



# The maximum of a tree-indexed random walk in the big jump domain

Pascal Maillard

Département de Mathématiques, Univ. Paris-Sud, CNRS,  
Université Paris-Saclay, 91405 Orsay Cedex, France.  
*E-mail address:* [pascal.maillard@u-psud.fr](mailto:pascal.maillard@u-psud.fr)

**Abstract.** We consider random walks indexed by arbitrary finite random or deterministic trees. We derive a simple sufficient criterion which ensures that the maximal displacement of the tree-indexed random walk is determined by a single large jump. This criterion is given in terms of four quantities: the tail and the expectation of the random walk steps, the height of the tree and the number of its vertices. The results are applied to critical Galton–Watson trees with offspring distributions in the domain of attraction of a stable law.

## 1. Introduction

Let  $S_n = X_1 + \dots + X_n$  be a symmetric random walk with  $\mathbf{P}(X_1 > x) \sim x^{-\alpha}$  as  $x \rightarrow \infty$ , for some  $\alpha > 0$  (here and throughout, we write  $a_n \sim b_n$  if  $a_n/b_n \rightarrow 1$  for two sequences  $(a_n)_{n \geq 0}$  and  $(b_n)_{n \geq 0}$  of positive numbers). It is well-known and easy to show that  $M_n^X = \max(X_1, \dots, X_n) \asymp n^{1/\alpha}$  as  $n \rightarrow \infty$ . On the other hand, standard random walk theory (see, e.g. [Feller, 1971](#)) gives that  $S_n = n^{1/(2 \wedge \alpha) + o(1)}$  as  $n \rightarrow \infty$ . In other words,  $S_n$  and  $M_n^X$  are (roughly) of the same order if and only if  $\alpha \leq 2$ .

Now consider a critical, finite variance branching random walk, whose random walk steps are distributed as above. This means that starting from a Galton–Watson tree with offspring distribution of mean 1 and finite variance, we assign iid random variables  $X_v$  to each non-root vertex  $v$ , distributed as above, and let  $S_v$  be the sum over all  $X_u$  where  $u$  runs through all the non-root vertices on the path from the root to  $v$ . Condition the tree on having  $n$  non-root vertices and let  $M_n^X$  and  $M_n^S$  be the maximum over all  $X_v$  and  $S_v$ , respectively. Of course,  $M_n^X \asymp n^{1/\alpha}$  as above. As for  $M_n^S$ , [Kesten \(1995\)](#) proved that if  $\alpha > 4$ , then  $M_n^S \asymp n^{1/4}$ .

---

*Received by the editors February 2, 2016; accepted June 20, 2016.*

*2010 Mathematics Subject Classification.* 60J10.

*Key words and phrases.* Tree-indexed random walk, branching random walk, heavy tails, extreme values.

Partially supported by Grant ANR-14-CE25-0014 (ANR GRAAL).

Subsequently, [Janson and Marckert \(2005\)](#) showed that in general,  $M_n^S \asymp n^{1/(4 \wedge \alpha)}$ , so that  $M_n^S \asymp M_n^X$  if and only if  $\alpha \leq 4$ .

It is well-known that a critical, finite variance Galton–Watson tree conditioned on having  $n$  vertices converges after rescaling to Aldous’ *continuum random tree*, see [Aldous \(1993\)](#), a random metric space of Hausdorff dimension 2. Looking at the above results, one question immediately comes to mind: for a random walk indexed by a large “ $D$ -dimensional” random tree, a notion to be made precise, is it true that  $M^S \asymp M^X$  if and only if  $\alpha \leq 2D$ , with  $M^S = \max_v S_v$  and  $M^X = \max_v X_v$ ? In this article, we provide a partial response to this question.

The meaning we give to the “dimension” of a tree is very crude and simple: we say that a tree is of dimension (at least)  $D > 1$  if its height is at most of order  $V^{1/D}$ , where  $V$  is the number of its vertices. We will make this definition precise in two ways, yielding two different settings:

- (1) through a growing sequence of (possibly random) trees of height  $H_n$  and number of vertices  $V_n$ , satisfying for each  $\varepsilon > 0$ ,  $H_n \leq V_n^{1/D+\varepsilon}$  with high probability, and
- (2) through a single random tree of height  $H$  and number of vertices  $V$  satisfying for each  $\varepsilon > 0$ ,  $\mathbf{P}(H \leq V^{1/D+\varepsilon}, V \geq n) \leq C_\varepsilon n^{-\kappa}$  for some large enough  $\kappa$ .

Under the condition  $\alpha < 2D$  (or,  $\alpha < D$  for non-centered random walk), we then prove the following:

- In Setting 1, we have  $M_n^S/M_n^X \rightarrow 1$  in probability as  $n \rightarrow \infty$ , with  $M_n^S$  and  $M_n^X$  being respectively the maximal displacement and the size of the maximal jump in the  $n$ -th process ([Theorem 2.1](#)).
- In Setting 2, we have  $\mathbf{P}(M^S > x) \sim \mathbf{P}(M^X > x)$  as  $x \rightarrow \infty$ , where  $M^S$  and  $M^X$  are as above ([Theorem 2.2](#)).

We thus have a very easy to verify sufficient criterion for  $M^S$  to be of the same order as  $M^X$  (in fact, they are approximately equal for reasons explained below).

We expect the notion of dimension used here to coincide with other notions in typical cases of interest. We illustrate this through the example of critical Galton–Watson trees whose offspring distribution is in the domain of attraction of a stable law (which have dimensions  $D \in [2, \infty)$ ).

We finish this introduction by a review of the existing literature on tree-indexed random walks with heavy tails (without pretending to be exhaustive). To the knowledge of the author, these have only been considered so far only on Galton–Watson trees, under the name of branching random walks. For critical, finite-variance Galton–Watson trees, the results of our [Theorem 2.1](#) were shown by [Janson and Marckert \(2005\)](#)<sup>1</sup>. [Lalley and Shao \(2016+\)](#) consider symmetric stable branching Lévy processes of index  $\alpha \in (0, 2)$ , with critical binary offspring distribution, for which they prove the analogue of our [Theorem 2.2](#) through analysis of a certain

<sup>1</sup>Much more is known for critical, finite-variance Galton–Watson trees: under the condition  $\mathbf{P}(|X| > x) = o(x^{-4})$ , [Janson and Marckert \(2005\)](#) showed that a certain exploration process of the branching random walk, called the *discrete snake*, converges after renormalization uniformly to a continuous process called the *Brownian snake* introduced by [Le Gall \(1993\)](#). This result was first proven for  $\alpha > 8$  by [Marckert and Mokkadem \(2003\)](#). On the other hand, for  $\alpha \leq 4$ , Janson and Marckert showed convergence of the discrete snake w.r.t. a certain topology similar to Skorokhod’s  $M_1$ -topology to a certain non-continuous process called the *jumping snake* or *hairy snake*.

pseudo-differential equation. This work was an inspiration to the current article. It is easy to show that their results can be recovered from our results applied to a discrete skeleton of the branching Lévy process.

On supercritical Galton–Watson trees (which correspond to  $D = \infty$ ), [Durrett \(1983\)](#) considers the maximal displacement at generation  $n$  of a tree-indexed random walk with regularly varying tails and shows that it approaches the maximal jump size until generation  $n$  as  $n \rightarrow \infty$ . This result is easily recovered by our [Theorem 2.1](#). [Bhattacharya et al. \(2016+\)](#) recently extended these results to the collection of extremal particles. [Bérard and Maillard \(2014\)](#) considered a supercritical branching random walk with regularly varying tails and with selection of the  $N$  maximal particles, for large  $N$ . Finally, [Gantert \(2000\)](#) studied the maximum of supercritical branching random walk with stretched exponential tails.

To conclude, our contributions in this article are the following:

- to unify previous results on the maximal displacement of heavy-tailed branching random walks,
- to give a simple, transparent proof of these results, and
- to generalize them to arbitrary trees satisfying an easily verifiable condition.

