Tail asymptotics for the supercritical Galton-Watson process in the heavy-tailed case

Vitali Wachtel, D. E. Denisov, Dmitry A. Korshunov

Angaben zur Veröffentlichung / Publication details:

Wachtel, Vitali, D. E. Denisov, and Dmitry A. Korshunov. 2013. "Tail asymptotics for the supercritical Galton-Watson process in the heavy-tailed case." *Proceedings of the Steklov Institute of Mathematics* 282 (1): 273–97. https://doi.org/10.1134/s0081543813060205.

Tail Asymptotics for the Supercritical Galton–Watson Process in the Heavy-Tailed Case

V. I. Wachtel[a], D. E. Denisov[b], and D. A. Korshunov[c]

Abstract—As is well known, for a supercritical Galton–Watson process Z_n whose offspring distribution has mean $m > 1$, the ratio $W_n := Z_n/m^n$ has almost surely a limit, say W. We study the tail behaviour of the distributions of W_n and W in the case where Z_1 has a heavy-tailed distribution, that is, $\mathbb{E}\, e^{\lambda Z_1} = \infty$ for every $\lambda > 0$. We show how different types of distributions of Z_1 lead to different asymptotic behaviour of the tail of W_n and W. We describe the most likely way in which large values of the process occur.

1. INTRODUCTION

Let Z_n be a supercritical Galton–Watson process with $Z_0 = 1$ and $m := \mathbb{E}\, Z_1 > 1$. By definition,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n)},$$

where $\xi_i^{(n)}$, $i, n = 0, 1, \ldots$, are independent identically distributed random variables with distribution F on $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$; by $\overline{F}(x)$ we denote the tail of F, $\overline{F}(x) := \mathbb{P}\{\xi > x\}$.

Put $W_n := Z_n/m^n$. As is well known (see, e.g., [2, Theorem 1.6.1]), $W_n \to W$ a.s. as $n \to \infty$. If $\mathbb{E}\, \xi \log \xi < \infty$, then $\mathbb{E}\, W = 1$, so $\mathbb{P}\{W > 0\} > 0$ (see [2, Theorem 1.10.1]).

Our goal is to consider asymptotic probabilities of large deviations for the martingale sequence $\{W_n\}$ and for its limit W. More precisely, we are going to find asymptotics for $\mathbb{P}\{W_n > x\}$ as $x \to \infty$ in the whole range of $n \ge 1$.

The tail behaviour of the martingale limit is one of the classical problems in the theory of supercritical Galton–Watson processes. The study of $\mathbb{P}\{W > x\}$ was initiated by Harris [14], who showed that if ξ is bounded, then

$$\log \mathbb{E}\, e^{uW} = u^\gamma H(u) + O(1) \qquad \text{as} \quad u \to \infty,$$

where H is a positive multiplicatively periodic function and γ is defined by the equality $m^\gamma = \max\{k \colon \mathbb{P}\{\xi = k\} > 0\}$. This information on the generating function can be translated into asymptotics of tail probabilities. It was done by Biggins and Bingham [4]:

$$\log \mathbb{P}\{W > x\} \sim -x^{\gamma/(\gamma-1)} M(x), \tag{1.1}$$

where M is also a positive multiplicatively periodic function; hereinafter we write $f(x) \sim g(x)$ as $x \to \infty$ if $f(x)/g(x) \to 1$. Bingham and Doney [5, 6] found asymptotics for $\mathbb{P}\{W > x\}$ in

[a] Mathematisches Institut, Ludwig-Maximilians-Universität München, Theresienstr. 39, D-80333 München, Germany.

[b] School of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, UK.

[c] Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, pr. Akademika Koptyuga 4, Novosibirsk, 630090 Russia.

the case when ξ is regularly varying with noninteger index $\alpha < -1$ (for the case of integer α see De Meyer [8]). In [4] one can find results similar to (1.1) for the left tail of W in the case when the minimum offspring size is at least 2. Fleischmann and Wachtel [11, 12] found exact (without logarithmic scaling) asymptotics for $\mathbb{P}\{W_n \in (0,x)\}$ and $\mathbb{P}\{W \in (0,x)\}$ as $x \to 0$. These two papers give a complete description of the asymptotic behaviour of the left tail of W. It is possible to adapt the method from [12] to upper deviation problems for processes with polynomial offspring generating functions. As a result one gets exact asymptotics for $\mathbb{P}\{W > x\}$ as $x \to \infty$ (see [12, Remark 3]).

In all the papers mentioned above, the proofs were based on the fact that $\varphi(u) := \mathbb{E}\, e^{-uW}$ satisfies the Poincaré functional equation, $\varphi(mu) = f(\varphi(u))$, where f stands for the offspring generating function. In the present paper we do not use that equation. Instead, we apply recently developed probabilistic techniques for sums of independent identically distributed random variables and for Galton–Watson processes with heavy tails.

We work with the following classes of distributions.

The distribution of a random variable ξ is called *heavy-tailed* if $\mathbb{E}\, e^{\lambda\xi} = \infty$ for every $\lambda > 0$.

We say that a distribution F on $\mathbb{R}$ is *dominated varying*, and write $F \in \mathcal{D}$, if

$$\sup_x \frac{\overline{F}(x/2)}{\overline{F}(x)} < \infty. \tag{1.2}$$

A distribution F on $\mathbb{R}$ is called *intermediate regularly varying* if

$$\lim_{\varepsilon\downarrow 0} \liminf_{x\to\infty} \frac{\overline{F}(x(1+\varepsilon))}{\overline{F}(x)} = 1.$$

Note that any regularly varying distribution is intermediate regularly varying. Any intermediate regularly varying distribution is dominated varying.

For any positive function $h(x) \to \infty$, we say that F is *h-insensitive* if $\overline{F}(x+h(x)) \sim \overline{F}(x)$ as $x \to \infty$. A distribution F is intermediate regularly varying if and only if F is h-insensitive for any positive function h such that $h(x) = o(x)$ as $x \to \infty$; in other words, if F is $o(x)$-insensitive (see [13, Theorem 2.47]).

We say that a distribution F on $\mathbb{R}^+$ with mean m is *strongly subexponential*, and write $F \in \mathcal{S}^*$, if

$$\int_0^x \overline{F}(x-y)\overline{F}(y)\,dy \sim 2m\overline{F}(x) \qquad \text{as} \quad x \to \infty.$$

Among strongly subexponential distributions are intermediate regularly varying, log-normal and Weibull distributions with parameter $\beta < 1$. Any dominated varying distribution is in $\mathcal{S}^*$ if it is long-tailed, that is, constant-insensitive.

A distribution F is called *rapidly varying* if, for any $\varepsilon > 0$,

$$\overline{F}(x(1+\varepsilon)) = o(\overline{F}(x)) \qquad \text{as} \quad x \to \infty.$$

Clearly, this class includes the Weibull distributions $\overline{F}(x) = e^{-x^\beta}$ with parameter $\beta > 0$. The log-normal distribution is also rapidly varying. This class does not include intermediate regularly varying distributions.

Theorem 1. *Let F be a dominated varying distribution such that, for some $\delta > 0$ and $c < \infty$,*

$$\overline{F}(xy) \le \frac{c\overline{F}(x)}{y^{1+\delta}} \qquad \textit{for all} \quad x,\ y > 1. \tag{1.3}$$

Then there exist constants $c_1 > 0$ and $c_2 < \infty$ such that

$$c_1\overline{F}(x) \le \mathbb{P}\{W_n > x\} \le c_2\overline{F}(x) \qquad \textit{for all} \quad x,\ n. \tag{1.4}$$

If, in addition, F is an intermediate regularly varying distribution, then, uniformly in n,

$$\mathbb{P}\{W_n > x\} \sim \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \qquad as \quad x \to \infty. \tag{1.5}$$

In particular,

$$\mathbb{P}\{W_n > x\} \sim \sum_{i=0}^{\infty} m^i \overline{F}(m^{i+1}x) \qquad as \quad x, n \to \infty \tag{1.6}$$

and

$$\mathbb{P}\{W > x\} \sim \sum_{i=0}^{\infty} m^i \overline{F}(m^{i+1}x) \qquad as \quad x \to \infty. \tag{1.7}$$

As follows from the proof of Lemma 9 (see Section 3 below),

$$\mathbb{P}\Big\{\max_{i \le Z_k} \xi_i^{(k)} \ge m^{k+1}x\Big\} \sim m^k \overline{F}(m^{k+1}x) \qquad \text{as} \quad x \to \infty$$

and the term $m^k \overline{F}(m^{k+1}x)$ in (1.5)–(1.7) describes the probability of the existence of a very productive particle in the kth generation. We can informally restate (1.5)–(1.7) as follows:

$$\{W_n > x\} \approx \bigcup_{k=0}^{n-1} \Big\{\max_{i \le Z_k} \xi_i^{(k)} \ge m^{k+1}x\Big\} \qquad \text{and} \qquad \{W > x\} \approx \bigcup_{k=0}^{\infty} \Big\{\max_{i \le Z_k} \xi_i^{(k)} \ge m^{k+1}x\Big\}.$$

Moreover, if $\overline{F}(x)$ is regularly varying with index $\alpha < -1$, then, uniformly in n,

$$\mathbb{P}\Big\{\max_{i \le Z_k} \xi_i^{(k)} \ge m^{k+1}x \,\Big|\, W_n > x\Big\} \to \frac{m^{-(\alpha-1)k}}{\sum_{j=0}^{n-1} m^{-(\alpha-1)j}} \qquad \text{as} \quad x \to \infty.$$

In the limit $n \to \infty$ we get the geometric distribution with the parameter $m^{-(\alpha-1)}$. Therefore, atypically large values of the limit W are caused by a very productive particle which lives in one of the initial generations, and the number of this generation is random with the geometric distribution mentioned above.

If we assume the second moment of ξ to be finite, then we may relax the regularity condition on F; namely, we may consider distributions which are not necessarily intermediate regularly varying as was assumed in Theorem 1.

Theorem 2. *Let F be a dominated varying distribution and condition* (1.3) *hold. If $\mathbb{E}\,\xi^2 < \infty$ and F is an x^γ-insensitive distribution for some $\gamma > 1/2$, then the asymptotics* (1.5)–(1.7) *hold.*

We next turn to the case of Weibull-type offspring distributions.

Theorem 3. *Let $\overline{F}(x) = e^{-R(x)}$ where $R(x)$ is regularly varying with index $\beta \in (0, 1)$. Additionally assume that $F \in \mathcal{S}^*$. Then, for every $\varepsilon > 0$,*

$$(1 + o(1))\overline{F}((m + \varepsilon)x) \le \mathbb{P}\{W_n > x\} \le (1 + o(1))\overline{F}((m - \varepsilon)x)$$

as $x \to \infty$ uniformly in n.

If $\beta < (3 - \sqrt{5})/2 \approx 0.382$, then $\mathbb{P}\{W_n > x\} \sim \overline{F}(mx)$ as $x \to \infty$ uniformly in n and $\mathbb{P}\{W > x\} \sim \overline{F}(mx)$ as $x \to \infty$.

If $\beta < 1/2$ and, in addition, for some $c_1 < \infty$,

$$R(k) - R(k-1) \le c_1 \frac{R(k)}{k}, \qquad k \ge 1, \tag{1.8}$$

then $\mathbb{P}\{W_n > x\} \sim \mathbb{P}\{W > x\} \sim \overline{F}(mx)$ as $x \to \infty$ uniformly in n.

