DOI: 10.1214/EJP.v16-933

Electronic Journal of Probability

Vol. 16 (2011), Paper no. 69, pages 1900–1933.

Journal URL
http://www.math.washington.edu/~ejpecp/

Upper large deviations for Branching Processes in Random Environment with heavy tails

Vincent Bansaye *   Christian Böinghoff †

Abstract

Branching Processes in Random Environment (BPREs) $(Z_n : n \geq 0)$ are the generalization of Galton-Watson processes where 'in each generation' the reproduction law is picked randomly in an i.i.d. manner. The associated random walk of the environment has increments distributed like the logarithmic mean of the offspring distributions. This random walk plays a key role in the asymptotic behavior. In this paper, we study the upper large deviations of the BPRE Z when the reproduction law may have heavy tails. More precisely, we obtain an expression for the limit of $-\log \mathbb{P}(Z_n \geq \exp(\theta n))/n$ when $n \to \infty$. It depends on the rate function of the associated random walk of the environment, the logarithmic cost of survival $\gamma := -\lim_{n\to\infty} \log \mathbb{P}(Z_n > 0)/n$ and the polynomial rate of decay β of the tail distribution of Z_1. This rate function can be interpreted as the optimal way to reach a given "large" value. We then compute the rate function when the reproduction law does not have heavy tails. Our results generalize the results of Böinghoff & Kersting (2009) and Bansaye & Berestycki (2008) for upper large deviations. Finally, we derive the upper large deviations for the Galton-Watson processes with heavy tails. .

Key words: Branching processes, random environment, large deviations, random walks, heavy tails.

AMS 2010 Subject Classification: Primary 60J80, 60K37, 60J05, 92D25.

Submitted to EJP on July 13, 2010, final version accepted September 2, 2011.

*Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, FRANCE, vincent.bansaye@polytechnique.edu, partially supported by ANR Manege

†Department of mathematics, Goethe-university, Fach 187, 60054 Frankfurt/Main, boeinghoff@math.uni-frankfurt.de, partially supported by the German Research Foundation (DFG), Grant DFG-RFBR 08-01-91954

1 Introduction

Branching processes in random environment have been introduced in [3] and [21]. In such processes, for each generation, an offspring distribution is chosen at random, independently of other generations. As an example, we can consider a population of plants in which each plant has a one year life-cycle. Every year the weather conditions (the environment) vary affecting the reproductive success of the plant. Given these conditions, all the plants reproduce independently according to the same mechanism.

BPREs have originally been studied under the assumption of i.i.d. geometric -or more generally linear fractional- offspring distributions [1, 17]. Later on, the case of general offspring distributions has attracted attention [2, 4, 8, 11].

Recently, several results about the large deviations of branching processes in random environment for offspring distributions with weak tails have been proved. More precisely, the exact asymptotics of $\mathbb{P}(Z_n \geq \exp(\theta n))$ for geometric offspring distributions are computed in [18] and [19]. In [5], the authors present a general upper bound for the rate function and compute it in the special case of each individual leaving at least one offspring, i.e. $\mathbb{P}(Z_1 = 0) = 0$. Finally, in [9] an expression of the upper rate function is derived when the reproduction laws have geometrically bounded tails. This obviously excludes heavy tails.

Upper large deviations of BPREs correspond to the exceptional growth of these processes and can be due to an exceptional environment and/or to the exceptional reproduction in a given environment. Thus the motivation is not only to compute the rate function but also to understand the effect of *environmental* and *demographical* stochasticity. The way both effects can contribute to atypical events is a challenging question in theoretical ecology.

In this paper, we focus on the large deviation probabilities when the offspring distributions may have heavy tails and the exceptional reproduction of a single individual can contribute to a large deviation event. For the proofs, new auxiliary power series and higher order derivatives of generating functions are used.

Let us now state the formal definition of the process $(Z_n : n \in \mathbb{N})$, $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, by considering a random probability generating function f and a sequence $(f_n : n \geq 1)$ of i.i.d. realizations of f which serve as the random environment. Conditioned on the environment $(f_n : n \geq 1)$, the individuals at generation n reproduce independently of each other and their offsprings have generating function f_{n+1}. Let Z_n denote the number of particles in generation n and Z_{n+1} is the sum of Z_n independent random variables with generating function f_{n+1}. That is, for every $n \geq 0$,

$$\mathbb{E}\big[s^{Z_{n+1}} | Z_0, \ldots, Z_n;\ f_1, \ldots, f_{n+1}\big] \ =\ f_{n+1}(s)^{Z_n} \text{ a.s.} \qquad (0 \leq s \leq 1).$$

In the whole paper, $\mathbb{P}_k$ denotes the probability associated with k initial particles. Then, for all $k \in \mathbb{N}$ and $n \in \mathbb{N}$, we have

$$\mathbb{E}_k[s^{Z_n} \mid f_1, ..., f_n] \ =\ [f_1 \circ \cdots \circ f_n(s)]^k \quad \text{a.s.} \qquad (0 \leq s \leq 1).$$

Unless otherwise specified, the initial population size is 1.

We introduce the exponential rate of decay of the survival probability

$$\gamma \quad := \quad \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}(Z_n > 0). \tag{1}$$

The fact that the limit exists and $0 \le \gamma < \infty$ is classical since the sequence $(-\log \mathbb{P}(Z_n > 0))_n$ is subadditive and nonnegative (see [15] p. 38, Lemma III.29). Essentially, if $\mathbb{E}[\log f'(1)] \ge 0$, then Z is supercritical or critical and $\gamma = 0$. Otherwise, Z is subcritical and

$$\gamma \;=\; -\log\Big(\inf\Big\{\mathbb{E}\big[(\log f'(1))^s\big] : s \in [0,1]\Big\}\Big) > 0.$$

More precisely, $\gamma = -\log(\mathbb{E}[f'(1)])$ in the strongly or intermediately subcritical case, i.e. $\mathbb{E}[f'(1)\log f'(1)] \le 0$, whereas $\gamma > -\log(\mathbb{E}[f'(1)])$ in the weakly subcritical case, i.e. $\mathbb{E}[f'(1)\log f'(1)] > 0$. We refer to [12] for more precise asymptotic results on the survival probability of subcritical BPREs. For large deviations without heavy tails, it has been already observed in [9] that γ is of importance in the limit theorems only in the strongly subcritical case.

Many properties of Z are mainly determined by the **random walk associated with the environment**

$$S_0 \;=\; 0, \qquad S_n - S_{n-1} \;=\; X_n \quad (n \ge 1),$$

where

$$X_n \quad := \quad \log f_n'(1) \quad (n \ge 1),$$

are i.i.d. copies of the logarithm of the mean number of offsprings

$$X \quad := \quad \log f'(1).$$

If $Z_0 = 1$, we get for the conditioned means of Z_n

$$\mathbb{E}[Z_n | f_1, \ldots, f_n] \quad = \quad e^{S_n} \quad \text{a.s.} \tag{2}$$

In the whole paper, we assume that there exists a $\lambda > 0$ such that the moment generating function $\mathbb{E}[\exp(\lambda X)]$ is finite. Then the rate function ψ of the random walk $(S_n : n \in \mathbb{N})$ is given by

$$\psi(\theta) \quad := \quad \sup_{\lambda \ge 0}\{\lambda\theta - \log(\mathbb{E}[\exp(\lambda X)])\}. \tag{3}$$

As ψ is convex and lower semicontinuous, there is at most one $\theta \ge 0$ with $\psi(\theta) \ne \psi(\theta+)$. In this case, $\psi(\theta+) = \infty$ (see e.g. [10], [15]). Usually, ψ is defined as the Legendre transform of $\log(\mathbb{E}[\exp(\lambda X)])$ and the supremum in (3) is taken over all $\lambda \in \mathbb{R}$. Here, we are only interested in upper deviations, thus setting $\psi(\theta) = 0$ for $\theta \le \mathbb{E}[X]$ is convenient.

Notations: In the whole paper, we denote by $\Pi := (f_1, f_2, \ldots)$ the entire sequence of environments. We write $L = L(f)$ for the random variable associated with the probability generating function f:

$$\mathbb{E}[s^L \mid f] \;=\; f(s) \quad (0 \le s \le 1) \quad \text{a.s.}$$

and by $m = m(f)$ we denote its expectation:

$$m \;:=\; f'(1) \;=\; \mathbb{E}[L|f] < \infty \quad \text{a.s.}$$

Unless specified otherwise, we start the branching process with a single individual and denote by $\mathbb{P}$ the underlying probability measure. We denote by $\mathbb{P}_k$ the probability measure when the initial size of the population is k. As a matter of fact, large deviation results do not depend on the initial number of individuals if the latter is fixed (or bounded).
For notational convenience, we use the symbol $\le_c$ to indicate that the inequality holds up to a multiplicative constant (which does not depend on any variable).
Throughout the paper, we use the convention $0 \cdot \infty = 0$.

2 Main results and interpretation

In this paper, we will describe the upper large deviations of the branching process $(Z_n : n \in \mathbb{N})$ when the offspring distributions may have heavy tails. This means that the probability that one individual gives birth to an exponential number of offsprings is 'only of exponentially small order'. Throughout this paper we will work with the following assumption. It ensures that the tail of the offspring distribution of an individual, conditioned to be positive, decays at least with exponent $\beta \in (1,\infty)$ uniformly with respect to the environments.

Assumption $\mathscr{H}(\beta)$**.** *There exists a constant* $0 < d < \infty$ *such that for every* $z \geq 0$*,*

$$\mathbb{P}(L > z \mid f, L > 0) \quad \leq \quad d \cdot (m \wedge 1) \cdot z^{-\beta} \quad \textit{a.s.}$$

The rate function χ that we establish and interpret below depends on γ, β and ψ and is defined by

$$\chi(\theta) := \inf_{t \in [0,1], s \in [0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\psi((\theta - s)/(1-t)) \Big\} \ \ (= \chi_{\gamma,\beta,\psi}(\theta)). \tag{4}$$

We will actually prove that, apart from the strongly subcritical case, χ simplifies to

$$\chi(\theta) \ = \ \inf_{s \in [0,\theta]} \{\beta s + \psi(\theta - s)\}.$$

Theorem 1. *Assume that for some* $\beta \in (1,\infty)$*,* $\log(\mathbb{P}(Z_1 > z))/\log(z) \stackrel{z\to\infty}{\longrightarrow} -\beta$ *and that additionally* $\mathscr{H}(\beta)$ *holds. Then for every* $\theta \geq 0$*,*

$$\frac{1}{n}\log(\mathbb{P}(Z_n \geq e^{\theta n})) \ \stackrel{n\to\infty}{\longrightarrow} \ -\chi(\theta).$$

The assumptions in this theorem essentially ensure that in a positive probability set of environments, the offspring distributions have polynomial tails with exponent $-\beta$, and no tail distribution exceeds this exponent.

The lower bound is proved in Section 3, while the proof of the upper bound is presented in Sections 4 and 5 by distinguishing the case $\beta \in (1,2]$ and the case $\beta > 2$. The proof for $\beta > 2$ is technically more involved since it requires higher order derivatives of generating functions to get the divergence of the power series $\frac{d^{\lceil\beta\rceil}}{ds^{\lceil\beta\lceil}} \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n \geq k)$, $s \to 1$. In Section 5, we adapt the arguments of the proof for $\beta \in (1,2]$ to the case $\beta > 2$.

Remark: Let us note that we can relax Assumption $\mathscr{H}(\beta)$ by letting d depend on the environment. But this would make the proof more tedious.
Moreover, Theorem 1 still holds if we just assume that there exists a slowly varying function l such that

$$\mathbb{P}(L > z \mid f, L > 0) \leq d \cdot (m \wedge 1) \cdot l(z) \cdot z^{-\beta} \quad \text{a.s.} \tag{5}$$

instead of Assumption $\mathscr{H}(\beta)$. Indeed, the properties of slowly varying functions (see [7], Proposition 1.3.6, p. 16) imply that for any $\epsilon > 0$, there exists a constant d_ϵ such that

$\mathbb{P}(L > z \mid f, L > 0) \le d_\epsilon \cdot (m \wedge 1) \cdot z^{-\beta+\epsilon}$ a.s. For fixed $\theta \ge 0$, $\chi_{\gamma,\beta,\psi}$ is continuous in β. So letting $\epsilon \to 0$ yields the upper bound in the theorem. Finally, the proof of the lower bound only requires that $\mathbb{E}[Z_1 \log^+(Z_1)/f_1'(1)] < \infty$ (see p. 1911) which is assured by (5). □

Let us state two consequences of this result. First, we will derive a large deviation result for offspring distributions without heavy tails by letting $\beta \to \infty$, which generalizes Theorem 1 in [9].

Corollary 1. *If Assumption $\mathscr{H}(\beta)$ is fulfilled for every $\beta > 1$, then for every $\theta \ge 0$,*

$$\varphi(\theta) \le \liminf_{n\to\infty} -\frac{1}{n}\log(\mathbb{P}(Z_n \ge e^{\theta n})) \le \limsup_{n\to\infty} -\frac{1}{n}\log(\mathbb{P}(Z_n \ge e^{\theta n})) \le \varphi(\theta+),$$

where $\varphi(\theta) := \inf_{t\in[0,1]}\left\{t\gamma + (1-t)\psi(\theta/(1-t))\right\}$.

For example, this result holds if the offspring distributions are bounded ($\mathbb{P}(L \ge a \mid f) = 0$ a.s. for some constant a) or if $\mathbb{P}(L > z \mid f,\ L > 0) \le c\exp(-z^\alpha)$ a.s. for some constants $c, \alpha > 0$.

Secondly, Theorem 1 also covers the Galton-Watson case, when the environment is not random and f is deterministic. We refer to [6, 20] for the precise large deviations results without heavy tails.

