

ECOLE POLYTECHNIQUE

CENTRE DE MATHÉMATIQUES APPLIQUÉES

UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 46 00. Fax: 01 69 33 46 46

<http://www.cmap.polytechnique.fr/>

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

V. Bansaye ^a, C. Boeinghoff ^b

R.I. 682

February 2010

^aCMAP, Ecole Polytechnique, Palaiseau

^bDepartment of mathematics, Goethe-university Frankfurt/Main

Abstract

Branching Processes in a Random Environment (BPRES) $(Z_n : n \geq 0)$ are a generalization of Galton Watson processes where in each generation the reproduction law is picked randomly in an i.i.d. manner. We determine here the upper large deviation of the process when the reproduction law may have heavy tails. The behavior of BPRES is related to the associated random walk of the environment, whose increments are distributed like the logarithmic mean of the offspring distributions. We obtain an expression of the upper rate function of $(Z_n : n \geq 0)$, that is the limit of $-\log \mathbb{P}(Z_n \geq e^{\theta n})/n$ when $n \rightarrow \infty$. It depends on the rate function of the associated random walk of the environment, the logarithmic cost of survival $\gamma := -\lim_{n \rightarrow \infty} \log \mathbb{P}(Z_n > 0)/n$ and the polynomial decay β of the tail distribution of Z_1 . We give interpretations of this rate function in terms of the least costly ways for the process $(Z_n : n \geq 0)$ of attaining extraordinarily large values and describe the phase transitions. We derive then the rate function when the reproduction law does not have heavy tails, which generalizes the results of Böinghoff and Kersting (2009) and Bansaye and Berestycki (2008) for upper large deviations. Finally, we specify the upper large deviations for the Galton Watson processes with heavy tails.

AMS 2000 Subject Classification. 60J80, 60K37, 60J05, 92D25

Key words and phrases. Branching processes, random environments, large deviations, random walks, heavy tails.

1 Introduction

Branching processes in a random environment have been introduced in [4] and [23]. In each generation, an offspring distribution is chosen at random, independently from one generation to the other. We can think of a population of plants which have a one year life-cycle. Each year the weather conditions (the environment) vary, which impacts the reproductive success of the plant. Given the climate, all the plants reproduce independently according to the same mechanism. Initially, these processes have mainly been studied under the assumption of i.i.d. offspring distributions which are geometric, or more generally, linear fractional [1, 20]. Then, the case of general offspring distributions has attracted attention [3, 6, 10, 13].

Recently, several results about large deviations of branching processes in random environment for offspring distributions with weak tails have been proved. More precisely, [21] ensures that $\mathbb{P}(Z_n \geq \exp(\theta n))$ is equivalent to $I(\theta)\mathbb{P}(S_n \geq \theta n)$ for geometric offspring distributions and θ large enough. In [7], the authors give a general upper bound for the rate function and compute it when each individual leaves at least one offspring, i.e. $\mathbb{P}(Z_1 = 0) = 0$. Finally [11] gives an expression of the upper rate function when the reproduction laws have at most geometric tails, which excludes heavy tails.

Exceptional growth of BPRES can be due to an exceptional environment and/or to exceptional reproduction in some given environment. In this paper, we focus on large deviation probabilities when the offspring distributions may have heavy tails and the exceptional reproduction of a single individual can now contribute to the large deviation event. This leads us to consider new auxiliary power series and higher order derivatives of generating functions for the proof.

Let us give now the formal definition of the process $(Z_n : n \in \mathbb{N})$, $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, by considering a random probability generating function f and a sequence $(f_n : n \geq 1)$ of i.i.d.

copies of f which serve as random environment. Conditionally on the environment ($f_n : n \geq 1$), individuals at generation n reproduce independently of each other and their offsprings have generating function f_{n+1} . We denote by Z_n 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 $n \geq 0$,

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

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

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

Unless otherwise specified, the initial population size is 1.