Let us make a remark on Weibull-type offspring distributions which are not $\sqrt{x}$-insensitive. If $\mathbb{P}\{\xi > x\} \sim e^{-x^\beta}$ with some $\beta \in (1/2, 1)$, then

$$\mathbb{P}\{W_n > x\} \geq \exp\biggl\{-(mx)^\beta + \frac{\beta^2\sigma_n^2}{2}(mx)^{2\beta-1}(1+o(1))\biggr\}, \qquad n \geq 2, \tag{1.9}$$

and

$$\mathbb{P}\{W > x\} \geq \exp\biggl\{-(mx)^\beta + \frac{\beta^2\sigma^2}{2}(mx)^{2\beta-1}(1+o(1))\biggr\}. \tag{1.10}$$

Here $\sigma_n^2 := \mathbb{E}(W_n - 1)^2$ and $\sigma^2 := \mathbb{E}(W-1)^2$. These bounds imply that, in contrast to the case $\beta < 1/2$, $\mathbb{P}\{W_n > x\} \gg \overline{F}(mx)$ for all $n \geq 2$. At the end of Section 3 we give arguments for (1.10).

In Theorem 3 we have, uniformly in n,

$$\{W_n > x\} \approx \{\xi_1^{(0)} > mx\}.$$

Thus, large values of all W_n are caused by a correspondingly large first generation.

The importance of initial generations for deviation probabilities can be explained by the multiplicative structure of supercritical Galton–Watson processes. As a consequence of this fact, it is "cheaper" to have some special type of behaviour at the very beginning of the process. In Theorems 1 and 3 we see a very strong time localization: only a few first generations are important. There are some examples in the literature where a weaker form of localization occurs. In the case of lower deviations which were studied in [11, 12], the optimal strategy looks as follows: In order for $\{Z_n = k_n\}$ to hold with some $k_n = o(m^n)$, every particle in the first a_n generations should have exactly $\mu := \min\{k\colon \mathbb{P}\{\xi = k\} > 0\}$ children. (Here we assume for simplicity that $\xi \geq 1$.) In all later generations we let Z_k grow without any restriction, i.e., geometrically with the rate m. Since we want to get k_n particles in the nth generation, a_n should satisfy $\mu_n^a m^{n-a_n} \approx k_n$. Recalling that $k_n = o(m^n)$, we see that the number of generations with nontypical behaviour tends to infinity. A similar strategy is behind asymptotics for $\mathbb{P}\{W < \varepsilon\}$ as $\varepsilon \to 0$ and behind asymptotics for upper deviations of processes with polynomial generating functions. This localization effect for Galton–Watson processes with vanishing limit, that is, Z_n conditioned on $\{W < \varepsilon\}$ with $\varepsilon \to 0$, was recently studied by Berestycki, Gantert, Mörters and Sidorova [3]. They showed that the genealogical tree coincides up to a certain generation with the regular μ-ary tree.

It turns out that this type of optimal strategies is not universal for supercritical Galton–Watson processes. The next result shows that if the offspring distribution has only the first power moment, then large values of W_n and W can be produced by the middle part of the tree.

Theorem 4. *Assume that* $\mathbb{E}\,\xi\log\xi < \infty$ *and* $\overline{F}(x)$ *is regularly varying with index* -1. *Then, uniformly in* $n \geq 1$,

$$\mathbb{P}\{W_n > x\} \sim \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \sim \frac{1}{m\log m}x^{-1}\int_x^{m^n x} \overline{F}(u)\,du \qquad \text{as} \quad x \to \infty. \tag{1.11}$$

For the limit W *we have*

$$\mathbb{P}\{W > x\} \sim \sum_{i=0}^{\infty} m^i \overline{F}(m^{i+1}x) \sim \frac{1}{m\log m}x^{-1}\int_x^{\infty} \overline{F}(u)\,du \qquad \text{as} \quad x \to \infty. \tag{1.12}$$

Relation (1.12) is a refinement of Theorem 1.4 in [5], which states that if $\mathbb{E}\{Z_1;\ Z_1 > x\} \sim L(x)$ for some slowly varying function L satisfying $\int_1^\infty (L(x)/x)\,dx < \infty$, then

$$\mathbb{E}\{W;\ W > x\} \sim \frac{1}{m\log m}\int_x^\infty \frac{L(y)}{y}\,dy.$$

Noting that $\overline{F}(x) = o\big(x^{-1}\int_x^\infty \overline{F}(u)\,du\big)$, we conclude from Theorem 4 that, for every $N \ge 1$,

$$\sum_{i=0}^{N} m^i \overline{F}(m^{i+1}x) = o\big(\mathbb{P}(W > x)\big) \qquad \text{as} \quad x \to \infty.$$

This means that "big jumps" in any fixed number of generations do not affect large values of W. Furthermore, the main contribution to $\sum_{i=0}^\infty m^i \overline{F}(m^{i+1}x)$ (and therefore to $\mathbb{P}\{W > x\}$) comes from indices i such that the ratio $\int_{m^i x}^\infty \overline{F}(u)\,du / \int_x^\infty \overline{F}(u)\,du$ is bounded away from 0 and 1. For finite values of n we have three different regimes depending on the relation between n and x. We illustrate them by the following example.

Example. Assume that $\overline{F}(x) \sim x^{-1}\log^{-p-1} x$ with some $p > 1$. Then

$$L(x) := \int_x^\infty \overline{F}(y)\,dy \sim \frac{1}{p}\log^{-p} x \qquad \text{as} \quad x \to \infty.$$

Therefore,

$$\mathbb{P}\{W > x\} \sim \frac{1}{m\log m} x^{-1} L(x) \sim \frac{1}{pm\log m} x^{-1}\log^{-p} x. \tag{1.13}$$

Consider now finite values of n. First, if n and x are such that $n/\log x \to \infty$, then, according to (1.11),

$$\mathbb{P}\{W_n > x\} \sim \frac{1}{m\log m} x^{-1}\big(L(x) - L(m^n x)\big) \sim \frac{1}{m\log m} x^{-1} L(x) \sim \frac{1}{pm\log m} x^{-1}\log^{-p} x.$$

Comparing this with (1.13), we see that the asymptotics of $\mathbb{P}\{W_n > x\}$ and $\mathbb{P}\{W > x\}$ are equal in this case.

Second, if n and x are such that $n/\log x \to t \in (0,\infty)$, then

$$L(m^n x) \sim \frac{1}{p}(\log x + n\log m)^{-p} \sim \frac{1}{p}\log^{-p} x(1 + t\log m)^{-p}.$$

Consequently,

$$\mathbb{P}\{W_n > x\} \sim \frac{1}{pm\log m} x^{-1}\log^{-p} x\big(1 - (1 + t\log m)^{-p}\big).$$

Here we see that $\mathbb{P}\{W_n > x\}$ and $\mathbb{P}\{W > x\}$ are still of the same order, but the constants are different.

Third, if $n/\log x \to 0$, then, noting that $\log y \sim \log x$ uniformly in $y \in [x, m^n x]$, we have

$$\mathbb{P}\{W_n > x\} \sim \frac{1}{m\log m} x^{-1}\int_x^{m^n x} \frac{dy}{y\log^{p+1} y} \sim \frac{1}{m\log m}\frac{1}{x\log^{p+1} x}\int_x^{m^n x} \frac{dy}{y} \sim n\overline{F}(mx).$$

Therefore, $\mathbb{P}\{W_n > x\}$ is much smaller than $\mathbb{P}\{W > x\}$ for these values of n.

The problem of describing tail asymptotics for supercritical Galton–Watson process is closely related to the problem of tail behaviour for a randomly stopped sum S_τ where the random number τ of summands has the same distribution as the summands ξ themselves. For random sums, the only well-studied case is when the distribution tail of τ is much lighter than that of ξ (see [10]); in this case the typical answer is $\mathbb{P}\{S_\tau > x\} \sim \mathbb{E}\,\tau\mathbb{P}\{\xi > x\}$ as $x \to \infty$. The present study may be considered as a step towards general problem for randomly stopped sums where the tails of the stopping time τ and of the summand ξ are comparable.

The rest of the paper is organized as follows. Section 2 is devoted to related upper bounds for the distribution tails of sums with zero drift in the large deviation zone. Later on in Section 4 we use them to derive upper bounds for $\mathbb{P}\{W_n > x\}$; more precisely, we reduce the problem of finding the asymptotic behaviour of $\mathbb{P}\{W_n > x\}$ to that for $\mathbb{P}\{W_N > x\}$ with some fixed N. Also, upper bounds of Section 2 help to compute the asymptotics for $\mathbb{P}\{W_N > x\}$ for every fixed N. Lower bounds for the distribution tail of the number of descendants in the nth generation are given in Section 3. In Section 6 we provide final proofs of Theorems 1, 2 and 3. Finally, for Theorem 4, in which only the first moment is finite, our approach based on describing and computing the most likely events that lead to large deviations of W_n does not work. Here we propose an analytic method adapted from [17] (see Section 7).

2. PRELIMINARY RESULTS FOR SUMS

We repeatedly make use of the following result which is a version of Theorem 2(i) in [10] with exactly the same proof. In what follows $\eta_1, \eta_2, \ldots$ are independent random variables with common distribution G and $T_n := \eta_1 + \ldots + \eta_n$.

Proposition 5. *Let the distribution G have negative mean $a := \mathbb{E}\,\eta_1 < 0$. If $G \in \mathcal{S}^*$, then*

$$\mathbb{P}\{T_n > x\} \le (1 + o(1))n\overline{G}(x)$$

as $x \to \infty$ uniformly in n.

This proposition helps to deduce exact asymptotics for $\mathbb{P}\{T_n > x\}$ in the case of zero mean if $x/n > c > 0$. If $x = o(n)$, then Proposition 5 is not useful for estimation of $\mathbb{P}\{T_n > x\}$ in the case of zero mean. So, in the following two propositions we derive rough upper bounds for the large deviation probabilities for sums with zero mean; these rough bounds will be appropriate for our purposes. The first proposition is devoted to distributions of regularly varying type, while the second one is devoted to Weibull-type distributions. Deriving rather rough bounds, we relax conditions on the distribution of jumps as compared to the asymptotic results of [9, Theorems 8.1, 8.3] and [7, Theorems 3.1.1, 4.1.2, 5.2.1].

Proposition 6. *Let $\mathbb{E}\,\eta_1 = 0$, $\mathbb{E}\{\eta_1^2;\ \eta_1 \le 0\} < \infty$ and G be a dominated varying distribution. If, for some $\delta \in (0, 1)$,*

$$\mathbb{E}\big\{\eta^{1+\delta};\ \eta > 0\big\} < \infty, \tag{2.1}$$

then, for every $\delta' \in (0, \delta)$, there exists a $c < \infty$ such that $\mathbb{P}\{T_n > x\} \le cn\overline{G}(x)$ for all $x > 0$ and $n \le x^{1+\delta'}$.

If, for some $\delta > 0$,

$$\mathbb{E}\big\{\eta^{2+\delta};\ \eta > 0\big\} < \infty, \tag{2.2}$$

then there exists a $c < \infty$ such that $\mathbb{P}\{T_n > x\} \le cn\overline{G}(x)$ for all $x > 0$ and $n \le x^2/(c\log x)$ (or, equivalently, $x \ge c\sqrt{n \log n}$).

Proof. Let $R(x)$ be the hazard function for G, that is, $\overline{G}(x) = e^{-R(x)}$. First we prove that dominated variation yields, for some $C < \infty$, the upper bound

$$R(x) \le C + C\log x, \qquad x \ge 1. \tag{2.3}$$

Indeed, there exists a $c < \infty$ such that $\overline{G}(x/2) \le e^c\overline{G}(x)$ for all x. Equivalently, $R(x/2) \ge R(x) - c$, which implies $R(x2^{-n}) \ge R(x) - cn$. For $n(x) := [\log_2 x] + 1$ we get

$$R(1) \ge R\big(x2^{-n(x)}\big) \ge R(x) - cn(x) = R(x) - c\log_2 x - c,$$

and the upper bound (2.3) follows.