Corollary 2. *Assume that Z is a Galton-Watson process. If $\log(\mathbb{P}(Z_1 > z))/\log(z) \stackrel{z\to\infty}{\longrightarrow} -\beta$, then for every $\theta \ge \mathbb{E}[Z_1] = m$*

$$\frac{1}{n}\log \mathbb{P}(Z_n \ge e^{\theta n}) \stackrel{n\to\infty}{\longrightarrow} \chi(\theta) = \begin{cases} -\beta\theta + \log m & , \ \textit{if } m < 1 \\ \beta(\log m - \theta) & , \ \textit{if } m \ge 1 \end{cases}.$$

Indeed, in the Galton-Watson case, for $\beta + \epsilon > \beta$, (5) results from $\log(\mathbb{P}(Z_1 > z))/\log(z) \stackrel{z\to\infty}{\longrightarrow} -\beta$. Then, we let $\epsilon \to 0$ to derive this corollary from Theorem 1. Moreover $\psi(\theta) = \infty$ for $\theta > \log m$ and $\psi(\log m) = 0$. For the rate of decay of the survival probability, it is well-known that $\gamma = -\log(f'(1)) = -\log m$ in the subcritical case ($m < 1$) , and $\gamma = 0$ in the critical ($m = 1$) and supercritical ($m > 1$) case. In the subcritical case, $\psi(s) = \infty$ for $s > 0$, implying $t = 1$ in (4). The only way to grow can come from an individual having exceptionally many offsprings. In the supercritical case, $\gamma = 0$ implies $t = 0$ in the infimum in (4), meaning that the process starts growing right from the beginning. It remains to minimize $\chi(\theta) = \inf_{s\in[0,\theta]}\{\beta s + \psi(\theta - s)\}$, where $\psi(\theta) = 0$ for $\theta \le \log m$ and $\psi(\theta) = \infty$ for $\theta > \log m$. Hence, $\chi(\theta) = \beta(\theta - \log m)$. The critical case is obtained similarly.

We give now some complements and interpretations of the results.

The quenched approach The asymptotic behavior of the large deviation probabilities is now considered conditionally on the environment.

Proposition 1. *If $\limsup_{z\to\infty} \log \mathbb{P}(Z_1 > z \mid f_1, Z_0 = 1)/\log z = -\beta$ a.s. for some $\beta \in (1,\infty)$, then for every $\theta \ge (\mathbb{E}[X] \vee 0)$,*

$$\lim_{n\to\infty} \frac{1}{n}\log \mathbb{P}(Z_n \ge e^{\theta n} \mid f_1, f_2, \ldots) \quad = \quad \begin{cases} -\beta\theta + \mathbb{E}[X] & , \ \textit{if } \mathbb{E}[X] < 0 \\ -\beta(\theta - \mathbb{E}[X]) & , \ \textit{if } \mathbb{E}[X] \ge 0 \end{cases} \quad \textit{a.s.}$$

Essentially, this is a modification of the rate function given in Corollary 2 for the Galton-Watson case. Here $\log m$ has been replaced by $\mathbb{E}[X]$, since $S_n/n \to \mathbb{E}[X]$ a.s. Indeed large deviations cannot rely on the stochasticity of the environments any longer and will essentially be realized by one individual having exponentially many offsprings. The upper bound in Proposition 1 follows directly from Theorems 2 and 3. The lower bound can also be easily proved by slightly adapting the proof in Section 3 with the help of the Paley-Zygmund inequality. We defer the details of the proof to Christian Böinghoff's thesis.

Path interpretation of the rate function. Following the terminology in [15], if for a measurable set A,

$$\mathbb{P}(Z_n \in A) = e^{-an+o(n)}$$

then a will be referred to as **cost for** A.
In this terminology, the rate function describes the cost of reaching exceptionally large values, namely

$$\mathbb{P}(Z_n \geq e^{\theta n}) \quad = \quad \exp(-\chi(\theta)n + o(n)).$$

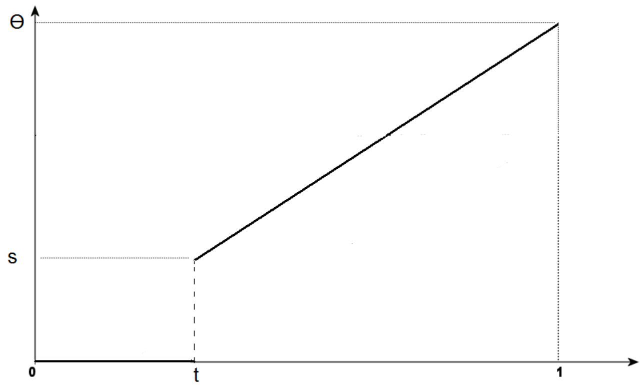
We now describe the paths which lead to exceptionally large values, i.e. paths which realize $\{Z_n \geq \exp(\theta n)\}$ for $n \gg 1$ and $\theta > \mathbb{E}[X]$. In the subcritical case, at the beginning, up to time $\lfloor tn \rfloor$, $t \in [0,1]$, there is a period without growth, during which the process just survives. The probability of this event decreases as $\exp(-\gamma\lfloor tn \rfloor)$. At time $\lfloor tn \rfloor$, there are very few individuals and one individual has exceptionally many offsprings, namely $\exp(sn)$-many, $s \in [0,\theta]$. The probability of this reproduction event is given by $\mathbb{P}(Z_1 \geq \exp(sn)) \sim \exp(-\beta sn)$. Then the process grows exponentially according to its expectation in an exceptionally good environment to reach $\exp(\theta n)$. That is S grows linearly such that $S_n - S_{\lfloor nt \rfloor} \approx [\theta - s]n$ and the probability of observing this exceptionally good environment sequence decreases as $\exp(-(1-t)\psi((\theta-s)/(1-t))n)$. The most probable path reaching exceptionally large values $\exp(\theta n)$ at time n is then obtained by minimizing the sum of these three costs γt, βs and $(1-t)\psi((\theta-s)/(1-t))$. The rate function χ results from this minimization:

$$\chi(\theta) = \inf_{t\in[0,1], s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\psi((\theta-s)/(1-t)) \Big\} = t_\theta\gamma + \beta s_\theta + (1-t_\theta)\psi((\theta-s_\theta)/(1-t_\theta))$$

and thus corresponds to a strategy described by the function

$$f_\theta(t) := \begin{cases} 0, & \text{if } t \leq t_\theta \\ \beta s_\theta + \frac{c}{1-t_\theta}(t-t_\theta), & \text{if } t > t_\theta. \end{cases}$$

Figure 1. *Representation of* $t \in [0,1] \to f_\theta(t)$.



More precisely, following [5], we expect that

$$\sup_{t\in[0,1]} \left\{ \left| \log(Z_{[tn]})/n - f_\theta(t) \right| \right\} \stackrel{n\to\infty}{\longrightarrow} 0$$

in probability. But the proof becomes very technical and cumbersome and we refrain from giving it here.

Actually, convexity arguments given below ensure that the jump occurs either at the beginning or at the end (except in a degenerated case which is explained below). Then $t = 0$ or $t = 1$ in the last picture and the corresponding paths are given by the four paths plotted in Figure 3. More generally, we will describe the most probable trajectories realizing the large deviation events and their depending on θ. For that purpose, we need a new characterization of the rate function χ.

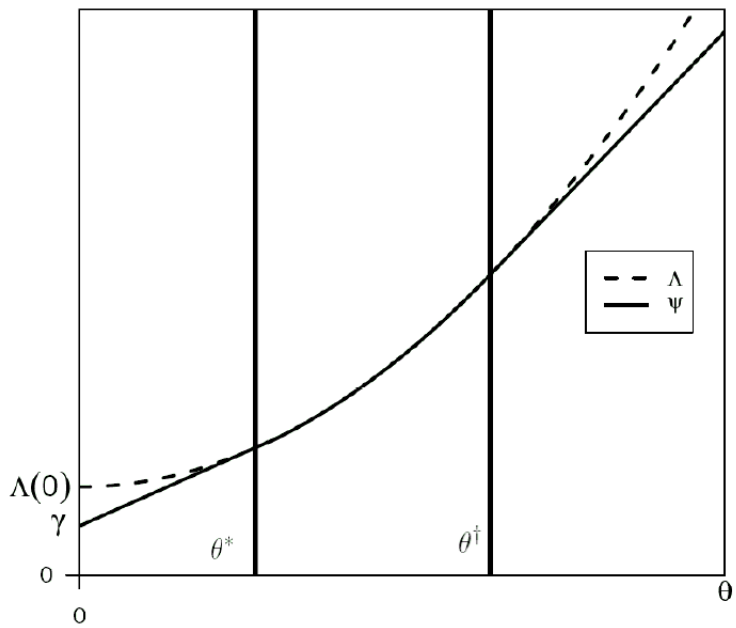
Graphical construction of the rate function. It turns out that χ is the largest convex function satisfying for all $x, \theta \geq 0$ (see appendix, Lemma 3)

$$\chi(0) = \gamma, \quad \chi(\theta) \leq \psi(\theta), \quad \chi(\theta + x) \leq \chi(\theta) + \beta x.$$

The first condition plays a role iff $\psi(0) > \gamma$, which corresponds to the strongly subcritical case. This can be seen in the following way: if $\mathbb{E}[X\exp(X)] < 0$, then the derivative of the map $s \to \mathbb{E}[\exp(sX)]$ in $s = 1$ is negative and $\psi(0) = \sup\{-\log(\mathbb{E}[\exp(sX)]) : s \geq 0\} > -\log(\mathbb{E}[\exp(X)]) = \gamma$. If $\mathbb{E}[X] < 0$ and $\mathbb{E}[X\exp(X)] \geq 0$, the results in [12] and the definition of ψ ensure that both γ and $\psi(0)$ are equal to $-\log(\mathbb{E}[\exp(\nu X)])$ and ν is characterized by $\mathbb{E}[X\exp(\nu X)] = 0$. Finally, if Z is critical or supercritical, i.e. $\mathbb{E}[X] \geq 0$, then $\psi(0) = \gamma = 0$.

Resulting from this characterization, χ can be constructed in three pieces separated by θ^* and $\theta^\dagger$ (see appendix, Lemma 4).

Figure 2. *Illustration of χ in the strongly subcritical case:*



More explicitly, let us define φ as the largest convex function that satisfies $\varphi(0) \leq \gamma$ and $\varphi(\theta) \leq \psi(\theta)$ for all $\theta \geq 0$. As proved in [9], this function is the rate function of Z in the case of the offspring distributions having geometrically bounded tails and is given by

$$\varphi(\theta) = \begin{cases} \gamma\left(1 - \frac{\theta}{\theta^*}\right) + \frac{\theta}{\theta^*}\psi(\theta^*) & , \text{ if } \theta \leq \theta^* \\ \psi(\theta) & , \text{ else} \end{cases}$$

where $0 \leq \theta^* \leq \infty$ is defined by

$$\frac{\psi(\theta^*) - \gamma}{\theta^*} = \inf_{\theta \geq 0} \frac{\psi(\theta) - \gamma}{\theta}. \tag{6}$$

Next we define

$$\theta^\dagger = \sup\left\{\theta \geq \max\{0, \mathbb{E}[X]\} : \varphi'(\theta) \leq \beta \text{ and } \varphi(\theta) < \infty\right\}. \tag{7}$$

Then

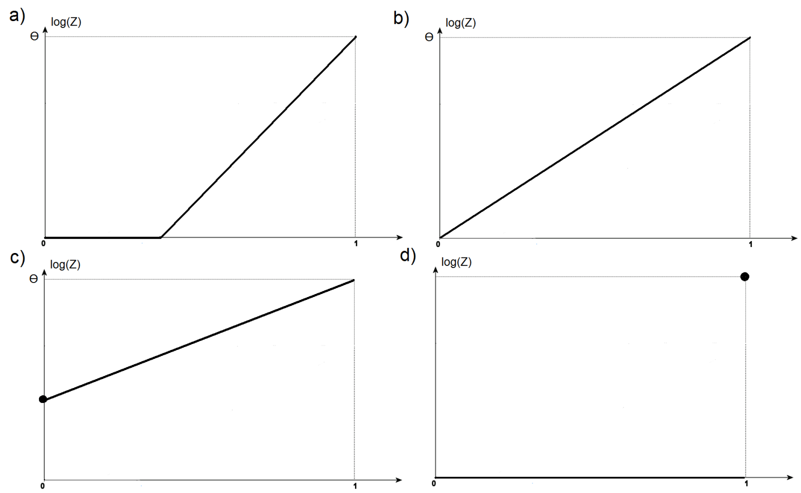
$$\chi(\theta) = \begin{cases} \varphi(\theta) & , \text{ if } \theta < \theta^\dagger \\ \beta\theta - \log(\mathbb{E}[e^{\beta X}]) & , \text{ else} \end{cases} \tag{8}$$

and we get the following expression (see appendix, Lemma 4):

$$\chi(\theta) = \begin{cases} \gamma(1 - \frac{\theta}{\theta^*}) + \frac{\theta}{\theta^*}\psi(\theta^*) & , \text{ if } \theta \leq \theta^* \\ \psi(\theta) & , \text{ if } \theta^* < \theta < \theta^\dagger \\ \beta(\theta - \theta^\dagger) + \psi(\theta^\dagger) & , \text{ if } \theta \geq \theta^\dagger \end{cases}.$$

Thus the trajectories have the following form.

Figure 3. *Representation of the possible trajectories of the path associated with upper large deviations.*



Phase Transitions Let us first describe the phase transitions (of order two) of the rate function χ, which correspond to discontinuities of the second derivative of χ and the associated strategies when $\theta^{\dagger} > 0$.

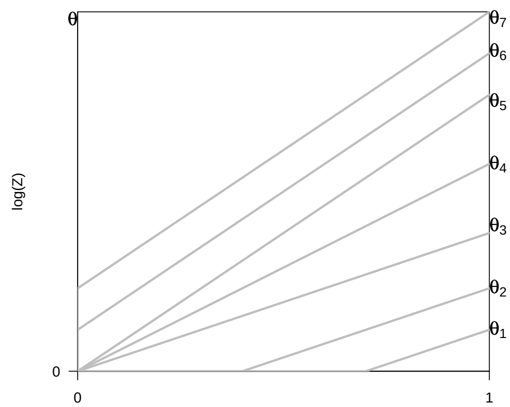
For $\theta < \theta^*$, the rate function χ is identical to φ, which is a convex combination of γ and ψ (and not β). Thus, conditioned on the event $\{Z_n \geq \exp(\theta n)\}$, the process first 'just survives with bounded values' until time $\lfloor t_\theta n \rfloor$ ($t_\theta \in (0,1)$). Then it grows in a good environment such that $S_n - S_{\lfloor t_\theta n \rfloor} \approx \theta n$ (see Figure 3 a)). When θ increases, the survival period decreases while the exponential growth rate of the process remains constant and is equal to θ^*. Large deviation events are typically not realized by a reproduction event with exponentially many offsprings.