## 2. Definitions and statements of results

The following notation will be used throughout the article. We fix a real-valued random variable  $X$ , whose law is, for now, arbitrary. We say that a sequence  $(b_n)_{n \in \mathbb{N}}$  is a *natural scale sequence*<sup>2</sup> if it is non-decreasing and if the family of random variables  $((X_1 + \dots + X_n)/b_n)_{n \in \mathbb{N}}$  is tight, where  $X_1, X_2, \dots$  are iid copies of  $X$ .

Let  $T$  be a finite tree, deterministic or random, with root  $\rho$ . Denote by  $\mathcal{V}$  the set of vertices and set  $\mathcal{V}^* = \mathcal{V} \setminus \{\rho\}$ . We denote by  $V = |\mathcal{V}^*|$  the number of non-root vertices of the tree  $T$  and by  $H$  its height/depth, i.e. the largest distance between  $\rho$  and another vertex.

Let  $(X_v)_{v \in \mathcal{V}^*}$  be iid of the same law as  $X$ . Set

$$\forall v \in \mathcal{V} : S_v = \sum_{\rho \neq u \leq v} X_u,$$

where  $u \leq v$  means that  $u$  lies on the path from the root to  $v$  (including  $v$  itself). The collection  $(S_v)_{v \in \mathcal{V}}$  is then called the *random walk indexed by the tree  $T$* . Let  $\mathcal{L} \subset \mathcal{V}$  be the subset of leaves (i.e. vertices without descendant) of the tree. We then define,

$$\begin{aligned} M^S &= \max_{v \in \mathcal{V}} S_v, & \widetilde{M}^S &= \max_{v \in \mathcal{L}} S_v, & M^X &= \max_{v \in \mathcal{V}^*} X_v \\ M^{|\mathcal{S}|} &= \max_{v \in \mathcal{V}} |S_v|, & \widetilde{M}^{|\mathcal{S}|} &= \max_{v \in \mathcal{L}} |S_v|, & M^{|\mathcal{X}|} &= \max_{v \in \mathcal{V}^*} |X_v|. \end{aligned}$$

Note that trivially,  $\widetilde{M}^S \leq M^S$  and  $\widetilde{M}^{|\mathcal{S}|} \leq M^{|\mathcal{S}|}$ . We further define

$$\Delta^{SX} = \max \left\{ |M^S - M^X|, |\widetilde{M}^S - M^X|, |M^{|\mathcal{S}|} - M^{|\mathcal{X}|}|, |\widetilde{M}^{|\mathcal{S}|} - M^{|\mathcal{X}|}| \right\}.$$

---

<sup>2</sup>This is the same definition as in [Denisov et al. \(2008\)](#) apart from the fact that we require furthermore that  $b_n$  is non-decreasing in  $n$ .

We will also consider sequences  $(T_n)_{n \in \mathbb{N}}$  of random trees, in which case we denote by  $V_n, H_n, M_n^X, M_n^S$ , etc. the objects introduced above corresponding to the tree  $T_n$ .

We now introduce the assumptions on the class of tree-indexed random walks we will focus on. The assumption on the law of  $X$  is

(XR) There exists  $\alpha > 0$ , such that  $\mathbf{P}(X > x)$  and  $\mathbf{P}(|X| > x)$  are regularly varying<sup>3</sup> at  $\infty$  with index  $-\alpha$ . In this case, we set

$$D_{\text{crit}}(X) = \begin{cases} \max(1, \frac{\alpha}{2}) & \text{if } \mathbf{E}[|X|] < \infty \text{ and } \mathbf{E}[X] = 0 \\ \max(1, \alpha) & \text{otherwise.} \end{cases}$$

As for the underlying tree, recall from the introduction that we consider two different settings, the first involving a sequence  $(T_n)_{n \in \mathbb{N}}$  of growing random trees, the second involving a fixed, random tree  $T$ . In the first setting, the assumptions on the sequence  $(T_n)_{n \in \mathbb{N}}$  are

(Tn1)  $V_n \rightarrow +\infty$  in probability, as  $n \rightarrow \infty$ .

(Tn2) There exists  $D > 1$ , such that for every  $\varepsilon > 0$ ,

$$\mathbf{P}(H_n > V_n^{\frac{1}{D} + \varepsilon}) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

We then have the following result:

**Theorem 2.1.** *Assume (XR), (Tn1) and (Tn2). Assume that  $D > D_{\text{crit}}(X)$ , with  $D$  and  $D_{\text{crit}}(X)$  from (Tn2) and (XR), respectively. Then*

$$\lim_{n \rightarrow \infty} \frac{M_n^S}{M_n^X} = \lim_{n \rightarrow \infty} \frac{\widetilde{M}_n^S}{M_n^X} = \lim_{n \rightarrow \infty} \frac{M_n^{|S|}}{M_n^{|X|}} = \lim_{n \rightarrow \infty} \frac{\widetilde{M}_n^{|S|}}{M_n^{|X|}} = 1, \quad \text{in probability.}$$

In particular, if  $V_n/n \rightarrow 1$  in probability as  $n \rightarrow \infty$  and if  $a_n$  and  $\tilde{a}_n$  are such that  $\mathbf{P}(X > a_n) \sim 1/n$  and  $\mathbf{P}(|X| > \tilde{a}_n) \sim 1/n$  as  $n \rightarrow \infty$ , then the random variables  $a_n^{-1}M_n^S$ ,  $a_n^{-1}\widetilde{M}_n^S$ ,  $\tilde{a}_n^{-1}M_n^{|S|}$  and  $\tilde{a}_n^{-1}\widetilde{M}_n^{|S|}$  converge in law as  $n \rightarrow \infty$  to a Pareto law on  $[0, \infty)$  with distribution function  $F(x) = \exp(-x^{-\alpha})$ .

In the second setting, the assumptions on the law of the tree  $T$  are the following. Here, the generating function of  $V$  is denoted by  $g_V(s) = \mathbf{E}[s^V]$ .

(T1) There exists  $\beta \in (0, 1]$ , such that  $1 - g_V(1 - s)$  is regularly varying at 0 with index  $\beta$ .

(T2) There exists  $D > 1$  and  $\bar{\gamma} > 1 \vee \alpha$ , with  $\alpha$  from (XR), such that for every  $\varepsilon > 0$ ,

$$\mathbf{P}(H > V^{\frac{1}{D} + \varepsilon}, V \geq n) \times n^{\beta \bar{\gamma}} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Here  $\beta$  is the constant from (T1).

Our result under the previous hypotheses is the following:

**Theorem 2.2.** *Assume (XR), (T1) and (T2), Assume that  $D > D_{\text{crit}}(X)$ , with  $D$  and  $D_{\text{crit}}(X)$  from (T2) and (XR), respectively. Then, with  $g_V$  from (T1),*

$$\mathbf{P}(M^S > x) \sim \mathbf{P}(\widetilde{M}^S > x) \sim \mathbf{P}(M^X > x) = 1 - g_V(1 - \mathbf{P}(X > x)) \quad \text{as } x \rightarrow \infty$$

and

$$\mathbf{P}(M^{|S|} > x) \sim \mathbf{P}(\widetilde{M}^{|S|} > x) \sim \mathbf{P}(M^{|X|} > x) = 1 - g_V(1 - \mathbf{P}(|X| > x)) \quad \text{as } x \rightarrow \infty.$$

<sup>3</sup>The definition and basic properties of regularly varying functions are recalled in Section A.

In fact, we have

$$\lim_{x \rightarrow \infty} \frac{\mathbf{P}(\Delta^{SX} > x)}{\mathbf{P}(M^X > x)} = \lim_{x \rightarrow \infty} \frac{\mathbf{P}(\Delta^{SX} > x)}{\mathbf{P}(M^{|X|} > x)} = 0.$$

The basic estimate needed in the proof of Theorems 2.1 and 2.2 is the following proposition, which holds in full generality.

**Proposition 2.3.** *Let  $(b_n)_{n \in \mathbb{N}}$  be a natural scale sequence. For every  $z \geq b_H$  and  $y \geq 0$ , we have*

$$\mathbf{P}(\Delta^{SX} > y) \leq \frac{HV}{2} \mathbf{P}(|X| > z)^2 + CV \exp(-y/z),$$

where the constant  $C$  depends only on the law of  $X$  and on the sequence  $(b_n)_{n \in \mathbb{N}}$ .

