

# The extremal point process for branching random walk with stretched exponential displacements

Piotr Dyszewski<sup>1,a</sup>, Nina Gantert<sup>2,b</sup>

<sup>1</sup>*Instytut Matematyczny Uniwersytetu Wrocławskiego, Pl. Grunwaldzki 2/4 50-384, Wrocław, Poland, <sup>a</sup>pdysz@math.uni.wroc.pl*

<sup>2</sup>*Fakultät für Mathematik, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany, <sup>b</sup>gantert@ma.tum.de*

**Abstract.** We study a branching random walk in which the displacements are independent of the branching mechanism and follow a stretched exponential distribution with parameter  $r \in (0, 1)$ . We describe the positions of the particles near the rightmost particle using point process convergence. Our analysis reveals that the model undergoes infinitely many phase transitions at  $r = k/(k+1)$  for integers  $k \geq 2$ . As a consequence, we derive a new limit theorem for the position of the rightmost particle. Our approach is based on deriving precise large deviation results for sums of i.i.d. random variables with stretched exponential distributions, particularly outside the so-called "one big jump" regime.

*MSC2020 subject classifications:* Primary 60F10, 60J80

*Keywords:* branching random walk, limit theorem, point processes, stretched exponential distribution, precise large deviations

## 1. Introduction

Branching random walk (BRW) has been studied for more than forty years and during this time, due to its relation with the KPP equation, a lot of effort was made to understand the behaviour of its extremes. In the existing literature one can distinguish two different regimes depending on the laws of the displacements, however not many results address BRW outside of this two classical cases. The first one treats displacements with thin tails, i.e. possessing some exponential moments. In this case the asymptotic behaviour of the extremes of BRW is determined by the contribution coming from many, typical particles, see [2, 4, 12, 13, 15]. In the second regime one investigates regularly varying steps, in this case the asymptotic behaviour of the extremes of BRW is determined by a few, atypically big displacements, see [3, 9]. The contrast between these two cases raises the question about a regime in which both big jumps and typical displacements contribute to the asymptotics of the extremes of BRW. It turned out that in the case of stretched exponential displacements one can see a clear transition between the two aforementioned regimes [10]. We aim to give a complete picture of the extremes of BRW in this case.

In what follows we study BRW which is a branching process with a spatial component that can be described as follows. The process will evolve at time epochs  $n \in \mathbb{N} = \{0, 1, 2, \dots\}$ . Suppose that at time  $n = 0$  one particle is placed at the origin of the real line  $\mathbb{R}$ . At time  $n = 1$  the initial particle splits into a random number of new particles which are randomly displaced from their place of birth. We will always assume that all displacements are independent copies of a given random variable  $X$ . From now on, all particles independently evolve in the same way as the initial particle. Denote by  $\Lambda_n$  the point process obtained by putting a unit mass in the position of each particle present in the system at time  $n$ . Note that  $\Lambda_n$  is not necessarily simple, since two particles can occupy the same position. The sequence  $\{\Lambda_n\}_{n \in \mathbb{N}}$  is called a branching random walk. We refer to Section 2 for a more detailed set-up and to the recent monograph [21] for an introduction to the topic.

For a point measure  $\eta = \sum_i \delta_{x_i}$  and  $a, b \in \mathbb{R}$ ,  $a > 0$  denote the translation of  $\eta$  and scaling of  $\eta$  via

$$\mathfrak{T}_b \eta = \sum_i \delta_{x_i + b} \quad \text{and} \quad \mathfrak{S}_a \eta = \sum_i \delta_{ax_i}$$

respectively. The aim of the present article is a description of the asymptotic behaviour of the extremes of  $\{\Lambda_n\}_{n \in \mathbb{N}}$  by finding sequences  $\{a_n\}_{n \in \mathbb{N}}$  and  $\{b_n\}_{n \in \mathbb{N}}$  for which  $\mathfrak{S}_{a_n^{-1}} \mathfrak{T}_{-b_n} \Lambda_n$  has a non-trivial limit in distribution in the sense of point process convergence, which we will make precise in Section 2. Such a choice is not unique by any means. We will pick the pair  $a_n$  and  $b_n$  which allows to describe the position of the rightmost particle. In other words, we will work with sequences  $a_n$  and  $b_n$  for which the limiting configuration of  $\mathfrak{S}_{a_n^{-1}} \mathfrak{T}_{-b_n} \Lambda_n$  does not contain a sequence diverging to infinity. Under some mild moment conditions on the number of offspring of the initial particle, the asymptotic behaviour of  $\{\Lambda_n\}_{n \in \mathbb{N}}$  is determined mainly by the law of the displacements, i.e. the law of  $X$ . In the classical case one works assuming Cramér's condition

$$(1) \quad \mathbb{E} \left[ e^{s|X|} \right] < \infty \quad \text{for some } s > 0.$$

Under condition (1) the contribution of a single displacement is negligible and the position of the rightmost particle is concentrated around its mean which moves at a linear speed. One can show that there are constants  $C_1, C_2$  such that

$$\mathfrak{T}_{-C_1 n - C_2 \log n} \Lambda_n$$

converges to a non-trivial limit, see [15]. If the condition (1) is not satisfied, some additional scaling is necessary. For example, consider the case of regularly varying tails, i.e. a random variable  $X$  such that as  $t$  tends to infinity

$$(2) \quad \mathbb{P}[X > t] \sim C \cdot t^{-\gamma},$$

for some constants  $\gamma, C > 0$ , see (3) at the end of this section for the notation " $\sim$ ". The regularly varying distributions follow the so called "one big jump principle" (see (9)). Using this fact one can observe a different limiting behaviour of  $\Lambda_n$ . Namely, if one denotes by  $m > 1$  the expected number of children of the initial particle, one can show that

$$\mathfrak{S}_{m^{-n/\gamma}} \Lambda_n$$

converges in law to a non-trivial limit, see [3]. The scaling present above is necessary to compensate the contribution of atypical, huge displacements which, in contrast to the previous case, is not negligible.

We will investigate an intermediate case between (1) and (2), namely stretched exponential, or Weibull, displacements with tails of the form

$$\mathbb{P}[X > t] \sim e^{-t^r \ell(t)},$$

for  $r \in (0, 1)$  and a slowly varying function  $\ell$ . Recall that a function  $\ell : \mathbb{R} \rightarrow (0, \infty)$  is slowly varying (at infinity) if for any constant  $C > 0$ ,  $\ell(Ct) \sim \ell(t)$  as  $t \rightarrow \infty$ . We will show in our main result that

$$\mathfrak{S}_{n^{1-1/r} \ell_2(n)} \mathfrak{T}_{-n^{1/r} \ell_1(n)} \Lambda_n$$

converges for an appropriate choice of slowly varying functions  $\ell_1$  and  $\ell_2$ . Moreover we will provide a detailed asymptotic expansion of  $n^{1/r} \ell_1(n)$ . As a consequence one gets a new limit theorem for  $M_n$ , the position of the rightmost particle in generation  $n$ . Namely, we will infer that

$$n^{1-1/r} \ell_2(n) \left( M_n - n^{1/r} \ell_1(n) \right)$$

converges in distribution to a random shift of the Gumbel law. Such results were obtained recently for constant  $\ell$  in the case  $r \leq 2/3$ , see [10]. Our proof relies on precise large deviation estimates for sums of i.i.d. random variables with stretched exponential tails. Since the corresponding results established in the sixties [17, 18] were too implicit for our needs, we revisit the problem and provide a more explicit description of the large deviation probabilities on the polynomial scale.

The paper is organized as follows. In Section 2 we give a precise description of the model and our assumptions. Our main results, Theorem 3.1 and Corollary 3.2, are presented in Section 3. In Sections 4 and 5 we present the main steps of our arguments, namely trimming and decoupling of the random weighted tree. Lastly, in Section 6 we give the final arguments of our proof and in Section 7 we prove precise large deviations for sums of stretched exponential random variables, which may be of independent interest.

Throughout the article we write  $g(t) \sim f(t)$  for two functions  $f, g: \mathbb{R} \rightarrow [0, \infty)$  whenever

$$(3) \quad \lim_{t \rightarrow \infty} f(t)/g(t) = 1$$

and we write  $f(t) \gg g(t)$  if  $g(t) = o(f(t))$ , i.e. if  $f(t)/g(t) \rightarrow 0$ . We will denote by  $C$  positive constants whose value is not important for us. Note that the value of "C" may change from line to line.

## 2. Preliminaries

Let us now give a precise definition of our model. Since we will assume that the displacements and the reproduction mechanism are independent, we will introduce both of these components separately. If we denote by  $Z_n = \Lambda_n(\mathbb{R})$  the number of particles present in the system at time  $n$ , then the sequence  $Z = \{Z_n\}_{n \in \mathbb{N}}$  forms a Galton-Watson process starting from  $Z_0 = 1$ , i.e. with one particle, and a reproduction law with mean

$$m = \mathbb{E}[Z_1].$$

**Assumption 1.** We have  $\mathbb{P}[Z_1 = 1] < 1$  and  $\mathbb{E}[Z_1 \log^+ Z_1] < \infty$  and the Galton-Watson process  $Z$  is supercritical, that is  $m > 1$ .

It is well known from the classical literature, see [1, Theorem I.A.5.1], that if  $\mathbb{P}[Z_1 = 1]$  is strictly less than 1,  $Z$  survives with positive probability if and only if  $m > 1$ . Since our objects become trivial on the set of extinction of the underlying Galton-Watson process, one can work under the conditional probability

$$\mathbb{P}^*[\cdot] = \mathbb{P}[\cdot \mid \forall n \in \mathbb{N}, Z_n > 0].$$

In fact one can recast our main results in terms of weak convergence under the probability measure  $\mathbb{P}^*$ . We will however, for notational simplicity, continue to work under  $\mathbb{P}$ . Note that by the martingale convergence theorem there is a random variable  $W$  such that  $\mathbb{P}$ -almost surely,

$$(4) \quad \lim_{n \rightarrow \infty} m^{-n} Z_n = W.$$

If Assumption 1 is satisfied then the convergence holds also in  $L^1$  and moreover  $\mathbb{P}^*[W > 0] = 1$ , see for example [21, Chapter 2]. This means that

$$\mathbb{P}[W = 0] = \mathbb{P}[\exists n \in \mathbb{N} : Z_n = 0].$$

The random variable  $W$  will play a crucial role in our results: it will be a shift parameter in the limit of  $\Lambda_n$ .

We will now describe the branching structure in more detail and introduce the displacements. Denote by  $\mathbb{T} \subseteq \mathbb{U} = \{\emptyset\} \cup \bigcup_{n \geq 1} \mathbb{N}^n$  the Galton-Watson tree corresponding to  $Z$  with Ulam-Harris labelling. That is, label the initial particle with  $\emptyset$  and its children with  $\{1, 2, \dots, Z_1\} \subset \mathbb{N}$  and for each particle with a label  $w \in \mathbb{T}$ , label its children with labels of the form  $(w, 1), (w, 2), \dots, (w, N(w))$ , where  $N(w)$  is the random variable denoting the number of children of the individual labelled with  $w$ . Write  $|w| = n$  if  $w$  is a label of a particle from generation  $w$ , in other words if  $w \in \mathbb{N}^n$ . Denote by  $(\emptyset, w]$  the unique vertex path from  $\emptyset$  to  $w$  (excluding  $\emptyset$  and including  $w$ ) and write  $w \wedge z$  to denote the last common ancestor of  $w$  and  $z$ , i.e. the unique element of  $\mathbb{T}$  for which  $(\emptyset, w] \cap (\emptyset, z] = (\emptyset, w \wedge z]$ . Finally, write  $w \leq z$  whenever  $w$  is an ancestor of  $z$ , that is  $w \in (\emptyset, z]$  and  $\mathbb{T}_n$  for the particles from the  $n$ th generation present in the system, i.e.  $\mathbb{T}_n = \{w \in \mathbb{T} : |w| = n\}$ .

To model the displacements consider a family  $\{X_w : w \in \mathbb{U}\}$  of independent copies of  $X$  which are independent of  $Z$ . The random variable  $X_w$  represents the displacement of the particle with label  $w$  from its place of birth. The position of the particle with label  $w$  is therefore represented via

$$V(w) = \sum_{z \in (\emptyset, w]} X_z.$$

To study the collection  $\{V(w)\}_{w \in \mathbb{T}_n}$  define a sequence of point processes by

$$\Lambda_n = \sum_{|w|=n} \delta_{V(w)}.$$

We call the sequence  $\{\Lambda_n\}_{n \in \mathbb{N}}$  a branching random walk. We will regard  $\Lambda_n$  as a point process, that is a random element of  $\mathcal{M}_p = \mathcal{M}_p((-\infty, \infty])$ , the space of point measures on  $(-\infty, \infty]$  equipped with the topology of vague convergence. Here  $(-\infty, +\infty]$  denotes the homomorphic image of  $(0, 1]$  where closed neighbourhoods of  $+\infty$  are compact. Recall that  $\eta$  is a point measure, i.e. an element of  $\mathcal{M}_p$ , if it can be written in the form  $\eta = \sum_i \delta_{x_i}$  for some sequence  $\{x_i\}_i \subset (-\infty, \infty]$  without an accumulation point. We refer to [19, Chapter 3.1] for a more in-depth description of  $\mathcal{M}_p$ . We equip  $\mathcal{M}_p$  with the topology of vague convergence, and say that  $\eta_n \rightarrow \eta$  in  $\mathcal{M}_p$  (or vaguely) if for any function  $f$  from the set  $\mathcal{C}_K^+ = \mathcal{C}_K^+((-\infty, \infty])$  of nonnegative, continuous functions with compact support on  $(-\infty, \infty]$ ,  $\int f(x) \eta_n(dx) \rightarrow \int f(x) \eta(dx)$ . In our results we use weak convergence of random elements of  $\mathcal{M}_p$  (denoted by  $\Rightarrow$ ), which is equivalent to the following condition, see [19, Proposition 3.19]. For point processes  $\Theta_n, \Theta$  we have

$$\Theta_n \Rightarrow \Theta \quad \text{in } \mathcal{M}_p,$$

if and only if for any  $f \in \mathcal{C}_K^+$ ,

$$\int f(x) \Theta_n(dx) \xrightarrow{d} \int f(x) \Theta(dx),$$

where  $\xrightarrow{d}$  denotes the convergence in law of random variables. Our goal is to study weak convergence of

$$\mathfrak{S}_{a_n^{-1} \mathfrak{T}_{-b_n}} \Lambda_n = \sum_{|w|=n} \delta_{a_n^{-1}(V(w)-b_n)}$$

in  $\mathcal{M}_p$ . Aiming to find a correct choice of  $b_n$ 's and  $a_n$ 's we first note that the convergence of  $\mathfrak{S}_{a_n^{-1} \mathfrak{T}_{-b_n}} \Lambda_n$  in  $\mathcal{M}_p$  describes the behaviour of  $\Lambda_n$  near its right frontier, that is the location of the rightmost particle, which we can write as

$$(5) \quad M_n = \max_{|w|=n} V(w)$$

with the convention that the maximum of an empty set is  $-\infty$ . The first order of magnitude of  $M_n$  can be predicted using the first moment method. Take  $X_1, X_2, \dots$  to be i.i.d. copies of  $X$  and let  $S_n = \sum_{k=1}^n X_k$  denote the corresponding random walk. For a sequence of real numbers  $\{x_n\}_{n \in \mathbb{N}}$  write

$$(6) \quad \mathbb{P}[M_n > x_n] = \mathbb{P}[\exists w \in \mathbb{T}_n, V(w) > x_n] \leq \mathbb{E}[Z_n] \mathbb{P}[S_n > x_n] = m^n \mathbb{P}[S_n > x_n].$$

For the estimate (6) to give any non-trivial information about the asymptotics of  $M_n$ , one needs to choose the sequence  $\{x_n\}_{n \in \mathbb{N}}$  such that the probabilities  $\mathbb{P}[S_n > x_n]$  decay as  $m^{-n}$ , that is exponentially fast. Here large deviation estimates for the random walk  $\{S_n\}_{n \in \mathbb{N}}$  come into play.