We introduce the exponential rate of decay of the survival probability

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

The fact that the limit exists and $0 \leq \gamma < \infty$ is classical (see [11]) since the sequence $(-\log \mathbb{P}(Z_n > 0))_n$ is subadditive and nonnegative (see [12]). Essentially, $\gamma = 0$ in the supercritical or critical case ($\mathbb{E}(X) \leq 0$) and

$$\gamma = -\log(\inf\{\mathbb{E}(\exp(sX)) : s \in [0, 1]\})$$

in the subcritical case. In this latter case, $\gamma = -\log(\mathbb{E}(f'(1)))$ in the strongly or intermediate subcritical case ($\mathbb{E}(X \exp(X)) \leq 0$) whereas $\gamma > -\log(\mathbb{E}(f'(1)))$ in the weakly subcritical case ($\mathbb{E}(X \exp(X)) > 0$). We refer to [14] for more precise asymptotic results on the survival probability in the subcritical case.

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

$$S_0 = 0, \quad S_n - S_{n-1} = X_n \quad (n \geq 1).$$

where

$$X_n := \log(f'_n(1)) \quad (n \geq 1),$$

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

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

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

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

In the whole paper, we assume that there exists $s > 0$ such that the moment generating function $\mathbb{E}[\exp(sX)]$ is finite and we introduce the rate function Λ of the random walk ($S_n : n \in \mathbb{N}$)

$$\Lambda(\theta) := \sup_{\lambda \geq 0} \{\lambda\theta - \log(\mathbb{E}[\exp(\lambda X)])\} \quad (3)$$

As Λ is convex and lower semicontinuous, there is at most one $\theta \geq 0$ with $\Lambda(\theta) \neq \Lambda(\theta+)$. In this case, $\Lambda(\theta+) = \infty$ (see e.g [18], [12]). Usually, Λ is defined as the Legendretransform 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 $\Lambda(\theta) = 0$ for $\theta \leq \mathbb{E}[X]$ is convenient. \square

We write $L = L(f)$ for the random variable associated with the probability generating function f :

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

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

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

2 Main results and interpretation

We describe here 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 may decrease 'only exponentially'. More precisely, we work with the following assumption, which 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 $\mathcal{H}(\beta)$. *There exists a constant $0 < d < \infty$ such that for every $z \geq 0$,*

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

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

$$\psi(\theta) := \inf_{t \in [0, 1], s \in [0, \theta]} \left\{ t\gamma + \beta s + (1-t)\Lambda((\theta-s)/(1-t)) \right\} \quad (= \psi_{\gamma, \beta, \Lambda}(\theta)). \quad (4)$$

Note that in the supercritical case (i.e. $\mathbb{E}[\log(f'(1))] > 0$), ψ simplifies to

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

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

$$-\frac{1}{n} \log(\mathbb{P}(Z_n \geq e^{\theta n})) \xrightarrow{n \rightarrow \infty} \psi(\theta).$$

The assumptions in this Theorem ensure that the offspring distributions associated to 'some environments' have polynomial tails with exponent $-\beta$, and no tail distribution exceeds this exponent.

The upper bound is proved in section 3, while the proof of the lower bound is given 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 and we adapt in section 5 the arguments of the proof for $\beta \in (1, 2]$.

Remark: This theorem still holds if we just assume that there exists a slowly varying function l such that

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

instead of assumption $\mathcal{H}(\beta)$. Indeed, by properties of slowly varying functions (see [9], proposition 1.3.6, page 16), for any $\epsilon > 0$, there exists a constant d_ϵ such that $\mathbb{P}(L > z \mid f, L > 0) \leq d_\epsilon \cdot (m \wedge 1) \cdot z^{-\beta+\epsilon}$ a.s. As for fixed $\theta \geq 0$, $\psi_{\gamma,\beta,\Lambda}$ is continuous in β , letting $\epsilon \rightarrow 0$ yields the claim. \square

Let us give two consequences of this result. First, we derive a large deviation result for offspring distributions without heavy tails by letting $\beta \rightarrow \infty$, which generalizes Theorem 1 in [11].