We comment in Section 3 below on the applicability of Proposition 2.3 and on possible improvements. For now, we just mention that the result is not at all optimal, but in the case where  $X$  has regularly varying tails, which is the focus of this article, it is more than enough for our purposes.

Theorems 2.1 and 2.2 apply to a large class of tree-indexed random walks. We consider now the particular case of (critical) *branching random walks*, i.e. random walks indexed by (critical) Galton–Watson trees. These have been well-studied in the literature (see below for a survey of existing results). Let  $\mathbf{p} = (p_n)_{n \in \mathbb{N}}$  be a probability distribution on  $\mathbb{N} = \{0, 1, 2, \dots\}$ . We recall that a Galton–Watson tree  $T$  with offspring distribution  $\mathbf{p}$  is a random rooted tree, such that the degree  $d_\rho$  of the root follows the law  $\mathbf{p}$  and the subtree of the neighbors of the root are independent copies of the tree  $T$ , independent of  $d_\rho$ . We say that  $T$  is *critical* if  $\mathbf{p}$  has mean 1, i.e.  $\sum np_n = 1$ . In this case  $T$  is finite almost surely. The following proposition now says that Theorems 2.1 and 2.2 can be applied for random walks indexed by critical Galton–Watson trees:

**Proposition 2.4.** *Assume that  $T$  is a critical Galton–Watson tree whose offspring distribution is in the domain of attraction of an  $\alpha_T$ -stable law,  $\alpha_T \in (1, 2]$ , and let  $D = \alpha_T/(\alpha_T - 1)$ . Then the following statements hold:*

- (1) *For  $n \in \mathbb{N}$ , let  $T_n$  be a random tree following the law of  $T$  conditioned on having  $n$  vertices. Then Assumptions (Tn1) and (Tn2) hold with the same  $D$ .*
- (2) *Assumptions (T1) and (T2) above hold with the same  $D$ ,  $\beta = 1/\alpha_T$  and every  $\bar{\gamma} \in \mathbb{R}$ . The asymptotic on  $1 - g_V(1 - s)$  as  $s \rightarrow 0$  is given in (6.5) and (6.6).*

We remark that the “dimension”  $D = \alpha_T/(\alpha_T - 1)$  in Proposition 2.4 above indeed coincides with the Hausdorff (or packing) dimension of the  $\alpha_T$ -stable tree, see Duquesne and Le Gall (2005), which is the scaling limit of the trees  $T_n$  from Proposition 2.4, see Duquesne (2003).

Apart from critical Galton–Watson trees, which cover the range  $D \in [2, \infty)$ , we remark that the results in this paper also allow to treat supercritical Galton–Watson trees, which correspond to  $D = \infty$ .

Overview of the remainder of the article. Section 3 presents the heuristics that underly the proofs of Theorems 2.1 and 2.2. It also discusses possible generalizations. Section 4 is devoted to the proof of Proposition 2.3. Theorems 2.1 and 2.2 are proven in Section 5. The case of Galton–Watson trees is considered in Section 6

where Proposition 2.4 is proven. Finally, an appendix, Section A, recalls standard properties of regularly varying functions.

### 3. Heuristics and discussion

In this article, we consider tree-indexed random walks with heavy tails in a regime where the total maximal displacement is attained by atypically large fluctuations on certain branches. At the heart of the results are therefore large deviation results for heavy-tailed random walks. Recall that a random variable  $X$  is said to follow a *subexponential distribution* if

$$\lim_{x \rightarrow \infty} \frac{\mathbf{P}(S_n > x)}{n\mathbf{P}(X > x)} = 1 \quad \text{for all } n \in \mathbb{N},$$

where  $S_n = X_1 + \dots + X_n$  with  $X_1, X_2, \dots$  iid copies of  $X$ . A huge body of literature is devoted to the problem of finding (optimal or nearly optimal) sequences  $x_n \rightarrow \infty$  such that  $\mathbf{P}(S_n > x) \approx n\mathbf{P}(X > x)$  for all  $x \geq x_n$ . The literature on this topic is quite overwhelming<sup>4</sup>, due in part to the large variety of subexponential distributions leading to a substantial number of treatments differing in results and/or techniques.

For our Proposition 2.3, we use an “exponential bound” from Denisov et al. (2008) on suitably truncated random walks, which we recall in Equation (4.1) below. This is the only random walk bound we use in this paper. We combine it with the simple observation that the maximal displacement of a tree-indexed random walk is approximately equal to the size of the maximal jump if the following two events happen:

- (1) No two large jumps occur on the same branch of the tree with high probability.
- (2) The contribution of the small jumps are asymptotically negligible.

Quantifying this leads to Proposition 2.3. Note that the precise bound given in the statement of Proposition 2.3 will turn out to be not so important, but rather the assumption that  $y \geq b_H$  which will need to be verified for the values of  $y$  we will be interested in. This assumption exactly corresponds to requiring that the typical value of the maximal displacement or of the biggest jump is typically much larger than the values of a random walk along a (fixed) branch of the tree.

For Theorems 2.1 and 2.2 we make explicit such a regime in the case of regularly varying displacement. Under Assumption (XR), this means that we require that  $V^{1/\alpha+o(1)} \gg b_H$ , where  $b_n$  is a natural scale sequence for the random walk. This amounts to the assumption  $D > D_{\text{crit}}(X)$  in Theorems 2.1 and 2.2.

We do not go into further details of the proofs of Theorems 2.1 and 2.2 here. However, we elaborate briefly on the possible improvements of Proposition 2.3. First, the exponential bound (4.1) might be replaced by a better bound. Indeed, although it is fairly efficient for values of  $x$  such that  $\mathbf{P}(S_n > x)$  is not too small, it

<sup>4</sup>Classical references are Linnik (1961b,a); Heyde (1967); Nagaev (1969a,b), for more see Embrechts et al. (1997, Section 8.6), Mikosch and Nagaev (1998) and Denisov et al. (2008). Treatments of general distributions with regularly varying tails appear in Durrett (1979); Cline and Hsing (1998). Denisov et al. (2008) give a uniform and fairly insightful treatment of general subexponential distributions.

is quite bad for those  $x$  for which  $\mathbf{P}(S_n > x)$  is very small<sup>5</sup>. Although this is not too much of a problem in the case of regularly varying tails, it is in fact disastrous in the case of stretched exponential tails  $\mathbf{P}(X > x) \approx \exp(-x^r)$  when  $r$  is small (see [Gantert \(2000, Theorem 3\)](#) for bounds on truncated random walks with stretched exponential tails). The reason why we chose (4.1) is its simplicity (both statement and proof) and its generality: we are not aware of any other work which allows to treat such a large class of distributions.

As a second improvement of Proposition 2.3, one might relax the first of the two points mentioned above. Namely, instead of just throwing away the event where two large jumps occur on the same branch of the tree, one might instead separate the jumps into large jumps and small jumps, ignore the small jumps and then consider the skeleton of the tree consisting of the large jumps. Then, by a certain induction, one might use bounds on this smaller tree to get bounds on the original, larger tree. Working out such an argument would give better quantitative bounds on the difference between the maximal displacement and the maximal jump size. For example it would give good bounds on the range of values  $x$  (in terms of  $n$ ) such that  $\mathbf{P}(M_n^S > x)/\mathbf{P}(M_n^X > x) \in [1 - \varepsilon, 1 + \varepsilon]$  in Theorem 2.1, for  $\varepsilon > 0$ . However, in order to keep the current proof as simple as possible, we did not pursue this argument and leave it open for future work.