In the classical case, under Cramér's condition (1), it is well known that if one takes the sequence of the form  $x_n = \rho n$  for  $\rho > \mathbb{E}[X]$ , by the Bahadur-Ranga-Rao theorem, see [7, Theorem 3.7.4],

$$(7) \quad \mathbb{P}[S_n > \rho n] \sim C(\rho) n^{-1/2} e^{-\psi^*(\rho)n},$$

where  $\psi^*(\rho) = \sup_s [s\rho - \psi(s)]$ , with  $\psi(t) = \log \mathbb{E}[e^{tX}]$  and a constant  $C(\rho)$  depending on  $\rho$ . Provided that  $\rho < \sup\{\psi'(s) : \psi(s) < \infty\}$ , under some mild regularity assumptions, there exists  $s_\rho$  such that  $\psi'(s_\rho) = \rho$  and then  $\psi^*(\rho) = \rho s_\rho - \psi(s_\rho)$ . Now, our estimate for  $\mathbb{P}[M_n > x_n]$  suggests that one should find a constant  $\rho$  such that  $\psi^*(\rho) = \rho s_\rho - \psi(s_\rho) = \log m$ . Then one could expect that, after establishing the lower bound corresponding to (6),

$$\lim_{n \rightarrow \infty} M_n/n = C_1 = \rho \quad \mathbb{P}^* \text{- a.s.}$$

This is indeed true and dates back to [4, 12, 13]. Since the seventies a lot of effort was devoted to a precise description of the asymptotics of  $M_n$  with the ultimate result stating that, with  $C_2 = 3/(2s_\rho)$ ,

$$M_n - C_1 n - C_2 \log n$$

converges in distribution to a random shift of the Gumbel law, see [2]. One can in fact argue that in this case the maximal particle at time  $n$  is most likely to be  $O(\sqrt{n})$  below the leading particle up to that point (see Figure 1). This is the scale to study the extremes of BRW. Indeed the point process

$$(8) \quad \mathfrak{T}_{-C_1 n - C_2 \log n} \Lambda_n$$

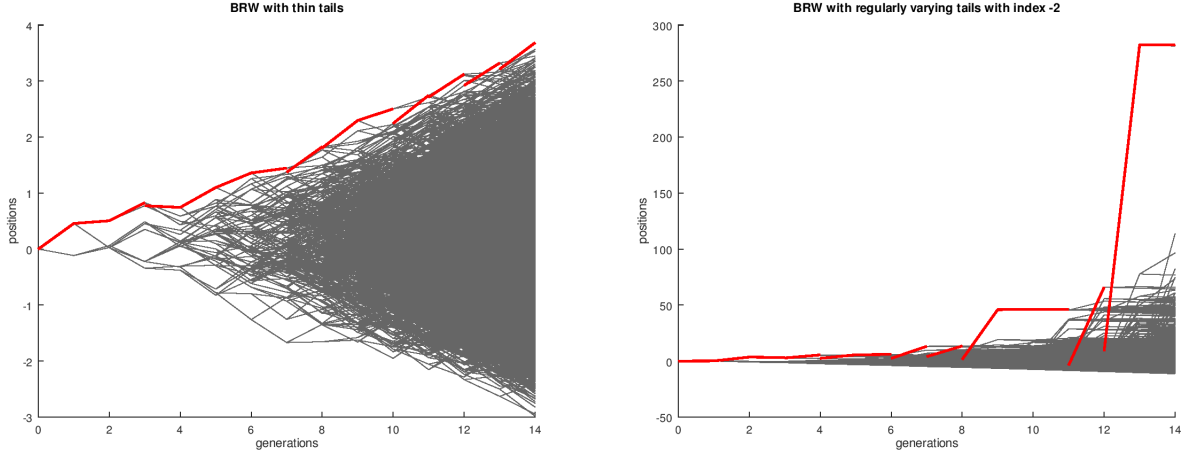


FIG 1. BRW with steps distributed uniformly on the interval  $(-1/2, 1/2)$  (left) and regularly varying steps with index  $\gamma = 2$  (right). The red lines indicate the displacements of the rightmost particles from its place of birth.

has a weak limit in  $\mathcal{M}_p$ , see [15] for details. It turns out that the limiting measure of (8) is a randomly shifted decorated Poisson process [23].

**Definition 2.1.**  $\Theta$  is a decorated Poisson point process of intensity  $v$  and decoration  $D$  if  $\Theta$  is distributed as  $\sum_{i \geq 1} \mathfrak{X}_{\zeta_i} D_i$  where  $\sum_{i \geq 1} \delta_{\zeta_i}$  is a Poisson point process with intensity  $v$  and  $D_i$  are i.i.d. copies of a point process  $D$ .

**Definition 2.2.**  $\Theta$  is a randomly shifted decorated Poisson point process of intensity  $v$ , decoration  $D$  and shift  $S$  if  $\Theta$  is distributed as  $\sum_{i \geq 1} \mathfrak{X}_{\zeta_i + S} D_i$  where  $\sum_{i \geq 1} \delta_{\zeta_i}$  is a Poisson point process with intensity  $v$  and  $D_i$  are i.i.d. copies of a point process  $D$ . One then writes  $\Theta \sim \text{SDPPP}(v, D, S)$ .

The weak limit of (8) is  $\text{SDPPP}(C_3 e^{-x} dx, N, -\log(D_\infty))$  for some positive  $C_3$ , some point process  $N$  and  $D_\infty$  denoting the a.s. limit of the so-called derivative martingale associated with the BRW [21, Chapter 5].

The behaviour of  $M_n$  under (2) is very different. In this case the law of  $X$  obeys the "one big jump principle", that is

$$(9) \quad \mathbb{P}[S_n > x_n] \sim n \mathbb{P}[X > x_n]$$

provided that  $x_n$  is large enough (for example  $x_n \gg \sqrt{n \log n}$  if  $\gamma > 2$ ). Going back to the estimate (6) one can expect, taking into account (9) and considering  $x_n = m^{n/\gamma}$ , that

$$m^{-n/\gamma} M_n$$

converges in law. Indeed, this sequence converges in law to a random shift of the Fréchet distribution, see [9]. In this case the rightmost particles are most likely to originate from the bulk of the population, i.e. the exponential number of typically placed particles (see Figure 1). Similarly as before, one has convergence of the associated point process and one can show that after an appropriate scaling  $\Lambda_n$  converges to so-called Cox cluster process, see [3]. Due to the similarity with our main result we now describe the limiting point process in detail. Assume for simplicity that the displacements are positive and let  $\{T_l\}_{l \in \mathbb{N}}$  be a collection of i.i.d. random variables with common distribution given by

$$(10) \quad \mathbb{P}[T_1 = k] = \frac{1}{v} \sum_{i=0}^{\infty} m^{-i} \mathbb{P}[Z_i = k], \quad k = 1, 2, \dots,$$

where

$$(11) \quad v = \sum_{i=0}^{\infty} m^{-i} \mathbb{P}[Z_i > 0].$$

Next let  $\{g_l\}_{l \in \mathbb{N}}$  denote the points of a Poisson point process with intensity  $\gamma x^{-\gamma-1} \mathbb{1}_{(0, \infty)}(x) dx$ . Then

$$\mathfrak{S}_{m^{-n/\gamma} \Lambda_n} \Rightarrow \sum_{l=0}^{\infty} T_l \delta_{(vW)^{1/\gamma} g_l},$$

in  $\mathcal{M}_p$ , where  $W$  is given by (4). The coefficients  $\{T_l\}_{l \in \mathbb{N}}$  represent the number of descendants of particles that made a big jump. As we will see, the limiting point process in the case of stretched exponential displacements has a very similar representation.

### 3. Main result

We turn our attention to a case between (1) and (2), namely the case where the displacements have a stretched exponential tail. We will impose some regularity condition on the regularly varying function  $R(x) = x^r \ell(x)$  that appears in the tail of the displacement law. One can usually circumvent this kind of restrictions using the smooth variation theorem [5, Theorem 1.8.2]. However this is not sufficient for our purposes since we would need an approximation up to a certain order. We will thus impose sufficient smoothness on  $R$  from the beginning. From now on we will assume that the law of the displacements is given as follows.

**Assumption 2.** *The random variable  $X$  is centred ( $\mathbb{E}[X] = 0$ ), has variance 1 ( $\mathbb{E}[X^2] = 1$ ) and has a stretched exponential upper tail with index  $r \in (0, 1)$ , that is for some positive  $x_0$ ,*

$$\mathbb{P}[X > x] = a(x) e^{-R(x)}, \quad x > x_0,$$

where  $a, R: (x_0, +\infty) \rightarrow [0, \infty)$ ,  $a(x) \rightarrow a > 0$  as  $x \rightarrow \infty$  and  $R$  is a  $\mathcal{C}^1(x_0, +\infty)$  function such that  $R'$  is regularly varying with index  $r - 1$  and differentiable with  $R''(x) = O(x^{r-2+\epsilon})$  for any  $\epsilon > 0$ . Moreover we assume that  $R(x)x^{-d}$  is eventually decreasing for some  $d \in (0, 1)$  and if  $r > 2/3$  we assume additionally that  $R(x)x^{-1/2-\epsilon}$  is eventually increasing and  $R(x)x^{-1+\epsilon}$  is eventually decreasing for some  $\epsilon > 0$ . Further, suppose that

$$\mathbb{E}[|X|^k] < \infty$$

for all  $k \leq \lfloor (2-r)/(1-r) \rfloor$ .

We note right away that since  $a(\cdot)$  is assumed to be bounded away from zero, the condition regarding regularity of  $R$  excludes discrete random variables, i.e. taking values in a closed, countable set. Recall that  $R'$  is regularly varying at infinity with index  $r - 1$ , if it can be written in the form

$$R'(x) = x^{r-1} \ell_{R'}(x), \quad x > x_0$$

for some slowly varying function  $\ell_{R'}$ . By Karamata's Theorem [19, Theorem 0.6] this implies that  $R$  is also regularly varying with index  $r$ , i.e. it is of the form

$$R(x) = x^r \ell_R(x), \quad x > x_0$$

for some slowly varying function  $\ell_R$ . There is a natural relation between  $\ell_R$  and  $\ell_{R'}$ , but it will be of little importance for us and thus we refrain from stating it explicitly. Note that under Assumption 2,  $R''(x)/R'(x)^2 \rightarrow 0$  as  $x \rightarrow \infty$  and thus  $e^{-R(x)}$  is a Von Mises function with an auxiliary function  $1/R'(x)$  [19, Proposition 1.1]. This further implies that under Assumption 2 the distribution of  $X$  lies in the max-domain of attraction of the Gumbel law [19, Proposition 1.4]. We will use this fact on the exponential scale. For sufficiently large  $n$  we put

$$(12) \quad d_n = \inf\{x \in (x_0, \infty) : R(x) > n \log m\}.$$

Note that by continuity of  $R$ , for sufficiently large  $n$ ,  $d_n$  is the smallest solution to  $R(d_n) = n \log m$ . The sequence  $\{d_n\}_{n \in \mathbb{N}}$  is regularly varying with index  $1/r$ , i.e.

$$d_n = n^{1/r} \ell_d(n)$$

for some slowly varying function  $\ell_d$  [19, Proposition 0.8]. Let also

$$(13) \quad a_n = 1/R'(d_n),$$

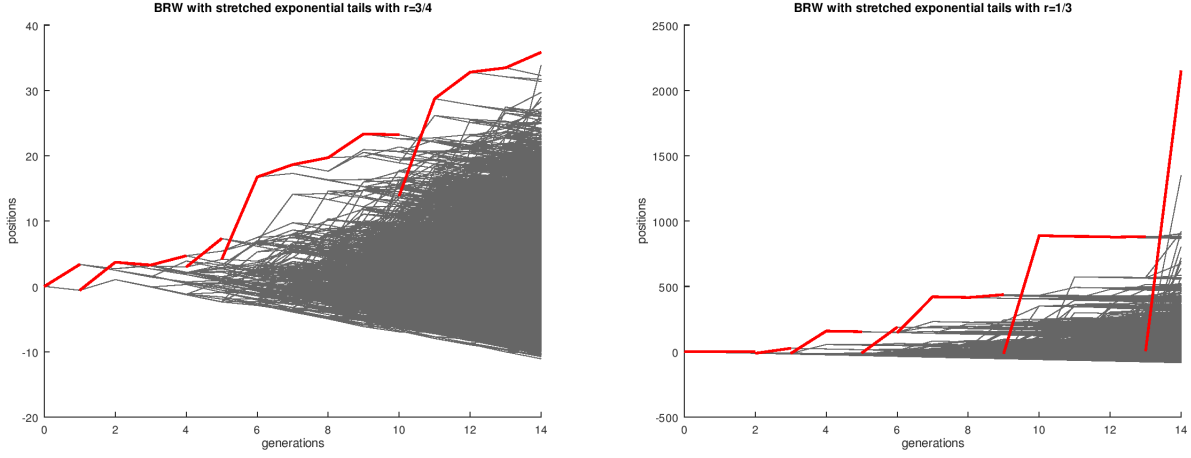


FIG 2. BRW with stretched exponential steps with  $r = 3/4$  (left) and  $r = 1/3$  (right). The red lines indicate the displacements of the rightmost particles from its place of birth.

for  $n$  large enough. Since a composition of two regularly varying function is regularly varying,  $a_n$  admits the representation

$$(14) \quad a_n = n^{1/r-1} \ell_a(n)$$

for some slowly varying function  $\ell_a$ . Extreme value theory asserts that, under Assumption 2, the maximum of  $m^n$  independent copies of  $X$  shifted by  $d_n$  and scaled by  $a_n$  converges in distribution to the Gumbel law, that is for any  $x \in \mathbb{R}$ , as  $n \rightarrow \infty$ ,

$$\mathbb{P}[X_k \leq d_n + a_n x, k \leq m^n] \rightarrow e^{-ae^{-x}}.$$

Note that here one uses the fact that  $R''(x)/(R'(x))^2 \rightarrow 0$  as  $x \rightarrow \infty$ . As we will soon see, the sequences  $\{d_n\}_{n \in \mathbb{N}}$  and  $\{a_n\}_{n \in \mathbb{N}}$  will drive the centring and scaling for  $\Lambda_n$ . In fact the latter sequence is the scaling for  $\Lambda_n$ .