For every $y < x$, we may estimate the tail distribution of the sum as follows:

$$\begin{aligned}\mathbb{P}\{T_n > x\} &\le \mathbb{P}\{T_n > x,\ \eta_k > y \text{ for some } k \le n\} + \mathbb{P}\{T_n > x,\ \eta_k \le y \text{ for all } k \le n\} \\ &\le n\overline{G}(y) + e^{-\lambda x}\big(\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\}\big)^n \qquad (2.4)\end{aligned}$$

for every $\lambda > 0$ by the exponential Chebyshev inequality. Fix an $\varepsilon \in (0,1)$. Take $y := \varepsilon x$ and $\lambda := 2R(x)/x$. Then $e^{-\lambda x} = \overline{G}(x)e^{-R(x)}$ and

$$\mathbb{P}\{T_n > x\} \le n\overline{G}(y) + \overline{G}(x)e^{-R(x)}\big(\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\}\big)^n.$$

Let us estimate the latter truncated exponential moment:

$$\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\} = \mathbb{E}\Big\{e^{\lambda\eta_1};\ \eta_1 \le \frac{1}{\lambda}\Big\} + \mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} \le \eta_1 \le y\Big\}. \qquad (2.5)$$

Since $e^u \le 1 + u + 2u^2$ for all $u \le 1$,

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \eta_1 \le \frac{1}{\lambda}\Big\} \le 1 + \lambda\,\mathbb{E}\Big\{\eta_1;\ \eta_1 \le \frac{1}{\lambda}\Big\} + 2\lambda^2\,\mathbb{E}\Big\{\eta_1^2;\ \eta_1 \le \frac{1}{\lambda}\Big\} \le 1 + 2\lambda^2\,\mathbb{E}\Big\{\eta_1^2;\ \eta_1 \le \frac{1}{\lambda}\Big\} \qquad (2.6)$$

owing to the mean zero for η_1.

In the case of finite second moment we get

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \eta_1 \le \frac{1}{\lambda}\Big\} \le 1 + c_1\lambda^2. \qquad (2.7)$$

Further,

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} < \eta_1 \le y\Big\} \le e^{\lambda y}\overline{G}\Big(\frac{1}{\lambda}\Big) \le e^{\lambda y}\,\mathbb{E}\{\eta^{2+\delta};\ \eta > 0\}\lambda^{2+\delta}$$

by condition (2.2) and the Chebyshev inequality. Choose $\varepsilon > 0$ so small that $\varepsilon C < \delta/4$. Then the upper bound (2.3) yields, for some $c_2 < \infty$,

$$e^{\lambda y} = e^{\varepsilon x 2R(x)/x} \le c_2 x^{\delta/2}$$

and consequently

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} < \eta_1 \le y\Big\} \le c_3\lambda^2. \qquad (2.8)$$

Together with (2.7) it implies that

$$\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\} \le 1 + \frac{c_4 R^2(x)}{x^2} \le e^{c_4 R^2(x)/x^2}$$

for some $c_4 < \infty$. Hence,

$$\mathbb{P}\{T_n > x\} \le n\overline{G}(y) + \overline{G}(x)e^{-R(x)}e^{c_4 n R^2(x)/x^2} \le n\overline{G}(y) + \overline{G}(x)e^{-R(x)+R(x)(c_5 n\log x/x^2)}$$

for some $c_5 < \infty$, due to (2.3). So, in the case of finite $2+\delta$ moment, the proposition follows for $n \le x^2/(c_5\log x)$ if we take into account (1.2).

In the case where condition (2.1) only holds,

$$\mathbb{E}\Big\{\eta_1^2;\ \eta_1 \le \frac{1}{\lambda}\Big\} \le \mathbb{E}\{\eta_1^2;\ \eta_1 \le 0\} + \frac{\mathbb{E}\{\eta_1^{1+\delta}; \eta_1 > 0\}}{\lambda^{1-\delta}}$$

and we deduce from estimate (2.6) that

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \eta_1 \le \frac{1}{\lambda}\Big\} \le 1 + c_6\lambda^{1+\delta}.$$

Similar to (2.8),

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} \le \eta_1 \le y\Big\} \le \frac{c_7}{x^{1+\delta'}}$$

by condition (2.1). Then

$$\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\} \le 1 + \frac{c_8}{x^{1+\delta'}} \le e^{c_8/x^{1+\delta'}},$$

because $R(x) \le c_9 \log x$ by (2.3). Hence,

$$\mathbb{P}\{T_n > x\} \le n\overline{G}(y) + \overline{G}(x)e^{-R(x)}e^{c_8 n/x^{1+\delta'}},$$

and the case of finite first moment follows. □

Proposition 7. *Let the distribution G have mean zero, $\mathbb{E}\,\eta_1 = 0$, and all moments be finite, $\mathbb{E}\,|\eta_1|^k < \infty$, $k = 1, 2, \dots$. Let $R(x)$ be the hazard function for G, that is, $\overline{G}(x) = e^{-R(x)}$. Suppose, for every $\varepsilon > 0$, there exists an x_0 such that*

$$\frac{R(x)}{x} \le \frac{(1+\varepsilon)R(z)}{z} \qquad \text{for all} \quad x \ge z \ge x_0. \tag{2.9}$$

Then, for every $0 < \varepsilon < 1$, there exists a $c = c(\varepsilon) < \infty$ such that

$$\mathbb{P}\{T_n > x\} \le (n+1)\overline{G}(y)$$

for all $x > 0$, $y \le (1-\varepsilon)x$ and n such that $nR(y)/x^2 \le 1/c$.

Proof. Take $\lambda := (1+\varepsilon)R(y)/x$. Then $e^{-\lambda x} = e^{-(1+\varepsilon)R(y)}$.

By condition (2.9),

$$\lambda z = (1+\varepsilon)\frac{R(y)}{y}\frac{y}{x}z \le (1-\varepsilon^2)\frac{R(y)}{y}z \le \Big(1 - \frac{\varepsilon^2}{2}\Big)R(z)$$

for all $z \le y$ sufficiently large. Therefore,

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} < \eta_1 \le y\Big\} \le \mathbb{E}\Big\{e^{(1-\varepsilon^2/2)R(\eta_1)};\ \frac{1}{\lambda} < \eta_1\Big\} \le -\int_{1/\lambda}^{\infty} e^{(1-\varepsilon^2/2)R(z)}\, de^{-R(z)}$$

$$= \int_{1/\lambda}^{\infty} e^{-\varepsilon^2 R(z)/2}\, dR(z) = \frac{2e^{-\varepsilon^2 R(1/\lambda)/2}}{\varepsilon^2}.$$

Taking into account that, for every $\alpha > 0$, $e^{-R(x)} = o(1/x^\alpha)$ as $x \to \infty$, we get

$$\mathbb{E}\Big\{e^{\lambda\eta_1};\ \frac{1}{\lambda} < \eta_1 \le y\Big\} = o(\lambda^2) \qquad \text{as} \quad y \to \infty. \tag{2.10}$$

Substituting (2.7) and (2.10) into (2.5), we obtain the inequality

$$\mathbb{E}\{e^{\lambda\eta_1};\ \eta_1 \le y\} \le 1 + \frac{cR^2(y)}{x^2} \le e^{cR^2(y)/x^2}$$

for some $c < \infty$. Hence,

$$\mathbb{P}\{T_n > x\} \le n\overline{G}(y) + e^{-(1+\varepsilon)R(y)} e^{c_1 n R^2(y)/x^2} \le n\overline{G}(y) + e^{-R(y)} = (n+1)\overline{G}(y)$$

in the range where $c_1 nR(y)/x^2 \le \varepsilon$, and the proof of the desired upper bound is complete. □

In the proof above the distribution G restricted to $(-\infty, 1/\lambda]$ comes into the upper bound through its second moment only. The tail of G influences the upper bound through its values to the right of the point $1/\lambda$. Having this observation in mind, we formulate the following uniform version of the previous proposition for a family of distributions whose tails are ultimately dominated by that of G.

Corollary 8. *Let all the conditions of Proposition 7 be fulfilled. Let $G^{(v)}$ be a family of distributions that depend on some parameter $v \in V$ and are such that, for some x_1, $\overline{G^{(v)}}(x) \le \overline{G}(x)$ for all $x > x_1$ and $v \in V$. Let every $G^{(v)}$ have mean zero, and let all the second moments be bounded. Then, for every $0 < \varepsilon < 1$, there exists a $c = c(\varepsilon) < \infty$ such that*

$$\overline{(G^{(v)})^{*n}}(x) \le (n+1)\overline{G^{(v)}}(y)$$

for all $v \in V$, $x > 0$, $y \le (1-\varepsilon)x$ and n such that $nR(y)/x^2 \le 1/c$.

3. LOWER BOUNDS

Lemma 9. *Let $\mathbb{E}\,\xi \log \xi < \infty$. Then, for every $\varepsilon > 0$,*

$$\mathbb{P}\{W_n > x\} \ge (1+o(1)) \sum_{i=0}^{n-1} m^i \overline{F}\big(m^{i+1}(1+\varepsilon)x\big)$$

as $x \to \infty$ uniformly in $n \ge 1$.

Proof. Consider the following decreasing sequence of events:

$$B_k(x) := \big\{Z_j \le m^j x \ \text{ for all } \ j \le k\big\}.$$

Since $Z_j/m^j \to W$ a.s. as $j \to \infty$,

$$\inf_{k \ge 1} \mathbb{P}\{B_k(x)\} \to 1 \qquad \text{as} \quad x \to \infty. \tag{3.1}$$

The events

$$A_k(x) := \big\{B_k(x),\ \xi_i^{(k)} > m^{k+1}(1+\varepsilon)x \ \text{ for some } \ i \le Z_k\big\}$$

are disjoint, which implies the lower bound

$$\mathbb{P}\{W_n > x\} \ge \sum_{k=0}^{n-1} \mathbb{P}\big\{Z_n > m^n x \mid A_k(x)\big\}\, \mathbb{P}\{A_k(x)\}. \tag{3.2}$$

First we estimate the probability

$$\begin{aligned}\mathbb{P}\{A_k(x)\} &= \sum_{j=0}^{m^k x} \mathbb{P}\{B_k(x),\ Z_k = j\}\, \mathbb{P}\big\{\xi_i^{(k)} > m^{k+1}(1+\varepsilon)x \ \text{ for some } \ i \le j\big\} \\ &= \sum_{j=0}^{m^k x} \mathbb{P}\{B_k(x),\ Z_k = j\}\big(1 - \big(1 - \overline{F}(m^{k+1}(1+\varepsilon)x)\big)^j\,\big).\end{aligned}$$

Since $\mathbb{E}\,\xi\log\xi<\infty$, by the Chebyshev inequality we have

$$\mathbb{P}\{\xi>m^{k+1}(1+\varepsilon)x\}\le\frac{\mathbb{E}\,\xi\log\xi}{m^{k+1}x\log x}=o\Big(\frac{1}{m^kx}\Big)\qquad\text{as}\quad x\to\infty\quad\text{uniformly in }\ k.$$

Hence,

$$\big(1-\overline{F}(m^{k+1}(1+\varepsilon)x)\big)^j=1-j\overline{F}(m^{k+1}(1+\varepsilon)x)(1+o(1))$$

as $x\to\infty$ uniformly in $k\ge 0$ and $j\le m^kx$. Therefore,

$$\mathbb{P}\{A_k(x)\}=(1+o(1))\overline{F}(m^{k+1}(1+\varepsilon)x)\sum_{j=0}^{m^kx}j\mathbb{P}\{B_k(x),\,Z_k=j\}$$

as $x\to\infty$ uniformly in $k\ge 0$. The Kesten–Stigum theorem (see, e.g., [1, Theorem 2.1]) states, in particular, that $\mathbb{E}\,\xi\log\xi<\infty$ if and only if the family of random variables $\{W_n,\,n\ge 0\}$ is uniformly integrable. Therefore, it follows from (3.1) that

$$\mathbb{E}\big\{W_k;\,B_k(x),\,W_k\le x\big\}\to 1\qquad\text{as}\quad x\to\infty\quad\text{uniformly in }\ k.$$