#### 4. Proof of Proposition 2.3

We start with a large deviation estimate for random walks. Let  $X_1, X_2, \dots$  be iid copies of the random variable  $X$  and define  $S_n = X_1 + \dots + X_n$  and for  $y \geq 0$ ,

$$S_n^{(y)} = \sum_{k=1}^n X_k \mathbb{1}_{(|X_k| \leq y)}.$$

We then have the following lemma:

**Lemma 4.1.** *Let  $(b_n)_{n \in \mathbb{N}}$  be a natural scale sequence. Then there exists  $C \in (0, \infty)$  (depending only on the law of  $X$  and on the sequence  $(b_n)_{n \in \mathbb{N}}$ ), such that*

$$\forall n \in \mathbb{N} \forall x \geq 0 \forall y \geq b_n : \mathbf{P}(|S_n^{(y)}| > x) \leq C \exp(-x/y).$$

*Proof:* The lemma is a simple extension of the following bound ([Denisov et al., 2008, Lemma 2.1](#)):

$$\begin{aligned} \exists C \in (0, \infty) \forall n \in \mathbb{N} \forall x \geq 0 \forall y \geq b_n : \\ \mathbf{P}(|S_n| > x, |X_1| \leq y, \dots, |X_n| \leq y) \leq C \exp(-x/y). \end{aligned} \tag{4.1}$$

In order to use (4.1), we decompose:

$$\begin{aligned} \mathbf{P}(|S_n^{(y)}| > x) &= \sum_{I \subset \{1, \dots, n\}} \mathbf{P}(|S_n^{(y)}| > x, \forall i \in I : |X_i| \leq y, \forall j \notin I : |X_j| > y) \\ &= \sum_{k=0}^n \binom{n}{k} \mathbf{P}(|S_{n-k}| > x, |X_1| \leq y, \dots, |X_{n-k}| \leq y) \mathbf{P}(|X| > y)^k. \end{aligned}$$

<sup>5</sup>For example, in case of regularly varying tails it is known that the sum over the jumps of size at most  $(1 - \varepsilon)x$ ,  $\varepsilon > 0$ , is negligible for large enough  $x$  compared to the non-truncated sum [Durrett \(1979\)](#); [Cline and Hsing \(1998\)](#), a fact which is not apparent from (4.1)

Since  $b_n$  is increasing by assumption, we have  $y \geq b_{n-k}$  for every  $k \in \{0, \dots, n\}$ . With (4.1), this gives for all  $n \in \mathbb{N}$ ,  $x \geq 0$ ,  $y \geq b_n$ :

$$\begin{aligned} \mathbf{P}(|S_n^{(y)}| > x) &\leq C \exp(-x/y) \sum_{k=0}^n \binom{n}{k} \mathbf{P}(|X| > y)^k \\ &\leq C \exp(-x/y) (1 + \mathbf{P}(|X| > b_n))^n \\ &\leq C \exp(-x/y) \exp(n\mathbf{P}(|X| > b_n)). \end{aligned}$$

From the proof of the lemma in Feller (1971, Section IX.7), we have  $\sup_n n\mathbf{P}(|X| > b_n) < \infty$ . This yields the lemma.  $\square$

*Proof of Proposition 2.3:* Fix  $z \geq b_H$  and  $y \geq 0$ . Define the events

$$\begin{aligned} G_1 &= \{\forall u, v \in \mathcal{V}^* : \text{if } |X_u| > z \text{ and } |X_v| > z, \text{ then } u \not\prec v\}, \\ G_2 &= \{\forall v \in \mathcal{V} : |S_v^{(z)}| \leq y\}. \end{aligned}$$

Here, similarly to the definition at the beginning of the section, we define  $S_v^{(z)} = \sum_{\rho \neq u \leq v} X_u \mathbb{1}_{(|X_u| \leq z)}$ . It is clear that

$$\text{on } G_1 \cap G_2, \quad \Delta^{SX} \leq y.$$

It therefore suffices to bound  $\mathbf{P}(G_1^c)$  and  $\mathbf{P}(G_2^c)$ . By a union bound, we have

$$\mathbf{P}(G_1^c) \leq \sum_{u, v \in \mathcal{V}^*, u < v} \mathbf{P}(|X| > z)^2 \leq \frac{HV}{2} \mathbf{P}(|X| > z)^2.$$

Furthermore, again by a union bound,

$$\mathbf{P}(G_2^c) \leq \sum_{v \in \mathcal{V}^*} \mathbf{P}(|S_v^{(z)}| > y) \leq CV \exp(-y/z),$$

where the last inequality follows from Lemma 4.1 and the fact that  $z \geq b_n$  for all  $n \leq H$  since  $b_n$  is non-decreasing in  $n$ . The two previous inequalities then yield

$$\mathbf{P}((G_1 \cap G_2)^c) \leq \mathbf{P}(G_1^c) + \mathbf{P}(G_2^c) \leq \frac{HV}{2} \mathbf{P}(|X| > z)^2 + CV \exp(-y/z).$$

This finishes the proof.  $\square$

### 5. Regularly varying displacement: proofs of Theorems 2.1 and 2.2

Throughout this section, we will make use of some classic results on regularly varying functions (in particular, Potter’s bounds), readers not familiar with this theory may refer to Section A where these results are recalled.

Assume from now on that Assumption (XR) holds. Standard results on triangular arrays (Feller, 1971, Section XI.7) or the domain of attraction of stable laws (Feller, 1971, Section XVII.5), together with Potter’s bounds easily give that for every  $\varepsilon > 0$  the following sequence is a natural scale sequence for the random walk  $(S_n)_{n \in \mathbb{N}}$ :

$$b_n^\varepsilon = \begin{cases} n^{1/(2 \wedge \alpha) + \varepsilon}, & \text{if } \mathbf{E}[|X|] < \infty \text{ and } \mathbf{E}[X] = 0 \\ n^{1/(1 \wedge \alpha) + \varepsilon}, & \text{otherwise.} \end{cases} \tag{5.1}$$

We now first simplify Lemma 4.1 to the current setting:



**Corollary 5.1.** *For every  $\varepsilon > 0$ , there exists  $C = C(\varepsilon) \in (0, \infty)$ , such that for every tree  $T$ , for every  $y \geq b_H^\varepsilon$ ,*

$$\mathbf{P}(\Delta^{SX} > y) \leq CHV y^{-(2-\varepsilon)\alpha}.$$

*Proof:* Fix  $\varepsilon \in (0, 1)$ . Let  $\delta \in (0, \varepsilon/2)$ . By the assumption on  $\mathbf{P}(|X| > x)$  and Potter's bounds, there exists  $C' = C'(\varepsilon, \delta)$ , such that for all  $y > 0$ ,

$$\mathbf{P}(|X| > y^{1-\delta})^2 + \exp(-y^\delta) \leq C' y^{-(2-\varepsilon)\alpha}.$$

Now choose  $\delta = \delta(\varepsilon, \alpha) \in (0, \varepsilon/2)$  such that  $(b_n^\varepsilon)^{1-\delta}$  is a natural scale sequence. By Proposition 2.3 applied with  $z = y^{1-\delta}$ , we then have for every tree  $T$  and for every  $y \geq b_H^\varepsilon$ ,

$$\mathbf{P}(\Delta^{SX} > y) \leq \frac{HV}{2} \mathbf{P}(|X| > y^{1-\delta})^2 + CV \exp(-y^\delta) \leq C'(C + H)V y^{-(2-\varepsilon)\alpha},$$

where  $C$  is the constant from Proposition 2.3. This proves the corollary for all trees of height  $H \geq 1$ . For  $H = 0$  the bound trivially holds. This finishes the proof.  $\square$

We are now ready for the proofs of Theorems 2.1 and 2.2.