Corollary 1. *If assumption $\mathcal{H}(\beta)$ is fulfilled for every $\beta > 0$, then for every $\theta \geq 0$,*

$$-\frac{1}{n} \log(\mathbb{P}(Z_n \geq e^{\theta n})) \xrightarrow{n \rightarrow \infty} \inf_{t \in [0,1]} \left\{ t\gamma + (1-t)\Lambda(\theta/(1-t)) \right\}.$$

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

Second we deal with the Galton Watson case, so the environment is not random and f is deterministic, meaning $\Lambda(\theta) = \infty$ for $\theta > \log m$ and $\Lambda(\log m) = 0$. We refer to [8, 22] for precise results for large deviations without heavy tails. For the decay rate of the survival probability, it is known that (see [5]) in the subcritical case ($m < 1$)

$$\gamma = -\log m$$

and $\gamma = 0$ in the critical ($m = 1$) and supercritical ($m > 1$) case. Thus, in the subcritical case,

$$\psi(\theta) = -\log m + \beta\theta.$$

In the critical and supercritical case, it remains to minimize

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

where $\Lambda(\theta) = 0$ for $\theta \leq \log m$ and $\Lambda(\theta) = \infty$ for $\theta > \log m$. Hence,

$$\psi(\theta) = \beta(\theta - \log m).$$

Path interpretation of the rate function. The rate function gives the exponential decay rate of the probability of reaching exceptionally large values, namely

$$\mathbb{P}(Z_n \geq \theta n) = \exp(-\psi(\theta)n + o(n)).$$

We consider the following 'natural paths' which reaches extraordinarily large values, i.e a path which realizes $\{Z_n \geq \exp(\theta n)\}$ for $n \gg 1$ and $\theta > \mathbb{E}[\log(f'(1))]$. At the beginning, up to time $\lfloor tn \rfloor$, there is a period without growth, that is 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. The probability of this event is given by $\mathbb{P}(Z_1 \geq \exp(sn))$ so it is of the order of $\exp(-\beta sn)$. Then the process grows exponentially according to its expectation in a 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 to observe this exceptionally good environment sequence decreases as $\exp(-(1-t)\Lambda((\theta - s)/(1-t))n)$. The most probable path to reach extraordinary large values $\exp(\theta n)$ at time n is then obtained by minimizing the sum of these three 'costs' γt , βs and

$(1-t)\Lambda((\theta-s)/(1-t))$, which gives the rate function ψ .

The optimal strategy to realize the large deviation event is given by the bivariate value (t_θ, s_θ) such that

$$\psi(\theta) = t_\theta\gamma + \beta s_\theta + (1-t_\theta)\Lambda((\theta-s_\theta)/(1-t_\theta)).$$

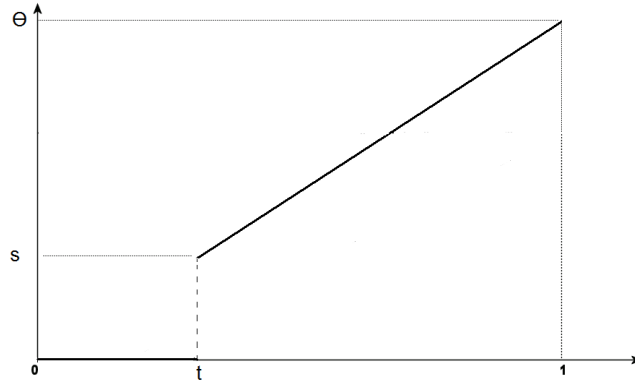
More formally, following the proof of [7], we should be able to prove the uniqueness of (t_θ, s_θ) (except for degenerated situations) and the forthcoming trajectorial result. But the proof become very heavy and technical. Conditionally on $Z_n \leq e^{cn}$, we expect that