For $\theta^* \leq \theta \leq \theta^{\dagger}$, χ is equal to ψ. Thus, conditionally on the large deviation event, the process grows exponentially (respectively linearly at the logarithmic scale) from the beginning to the end (see Figure 3 b)). This exceptional growth is due to a favorable environment such that $S_n \approx \theta n$.

For $\theta > \theta^{\dagger}$, the trajectory associated with the optimal strategy begins with one individual having exponentially many offsprings: $Z_1 \approx \exp(sn)$. Then it grows exponentially in a favorable environment such that $S_n \approx (\theta - s)n$ (see Figure 3 c)). When θ increases, the initial jump increases while the rate of the exponential growth is still equal to $\theta^{\dagger}$.

The case $\theta^{\dagger} = 0$ corresponds to $\chi(\theta) = \gamma + \beta\theta$. Here the optimal strategy consists in just surviving until the end and one individual having $\exp(\theta n)$-many offsprings in one of the last generations (see Figure 3 d)).
Finally, we note that in the case $0 < \theta = \theta^* = \theta^{\dagger}$, the optimal strategy is no longer unique. Indeed, for any $t \in (0,1]$, there exists an $s \in [0,\theta]$ such that all the following trajectories have the same probability: First, the process remains positive and bounded until time $\lfloor tn \rfloor$ (survival period), then it 'jumps' to $\exp(sn)$ and grows exponentially with a constant rate (see Figure 1).

Figure 4. *Representation of $t \in [0,1] \to f_\theta(t)$ in the strongly subcritical case for $\theta_1 < \theta_2 < \theta_3 = \theta^* < \theta_4 < \theta_5 = \theta^{\dagger} < \theta_6 < \theta_7$.*



In the next sections, we will prove the results. In Sections 3, 4, 5 and 6, relations stated conditioned on the environment hold a.s. with respect to the underlying probability measure. We refrain from indicating this in every equation. Section 7 is the appendix which contains several technical results used in the proof.

3 Proof of the lower bound of Theorem 1

For the proof of the lower bound of Theorem 1, we need the following result. It ensures that exceptional growth of the population can at least be achieved by some good environment sequences, whose probability decreases exponentially according to the rate function of the random walk $(S_n : n \in \mathbb{N})$. This result generalizes Proposition 1 in [5] for an exponential initial number of individuals. With a slight abuse of notation, we will write below $\exp(sn)$ for the initial number of individuals instead of the integer part of $\exp(sn)$.

Proposition 2. *Under Assumption $\mathscr{H}(\beta)$, for all $\theta \geq 0$ and $0 \leq s \leq \theta$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\exp(sn)}(Z_n \geq \exp(\theta n)) \quad \geq \quad -\psi((\theta - s)+).$$

Proof. Without loss of generality, we restrict ourselves to the case $\psi((\theta - s)+) < \infty$. Recall that for every $\theta' > 0$,

$$\psi(\theta') \ = \ \sup_{\lambda \geq 0} \Big\{ \lambda\theta' - \log \mathbb{E}[\exp(\lambda X)] \Big\}.$$

First, we assume that $\mathbb{E}[\exp(\lambda X)] < \infty$ for every $\lambda \geq 0$. Then the derivative of the map $\lambda \to \mathbb{E}[\exp(\lambda X)]$ exists for every $\lambda \geq 0$. The supremum above is achieved at $\lambda = \lambda_{\theta'}$ satisfying

$$\theta' \ = \ \frac{\mathbb{E}[X \exp(\lambda_{\theta'} X)]}{\mathbb{E}[\exp(\lambda_{\theta'} X)]}.$$

Pursuing classical large deviation techniques (or more specifically the proof in [5]), we introduce the probability $\widetilde{\mathbb{P}}$ defined by

$$\widetilde{\mathbb{P}}(X \in \mathrm{d}x) \;=\; \frac{\exp(\lambda_{\theta'} x)}{\mathbb{E}[\exp(\lambda_{\theta'} X)]}\mathbb{P}(X \in \mathrm{d}x).$$

Under this new probability measure, $(S_n : n \in \mathbb{N})$ is a random walk with drift $\widetilde{\mathbb{E}}[X] = \theta' > 0$ and Z_n is a supercritical BPRE.

Let us fix $\theta \geq 0$ and $0 \leq s \leq \theta$. For all $n \geq 1$, $\theta' > \theta - s$ and $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}_{\exp(sn)}\big(Z_n \geq \exp(\theta n)\big)\\
&\quad \geq \mathbb{P}_{\exp(sn)}\big(Z_n \geq \exp(\theta n); S_n \leq (\theta' + \epsilon)n\big)\\
&\quad = \mathbb{E}[\exp(\lambda_{\theta'} X)]^n \widetilde{\mathbb{E}}_{\exp(sn)}\big[\exp(-\lambda_{\theta'} S_n)\mathbb{1}_{\{\, Z_n \geq \exp(\theta n), S_n \leq (\theta'+\epsilon)n\}}\big]\\
&\quad \geq \exp\big(n\big[\log(\mathbb{E}[\exp(\lambda_{\theta'} X)] - \lambda_{\theta'}(\theta' + \epsilon)\big]\big)\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq \exp(\theta n),\ S_n \leq (\theta' + \epsilon)n\big)\\
&\quad \geq \exp(n[-\psi(\theta') - \lambda_{\theta'}\epsilon])\big[\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq \exp(\theta n)\big) - \widetilde{\mathbb{P}}\big(S_n > (\theta' + \epsilon)n\big)\big].
\end{aligned}$$

As $\widetilde{\mathbb{P}}\big(S_n > (\theta' + \epsilon)n\big) \to 0$ for $n \to \infty$, it remains to prove that

$$\liminf_{n\to\infty} \widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq \exp(\theta n)\big) \;>\; 0. \tag{9}$$

The statement of the proposition results from letting $\epsilon \to 0$ and $\theta' \to \theta - s$.

Relation (9) results from the fact that under $\widetilde{\mathbb{P}}$ the population Z_n starting from a single individual grows roughly as $\exp(S_n) = \exp(\theta' n + o(n))$ on the nonextinction event. To prove (9), we label the individuals of the initial population and denote the number of descendants in generation n of individual i by $Z_n^{(i)}$. Let us fix $N \in \mathbb{N}$ and introduce the 'success' probability p_n:

$$p_n \;:=\; \mathbb{P}_1(Z_n \geq N\exp(n(\theta - s)) \mid \Pi) \quad \text{a.s.}$$

Then, conditionally on Π, the number of initial individuals whose number of descendants in generation n is larger than $N\exp(n(\theta - s))$, given by

$$N_n \;:=\; \#\{1 \leq i \leq \exp(sn) : Z_n^{(i)} \geq N\exp(n(\theta - s))\},$$

follows a binomial distribution with parameters $(\exp(sn), p_n)$. Moreover, since $\mathbb{E}[N_n \mid \Pi] = \exp(sn)p_n$ a.s., we obtain

$$\widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq \exp(\theta n)\big) \;\geq\; \widetilde{\mathbb{P}}_{\exp(sn)}\big(N_n \geq \exp(sn)/N\big) \geq \widetilde{\mathbb{P}}_{\exp(sn)}\Big(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{Np_n}\Big).$$

The classical Paley and Zygmund inequality (see e.g. [16] p. 63) given for $r \in [0,1]$ by

$$\mathbb{P}(Y \geq r\mathbb{E}[Y]) \;\;\geq\;\; (1-r)^2\frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]}\,, \tag{10}$$

and the fact that $\mathbb{E}[N_n^2 \mid \Pi] = e^{2sn}p_n^2 + e^{sn}p_n(1-p_n)$ a.s., ensure

$$\widetilde{\mathbb{P}}_{\exp(sn)}\Big(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{Np_n} \;\Big|\; \Pi\Big) \geq \Big[1 - 1 \wedge \frac{1}{Np_n}\Big]^2 \frac{\mathbb{E}[N_n \mid \Pi]^2}{\mathbb{E}[N_n^2 \mid \Pi]} \geq \frac{\Big[1 - 1 \wedge \frac{1}{Np_n}\Big]^2}{1 + \frac{e^{-sn}}{p_n}} \quad \text{a.s.}$$

Observe that under Assumption $\mathscr{H}(\beta)$,

$$\mathbb{E}[Z_1 \log^+(Z_1)/f_1'(1)] < \infty.$$

Theorem 1 in [3] ensures that $Z_n \exp(-S_n)$ converges a.s. to a random variable which is positive on the nonextinction event. So for every $N \in \mathbb{N}$,

$$\widetilde{\mathbb{E}}[p_n] \;=\; \widetilde{\mathbb{P}}_1(Z_n \geq N\exp(\theta n)) \overset{n\to\infty}{\longrightarrow} \widetilde{\mathbb{P}}_1(\forall n \in \mathbb{N} : Z_n > 0) > 0.$$

As the right-hand side does not depend on N and $\widetilde{\mathbb{E}}[p_n] \leq \widetilde{\mathbb{P}}(p_n \geq \alpha) + \alpha$ for every $\alpha > 0$, it follows that for N large enough that

$$\delta \;:=\; \liminf_{n\to\infty} \widetilde{\mathbb{P}}(p_n \geq 2/N) > 0.$$

Thus we get for such N that

$$\liminf_{n\to\infty} \widetilde{\mathbb{P}}_{\exp(sn)}\big(Z_n \geq \exp(\theta n)\big) \;\geq\; \liminf_{n\to\infty} \widetilde{\mathbb{E}}\Big[\frac{\big[1 - 1 \wedge 1/Np_n\big]^2}{1 + 1/p_n}\Big] \geq \frac{\delta(1-1/2)^2}{1+N/2} > 0,$$

which proves (9) and finishes the proof when $\mathbb{E}[\exp(\lambda X)] < \infty$ for every $\lambda \geq 0$. The general case follows by a standard approximation argument, see e.g. [9] p. 2075/2076 □

Proof of the lower bound in Theorem 1. The proof now amounts to exhibiting good trajectories which realize the large deviation event $\{Z_n \geq \exp(\theta n)\}$. By Markov property, for every $t \in (0,1)$ and $s \in [0, \theta]$,

$$\mathbb{P}(Z_n \geq \exp(\theta n)) \;\geq\; \mathbb{P}(Z_{[tn]} > 0)\mathbb{P}(Z_1 \geq \exp(sn))\mathbb{P}_{\exp(sn)}(Z_{n-[tn]} \geq \exp(\theta n)).$$

By (1), we get that

$$\frac{1}{tn}\log(\mathbb{P}(Z_{[tn]} > 0)) \;\overset{n\to\infty}{\longrightarrow}\; -\gamma.$$

Using that $\log(\mathbb{P}(Z_1 > z))/\log(z) \overset{z\to\infty}{\longrightarrow} -\beta$ yields

$$\frac{1}{n}\log(\mathbb{P}(Z_1 \geq \exp(sn))) \;\overset{n\to\infty}{\longrightarrow}\; -s\beta.$$

Finally, by Proposition 2, we obtain that

$$\liminf_{n\to\infty} \frac{1}{(1-t)n}\log(\mathbb{P}_{\exp(sn)}(Z_{n-[tn]} \geq \exp(\theta n))) \;\geq\; -\psi((\theta - s)/(1-t)+)$$

since

$$\mathbb{P}_{\exp(sn)}(Z_{n-[tn]} \geq \exp(\theta n)) \;=\; \mathbb{P}_{\exp((1-t)ns/(1-t))}(Z_{n-[tn]} \geq \exp(n(1-t)\theta/(1-t))).$$

Combining the first inequality and the last three limits ensures that

$$\liminf_{n\to\infty} \frac{1}{n} \log(\mathbb{P}(Z_n \geq \exp(\theta n))) \;\geq\; -\inf_{t\in[0,1],s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\psi((\theta-s)/(1-t)+) \Big\}.$$

The convex nonnegative function, ψ can have at most one jump and this jump goes to infinity. Thus the infimum above is $\chi(\theta)$. To see this, we only have to consider the jump point. Suppose that exist $s_\theta \in [0,\theta]$ and $t_\theta \in [0,1)$ such that

$$t_\theta\gamma + \beta s_\theta + (1-t_\theta)\psi((\theta-s_\theta)/(1-t_\theta)) \;=\; \chi(\theta) < \infty$$

and $\psi((\theta-s_\theta)/(1-t_\theta)+) = \infty$. Then, as $(\theta-s_\theta)/(1-t_\theta)$ is the only jump point, for any $\epsilon > 0$ there is a $\delta = \delta(\epsilon) > 0$ such that

$$\begin{aligned} \chi(\theta) - \epsilon &\leq t_\theta\gamma + \beta(s_\theta - \delta) + (1-t_\theta)\psi((\theta - s_\theta - \delta)/(1-t_\theta)+) \\ &= t_\theta\gamma + \beta(s_\theta - \delta) + (1-t_\theta)\psi((\theta - s_\theta - \delta)/(1-t_\theta)). \end{aligned}$$

Letting $\epsilon, \delta(\epsilon) \to 0$ proves the result and hence the lower bound of Theorem 1. □

4 Proof of the upper bound of Theorem 1 for $\beta \in (1,2]$

Let us now introduce the minimum and maximum of the associated random walk up to time n:

$$L_n := \min_{0\leq k\leq n} S_k, \qquad M_n := \max_{0\leq k\leq n} S_k$$

By Markov inequality,

$$\mathbb{P}(Z_n > 0|\Pi) \;\leq\; \mathbb{E}[Z_n|\Pi] = \exp(S_n) \quad \text{a.s.}$$

Furthermore $\mathbb{P}(Z_n > 0|\Pi)$ is decreasing a.s. and we get the following classical inequality, which will indeed be a good estimate of the rate of decay of the survival probability

$$\mathbb{P}(Z_n > 0|\Pi) \leq e^{L_n} \quad \text{a.s.} \tag{11}$$

For the proof of the upper bound of Theorem 1, we require the following key result for the tail probability of Z_n.