*Proof of Theorem 2.1:* We trivially have  $M_n^X \leq M_n^{|X|}$ . In order to prove the first statement of the theorem, it is therefore enough to show that

$$\lim_{n \rightarrow \infty} \frac{\Delta_n^{SX}}{M_n^X} = 0, \quad \text{in probability.} \tag{5.2}$$

We first give a lower bound on  $M_n^X$ . By independence, we have for all  $x$ ,

$$\mathbf{P}(M_n^X \leq x | T_n) = \mathbf{P}(X \leq x)^{V_n}. \tag{5.3}$$

Potter's bounds then give for every  $\varepsilon > 0$  and for large  $x$ ,

$$\mathbf{P}(M_n^X \leq x | T_n) = (1 - x^{-(\alpha+\varepsilon)})^{V_n} \leq \exp(-V_n x^{-(\alpha+\varepsilon)}).$$

Applying this with  $x = V_n^{1/(\alpha+2\varepsilon)}$  gives for every  $\varepsilon > 0$ ,

$$\mathbf{P}(M_n^X \leq V_n^{1/(\alpha+\varepsilon)}) \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{5.4}$$

Now let  $b_n^\varepsilon$  be as in (5.1), so that  $b_n^\varepsilon = n^{\eta+\varepsilon}$ , where

$$\eta = \begin{cases} 1/(2 \wedge \alpha) & \text{if } \mathbf{E}[|X|] < \infty \text{ and } \mathbf{E}[X] = 0 \\ 1/(1 \wedge \alpha) & \text{otherwise.} \end{cases} \tag{5.5}$$

Note that by the assumption on  $\alpha$  and the assumption  $D > 1$ , we have  $\eta/D < 1/\alpha$ . In particular, for every  $\varepsilon > 0$  small enough, we have for every  $h, v \in \mathbb{N}$

$$h \leq v^{1/D+\varepsilon} \Rightarrow b_h^\varepsilon \leq v^{1/(\alpha+\varepsilon)} \quad \text{for } v \text{ large enough.} \tag{5.6}$$

By Corollary 5.1 (applied to the tree-indexed random walk conditioned on the tree  $T_n$ ), we now have for every  $\varepsilon > 0$ , for some  $C = C(\varepsilon) < \infty$ ,

$$\begin{aligned} & \mathbf{P}(\Delta_n^{SX} > V_n^{1/(\alpha+\varepsilon)} | T_n) \mathbb{1}_{(H_n \leq V_n^{1/D+\varepsilon}, b_{H_n}^\varepsilon \leq V_n^{1/(\alpha+\varepsilon)})} \\ & \leq CH_n V_n V_n^{-(2-\varepsilon)\alpha/(\alpha+\varepsilon)} \mathbb{1}_{(H_n \leq V_n^{1/D+\varepsilon})} \\ & \leq CV_n^{1/D+1-2+C'\varepsilon}, \end{aligned} \tag{5.7}$$

for some constant  $C' = C'(\alpha)$ . Since  $D > 1$  by (Tn2) and  $V_n \rightarrow +\infty$  in probability by (Tn1), the previous equations (5.7) and (5.6) show that for every  $\varepsilon > 0$  small enough,

$$\mathbf{P}(\Delta_n^{SX} > V_n^{1/(\alpha+\varepsilon)}, H_n \leq V_n^{1/D+\varepsilon}) \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{5.8}$$

Summing up the above equations, we have for some  $\varepsilon > 0$  small enough,

$$\begin{aligned} & \mathbf{P}(\Delta_n^{SX} > (M_n^X)^{(\alpha+\varepsilon/2)/(\alpha+\varepsilon)}) \leq \\ & \mathbf{P}(\Delta_n^{SX} > V_n^{1/(\alpha+\varepsilon)}, H_n \leq V_n^{1/D+\varepsilon}) + \mathbf{P}(M_n^X \leq V_n^{1/(\alpha+\varepsilon/2)}) + \mathbf{P}(H_n > V_n^{1/D+\varepsilon}), \end{aligned}$$

and all three terms go to zero as  $n \rightarrow \infty$  by (5.8), (5.4) and assumption (Tn2), respectively. Since  $M_n^X \rightarrow \infty$  in probability by assumption (Tn1), this proves (5.2) and therefore finishes the proof of the first statement of the theorem. The second statement follows easily from (5.3) and standard arguments.  $\square$

*Proof of Theorem 2.2:* We only prove the statements involving  $M^X$ . The first statement involving  $M^{|X|}$  then immediately follows since  $M^X \leq M^{|X|}$  by definition, and the second statement involving  $M^{|X|}$  is proven similarly.

As in (5.3) we have by independence,

$$\mathbf{P}(M^X > x) = 1 - \mathbf{E}[\mathbf{P}(X \leq x)^V] = 1 - g_V(1 - \mathbf{P}(X > x)). \tag{5.9}$$

With Assumptions (XR) and (T1), this implies that  $\mathbf{P}(M^X > x)$  is regularly varying at  $\infty$  (with index  $-\beta\alpha$ ), such that in particular, for every  $\delta(x) \rightarrow 0$  as  $x \rightarrow \infty$ , by the uniform convergence theorem for regularly varying functions (Bingham et al., 1987, Theorem 1.2.1),

$$\mathbf{P}(M^X > (1 - \delta(x))x) \sim \mathbf{P}(M^X > x) \sim \mathbf{P}(M^X > (1 + \delta(x))x), \quad \text{as } x \rightarrow \infty. \tag{5.10}$$

Fix<sup>6</sup>  $\gamma \in (\alpha/\bar{\gamma}, 1 \wedge \alpha)$ . For  $\varepsilon > 0$  and  $x > 0$ , set

$$G_T := \{V \leq x^\gamma\} \cup \{H \leq V^{\frac{1}{D}+\varepsilon}, V \leq x^{\alpha+\varepsilon}\}$$

(note that  $G_T$  depends on  $\varepsilon$  and  $x$  but that we suppress this from the notation for readability). Fix a positive function  $\delta(x)$  converging to 0 slower than polynomially, e.g.  $\delta(x) = 1/\log(2+x)$ . We claim that for small enough  $\varepsilon$ ,

$$\mathbf{P}(G_T^c) = o(\mathbf{P}(M^X > x)) \quad \text{as } x \rightarrow \infty, \tag{5.11}$$

$$\mathbf{P}(\Delta^{SX} > \delta(x)x, G_T) = o(\mathbf{P}(M^X > x)) \quad \text{as } x \rightarrow \infty. \tag{5.12}$$

Let us show how Equations (5.9), (5.10), (5.11) and (5.12) together imply the theorem. First, by (5.11) and (5.12), we have

$$\mathbf{P}(\Delta^{SX} > \delta(x)x) \leq \mathbf{P}(\Delta^{SX} > \delta(x)x, G_T) + \mathbf{P}(G_T^c) = o(\mathbf{P}(M^X > x)), \tag{5.13}$$

which implies in particular the last statement of the theorem. As for the other statements, note that we have

$$\text{on } \{\Delta^{SX} \leq \delta(x)x\}, \quad \{M^X > (1 + \delta(x))x\} \subset \{M^S > x\} \subset \{M^X > (1 - \delta(x))x\}.$$

Hence, by (5.13), (5.10) and (5.9),

$$\mathbf{P}(M^S > x) \sim \mathbf{P}(M^X > x) = 1 - g_V(1 - \mathbf{P}(X > x)).$$

The remaining statements of the theorem follow similarly.

<sup>6</sup>Note that  $\gamma$  exists, since by Assumption (T2),  $(1 \wedge \alpha)\bar{\gamma} > (1 \wedge \alpha)(1 \vee \alpha) = \alpha$ .

It remains to prove (5.11) and (5.12). Let us start with (5.11). We have

$$\mathbf{P}(G_T^c) \leq \mathbf{P}(V > x^{\alpha+\varepsilon}) + \mathbf{P}(H > V^{\frac{1}{D}+\varepsilon}, V > x^\gamma)$$

by Assumption (T1) and Karamata's Tauberian theorem<sup>7</sup> (Fact A.2),  $\mathbf{P}(V \geq y)$  is dominated by a regularly varying function with index  $-\beta$  at  $+\infty$ . Potter's bounds then give that the first summand is bounded by  $x^{-\beta\alpha-\beta\varepsilon/2}$  for large  $x$ . As for the second summand, by Assumption (T2) it is bounded by  $x^{-\beta\gamma\bar{\gamma}}$  for large  $x$ , and  $\gamma\bar{\gamma} > \alpha$  by definition of  $\gamma$ . Since  $\mathbf{P}(M^X > x)$  is regularly varying with index  $-\beta\alpha$ , this readily implies (5.11).