Large deviations for random walks under Assumption 2 date back to the late sixties, see [14, 17, 18, 20] or [8] for a general set-up. As it turns out (9) is satisfied also in this case provided that  $x_n \gg n^{1/(2-2r)+\varepsilon}$  for some  $\varepsilon > 0$ . Such deviations are said to lie in the one big jump domain. Once again the estimate (6) suggests that in this case  $M_n$  should be of order  $d_n = n^{1/r} \ell_d(n)$ . This causes some technical problems since  $d_n \gg n^{1/(2-2r)+\varepsilon}$  if and only if  $r < 2/3$ . Indeed for  $r \geq 2/3$  and  $x_n$  of the order  $d_n$ , the asymptotic of  $\mathbb{P}[S_n > x_n]$  differs from the one of  $n\mathbb{P}[X > x_n]$ . However the latter still provides the leading term in the asymptotic of the former. More precisely, for any  $r \in (0, 1)$ ,

$$(15) \quad \log \mathbb{P}[S_n > d_n] = -(\log m)n(1 + o(1)),$$

see [11, Theorem 3]. In the sequel we will revisit this problem to provide a new description of the classical results [17] on the precise asymptotic of  $\mathbb{P}[S_n > x_n]$  with  $x_n \sim d_n$ . The logarithmic asymptotics (15) suffices to predict the law of large numbers for  $M_n$ . Indeed, as shown in [11] we have

$$\lim_{n \rightarrow \infty} M_n/d_n = 1 \quad \mathbb{P}^* - \text{a.s.}$$

The behaviour of  $M_n - d_n$  was recently studied in [10] for the special case  $R(x) = \lambda x^r$  for some  $\lambda > 0$ . In this case

$$d_n = ((\log m)/\lambda)^{1/r} n^{1/r} \quad \text{and} \quad a_n = r((\log m)/\lambda)^{1/r-1} n^{1/r-1}.$$

It transpires that for  $r \leq 2/3$  the rightmost particle is most likely to originate from the bulk of the population (see Figure 2) and for  $r > 2/3$  it typically originates from a subexponential number of particles that deviate from the bulk (located around  $O(n^{2-1/r})$ ), see Figure 2. This in turn leads to a more balanced behaviour of  $M_n$  for big values of  $r \in (0, 1)$ . More precisely, for this choice of  $R$  we have, for  $r > 2/3$ ,

$$\lim_{n \rightarrow \infty} \frac{M_n - ((\log m)/\lambda)^{1/r} n^{1/r}}{n^{2-1/r}} = \frac{r\lambda^{1/r} (\log m)^{1-1/r}}{2} \quad \mathbb{P}^* - \text{a.s.}$$

and for  $r \leq 2/3$ , with respect to  $\mathbb{P}^*$ ,

$$\frac{M_n - ((\log m)/\lambda)^{1/r} n^{1/r}}{n^{1/r-1}}$$

converges in distribution to a random shift of the Gumbel law [10, Theorem 3.1]. Here, we extend these results to distributions satisfying Assumption 2 and prove point process convergence. The limit object of  $\Lambda_n$  will be described with the help of an auxiliary polynomial. However it comes into play only in the case  $r \geq 2/3$ . Let the sequence of cumulants of  $X$ ,  $\{k_j\}_{j \leq \kappa}$  be defined recursively via  $k_1 = 0$  and for  $j \geq 1$ ,

$$(16) \quad k_j = \mathbb{E}[X^j] - \sum_{i=1}^{j-1} \binom{j-1}{i} k_{j-i} \mathbb{E}[X^i].$$

Define the truncated cumulant generating function via

$$(17) \quad K(x) = \sum_{j=2}^{\kappa} \frac{k_j}{j!} x^j, \quad \kappa = \left\lfloor \frac{2-r}{1-r} \right\rfloor.$$

The name truncated cumulant generating function will become clear after we state Lemma 4.1 below. For  $n \in \mathbb{N}$  consider a function  $\Psi_n$  given for  $z \in \mathbb{R}$  and sufficiently large  $n$  via

$$\Psi_n(z) = \inf_{s \in [0,1]} \{R(d_n + z - nK'(s)) + n(sK'(s) - K(s))\}.$$

One can then check that  $\Psi_n$  is a nondecreasing and continuous function such that for any  $n \in \mathbb{N}$ ,  $\Psi_n(z) \rightarrow \infty$  as  $z \rightarrow \infty$ . Finally, since  $\Psi_n(0) \leq R(d_n)$ , for sufficiently large  $n \in \mathbb{N}$  we can define  $\tau_n$  to be the smallest real number solving

$$(18) \quad \Psi_n(\tau_n) = R(d_n).$$

By a direct calculation using the regular variation of  $R$  and  $R'$ , one gets that for some slowly varying function  $\ell$ ,

$$\tau_n = nR'(d_n)/2 + O(n^{3-2/r}\ell(n)).$$

It follows that  $\tau_n$  is regularly varying with index  $2 - 1/r$  and thus, recalling (14),  $\tau_n = o(a_n)$  if  $r < 2/3$ . In fact further examination of  $\tau_n$  yields that

$$\tau_n = \sum_{j=1}^{\kappa_1} n^{1+(1-1/r)j} \ell_{\tau,j}(n) + o(a_n), \quad \kappa_1 = \left\lfloor \frac{2r-1}{1-r} \right\rfloor$$

for some slowly varying functions  $\ell_{\tau,j}$ . It transpires that there is a phase transition in the behaviour of  $\Lambda_n$  at  $r$  of the form  $k/(k+1)$  for some integer  $k \geq 2$ . We will now describe the limiting measure. Let  $\{T_k\}_{k \in \mathbb{N}}$  be i.i.d. with distribution given in (10) and take  $\{\iota_k\}_{k \in \mathbb{N}}$  to be the points of a Poisson point process with intensity  $ae^{-x} dx$ . Define a random element of  $\mathcal{M}_p$  by

$$(19) \quad \Lambda = \mathbf{1}_{\{W > 0\}} \sum_{k=0}^{\infty} T_k \delta_{\iota_k - \log(vW)},$$

recalling (11). Then  $\Lambda$  is a Cox cluster process with the Laplace functional given via

$$\mathbb{E} \left[ e^{-\int f(x) \Lambda(dx)} \right] = \mathbb{E} \left[ \exp \left\{ -aW \sum_{j=0}^{\infty} m^{-j} \int_{\mathbb{R}} \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] e^{-x} dx \right\} \right]$$

for  $f \in \mathcal{C}_K^+$ . Note that even though  $W$  is the limiting random variable of  $Z_j/m^j$ , the  $Z_j$ 's present in the above formula can be taken independent from  $W$  due to the presence of the expectation in the exponent. If we denote by  $z: [0,1] \rightarrow [0,1]$  the probability generating function of the underlying reproduction law

$$z(s) = \mathbb{E} [s^{Z_1}]$$

then a standard application of the branching property yields that the probability generating function of  $Z_n$  is an  $n$ -fold composition

$$\mathbb{E} [s^{Z_n}] = z^{(\circ n)}(s) = z \circ z \circ \dots \circ z(s), \quad s \in [0, 1].$$

The formula for the Laplace functional of  $\Lambda$  can be thus recast as

$$\mathbb{E} \left[ e^{-\int f(x) \Lambda(dx)} \right] = \mathbb{E} \left[ \exp \left\{ -aW \sum_{j=0}^{\infty} m^{-j} \int_{\mathbb{R}} \left( 1 - z^{(\circ j)} \left( e^{-f(x)} \right) \right) e^{-x} dx \right\} \right].$$

**Theorem 3.1.** *Suppose that Assumptions 1 and 2 are in force. Let  $d_n$  and  $a_n$  be given by (12) and (13) respectively and take  $b_n = d_n + \tau_n$ , where  $\tau_n$  is given in (18). Then*

$$\mathfrak{S}_{a_n^{-1}} \mathfrak{T}_{-b_n} \Lambda_n \Rightarrow \Lambda \quad \text{in } \mathcal{M}_p,$$

with  $\Lambda$  given by (19).

The strategy of the proof is as follows. First, one needs to identify the candidates for the extremal particles. We establish these in Lemmas 4.2, 4.3, 4.4, and 7.4. These are the particles that deviate moderately from the majority of the population around  $nK'(s)$  for some  $s \in [0, 1]$  at time  $n - o(n)$  and then make a jump of order  $d_n + \tau_n - nK'(s)$  in the time interval  $[n - o(n), n]$ . Optimizing over  $s \in [0, 1]$  yields that the particle resides in the neighborhood of  $d_n + \tau_n$  with probability at least  $e^{-\Psi(\tau_n)}$ . The choice of  $\tau_n$  then guarantees the existence of such particles, since we have  $e^{R(d_n)}$  particles. Next, one needs to argue that the particles around  $d_n + \tau_n$  are asymptotically independent (Lemma 5.2). Observing that the positions of these particles fluctuate as the largest jump (i.e., the one of the order  $d_n + \tau_n - nK'(s)$ ), the final steps of the argument reduce to counting the offspring of the particle that made a big jump and using classical extreme value theory. The penultimate step results in the weights  $T_k$  in the definition of the limiting point process.

Note that for any  $x \in \mathbb{R}$ ,  $\mathbb{P}[M_n \leq x] = \mathbb{P}[\Lambda_n(x, +\infty) = \emptyset]$ . As a by-product of Theorem 3.1 we obtain a new result for  $M_n$  which complements the results of [10, Theorem 3.1].

**Corollary 3.2.** *Let Assumptions 1 and 2 hold. Then for  $M_n = \max_{|w|=n} V(w)$  and  $b_n$  and  $a_n$  as in the statement of Theorem 3.1,*

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} V$$

with respect to  $\mathbb{P}$ , where the random variable  $V \in [-\infty, \infty)$  is distributed according to

$$\mathbb{P}[V \leq x] = \mathbb{E} \left[ \exp \left\{ -avW e^{-x} \right\} \right].$$

In particular if  $2/3 < r < 3/4$  then

$$\frac{M_n - d_n - nR'(d_n)/2}{a_n} \xrightarrow{d} V.$$

Note that  $V$  can have an atom at  $-\infty$  if  $W$  has an atom at 0. The point process convergence allows to treat upper order statistics of the positions of particles in generation  $n$ . For  $k \in \mathbb{N}$ , let  $M_n^{(k)}$  denote the  $k$ -th order statistic of  $\{V(w)\}_{w \in \mathbb{T}_n}$ . That is for any  $n \in \mathbb{N}$   $\{M_n^{(k)}\}_{k \leq \#\mathbb{T}_n}$  is the non-decreasing enumeration of  $\{V(w)\}_{w \in \mathbb{T}_n}$ . Then we have

$$\mathbb{P} \left[ M_n^{(k)} \leq b_n + a_n x \right] = \mathbb{P} \left[ \mathfrak{S}_{1/a_n} \mathfrak{T}_{-b_n} \Lambda_n(x, +\infty) \leq k - 1 \right] \rightarrow \mathbb{P}[\Lambda(x, +\infty) \leq k - 1].$$

Following [3] we can express the limiting quantity on the right hand side in the following way. Consider a marked point process

$$\tilde{\Lambda} = \sum_{k=0}^{\infty} \delta_{T_k} \otimes \delta_{\iota_k - \log(vW)},$$

which conditioned on  $W$  is a Poisson point process with intensity

$$avW \mathbb{P}[T_1 \in dx] \otimes e^{-y} dy.$$

Denote by  $\pi$  a generic partition of some given integer, say  $l$ , of the form  $l = i_1 y_1 + \dots + i_{|\pi|} y_{|\pi|}$ , where each  $i_j = i_j(\pi)$  repeats  $y_j = y_j(\pi)$  times in the partition and the  $i_j$  are increasing. Here  $|\pi|$  denotes the length of the partition, i.e. the number of its distinct elements. Write  $\Pi_l$  for the collection of all possible partitions of  $l \in \mathbb{N}$ . Then  $\mathbb{P}[\Lambda(x, +\infty) \leq k-1]$  is equal to

$$\sum_{l=1}^{k-1} \sum_{\pi \in \Pi_l} \mathbb{P} \left[ \bigcap_{j=1}^{|\pi|} \{\tilde{\Lambda}(\{i_j\} \times (x + \infty)) = y_j\} \right] = \sum_{l=1}^{k-1} \sum_{\pi \in \Pi_l} \mathbb{E} \left[ \prod_{j=1}^{|\pi|} (avW e^{-x} \mathbb{P}[T_1 = i_j])^{y_j} \frac{1}{y_j!} e^{-avW e^{-x} \mathbb{P}[T_1 = i_j]} \right].$$

The right hand side is therefore the limit of  $\mathbb{P}[M_n^{(k)} \leq b_n + a_n x]$ .

**Remark 3.3.** *The limiting point process (19) falls into the class of randomly shifted decorated Poisson point processes. Comparing the Definition 2.2 with (19) one readily sees that*

$$\Lambda \sim \text{SDPPP} (ae^{-x} dx, T_1 \delta_0, -\log(vW)).$$

**Remark 3.4.** *The limiting point process  $\Lambda$  enjoys the superposability [6, subsection 3.4]: the distances from the rightmost position of a union of realizations of  $\Lambda$  have the same distribution as those of a single realization. To state this in a more rigorous way let  $\Lambda^{(i)} = \sum_{k=0}^{\infty} T_k^{(i)} \delta_{\ell_k^{(i)} - \log(vW_i)}$  denote i.i.d. copies of  $\Lambda$ . Since the martingale limit  $W$  is known to satisfy*

$$\sum_{i=1}^{Z_1} W_i \stackrel{d}{=} mW$$

where  $W_1, W_1, \dots$  are i.i.d. copies of  $W$  independent of  $Z_1$ , one sees from studying the corresponding Laplace functionals that

$$\sum_{i=1}^{Z_1} \Lambda_i \stackrel{d}{=} \mathfrak{T}_{\log m} \Lambda.$$

Now consider a mapping  $\mathfrak{V}$  on  $\mathcal{M}_p$  given by

$$\mathfrak{V} \sum_i \delta_{x_j} = \begin{cases} \sum_i \delta_{x_i - \max_j x_j} & \max_j x_j < \infty \\ o & \max_j x_j = \infty, \end{cases}$$

where  $o$  denotes the null measure. Then  $\mathfrak{V}\eta$  is the point process  $\eta$  seen from the rightmost particle. Since  $\mathfrak{V} \circ \mathfrak{T}_x = \mathfrak{V}$  for any  $x \in \mathbb{R}$ , one readily sees that

$$\mathfrak{V} \sum_{i=1}^{Z_1} \Lambda_i \stackrel{d}{=} \mathfrak{V}\Lambda.$$

**Remark 3.5.** *One can show in a similar fashion that  $\Lambda$  is exponentially stable provided that  $W = 1$  a.s. which is the case when the branching is deterministic, i.e.  $\mathbb{P}[Z_1 = k] = 1$  for some  $k \geq 2$ . More precisely  $\Lambda$  is exponentially 1-stable if for any choice of  $\alpha, \beta \in \mathbb{R}$  such that  $e^\alpha + e^\beta = 1$  we have*

$$\mathfrak{T}_\alpha \Lambda + \mathfrak{T}_\beta \Lambda' \stackrel{d}{=} \Lambda,$$

where  $\Lambda'$  is an independent copy of  $\Lambda$ . One can check that this is true since in the case  $W = 1$  a.s. one has

$$\mathbb{E} \left[ e^{-\int f(x) \mathfrak{T}_\alpha \Lambda(dx)} \right] = \exp \left\{ -a(k-1)e^\alpha \int_{\mathbb{R}} \sum_{j=0}^{\infty} k^{-j-1} (1 - e^{-k^j f(x)}) e^{-x} dx \right\}$$

for any  $\alpha \in \mathbb{R}$ . It has been conjectured in [6, subsection 6.1] and proved in [16] that a exponentially 1-stable point process is a decorated exponential Poisson process i.e. with intensity of the form  $Ce^{-x} dx$  and a certain decoration. This can be checked quite easily in our context since the aforementioned SDPPP representation of  $\Lambda$  boils down to a DPPP representation

$$\Lambda = \sum_k \mathfrak{T}_{L_k} D_k, \quad D_k = T_k \delta_{-\log v}.$$

with  $v$  defined in (11).