For this reason,

$$\sum_{j=0}^{m^kx}j\mathbb{P}\big\{B_k(x),\,Z_k=j\big\}=\mathbb{E}\big\{Z_k;\,B_k(x),\,Z_k\le m^kx\big\}=m^k\,\mathbb{E}\big\{W_k;\,B_k(x),\,W_k\le x\big\}\sim m^k$$

as $x\to\infty$ uniformly in $k\ge 0$. Thus, uniformly in $k\ge 0$,

$$\mathbb{P}\{A_k(x)\}=(1+o(1))m^k\overline{F}(m^{k+1}(1+\varepsilon)x)\qquad\text{as}\quad x\to\infty. \tag{3.3}$$

Second we prove that

$$\inf_{n\ge 1,\,k\le n-1}\mathbb{P}\big\{Z_n>m^n(1+\varepsilon)x\mid A_k(x)\big\}\to 1\qquad\text{as}\quad x\to\infty. \tag{3.4}$$

Indeed, by the Markov property,

$$\mathbb{P}\big\{Z_n>m^nx\mid A_k(x)\big\}\ge\mathbb{P}\Bigg\{\sum_{j=1}^{m^{k+1}(1+\varepsilon)x}Z_{n-k-1,j}>m^nx\Bigg\}=\mathbb{P}\Bigg\{\sum_{j=1}^{m^{k+1}(1+\varepsilon)x}W_{n-k-1,j}>m^{k+1}x\Bigg\},$$

where $Z_{n-k-1,j}$ are independent copies of Z_{n-k-1} and $W_{n-k-1,j}$ are independent copies of W_{n-k-1}. Since the family $\{W_n\}$ is uniformly integrable and $\mathbb{E}\,W_n=1$ for every n, we can apply the law of large numbers, which ensures that

$$\frac{1}{m^{k+1}x}\sum_{j=1}^{m^{k+1}(1+\varepsilon)x}W_{n-k-1,j}\xrightarrow{\text{p}}(1+\varepsilon)\,\mathbb{E}\,W_{n-k-1}=1+\varepsilon$$

as $x\to\infty$ uniformly in $n\ge 1$ and $k\le n-1$. Therefore,

$$\mathbb{P}\Bigg\{\sum_{j=1}^{m^{k+1}(1+\varepsilon)x}W_{n-k-1,j}>m^{k+1}x\Bigg\}\to 1,$$

which justifies the convergence (3.4). Substituting (3.3) and (3.4) into (3.2), we deduce the desired lower bound uniform in n. $\square$

Lemma 10. *Let the distribution F have the second moment finite, $\sigma^2 := \operatorname{Var}\xi_1 < \infty$. Then, for every $A > 0$,*

$$\mathbb{P}\{W_n > x\} \geq \left(1 - \frac{\sigma^2}{(m^2-m)A^2} + o(1)\right) \sum_{i=0}^{n-1} m^i \overline{F}\big(m^{i+1}x + A\sqrt{m^{i+1}x}\big) \tag{3.5}$$

as $x \to \infty$ uniformly in n.

In particular, if additionally the distribution F is $\sqrt{x}$-insensitive, then

$$\mathbb{P}\{W_n > x\} \geq (1+o(1)) \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \qquad \text{as} \quad x \to \infty \quad \text{uniformly in } n. \tag{3.6}$$

Proof. Let the events $B_k(x)$ be defined as above and

$$A_k(x) := \big\{B_k(x),\ \xi_i^{(k)} > m^{k+1}x + A\sqrt{m^{k+1}x} \text{ for some } i \leq Z_k\big\},$$

which are again disjoint, which implies the lower bound (3.2). The same calculations as in the previous proof lead to the following relation uniform in $k \geq 0$:

$$\mathbb{P}\{A_k(x)\} = (1+o(1)) m^k \overline{F}\big(m^{k+1}x + A\sqrt{m^{k+1}x}\big) \qquad \text{as} \quad x \to \infty. \tag{3.7}$$

Then it remains to prove that

$$\liminf_{x\to\infty} \inf_{n\geq 1,\, k\leq n-1} \mathbb{P}\big\{Z_n > m^n x \mid A_k(x)\big\} \geq 1 - \frac{\sigma^2}{(m^2-m)A^2}. \tag{3.8}$$

Indeed,

$$\begin{aligned}
\mathbb{P}\big\{Z_n > m^n x \mid A_k(x)\big\} &\geq \mathbb{P}\left\{\sum_{j=1}^{m^{k+1}x + A\sqrt{m^{k+1}x}} Z_{n-k-1,j} > m^n x\right\} \\
&= \mathbb{P}\left\{\sum_{j=1}^{m^{k+1}x + A\sqrt{m^{k+1}x}} W_{n-k-1,j} > m^{k+1} x\right\} \\
&= \mathbb{P}\left\{\sum_{j=1}^{m^{k+1}x + A\sqrt{m^{k+1}x}} (W_{n-k-1,j} - 1) > -A\sqrt{m^{k+1}x}\right\},
\end{aligned}$$

since $\mathbb{E}\, W_n = 1$. Applying the Chebyshev inequality, we deduce

$$\mathbb{P}\big\{Z_n > m^n x \mid A_k(x)\big\} \geq 1 - \frac{\operatorname{Var} W_{n-k-1}}{A^2} \frac{m^{k+1}x + A\sqrt{m^{k+1}x}}{m^{k+1}x} = 1 - \frac{\operatorname{Var} W_{n-k-1}}{A^2}(1+o(1))$$

as $x \to \infty$ uniformly in $n \geq 1$ and $k \leq n-1$. As calculated in [15, Theorem 1.5.1],

$$\operatorname{Var} W_n = \frac{\sigma^2(1-m^{-n})}{m^2-m} \uparrow \frac{\sigma^2}{m^2-m} = \operatorname{Var} W \qquad \text{as} \quad n \to \infty,$$

which completes the proof of (3.8). Substituting (3.7) and (3.8) into (3.2), we deduce the lower bound (3.5).

If F is $\sqrt{x}$-insensitive, then letting $A \to \infty$ we derive the second lower bound of the lemma. □

As clearly seen from the proof of Lemma 10, in the case of Weibull distribution with parameter $\beta \in (1/2,1)$ the tail of W_n is definitely heavier than $\overline{F}(mx)$. Now let us explain why the more accurate lower bound (1.10) given in the Introduction holds. Recalling that

$$W = \frac{1}{m}\sum_{i=1}^{\xi} W^{(i)},$$

where $W^{(i)}$ are independent copies of W that do not depend on ξ, we derive

$$\mathbb{P}\{W > x\} \geq \mathbb{P}\left\{\sum_{i=1}^{\xi} W^{(i)} > mx;\ \xi \geq N_x\right\} \geq \mathbb{P}\{\xi > N_x\}\,\mathbb{P}\left\{\sum_{i=1}^{N_x} W^{(i)} > mx\right\}, \tag{3.9}$$

where $N_x := [mx - z(mx)^\beta]$, $z > 0$. It is easy to see that

$$\mathbb{P}\{\xi > N_x\} \sim e^{-(mx - z(mx)^\beta)^\beta} = e^{-(mx)^\beta + \beta z(mx)^{2\beta-1} + O(x^{3\beta-2})}. \tag{3.10}$$

In view of log-scaled asymptotics for $\mathbb{P}\{W > x\}$ (see the first assertion of Theorem 3), we have $\mathbb{E}\, e^{(1-\varepsilon)m^\beta W^\beta} < \infty$ for every $\varepsilon > 0$. Moreover, $x^{2\beta} \ll N_x(x^\beta)^\beta$. Consequently, we may apply Nagaev's theorem [16, Theorem 3]:

$$\mathbb{P}\left\{\sum_{i=1}^{N_x} W^{(i)} > mx\right\} \geq \mathbb{P}\left\{\sum_{i=1}^{N_x}\big(W^{(i)} - 1\big) > z(mx)^\beta\right\} = \exp\left\{-\frac{z^2}{2\sigma^2}(mx)^{2\beta-1}(1+o(1))\right\}. \tag{3.11}$$

Combining (3.9)–(3.11), we get

$$\mathbb{P}\{W > x\} \geq \exp\left\{-(mx)^\beta + \left(\beta z - \frac{z^2}{2\sigma^2}\right)(mx)^{2\beta-1}(1+o(1))\right\}.$$

Maximizing $\beta z - z^2/2\sigma^2$, we obtain (1.10).

4. UPPER BOUNDS: REDUCTION TO A FINITE TIME HORIZON

Lemma 11. *Let the distribution F be dominated varying and satisfy condition* (1.3). *Then, for every $\varepsilon > 0$, there exists an N such that, for all $n > N$ and all sufficiently large x,*

$$\mathbb{P}\{W_n > x\} \leq (1+\varepsilon)\,\mathbb{P}\big\{W_N > (1-\varepsilon)x\big\}.$$

Proof. In order to derive this upper bound, we write, for $z < y$,

$$\mathbb{P}\left\{\sum_{i=1}^{Z_{n-1}} \xi_i > my\right\} \leq \mathbb{P}\{Z_{n-1} > z\} + \mathbb{P}\left\{\sum_{i=1}^{Z_{n-1}} \xi_i > my;\ Z_{n-1} \leq z\right\}, \tag{4.1}$$

where the ξ_i are independent of Z_{n-1}. It follows from Proposition 6 (under condition (2.1)) for sums with zero mean, $\eta_i = \xi_i - m$, that, for some $c < \infty$,

$$\mathbb{P}\left\{\sum_{i=1}^{k} \xi_i > my\right\} = \mathbb{P}\left\{\sum_{i=1}^{k}(\xi_i - m) > m(y-k)\right\} \leq ck\overline{F}(m(y-k)) \qquad \text{for all} \quad k \leq z$$

provided $z \le (y-z)^{1+\delta/2}$. Therefore,

$$\begin{aligned}\mathbb{P}\Big\{\sum_{i=1}^{Z_{n-1}} \xi_i > my;\ Z_{n-1} \le z\Big\} &= \sum_{k=1}^{z} \mathbb{P}\{Z_{n-1}=k\}\,\mathbb{P}\Big\{\sum_{i=1}^{k}\xi_i > my\Big\}\\ &\le c\sum_{k=1}^{z}\mathbb{P}\{Z_{n-1}=k\}k\overline{F}(m(y-k))\\ &\le c\,\mathbb{E}\,Z_{n-1}\overline{F}(m(y-z)) = cm^{n-1}\overline{F}(m(y-z)). \qquad (4.2)\end{aligned}$$

Substituting this into (4.1) with $y = m^{n-1}x$ and $z = m^{n-1}(x-x_n)$, we obtain

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\{W_{n-1} > x - x_n\} + cm^{n-1}\overline{F}(m^n x_n)$$

provided $x - x_n \le m^{(n-1)\delta/2}x_n^{1+\delta/2}$. Iterating this upper bound $n-N$ times, we arrive at the following inequality:

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > x - x_n - \ldots - x_{N+1}\big\} + c\sum_{k=N+1}^{n} m^{k-1}\overline{F}(m^k x_k) \qquad (4.3)$$

provided $x \le m^{(k-1)\delta/2}x_k^{1+\delta/2}$ for all k. Consider the decreasing sequence $x_k = x/k^2$ and choose N so large that $m^{(k-1)\delta/2} \ge k^{2+\delta}$ for all $k \ge N+1$. Then (4.3) holds for all $n \ge N+1$, and we have

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\Big\{W_N > \Big(1 - \frac{1}{(N+1)^2} - \ldots - \frac{1}{n^2}\Big)x\Big\} + c\sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\frac{m^k x}{k^2}\Big).$$

Choose N so large that additionally $\sum_{k=N}^{\infty} 1/k^2 \le \varepsilon$. Then

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > (1-\varepsilon)x\big\} + c\sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\frac{m^k x}{k^2}\Big).$$

Owing to condition (1.3),

$$\sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\frac{m^k x}{k^2}\Big) \le c_1\overline{F}(mx)\sum_{k=N+1}^{\infty}\frac{m^{k-1}}{(m^{k-1}/k^2)^{1+\delta}}.$$

Now we may increase N so that

$$c\sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\frac{m^k x}{k^2}\Big) \le \frac{\varepsilon\overline{F}(mx)}{3},$$

which implies

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > (1-\varepsilon)x\big\} + \frac{\varepsilon\overline{F}(mx)}{2}.$$

Applying here Lemma 9, we deduce $\overline{F}(mx) \le (1+o(1))\,\mathbb{P}\{W_N > (1-\varepsilon)x\}$ as $x \to \infty$, so

$$\mathbb{P}\{W_n > x\} \le (1+\varepsilon)\,\mathbb{P}\big\{W_N > (1-\varepsilon)x\big\}$$

for all sufficiently large x, and the proof is complete. □

The calculations above imply the following

Corollary 12. *Let the distribution F be dominated varying and satisfy condition* (1.3). *Then there exists a constant $c<\infty$ such that $\mathbb{P}\{W_n>x\}\le c\overline{F}(x)$ for all n and x.*

For dominated varying distributions it is possible to obtain a more accurate bound which will be useful for a wider class of distributions than the intermediate regularly varying ones. We do it in the next lemma, in which the bound provided by the previous corollary serves as the first-step preliminary bound.