It remains to show (5.12). We want to apply Corollary 5.1 and thus need to show:

$$\forall \varepsilon, \varepsilon' > 0 \text{ small enough, } \forall x \geq 1, \text{ we have } x \geq b_H^{\varepsilon'} \text{ on the event } G_T, \quad (5.14)$$

where  $b_H^{\varepsilon'}$  is defined in (5.1). Let  $\eta$  be as in (5.5), such that  $b_H^{\varepsilon'} = H^{\eta+\varepsilon'}$ . Then, on the event  $\{V \leq x^\gamma\}$ , since  $H \leq V$ , we have  $b_H^{\varepsilon'} \leq x^{\gamma(\eta+\varepsilon')}$ . Now note that  $\gamma\eta < (1 \wedge \alpha)/(1 \wedge \alpha) = 1$ . Hence, for  $\varepsilon'$  small enough and  $x \geq 1$ , we have  $x \geq b_H^{\varepsilon'}$  on the event  $\{V \leq x^\gamma\}$ .

Now suppose that  $H \leq V^{\frac{1}{D}+\varepsilon}$  and  $V \leq x^{\alpha+\varepsilon}$ , such that  $H \leq x^{\alpha/D+C\varepsilon}$ , with  $C = 1/D + \alpha + \varepsilon$ . By the assumption on  $\alpha$  and the assumption  $D > 1$ , we have  $(\alpha/D)\eta < 1$ . In particular, for every  $\varepsilon', \varepsilon > 0$  small enough, we have  $b_H^{\varepsilon'} \leq x^{(\alpha/D+C\varepsilon)(\eta+\varepsilon')} \leq x$  for  $x \geq 1$ . This proves (5.14).

Now suppose for the rest of the proof that  $\varepsilon, \varepsilon' > 0$  are such that both (5.11) and (5.14) hold. Then  $\delta(x)x \geq b_H^{\varepsilon''}$  on the event  $G_T$  for large  $x$ , for every  $\varepsilon'' < \varepsilon'$ . Corollary 5.1 then yields for every  $\varepsilon'' < \varepsilon'$ , for large  $x$ , on the event  $G_T$ , with  $C = C(\varepsilon')$ ,

$$\mathbf{P}(\Delta^{SX} > \delta(x)x | T) \leq CHV(\delta(x)x)^{-(2-\varepsilon'')\alpha} \leq CHVx^{-(2-2\varepsilon'')\alpha}.$$

Integrating over  $T$  and using the fact that  $H \leq V$ , this gives for every  $\varepsilon'' < \varepsilon'/2$ , for large  $x$ ,

$$\mathbf{P}(\Delta^{SX} > \delta(x)x, G_T) \leq C\mathbf{E} \left[ V^2 \mathbb{1}_{(V \leq x^\gamma)} + V^{\frac{1}{D}+1+\varepsilon} \mathbb{1}_{(V \leq x^{\alpha+\varepsilon})} \right] x^{-(2-\varepsilon'')\alpha}. \quad (5.15)$$

Since  $\mathbf{P}(V > y)$  is dominated by a regularly varying function with index  $-\beta$  at  $+\infty$  (see above), this gives for every  $\tilde{\varepsilon} > 0$  and every  $y > 0$  and  $r > \beta$ ,

$$\mathbf{E}[V^r \mathbb{1}_{(V \leq y)}] \leq r \int_0^y z^{r-1} \mathbf{P}(V > z) dz \leq Cr \int_0^y z^{r-1-\beta+\tilde{\varepsilon}} dz \leq C \frac{r}{r-\beta} y^{r-\beta+\tilde{\varepsilon}}.$$

Together with (5.15), this gives for all  $\varepsilon > 0$  small enough, all  $\varepsilon' > 0$  and all  $x$  large enough,

$$\mathbf{P}(\Delta^{SX} > \delta(x)x, G_T) \leq C(x^{\gamma(2-\beta)-2\alpha} + x^{\alpha(\frac{1}{D}+1-\beta-2)+C'\varepsilon})x^{\varepsilon'},$$

with  $C' = 1/D + 1 + \alpha + \varepsilon$ . Now, since  $\gamma < 1 \wedge \alpha \leq \alpha$ , we have  $\gamma(2-\beta) - 2\alpha < -\beta\alpha$ . Furthermore, since  $D > 1$ , we have  $\alpha(\frac{1}{D} + 1 - \beta - 2) < -\beta\alpha$  as well. Choosing  $\varepsilon$  and  $\varepsilon'$  small enough in the previous inequality and using again the fact that  $\mathbf{P}(M^X > x)$  is regularly varying with index  $-\beta\alpha$  yields (5.12) and thus finishes the proof of the theorem.  $\square$

<sup>7</sup>In the case  $\beta = 1$ , use that  $\mathbf{E}[\min(X, x)] = \int_0^x \mathbf{P}(X > y) dy$  and that  $\mathbf{P}(X > y)$  is decreasing in  $y$ .

## 6. Galton–Watson trees: proof of Proposition 2.4

Throughout the section, we denote by  $T$  a critical Galton–Watson tree with offspring distribution  $\mathbf{p} = (p_n)_{n \in \mathbb{N}}$  in the domain of attraction of an  $\alpha_T$ -stable law,  $\alpha_T \in (1, 2]$ . We also denote by  $H$  its height and by  $V$  the number of its vertices<sup>8</sup>. Since  $\mathbf{p} = (p_n)_{n \in \mathbb{N}}$  is in the domain of attraction of an  $\alpha_T$ -stable law, there exists a sequence  $a_n = L(n)n^{1/\alpha_T}$  with a slowly varying function  $L$  such that  $a_n^{-1}(Z_1 + \cdots + Z_n - n)$  converges in distribution (to an  $\alpha_T$ -stable law), where  $Z_1, Z_2, \dots$  are iid copies of a random variable  $Z$  with law  $\mathbf{p}$ . Note that this means that there exists<sup>9</sup>  $\lambda > 0$ , such that Zolotarev (1986, Theorem 2.6.1)

$$\forall t \geq 0 : \lim_{n \rightarrow \infty} \mathbf{E}[e^{-t(Z-1)/a_n}]^n = e^{\lambda t^{\alpha_T}}. \quad (6.1)$$

For the second part of Proposition 2.4, we will need asymptotics for the generating function  $g_V$  of  $V$ . These are well-known, but we establish them here for completeness. Recall the following formula<sup>10</sup> due to Good (1949):

$$g_V(s) = sg_Z(g_V(s)), \quad \text{for } |s| \leq 1, \quad (6.2)$$

where  $g_Z$  is the generating function of the random variable  $Z$ . It will be more useful to translate this formula in terms of log-Laplace transforms. For  $t \geq 0$ , let  $\kappa_V(t) = \log \mathbf{E}[e^{-tV}]$  and  $\kappa_{Z-1}(t) = \log \mathbf{E}[e^{-t(Z-1)}]$ . Then (6.2) becomes

$$\kappa_V = -\kappa_{Z-1}^{-1} \quad \text{on } [0, \infty), \quad (6.3)$$

where the existence of the inverse follows from simple convexity arguments.

Equation (6.1) now gives

$$\forall t \geq 0 : \quad \kappa_{Z-1}(t/a_n) \sim \lambda t^{\alpha_T}/n, \quad \text{as } n \rightarrow \infty,$$

from which one easily sees that  $\kappa_{Z-1}$  is regularly varying at zero<sup>11</sup> with index  $\alpha_T$ . Furthermore, setting  $t = \lambda^{-1/\alpha_T}$  in the above equation and taking inverses on both sides of the equation gives

$$\kappa_{Z-1}^{-1}(t) \sim \frac{t^{1/\alpha_T}}{\lambda^{1/\alpha_T} L(1/t)}, \quad \text{as } t \rightarrow \infty. \quad (6.4)$$

<sup>8</sup>We include the root here because it makes the formulae below simpler. It is clear that this will not affect the validity of Assumptions (T1) and (T2).

<sup>9</sup>We could of course assume w.l.o.g.  $\lambda = 1$ , but we keep the general form for convenience.

<sup>10</sup>The same formula is satisfied by the generating function of the hitting time  $\tau$  of  $-1$  of the left-continuous random walk  $S_n = Z_1 + \cdots + Z_n - n$ , see e.g. Spitzer (1976, p234). In fact, it is well-known that  $V$  and  $\tau$  are equal in law; this follows from an encoding of the Galton–Watson tree through its *Lukasiewicz path*, see e.g. Le Gall and Miermont (2012).