#### 4. Trimming lemmas

In our first steps towards the proof of Theorem 3.1 we will show which particles in  $\mathbb{T}_n$  contribute to the extremes of  $\Lambda_n$ . To facilitate this we will use several deviation estimates for sums of i.i.d. random variables with stretched exponential tails. We will often use the following expansion for  $R$ ,

$$R(d_n + x) = R(d_n) + R'(d_n)x + x^2 R''(\xi_{d_n, d_n+x})/2$$

for  $x \in \mathbb{R}$  and some  $\xi_{d_n, d_n+x}$  that lies between  $d_n$  and  $d_n + x$ .

The first few steps of our arguments concentrate on pinpointing the particles that contribute to the point process which is the limit of  $\mathfrak{S}_{a_n^{-1}\mathfrak{T}_{-b_n}}\Lambda_n$ . As we will see the positions of those particles will be asymptotically independent. We will partition  $\mathbb{T}_n$  into four classes of particles and show which class contributes to the asymptotics of  $\mathfrak{S}_{a_n^{-1}\mathfrak{T}_{-b_n}}\Lambda_n$ . The first one consists of those particles with no big jumps along their ancestral line, i.e.

$$\mathcal{A}_n = \left\{ w \in \mathbb{T}_n : \max_{z \in (\emptyset, w]} X_z \leq \varepsilon_n \right\},$$

where  $\varepsilon_n = \varepsilon d_n$  for some fixed  $\varepsilon \in (2^{-r}, 1)$ . The positions of particles from  $\mathcal{A}_n$  consist of individual displacements of typical particles and are therefore well controlled. The point process given by the positions of the particles from  $\mathcal{A}_n$  will be denoted by

$$\Lambda_n^{\mathcal{A}} = \sum_{w \in \mathcal{A}_n} \delta_{a_n^{-1}(V(w)-b_n)}.$$

As we will see in Lemma 4.2 below, the contribution of particles from  $\mathcal{A}_n$  is negligible. To control the contribution of  $\Lambda_n^{\mathcal{A}}$  we will need a deviation results for typical particles. For  $L > 0$  take  $\mathbb{P}_L$  to be a probability measure such that the  $X_k$ 's are i.i.d. under  $\mathbb{P}_L$  with distribution given by

$$\mathbb{P}_L[X_k \leq t] = \mathbb{P}[X_k \leq t \mid X_k \leq L].$$

One can use classical large deviation techniques to control the sum of the  $X_k$ 's under  $\mathbb{P}_L$ . More precisely one can exploit the exponential change of measure and use the cumulant generating function of the truncated variables

$$K_L(s) = \log \mathbb{E}_L[\exp\{sX_1\}], \quad s > 0,$$

where  $\mathbb{E}_L$  denotes the expectation corresponding to  $\mathbb{P}_L$ . In order to be able to control the particles from  $\mathcal{A}_n$  we need to take  $L = \varepsilon_n$  in the above construction. We will first check the asymptotic expansion of  $K_L = K_{\varepsilon_n}$  near 0. For functions  $f$ , we write  $f^{(0)}$ ,  $f^{(1)}$  and  $f^{(2)}$  for  $f$ ,  $f'$  and  $f''$  respectively. Recall the truncated cumulant generating function  $K$  given via (17).

**Lemma 4.1.** *Let Assumption 2 be in force and let  $\varepsilon_n = \varepsilon d_n$  for some  $\varepsilon \in (2^{-r}, 1)$ . Let  $d \in (0, 1)$  be such that  $R(x)x^{-d}$  is decreasing. Then for  $\chi \in (0, 1/d)$  and  $i \in \{0, 1, 2\}$ ,*

$$\sup_{0 \leq s \leq \chi R'(d_n)} \left| K_{\varepsilon_n}^{(i)}(s) - K^{(i)}(s) \right| = o(n^{-1}).$$

The proof of Lemma 4.1 is quite standard and utilizes the fact that the  $k_m$ 's obey a similar recursive formula (16) as the cumulants of  $X_1$  under  $\mathbb{P}_{\varepsilon_n}$ . The details are given in the appendix, where it is restated as Lemma A.2. One of the consequences of Lemma 4.1 is the fact that since  $K(s) \sim s^2/2$  as  $s \rightarrow 0^+$  we have

$$nK_{\varepsilon_n}(s) = (1 + o(1))ns^2/2,$$

as  $n \rightarrow \infty$  and  $s \rightarrow 0^+$  provided that  $s \leq \chi R'(d_n)$ . For future reference denote

$$(20) \quad N_n = \max_{1 \leq k \leq n} X_k.$$

With Lemma 4.1 at hand we can state and prove the aforementioned lemma regarding the contribution of particles from  $\mathcal{A}_n$ .

**Lemma 4.2.** *Under Assumptions 1 and 2 for any  $r \in (0, 1)$ , if  $b_n \sim d_n$  and  $a_n = 1/R'(d_n)$ , then*

$$\Lambda_n^A \Rightarrow o$$

in  $\mathcal{M}_p$ , where  $o$  denotes the null measure.

**Proof.** Since the convergence is tested by integrating functions of compact support, it is enough to show that

$$\Lambda_n^A[L, \infty] \xrightarrow{\mathbb{P}} 0$$

for any constant  $L \in \mathbb{R}$ . Since  $\Lambda_n^A[L, \infty]$  is integer valued it is sufficient to estimate

$$\mathbb{P}[\Lambda_n^A[L, \infty] \geq 1] \leq \mathbb{E}[\Lambda_n^A[L, \infty]] = \mathbb{E}[Z_n] \mathbb{P}[S_n > b_n + La_n, N_n \leq \varepsilon_n].$$

To tackle the second term on the right hand side note that

$$\mathbb{P}[S_n > b_n + La_n, N_n \leq \varepsilon_n] \leq \mathbb{P}_{\varepsilon_n}[S_n > b_n + La_n].$$

Now use the exponential Markov inequality to get for  $s > 0$ ,

$$\mathbb{P}_{\varepsilon_n}[S_n > b_n + La_n] \leq \exp\{-s(b_n + La_n) + nK_{\varepsilon_n}(s)\}.$$

Choose  $s = (n \log m + n^\xi)/(b_n + La_n)$  with  $\xi \in (3 - 2/r, 1) \cap (0, 1)$ . By Lemma 4.1,  $nK_{\varepsilon_n}(s) = o(n^\xi)$  and thus

$$(21) \quad \exp\{-s(b_n + La_n) + nK_{\varepsilon_n}(s)\} \leq m^{-n} \exp\{-n^\xi/C\}$$

which yields our claim. □

The next class of particles consist of those that had at least two big jumps along the ancestral line, that is

$$\mathcal{B}_n = \{w \in \mathbb{T}_n : \exists v, u \in (\emptyset, w), v \neq u, \min\{X_v, X_u\} > \varepsilon_n\}.$$

In the next lemma we treat

$$\Lambda_n^B = \sum_{w \in \mathcal{B}_n} \delta_{a_n^{-1}(V(w) - b_n)}.$$

**Lemma 4.3.** *Under Assumptions 1 and 2 for any  $r \in (0, 1)$ , if  $b_n \sim d_n$  and  $a_n = 1/R'(d_n)$ ,*

$$\Lambda_n^B \Rightarrow o$$

in  $\mathcal{M}_p$ , where  $o$  denotes the null measure.

**Proof.** By the choice of  $\varepsilon \in (2^{-r}, 1)$ ,  $\mathcal{B}_n = \emptyset$  almost surely for sufficiently large  $n$ . Indeed, by a direct first moment bound

$$(22) \quad \mathbb{P}[|\mathcal{B}_n| \geq 1] \leq C n^2 m^n \mathbb{P}[X > \varepsilon_n]^2 = o(1),$$

since  $2R(\varepsilon_n) \sim 2\varepsilon^r R(d_n)$  and  $2\varepsilon^r > 1$ . Thus

$$\Lambda_n^B = \sum_{w \in \mathcal{B}_n} \delta_{a_n^{-1}(V(w) - b_n)} \Rightarrow o$$

in  $\mathcal{M}_p$ . □

Particles that are not in  $\mathcal{A}_n \cup \mathcal{B}_n$  have exactly one big jump along their ancestral line. We will need to distinguish whether this jump is greater or smaller than

$$(23) \quad y_n = \begin{cases} d_n - (\log n)^2 \frac{d_n^2 + nR(d_n)^2}{d_n R(d_n)} & r \in (0, \frac{2}{3}] \\ d_n - nR'(d_n) & r \in (\frac{2}{3}, 1). \end{cases}$$

For  $w \in \mathbb{T}$  define  $v^*(w)$  to be the element of  $(\emptyset, w]$  closest to the root such that  $X_{v^*(w)} = \max_{z \in (\emptyset, w]} X_z$ . The next class of particles we will consider is given by

$$\mathcal{C}_n = \left\{ w \in \mathbb{T}_n : X_{v^*(w)} \in (\varepsilon_n, y_n] \text{ and } \max_{u \in (\emptyset, w] \setminus \{v^*(w)\}} X_u \leq \varepsilon_n \right\}.$$

Define the corresponding point process via

$$\Lambda_n^{\mathcal{C}} = \sum_{w \in \mathcal{C}_n} \delta_{a_n^{-1}(V(w) - b_n)}.$$

The second truncation sequence  $\{y_n\}_{n \in \mathbb{N}}$  is chosen such that the contribution of the particles in  $\mathcal{C}_n$  is also negligible.

**Lemma 4.4.** *Suppose that Assumptions 1 and 2 are satisfied. Then for sequences  $\{a_n\}_{n \in \mathbb{N}}$  and  $\{b_n\}_{n \in \mathbb{N}}$  chosen as in Theorem 3.1 we have*

$$\Lambda_n^{\mathcal{C}} \Rightarrow o$$

in  $\mathcal{M}_p$ .

**Proof.** We will proceed similarly as in the proof of Lemma 4.2. In order to carry out this program we will first estimate

$$P_n = \mathbb{P}[S_{n-1} + X_n > b_n + La_n, N_{n-1} \leq \varepsilon_n, X_n \in (\varepsilon_n, y_n]],$$

for  $L \in \mathbb{R}$  and  $N_n$  given in (20). We will first treat the case  $r \leq 2/3$ . Note that with  $h_n = R(d_n)/d_n$ , using the exponential Markov inequality,

$$(24) \quad P_n \leq Ce^{-h_n(b_n + La_n) + nK_{\varepsilon_n}(h_n)} \mathbb{E} \left[ e^{h_n X} \mathbb{1}_{\{X \in (\varepsilon_n, y_n]\}} \right].$$

By Lemma 4.1,  $nK_{\varepsilon_n}(h_n) = O(nR(d_n)^2/d_n^2)$  and the expectation on the r.h.s. of (24), by Lemma A.3 from the appendix, has an upper bound of the form

$$\mathbb{E} \left[ e^{h_n X} \mathbb{1}_{\{X \in (\varepsilon_n, y_n]\}} \right] \leq \exp \left\{ -(1 + nR(d_n)^2/d_n^2)(\log n)^2/C \right\}.$$

Finally for sufficiently large  $n$ ,  $h_n(b_n + La_n) \geq R(d_n) - |L/d|$  as  $R(x)x^{-d}$  is assumed to be increasing for some  $d \in (0, 1)$ . Gathering all the estimates together shows that for  $r \leq 2/3$ , and sufficiently large  $n$ ,

$$P_n \leq m^{-n} e^{-(\log n)^2/C}.$$

For the case  $r > 2/3$  we proceed in the same way and get (24) with  $nK_{\varepsilon_n}(h_n) = (1/2 + o(1))nR(d_n)^2/d_n^2$  and by Lemma A.4 from the appendix

$$\mathbb{E} \left[ e^{h_n X} \mathbb{1}_{\{X \in (\varepsilon_n, y_n]\}} \right] \leq \exp \left\{ \frac{nR(d_n)R'(d_n)}{d_n} \left( \frac{R'(d_n)d_n}{R(d_n)} - 1 \right) (1 + o(1)) \right\}.$$

Recall that  $b_n = d_n + \tau_n$  with  $\tau_n = (1 + o(1))nR'(d_n)/2$  and write

$$h_n(b_n + La_n) = -R(d_n) - (1 + o(1))nR(d_n)R'(d_n)/(2d_n).$$

Combining all the estimates yields

$$P_n \leq m^{-n} \exp \left\{ \frac{nR(d_n)R'(d_n)}{d_n} \left( \frac{R'(d_n)d_n}{R(d_n)} - \frac{3}{2} + \frac{R(d_n)}{2R'(d_n)d_n} \right) (1 + o(1)) \right\}.$$

Since we assume that for some  $\epsilon > 0$  the functions  $R(x)x^{-1/2-\epsilon}$  and  $R(x)x^{-1+\epsilon}$  are eventually increasing and decreasing respectively,  $y = xR'(x)/R(x) \in [1/2 + \epsilon, 1 - \epsilon]$  for sufficiently large  $x$  and thus the expression  $y - 3/2 + 1/(2y)$  remains negative and bounded away from 0. We can therefore conclude that for  $r > 2/3$ ,

$$P_n \leq m^{-n} \exp \left\{ -nR(d_n)R'(d_n)/(Cd_n) \right\}.$$

Note that the sequence  $nR(d_n)R'(d_n)/d_n$  is regularly varying with index  $3 - 2/r > 0$  and thus it diverges to  $+\infty$  as  $n \rightarrow \infty$ .

Finally to conclude our claim note that we have proven that for any  $r \in (0, 1)$ , and  $L \in \mathbb{R}$ ,

$$\mathbb{P}[\Lambda_n^c[L, \infty] \geq 1] \leq \mathbb{E}[\Lambda_n^c[L, \infty)] \leq Cnm^n P_n \leq e^{-(\log n)^2/C} \rightarrow 0.$$

□

Our considerations up to this point show that the particles from  $\mathcal{A}_n \cup \mathcal{B}_n \cup \mathcal{C}_n$  do not affect the asymptotic behaviour of  $\mathfrak{S}_{a_n^{-1}\mathfrak{T}_{-b_n}}\Lambda_n$  provided that the latter has a non-trivial limit. In the rest of the article we will show that for  $\mathcal{D}_n = \mathbb{T}_n \setminus (\mathcal{A}_n \cup \mathcal{B}_n \cup \mathcal{C}_n)$  the point process

$$\Lambda_n^{\mathcal{D}} = \sum_{w \in \mathcal{D}_n} \delta_{a_n^{-1}(V(w)-b_n)}$$

has a non-trivial limit in distribution. For this reason we need to show that the extremes on  $\mathcal{D}_n$  decouple.