Lemma 13. *Let $\mathbb{E}\,\xi^2<\infty$ and the distribution F be dominated varying and satisfy condition* (1.3). *Then, for every $\gamma>1/2$ and $\varepsilon>0$, there exists an N such that, for all $n>N$ and all sufficiently large x,*

$$\mathbb{P}\{W_n>x\}\le(1+\varepsilon)\,\mathbb{P}\{W_N>x-x^\gamma\}.$$

Proof. Here we need more accurate upper bounds based on (4.2). Take $\delta\in(1/\gamma-1,1)$. First note that, as follows from Proposition 6 under condition (2.1) (which is fulfilled because $\mathbb{E}\,\xi^2<\infty$), the bound (4.2) now holds within a larger time range where $z\le(y-z)^{1+\delta}$. For those z,

$$\mathbb{P}\left\{\sum_{i=1}^{Z_{n-1}}\xi_i>my;\ Z_{n-1}\le z\right\}\le c\left(\sum_{k=1}^{z/2}+\sum_{k=z/2}^{z}\right)\mathbb{P}\{Z_{n-1}=k\}k\overline{F}(m(y-k))=:c(\Sigma_1+\Sigma_2).$$

We have

$$\Sigma_1\le\overline{F}\Big(m\Big(y-\frac{z}{2}\Big)\Big)\sum_{k=1}^{y/2}\mathbb{P}\{Z_{n-1}=k\}k\le\mathbb{E}\,Z_{n-1}\overline{F}\Big(\frac{my}{2}\Big)\le c_1m^{n-1}\overline{F}(my)$$

for some $c_1<\infty$, by the dominated variation of F. Further,

$$\Sigma_2\le\mathbb{P}\Big\{Z_{n-1}>\frac{z}{2}\Big\}z\overline{F}(m(y-z))\le c_2\overline{F}\Big(\frac{z}{2m^{n-1}}\Big)z\overline{F}(m(y-z))\le c_2c_1\overline{F}\Big(\frac{z}{m^{n-1}}\Big)z\overline{F}(m(y-z))$$

by Corollary 12 and the dominated variation of F. Collecting the bounds for Σ_1 and Σ_2 with $y=m^{n-1}x$ and $z=m^{n-1}(x-x_n)$, we find from (4.1) that

$$\mathbb{P}\{W_n>x\}\le\mathbb{P}\{W_{n-1}>x-x_n\}+c_1m^{n-1}\overline{F}(m^nx)+c_3\overline{F}(x-x_n)m^{n-1}x\overline{F}(m^nx_n)$$

provided $x-x_n\le m^{(n-1)\delta}x_n^{1+\delta}$. Iterating this upper bound $n-N$ times, we arrive at the inequality

$$\begin{aligned}\mathbb{P}\{W_n>x\}\le\mathbb{P}\big\{W_N>x-x_n-\ldots-x_{N+1}\big\}+c_1\sum_{k=N+1}^{n}m^{k-1}\overline{F}\big(m^k(x-x_n-\ldots-x_{k+1})\big)\\+c_3\sum_{k=N+1}^{n}\overline{F}(x-x_n-\ldots-x_k)m^{k-1}x\overline{F}(m^kx_k)\end{aligned}\tag{4.4}$$

provided $x\le m^{(k-1)\delta}x_k^{1+\delta}$ for all $k=n,\ldots,N+1$.

Now take the decreasing sequence $x_k=x^\gamma/k^2$. Since $\gamma>1/2$ and $\delta\in(1/\gamma-1,1)$, it follows that $x^{\gamma(1+\delta)}>x$. Then (4.4) holds for every $n\ge N+1$ and we have

$$\begin{aligned}\mathbb{P}\{W_n>x\}&\le\mathbb{P}\left\{W_N>x-\left(\frac{1}{(N+1)^2}+\ldots+\frac{1}{n^2}\right)x^\gamma\right\}\\&+c_1\sum_{k=N+1}^{n}m^{k-1}\overline{F}\left(m^k\left(x-\left(\frac{1}{n^2}+\ldots+\frac{1}{(k+1)^2}\right)x^\gamma\right)\right)\\&+c_3\sum_{k=N+1}^{n}\overline{F}\left(x-\left(\frac{1}{n^2}+\ldots+\frac{1}{k^2}\right)x^\gamma\right)m^{k-1}x\overline{F}\left(\frac{m^kx^\gamma}{k^2}\right).\end{aligned}$$

Choose N so large that $\sum_{k=N+1}^{\infty} 1/k^2 \le 1$. Then

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > x - x^\gamma\big\} + c_1\overline{F}(x - x^\gamma) \sum_{k=N+1}^{n} m^{k-1} \frac{\overline{F}(m^k(x - x^\gamma))}{\overline{F}(x - x^\gamma)}$$
$$+ c_3\overline{F}(x - x^\gamma) \sum_{k=N+1}^{n} m^{k-1} x\overline{F}\bigg(\frac{m^k x^\gamma}{k^2}\bigg).$$

Owing to condition (1.3),

$$\sum_{k=N+1}^{n} m^{k-1} \frac{\overline{F}(m^k(x - x^\gamma))}{\overline{F}(x - x^\gamma)} \le c_4 \sum_{k=N+1}^{\infty} \frac{m^{k-1}}{m^{k(1+\delta)}} \to 0$$

and

$$\sum_{k=N+1}^{n} m^{k-1} x\overline{F}\bigg(\frac{m^k x^\gamma}{k^2}\bigg) \le c_4 x\overline{F}(x^\gamma) \sum_{k=N+1}^{n} \frac{m^{k-1}}{(m^k/k^2)^{1+\delta}} \le c_4\, \mathbb{E}\,\xi^2 x^{1-2\gamma} \sum_{k=N+1}^{\infty} \frac{m^{k-1}}{(m^k/k^2)^{1+\delta}} \to 0$$

as $N \to \infty$. Taking into account that $\overline{F}(x - x^\gamma) \le c_5\overline{F}(mx)$ and further increasing N, we derive the following bound:

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > x - x^\gamma\big\} + \frac{\varepsilon\overline{F}(mx)}{2}.$$

Applying here Lemma 9, we deduce $\overline{F}(mx) \le (1 + o(1))\,\mathbb{P}\{W_N > x\}$ as $x \to \infty$, so

$$\mathbb{P}\{W_n > x\} \le (1 + \varepsilon)\,\mathbb{P}\big\{W_N > x - x^\gamma\big\}$$

for all sufficiently large x, and the proof is complete. □

Note that the assertion of Lemma 11 holds not only for intermediate regularly varying distributions but for Weibull distributions as well; more precisely, the following result holds.

Lemma 14. *Let $\overline{F}(x) = e^{-R(x)}$ where $R(x)$ satisfies condition* (2.9) *and $R(x)/x \to 0$. Let condition* (1.3) *hold. Then, for every $\varepsilon > 0$, there exists an N such that*

$$\mathbb{P}\{W_n > x\} \le (1 + \varepsilon)\,\mathbb{P}\big\{W_N > (1 - \varepsilon)x\big\}$$

for all $n > N$ and all sufficiently large x.

Proof. It is similar to the proof of Lemma 11. We start again with inequality (4.1). It follows from Proposition 7 for sums with zero mean, $\eta_i = \xi_i - m$, that, for some $c < \infty$,

$$\mathbb{P}\Bigg\{\sum_{i=1}^{k} \xi_i > my\Bigg\} = \mathbb{P}\Bigg\{\sum_{i=1}^{k} (\xi_i - m) > m(y - k)\Bigg\} \le ck\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)(y - k)\Big) \qquad \text{for all} \quad k \le z$$

provided $z \le (y - z)^2/(cR(y - z))$. Therefore,

$$\mathbb{P}\Bigg\{\sum_{i=1}^{Z_{n-1}} \xi_i > my;\ Z_{n-1} \le z\Bigg\} \le c\,\mathbb{E}\,Z_{n-1}\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)(y - z)\Big) = cm^{n-1}\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)(y - z)\Big).$$

Substituting this into (4.1) with $y = m^{n-1}x$ and $z = m^{n-1}(x - x_n)$, we obtain

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_{n-1} > x - x_n\big\} + cm^{n-1}\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)^n x_n\Big)$$

provided $x - x_n \le m^{n-1}x_n^2/(cR(m^{n-1}x_n))$. Iterating this upper bound $n-N$ times, we arrive at the following inequality:

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{W_N > x - x_n - \ldots - x_{N+1}\big\} + c \sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)^k x_k\Big) \qquad (4.5)$$

provided $x - x_k \le m^{k-1}x_k^2/(cR(m^{k-1}x_k))$ for all $k = n, \ldots, N+1$. Consider the decreasing sequence $x_k = x/k^2$ and choose N so large that $m^{k-1}x/k^2 \ge R(m^{k-1}x/k^2)$ for all $k \ge N+1$; it is possible because $R(z)/z \to 0$ as $z \to \infty$. Then (4.5) holds for every $n \ge N+1$, and we have

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\Big\{W_N > \Big(1 - \frac{1}{(N+1)^2} - \ldots - \frac{1}{n^2}\Big)x\Big\} + c \sum_{k=N+1}^{n} m^{k-1}\overline{F}\Big(\Big(m - \frac{\varepsilon}{2}\Big)^k \frac{x}{k^2}\Big).$$

Choose $\varepsilon > 0$ so small that $m < (m - \varepsilon/2)^{1+\delta}$ where $\delta > 0$ is taken from condition (1.3). Then the rest of the proof is the same as the proof of Lemma 11. $\square$

5. FINITE TIME HORIZON ASYMPTOTICS

As follows from [10, Sect. 6], for intermediate regularly varying distribution F, for every fixed n,

$$\mathbb{P}\{W_n > x\} \sim \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \qquad \text{as} \quad x \to \infty. \qquad (5.1)$$

For the case where the second moment of ξ is finite, we extend this result to a wider class of distributions as follows.

Lemma 15. *Let $\mathbb{E}\,\xi^2 < \infty$ and the distribution F be dominated varying. If F is x^γ-insensitive for some $\gamma > 1/2$, then the equivalence* (5.1) *holds for every fixed n.*

Proof. First, Lemma 10 guarantees an appropriate lower bound. The upper bound will be proved by induction. It is true for $n = 1$. Assume, for some n,

$$\mathbb{P}\{W_n > x\} \le (1 + o(1)) \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \qquad \text{as} \quad x \to \infty. \qquad (5.2)$$

Let us prove that then (5.2) holds for $n+1$. We start with the inequality

$$\begin{aligned}
\mathbb{P}\{W_{n+1} > x\} &= \mathbb{P}\Big\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x\Big\} \\
&\le \mathbb{P}\big\{Z_n > m^n(x - x^\gamma)\big\} + \mathbb{P}\Big\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x;\ \frac{m^n x}{2} < Z_n \le m^n(x - x^\gamma)\Big\} \\
&\quad + \mathbb{P}\Big\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x;\ m^n x^\gamma < Z_n \le \frac{m^n x}{2}\Big\} + \mathbb{P}\Big\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x;\ Z_n \le m^n x^\gamma\Big\} \\
&=: P_1 + P_2 + P_3 + P_4,
\end{aligned}$$

where the ξ_i are independent of Z_n. Due to the induction hypothesis and to the fact that F is x^γ-insensitive,

$$P_1 = \mathbb{P}\big\{W_n > x - x^\gamma\big\} \le (1 + o(1)) \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}(x - x^\gamma)) \sim \sum_{i=0}^{n-1} m^i \overline{F}(m^{i+1}x) \qquad \text{as} \quad x \to \infty.$$

Take $\delta \in (1/\gamma - 1, 1)$. All the values of k not greater than $m^n x$ are negligible compared to $y^{1+\delta}$ where $y = m^{n+1}x^\gamma$, $\gamma > 1/2$. Therefore, by Proposition 6 there exists a $c < \infty$ such that

$$\mathbb{P}\left\{\sum_{i=1}^{k}(\xi_i - m) > m^{k+1}x - km\right\} \le ck\overline{F}(m^{n+1}x - km)$$

for sufficiently large x and all $k \le m^n(x - x^\gamma)$.