<sup>11</sup>This can also be obtained from the characterization of the domain of attraction of stable laws in terms of truncated second moments and a tail balance condition, see Feller (1971, Section XVII.5, Theorem 2), together with a Tauberian theorem, see Bingham et al. (1987, Theorem 8.1.6).

Equations (6.3) and (6.4) now readily yield<sup>12</sup>

$$1 - g_V(1 - s) \sim \frac{s^{1/\alpha_T}}{\lambda^{1/\alpha_T} L(1/s)}, \quad \text{as } s \rightarrow 0. \tag{6.5}$$

In particular, if  $\sigma^2 = \text{Var}(Z_1) < \infty$ , then setting  $\alpha_T = 2$ ,  $L \equiv \sigma$  and  $\lambda = 1/2$ , we have

$$1 - g_V(1 - s) \sim \frac{\sqrt{2}}{\sigma} \sqrt{s}, \quad \text{as } s \rightarrow 0. \tag{6.6}$$

We can now turn to the proof of Proposition 2.4.

*Proof of Proposition 2.4:* The first part follows directly from existing results in the literature: It is known<sup>13</sup> that conditioned on  $V = n$ , the random variable  $(a_n/n)H$  converges in law to a non-degenerate random variable, as  $n \rightarrow \infty$ . Since  $n/a_n = n^{1-1/\alpha_T+o(1)} = n^{1/D+o(1)}$ , this immediately proves the first part of the proposition.

We now turn to the second part of the proposition. Assumption (T1) follows directly from (6.5). It remains to check that (T2) holds for every  $\bar{\gamma} \in \mathbb{R}$ . This is in fact a direct consequence of the results in Kortchemski (2015) on the height of Galton–Watson trees conditioned on its number of vertices. However, since we only need weaker results, we present here for completeness a simple and transparent proof. We recall the following construction due to Geiger (1999) of the Galton–Watson tree conditioned on the event that its height is at least  $k \in \mathbb{N}$ . Define  $c_i = \mathbf{P}(H \geq i - 1)/\mathbf{P}(H \geq i)$ . Let  $(A_1, B_1), \dots, (A_k, B_k)$  be independent pairs of integer r.v. with law

$$\mathbf{P}(A_i = a, B_i = b) = c_i p_b \mathbf{P}(H < i - 1)^{a-1} \mathbb{1}_{(1 \leq a \leq b)}.$$

Then the tree  $T$  conditioned on  $H \geq k$  can be constructed as follows:

- There exists a ray of length  $k$ , called the *spine*, starting from the root.
- The vertex on the spine at generation/depth/height  $i$ ,  $i = 0, \dots, k - 1$ , has  $B_{k-i} - 1$  children off the spine, out of which, independently,
  - $A_{k-i} - 1$  children spawn copies of  $T$  conditioned on  $H < k - (i + 1)$ ,
  - $B_{k-i} - A_{k-i}$  children spawn copies of  $T$ .
- The vertex on the spine at generation  $k$  spawns a copy of  $T$ .

Ignoring in the above construction the  $A_{k-i} - 1$  copies of  $T$  conditioned on  $H < k - (i + 1)$  in each generation, we obtain that the number of vertices in the above tree is stochastically bounded from below by the sum of  $\sum_{i=1}^k B_i - A_i$  copies of  $V$ . Hence, if  $V^{(1)}, V^{(2)}, \dots$  denote iid copies of  $V$ , then

$$\mathbf{P}(V \leq n | H \geq k) \leq \mathbf{P}\left(\max_{j=1, \dots, \sum_{i=1}^k B_i - A_i} V^{(j)} \leq n\right) = g_{\sum_{i=1}^k B_i - A_i}(\mathbf{P}(V \leq n)), \tag{6.7}$$

<sup>12</sup>This formula can also be obtained by the so-called Kemperman’s formula for the hitting time of a left-continuous random walk on  $\mathbb{Z}$  (see Spitzer, 1976, p234 or Le Gall and Miermont, 2012) together with local limit theorems for random walks (Ibragimov and Linnik, 1971, Theorem 4.2.1), an explicit expression of the density at 0 of the  $\alpha_T$ -stable law (see e.g. Zolotarev, 1986, Section 2.2, Corollary 2) and a Tauberian theorem. This way, one obtains several multiplicative factors involving the Gamma function, the sine function and the constant  $\pi$  which mysteriously cancel by Euler’s reflection formula for the Gamma function. In our opinion, the approach presented here is more transparent.

<sup>13</sup>See Duquesne (2003, Theorem 3.1). The assumption of aperiodicity in that paper is not needed.

where for a random variable  $Y$  we denote by  $g_Y$  its generating function. Now let  $(A, B)$  be a pair of integer random variables such that

$$\mathbf{P}(A = a, B = b) = p_b \mathbb{1}_{(1 \leq a \leq b)},$$

so that  $B$  is distributed according to the size-biased distribution of  $\mathbf{p}$  and conditionally on  $B$ ,  $A$  is uniform in  $\{1, \dots, B\}$  (this law also appears in Geiger, 1999). For every  $s \in [0, 1]$ , we then have by the definition of  $(A_i, B_i)_{i=1, \dots, k}$ ,

$$\begin{aligned} g_{\sum_{i=1}^k B_i - A_i}(s) &= \prod_{i=1}^k \sum_{j=0}^{\infty} \mathbf{P}(B_i - A_i = j) s^j \\ &\leq \left( \prod_{i=1}^k c_i \right) \left( \sum_{j=0}^{\infty} \mathbf{P}(B - A = j) s^j \right)^k \\ &= \frac{1}{\mathbf{P}(H \geq k)} g_{B-A}(s)^k. \end{aligned} \tag{6.8}$$

We have,

$$\begin{aligned} g_{B-A}(s) &= \sum_{j=0}^{\infty} \mathbf{P}(B - A = j) s^j = \sum_{j=0}^{\infty} \left( \sum_{l=j+1}^{\infty} p_l \right) s^j = (1-s)^{-1} \sum_{l=0}^{\infty} p_l (1-s^l) \\ &= (1-s)^{-1} (1 - g_Z(s)). \end{aligned}$$

Now, since  $\kappa_{Z-1}(t)$  is regularly varying at zero with index  $\alpha_T$  (see above), we have

$$g_Z(s) = s + (1-s)^{\alpha_T + o(1)} = 1 - (1-s) + (1-s)^{\alpha_T + o(1)}, \quad \text{as } s \rightarrow 1.$$

Hence,

$$g_{B-A}(s) = 1 - (1-s)^{\alpha_T - 1 + o(1)}, \quad \text{as } s \rightarrow 1. \tag{6.9}$$

Equations (6.7), (6.8) and (6.9) now give for large  $n$ ,

$$\begin{aligned} \mathbf{P}(V \leq n, H \geq k) &\leq g_{B-A}(\mathbf{P}(V \leq n))^k = (1 - \mathbf{P}(V > n))^{\alpha_T - 1 + o(1)k} \\ &\leq \exp(-k \mathbf{P}(V > n)^{\alpha_T - 1 + o(1)}), \end{aligned}$$

such that by (6.5) and Fact A.2,

$$\mathbf{P}(V \leq n, H \geq k) \leq \exp(-kn^{-1/D + o(1)}), \quad \text{as } n \rightarrow \infty. \tag{6.10}$$

Now fix  $\varepsilon > 0$  and  $\rho > 1$ . We have with  $k = k(n) = n^{1/(1/D + \varepsilon)}$ ,

$$\begin{aligned} \mathbf{P}(H > V^{\frac{1}{D} + \varepsilon}, V \geq n) &\leq \mathbf{P}(H > V^{\frac{1}{D} + \varepsilon}, H \geq k) \\ &= \sum_{i=0}^{\infty} \mathbf{P}(H > V^{\frac{1}{D} + \varepsilon}, k^{\rho^i} \leq H < k^{\rho^{i+1}}) \\ &\leq \sum_{i=0}^{\infty} \mathbf{P}(V < k^{\rho^{i+1}/(1/D + \varepsilon)}, k^{\rho^i} \leq H < k^{\rho^{i+1}}) \\ &\leq \sum_{i=0}^{\infty} \mathbf{P}(V < (k^{\rho^i})^{\rho/(1/D + \varepsilon)}, k^{\rho^i} \leq H) \\ &\leq \sum_{i=0}^{\infty} \exp(-k^{\rho^i} k^{-\rho^i(\rho/(1/D + \varepsilon) + o(1))}) \end{aligned} \tag{by (6.10)}.$$

Choosing  $\rho < 1 + D\varepsilon$  yields for some  $\eta > 0$ ,

$$\mathbf{P}(H > V^{\frac{1}{D}+\varepsilon}, V \geq n) \leq \exp(-n^\eta), \quad \text{for large } n.$$

This proves (T2) for every  $\bar{\gamma} \in \mathbb{R}$ .  $\square$

## Appendix A. Regularly varying functions

We recall here known facts about regularly varying functions.