## 5. Decoupling

Until now we investigated only the contribution of big jumps which as it will soon turn out is not enough for our needs. Now we will be also investigating the contribution coming from typical particles. This is when the precise large deviations beyond the so called one big jump domain come into play. In this Section we will state the large deviation result in Theorem 5.1 and then use it to show that the extremal positions of particles from  $\mathcal{D}_n$  decouple in Lemma 5.2.

For functions  $f_n, g_n : \mathbb{R} \rightarrow \mathbb{R}$  and sets  $J_n, H_n \subseteq \mathbb{R}$  we will say that  $f_{n-j_n}(z_n) \sim g_{n-j_n}(z_n)$  uniformly with respect to  $j_n \in J_n \subseteq \mathbb{R}$  and  $z_n \in H_n \subseteq \mathbb{R}$  if

$$\lim_{n \rightarrow \infty} \sup_{j_n \in J_n} \sup_{z_n \in H_n} \left| \frac{f_{n-j_n}(z_n)}{g_{n-j_n}(z_n)} - 1 \right| = 0.$$

Equivalently,  $f_{n-j_n}(z_n) \sim g_{n-j_n}(z_n)$  for any choice of  $z_n \in H_n$  and  $j_n \in J_n$ . Put

$$(25) \quad H_n = [d_n + nR'(d_n)/2 - z_n^*, d_n + nR'(d_n)/2 + z_n^*],$$

where  $z_n^*$  is any sequence of real numbers that is regularly varying with index strictly less than  $2 - 1/r$  if  $r > 2/3$  and less than  $1/r - 1/2$  if  $r \leq 2/3$ . Finally for  $x_n \in H_n$  we write

$$(26) \quad I(x_n) = \inf_{s \in [0,1]} \{R(x_n - nK'(s)) + n(sK'(s) - K(s))\}.$$

Theorem 5.1 given below states that the probability that  $S_n$  exceeds  $x_n \in H_n$  decays as  $ne^{-I(x_n)}$ . The form of  $I(x_n)$  can be heuristically derived in the following way. Write  $S_n = N_n + S_n - N_n$ . Then  $S_n - N_n$  can be seen as the sum of  $n - 1$  of the  $X'_k$ 's that are not the maximal among  $X_1, \dots, X_n$ . In other words,  $S_n - N_n$  is the sum of typical values  $X_1, \dots, X_n$  and  $N_n$  is the one atypical value among  $X_1, \dots, X_n$ . Now the event  $\{S_n > x_n\}$  is a union of  $\{N_n > a\} \cap \{S_n - N_n > b\}$  along  $a + b = x_n$ . For technical reasons it is most convenient to write  $a = x_n - nK'(s)$  and  $b = nK'(s)$  for some  $s > 0$ . Then the probability of  $\{N_n = x_n - nK'(s)\}$  decays as  $ne^{-R(x_n - nK'(s))}$  and  $\{S_n - N_n > nK'(s)\}$  decays as  $n^{-1/2}e^{n(K(s) - sK'(s))}$  (see (7)). Finally the infimum present in the definition of  $I(x_n)$  should be interpreted as taking the most probable choice of  $s$  and thus of  $a = x_n - nK'(s)$  and  $b = nK'(s)$ .

Note that the infimum in (26) is attained for  $s = s_n^*$  which is the unique solution to

$$s_n^* = R'(x_n - nK'(s_n^*))$$

in the interval  $[0, 1]$ . By computing the derivative one can easily check that, for a sufficiently large  $n \in \mathbb{N}$ , the function  $s \mapsto R'(x_n - nK'(s))$  is a contraction on  $[0, 1]$  and thus the existence and uniqueness of  $s_n^*$  follows from the Banach fixed point theorem. We thus have the following representation

$$I(x_n) = R(x_n - nK'(s_n^*)) + n(s_n^*K'(s_n^*) - K(s_n^*)).$$

One then checks, using the regular variation of  $R'$  that  $s_n^*$  is bounded from above by  $(1 + o(1))R'(d_n)$ . It transpires that if  $r < 2/3$  then  $I(x_n) = R(x_n) + o(1)$  uniformly in  $x_n \in H_n$  and for  $r \geq 2/3$  we have  $I(x_n) = R(d_n) + O(n^{4-3/r}\ell(n))$  uniformly in  $x_n \in H_n$  for some slowly varying function  $\ell$ .

**Theorem 5.1.** *Let Assumption 2 be satisfied and let  $J_n = [0, j_n^*] \cap \mathbb{N}$  where  $j_n^*$  is any sequence of natural numbers which is regularly varying with index strictly less than  $\min\{1, 3/r - 3\}$ . Then uniformly in  $x_n \in H_n$  and  $j \in J_n$ ,*

$$\mathbb{P}[S_{n-j_n} > x_n] \sim ne^{-I(x_n) + O(j_n^* R'(x_n)^2)}.$$

The case  $r \leq 2/3$  corresponds to [17, Theorem 3], the case  $r = 2/3$  to [17, Theorem 5] and the case  $r > 2/3$  to [17, Theorem 2]. However the statement of the last one was too implicit for our needs and therefore we provide a new proof with a more explicit statement. We will present a self-contained proof of Theorem 5.1 in Section 7. Note that the case  $r < 2/3$  is covered in [8].

We will use Theorem 5.1 in the proof of our decoupling lemma which we will state below. For  $k < j$  and  $w \in \mathbb{T}$  define

$$V_{k,j}(w) = \sum_{u \leq w, k < |u| \leq j} X_u.$$

The next lemma asserts that the particles in  $\mathcal{D}_n$  whose positions are around  $b_n$  are asymptotically independent. The idea of the proof is the following. First, note that for  $w \in \mathbb{T}_n \setminus (\mathcal{A}_n \cup \mathcal{B}_n \cup \mathcal{C}_n)$  we have  $X_{v^*(w)} > y_n$ . One can use this information to infer estimates on  $V_{k_n^{(1)}, n}(w)$  for sufficiently big  $k_n^{(1)}$ , which will be determined later. Then, using the branching property and the aforementioned estimates, one can get bounds on  $V_{k_n^{(2)}, n}(w)$  for  $k_n^{(2)} < k_n^{(1)}$ . One can then iterate this procedure  $j$  times until  $k_n^{(j)}$  gets small enough. We will specify the choice of  $k_n^{(j)}$ 's in the proof of the next lemma. To state the next lemma we will need to define a sequence of events  $\{A_n\}_{n \in \mathbb{N}}$  depending on a sequence of natural numbers  $\{k_n\}_{n \in \mathbb{N}}$  and a sequence of real numbers  $\{l_n\}_{n \in \mathbb{N}}$ . If  $r \leq 2/3$  let

$$A_n = A_n(k_n, l_n) = \{\exists w, z \in \mathcal{D}_n, \min\{V_{k_n, n}(w), V_{k_n, n}(z)\} > b_n - l_n, v^*(w) \neq v^*(z), |w \wedge z| > k_n\}.$$

If  $r > 2/3$  let

$$A_n = A_n(k_n, l_n) = \{\exists w, z \in \mathcal{D}_n, \min\{V_{k_n, n}(w), V_{k_n, n}(z)\} > b_n - l_n, v^*(w) \neq v^*(z), |w \wedge z| > k_n\} \\ \cap \{\exists w \in \mathcal{D}_n, V_{k_n, n}(v^*(w)) > b_n - l_n \text{ and } |v^*(w)| \leq n - k_n\}.$$

**Lemma 5.2** (Decoupling lemma). *Suppose that Assumptions 1 and 2 are satisfied. Then there exists a sequence of natural numbers  $k_n$  such that  $k_n = o(n^{1-r-\epsilon})$  for some  $\epsilon > 0$  and a sequence  $l_n = o(b_n)$  such that  $a_n = o(l_n)$  such that*

$$\lim_{n \rightarrow \infty} \mathbb{P}[A_n] = 0.$$

**Proof.** We will first treat the case of  $r \leq 2/3$ . Take  $k_n$  to be the smallest integer greater than

$$(\log n)^3 R'(d_n)(d_n - y_n + 1),$$

take  $l_n = 0$  and write, for sufficiently large  $n$ , a simple estimate

$$\mathbb{P}[A_n] \leq \mathbb{E} \left[ \sum_{w, z \in \mathcal{D}_n} \mathbb{1}_{\{\min\{V_{k_n, n}(w), V_{k_n, n}(z)\} > b_n - l_n, v^*(w) \neq v^*(z), |w \wedge z| > k_n\}} \right] \\ \leq Cm^{2n-k_n} \mathbb{P}[X > y_n]^2 \leq e^{-(\log n)^3/C}$$

which secures our claim for  $r \leq 2/3$ .

The proof for  $r > 2/3$  is more involved. In this case the event  $A_n$  is determined by both conditions mentioned in the statement. For  $j \geq 1$  define the sequence of indexes  $\{q_j\}_{j \geq 1}$ , where  $q_1 \in (3 - 2/r, 1)$  and  $q_{j+1} = 1 + (q_j - 1)/r$ . Then  $q_{j+1} = 1 - 2r^{-j}(1 - q_1) \rightarrow -\infty$ . Let  $j_0$  to be the smallest positive integer for which  $q_{j_0} < (1 - r)$ . Note that by the recursive formula for the  $q_j$ 's,  $q_{j_0} \geq 0$ . We can always choose the value of  $q_1$  such that the last inequality is strict, so we have  $q_{j_0} > 0$ . Next define the approximating sequences  $\{k_n^{(j)}\}_{n \in \mathbb{N}}$  and  $\{\ell_n^{(j)}\}_{n \in \mathbb{N}}$  via  $k_n^{(1)} = n^{q_1}$ ,  $k_n^{(j)} = d_{k_n^{(j-1)}} R'(d_n) \log n$  and  $\ell_n^{(j)} = d_{k_n^{(j)}} R'(d_n) (\log n)^{1/2}$ . Then for each  $j$ ,  $\{k_n^{(j)}\}_{n \in \mathbb{N}}$  is regularly varying with index  $q_j$  and

$\{\ell_n^{(j)}\}_{n \in \mathbb{N}}$  is regularly varying with index  $q_j/r$ . In what follows we will make use of three asymptotic relations between the two sequences which follow from the construction. Namely for each  $j \geq 2$ , as  $n \rightarrow \infty$ ,

$$(27) \quad k_n^{(j)} \gg R'(d_n)\ell_n^{(j-1)}, \quad \ell_n^{(j)} \gg d_{k_n^{(j)}}, \quad \text{and} \quad k_n^{(1)} \gg nR'(d_n)^2.$$

Define  $l_n = l_n^{(j_0)}$  and  $k_n = k_n^{(j_0)}$ . Consider the subsets of  $\mathbb{T}_n$  given via  $\mathcal{F}_n^{(0)} = \mathcal{D}_n$  and for  $j \geq 1$ , let

$$\begin{aligned} \mathcal{G}_n^{(j)} &= \left\{ w \in \mathcal{F}_n^{(j-1)} : |v^*(w)| \leq n - k_n^{(j)} \right\}, \\ \mathcal{H}_n^{(j)} &= \left\{ w \in \mathcal{F}_n^{(j-1)} : |v^*(w)| > n - k_n^{(j)}, V_{k_n^{(j)}, n}(|v^*(w)|) \leq b_n - \ell_n^{(j)} \right\}, \\ \mathcal{F}_n^{(j)} &= \left\{ w \in \mathcal{F}_n^{(j-1)} : |v^*(w)| > n - k_n^{(j)}, V_{k_n^{(j)}, n}(|v^*(w)|) > b_n - \ell_n^{(j)} \right\} \\ &= \left\{ w \in \mathcal{D}_n : |v^*(w)| > n - k_n^{(j)}, V_{k_n^{(i)}, n}(|v^*(w)|) > b_n - \ell_n^{(i)}, i < j \right\}. \end{aligned}$$

We have

$$\mathbb{P} \left[ \mathcal{G}_n^{(1)} \neq \emptyset \right] \leq Cm^{n-k_n^{(1)}} \mathbb{P}[X > y_n] \leq \exp \{-n^{q_1}/C\}$$

and

$$\mathbb{P} \left[ \exists w \in \mathcal{H}_n^{(1)}, V_{k_n, n}(w) > b_n - l_n \right] \leq Cm^n \mathbb{P}[X > y_n] \mathbb{P} \left[ S_{k_n^{(1)} - k_n} > l_n^{(1)}(1 + o(1)) \right] \leq \exp \{-n^{q_1}/C\}.$$

Turning our attention to  $\mathcal{G}_n^{(m)}$  and  $\mathcal{H}_n^{(m)}$  with  $j_0 > m \geq 2$  we use the uniform estimates in Theorem 5.1 and the first condition in (27) to write

$$\begin{aligned} \mathbb{P} \left[ \mathcal{G}_n^{(m)} \neq \emptyset \right] &\leq \mathbb{P} \left[ \exists w \in \mathbb{T}, n - k_n^{(m-1)} < |w| \leq n - k_n^{(m)}, V_{k_n, n}(w) > b_n - l_n^{(m-1)} \right] \\ &\leq \sum_{j=k_n^{(m)}}^{k_n^{(m-1)}} m^{n-j} \mathbb{P}[S_{n-j-k_n} > b_n - l_n^{(m-1)}] \leq \exp \left\{ -k_n^{(m)}/C \right\}. \end{aligned}$$

In a similar way, using the first two conditions in (27) and Theorem 5.1,

$$\begin{aligned} \mathbb{P} \left[ \exists w \in \mathcal{H}_n^{(m)}, V_{k_n, n}(w) > b_n - l_n \right] &\leq \\ &\leq m^n \sum_{j=0}^{k_n^{(m)}} \mathbb{P} \left[ S_{n-j-k_n} > b_n + La_n - l_n^{(m-1)} \right] \mathbb{P}_{\varepsilon_n} \left[ S_j > l_n^{(m)} \right] \leq \exp \left\{ -k_n^{(m)}/C \right\}. \end{aligned}$$

From the above considerations

$$(28) \quad \mathbb{P} \left[ \exists w \in \bigcup_{j=0}^{j_0-1} \mathcal{F}_n^{(j)}, V_{k_n, n}(|v^*(w)|) > b_n - l_n, |v^*(w)| \leq n - k_n \right] \rightarrow 0.$$

Next consider the events

$$A_n^{(j)} = \left\{ \exists w, z \in \mathcal{F}_n^{(j_0)} : \min\{V_{k_n, n}(w), V_{k_n, n}(z)\} > b_n - l_n, v^*(w) \neq v^*(z), |v^*(w) \wedge v^*(z)| \geq k_n^{(j)} \right\}.$$