Therefore, for sufficiently large x,

$$\begin{aligned} P_2 &= \sum_{k=m^n x/2}^{m^n(x-x^\gamma)} \mathbb{P}\{Z_n = k\}\,\mathbb{P}\left\{\sum_{i=1}^{k}(\xi_i - m) > m^{n+1}x - km\right\} \\ &\le c\sum_{k=m^n x/2}^{m^n(x-x^\gamma)} \mathbb{P}\{Z_n = k\}k\overline{F}(m^{n+1}x - km) \\ &\le cm^n x\mathbb{P}\left\{Z_n \ge \frac{m^n x}{2}\right\}\overline{F}(m^{n+1}x^\gamma). \end{aligned}$$

Since $\mathbb{E}\,\xi^2 < \infty$ and $\gamma > 1/2$, it follows that $x\overline{F}(m^{n+1}x^\gamma) \to 0$ as $x \to \infty$. Hence, as $x \to \infty$,

$$P_2 = o\Big(\mathbb{P}\Big\{W_n \ge \frac{x}{2}\Big\}\Big) = o\Big(\overline{F}\Big(\frac{mx}{2}\Big)\Big) = o(\overline{F}(mx))$$

owing to the induction hypothesis (5.2) and the dominated variation of F.

Further, for sufficiently large x,

$$\begin{aligned} P_3 &\le c\sum_{k=m^n x^\gamma}^{m^n x/2} \mathbb{P}\{Z_n = k\}k\overline{F}(m^{n+1}x - km) \le c\,\mathbb{E}\{Z_n;\, Z_n > m^n x^\gamma\}\overline{F}\left(\frac{m^{n+1}x}{2}\right) \\ &= o(\overline{F}(mx)) \qquad \text{as} \quad x \to \infty \end{aligned}$$

again because of the dominated variation of F.

Finally,

$$P_4 = \sum_{k=1}^{m^n x^\gamma} \mathbb{P}\{Z_n = k\}\,\mathbb{P}\left\{\sum_{i=1}^{k}(\xi_i - 2m) > m^{n+1}x - 2km\right\}.$$

The distribution F is dominated varying and long-tailed (constant-insensitive), which implies that it belongs to the class $\mathcal{S}^*$ (see, e.g., [13, Theorem 3.29]). Also, the expression $m^{n+1}x - 2km$ tends to infinity as $x \to \infty$ uniformly in $k \le mx^\gamma$. This allows us to apply here Proposition 5 to the random variables $\eta_i := \xi_i - 2m$ with negative mean; it ensures that, uniformly in $k \le mx^\gamma$,

$$\mathbb{P}\left\{\sum_{i=1}^{k}(\xi_i - 2m) > m^{n+1}x - 2km\right\} \le (1 + o(1))k\overline{F}(m^{n+1}x - 2km) \sim k\overline{F}(m^{n+1}x) \qquad \text{as} \quad x \to \infty,$$

because F is x^γ-insensitive. Thus,

$$P_4 \sim \overline{F}(m^{n+1}x)\sum_{k=1}^{m^n x^\gamma} \mathbb{P}\{Z_n = k\}k \sim \overline{F}(m^{n+1}x)\,\mathbb{E}\,Z_n = m^n\overline{F}(m^{n+1}x).$$

Combining the bounds for $P_1, \ldots, P_4$, we deduce that

$$\mathbb{P}\{W_{n+1} > x\} \le \mathbb{P}\{W_n > mx\} + m^n\overline{F}(m^{n+1}x) + o(\overline{F}(mx)) \qquad \text{as} \quad x \to \infty,$$

and the induction hypothesis (5.2) completes the proof. □

If the distribution F is rapidly varying, then

$$\sum_{i=0}^{\infty} m^i \overline{F}(m^{i+1}x) \sim \overline{F}(mx) \qquad \text{as} \quad x \to \infty. \tag{5.3}$$

Indeed, fix $\varepsilon > 0$ and choose $x(\varepsilon)$ such that $\overline{F}(mx) \le \varepsilon \overline{F}(x)$ for every $x > x(\varepsilon)$. Then, for $x > x(\varepsilon)$,

$$\sum_{i=1}^{\infty} m^i \overline{F}(m^{i+1}x) \le \sum_{i=1}^{\infty} (m\varepsilon)^i \overline{F}(mx) = \frac{m\varepsilon}{1 - m\varepsilon} \overline{F}(mx).$$

The constant multiplier on the right-hand side can be made as small as we please by an appropriate choice of ε, so the equivalence (5.3) follows.

Lemma 16. *Let $\overline{F}(x) = e^{-R(x)}$ where $R(x)$ is regularly varying with index $\beta \in (0, 1/2)$. In the case $\beta \in [(3 - \sqrt{5})/2, 1/2)$ assume also that condition (1.8) holds. Additionally assume that $F \in \mathcal{S}^*$. Then, for every fixed n,*

$$\mathbb{P}\{W_n > x\} \sim \overline{F}(mx) \qquad \textit{as} \quad x \to \infty.$$

Proof. Since $\beta < 1/2$, the distribution F is $\sqrt{x}$-insensitive, which by Lemma 10 implies the lower bound $\mathbb{P}\{W_n > x\} \ge (1 + o(1))\overline{F}(mx)$ as $x \to \infty$.

To prove the upper bound, we apply induction arguments. For $n = 1$, we have the equality $\mathbb{P}\{W_1 > x\} = \overline{F}(mx)$. Assume now $\mathbb{P}\{W_n > x\} \sim \overline{F}(mx)$ for some $n \ge 1$. Let us prove that then it holds for $n + 1$.

If $\beta < (3 - \sqrt{5})/2$, then the interval $(1/(2 - \beta), 1 - \beta)$ is not empty; in this case we take $\gamma_1 = \gamma_2 \in (1/(2 - \beta), 1 - \beta)$. If $\beta \in [(3 - \sqrt{5})/2, 1/2)$, then $1/(2 - \beta) \ge 1 - \beta$ and we take $\gamma_1 \in (1/2, 1 - \beta)$ and $\gamma_2 > 1/(2 - \beta)$ so that $\gamma_2 \ge \gamma_1$. Since $\gamma_1 < 1 - \beta$, the distribution F is x^{γ_1}-insensitive. We start with the inequality

$$\begin{aligned}\mathbb{P}\biggl\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x\biggr\} &\le \mathbb{P}\bigl\{Z_n > m^n(x - x^{\gamma_1})\bigr\} + \mathbb{P}\biggl\{\sum_{i=1}^{Z_n} \xi_i > m^{n+1}x;\ Z_n \le m^n(x - x^{\gamma_1})\biggr\} \\ &=: P_1 + P_2,\end{aligned}$$

where the ξ_i do not depend on Z_n. By the induction hypothesis and since F is x^{γ_1}-insensitive,

$$P_1 \sim \overline{F}(m(x - x^{\gamma_1})) \sim \overline{F}(mx) \qquad \text{as} \quad x \to \infty.$$

It remains to prove that $P_2 = o(\overline{F}(mx))$ as $x \to \infty$. We start with the following decomposition:

$$P_2 = \Biggl(\sum_{k=0}^{m^n(x - x^{\gamma_2}) - 1} + \sum_{k=m^n(x - x^{\gamma_2})}^{m^n(x - x^{\gamma_1})}\Biggr) \mathbb{P}\{Z_n = k\}\, \mathbb{P}\biggl\{\sum_{i=1}^{k} (\xi_i - m) > m^{n+1}x - km\biggr\} =: P_{21} + P_{22}.$$

In the first sum P_{21} we have $m^{n+1}x - mk \ge m^{n+1}x^{\gamma_2} \gg x^{1/(2-\beta)}$ due to the choice $\gamma_2 > 1/(2 - \beta)$. The function $R(x)/x^2$ is regularly varying with index $\beta - 2$. Hence, as $x \to \infty$, we have $kR(m^{n+1}x - km)/(m^{n+1}x - km)^2 \to 0$ uniformly in $k \le m^n(x - x^{\gamma_2})$. This observation, together with the regular variation of $R(x)$, allows us to apply Proposition 7 with $y = (1 - \varepsilon)x$, which ensures that

$$\mathbb{P}\biggl\{\sum_{i=1}^{k} (\xi_i - m) > m^{n+1}x - mk\biggr\} \le k\overline{F}\bigl((m^{n+1}x - mk)(1 - \varepsilon)\bigr)$$

for sufficiently large x and all $k \le m^n(x - x^{\gamma_2})$. Thus, for sufficiently large x,

$$P_{21} \le \sum_{k=0}^{m^n(x-x^{\gamma_2})} \mathbb{P}\{Z_n = k\} k \overline{F}\big((m^n x - k)m(1-\varepsilon)\big).$$

Take $\varepsilon > 0$ so small that $m(1-\varepsilon) > 1$. Then, due to the rapid variation of F, as $x \to \infty$,

$$\overline{F}\big((m^n x - k)m(1-\varepsilon)\big) = o(\overline{F}(m^n x - k)) \qquad \text{uniformly in } \ k \le m^n(x - x^{\gamma_2}).$$

In addition, owing to the induction hypothesis,

$$\mathbb{P}\{Z_n = k\} \le \mathbb{P}\Big\{W_n \ge \frac{k}{m^n}\Big\} \le c\overline{F}\Big(\frac{k}{m^{n-1}}\Big)$$

for some $c < \infty$. Thus, as $x \to \infty$,

$$P_{21} \le o(1) \int_0^{m^n(x-x^{\gamma_2})} y\overline{F}\Big(\frac{y}{m^{n-1}}\Big)\overline{F}(m^n x - y)\, dy = o(1) \int_0^{m(x-x^{\gamma_2})} y\overline{F}(y)\overline{F}(m^{n-1}(mx - y))\, dy.$$

Since $m^{n-1} \ge m > 1$ and $\beta > 0$,

$$y\overline{F}(m^{n-1}(mx - y)) = o(\overline{F}(mx - y))$$

as $x \to \infty$ uniformly in $y \le m(x - x^{\gamma_2})$. Therefore,

$$P_{21} \le o(1) \int_0^{mx} \overline{F}(y)\overline{F}(mx - y)\, dy.$$

The inclusion $F \in \mathcal{S}^*$ means that

$$\int_0^{mx} \overline{F}(y)\overline{F}(mx - y)\, dy \sim 2\overline{F}(mx) \int_0^{\infty} \overline{F}(y)\, dy \qquad \text{as} \quad x \to \infty,$$

which finally implies $P_{21} = o(\overline{F}(mx))$. In the case $\beta < (3 - \sqrt{5})/2$ this completes the proof because then $\gamma_1 = \gamma_2$ and $P_{22} = 0$.