Theorem 2. *Under Assumption $\mathscr{H}(\beta)$ for some $\beta \in (1,2]$, there exist a constant $0 < c < \infty$ and a positive nondecreasing slowly varying function Υ such that for all $k \geq 1$ and $n \geq 1$,*

$$\mathbb{P}(Z_n > k|\Pi) \;\leq\; cn\Upsilon(kn^{2/(\beta-1)}e^{M_n-L_n})e^{L_n}(e^{S_n-L_n}/k)^\beta \quad a.s.$$

Let us briefly explain this result. The probability to survive until time n evolves as e^{L_n}. Conditioned on survival, a good environment sequence corresponds to large values of $(S_n - L_n)$. The possibility of high reproduction of the initial individual is reflected by the last term, $k^{-\beta}$. Conditioned on the environment and survival, the expected size of the process at time n is of order $e^{S_n-L_n}$, which corresponds to a period of exponential growth of the process. Thus, this theorem essentially says that conditioned on $Z_n > 0$, the tail distribution of $Z_n/e^{S_n-L_n}$ decays at least polynomially with exponent

$-\beta$.

Recalling that $\Pi = (f_1, f_2, ...)$ and $f_n(s)$ is the probability generating function of the offspring distribution of an individual in generation $n-1$, we have

$$f_{0,n}(s) \;:=\; \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n = k|\Pi) = \mathbb{E}[s^{Z_n}|\Pi] \quad \text{a.s.} \quad (0 \le s \le 1). \tag{12}$$

For the proofs, it is suitable to work with an alternative expression, namely for every $n \ge 1$,

$$g_n(s) \;:=\; \frac{1 - f_n(s)}{1-s} \quad \text{a.s.} \quad (0 \le s < 1).$$

The function g_n has already been introduced in [9]. As $\lim_{s \nearrow 1} g_n(s) = f_n'(1)$ exists, we define $g_n(1) := f_n'(1)$. Using the representation of the function as power series, let us define

$$g_{0,n}(s) \;\;:=\;\; \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n > k|\Pi) = \frac{1 - f_{0,n}(s)}{1-s} \quad \text{a.s.} \quad (0 \le s \le 1). \tag{13}$$

Moreover we need the following auxiliary function defined for every $\mu \in (0,1]$ by

$$h_{\mu,k}(s) \;:=\; \frac{1}{(1-f_k(s))^{\mu}} - \frac{1}{(f_k'(1)(1-s))^{\mu}} = \frac{g_k(1)^{\mu} - g_k(s)^{\mu}}{(g_k(1) g_k(s)(1-s))^{\mu}} \quad \text{a.s.} \quad (0 \le s \le 1). \tag{14}$$

Finally, we define for all $0 \le k \le n$,

$$\begin{aligned} U_k \;&:=\; \big(f_1'(1) \cdots f_k'(1)\big)^{-1} \;=\; f_{0,k}'(1)^{-1} \;=\; e^{-S_k}, \quad 0 < k \le n;\; U_0 := 1 \\ f_{k,n} \;&:=\; f_{k+1} \circ f_{k+2} \circ \cdots \circ f_n, \quad 0 \le k < n;\; f_{n,n} := id \quad \text{a.s.} \end{aligned}$$

By a telescoping summation argument similar to the one used in [11], we get that

$$\begin{aligned} \frac{1}{(1-f_{0,n}(s))^{\mu}} &= \frac{U_0^{\mu}}{(1-f_{0,n}(s))^{\mu}} \\ &= \frac{U_n^{\mu}}{(1-f_{n,n}(s))^{\mu}} + \sum_{k=0}^{n-1} \left(\frac{U_k^{\mu}}{(1-f_{k,n}(s))^{\mu}} - \frac{U_{k+1}^{\mu}}{(1-f_{k+1,n}(s))^{\mu}} \right) \\ &= \frac{U_n^{\mu}}{(1-s)^{\mu}} + \sum_{k=0}^{n-1} U_k^{\mu} \left(\frac{1}{(1-f_{k+1}(f_{k+1,n}(s)))^{\mu}} - \frac{1}{(f_{k+1}'(1)(1-f_{k+1,n}(s)))^{\mu}} \right) \\ &= \frac{U_n^{\mu}}{(1-s)^{\mu}} + \sum_{k=0}^{n-1} U_k^{\mu} h_{\mu,k+1}(f_{k+1,n}(s)), \quad s \ge 0. \end{aligned} \tag{15}$$

Proof of Theorem 2. In the same vein as in [9], we will obtain an upper bound for $\mathbb{P}(Z_n > z|\Pi)$ from the divergence of $g_{0,n}'(s) = \sum_{j=0}^{\infty} j \mathbb{P}(Z_n > j|\Pi) s^{j-1}$ as $s \to 1$. More precisely, for all $k \ge 1$ and

$s \in [0,1]$,

$$
\begin{aligned}
g'_{0,n}(s) &\geq \sum_{j=1}^{k} j\mathbb{P}(Z_n > j|\Pi)s^{j-1} \\
&\geq s^k \frac{k^2}{2}\mathbb{P}(Z_n > k|\Pi). \qquad (16)
\end{aligned}
$$

To get an upper estimate for $g'_{0,n}(s)$, we will use (15) with $\mu = \beta - 1$. This yields

$$
g_{0,n}(s) \;=\; \Big(U_n^{\beta-1} + (1-s)^{\beta-1}\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s))\Big)^{-1/(\beta-1)} \quad (0 \leq s \leq 1).
$$

Calculating the first derivative of $g_{0,n}$, we get that:

$$
\begin{aligned}
& g'_{0,n}(s) \\
={}& -(\beta-1)^{-1}\Big(U_n^{\beta-1} + (1-s)^{\beta-1}\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s))\Big)^{-1-1/(\beta-1)} \\
& \times\Big(-(\beta-1)(1-s)^{\beta-2}\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s)) \\
& \qquad + (1-s)^{\beta-1}\sum_{k=0}^{n-1} U_k^{\beta-1} h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)\Big) \\
={}& \frac{\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s)) - (\beta-1)^{-1}(1-s)\sum_{k=0}^{n-1} U_k^{\beta-1} h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)}{(1-s)^{2-\beta}\Big(U_n^{\beta-1} + (1-s)^{\beta-1}\sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s))\Big)^{1+1/(\beta-1)}} \\
\leq{}& \frac{\sum_{k=0}^{n-1} U_k^{\beta-1}\Big(h_{\beta-1,k+1}(f_{k+1,n}(s)) - (\beta-1)^{-1} h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)(1-s)\Big)}{U_n^{\beta}(1-s)^{2-\beta}}. \qquad (17)
\end{aligned}
$$

In the last step, we have used (14) to get $h_{\beta-1,k+1}(s) \geq 0$ for all $s \geq 0$ to estimate the denominator. Then Lemma 5 in the appendix ensures that there exist a $c > 0$ and a slowly varying function Υ such that for every $s \in [0,1)$,

$$
\begin{aligned}
h_{\beta-1,k}(s) &\leq c\Upsilon(1/(1-s)), \qquad (18) \\
-h'_{\beta-1,k}(s) &\leq c\Upsilon(1/(1-s))/(1-s) \quad \text{a.s.} \qquad (19)
\end{aligned}
$$

Moreover, adapting (15) to $f_{k+1,n}(s)$ instead of $f_{0,n}(s)$ yields

$$
\frac{1}{(1-f_{k+1,n}(s))^{\mu}} \;=\; \frac{e^{-\mu(S_n - S_{k+1})}}{(1-s)^{\mu}} + \sum_{j=k+1}^{n-1} e^{\mu(S_j - S_{k+1})} h_{\mu,k+1}(f_{k+1,n}(s)), \quad s \geq 0.
$$

Applying (45) in the appendix for $0 < \mu < \beta - 1$ to bound $h_{\mu,k}$, we can conclude that there exists a $c \geq 1$ such that for every $s \in [0,1)$,

$$
\frac{1}{(1-f_{k+1,n}(s))^{\mu}} \;\leq\; \frac{e^{-\mu(S_n - S_{k+1})}}{(1-s)^{\mu}} + n\; c\; e^{\mu \max_{k+1 \leq j \leq n}(S_n - S_j)} \leq c\; e^{\mu(M_n - L_n)}(n+1)/(1-s)^{\mu}.
$$

Combining this inequality with (18) yields

$$h_{\beta-1,k+1}(f_{k+1,n}(s)) \quad \leq \quad c\Upsilon\big((n+1)^{1/\mu}e^{M_n-L_n}(1-s)^{-1}\big) \quad (0\leq s<1).$$

for some $c>0$. Moreover, $f_{k+1,n}(s)\leq 1-f'_{k+1,n}(s)(1-s)$ by convexity of $f_{k+1,n}$ and thus (19) ensures that

$$\begin{aligned} -h'_{\beta-1,k+1}(f_{k+1,n}(s))f'_{k+1,n}(s)(1-s) \quad &\leq \quad c\, f'_{k+1,n}(s)(1-s)\Upsilon\big(1/(1-f_{k+1,n}(s))\big)\frac{1}{1-f_{k+1,n}(s)} \\ &\leq \quad c\,\Upsilon\big((n+1)^{1/\mu}e^{M_n-L_n}(1-s)^{-1}\big) \quad \text{a.s.} \quad (0\leq s<1). \end{aligned}$$

From the two last estimates with $\mu=(\beta-1)/2$ in (17), we get that for every $s\in[0,1]$,

$$g'_{0,n}(s) \quad \leq \quad \frac{c\;n\;e^{-(\beta-1)L_n}\Upsilon\big((n+1)^{2/(\beta-1)}e^{M_n-L_n}(1-s)^{-1}\big)}{U_n^{\beta}(1-s)^{2-\beta}}.$$

By setting $s=1-1/k$ in the estimate above and using (16), we obtain that

$$\Big(1-\frac{1}{k}\Big)^{k}\frac{k^2}{2}\mathbb{P}(Z_n>k|\Pi) \quad \leq \quad c\frac{n\;e^{-(\beta-1)L_n}k^{2-\beta}\;\Upsilon\big(k(n+1)^{2/(\beta-1)}e^{M_n-L_n}\big)}{U_n^{\beta}},$$

which completes the proof since $U_n=\exp(-S_n)$. □

For the proof of the upper bound of Theorem 1, we also need the following characterization of the cost of survival γ:

Lemma 1. *Under Assumption $\mathscr{H}(\beta)$, for all $\theta\geq 0$, $b>0$ and Υ a positive nondecreasing slowly varying function at infinity, we have*

$$\gamma \quad = \quad -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[\Upsilon(n^be^{\theta n}e^{-L_n})e^{L_n}\big].$$

Proof of Lemma 1. First let $\Upsilon=1$. Then using (11) as the upper bound, and (15) with some $0<\mu<\beta-1$ along with (45) as the lower bound ensures that

$$\begin{aligned} e^{L_n} \;\geq\; \mathbb{P}(Z_n>0|\Pi) \quad &\geq \quad \frac{1}{(e^{-\mu S_n}+\sum_{k=0}^{n-1}e^{-\mu S_k}h_{\mu,k+1}(f_{k+1,n}(0)))^{1/\mu}} \\ &\geq \quad c^{-1}n^{-1/\mu}e^{L_n}. \end{aligned}$$

This yields

$$\gamma \quad = \quad -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n>0)=-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[e^{L_n}\big].$$

As Υ is nondecreasing and $-L_n$ (and thus $n^be^{\theta n}e^{-L_n}$) is also nondecreasing,

$$\gamma \;=\lim_{n\to\infty}-\frac{1}{n}\log\mathbb{E}\big[e^{L_n}\big]\geq\;\limsup_{n\to\infty}-\frac{1}{n}\log\mathbb{E}[\Upsilon(n^be^{\theta n}e^{-L_n})e^{L_n}].$$

For the converse inequality, we use that $\mathbb{E}\left[e^{tL_n}\right]$ is nonincreasing in n to define

$$\xi(t) \;\; := \;\; -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[e^{tL_n}\right].$$

Note that $\xi(t)\geq 0$ and by Lemma V.4 in [15], $\xi(t)$ is finite and convex. Thus ξ is continuous. Now by properties of slowly varying sequences, for any $\delta>0$, $x^{-\delta}\Upsilon(x)\to 0$ as $x\to\infty$ (see [7], Proposition 1.3.6, p. 16) and thus

$$\liminf_{n\to\infty} -\frac{1}{n}\log\mathbb{E}\left[\Upsilon(n^b e^{\theta n}e^{-L_n})e^{L_n}\right] \;\geq\; -\delta\theta - \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\left[e^{(1+\delta)L_n}\right].$$

Letting $\delta\to 0$ and using the continuity of ξ finishes the proof. □

Proof of the upper bound of Theorem 1. First, recall the following classical large deviation inequality:

$$\mathbb{P}(S_n\geq\theta n) \;\leq\; e^{-\psi(\theta)n}. \tag{20}$$

Let us define the first time τ_n when the random walk $(S_i : i\leq n)$ reaches its minimum value on $[0,n]$:

$$\tau_n \;:=\; \inf\{0\leq k\leq n:\; S_k=L_n\}.$$

We decompose the probability of having an exponentially large population according to the values S_n-L_n. To control the term in the slowly varying function in Theorem 2, we also add a term bounding the maximum of the random walk up to time n. Let $r\in\mathbb{N}$. Then

$$\begin{aligned}\mathbb{P}(Z_n\geq e^{\theta n}) \;\;=\;\; & \mathbb{P}(Z_n\geq e^{\theta n}, S_n-L_n\geq\theta n)+\mathbb{E}[\mathbb{P}(Z_n\geq e^{\theta n}|\Pi); S_n-L_n<\theta n, M_n\leq r\theta n]\\ & +\mathbb{E}[\mathbb{P}(Z_n\geq e^{\theta n}|\Pi); S_n-L_n<\theta n, M_n>r\theta n].\end{aligned} \tag{21}$$