$$\sup_{t \in [0,1]} \{|\log(Z_{[tn]})/n - f_\theta(t)|\} \xrightarrow{n \rightarrow \infty} 0$$

in probability in the sense of the uniform norm where

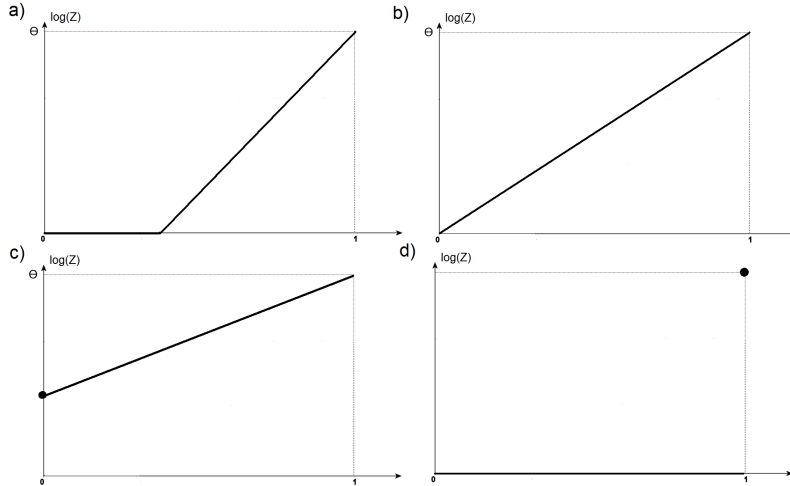
$$f_\theta(t) := \begin{cases} 0, & \text{if } t \leq t_\theta \\ \beta s_\theta + \frac{c}{1-t_\theta}(t-t_\theta), & \text{if } t > t_\theta. \end{cases}$$

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



As detailed in the next paragraph, several strategies may occur following the regime of the process and the value of θ . Except in degenerated cases when the associated path is not unique, we prove below using convexity arguments that the jump occurs at the beginning ($s_\theta > 0 \Rightarrow t_\theta = 0$) or at the end ($t_\theta = 1$) of the trajectory. Thus upper large deviation events correspond to one of the following trajectories.

Figure 2. Representation of the possible trajectories of the path associated to upper large deviations.



Obviously, keeping the population size small during a first period ($t_\theta > 0$) and growing later (Figure 2 a)) can be relevant only in the subcritical case. Actually we see below that the situation from Figure 2 a) only occurs if the process is strongly subcritical, as previously observed in [11] without heavy tails. In the subcritical case, if $\Lambda'(0) > \beta$, the associated optimal way is to keep the population size small by just surviving until the final time and then jump to the final value (Figure 2 d)). Then the population of size $\exp(\theta n)$ comes from a single parent of one of the last generations. The phase transitions are described later.

In the supercritical case ($\mathbb{E}[\log(f'(1))] > 0$), the process does not stay at zero ($t_\theta = 0$) but may jump at time $t = 0$, and then goes in straight line to reach θ . This corresponds to Figure 2 b) and c).

In the Galton Watson case with mean offspring m , the good strategy is either to survive until the final time and jump to the desired value θ (if $m \leq 1$), or to jump to $\theta - \log(m)$ and then grow normally (if $m > 1$).