Then for  $j = 1$ , and sufficiently large  $n$ ,

$$\mathbb{P} \left[ A_n^{(1)} \right] \leq Cm^n m^{n-k_n^{(1)}} \mathbb{P}[X > y_n]^2 \leq \exp \{-n^{q_1}/C\}$$

and for  $j > 1$ ,

$$(29) \quad \mathbb{P} \left[ A_n^{(j)} \right] \leq \mathbb{P} \left[ A_n^{(j-1)} \right] + \mathbb{P} \left[ A_n^{(j)} \setminus A_n^{(j-1)} \right].$$

On the event  $\{k_n^{(j-1)} > |v^*(w) \wedge v^*(z)|\}$  the random variables  $V_{k_n^{(j-1)},n}(w)$  and  $V_{k_n^{(j-1)},n}(z)$  are independent given  $\sigma(X_v : |v| \leq k_n^{(j-1)})$ . Now by the definition of  $\mathcal{F}_n^{(j-1)}$  we use the aforementioned conditional independence and write

$$\begin{aligned} \mathbb{P}\left[A_n^{(j)} \setminus A_n^{(j-1)}\right] &\leq \mathbb{P}\left[\exists w, z \in \mathcal{F}_n^{(j_0)}, V_{k_n,n}(w) > b_n - l_n, \right. \\ &\quad \left. V_{k_n^{j-1},n}(z^*(y)) > b_n - l_n^{(j-1)}, k_n^{(j-1)} > |v^*(w) \wedge v^*(z)| \geq k_n^{(j)}\right] \\ &\leq Cm^n m^{n-k_n^{(j)}} \mathbb{P}\left[S_{n-k_n^{(j-1)}-k_n} > b_n - l_n^{(j-1)}\right] \mathbb{P}[S_{n-k_n} > b_n - l_n]. \end{aligned}$$

The large deviation result for stretched exponential random variables in Theorem 5.1 implies that for sufficiently large  $n$ ,

$$\mathbb{P}[S_{n-k_n^{(j-1)}} > b_n + La_n - l_n^{(j-1)}] \leq m^n \exp\left\{CR'(d_n)l_n^{(j-1)}\right\}.$$

By our choice of sequences,  $R'(d_n)l_n^{(j-1)} \ll k_n^{(j)}$  as stated in (27). Going back to (29) gives

$$\mathbb{P}\left[A_n^{(j)}\right] \leq \mathbb{P}\left[A_n^{(j-1)}\right] + \exp\left\{-k_n^{(j)}/C\right\}$$

which by induction implies that for any  $j \leq j_0$ ,  $\mathbb{P}[A_n^{(j)}] \rightarrow 0$ . The claim follows since  $A_n$  is a union of  $A_n^{(j_0)}$  and the event considered in (28).  $\square$

Lemma 5.2 will allow for an efficient decoupling. The position of each  $w \in \mathcal{D}_n$  can be written in the following way:

$$(30) \quad V(w) = H_1(w) + T(w) + X_{v^*(w)} + H_2(w),$$

where

$$\begin{aligned} H_1(w) &= \sum_{z \leq w} X_z \mathbf{1}_{\{|z| \leq k_n\}}, & H_2(w) &= \sum_{z \leq w} X_z \mathbf{1}_{\{X_z \leq \varepsilon_n, |z| > n - k_n\}}, \\ T(w) &= \sum_{\substack{|z| \leq w \\ |z| \in [k_n, n - k_n]}} X_z \mathbf{1}_{\{|z| \in [k_n, n - k_n]\}}. \end{aligned}$$

As one may expect, the first term on the r.h.s. of (30) is negligible.

**Lemma 5.3.** *For any  $r \in (0, 1)$ ,  $\mathbb{P}$ -a.s.*

$$\lim_{n \rightarrow \infty} \max_{w \in \mathcal{D}_n} H_1(w)/a_n = 0.$$

**Proof.** One can use the quite generous upper bound

$$\max_{w \in \mathcal{D}_n} |H_1(w)| \leq M_{k_n}.$$

Since  $d_{k_n} = o(n^{1/r-1-\epsilon})$  for some  $\epsilon > 0$  as in the claim of Lemma 5.2, we get  $M_{k_n}/a_n \rightarrow 0$  as  $a_n$  is regularly varying with index  $1/r - 1$ .  $\square$

As a consequence of Lemma 5.3 in order to establish convergence of  $\Lambda_n^{\mathcal{D}}$  it is sufficient to study

$$\Theta_n = \sum_{w \in \mathcal{D}_n} \delta_{a_n^{-1}(T(w) + X_{v^*(w)} + H_2(w) - b_n)}.$$

**Lemma 5.4.** *Suppose that Assumptions 1 and 2 are satisfied. Let  $a_n$  and  $b_n$  be chosen as in the statement of Theorem 3.1. Then as  $n \rightarrow \infty$ ,*

$$\Lambda_n^{\mathcal{D}} - \Theta_n \Rightarrow o$$

in  $\mathcal{M}_p$ .

We will present the proof of this lemma once we establish the convergence of  $\Theta_n$ .

## 6. Convergence of the point processes

We now turn to the final arguments in the proof of Theorem 3.1. We will first show that the decoupled point process  $\Theta_n$  converges to the desired limit. Then, using the tightness of  $\Theta_n$ , we will show Lemma 5.4 which will yield immediately Theorem 3.1. Note first that we can rewrite  $\Theta_n$  in the following way

$$\Theta_n = \sum_{n-k_n < |v| \leq n} \mathbb{1}_{\{X_v > y_n\}} \sum_{z \geq v, |z|=n} \delta_{a_n^{-1}(T(z)+X_v+H_2(z)-b_n)} \mathbb{1}_{E_n(v,z)},$$

where  $E_n(v, z) = \{X_w \leq \varepsilon_n \text{ for all } w \in (\emptyset, z] \setminus \{v\}, |w| \geq k_n\}$ .

**Proposition 6.1.** *Suppose that Assumptions 1 and 2 are satisfied. Then for sequences  $\{b_n\}_{n \in \mathbb{N}}$  and  $\{a_n\}_{n \in \mathbb{N}}$  as in the statement of Theorem 3.1, and  $\Lambda$  given by (19),*

$$\Theta_n \Rightarrow \Lambda$$

in  $\mathcal{M}_p$ .

**Proof.** By the choice of the centring sequence  $\{b_n\}_{n \in \mathbb{N}}$  and the scaling sequence  $\{a_n\}_{n \in \mathbb{N}}$ , by an appeal to Lemma 4.2, Lemma 4.4 and Theorem 5.1 we have for any  $x \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} m^n \mathbb{P}[S_{n-2k_n} + X > b_n + xa_n, N_{n-2k_n} \leq \varepsilon_n, X > y_n] = ae^{-x}.$$

Therefore if we denote

$$\eta_n(\cdot) = \mathbb{P}[a_n^{-1}(S_{n-2k_n} + X - b_n) \in \cdot, N_{n-2k_n} \leq \varepsilon_n, X > y_n]$$

the above observation implies that vaguely

$$m^n \eta_n(dx) \rightarrow \eta(dx) = ae^{-x} dx.$$

Fix any  $f \in \mathcal{C}_K^+$  and take  $n \in \mathbb{N}$  sufficiently small such that  $\text{supp}(f) \subseteq [-l_n/a_n, \infty]$ . Recall the statement of Lemma 5.2 and the definition of the event  $A_n$ . By conditioning on  $\mathbb{T}$  and  $\mathcal{E}_n = \sigma(X_u, |u| \geq n - k_n)$  we can write on the set  $A_n^c$ ,

$$\begin{aligned} \mathbb{E} \left[ \exp \left\{ - \int f(x) \Theta_n(dx) \right\} \middle| \mathbb{T}, \mathcal{E}_n \right] = \\ \prod_{|v| \in (n-k_n, n]} \mathbb{E} \left[ \exp \left\{ - \mathbb{1}_{E_n} \sum_{w \geq v, |w|=n} \mathbb{1}_{E'_n(v,w)} f \left( \frac{S_{n-k_n-1} + H_2(w) + X_v}{a_n} \right) \right\} \middle| \mathbb{T}, \mathcal{E}_n \right] \end{aligned}$$

where  $S_{n-k_n-1}$  is independent from everything else,

$$E'_n(v, w) = \{X_z \leq \varepsilon_n \text{ for all } z \in (\emptyset, w] \setminus \{v\}, |z| \geq n - k_n, X_v > y_n\} \in \mathcal{E}_n$$

and  $E_n = \{N_{n-k_n-1} \leq \varepsilon_n\}$ . Integrating the conditional expectation over the set  $A_n^c$  and using the fact that  $\mathbb{P}[A_n] = o(1)$  we get

$$\mathbb{E} \left[ \exp \left\{ - \int f(x) \Theta_n(dx) \right\} \right] = \mathbb{E} \left[ \prod_{k=n-k_n}^n \mathbb{E} \left[ \exp \left\{ - \int f(x) \Theta_{n,k}(dx) \right\} \right]^{Z_k} \right] + o(1)$$

where

$$\Theta_{n,k} = \mathbb{1}_{E_n} \mathbb{1}_{\{X > y_n\}} \sum_{w \in \mathcal{A}_{n-k}} \delta_{a_n^{-1}(S_{n-2k_n-1} + X + V(w) - b_n)}.$$

We will argue that for  $j \in \mathbb{N}$ ,

$$(31) \quad Z_{n-j} \mathbb{E} \left[ 1 - \exp \left\{ - \int f(x) \Theta_{n,n-j}(dx) \right\} \right] \rightarrow m^{-j} W \int \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] \eta(dx)$$

a.s. which implies

$$\prod_{j=0}^{k_n} \mathbb{E}^* \left[ \exp \left\{ - \int f(x) \Theta_{n,n-j}(dx) \right\} \right]^{Z_{n-j}} \rightarrow \mathbb{E}^* \left[ \exp \left\{ -W \sum_{j=0}^{\infty} m^{-j} \int \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] \eta(dx) \right\} \right].$$

To show (31) define the modulus of continuity of  $f$  via

$$(32) \quad \omega_f(\epsilon) = \sup_{|x-y| \leq \epsilon} |f(x) - f(y)|, \quad \epsilon > 0.$$

Denote

$$\Delta_{n-j} = \int f(x) \Theta_{n,n-j}(dx) - Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))$$

and notice that the law  $|\Delta_{n-j}|$  is dominated in the sense of the stochastic order via

$$|\Delta_{n-j}| \stackrel{d}{\leq} Z_j \omega_f(M_{k_n}/a_n)$$

where the second term on the right hand side goes to 0 in probability as  $n \rightarrow \infty$ , since  $M_{k_n}/a_n \xrightarrow{\mathbb{P}} 0$  (see the proof of Lemma 5.3). Note that since  $f$  is necessarily bounded, the aforementioned convergence holds also in  $L^1$ . Now write

$$\begin{aligned} Z_{n-j} \mathbb{E} \left[ 1 - e^{-\int f(x) \Theta_{n,n-j}(dx)} \right] &= Z_{n-j} \mathbb{E} \left[ 1 - e^{-Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))} \right] \\ &\quad + Z_{n-j} \mathbb{E} \left[ e^{-Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))} (1 - e^{-\Delta_{n-j}}) \right]. \end{aligned}$$

The first term has the desired asymptotics, i.e. by an appeal to the vague convergence  $m^n \eta_n(\cdot) \rightarrow \eta(\cdot)$ , we can write

$$\begin{aligned} Z_{n-j} \mathbb{E} \left[ 1 - e^{-Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))} \right] &= Z_{n-j} \int \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] \eta_n(dx) \\ &= \frac{Z_{n-j}}{m^{n-j}} m^{-j} \int \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] m^n \eta_n(dx) \rightarrow m^{-j} W \int \mathbb{E} \left[ 1 - e^{-Z_j f(x)} \right] \eta(dx). \end{aligned}$$

The second term is negligible since

$$\begin{aligned} Z_{n-j} \mathbb{E} \left[ e^{-Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))} (1 - e^{-\Delta_{n-j}}) \right] \\ \leq Z_{n-j} \mathbb{E} \left[ e^{-Z_j f(a_n^{-1}(S_{n-2k_n} + X - b_n))} \Delta_{n-j} \right] \\ \leq \frac{Z_{n-j}}{m^{n-j}} \mathbb{E} \left[ \frac{Z_j}{m^j} \omega_f \left( \frac{M_{k_n}}{R'(d_n)} \right) \right] m^n \eta_n((\epsilon, \infty]). \end{aligned}$$

The convergence in (31) follows and so does the claim.  $\square$

**Proof of Lemma 5.4.** Fix  $f \in \mathcal{C}_K^+$  and recall (32). As an element of  $\mathcal{C}_K^+$  the function  $f$  must have a limit at  $+\infty$  and thus  $\omega_f(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . For any fixed  $\epsilon > 0$ , on the event  $B_n(\epsilon) = \{\max_{w \in \mathcal{D}_n} |H_1(w)| < \epsilon/R'(d_n)\}$ ,

$$\left| \int f(x) \Lambda_n^{\mathcal{D}}(dx) - \int f(x) \Theta_n(dx) \right| \leq \Theta_n((\epsilon, \infty) + \text{supp}(f)) \omega_f(\epsilon).$$

The claim follows since the events  $B_n(\epsilon)$  have probability close to 1 and  $\omega_f(\epsilon) \rightarrow 0$  while the measure of  $(\epsilon, \infty) + \text{supp}(f)$  with respect to  $\Theta_n$  remains bounded in probability as  $\epsilon \rightarrow 0$ .  $\square$

**Proof of Theorem 3.1.** The claim is an immediate consequence of Lemma 5.4 and Proposition 6.1.  $\square$

## 7. Large deviations for stretched exponential tails

In this section we will revisit precise large deviations for sums of i.i.d. random variables with stretched exponential tails. This problem was addressed previously, see [17, 18]. However the general results in [17] were not sufficiently explicit for our purposes. For this reason we decided to revisit the problem and study the deviations on the scale  $d_n$ , where  $d_n$  is any regularly varying function with index  $1/r$ . We will study deviations uniformly in the time interval

$$n_0 = n - j_n, \quad j_n \in J_n = [0, j_n^*],$$

where  $j_n^*$  is any sequence of natural number which is regularly varying with index strictly less than  $\min\{1, 3/r - 3\}$ . Recall  $H_n \subseteq \mathbb{R}$  given in (25) and the rate function  $I(x_n)$  given in (26).

The first lemma asserts that whenever  $S_{n_0} > x_n$  there has to be at least one jump greater than  $\varepsilon_n = \varepsilon d_n$  for some fixed  $\varepsilon \in (2^{-r}, 1)$ . The proof goes along the exact same lines as the proof of Lemma 4.2 (see (21)) and is therefore omitted.

**Lemma 7.1.** *Suppose that Assumption 2 is satisfied. Then*

$$\sup_{x_n \in H_n} \sup_{j_n \in J_n} \mathbb{P}[S_{n_0} > x_n, N_{n_0} \leq \varepsilon_n] = o\left(ne^{-I(x_n)}\right).$$

Next we will treat the possibility of having two jumps bigger than  $\varepsilon_n$ . The proof is straightforward, see (22), and is therefore also omitted.