The asymptotic of the first term can be found using (20) (see [9]):

$$\begin{aligned}\mathbb{P}(Z_n\geq e^{\theta n},\; S_n-L_n\geq\theta n) \;\;\leq\;\; & \sum_{i=1}^{n}\mathbb{P}(Z_i>0)\mathbb{P}(S_n-S_i\geq\theta n)\\ \leq\;\; & \sum_{i=1}^{n}\mathbb{P}(Z_i>0)\exp(-(n-i)\psi(\theta n/(n-i))).\end{aligned}$$

This ensures that

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(Z_n\geq e^{\theta n}, S_n-L_n\geq\theta n) \;\leq\; -\varphi(\theta), \tag{22}$$

where we recall that

$$\varphi(\theta) \;\;=\;\; \inf_{0<t\leq 1}\left\{t\gamma+(1-t)\psi(\theta/(1-t))\right\}.$$

For the second term in (21), we use Theorem 2 and the Markov property for $(S_n : n \geq 0)$. For every $r \in \mathbb{N}$,

$$
\begin{aligned}
&\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi);\ S_n - L_n < \theta n,\ M_n \leq r\theta n] \\
&\quad \leq c\ n\ \mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)}e^{M_n - L_n}e^{\theta n})e^{L_n}e^{\beta(S_n - L_n - \theta n)}; S_n - L_n < \theta n, M_n \leq r\theta n\Big] \\
&\quad \leq c\ n\ \sum_{k=0}^{n}\mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)}e^{-S_k}e^{(r+1)\theta n})e^{S_k}e^{\beta(S_n - S_k - \theta n)}; S_n - L_n < \theta n, \tau_n = k\Big] \\
&\quad \leq c\ n\ \sum_{k=0}^{n}\mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-S_k}e^{(r+1)\theta n})e^{S_k}; \tau_k = k]\mathbb{E}[e^{-\beta(\theta n - S_{n-k})}; S_{n-k} < \theta n, L_{n-k} \geq 0].
\end{aligned}
$$

Let $\epsilon = 1/n^2$ and $m_\epsilon = \lceil \theta/\epsilon \rceil$. Note that

$$
\begin{aligned}
\mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-S_k}e^{(r+1)\theta n})e^{S_k}; \tau_k = k] &= \mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-L_k}e^{(r+1)\theta n})e^{L_k}, \tau_k = k] \\
&\leq \mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-L_k}e^{(r+1)\theta n})e^{L_k}],
\end{aligned}
$$

and hence from (20), we deduce that

$$
\begin{aligned}
&\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n] \\
\leq\ & c\ n \sum_{k=1}^{n}\mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-L_k}e^{(r+1)\theta n})e^{L_k}]\sum_{j=0}^{m_\epsilon} e^{-\beta(\theta-(j+1)\epsilon)n}\mathbb{P}\big(S_{n-k} \in [nj\epsilon, n(j+1)\epsilon), L_{n-k} \geq 0\big) \\
\leq\ & c\ n \sum_{k=1}^{n}\mathbb{E}[\Upsilon(n^{2/(\beta-1)}e^{-L_k}e^{(r+1)\theta n})e^{L_k}]\sum_{j=0}^{m_\epsilon} e^{-\beta(\theta-(j+1)\epsilon)n}e^{-\psi(j\epsilon n/(n-k))(n-k)} \\
\leq\ & c\ \theta\ n^4 \sup_{0<t\leq 1, 0\leq s\leq \theta}\Big\{\mathbb{E}\Big[\Upsilon(n^{2/(\beta-1)}e^{-L_{\lfloor tn\rfloor}}e^{(r+1)\theta n})e^{L_{\lfloor tn\rfloor}}\Big]\cdot e^{-(\beta s + (1-t)\psi((\theta-s)/(1-t)))n}\Big\}.
\end{aligned}
$$

Together with Lemma 1, this yields that for every $r \in \mathbb{N}$,

$$
\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n \leq r\theta n] \quad \leq \quad -\chi(\theta), \tag{23}
$$

where

$$
\chi(\theta) \quad = \quad \inf_{0<t\leq 1, 0\leq s\leq\theta}\Big\{\gamma t + \beta s + (1-t)\psi((\theta-s)/(1-t))\Big\}.
$$

As to the third term in (21), by duality,

$$
\begin{aligned}
\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n}|\Pi); S_n - L_n < \theta n, M_n > r\theta n] &\leq \mathbb{P}(M_n > r\theta n) \\
&= \mathbb{P}(\max_{k=0,\ldots,n}(S_n - S_k) > r\theta n) = \mathbb{P}(S_n - L_n > r\theta n).
\end{aligned} \tag{24}
$$

It has been proved in [9] (see p. 2068) that,

$$
\varphi_0(x) := -\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}(S_n - L_n > xn) = \inf_{0<t\leq 1}\{(1-t)\psi(x/(1-t))\} \overset{x\to\infty}{\longrightarrow} \infty.
$$

Combining this result with (21), (22), (23) and (24) shows that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \;\leq\; -\min\{\varphi(\theta); \chi(\theta); \varphi_0(r\theta)\}.$$

Observe that $\chi(\theta) \leq \varphi(\theta)$ since the infimum is considered on a larger set for χ than for φ. Adding that $\varphi_0(x) \to \infty$ as $x \to \infty$, we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \;\leq\; -\chi(\theta)$$

by letting $r \to \infty$. This proves the upper bound of Theorem 1. □

5 Adaptation of the proof of the upper bound for $\beta > 2$

First, note that Lemma 1 still holds for $\beta > 2$ by following the same proof. Indeed, using (15) for $\mu = 1$ together with Lemma 5 stated in the appendix ensures that

$$\mathbb{P}(Z_n > 0|\Pi) = 1 - f_{0,n}(0) \;\;\geq\;\; \frac{1}{e^{-S_n} + \sum_{k=0}^{n-1} e^{-S_k} h_{1,k+1}(f_{k+1,n}(0))} \;\geq\; n^{-1}\, c^{-1} e^{L_n}.$$

The main difficulty here is to obtain the equivalent of Theorem 2. For this, one has to calculate the higher order derivatives of $g_{0,n}$. Then, the upper bound for the tail probability of Z_n contains an additional term:

Theorem 3. *Under Assumption $\mathscr{H}(\beta)$ for some $\beta > 2$, there are a constant $0 < c < \infty$ and a positive nondecreasing slowly varying function Υ such that for every $k \geq 1$,*

$$P(Z_n > k|\Pi) \;\;\leq\;\; c\; e^{S_n} n^{\beta} \Upsilon(n^2 e^{M_n - L_n} k) \max\left\{k^{-\beta} e^{(\beta-1)(S_n - L_n)}; k^{-\lceil\beta\rceil-1} e^{\lceil\beta\rceil(S_n-L_n)}\right\} \quad a.s.$$

For the proof, we will use the functions (compare with (14))

$$h_k(s) \;:=\; h_{1,k}(s) \;=\; \frac{1}{(1-f_k(s))} - \frac{1}{f_k'(1)(1-s)} = \frac{g_k(1) - g_k(s)}{g_k(1)g_k(s)(1-s)} \quad (0 \leq s < 1) \quad \text{a.s.}$$

and

$$H(s) \;:=\; \sum_{k=0}^{n-1} U_k h_{k+1}(f_{k+1,n}(s)) \quad (0 \leq s < 1) \quad \text{a.s.} \tag{25}$$

Applying (15) with $\mu = 1$ yields

$$g_{0,n}(s)^{-1} \;\;=\;\; \frac{1-s}{1-f_{0,n}(s)} = U_n + (1-s)H(s) \quad (0 \leq s < 1) \quad \text{a.s.}$$

Calculating the l-th derivative of the above equation, we get that for all $l \geq 1$ and $s \in [0, 1)$,

$$\frac{d^l}{ds^l} g_{0,n}(s)^{-1} \;\;=\;\; (1-s)H^{(l)}(s) - lH^{(l-1)}(s) \quad (0 \leq s < 1) \quad \text{a.s.} \tag{26}$$

The rest of this section is organized as follows. First, we provide a technical lemma which gives useful bounds for the functions defined above. We then prove Theorem 3. Finally, we sketch the main steps of the proof of the upper bound of Theorem 1 for $\beta > 2$. For notational simplicity, we introduce $\leq_c$ to mean that the inequality is fulfilled up to a multiplicative constant c which does not depend on s, k or Π.

Lemma 2. *Under Assumption $\mathscr{H}(\beta)$, for every $l \leq \lceil\beta\rceil - 1$ and $0 \leq k \leq n$,*

$$f_{k,n}^{(l)}(1) \quad \leq_c \quad n^{l-1}\ e^{S_n - S_k} e^{(l-1)(S_n - L_n)} \quad a.s. \tag{27}$$

Moreover the following estimates hold a.s. for every $s \in [0,1)$ respectively for $l < \lceil\beta\rceil - 2$, $l = \lceil\beta\rceil - 2$ and $l = \lceil\beta\rceil - 1$,

$$|H^{(l)}(s)| \quad \leq_c \quad n^l\ e^{l(S_n - L_n)} e^{-L_n}, \tag{28}$$

$$\begin{aligned} |H^{(l)}(s)| \quad \leq_c \quad & n^l e^{(\lceil\beta\rceil - 2)(S_n - L_n)} e^{-L_n} \\ & + n\Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1})(1-s)^{-(\lceil\beta\rceil - \beta)} e^{-S_n} e^{(\beta-1)(S_n - L_n)} \end{aligned} \tag{29}$$

$$\begin{aligned} |H^{(l)}(s)| \quad \leq_c \quad & n^l e^{(\lceil\beta\rceil - 1)(S_n - L_n)} e^{-L_n} + n\Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1})\ e^{-S_n} e^{\beta(S_n - L_n)}(1-s)^{-(\lceil\beta\rceil - \beta)} \\ & + n\Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1})\ e^{-S_n} e^{(\beta-1)(S_n - L_n)}(1-s)^{-1-(\lceil\beta\rceil - \beta)} \quad a.s. \end{aligned} \tag{30}$$

Proof. We prove the lemma by induction with respect to l. All the following relations hold a.s. for every $s \in [0,1)$. For $l = 1$, (27) is trivially satisfied since $f'_{0,n}(1) = e^{S_n}$ and $f'_{k,n}(1) = e^{S_n - S_k}$. We first consider $l < \lceil\beta\rceil - 2$ and assume that then (27) holds for every $i \leq l$. In the first step, it will be proved that (28) holds for l. In the second step, we will show that (27) holds for $l + 1$.

By the induction assumptions and the monotonicity of the generating functions and their derivatives, for all $i \leq l$ and $s \in [0,1]$, we have

$$f_{k+1,n}^{(i)}(s) \ \leq\ f_{k+1,n}^{(i)}(1) \quad \leq_c \quad n^{i-1}\ e^{S_n - S_k}\ e^{(i-1)(S_n - L_n)}. \tag{31}$$

We will use this to bound the l-th derivative of $h_{k+1} \circ f_{k+1,n}$. Lemma 7 (see appendix) ensures that

$$\left|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\right| \quad = \quad \left|\sum_{j=1}^{l} h_{k+1}^{(j)}(f_{k+1,n}(s)) u_{j,l}(s)\right|,$$

where

$$u_{j,l}(s) \quad = \quad \sum_{(i_1,\ldots,i_{2j}) \in \mathscr{C}(j,l)} c_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j-1})})(s))^{i_{2j}}$$

and $\mathscr{C}(j,l) = \{(i_1,\ldots,i_{2j}) \in \mathbb{N}^{2j} \big| i_1 i_2 + i_3 i_4 + \ldots = l \text{ and } i_2 + i_4 + \ldots = j\}$. Using (53) and (31), these functions statisfy

$$u_{j,l}(s) \quad \leq_c \quad n^{l-j}\ e^{j(S_n - S_k)}\ e^{(l-j)(S_n - L_n)} \ \leq_c\ n^{l-1}\ e^{S_n - S_k}\ e^{(l-1)(S_n - L_n)}.$$

For $j < \lceil\beta\rceil - 2$, the derivatives $h_k^{(j)}$ are bounded by a constant that does not depend on Π (see Lemma 6, appendix). Thus

$$\left|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\right| \quad \leq_c \quad n^{l-1}\ e^{S_n - S_k}\ e^{(l-1)(S_n - L_n)}.$$

By definition of H (see (25)), we get that

$$|H^{(l)}(s)| \quad \leq_c \quad \sum_{k=0}^{n-1} n^{l-1}\; n^{l-1}\; e^{S_n - S_k}\; e^{(l-1)(S_n-L_n)}\; e^{-S_k} \;\leq_c\; n^l e^{l(S_n-L_n)} e^{-L_n},$$

which proves (28) for $l < \lceil \beta \rceil - 2$.