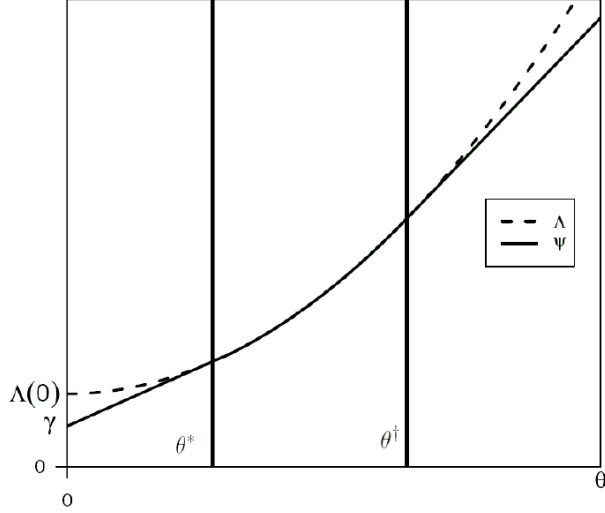
Graphical construction of the rate function. Here, we give another characterization of ψ , which will be useful to describe the strategy for upper large deviations in function of θ . As proved in Lemma 3 (see appendix), ψ is the largest convex function which satisfies for all $x, \theta \geq 0$

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

The first condition plays a role iff $\Lambda(0) > \gamma$, which corresponds to the strongly subcritical case (i.e. $\mathbb{E}[f'(1)\log(f'(1))] < 0$, see [14]). Indeed if $\mathbb{E}[X \exp(X)] < 0$, then the differentiation of $s \rightarrow \mathbb{E}[\exp(sX)]$ in $s = 1$ is negative and $\Lambda(0) = \sup\{-\log(\mathbb{E}[\exp(sX)]) : s \geq 0\} > -\log(\mathbb{E}[\exp(X)]) = \gamma$. If $\mathbb{E}[X \exp(X)] \geq 0$, [14] and the definition of Λ ensure that both γ and $\Lambda(0)$ are equal to $-\log(\mathbb{E}[\exp(\nu X)])$ where ν is characterized by $\mathbb{E}[X \exp(\nu X)] = 0$.

This characterization leads us to construct ψ by three pieces separated by θ^* and θ^\dagger .

Figure 3. *The following picture gives ψ in the strongly subcritical case:*



More explicitly, we define χ as the largest convex function which satisfies

$$\chi(0) \leq \gamma, \quad \chi(\theta) \leq \Lambda(\theta)$$

for all $\theta \geq 0$. This function is the rate function of Z in case of offspring distributions having at most geometric tails (see [11]) and is given by

$$\chi(\theta) = \begin{cases} \gamma(1 - \frac{\theta}{\theta^*}) + \frac{\theta}{\theta^*} \Lambda(\theta^*) & , \text{ if } \theta < \theta^* \\ \Lambda(\theta) & , \text{ else} \end{cases}$$

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

$$\frac{\Lambda(\theta^*) - \gamma}{\theta^*} = \inf_{\theta \geq 0} \frac{\Lambda(\theta) - \gamma}{\theta}. \quad (5)$$

Now define

$$\theta^\dagger = \sup \left\{ \theta \geq \max\{0, \mathbb{E}[X]\} : \chi'(\theta) \leq \beta \text{ and } \chi(\theta) < \infty \right\}. \quad (6)$$

Then

$$\psi(\theta) = \begin{cases} \chi(\theta) & , \text{ if } \theta \leq \theta^\dagger \\ \beta\theta - \log(\mathbb{E}[e^{\beta X}]) & , \text{ else} \end{cases}. \quad (7)$$

Phase Transitions Let us first describe the phase transitions (of order two) of the rate function and the strategies associated with when $\theta^\dagger > 0$. For that we use the following expression,

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

which can be guessed from the previous picture and is also proved in the first section of the Appendix.

For $\theta < \theta^*$, the rate function ψ is identical with χ . This means that no jump occurs. Conditionally 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 within a good environment such that $S_n - S_{\lfloor t_\theta n \rfloor} \approx \theta n$

(see Figure 2 a)). When θ increases, the survival period decreases whereas the geometric growth rate of the process remains constant and is equal to θ^* .

