_n + O\left(\sum_{i \geq n+1} \frac{w_i^2}{W_i^2}\right) = \frac{n^{1-\alpha}}{W_\infty(\alpha-1)} \cdot (1 + o(1)).$$

Hence, the hypotheses of Theorem 1.4 are satisfied with  $\beta = 1$  and  $J(x) = 1/(W_\infty(\alpha-1))$ .

## Acknowledgments

Delphin Sénizergues was a post-doctoral fellow at the University of British Columbia when the research that led to this paper was conducted.

## References

- [1] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [2] E.-S. Boci and C. Mailler. Large deviations principle for a stochastic process with random reinforced relocations. *arXiv:2105.02633*.
- [3] K. A. Borovkov and V. Vatutin. On the asymptotic behaviour of random recursive trees in random environments. *Adv. in Appl. Probab.*, 38(4):1047–1070, 2006.
- [4] A. Bovier and L. Hartung. The extremal process of two-speed branching Brownian motion. *Electron. J. Probab.*, 19:no. 18, 28, 2014.
- [5] A. Bovier and L. Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1):261–291, 2015.
- [6] A. Bovier and I. Kurkova. Derrida’s generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):481–495, 2004.
- [7] F. Caravenna and L. Chaumont. An invariance principle for random walk bridges conditioned to stay positive. *Electron. J. Probab.*, 18:no. 60, 32, 2013.
- [8] L. Eslava, B. Lodewijks, and M. Ortgièse. Fine asymptotics for the maximum degree in weighted recursive trees with bounded random weights. *arXiv:2109.15270*.
- [9] M. Fang. Tightness for maxima of generalized branching random walks. *J. Appl. Probab.*, 49(3):652–670, 2012.
- [10] M. Fang and O. Zeitouni. Branching random walks in time inhomogeneous environments. *Electron. J. Probab.*, 17:no. 67, 18, 2012.
- [11] M. Fang and O. Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. *J. Stat. Phys.*, 149(1):1–9, 2012.
- [12] N. Fountoulakis, T. Iyer, C. Mailler, and H. Sulzbach. Dynamical models for random simplicial complexes. *arXiv:1910.12715*.
- [13] E. Hiesmayr and U. Işlak. Asymptotic results on Hoppe trees and their variations. *J. Appl. Probab.*, 57(2):441–457, 2020.
- [14] T. Iyer. Degree distributions in recursive trees with fitnesses. *arXiv: 2005.02197*.

- [15] B. Lodewijks. Location of high-degree vertices in weighted recursive graphs with bounded random weights and the random recursive tree. *arXiv:2110.00522*.
- [16] B. Lodewijks and M. Ortgiese. The maximal degree in random recursive graphs with random weights. *arXiv: 2007.05438*.
- [17] P. Maillard and O. Zeitouni. Slowdown in branching Brownian motion with inhomogeneous variance. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1144–1160, 2016.
- [18] C. Mailler and G. Uribe Bravo. Random walks with preferential relocations and fading memory: a study through random recursive trees. *J. Stat. Mech. Theory Exp.*, (9):093206, 49, 2019.
- [19] B. Mallein. Maximal displacement of a branching random walk in time-inhomogeneous environment. *Stochastic Process. Appl.*, 125(10):3958–4019, 2015.
- [20] M. Pain and D. Sénizergues. Correction terms for the height of weighted recursive trees. *Ann. Appl. Probab.*, 2022. to appear.
- [21] D. Sénizergues. Geometry of weighted recursive and affine preferential attachment trees. *Electron. J. Probab.*, 26:Paper No. 80, 56, 2021.