We now show that (27) is fulfilled for $l+1 < \lceil \beta \rceil - 1$. Using Lemma 7 again with $f = g_{0,n}$ and $h(x) = 1/x$, we get that

$$\begin{aligned} \frac{d^l}{ds^l} g_{0,n}(s)^{-1} &= \sum_{j=1}^{l} (-1)(-2)\cdots(-j) g_{0,n}(s)^{-(j+1)} \tilde{u}_{j,l}(s) \\ &= -g_{0,n}(s)^{-2} g_{0,n}^{(l)}(s) + \sum_{j=2}^{l} (-1)(-2)\cdots(-j) g_{0,n}(s)^{-(j+1)} \tilde{u}_{j,l}(s), \end{aligned} \tag{32}$$

where

$$\tilde{u}_{j,l}(s) \quad = \sum_{i=(i_1,\ldots,i_{2j})\in \mathscr{C}(j,l)} c_i (g_{0,n}^{(i_1)}(s))^{i_2} \cdots (g_{0,n}^{(i_{2j-1})})(s))^{i_{2j}}, \tag{33}$$

and $\mathscr{C}(j,l) = \left\{(i_1,\ldots,i_{2j}) \in \mathbb{N}^{2j} \middle| i_1 i_2 + i_3 i_4 + \ldots = l \text{ and } i_2 + i_4 + \ldots = j\right\}$.
Moreover, $f^{(l)}(1) = l g^{(l-1)}(1)$ (see (39)). Thus, using the induction assumption (27) yields for every $i \leq l-1$, $g_{0,n}^{(i)}(1) \leq_c n^i e^{S_n} e^{i(S_n-L_n)}$ and

$$\tilde{u}_{j,l}(1) \leq_c n^l\; e^{jS_n} e^{l(S_n-L_n)}.$$

By (25), the left-hand side of (32) is equal to $(1-s)H^{(l)}(s) - lH^{(l-1)}(s)$. By (28), for $l < \lceil \beta \rceil - 2$, $(1-s)H^{(l)}(s)$ vanishes for $s = 1$. Thus letting $s \to 1$ and recalling that $g_{0,n}(1) = e^{S_n}$ yields

$$\begin{aligned} g_{0,n}^{(l)}(1) &\leq_c e^{2S_n} \Big(\sum_{j=2}^{l} (-1)(-2)\cdots(-j) e^{-(j+1)S_n}\; n^l\; e^{jS_n} e^{l(S_n-L_n)} \;+ l|H^{(l-1)}(1)| \Big) \\ &\leq_c e^{S_n}\; n^l\; e^{l(S_n-L_n)} + e^{2S_n} |H^{(l-1)}(1)|. \end{aligned}$$

As (28) is already proved for $l < \lceil \beta \rceil - 2$, we get that

$$\begin{aligned} g_{0,n}^{(l)}(1) &\leq_c n^l\; e^{S_n} e^{l(S_n-L_n)} + n^{l-1}\; e^{2S_n} e^{(l-1)(S_n-L_n)} e^{-L_n} \\ &\leq_c n^l\; e^{S_n} e^{l(S_n-L_n)}. \end{aligned}$$

Using (39), we get (27) for $l+1$, which completes the induction step and proves (27) for $l < \lceil \beta \rceil - 1$.
The proof for $f_{k,n}^{(l)}(1)$ instead of $f_{0,n}^{(l)}(1)$ is the same. Here, just note that for all $0 \leq k \leq n$, it holds that $S_n - S_k - \min_{j \geq k}\{S_j - S_k\} \leq S_n - L_n$.

Let us now prove the bound on $H^{(l)}(s)$ for $l = \lceil\beta\rceil - 2$. Using Lemma 6 and (27) yields

$$\begin{aligned}
&\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \\
&= \Big|\sum_{j=1}^{l-1} h_{k+1}^{(j)}(f_{k+1,n}(s))u_{j,l}(s) + h_{k+1}^{(l)}(f_{k+1,n}(s))(f'_{k+1,n}(s))^l\Big| \\
&\le_c n^{l-1} e^{S_n - S_k} e^{(\lceil\beta\rceil-3)(S_n-L_n)} \\
&\qquad + \Upsilon(1/(1-f_{k+1,n}(s)))(1-f_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{\lceil\beta\rceil-2}. \qquad (34)
\end{aligned}$$

Now by the same arguments as in the proof of Theorem 2,

$$\Upsilon(1/(1-f_{k+1,n}(s))) \le \Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1}).$$

The convexity of $f_{k+1,n}$ ensures that

$$(1-f_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)} \ \le\ (1-s)^{-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)}.$$

Recalling that $\beta > 2$ and $f'_{k+1,n}(s) \le e^{S_n - L_n}$, and applying (34) yields

$$\begin{aligned}
\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \ \le_c\ & n^{l-1} e^{S_n-L_n} e^{(\lceil\beta\rceil-3)(S_n-L_n)} \\
&+\Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})(1-s)^{-(\lceil\beta\rceil-\beta)} e^{(\beta-2)(S_n-L_n)}.
\end{aligned}$$

Combining this inequality with the estimate (recall (25)),

$$|H^{(l)}(s)| \le \sum_{k=0}^{n-1} e^{-S_k}\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \le n e^{-L_n}\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big|, \qquad (35)$$

proves (29).

This implies that $(1-s)H^{(l)}(s) \to 0$ as $s \to 1$ for $l = \lceil\beta\rceil - 2$. Thus we can apply the same arguments to get an upper bound for $g_{0,n}^{(l)}(1)$ to prove (27) for $l = \lceil\beta\rceil - 1$.

Finally, let $l = \lceil\beta\rceil - 1$. Applying the same arguments as before, Lemmas 6 yields

$$\begin{aligned}
&\Big|\frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s))\Big| \\
&= \Big|\sum_{j=1}^{l-2} h_{k+1}^{(j)}(f_{k+1,n}(s))u_{j,l}(s) + l h_{k+1}^{(l-1)}(f_{k+1,n}(s)) f_{k+1,n}^{(2)}(s)(f'_{k+1,n}(s))^{l-2} \\
&\qquad\qquad + h_{k+1}^{(l)}(f_{k+1,n}(s))(f'_{k+1,n}(s))^l\Big| \\
&\le_c n^{l-1} e^{(\lceil\beta\rceil-1)(S_n-L_n)} \\
&\quad + \Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})(1-s)^{-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{-(\lceil\beta\rceil-\beta)} f_{k+1,n}^{(2)}(s)(f'_{k+1,n}(s))^{\lceil\beta\rceil-3} \\
&\quad + \Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})(1-s)^{-1-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{-1-(\lceil\beta\rceil-\beta)}(f'_{k+1,n}(s))^{\lceil\beta\rceil-1} \\
&\le_c n^{l-1} e^{S_n} e^{(\lceil\beta\rceil-2)(S_n-L_n)} + \Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1}) \\
&\qquad\qquad \cdot \Big(e^{(\beta-1)(S_n-L_n)}(1-s)^{-(\lceil\beta\rceil-\beta)} + e^{(\beta-2)(S_n-L_n)}(1-s)^{-1-(\lceil\beta\rceil-\beta)}\Big).
\end{aligned}$$

Using (35) again, we get (30) and the proof is complete. □

Proof of Theorem 3 for $\beta > 2$. Let $l = \lceil \beta \rceil - 1$. Without loss of generality, we can assume that $\Upsilon \geq 1$. The following relations hold a.s. Combining (26) and (32), we get that

$$g_{0,n}^{(l)}(s) \;=\; g_{0,n}(s)^2\Big(-(1-s)H^{(l)}(s) + lH^{(l-1)}(s) + \sum_{j=2}^{l}(-1)(-2)\cdots(-j)g_{0,n}(s)^{-(j+1)}\tilde{u}_{j,l}(s)\Big),$$

where $\tilde{u}_{j,l}$ is defined in (33). Using (29), (30), and $g_{0,n}(s) \leq e^{S_n}$ for the first terms yields

$$\begin{aligned} g_{0,n}^{(l)}(s) \leq_c\; & e^{2S_n} n^l \Upsilon(n^2 e^{M_n - L_n}(1-s)^{-1})\Big((1-s)e^{(\lceil\beta\rceil-1)(S_n-L_n)}e^{-L_n} + e^{-S_n}e^{\beta(S_n-L_n)}(1-s)^{1-(\lceil\beta\rceil-\beta)} \\ & + e^{-S_n}e^{(\beta-1)(S_n-L_n)}(1-s)^{-(\lceil\beta\rceil-\beta)} + e^{(\lceil\beta\rceil-2)(S_n-L_n)}e^{-L_n} + e^{-S_n}e^{(\beta-1)(S_n-L_n)}(1-s)^{-(\lceil\beta\rceil-\beta)}\Big) \\ & + g_{0,n}^{-(j-1)}(s)\tilde{u}_{j,l}(s). \end{aligned}$$

Using that for every $i \in \mathbb{N}$, $g^{(i)}(s)/(g(s))^i \leq g^{(i)}(1)/(g(1))^i$ (see (41), appendix), the definition of $\tilde{u}_{j,l}$, (27) and (39), we get that

$$g_{0,n}^{-(j-1)}(s)\tilde{u}_{j,l}(s) \leq_c n^{l-1}e^{S_n}e^{(l-1)(S_n-L_n)}.$$

Thus we get that

$$\begin{aligned} g_{0,n}^{(l)}(s) \leq_c\; & e^{S_n} n^l \Upsilon(n^2 e^{M_n-L_n}(1-s)^{-1})\Big((1-s)^{-(\lceil\beta\rceil-\beta)}e^{(\beta-1)(S_n-L_n)} + (1-s)^{1-(\lceil\beta\rceil-\beta)}e^{\beta(S_n-L_n)} \\ & + (1-s)e^{\lceil\beta\rceil(S_n-L_n)} + e^{(\lceil\beta\rceil-1)(S_n-L_n)}\Big) + e^{S_n}n^l e^{(\lceil\beta\rceil-1)(S_n-L_n)}. \end{aligned}$$

As in (16), we get the following estimate for every $1/2 \leq s < 1$,

$$g_{0,n}^{(l)}(s) \;\geq_c\; s^k k^{l+1}\mathbb{P}(Z_n > k|\Pi).$$

Choosing $s = 1 - 1/k$ yields

$$\begin{aligned} \mathbb{P}(Z_n > k|\Pi) \;\leq_c\; & e^{S_n} n^l \Upsilon(n^2 e^{M_n-L_n}k)\Big(k^{-\beta}e^{(\beta-1)(S_n-L_n)} + k^{-(\beta+1)}e^{\beta(S_n-L_n)} \\ & + k^{-\lceil\beta\rceil-1}e^{\lceil\beta\rceil(S_n-L_n)} + k^{-\lceil\beta\rceil}e^{(\lceil\beta\rceil-1)(S_n-L_n)}\Big). \end{aligned}$$

Using the monotonicity of the function $x \to a^{-x}\exp((x-1)b)$ for all $a \geq 1$ and $b \geq 0$, and that $\beta \leq \lceil\beta\rceil < \beta + 1 \leq \lceil\beta\rceil + 1$, we get that for all $k \geq 1$,

$$\max\{k^{-\beta-1}e^{\beta(S_n-L_n)}; k^{-\lceil\beta\rceil}e^{(\lceil\beta\rceil-1)(S_n-L_n)}\} \;\leq\; \max\{k^{-\beta}e^{(\beta-1)(S_n-L_n)}; k^{-\lceil\beta\rceil-1}e^{\lceil\beta\rceil(S_n-L_n)}\}.$$

Combining the two last inequalities leads to

$$P(Z_n > k|\Pi) \;\leq_c\; e^{S_n}n^l\Upsilon(n^2e^{M_n-L_n}k)\max\{k^{-\beta}e^{(\beta-1)(S_n-L_n)}; k^{-\lceil\beta\rceil-1}e^{\lceil\beta\rceil(S_n-L_n)}\},$$

and this completes the proof. □

Proof of the upper bound of Theorem 1 *for* $\beta > 2$. The proof is in the same spirit as the proof for $\beta \in (1,2]$. Theorem 3 yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad -\min\{\chi_{\gamma,\beta,\psi}(\theta), \chi_{\gamma,\lceil\beta\rceil+1,\psi}(\theta)\},$$

where χ is defined in (4). Using the characterization of χ (see Lemma 3, appendix), we deduce that for any $\theta \geq 0$,

$$\chi_{\gamma,\beta,\psi}(\theta) \leq \chi_{\gamma,\lceil\beta\rceil+1,\psi}(\theta).$$

Thus

$$\min\{\chi_{\gamma,\beta,\psi}(\theta), \chi_{\gamma,\lceil\beta\rceil+1,\psi}(\theta)\} = \chi_{\gamma,\beta,\psi}(\theta) = \chi(\theta),$$

which yields the upper bound. □

6 Proof of Corollary 1

By assumption, there exists a constant $d < \infty$ such that for every $\beta > 0$,

$$\mathbb{P}(L > z | f, L > 0) \quad \leq \quad d \cdot (m \wedge 1) \cdot z^{-\beta} \quad \text{a.s.}$$

Thus we may apply the upper bound in Theorem 1 for every $\beta > 0$. Thus, for all $\beta > 0$ and $\theta \geq 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad -\chi_{\gamma,\beta,\psi}(\theta).$$

Taking the limit $\beta \to \infty$, monotone convergence of $\chi_{\gamma,\beta,\psi}$ yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \quad \leq \quad -\chi_{\gamma,\infty,\psi}(\theta),$$

where

$$\begin{aligned} \chi_{\gamma,\infty,\psi}(\theta) \quad &:= \quad \lim_{\beta\to\infty} \inf_{t\in[0,1],s\in[0,\theta]} \Big\{ t\gamma + \beta s + (1-t)\psi((\theta - s)/(1-t)) \Big\} \\ &= \quad \inf_{t\in[0,1]} \Big\{ t\gamma + (1-t)\psi(\theta/(1-t)) \Big\}. \end{aligned}$$

This gives the upper bound. For the proof of the lower bound, we can apply the same arguments as in Section 3.

7 Appendix

In this section, we present several technical results for real functions which are required for the proofs.

7.1 Characterization of the rate function χ

Lemma 3. *Let ψ be a nonnegative convex function, $0 \le \gamma \le \psi(0)$ and $\beta > 0$. Then the function χ defined for $\theta \ge 0$ by*

$$\chi(\theta) \;=\; \inf_{t\in[0,1], s\in[0,\theta]} \left\{ t\gamma + \beta s + (1-t)\psi((\theta - s)/(1-t)) \right\}$$

is the largest convex function such that for all $x, \theta \ge 0$,

$$\chi(0) = \gamma, \quad \chi(\theta) \le \psi(\theta), \quad \chi(\theta + x) \le \chi(\theta) + \beta x. \tag{36}$$

Proof. We first prove that that χ is convex. Using the definition of χ and the convexity of ψ, for any $\theta', \theta'' \ge 0$ and $\epsilon > 0$ there exist $t', t'' \in [0,1)$, $s' \in [0, \theta']$ and $s'' \in [0, \theta'']$ such that for every $\lambda \in [0,1]$,

$$\begin{aligned}
&\lambda\chi(\theta') + (1-\lambda)\chi(\theta'') \\
&\quad \ge \lambda[t'\gamma + \beta s' + (1-t')\psi((\theta' - s')/(1-t'))] \\
&\qquad\qquad + (1-\lambda)[t''\gamma + \beta s'' + (1-t'')\psi((\theta'' - s'')/(1-t''))] - \epsilon \\
&\quad \ge [\lambda t' + (1-\lambda)t'']\gamma + [\lambda s' + (1-\lambda)s'']\beta \\
&\qquad + (\lambda(1-t') + (1-\lambda)(1-t''))\tfrac{\lambda(1-t')}{\lambda(1-t')+(1-\lambda)(1-t'')}\psi((\theta' - s')/(1-t')) \\
&\qquad + (\lambda(1-t') + (1-\lambda)(1-t''))\tfrac{(1-\lambda)(1-t'')}{\lambda(1-t')+(1-\lambda)(1-t'')}\psi((\theta'' - s'')/(1-t'')) - \epsilon \\
&\quad \ge [\lambda t' + (1-\lambda)t'']\gamma + [\lambda s' + (1-\lambda)s'']\beta \\
&\qquad\qquad + \Big(1 - [\lambda t' + (1-\lambda)t'']\Big)\psi\Big(\frac{\lambda\theta' + (1-\lambda)\theta'' - (\lambda s' + (1-\lambda)s'')}{1 - [\lambda t' + (1-\lambda)t'']}\Big) - \epsilon \\
&\quad \ge \chi\Big(\lambda\theta' + (1-\lambda)\theta''\Big) - \epsilon.
\end{aligned}$$

Letting $\epsilon \to 0$ entails that χ is convex.