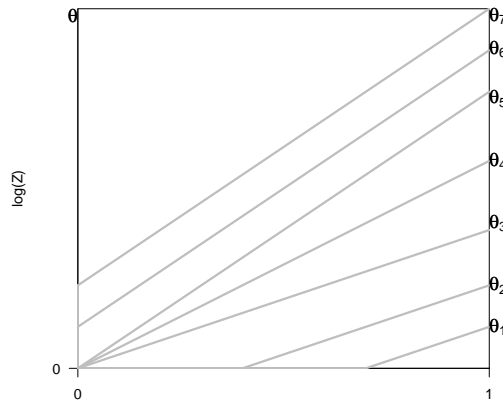
For $\theta^* \leq \theta \leq \theta^\dagger$, ψ is equal to Λ . Thus, conditionally on the large deviation event, the process grows exponentially (respectively linearly at the logarithmic scale) from the beginning to the end (see Figure 2 b)). This exceptional growth is due to a favorable environment such that $S_n \approx \theta n$.

For $\theta > \theta^\dagger$, the trajectory associated with begins now with a jump : $Z_1 \approx \exp(sn)$. Then it follows an exponential growth which corresponds to a favorable environment $S_n \approx (\theta - s)n$ (see Figure 2 c)). When θ increases, the initial jump increases whereas the rate of the exponential growth is still equal to θ^\dagger .

The case $\theta^\dagger = 0$ corresponds to $\psi(\theta) = \gamma + \beta\theta$. Here the good strategy consists in just surviving until the end and in one of the prelast generations, one individual has $\exp(\theta n)$ -many offsprings (see Figure 2 d)).

Finally, we note that in the case $0 < \theta = \theta^* = \theta^\dagger$, the best strategy is no longer unique. Indeed, for any $t \in (0, 1]$, there exists $s \in [0, \theta]$ such that all the following trajectories have the same cost. 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 $t \in [0, 1] \rightarrow f_\theta(t)$ in the strongly subcritical case for $\theta_1 < \theta_2 < \theta_3 = \theta^* < \theta_4 < \theta_5 = \theta^\dagger < \theta_6 < \theta_7$.



Notations: Unless otherwise is specified, we start the branching process from one single individual and denote by \mathbb{P} the probability associated with. We denote by \mathbb{P}_k the probability when the initial size of the population is equal to k . Large deviations results actually do not depend on the initial number of individuals if this latter is fixed (or bounded).

In the whole paper, we denote by $\Pi := (f_1, f_2, \dots)$ the complete environment.

For simplicity of notations, we are using several times \leq_c to indicate that the inequality holds up to some multiplicative constant (which does not depend on any variable).

Acknowledgements: *The authors are grateful to Götz Kersting and Julien Berestycki for fruitful discussions. The research was supported in part by the German Research Foundation (DFG), Grant*

3 Proof of the upper bound of Theorem 1

For the proof of the upper bound of Theorem 1, we need the following result. It ensures that exceptional growth of the population can at least be achieved thanks to some suitable good environment sequences, whose probability decreases exponentially following the rate function of the random walk $(S_n : n \in \mathbb{N})$. This result generalizes Proposition 1 in [7] for an exponential initial number of individuals. With a slight abuse, we denote below by $\exp(sn)$ the initial number of individuals instead of the integer part of $\exp(sn)$.

Proposition 1. *Under assumption $\mathcal{H}(\beta)$, for all $\theta \geq 0$ and $0 \leq s \leq \theta$,*

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}_{\exp(sn)}(Z_n \geq \exp(\theta n)) \leq \Lambda((\theta - s)_+).$$

Proof. For every $\theta' > 0$, we recall that

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

First, we assume that $\mathbb{E}[\exp(\lambda X)] < \infty$ for every $\lambda \geq 0$. Then the derivative of $\lambda \rightarrow \mathbb{E}[\exp(\lambda X)]$ exists for every $\lambda \geq 0$ and the supremum is reached in $\lambda = \lambda_{\theta'}$ such that

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

Following classical large deviations methods and more specifically [7], we introduce the probability $\tilde{\mathbb{P}}$ defined by

$$\tilde{\mathbb{P}}(X \in dx) = \frac{\exp(\lambda_{\theta'} x)}{\mathbb{E}[\exp(\lambda_{\theta'} X)]} \mathbb{P}(X \in dx).$$