**Lemma 7.2.** *Suppose that Assumption 2 is satisfied. Then*

$$\sup_{x_n \in H_n} \sup_{j_n \in J_n} \mathbb{P}[S_{n_0} > x_n, \exists i, j \leq n_0, \min\{X_i, X_j\} > \varepsilon_n, i \neq j] = o\left(ne^{-I(x_n)}\right).$$

Taking the last two lemmas into account, we see that in order to establish the asymptotics of  $\mathbb{P}[S_{n_0} > x_n]$  it is sufficient to study

$$\mathbb{P}[S_{n_0} + X > x_n, N_{n_0} \leq \varepsilon_n, X > \varepsilon_n].$$

In our next step we will show that  $X$  has to be sufficiently big. Recall the sequence  $\{y_n\}_{n \in \mathbb{N}}$  given by (23).

**Lemma 7.3.** *Suppose that Assumption 2 is satisfied. Then*

$$\sup_{x_n \in H_n} \sup_{j_n \in J_n} \mathbb{P}[S_{n_0} + X > x_n, N_{n_0} \leq \varepsilon_n, X \in (\varepsilon_n, y_n]] = o\left(ne^{I(x_n)}\right).$$

The proof goes along the same lines as the proof of Lemma 4.4 and is therefore omitted. In what follows we will investigate

$$P_n = \mathbb{P}[S_{n_0} + X > x_n, N_{n_0} \leq \varepsilon_n, X > y_n].$$

Consider  $w_n > y_n$  given via

$$(33) \quad w_n = \begin{cases} d_n + (\log n)^2 / R'(d_n) & r \in (0, 2/3], \\ d_n + nR'(d_n)/3 & r \in (2/3, 1). \end{cases}$$

We will start by writing

$$\begin{aligned} P_n &= \mathbb{P}[S_{n_0} > x_n - y_n, N_{n_0} \leq \varepsilon_n, X > y_n] \\ &\quad + \mathbb{P}[S_{n_0} < x_n - w_n, N_{n_0} \leq \varepsilon_n, S_{n_0} + X > x_n, X > y_n] \\ &\quad + \mathbb{P}[x_n - S_{n_0} \in [y_n, w_n], N_{n_0} \leq \varepsilon_n, X > x_n - S_{n_0}] =: A_n + B_n + C_n. \end{aligned}$$

Recall that  $n_0 = n - j_n$ ,  $j_n \in J_n = [0, j_n^*]$ , where  $j_n^*$  is any regularly varying sequence with index less than  $\min\{1, 3/r - 3\}$ . Recall also  $H_n$  given in (25).

**Lemma 7.4.** *Suppose that Assumption 2 is satisfied. We have*

$$\sup_{x_n \in H_n} \sup_{j_n \in J_n} (A_n + B_n) = o\left(ne^{-I(x_n)}\right).$$

**Proof.** We will first treat  $A_n$ . Simply write

$$A_n \leq \mathbb{P}[X > y_n] \mathbb{P}_{\varepsilon_n}[S_{n_0} > x_n - y_n].$$

If  $r \leq 2/3$  the above estimate combined with an exponential Markov's inequality with  $s = (1 + \epsilon)R'(d_n)$  constitutes

$$A_n \leq \mathbb{P}[X > y_n] \exp\{n_0 K_{\varepsilon_n}(s) - s(x_n - y_n)\}.$$

Provided that  $\epsilon > 0$  is sufficiently small, we can invoke Lemma 4.1 and conclude that

$$A_n \leq \exp\{-R(d_n) - \epsilon(\log n)^2\}.$$

For  $r > 2/3$  we apply the same argument (with the same choice of  $s$  in the exponential Markov's inequality) and get

$$A_n \leq \exp\{-R(d_n) - \epsilon n R'(d_n)^2\}.$$

This proves our first claim. For the second term we have a simple bound  $B_n \leq \mathbb{P}[X > w_n]$  which implies our claim.  $\square$

To treat  $C_n$  firstly condition on  $X_1, \dots, X_{n_0}$  and write

$$(34) \quad C_n = \mathbb{E}[e^{-R(x_n - S_{n_0})} \mathbb{1}_{\{N_{n_0} \leq \varepsilon_n, x_n - S_{n_0} \in [y_n, w_n]\}}].$$

Recall the measure  $\mathbb{P}_{\varepsilon_n}$  introduced in Section 4 and that we write  $\mathbb{E}_{\varepsilon_n}$  for its expectation. We have uniformly in  $x_n \in H_n$  and  $j_n \in J_n$ ,

$$C_n = (1 + o(1)) \mathbb{E}_{\varepsilon_n}[e^{-R(x_n - S_{n_0})} \mathbb{1}_{\{x_n - S_{n_0} \in [y_n, w_n]\}}].$$

We will use yet another change of measure. For  $s > 0$  denote by  $\mathbb{P}_{s, \varepsilon_n}$  a probability measure with respect to which  $X_k$  are i.i.d. distributed according to

$$\mathbb{P}_{s, \varepsilon_n}[X_k \leq t] = e^{-K_{\varepsilon_n}(s)} \mathbb{E}_{\varepsilon_n}[\mathbb{1}_{\{X_k \leq t\}} e^{sX_k}].$$

Denoting by  $\mathbb{E}_{s, \varepsilon_n}$  the expected value corresponding to  $\mathbb{P}_{s, \varepsilon_n}$  we have  $\mathbb{E}_{s, \varepsilon_n}[X_1] = K'_{\varepsilon_n}(s)$  and  $\mathbb{E}_{s, \varepsilon_n}[(X_1 - \mathbb{E}_{s, \varepsilon_n}[X_1])^2] = K''_{\varepsilon_n}(s)$ . Consider the distribution of normalized sums

$$F_n^{(s)}(y) = \mathbb{P}_{s, \varepsilon_n}[\theta_n(S_{n_0}) \leq y], \quad \text{where} \quad \theta_n(y) = (y - n_0 K'_{\varepsilon_n}(s)) / \sqrt{n_0 K''_{\varepsilon_n}(s)}.$$

Then  $F_n^{(s)}$  converge point wise to the cumulative distribution function of standard normal distribution. and observe that for  $D_n = [\theta_n(x_n - w_n), \theta_n(x_n - y_n)]$  we have

$$\begin{aligned} \mathbb{E}_{\varepsilon_n}[e^{-R(x_n - S_{n_0})} \mathbb{1}_{\{x_n - S_{n_0} \in [y_n, w_n]\}}] &= \\ e^{n_0 K_{\varepsilon_n}(s)} \mathbb{E}_{s, \varepsilon_n} \left[ e^{-s S_{n_0} - R(x_n - S_{n_0})} \mathbb{1}_{\{x_n - S_{n_0} \in [y_n, w_n]\}} \right] &= \\ = \exp\{n_0 K_{\varepsilon_n}(s)\} \int_{D_n} e^{-s \theta_n^{-1}(y) - R(x_n - \theta_n^{-1}(y))} F_n^{(s)}(dy). \end{aligned}$$

Now use Taylor expansion for  $R$  to write, for  $y \in \theta_n([x_n - w_n, x_n - y_n])$ , and  $x_n(s) = x_n - n_0 K'_{\varepsilon_n}(s)$ ,

$$R(x_n - \theta_n^{-1}(y)) = R(x_n(s)) - y R'(x_n(s)) \sqrt{n_0 K''_{\varepsilon_n}(s)} + y^2 R''(\xi_y) n_0 K''_{\varepsilon_n}(s) / 2$$

for some  $\xi_y$  that lies between  $x_n(s)$  and  $x_n(s) - y \sqrt{n_0 K''_{\varepsilon_n}(s)}$ . Thus we arrive at

$$C_n = (1 + o(1)) e^{n_0 K_{\varepsilon_n}(s) - n_0 s K'_{\varepsilon_n}(s) - R(x_n(s))} \int_{D_n} e^{y e_n(s) - y^2 R''(\xi_y) n_0 K''_{\varepsilon_n}(s) / 2} F_n^{(s)}(dy)$$

with

$$e_n(s) = (R'(x_n(s)) - s) \sqrt{n_0 K''_{\varepsilon_n}(s)}.$$

We want to take  $s = s^*$  in the above construction, where  $s^*$  is characterised by being the unique solution to (5). Then, by the merit of regular variation of  $R''$ , for some slowly varying function  $\ell_n$ ,

$$|e_n(s^*)| \leq |R'(x_n(s^*)) - R'(x_n - nK'(s^*))| \sqrt{nK''_{\varepsilon_n}(s^*)} \leq n^{-1/2} \ell_n.$$

Note that we used here the bound on  $j_n^*$ . In the sequel we will write  $e_n = e_n(s^*)$ . Finally by the regular variation of  $R'''$  for some slowly varying sequence  $\ell_n$

$$(35) \quad \sup_{z \in [x_n - w_n, x_n - y_n]} |R''(\xi_{x_n, x_n + z}) - R''(x_n)| \leq \epsilon_n = n^{3-4/r} \ell_n.$$

It is therefore sufficient to show the claim of our next lemma.

**Lemma 7.5.** *Suppose that Assumption 2 is satisfied. Let  $f_n = (R''(x_n(s^*)) + \epsilon_n)n_0K''_{\varepsilon_n}(s^*)$  where  $\{\epsilon_n\}_{n \in \mathbb{N}}$  is regularly varying with index  $3 - 4/r$ . Then, as  $n \rightarrow \infty$ ,*

$$\sup_{x_n \in H_n} \sup_{j_n \in J_n} \left| \int_{D_n} e^{ye_n - y^2 f_n/2} F_n^{(s^*)}(dy) - 1 \right| \rightarrow 0.$$

**Proof.** Firstly, for  $r \geq 2/3$ ,  $ye_n - y^2 f_n/2 \rightarrow 0$  uniformly in  $y \in D_n$ ,  $x_n \in H_n$ , and  $j_n \in J_n$  and so the claim follows from the central limit theorem. We thus treat only the case  $r > 2/3$ . We will utilize the following estimate, for  $y \in [0, \theta_n(x_n - y_n)]$  and  $h > 0$  we have

$$1 - F_n^{(s^*)}(y) = e^{-n_0 K'_{\varepsilon_n}(s^*)} \mathbb{E}_{\varepsilon_n} [e^{s^* S_{n_0}} \mathbf{1}_{\{\theta_n(S_{n_0}) > y\}}] \leq (1 + o(1)) \exp \left\{ n_0 (K_{\varepsilon_n}(s^* + h) - K_{\varepsilon_n}(s^*) - hK'_{\varepsilon_n}(s^*)) - \sqrt{n_0 K''_{\varepsilon_n}(s^*)} h y \right\}.$$

Take  $h = \epsilon y / \sqrt{n_0}$  for  $\epsilon > 0$  sufficiently small so that we can invoke Lemma 4.1 to get

$$(36) \quad 1 - F_n^{(s^*)}(y) \leq C \exp \{-y^2 / C_\epsilon\},$$

for some constant  $C_\epsilon$  that depends on  $\epsilon$ . Similarly, for  $y \in [\theta_n(x_n - w_n), 0]$

$$F_n^{(s^*)}(y) \leq C \exp \{-y^2 / C_\epsilon\}.$$

Denote  $D_n^+ = [0, \theta_n(x_n - y_n)]$ . We will first show that

$$\sup_{z_n \in I_n} \sup_{j_n \in J_n} \left| \int_{D_n^+} e^{ye_n - y^2 f_n/2} F_n^{(s^*)}(dy) - \frac{1}{2} \right| \rightarrow 0.$$

Using the integration by parts formula we arrive at

$$\begin{aligned} \int_{D_n^+} e^{ye_n - y^2 f_n/2} F_n^{(s^*)}(dy) &= \int_{D_n^+} (e_n - y f_n) e^{ye_n - y^2 f_n/2} (1 - F_n^{(s^*)}(y)) dy \\ &\quad - e^{\theta_n(x_n - y_n) e_n - \theta_n(x_n - y_n)^2 f_n/2} (1 - F_n^{(s^*)}(\theta_n(x_n - y_n))) + (1 - F_n^{(s^*)}(0)). \end{aligned}$$

One can show that the first two terms on the right hand side are negligible using the estimate (36). Indeed the behaviour of the second term follows directly from (36). The integral can be treated by considering  $y > n^{2/r-2-\epsilon}$  and  $y < n^{2/r-2-\epsilon}$ . In the first case the pre factor  $e_n + f_n y$  is bounded via  $n^{-\epsilon}$ . In the second case the exponent tends to  $-\infty$ . Our claim follows since  $F_n^{(s^*)}(0) \rightarrow 1/2$ . In a similar fashion we treat  $D_n^- = [\theta_n(x_n - w_n), 0]$ , we write

$$\begin{aligned} \int_{D_n^-} e^{ye_n - y^2 f_n/2} F_n^{(s^*)}(dy) &= \int_{D_n^-} (e_n - y f_n) e^{ye_n - y^2 f_n/2} (1 - F_n^{(s^*)}(y)) dy \\ &\quad - e^{\theta_n(x_n - w_n) e_n - \theta_n(x_n - w_n)^2 f_n/2} (F_n^{(s^*)}(\theta_n(x_n - w_n))) + F_n^{(s^*)}(0). \end{aligned}$$

and conclude that the limit of the right hand side is  $1/2$ . □

**Proof of Theorem 5.1.** Our considerations up to this point show that it is enough to justify that  $C_n$  given in (34) has the desired asymptotics. Note that with  $\varepsilon_n^*$  be given via (35) one has

$$\int_{D_n} e^{ye_n + y^2 f_n^\downarrow/2} F_n^{(s^*)}(dy) \leq C_n e^{n_0 s^* K'_{\varepsilon_n}(s^*) - n_0 K_{\varepsilon_n}(s^*) + R(x_n(s^*))} \leq \int_{D_n} e^{ye_n + y^2 f_n^\uparrow/2} F_n^{(s^*)}(dy)$$

where  $f_n^\downarrow = (R''(x_n(s^*)) + \varepsilon_n^*)n_0 K''_{\varepsilon_n}(s^*)$  and  $f_n^\uparrow = (R''(x_n(s^*)) - \varepsilon_n^*)n_0 K''_{\varepsilon_n}(s^*)$ . The claim now follows from Lemma 7.5 and the fact that

$$n_0 s^* K'_{\varepsilon_n}(s^*) - n_0 K_{\varepsilon_n}(s^*) + R(x_n(s^*)) = I(x_n) + O(j_n^* R'(x_n)^2).$$

□

## Appendix A: Auxiliary estimates

Here we present the proof of the auxiliary lemmas used in the article. On a few occasions will make use of the following integration by parts formula: for  $A < B$  and a continuously differentiable function  $\psi$ ,

$$(37) \quad \mathbb{E}[\psi(X)\mathbb{1}_{\{X \in (A, B]\}}] = \int_A^B \psi'(s)\mathbb{P}[X > s]ds + \psi(A)\mathbb{P}[X > A] - \psi(B)\mathbb{P}[X > B].$$

In the proof of Lemma 4.1 it will be convenient to first analyse the moment generating function of the truncated random variables

$$F_{\varepsilon_n}(s) = \mathbb{E}_{\varepsilon_n}[\exp\{sX_1\}],$$

where  $\varepsilon_n = \varepsilon d_n$ , with  $\varepsilon \in (2^{-r}, 1)$ . We begin with a result which will allow to control  $F_{\varepsilon_n}$ . Recall that  $\kappa$  is the smallest integer greater than or equal to  $(2-r)/(1-r)$ . Recall that under Assumption 2 the function  $R(x)x^{-d}$  is decreasing for some  $d \in (0, 1)$ .