If $\beta < 1/2$, then it remains to prove that $P_{22} = o(\overline{F}(mx))$ as well. We have

$$P_{22} = \sum_{k=m^n x^{\gamma_1}}^{m^n x^{\gamma_2}} \mathbb{P}\big\{Z_n = m^n x - k\big\}\, \mathbb{P}\Bigg\{\sum_{i=1}^{m^n x - k} (\xi_i - m) > mk\Bigg\}.$$

By the induction hypothesis

$$\mathbb{P}\big\{Z_n = m^n x - k\big\} \le \mathbb{P}\Big\{W_n \ge x - \frac{k}{m^n}\Big\} \sim \overline{F}\Big(mx - \frac{k}{m^{n-1}}\Big),$$

so that

$$P_{22} \le c_1 \sum_{k=m^n x^{\gamma_1}}^{m^n x^{\gamma_2}} \overline{F}\Big(mx - \frac{k}{m^{n-1}}\Big)(m^n x - k + 1)\overline{F}(y_k)$$

for any y_k satisfying the inequalities $y_k \le mk/2$ and $m^n x - k \le (mk)^2/cR(y_k)$, where $c = c(1/2)$ is defined in Proposition 7. Choose $\gamma \in (2\beta, 1)$ such that

$$\frac{1}{\gamma_1} - 1 < \gamma < \frac{1}{\gamma_2} - 1 + \beta, \tag{5.4}$$

which is possible if we choose $\gamma_2 > 1/(2-\beta)$ sufficiently close to $1/(2-\beta)$. Then we take y_k which solves $R(y_k) = m^{2-n}k^{1+\gamma}/cx = c_2 k^{1+\gamma}/x$. With this choice, $y_k \le mk/2$ for $k \le m^n x^{\gamma_2}$ and sufficiently large x (by the right inequality in (5.4)) and $m^n x - k \le (mk)^2/(cR(y_k))$.

Further, since $\overline{F}(y_k) = e^{-R(y_k)}$,

$$P_{22} \le c_3 x \sum_{k=m^n x^{\gamma_1}}^{m^n x^{\gamma_2}} \overline{F}\left(mx - \frac{k}{m^{n-1}}\right)\overline{F}(y_k) \le c_3 x \overline{F}(mx) \sum_{k=m^n x^{\gamma_1}}^{m^n x^{\gamma_2}} e^{R(mx)-R(mx-k/m^{n-1})-R(y_k)}.$$

By condition (1.8) on the increments of R and by the regular variation of R, we have

$$\frac{R(x)-R(y)}{x-y} \le c_4 \frac{R(x)}{x}, \qquad x \ge y \ge 1,$$

which implies

$$R(mx) - R\left(mx - \frac{k}{m^{n-1}}\right) - R(y_k) \le \frac{c_5 k R(mx)}{x} - \frac{c_2 k^{1+\gamma}}{x} = \frac{(c_5 R(mx) - c_2 k^{\gamma})k}{x}.$$

Since $R(mx)$ is regularly varying with index $\beta < 1/2$ and $k \ge m^n x^{\gamma_1}$, the choice $\gamma_1 \in (1/2, 1-\beta)$ and $\gamma \in (2\beta, 1)$ ensures $R(mx) = o(k)$. Hence,

$$R(mx) - R\left(mx - \frac{k}{m^{n-1}}\right) - R(y_k) \le -\frac{c_6 k^{1+\gamma}}{x},$$

which yields

$$P_{22} \le c_4 x \overline{F}(mx) \sum_{k=m^n x^{\gamma_1}}^{\infty} e^{-c_6 k^{1+\gamma}/x} = o(\overline{F}(mx)) \qquad \text{as} \quad x \to \infty$$

due to $\gamma_1(1+\gamma) > 1$ and the left inequality in (5.4). Combining all the bounds, we deduce that $P_2 = o(\overline{F}(mx))$ and consequently $\mathbb{P}\{W_{n+1} > x\} \sim P_1 \sim \overline{F}(mx)$ as $x \to \infty$, and the proof is complete. □

6. PROOFS OF THEOREMS 1, 2 AND 3

Proof of Theorem 1. The bounds (1.4) follow from Lemma 9 and Corollary 12. All the other assertions follow from the equivalence (5.1) and from Lemmas 9 and 11. □

Proof of Theorem 2. It follows from Lemmas 15, 10 and 13. □

Proof of Theorem 3. The lower bound for the general case $\beta < 1$ follows from Lemma 9. The upper bound follows from Lemma 14, which reduces the problem to the finite time horizon N and further induction arguments like

$$\mathbb{P}\{W_N > x\} = \mathbb{P}\left\{\sum_{i=1}^{\xi} W_{N-1}^{(i)} > mx\right\} \le \mathbb{P}\{\xi > mx(1-\varepsilon)\} + \mathbb{P}\left\{\sum_{i=1}^{\xi} W_{N-1}^{(i)} > mx;\ \xi \le mx(1-\varepsilon)\right\},$$

where $W_{N-1}^{(1)}, W_{N-1}^{(2)}, \dots$ are independent copies of W_{N-1}. Assuming that W_{N-1} has a tail not heavier than $c\overline{F}((1-\varepsilon)x)$, me can estimate here the second probability as follows:

$$\mathbb{P}\Big\{\sum_{i=1}^{\xi} W_{N-1}^{(i)} > mx;\ \xi \le mx(1-\varepsilon)\Big\} = \sum_{k=1}^{mx(1-\varepsilon)} \mathbb{P}\{\xi = k\}\,\mathbb{P}\Big\{\sum_{i=1}^{k}\big(W_{N-1}^{(i)} - k\big) > mx - k\Big\}.$$

By Proposition 7,

$$\mathbb{P}\Big\{\sum_{i=1}^{k}\big(W_{N-1}^{(i)} - k\big) > mx - k\Big\} \le (k+1)\overline{F}\big((1-\varepsilon)(mx-k)\big)$$

as $x \to \infty$ uniformly in $k \le mx(1-\varepsilon)$; note that the condition $k \le mx(1-\varepsilon)$ implies that $mx - k \ge mx\varepsilon$ and hence covers both conditions of Proposition 7. Thus,

$$\begin{aligned}\sum_{k=1}^{mx(1-\varepsilon)} \mathbb{P}\{\xi = k\}\,\mathbb{P}\Big\{\sum_{i=1}^{k}\big(W_{N-1}^{(i)} - k\big) > mx - k\Big\} &\le 2\sum_{k=1}^{mx(1-\varepsilon)} \mathbb{P}\{\xi = k\}k\overline{F}\big((1-\varepsilon)(mx-k)\big)\\ &= o\big(\overline{F}(m(1-2\varepsilon)x)\big)\end{aligned}$$

as $x \to \infty$, by the standard properties of Weibull-type distributions. This completes the proof of the upper bound for the case $\beta < 1$.

In the case $\beta < 1/2$ the distribution F is $\sqrt{x}$-insensitive, which by Lemma 10 implies the lower bound $\mathbb{P}\{W_n > x\} \ge (1+o(1))\overline{F}(mx)$ as $x \to \infty$.

Now let us prove the upper bound in the case $\beta < 1/2$. Fix an $\varepsilon > 0$. Owing to Lemma 14 we find N such that, for all $n > N$ and all sufficiently large x,

$$\mathbb{P}\{W_n > x\} \le (1+\varepsilon)\,\mathbb{P}\big\{W_N > (1-\varepsilon)x\big\}.$$

As in the proof of Lemma 16, we take $\gamma \in (1/(2-\beta), 1-\beta)$, so F is x^γ-insensitive. We make use of the following decomposition for $n > N+1$:

$$\mathbb{P}\{W_n > x\} \le \mathbb{P}\big\{\xi > m(x - x^\gamma)\big\} + \mathbb{P}\Big\{\sum_{i=1}^{\xi} W_{n-1}^{(i)} > mx;\ \xi \le m(x - x^\gamma)\Big\} =: P_1 + P_2.$$

Since F is x^γ-insensitive,

$$P_1 = \mathbb{P}\big\{\xi > m(x - x^\gamma)\big\} \sim \overline{F}(mx) \qquad \text{as} \quad x \to \infty.$$

Further, we use Lemma 14, which is applicable because $n - 1 > N$: ultimately in y,

$$\mathbb{P}\{W_{n-1} > y\} \le (1+\varepsilon)\,\mathbb{P}\big\{W_N > (1-\varepsilon)y\big\} \le (1+2\varepsilon)\overline{F}((1-\varepsilon)my)$$

by virtue of Lemma 16. Choose $\varepsilon > 0$ so small that $m_* := (1-\varepsilon)m > 1$, which is possible since $m > 1$. The family $\{W_{n-1} - 1,\ n > N+1\}$ satisfies the conditions of Corollary 8, which further allows us to prove that $P_2 = o(\overline{F}(mx))$ as $x \to \infty$ uniformly in $n > N+1$ in the same way as in the proof of Lemma 16. $\square$

7. THE CASE OF REGULARLY VARYING TAIL WITH INDEX −1; PROOF OF THEOREM 4

As proved in Lemma 9, for every $\varepsilon > 0$,

$$\mathbb{P}\{W_n > x\} \geq (1 + o(1)) \sum_{k=0}^{n-1} m^k \overline{F}\big(m^{k+1}(1+\varepsilon)x\big)$$

as $x \to \infty$ uniformly in $n \geq 1$. Since F is regularly varying, we deduce from here that

$$\mathbb{P}\{W_n > x\} \geq (1 + o(1)) \sum_{k=0}^{n-1} m^k \overline{F}(m^{k+1}x)$$

as $x \to \infty$ uniformly in $n \geq 1$. Then it remains to prove the following upper bound: for every fixed $\varepsilon > 0$,

$$\mathbb{P}\{W_n > x\} \leq (1 + o(1)) \sum_{k=0}^{n-1} m^k \overline{F}\big(m^{k+1}x(1-\varepsilon)\big). \tag{7.1}$$

The method for proving upper bounds based on Lemma 11 does not work here because it relies heavily on condition (1.3). For this reason we proceed in a different way. Define events

$$A_k(x) := \big\{\xi_i^{(k)} > m^{k+1}x(1-\varepsilon) \text{ for some } i \leq Z_k\big\}.$$

Clearly,

$$\mathbb{P}\big\{A_k(x) \,|\, Z_k = j\big\} \leq j\overline{F}\big(m^{k+1}x(1-\varepsilon)\big), \qquad j \geq 1.$$

Therefore, $\mathbb{P}\{A_k(x)\} \leq m^k \overline{F}(m^{k+1}x(1-\varepsilon))$ and

$$\mathbb{P}\left\{\bigcup_{k=0}^{n-1} A_k(x)\right\} \leq \sum_{k=0}^{n-1} \mathbb{P}\{A_k(x)\} \leq \sum_{k=0}^{n-1} m^k \overline{F}\big(m^{k+1}x(1-\varepsilon)\big).$$

Owing to this and the upper bound

$$\mathbb{P}\{W_n > x\} \leq \mathbb{P}\left\{W_n > x,\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} + \mathbb{P}\left\{\bigcup_{k=0}^{n-1} A_k(x)\right\},$$

we conclude that (7.1) will be implied by the following relation: for every fixed $\varepsilon > 0$,

$$\mathbb{P}\left\{W_n > x,\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} = o\left(\sum_{k=0}^{n-1} m^k \overline{F}(m^{k+1}x)\right) \tag{7.2}$$

as $x \to \infty$ uniformly in $n \geq 1$. By the Chebyshev inequality, for every $\lambda > 0$

$$\begin{aligned}\mathbb{P}\left\{W_n > x,\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} &\leq \frac{\mathbb{E}\big\{e^{\lambda Z_n} - 1;\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\big\}}{e^{\lambda m^n x} - 1} \\ &= \frac{\mathbb{E}\big\{e^{\lambda Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\big\} - 1}{e^{\lambda m^n x} - 1} + \frac{\mathbb{P}\big\{\bigcup_{k=0}^{n-1} A_k(x)\big\}}{e^{\lambda m^n x} - 1},\end{aligned}$$

so that relation (7.2) will follow if we find $\lambda = \lambda_n(x)$ such that

$$\lambda m^n x \to \infty \tag{7.3}$$

and

$$\frac{\mathbb{E}\big\{e^{\lambda Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\big\} - 1}{e^{\lambda m^n x} - 1} = o\left(\sum_{k=0}^{n-1} m^k \overline{F}(m^{k+1}x)\right). \tag{7.4}$$