Following the previous computation, we can verify that χ fulfills (36). For any $\theta \ge 0$ and $\epsilon > 0$, there exist $t' \in [0,1)$ and $s' \in [0,\theta]$ such that

$$\begin{aligned}
\chi(\theta) &\ge t'\gamma + \beta s' + (1-t')\psi\big((\theta - s')/(1-t')\big) - \epsilon \\
&= t'\gamma + \beta(s' + x) + (1-t')\psi\big((\theta + x - (s' + x))/(1-t')\big) - \beta x - \epsilon \\
&\ge \inf_{t\in[0,1], \tilde{s}\in[0,\theta+x]} \left\{ t\gamma + \beta\tilde{s} + (1-t)\psi((\theta + x - \tilde{s})/(1-t)) \right\} - \beta x - \epsilon.
\end{aligned}$$

Taking the limit $\epsilon \to 0$ yields the third property in (36). Furthermore, setting $t = 0$, $s = 0$ implies $\chi(\theta) \le \psi(\theta)$ and letting $t \to 1$ shows that $\chi(0) \le \gamma$. This completes the proof of (36).

Finally, let κ be any convex function which satisfies (36). Then for all $t \in [0,1)$ and $0 \le s \le \theta$,

$$\begin{aligned}
t\gamma + \beta s + (1-t)\psi((\theta - s)/(1-t)) &\ge t\kappa(0) + \beta s + (1-t)\kappa((\theta - s)/(1-t)) \\
&\ge \beta s + \kappa\big(t0 + (1-t)(\theta - s)/(1-t)\big) \\
&= \beta s + \kappa(\theta - s) \\
&\ge \kappa(\theta).
\end{aligned}$$

Taking the infimum over s and t yields $\chi(\theta) \geq \kappa(\theta)$ and this completes the proof. □

Another characterization of χ results from Lemma 3 (see Figure 2):

Lemma 4. *Let χ, θ^* and $\theta^\dagger$ be defined as in (4), (6) and (7) and assume $0 < \theta^* < \theta^\dagger < \infty$. Then*

$$\chi(\theta) \;=\; \begin{cases} \gamma(1-\frac{\theta}{\theta^*}) + \frac{\theta}{\theta^*}\psi(\theta^*) & , \;\textit{if } \theta \leq \theta^* \\ \psi(\theta) & , \;\textit{if } \theta^* < \theta < \theta^\dagger \\ \beta(\theta-\theta^\dagger) + \psi(\theta^\dagger) & , \;\textit{if } \theta \geq \theta^\dagger \end{cases} .$$

Proof. Note that the convex and monotone function ψ has at most one jump (to infinity). Let this jump be in $0 < \theta_j \leq \infty$. For $\theta < \theta_j$, $\psi(\theta)$ is differentiable. As ψ is also continuous from below, $\psi(\theta_j) < \infty$. By the preceding lemma, χ is the largest convex function starting at $\chi(0) = \gamma$, being at most as large as ψ and having at most slope β.
The largest convex function through the point $(0, \gamma)$ being smaller/equal than ψ has to be linear and has to be a tangent to ψ. By the definition of θ^*, the tangent to ψ at θ^* touches the point $(0, \gamma)$. Thus χ is linear for $\theta < \theta^*$ and follows this tangent. For $\theta > \theta^*$, χ is identical to ψ until the slope of ψ is exactly β (or until ψ jumps to infinity). At this point $\theta^\dagger$, the last condition in (36) becomes important and χ is linear with slope β for $\theta > \theta^\dagger$. Summing up,

$$\chi(\theta) \;=\; \begin{cases} \gamma(1-\frac{\theta}{\theta^*}) + \frac{\theta}{\theta^*}\psi(\theta^*) & , \;\text{if } \theta \leq \theta^* \\ \psi(\theta) & , \;\text{if } \theta^* < \theta < \theta^\dagger \\ \beta(\theta-\theta^\dagger) + \psi(\theta^\dagger) & , \;\text{if } \theta \geq \theta^\dagger \end{cases} .$$

□

If $\gamma = \psi(0)$, then $\theta^* = 0$. If $\psi'(0) > \beta$, then $\theta^\dagger = 0$ and $\chi(\theta) = \gamma + \beta\theta$. We refrain from describing other degenerated cases.

7.2 Slowly varying functions

In this section, we will recall some properties of slowly varying functions. We refer to [7] for details. The function $\Upsilon : (0,\infty) \to (0,\infty)$ is a called *slowly varying* if for every $a > 0$,

$$\lim_{x\to\infty} \frac{\Upsilon(ax)}{\Upsilon(x)} \;=\; 1.$$

In the following, a Tauberian result from [13], p. 423 (see also [7], Theorem 1.5.11, p. 28) is used: For any $\alpha > -1$, the function $g(s) := \sum_{k=0}^{\infty} s^k k^\alpha$ satisfies

$$g(s) \;\sim\; \Gamma(\alpha+1)(1-s)^{-1-\alpha} \quad (s \to 1-).$$

Then the function $\xi = s \to (1-s)^{1+\alpha} g(s)$ is continuous on $[0,1)$ and has a finite left limit at 1. Denoting the supremum of this function extended to $[0,1]$ by M, we get that

$$\sum_{k=1}^{\infty} s^k k^\alpha \;\leq\; M(1-s)^{-1-\alpha} \qquad (0 \leq s < 1).$$

For $\alpha=-1$, $\sum_{k=1}^{\infty} s^k/k = -\log(1-s)$. As the logarithm is a slowly varying function, we may rewrite the previous results in the following way, which will be convenient in the proofs:
There exists a nondecreasing positive slowing varying function Υ such that for all $\alpha \geq -1$ and $s \in [0,1)$

$$\sum_{k=1}^{\infty} s^k k^{\alpha} \quad \leq \quad \Upsilon(1/(1-s))(1-s)^{-1-\alpha}. \tag{37}$$

7.3 Bounds for generating functions

We use the convention $0 \cdot \infty = 0$. Let L be a random variable with values in $\{0,1,2,...\}$ with expectation m, distribution $(p_k)_{k\in\mathbb{N}}$ and generating function f. Let us define

$$q_k \ := \ \mathbb{P}(L > k|f)$$

and,

$$g(s) \ := \ \sum_{k=0}^{\infty} s^k q_k \ = \ \frac{1-f(s)}{1-s}. \tag{38}$$

The last identity results from the Cauchy product formula of power series (see also [9]). Recall that the $l-$th derivative of a function f is denoted by $f^{(l)}$ and that $f^{(l)}(s)$, $g^{(l)}(s)$ exist for every $s \in [0,1)$. Since

$$f^{(l)}(s) \ = \ \sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1)s^{k-l}p_k, \qquad g^{(l)}(s) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1)s^{k-l}q_k,$$

all derivatives of f and g are nonnegative, nondecreasing functions. For the proofs, we will use g instead of f since the associated sequence $(q_k)_{k\in\mathbb{N}}$ is monotone, which is more convenient. It is straightforward to see by induction that

$$f^{(l)}(1) \quad = \quad l g^{(l-1)}(1). \tag{39}$$

Calculating the l-th derivative of $f(s) = 1-(1-s)g(s)$ yields

$$f^{(l)}(s) \quad = \quad l g^{(l-1)}(s) - (1-s)g^{(l)}(s). \tag{40}$$

Thus $g^{(l-1)}(1)$ and $f^{(l)}(1)$ both essentially describe the l-th moment of the corresponding probability distribution. Next, we prove that for every $i \in \mathbb{N}$,

$$g^{(i)}(s)\cdot(g(1))^i \leq (g(s))^i \cdot g^{(i)}(1). \tag{41}$$

We will prove the result by induction. For $i=1$, define a random variable Y with distribution $(q_k/g(1))_{k\in\mathbb{N}_0}$. Then

$$g'(s)\cdot g(1) = \mathbb{E}[s^Y Y]g(1)^2 \leq \mathbb{E}[s^Y]\mathbb{E}[Y]g(1)^2 = g(s)g'(1). \tag{42}$$

as s^Y and Y are obviously negatively correlated for $s \in [0,1]$. Note that the above inequality remains true if g' is replaced by $g^{(i)}$, $i \in \mathbb{N}$. Next, let (41) be fulfilled for i. Thus, using (40) and the induction assumption and monotonicity of $g^{(i)}$ yield

$$\frac{g^{(i+1)}(s)}{(g(s))^{i+1}} = \frac{(i+1)g^{(i)}(s) - f^{(i+1)}(s)}{(1-s)(g(s))^{i+1}} \le \frac{(i+1)g^{(i)}(1)}{(1-s)(g(1))^i g(s)} - \frac{f^{(i+1)}(s)}{(1-s)(g(1))^i g(s)}$$
$$= \frac{f^{(i+1)}(1) - f^{(i+1)}(s)}{(1-s)(g(1))^i g(s)} = \frac{1}{(g(1))^i g(s)} \sum_{k=0}^{\infty} k(k-1)\cdots(k-i)\frac{1-s^{k-i-1}}{1-s} p_k$$
$$= \frac{g^{(i+1)}(s)}{(g(1))^i g(s)} \le \frac{g^{(i+1)}(1)}{(g(1))^{i+1}}.$$

In the last step, we used (42) with g' replaced by $g^{(i+1)}$.

For $\mu \in (0,1]$, let us define the function

$$h_\mu(s) \;:=\; \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu}. \tag{43}$$

The following useful lemmas give the versions of Assumption $\mathcal{H}(\beta)$ in terms of the function h_μ. Noting that $g(0) = q_0 = \mathbb{P}(L > 0|f)$ and $g(1) = m$, we may rewrite Assumption $\mathcal{H}(\beta)$ in the following way

$$q_k \;\le\; d\; g(0)\,(g(1) \wedge 1)\, k^{-\beta} \qquad (k \ge 1). \tag{44}$$

Lemma 5. *Let $\beta > 1$ and assume that (44) holds for some constant $0 < d < \infty$. Then for every $0 < \mu < (\beta - 1) \wedge 1$, there exists a constant $c = c(\beta, d, \mu)$ such that for every $s \in [0,1]$,*

$$h_\mu(s) \;\;\le\;\; c. \tag{45}$$

The above bound also holds for $\mu = 1$ if $\beta > 2$. Moreover, if $\beta \in (1,2]$, there exists a nondecreasing positive slowly varying function $\Upsilon = \Upsilon(\beta, d)$ such that for every $s \in [0,1)$,

$$h_{\beta-1}(s) \;\;\le\;\; \Upsilon(1/(1-s)) \tag{46}$$
$$-h'_{\beta-1}(s) \;\;\le\;\; \Upsilon(1/(1-s))/(1-s). \tag{47}$$

Note that Υ depends on L (or g) only through the values of d and β. Thus under Assumption $\mathcal{H}(\beta)$, we can derive a nonrandom constant bound from this lemma.

In the proofs, we will use the notation $\le_c$ again, which means that the inequality is satisfied up to a multiplicative constant which depends on β and μ but is independent of s.

Proof. Let $s \in [0,1)$. Using $g(s) \ge g(0)$, we get that

$$\begin{aligned} h_\mu(s) &= \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu} \\ &\le \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(0)(1-s))^\mu} \\ &\le (g(1)\wedge 1)^{-1}\frac{(\sum_{k=0}^\infty g(0)^{-1}q_k)^\mu - (\sum_{k=0}^\infty s^k q_k g(0)^{-1})^\mu}{(1-s)^\mu}. \end{aligned} \tag{48}$$

Since $\mu \in (0,1]$, the function $x \to x^{\mu}$ is concave, such that $a^{\mu} - x^{\mu} \leq \mu x^{\mu-1}(a-x)$ for all $0 \leq x \leq a$. Moreover

$$1 = q_0/g(0) \leq x := \sum_{k=0}^{\infty} s^k q_k g(0)^{-1} \leq a := \sum_{k=0}^{\infty} q_k g(0)^{-1}. \tag{49}$$

Then $x^{\mu-1} \leq 1$ and using the inequality of concavity in (48) with $q_k \leq d g(0) \cdot (g(1) \wedge 1) \cdot k^{-\beta}$ leads to

$$\begin{aligned}
h_{\mu}(s) \quad &\leq \quad \mu(g(1) \wedge 1)^{-1} x^{\mu-1} \frac{\sum_{k=0}^{\infty} g(0)^{-1} q_k [1-s^k]}{(1-s)^{\mu}} \\
&\leq_c \quad \frac{\sum_{k=1}^{\infty} (1-s^k) k^{-\beta}}{(1-s)^{\mu}} \\
&= \quad (1-s)^{1-\mu} \sum_{k=1}^{\infty} \frac{1-s^k}{1-s} k^{-\beta} \\
&= \quad (1-s)^{1-\mu} \sum_{k=1}^{\infty} k^{-\beta} \sum_{j=0}^{k-1} s^j \\
&= \quad (1-s)^{1-\mu} \sum_{j=0}^{\infty} s^j \sum_{k=j+1}^{\infty} k^{-\beta} \\
&\leq_c \quad (1-s)^{1-\mu} \sum_{j=0}^{\infty} s^j (j+1)^{-\beta+1}.
\end{aligned}$$

The estimates (45) and (46) on h_{μ} for $0 < \mu < (\beta-1) \wedge 1$ and $\mu = \beta - 1$ now follow directly from (37). For $\mu = 1$, $\beta > 2$ and $s = 1$, the sum is finite and (45) also holds in this case.