*Definition.* A function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to *vary regularly at  $\infty$  (at zero) with index  $\lambda \in \mathbb{R}$*  if for every  $y > 0$ ,

$$\frac{f(xy)}{f(x)} \rightarrow y^\lambda, \quad \text{as } x \rightarrow \infty \text{ (} x \rightarrow 0 \text{)}.$$

If  $\lambda = 0$ , the function is also called *slowly varying*. The definition is extended to sequences by linear interpolation (say).

It follows readily from the definition that the composition of regularly varying functions is again a regularly varying function (whose index is the product of the indices).

The following result is often used in this article, sometimes without mentioning it explicitly:

**Fact A.1** (Potter's bounds, [Bingham et al., 1987](#), Theorem 1.5.6). *If  $f$  is regularly varying of index  $\lambda$  at  $\infty$ , then for every  $C > 1$  and  $\delta > 0$  there exists  $x_0 = x_0(C, \delta)$ , such that*

$$\frac{f(y)}{f(x)} \leq C \max((y/x)^{\lambda+\delta}, (y/x)^{\lambda-\delta}), \quad \text{for all } x \geq x_0, y \geq x_0.$$

Regularly varying functions play an important role in Tauberian theorems, of which we will use the following form:

**Fact A.2** (Karamata's Tauberian theorem, extended form, [Bingham et al., 1987](#), Corollary 8.1.7). *Let  $X$  be a random variable taking values in  $[0, \infty)$  and define  $g_X(s) = \mathbf{E}[s^X]$ ,  $s \in [0, 1]$ . Let  $\beta \in [0, 1]$  and  $L$  be a slowly varying function at  $\infty$ . Then the following are equivalent:*

- $1 - g_X(s) \sim (1 - s)^\beta L(1/(1 - s))$ , as  $s \uparrow 1$ .
- If  $\beta \in [0, 1)$ ,

$$P(X > x) \sim \frac{L(x)}{x^\beta \Gamma(1 - \beta)}, \quad \text{as } x \rightarrow \infty.$$

*If  $\beta = 1$ , either  $\mathbf{E}[\min(X, x)] \sim L(x)$  or  $\mathbf{E}[X \mathbb{1}_{(X \leq x)}] \sim L(x)$ , as  $x \rightarrow \infty$  (in which case both occur).*

## Acknowledgements

I thank Nicolas Curien for pointing out reference [Janson and Marckert \(2005\)](#) and an anonymous referee for several useful suggestions that improved the presentation.

## References

- D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1), 248–289 (1993). [MR1207226](#).
- J. Bérard and P. Maillard. The limiting process of  $N$ -particle branching random walk with polynomial tails. *Electron. J. Probab.* **19**, no. 22, 17 (2014). [MR3167886](#).
- A. Bhattacharya, R. S. Hazra and P. Roy. Point process convergence for branching random walks with regularly varying steps (2016+). To appear in *Ann. de l'Inst. Henri Poincaré (B) Probabilités et Statistiques*.
- 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 (1987). ISBN 0-521-30787-2. [MR898871](#).
- Darren B. H. Cline and Tailen Hsing. Large deviation probabilities for sums of random variables with heavy or subexponential tails. Technical report, Statistics Department, Texas A& M University (1998). Available online at <http://www.stat.tamu.edu/~dcline/Papers/large5.pdf>.
- D. Denisov, A. B. Dieker and V. Shneer. Large deviations for random walks under subexponentiality: the big-jump domain. *Ann. Probab.* **36** (5), 1946–1991 (2008). [MR2440928](#).
- T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* **31** (2), 996–1027 (2003). [MR1964956](#).
- T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** (4), 553–603 (2005). [MR2147221](#).
- R. Durrett. Maxima of branching random walks vs. independent random walks. *Stochastic Process. Appl.* **9** (2), 117–135 (1979). [MR548832](#).
- R. Durrett. Maxima of branching random walks. *Z. Wahrsch. Verw. Gebiete* **62** (2), 165–170 (1983). [MR688983](#).
- P. Embrechts, C. Klüppelberg and T. Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin (1997). ISBN 3-540-60931-8. For insurance and finance. [MR1458613](#).
- W. Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney (1971). [MR0270403](#).
- N. Gantert. The maximum of a branching random walk with semiexponential increments. *Ann. Probab.* **28** (3), 1219–1229 (2000). [MR1797310](#).
- J. Geiger. Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Probab.* **36** (2), 301–309 (1999). [MR1724856](#).
- I. J. Good. The number of individuals in a cascade process. *Proc. Cambridge Philos. Soc.* **45**, 360–363 (1949). [MR0029489](#).
- C. C. Heyde. A contribution to the theory of large deviations for sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **7**, 303–308 (1967). [MR0216549](#).
- I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen (1971). With a supplementary chapter by I. A. Ibragimov and V. V. Petrov. Translation from the Russian edited by J. F. C. Kingman. [MR0322926](#).
- S. Janson and J.-F. Marckert. Convergence of discrete snakes. *J. Theoret. Probab.* **18** (3), 615–647 (2005). [MR2167644](#).



- H. Kesten. Branching random walk with a critical branching part. *J. Theoret. Probab.* **8** (4), 921–962 (1995). [MR1353560](#).
- I. Kortchemski. Limit theorems for conditioned non-generic Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2), 489–511 (2015). [MR3335012](#).
- S. P. Lalley and Y. Shao. Maximal displacement of critical branching symmetric stable processes (2016+). To appear in *Ann. de l'Inst. Henri Poincaré (B) Probabilités et Statistiques*.
- J.-F. Le Gall. A class of path-valued Markov processes and its applications to superprocesses. *Probab. Theory Related Fields* **95** (1), 25–46 (1993). [MR1207305](#).
- J.-F. Le Gall and G. Miermont. Scaling limits of random trees and planar maps. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 155–211. Amer. Math. Soc., Providence, RI (2012). [MR3025391](#).
- Ju. V. Linnik. Limit theorems for sums of independent variables, taking large deviations into account. II. *Teor. Veroyatnost. i Primenen.* **6**, 377–391 (1961a). [MR0150817](#).
- Ju. V. Linnik. Limit theorems for the sums of independent variables taking into account the large deviations. I. *Teor. Veroyatnost. i Primenen.* **6**, 145–163 (1961b). [MR0138125](#).
- J.-F. Marckert and A. Mokkadem. States spaces of the snake and its tour—convergence of the discrete snake. *J. Theoret. Probab.* **16** (4), 1015–1046 (2004) (2003). [MR2033196](#).
- T. Mikosch and A. V. Nagaev. Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1** (1), 81–110 (1998). [MR1652936](#).
- A. V. Nagaev. Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. I. *Teor. Veroyatnost. i Primenen.* **14**, 51–63 (1969a). [MR0247651](#).
- A. V. Nagaev. Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. II. *Teor. Veroyatnost. i Primenen.* **14**, 203–216 (1969b). [MR0247652](#).
- F. Spitzer. *Principles of random walk*. Springer-Verlag, New York-Heidelberg, second edition (1976). Graduate Texts in Mathematics, Vol. 34. [MR0388547](#).
- V. M. Zolotarev. *One-dimensional stable distributions*, volume 65 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI (1986). ISBN 0-8218-4519-5. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. [MR854867](#).