In order to find $\lambda = \lambda_n(x)$ satisfying (7.3) and (7.4), we proceed with suitable exponential bounds for bounded random variables. Take $\lambda_{nn} > 0$ and consider the following exponential moment:

$$\mathbb{E}\left\{e^{\lambda_{nn} Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} = \sum_{i=1}^{\infty} \mathbb{E}\left\{e^{\lambda_{nn} Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)},\ Z_{n-1} = i\right\}$$
$$= \sum_{i=1}^{\infty} \mathbb{E}\left\{e^{\lambda_{nn}(\xi_1^{(n-1)}+\ldots+\xi_i^{(n-1)})};\ \overline{A_{n-1}(x)},\ \bigcap_{k=0}^{n-2} \overline{A_k(x)},\ Z_{n-1} = i\right\}.$$

Note that the events $\bigcap_{k=0}^{n-2} \overline{A_k(x)}$ and $Z_{n-1} = i$ do not depend on the $\xi^{(n-1)}$. Therefore,

$$\mathbb{E}\left\{e^{\lambda_{nn} Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} = \sum_{i=1}^{\infty} \mathbb{E}\left\{e^{\lambda_{nn}(\xi_1^{(n-1)}+\ldots+\xi_i^{(n-1)})};\ \overline{A_{n-1}(x)}\right\} \mathbb{P}\left\{\bigcap_{k=0}^{n-2} \overline{A_k(x)}, Z_{n-1} = i\right\}$$
$$= \sum_{i=1}^{\infty} \left(\mathbb{E}\left\{e^{\lambda_{nn}\xi};\ \xi \le m^n x(1-\varepsilon)\right\}\right)^i \mathbb{P}\left\{\bigcap_{k=0}^{n-2} \overline{A_k(x)},\ Z_{n-1} = i\right\}.$$

If we put

$$\lambda_{n,n-1} := \log \mathbb{E}\left\{e^{\lambda_{nn}\xi};\ \xi \le m^n x(1-\varepsilon)\right\},$$

then we obtain the recursive equality

$$\mathbb{E}\left\{e^{\lambda_{nn} Z_n};\ \bigcap_{k=0}^{n-1} \overline{A_k(x)}\right\} = \mathbb{E}\left\{e^{\lambda_{n,n-1} Z_{n-1}};\ \bigcap_{k=0}^{n-2} \overline{A_k(x)}\right\}.$$

We iterate this recursion n times. Let us estimate $\lambda_{n,n-1}$ via λ_{nn}.

For every $z > 0$ and $y \le z$, it holds that $e^y \le 1 + y + y^2 e^z/2$. Therefore,

$$\mathbb{E}\left\{e^{\lambda_{nn}\xi};\ \xi \le m^n x(1-\varepsilon)\right\} \le 1 + \lambda_{nn} m + \frac{1}{2}\lambda_{nn}^2\, \mathbb{E}\left\{\xi^2;\ \xi \le m^n x\right\} e^{\lambda_{nn} m^n x(1-\varepsilon)}.$$

Since F is regularly varying with index -1, for sufficiently large x we have

$$\mathbb{E}\left\{\xi^2;\ \xi \le m^n x\right\} \le \frac{3}{2}(m^n x)^2 \overline{F}(m^n x).$$

Hence,

$$\mathbb{E}\left\{e^{\lambda_{nn}\xi};\ \xi \le m^n x(1-\varepsilon)\right\} \le 1 + \lambda_{nn}\left(m + \frac{3}{4}(\lambda_{nn} m^n x) m^n x \overline{F}(m^n x) e^{\lambda_{nn} m^n x(1-\varepsilon)}\right). \tag{7.5}$$

Define

$$p_n(x) := \sum_{k=0}^{n-1} m^k \overline{F}(m^{k+1}x)$$

and make a special choice of the initial λ_n:

$$\lambda_{nn} = \lambda_n(x) := (1+\varepsilon)\frac{\log \frac{1}{p_n(x)x} - 2\log\log \frac{1}{p_n(x)x}}{x\prod_{k=0}^{n-1}\left(m + \frac{m^{k+1}x\overline{F}(m^{k+1}x)}{(p_n(x)x)^{1-\varepsilon^2}}\right)}.$$

For the product, we have the following inequalities:

$$m^n \le \prod_{k=0}^{n-1}\biggl(m+\frac{m^{k+1}x\overline{F}(m^{k+1}x)}{(p_n(x)x)^{1-\varepsilon^2}}\biggr) = m^n\prod_{k=0}^{n-1}\biggl(1+\frac{m^k\overline{F}(m^{k+1}x)}{p_n(x)}(p_n(x)x)^{\varepsilon^2}\biggr)$$

$$\le m^n\exp\biggl\{(p_n(x)x)^{\varepsilon^2}\sum_{k=0}^{n-1}\frac{m^k\overline{F}(m^{k+1}x)}{p_n(x)}\biggr\} = m^n e^{(p_n(x)x)^{\varepsilon^2}}.$$

Note that then this product is asymptotically equivalent to m^n because $p_n(x)x\to 0$. Note also that

$$\frac{1}{e^{\lambda_{nn}m^n x}-1} \le \frac{1}{\bigl(\frac{1}{p_n(x)x}\bigr)^{1+\varepsilon+o(1)}\log^{-2(1+\varepsilon+o(1))}\frac{1}{p_n(x)x}-1}$$

$$\sim (p_n(x)x)^{1+\varepsilon+o(1)}\log^{2(1+\varepsilon+o(1))}\frac{1}{p_n(x)x} \le c_1(p_n(x)x)^{1+\varepsilon/2} \tag{7.6}$$

ultimately in x and uniformly in n. In particular, it goes to zero and relation (7.3) follows.

Now we estimate all λ_{nk}, $k\le n-1$. With the choice of λ_{nn} made, it follows from (7.5) that

$$\mathbb{E}\{e^{\lambda_{nn}\xi};\ \xi\le m^n x(1-\varepsilon)\}$$

$$\le 1+\lambda_{nn}\biggl(m+\frac{3(1+\varepsilon)}{4}\log\frac{1}{p_n(x)x}m^n x\overline{F}(m^n x)\exp\biggl\{(1-\varepsilon^2)\biggl(\log\frac{1}{p_n(x)x}-2\log\log\frac{1}{p_n(x)x}\biggr)\biggr\}\biggr)$$

$$\le 1+\lambda_{nn}\biggl(m+m^n x\overline{F}(m^n x)\exp\biggl\{(1-\varepsilon^2)\log\frac{1}{p_n(x)x}\biggr\}\biggr)$$

provided $1+\varepsilon<4/3$ and $2(1-\varepsilon^2)>1$. Thus,

$$\mathbb{E}\{e^{\lambda_{nn}\xi};\ \xi\le m^n x(1-\varepsilon)\} \le 1+\lambda_{nn}\biggl(m+\frac{m^n x\overline{F}(m^n x)}{(p_n(x)x)^{1-\varepsilon^2}}\biggr) \le \exp\biggl\{\lambda_{nn}\biggl(m+\frac{m^n x\overline{F}(m^n x)}{(p_n(x)x)^{1-\varepsilon^2}}\biggr)\biggr\},$$

which yields

$$\lambda_{n,n-1} \le \lambda_{nn}\biggl(m+\frac{m^n x\overline{F}(m^n x)}{(p_n(x)x)^{1-\varepsilon^2}}\biggr) = (1+\varepsilon)\frac{\log\frac{1}{p_n(x)x}-2\log\log\frac{1}{p_n(x)x}}{x\prod_{k=0}^{n-2}\Bigl(m+\frac{m^{k+1}x\overline{F}(m^{k+1}x)}{(p_n(x)x)^{1-\varepsilon^2}}\Bigr)}.$$

Iterating this estimate n times, we finally deduce that

$$\mathbb{E}\biggl\{e^{\lambda_{nn}Z_n};\ \bigcap_{k=0}^{n-1}\overline{A_k(x)}\biggr\} \le \mathbb{E}\,e^{\lambda_{n0}} = e^{\lambda_{n0}} = \exp\biggl\{\frac{1+\varepsilon}{x}\log\frac{1}{p_n(x)x}\biggr\}.$$

From here and (7.6),

$$\frac{\mathbb{E}\{e^{\lambda_{nn}Z_n};\ \bigcap_{k=0}^{n-1}\overline{A_k(x)}\}-1}{e^{\lambda_{nn}m^n x}-1} \le c_2x^{-1}(p_n(x)x)^{1+\varepsilon/2}\log\frac{1}{p_n(x)x} \le c_3x^{-1}(p_n(x)x)^{1+\varepsilon/4}$$

$$= c_3p_n(x)(p_n(x)x)^{\varepsilon/4} = o(p_n(x)),$$

and (7.4) is also proved. This completes the proof of Theorem 4.

ACKNOWLEDGMENTS

We are very grateful to both referees, who pointed out some incorrectness in the first version of the paper and made very constructive comments.

The work was supported by the DFG and the Russian Foundation for Basic Research (project no. 11-01-00285).

REFERENCES

1. S. Asmussen and H. Hering, *Branching Processes* (Birkhäuser, Boston, 1983).
2. K. B. Athreya and P. E. Ney, *Branching Processes* (Springer, Berlin, 1972).
3. N. Berestycki, N. Gantert, P. Mörters, and N. Sidorova, "Galton–Watson trees with vanishing martingale limit," arXiv: 1204.3080 [math.PR].
4. J. D. Biggins and N. H. Bingham, "Large deviations in the supercritical branching process," Adv. Appl. Probab. **25**, 757–772 (1993).
5. N. H. Bingham and R. A. Doney, "Asymptotic properties of supercritical branching processes. I: The Galton–Watson process," Adv. Appl. Probab. **6**, 711–731 (1974).
6. N. H. Bingham and R. A. Doney, "Asymptotic properties of supercritical branching processes. II: Crump–Mode and Jirina processes," Adv. Appl. Probab. **7**, 66–82 (1975).
7. A. A. Borovkov and K. A. Borovkov, *Asymptotic Analysis of Random Walks: Heavy-Tailed Distributions* (Cambridge Univ. Press, Cambridge, 2008).
8. A. De Meyer, "On a theorem of Bingham and Doney," J. Appl. Probab. **19**, 217–220 (1982).
9. D. Denisov, A. V. Dieker, and V. Shneer, "Large deviations for random walks under subexponentiality: The big-jump domain," Ann. Probab. **36**, 1946–1991 (2008).
10. D. Denisov, S. Foss, and D. Korshunov, "Asymptotics of randomly stopped sums in the presence of heavy tails," Bernoulli **16**, 971–994 (2010).
11. K. Fleischmann and V. Wachtel, "Lower deviation probabilities for supercritical Galton–Watson processes," Ann. Inst. Henri Poincaré, Probab. Stat. **43**, 233–255 (2007).
12. K. Fleischmann and V. Wachtel, "On the left tail asymptotics for the limit law of supercritical Galton–Watson processes in the Böttcher case," Ann. Inst. Henri Poincaré, Probab. Stat. **45**, 201–225 (2009).
13. S. Foss, D. Korshunov, and S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions* (Springer, New York, 2011).
14. T. E. Harris, "Branching processes," Ann. Math. Stat. **19**, 474–494 (1948).
15. T. E. Harris, *The Theory of Branching Processes* (Springer, Berlin, 1963).
16. S. V. Nagaev, "Some limit theorems for large deviations," Teor. Veroyatn. Primen. **10** (2), 231–254 (1965) [Theory Probab. Appl. **10**, 214–235 (1965)].
17. S. V. Nagaev and V. I. Vakhtel, "Probability inequalities for a critical Galton–Watson process," Teor. Veroyatn. Primen. **50** (2), 266–291 (2005) [Theory Probab. Appl. **50**, 225–247 (2006)].

Translated by the authors