For the second part of the lemma, we explicitly compute the first derivative of $h_{\beta-1}$, using the formula

$$h_{\beta-1}(s) g(s)^{\beta-1} \quad = \quad \frac{g(1)^{\beta-1} - g(s)^{\beta-1}}{g(1)^{\beta-1}(1-s)^{\beta-1}}.$$

By the differentiation of both sides, we get

$$h'_{\beta-1}(s) g(s)^{\beta-1} + (\beta-1) h_{\beta-1}(s) g(s)^{\beta-2} g'(s) \;=\; \frac{(\beta-1)([g(1)^{\beta-1} - g(s)^{\beta-1}] - (1-s) g(s)^{\beta-2} g'(s))}{g(1)^{\beta-1}(1-s)^{\beta}}$$

and thus

$$-h'_{\beta-1}(s) \;=\; (\beta-1)\Big(\frac{h_{\beta-1}(s) g'(s)}{g(s)} + \frac{g'(s)}{g(s) g(1)^{\beta-1}(1-s)^{\beta-1}} - \frac{g(1)^{\beta-1} - g(s)^{\beta-1}}{g(s)^{\beta-1} g(1)^{\beta-1}(1-s)^{\beta}}\Big)$$

As g is nondecreasing, we can skip the last term which is negative. Using (44) and (46), we get that

$$-h'_{\beta-1}(s) \;\leq_c\; \frac{g(0) \cdot (g(1) \wedge 1) \cdot \sum_{k=1}^{\infty} k s^{k-1} k^{-\beta}}{g(s)} \Big(\Upsilon(1/(1-s)) + \frac{1}{g(1)^{\beta-1}(1-s)^{\beta-1}}\Big).$$

Moreover $g(s) \geq g(0)$ and $g(1)^{-(\beta-1)} \cdot (g(1) \wedge 1) \leq 1$ for $\beta - 1 \in (0,1]$. Hence,

$$-h'_{\beta-1}(s) \leq_c \sum_{k=1}^{\infty} s^{k-1} k^{-\beta+1} \Big(\Upsilon(1/(1-s)) + \frac{1}{(1-s)^{\beta-1}} \Big).$$

The result now follows from (37) and the fact that the product of two slowly varying functions is also slowly varying. □

Let us now consider the function

$$h(s) = h_1(s) = \frac{g(1) - g(s)}{g(1)g(s)(1-s)} \qquad (0 < s \leq 1).$$

Lemma 6. *Assume that (44) holds for some $\beta > 1$. Then there exists a finite constant $c = c(\beta, d) < \infty$ such that for every $s \in [0,1)$,*

$$\begin{aligned} |h^{(l)}(s)| &\leq c && \text{if} \quad 0 \leq l < \beta - 2 \\ |h^{(\lceil\beta\rceil-2)}(s)| &\leq c\ \Upsilon(1/(1-s))\ (1-s)^{-(\lceil\beta\rceil-\beta)} && \text{if} \quad \beta \geq 2 \\ |h^{(\lceil\beta\rceil-1)}(s)| &\leq c\ \Upsilon(1/(1-s))\ (1-s)^{-1-(\lceil\beta\rceil-\beta)}. && \end{aligned} \tag{50}$$

Proof. By (43) and the Cauchy product of power series, for every $s \in [0,1)$,

$$g(s)g(1)h(s) = \frac{g(1)-g(s)}{1-s} = \sum_{k=0}^{\infty} s^k (q_{k+1} + q_{k+2} + \ldots).$$

Thus, the l-th derivative of $g(s)h(s)$ is

$$\sum_{j=0}^{l} \binom{l}{j} g^{(j)}(s) h^{(l-j)}(s) = g(1)^{-1} \sum_{k=0}^{\infty} k(k-1)\cdots(k-l+1) s^{k-l} (q_{k+1} + q_{k+2} + \ldots).$$

Moreover, (44) ensures that for all $s \in [0,1)$ and $j < \beta - 2$ (and even $j < \beta - 1$),

$$g^{(j)}(s) \leq g^{(j)}(1) \leq \sum_{k=0}^{\infty} k^j q_k \leq_c g(0)(g(1) \wedge 1)$$

Combining the last two expressions and using $g(s)^{-1} \leq g(0)^{-1}$ yields

$$\begin{aligned} |h^{(l)}(s)| &\leq_c g(s)^{-1} \Big(g(1)^{-1} g(0) \cdot (g(1) \wedge 1) \cdot \sum_{k=0}^{\infty} k^l s^{k-l} \sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^{l} \binom{l}{j} g^{(j)}(1) |h^{(l-j)}(s)| \Big) \\ &\leq_c \sum_{k=0}^{\infty} k^l s^{k-l} \sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^{l} |h^{(l-j)}(s)| \Big) \end{aligned} \tag{51}$$

The first statement of the lemma is proved by induction on l. For $l = 0$, the result comes from Lemma 5. Assuming that the bound holds for $l' < l < \beta - 2$, the previous inequality ensures that

$$|h^{(l)}(s)| \leq_c 1 + \sum_{j=0}^{l-1} |h^{(j)}(s)|$$

since $\sum_{k=0}^{\infty} k^l \sum_{j=k+1}^{\infty} j^{-\beta} < \infty$. This finishes the induction step and proves the first estimate in (50).

Next, we consider $l = \lceil\beta\rceil - 2$. Using the bound on $h^{(l)}$ for $l < \beta - 2$ and (51) yields

$$|h^{(l)}(s)| \leq_c \sum_{k=0}^{\infty} k^l s^k \sum_{j=k+1}^{\infty} j^{-\beta} + 1 \leq_c \sum_{k=1}^{\infty} s^k k^{\lceil\beta\rceil-2} k^{-\beta+1} + 1 \leq_c \sum_{k=1}^{\infty} s^k k^{-(1-(\lceil\beta\rceil-\beta))}$$

Then the second estimate of the lemma follows from (37).

Finally, the bound for $l = \lceil\beta\rceil - 1$ is proved in the same way. By (51),

$$\begin{aligned} |h^{(l)}(s)| &\leq_c \sum_{k=0}^{\infty} k^l s^k \sum_{j=k+1}^{\infty} j^{-\beta} + l g^{(2)}(1)|h^{(l-1)}(s)| + 1 \\ &\leq_c \sum_{k=1}^{\infty} s^k k^{\lceil\beta\rceil-\beta} + \sum_{k=1}^{\infty} s^k k^{-(1-(\lceil\beta\rceil-\beta))} + 1 \\ &\leq_c \sum_{k=1}^{\infty} s^k k^{\lceil\beta\rceil-\beta} . \end{aligned}$$

Using (37) yields the claim. □

7.4 Successive differentiation for composition of functions

For the proof of the upper bound on the tail probabilities when $\beta > 2$, we have to calculate higher order derivatives of composition of functions. Here, a useful formula for the l-th derivative of a composition of two functions is proved. This can also be derived from the combinatorial form of Faà di Bruno's formula.

Lemma 7. *Let f and h be real-valued, l-times differentiable functions. Then*

$$\frac{d^l}{ds^l} h(f(s)) = \sum_{j=1}^{l} h^{(j)}(f(s)) u_{j,l}(s), \tag{52}$$

where $u_{j,l}(s)$ is given by

$$u_{j,l}(s) = \sum_{(i_1,\ldots,i_{2j}) \in \mathscr{C}(j,l)} c_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j-1})}(s))^{i_{2j}}, \tag{53}$$

for some constants $0 \leq c_i < \infty$ and $\mathscr{C}(j,l)$ defined by

$$\mathscr{C}(j,l) := \left\{(i_1, \ldots, i_{2j}) \in \mathbb{N}^{2j} \middle| i_1 i_2 + i_3 i_4 + \ldots = l \text{ and } i_2 + i_4 + \ldots = j\right\}.$$

Proof. The formula is proved by induction with respect to l. For $l = 1$, (52) is satisfied by the chain rule for differentiation. Assume that (52) and (53) hold for l. Then by the product rule for differentiation,

$$\frac{d^{l+1}}{ds^{l+1}} h(f(s)) \quad = \quad \sum_{j=1}^{l} \left(h^{(j)}(f(s)) \frac{d}{ds} u_{j,l}(s) + u_{j,l}(s) f'(s) h^{(j+1)}(f(s)) \right) .$$

Thus

$$\begin{aligned} u_{j,l}(s) f'(s) &= \sum_{i \in \mathscr{C}(j,l)} c_i \left(f^{(1)}(s)\right)^1 (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j-1})}(s))^{i_{2j}} \\ &= \sum_{i \in \mathscr{C}(j+1,l+1)} \tilde{c}_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j+1})}(s))^{i_{2(j+1)}} , \end{aligned}$$

with new constants defined by

$$\tilde{c}_{i_1,i_2,i_3,\ldots,i_{2(j+1)}} \quad := \quad \begin{cases} c_{i_3,\ldots,i_{2(j+1)}} & , \quad \text{if } i_1 = i_2 = 1 \\ 0 & , \quad \text{else} \end{cases} .$$

Furthermore,

$$\begin{aligned} \frac{d}{ds} u_{j,l}(s) &= \sum_{i \in \mathscr{C}(j,l)} \sum_{k=1}^{l} c_i (f^{(i_1)}(s))^{i_2} \cdots \ i_{2k} (f^{(i_{2k-1})}(s))^{i_{2k}-1} f^{(i_{2k-1}+1)}(s) \cdots (f^{(i_{2j-1})}(s))^{i_{2j}} \\ &= \sum_{i \in \mathscr{C}(j,l+1)} \hat{c}_i (f^{(i_1)}(s))^{i_2} \cdots (f^{(i_{2j+1})}(s))^{i_{2(j+1)}}, \end{aligned}$$

with some new constants $0 \le \hat{c}_i < \infty$. This completes the induction step. □

Acknowledgements: *The authors are grateful to Götz Kersting and Julien Berestycki for fruitful discussions. They are also grateful to the referees for their careful reading of the two previous versions and the improvement of the paper. The research was supported in part by the German Research Foundation (DFG), Grant DFG-RFBR 08-01-91954 and by ANR Manege. Moroever this research benefited from the support of the Chair Modelisation Mathematique et Biodiversite VEOLIA-Ecole Polytechnique-MNHN-F.X.*

References

[1] V.I. Afanasyev. Limit theorems for a conditional random walk and some applications. MSU. *Diss. Cand. Sci.* Moscow (1980).

[2] V.I. Afanasyev, J. Geiger, G. Kersting and V.A. Vatutin. Functional limit theorems for strongly subcritical branching processes in random environment. *Stochastic Process. Appl.* **115** (2005) 1658-1676. MR2165338

[3] K.B. Athreya and S. Karlin. On branching processes with random environments: I, II. *Ann. Math. Stat.* **42** (1971) 1499-1520, 1843-1858.

[4] V.I. Afanasyev, J. Geiger, G. Kersting and V.A. Vatutin. Criticality for branching processes in random environment. *Ann. Probab.* **33** (2005) 645-673. MR2123206

[5] V. Bansaye and J. Berestycki. Large deviations for branching processes in random environment. *Markov Process. Related Fields* **15** (2009) 493-524. MR2598126

[6] J. D. Biggins and N. H. Bingham. Large deviations in the supercritical branching process. *Adv. in Appl. Probab.* **25** (1993) 757-772. MR1241927

[7] N. H. Bingham and C. M. Goldie, J. L. Teugels. Regular Variation. Cambridge University Press, Cambridge (1987). MR0898871

[8] M. Birkner, J. Geiger and G. Kersting. Branching processes in random environment - a view on critical and subcritical cases. Springer-Verlag, Berlin (2005) 265-291. MR2118578

[9] C. Böinghoff and G. Kersting. Upper large deviations of branching in a random environment - Offspring distributions with geometrically bounded tails. *Stoch. Proc. Appl.* **120** (2010) 2064-2077. MR2673988

[10] A. Dembo and O. Zeitouni. Large Deviations Techniques and Applications. Jones and Barlett Publishers International, London (1993). MR1202429

[11] J. Geiger and G. Kersting. The survival probability of a critical branching process in random environment. *Theory Probab. Appl.* **45** (2000) 517-525. MR1967796

[12] J. Geiger, G. Kersting and V.A. Vatutin. Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **39** (2003) 593-620. MR1983172

[13] W. Feller. An Introduction to Probability Theory and Its Applications- Volume II. John Wiley & Sons, Inc., New York (1966). MR0210154

[14] Y. Guivarc'h, Q. Liu. Asymptotic properties of branching processes in random environment. *C.R. Acad. Sci. Paris,* **332**, Serie I (2001) 339-344. MR1821473

[15] F. den Hollander. Large Deviations. American Mathematical Society, Providence, RI (2000). MR1739680

[16] O. Kallenberg. Foundations of Modern Probability. Springer-Verlag, London (2001). MR1876169

[17] M. V. Kozlov. On the asymptotic behavior of the probability of non-extinction for critical branching processes in a random environment. *Theory Probab. Appl.* **21** (1976) 791-804. MR0428492

[18] M. V. Kozlov. On large deviations of branching processes in a random environment: geometric distribution of descendants. *Discrete Math. Appl.* **16** (2006) 155-174. MR2283329

[19] M. V. Kozlov. On Large Deviations of Strictly Subcritical Branching Processes in a Random Environment with Geometric Distribution of Progeny. *Theory Probab. Appl.* **54** (2010) 424-446. MR2766343

[20] A. Rouault. Large deviations and branching processes. *Proceedings of the 9th International Summer School on Probability Theory and Mathematical Statistics* (Sozopol, 1997). *Pliska Stud. Math. Bulgar.* **13** (2000) 15-38. MR1800359

[21] W. L. Smith and W.E. Wilkinson. On branching processes in random environments. *Ann. Math. Stat.* **40** (1969) 814-824. MR0246380