**Lemma A.1.** *Let Assumption 2 be in force and let  $d \in (0, 1)$  be such that  $R(x)x^{-d}$  is decreasing. Then for any  $k > \kappa$  such that  $\mathbb{E}[|X|^k] < \infty$  and  $\chi \in (0, 1/d)$  we have, as  $n \rightarrow \infty$ ,*

$$\sup_{0 \leq s \leq \chi R'(\varepsilon_n)} s^k F_{\varepsilon_n}^{(k)}(s) = o(n^{-1}).$$

**Proof.** As  $X$  is bounded under  $\mathbb{P}_{\varepsilon_n}$ ,

$$F_{\varepsilon_n}^{(k)}(s) = \mathbb{E}_{\varepsilon_n}[X^k e^{sX}].$$

Since  $\mathbb{P}[X \leq \varepsilon_n] \rightarrow 1$  it is sufficient to estimate

$$\mathbb{E}[X^k e^{sX} \mathbb{1}_{\{X \leq \varepsilon_n\}}] = \mathbb{E}[X^k e^{sX} \mathbb{1}_{\{X \leq 0\}}] + \mathbb{E}[X^k e^{sX} \mathbb{1}_{\{0 < X \leq \varepsilon_n\}}].$$

The first integral can be bounded from above via  $\mathbb{E}[|X|^k]$ . We will apply the integration by parts formula (37) for the second expectation with  $\psi(x) = x^k e^{sx}$ ,  $A = 0$  and  $B = \varepsilon_n$ . Before we present the next bound, let us note that

$$(38) \quad \sup_{0 \leq s \leq \chi R'(\varepsilon_n)} [s\varepsilon_n - R(\varepsilon_n)] \leq (\chi d - 1)R(\varepsilon_n).$$

With this estimate we infer that for sufficiently large  $n$ ,

$$\psi(\varepsilon_n)\mathbb{P}[X > \varepsilon_n] \leq e^{-R(d_n)/C}.$$

For the integral we can write

$$\int_0^{\varepsilon_n} \psi'(t)\mathbb{P}[X > t]dt \leq C \cdot d_n^k \int_0^{\varepsilon_n} \exp\{st - R(t)\} dt \leq C \cdot d_n^{k+1} \int_{\delta+o(1)}^1 \exp\{s\varepsilon_n t - R(\varepsilon_n t)\} dt.$$

As  $R$  is regularly varying with index  $r \in (0, 1)$ ,  $R(\varepsilon_n t) \sim t^r R(\varepsilon_n)$  almost uniformly in  $t \in (0, \infty)$ , i.e. uniformly in  $t \in G$  for any compact  $G \subseteq (0, +\infty)$ . Thus for arbitrarily small  $\varepsilon > 0$  there is  $n$  large enough such that  $R(\varepsilon_n t) \geq t^r R(\varepsilon_n)(1 - \varepsilon)$  for all  $t \in [\delta/2, 1]$ . Thus the integral can be bounded in the following way

$$\int_0^{\varepsilon_n} \psi'(t) \mathbb{P}[X > t] dt \leq C d_n^{k+1} \exp\{(\chi d - 1 + \varepsilon)R(\varepsilon_n)\}.$$

Provided that  $\varepsilon$  is sufficiently small the right hand side decays exponentially fast in  $n$  which in turn secures our claim.  $\square$

We now turn to the proof of Lemma 4.1 which we recall here for convenience.

**Lemma A.2.** *Let Assumption 2 be in force. Let  $d \in (0, 1)$  be such that  $R(x)x^{-d}$  is decreasing and let  $\chi \in (0, 1/d)$ . Then for any  $i \in \{0, 1, 2\}$ ,*

$$\sup_{0 \leq s \leq \chi R'(\varepsilon_n)} \left| K_{\varepsilon_n}^{(i)}(s) - K^{(i)}(s) \right| = o\left(\frac{1}{n}\right).$$

**Proof of Lemma A.2.** Since under  $\mathbb{P}_{\varepsilon_n}$ ,  $X$  is bounded a.s. for any fixed  $s$ ,

$$F_{\varepsilon_n}(s) = \sum_{j=0}^{\kappa} \frac{s^j}{j!} \mathbb{E}_{\varepsilon_n} [X^j] + \int_0^s \frac{(s-t)^\kappa}{(\kappa+1)!} F_{\varepsilon_n}^{(\kappa+1)}(t) dt,$$

where for any fixed  $j$ ,  $\mathbb{E}_{\varepsilon_n} [X^j] = \mathbb{E}[X^j] + o(1/n)$ . Let

$$F(s) = \sum_{j=0}^{\kappa} \frac{s^j}{j!} \mathbb{E} [X^j].$$

By Lemma A.1 and the choice of  $\kappa$ , for any  $i \in \{0, 1, 2\}$ ,

$$(39) \quad \sup_{0 \leq s \leq \chi R'(\varepsilon_n)} |F^{(i)}(s) - F_{\varepsilon_n}^{(i)}(s)| = o(1/n).$$

We continue with

$$(40) \quad K_{\varepsilon_n}(s) = \log(F_{\varepsilon_n}(s)) = \log(F(s)) + \widehat{\Delta}_n(s),$$

with  $\widehat{\Delta}_n(s)$  such that  $\sup_{0 \leq s \leq \chi R'(\varepsilon_n)} |\widehat{\Delta}_n(s)| = o(1/n)$ . Define  $\phi(s) = \log(F(s))$ . Following [22] we have  $\phi(0) = 0$  and

$$\frac{d}{ds} F(s) = \frac{d}{ds} e^{\phi(s)} = \phi'(s) e^{\phi(s)} = \phi'(s) F(s).$$

We see that  $\phi'(0) = 0$  and after differentiating  $m - 1$  times, where  $m \leq \kappa$ , we get

$$\frac{d^m}{ds^m} \phi(0) = \mathbb{E}[X^m] - \sum_{j=1}^{m-1} \binom{m-1}{j} \frac{d^{m-j}}{ds^{m-j}} \phi(0) \cdot \frac{d^j}{ds^j} F(0).$$

Comparing the above relation with (16) allows us to conclude that  $\frac{d^m}{ds^m} \phi(0) = k_m$  for  $m \leq \kappa$ . Using the Taylor expansion of  $\phi$  we get finally that

$$\log F(s) = \phi(s) = \sum_{j=2}^{\kappa} \frac{k_j}{j!} s^j + \frac{s^{\kappa+1}}{(\kappa+1)!} \phi^{(\kappa+1)}(\vartheta s) = K(s) + \frac{s^{\kappa+1}}{(\kappa+1)} \phi^{(\kappa+1)}(\vartheta s)$$

for some  $\vartheta \in [0, 1]$  with

$$\sup_{0 \leq s \leq \chi R'(\varepsilon_n)} \left| \frac{s^{\kappa+1}}{(\kappa+1)} \phi^{(\kappa+1)}(s) \right| = o\left(\frac{1}{n}\right)$$

by the choice of  $\kappa$ . This in combination with (40) yields our claim for  $i = 0$ . The claims for  $i = 1$  and  $i = 2$  follow from a direct comparison of  $K^{(i)}$  and  $K_{\varepsilon_n}^{(i)}$  and an appeal to (39).  $\square$

**Lemma A.3.** *Let Assumption 2 be in force for some  $r \leq 2/3$ . Then for  $y_n = d_n - (\log n)^2(d_n^2 + nR(d_n)^2)/(d_n R(d_n))$ ,  $h_n = R(d_n)/d_n$  and  $\varepsilon_n = \delta d_n$  for  $\varepsilon \in (2^{-r}, 1)$  the following inequality holds true for sufficiently large  $n$ ,*

$$\mathbb{E} \left[ e^{h_n X} \mathbf{1}_{\{X \in (\varepsilon_n, y_n)\}} \right] \leq \exp \left\{ -(\log n)^2 (1 + nR(d_n)^2/d_n^2)/C \right\}.$$

**Proof.** We will apply the integration by parts formula (37) with  $\psi(s) = e^{h_n s}$ ,  $A = \varepsilon_n$  and  $B = y_n$ . We have

$$(41) \quad \psi(\varepsilon_n) \mathbb{P}[X > \varepsilon_n] = \exp(\delta R(d_n) - R(\delta d_n)) \leq \exp\{-n/C\},$$

since  $R(\varepsilon d_n) \sim \varepsilon^r R(d_n) \sim \varepsilon^r n$  and  $\varepsilon \in (0, 1)$ . Now note that since  $R(x)x^{-d}$  is assumed to be eventually decreasing for some  $d \in (0, 1)$ , we have  $R(x)/(R'(x)x) > 1/d > 1$  for sufficiently large  $x$  and thus

$$(42) \quad h_n y_n - R(y_n) \leq (1 - 1/d)(\log n)^2 (1 + nR(d_n)^2/d_n^2) + o(1)$$

which in turn gives

$$\psi(y_n) \mathbb{P}[X > y_n] \leq C \exp \left\{ -(\log n)^2 (1 + nR(d_n)^2/d_n^2)/C \right\}.$$

We finally consider the integral on the r.h.s. of (37) for which we have

$$\begin{aligned} \int_{\varepsilon_n}^{y_n} \psi'(s) \mathbb{P}[X > s] ds &= h_n \int_{\varepsilon_n}^{y_n} \exp\{h_n s - R(s)\} ds \\ &\leq C \cdot n \int_{\delta+o(1)}^1 \exp\{h_n y_n s - R(y_n s)\} ds \\ &\leq C \cdot n \int_{\delta+o(1)}^1 \exp\{s(h_n y_n - R(y_n))\} ds, \end{aligned}$$

where in the last line we used the fact that  $R(sy_n) \sim s^r R(y_n)$  locally uniformly in  $s \in (0, \infty)$ . We see that if we plug in (42) into the integral and use the fact that  $s$  is bounded away from 0 we will arrive at

$$\int_{\varepsilon_n}^{y_n} \psi'(s) \mathbb{P}[X > s] ds \leq \exp \left\{ -(\log n)^2 (1 + nR(d_n)^2/d_n^2)/C \right\}$$

which in turn yields our claim.  $\square$

**Lemma A.4.** *Let Assumption 2 be in force for  $r > 2/3$ . Then for  $y_n = d_n - nR'(d_n)$  and  $h_n = R(d_n)/d_n$  the following inequality holds true*

$$\mathbb{E} \left[ e^{h_n X} \mathbf{1}_{\{X \in (\varepsilon_n, y_n)\}} \right] \leq \exp \left\{ \frac{nR(d_n)R'(d_n)}{d_n} \left( \frac{R'(d_n)d_n}{R(d_n)} - 1 \right) (1 + o(1)) \right\}.$$

**Proof.** The argument goes along the exact same lines as for Lemma A.3. Once again apply formula (37) with  $\psi(s) = e^{h_n s}$ ,  $A = \varepsilon_n$  and  $B = y_n$ . The estimate (41) is still sufficient for our purposes. To estimate the two other terms note that

$$R(y_n) = R(d_n) - nR'(d_n)^2(1 + o(1)),$$

which implies

$$h_n y_n - R(y_n) = \frac{nR(d_n)R'(d_n)}{d_n} \left( \frac{R'(d_n)d_n}{R(d_n)} - 1 \right) (1 + o(1)).$$

Using the same arguments as in the previous lemma

$$\psi(y_n) \mathbb{P}[X > y_n] + \int_{\varepsilon_n}^{y_n} \psi'(s) \mathbb{P}[X > s] ds \leq \exp \left\{ \frac{nR(d_n)R'(d_n)}{d_n} \left( \frac{R'(d_n)d_n}{R(d_n)} - 1 \right) (1 + o(1)) \right\}.$$

$\square$

## Acknowledgments

We thank two anonymous referees for reading very carefully the first version and suggesting numerous improvements.

## Funding

The research of PD was partially supported by the National Science Centre, Poland (Sonata, grant number 2020/39/D/ST1/00258).

## References

- [1] K. B. Athreya and P. E. Ney, *Branching processes*, 1972.
- [2] E. Aidékon, *Convergence in law of the minimum of a branching random walk*, *Annals of Probability* **41** (2013), 1362–1426.
- [3] A. Bhattacharya, R. S. Hazra, and P. Roy, *Point process convergence for branching random walks with regularly varying steps*, *Annales de l’institut Henri Poincaré (B) Probability and Statistics* **53** (2017), 802–818.
- [4] J. D. Biggins, *The first- and last-birth problems for a multitype age-dependent branching process*, *Advances in Applied Probability* (1976).
- [5] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, 1987.
- [6] É. Brunet and B. Derrida, *A branching random walk seen from the tip*, *Journal of Statistical Physics* **143** (2011), no. 3, 420–446.
- [7] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 1998.
- [8] D. Denisov, A. B. Dieker, and V. Shneer, *Large deviations for random walks under subexponentiality: The big-jump domain*, *Annals of Probability* **36** (2008), 1946–1991.
- [9] R. Durrett, *Maxima of branching random walks*, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **62** (1983), 165–170.
- [10] P. Dyszewski, N. Gantert, and T. Höfelsauer, *The maximum of a branching walk with stretched exponential tails*, to appear in *Annales de l’Institut Henri Poincaré (B) Probabilités et Statistiques*.
- [11] N. Gantert, *The maximum of a branching random walk with semiexponential increments*, *Annals of Probability* **28** (2000), 1219–1229.
- [12] J. M. Hammersley, *Postulates for subadditive processes*, *The Annals of Probability* **2** (1974), 652–680.
- [13] J. F. C. Kingman, *The first birth problem for an age-dependent branching process*, *The Annals of Probability* **3** (1975), 790–801.
- [14] V. Y. Linnik, *Limit theorems for sums of independent variables taking into account large deviations. i*, *Theory of Probability & Its Applications* **6** (1961), no. 2, 131–148.
- [15] T. Madaule, *Convergence in law for the branching random walk seen from its tip*, *Journal of Theoretical Probability* **30** (2017), 27–63.
- [16] Pascal Maillard, *A note on stable point processes occurring in branching Brownian motion*, *Electronic Communications in Probability* **18** (2013), no. none, 1 – 9.
- [17] A. V. Nagaev, *Integral limit theorems taking large deviations into account when Cramér’s condition does not hold. i*, *Theory of Probability & Its Applications* **14** (1969), 51–64.
- [18] ———, *Integral limit theorems taking large deviations into account when Cramér’s condition does not hold. ii*, *Theory of Probability & Its Applications* **14** (1969), 193–208.
- [19] S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer New York, New York, NY, 1987.
- [20] L. V. Rozovskii, *Probabilities of large deviations on the whole axis*, *Theory of Probability & Its Applications* **38** (1994), no. 1, 53–79.
- [21] Z. Shi, *Branching random walks*, Springer, 2015.
- [22] P. J. Smith, *A recursive formulation of the old problem of obtaining moments from cumulants and vice versa*, *American Statistician* **49** (1995), 217–218.
- [23] E. Subag and O. Zeitouni, *Freezing and decorated Poisson point processes*, *Communications in Mathematical Physics* **337** (2015), no. 1, 55–92.