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

For all $n \geq 1$, $\theta \in [0, \theta']$ and $\epsilon > 0$,

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

As $\tilde{\mathbb{P}}(S_n > (\theta' + \epsilon)n) \rightarrow 0$ when $n \rightarrow \infty$, we just need to prove that

$$\liminf_{n \rightarrow \infty} \tilde{\mathbb{P}}_{\exp(sn)}(Z_n \geq \exp([\theta + s]n)) > 0. \quad (8)$$

so that we can conclude the proof by letting $\epsilon \rightarrow 0, \theta' \rightarrow \theta$.

Relation (8) results from the fact that under $\tilde{\mathbb{P}}$ the population Z_n starting from one single individual grows as $\exp(S_n) \asymp n\theta'$ on the non-extinction event. More precisely, individuals of the initial population are labeled and the number of descendants in generation n of individual i is denoted by $Z_n^{(i)}$. Introduce then the 'success' probability p_n :

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

Then, conditionally on Π , for $N \geq 1$, the number of initial individuals whose number of descendants in generation n is larger than $N \exp(n\theta)$,

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

follows a binomial distribution of parameters $(\exp(sn), p_n)$. Moreover, as $\mathbb{E}[N_n \mid \Pi] = e^{sn} p_n$ a.s.,

$$\tilde{\mathbb{P}}_{\exp(sn)}(Z_n \geq \exp([\theta + s]n)) \geq \tilde{\mathbb{P}}_{\exp(sn)}(N_n \geq \exp(sn)/N) \geq \tilde{\mathbb{P}}_{\exp(sn)}\left(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{N p_n}\right).$$

Using the classical inequality due to Paley and Zygmund for $r \in [0, 1]$ (see e.g. [19] page 63),

$$\mathbb{P}(Y \geq r\mathbb{E}[Y]) \geq (1-r)^2 \frac{\mathbb{E}[Y]^2}{\mathbb{E}[Y^2]}, \quad (9)$$

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

$$\tilde{\mathbb{P}}_{\exp(sn)}\left(N_n \geq \frac{\mathbb{E}[N_n \mid \Pi]}{N p_n} \mid \Pi\right) \geq \left[1 - 1 \wedge \frac{1}{N p_n}\right]^2 \frac{\mathbb{E}[N_n \mid \Pi]^2}{\mathbb{E}[N_n^2 \mid \Pi]} \geq \frac{\left[1 - 1 \wedge \frac{1}{N p_n}\right]^2}{1 + \frac{e^{-sn}}{p_n}} \quad \text{a.s.}$$

Now, we use that under assumption $\mathcal{H}(\beta)$,

$$\tilde{\mathbb{E}}\left[\sum_{k \in \mathbb{N}} k^s \mathbb{P}(L = k \mid f)/m\right] \leq \tilde{\mathbb{E}}\left[\sum_{k \in \mathbb{N}} k^s \mathbb{P}(L = k \mid f, L > 0)\right] < \infty,$$

for every $1 < s < \beta$. So Theorem 3 in [17] ensures that for every $N \in \mathbb{N}$,

$$\tilde{\mathbb{E}}[p_n] = \tilde{\mathbb{P}}_1(Z_n \geq N \exp(\theta n)) \xrightarrow{n \rightarrow \infty} \tilde{\mathbb{P}}_1(\forall n \in \mathbb{N} : Z_n > 0) > 0.$$

As the right hand side does not depend on $N \geq 1$, we have for N large enough

$$\delta := \liminf_{n \rightarrow \infty} \tilde{\mathbb{P}}(p_n \geq 2/N) > 0$$

and get

$$\liminf_{n \rightarrow \infty} \mathbb{P}_{\exp(sn)}(Z_n \geq \exp([\theta + s]n)) \geq \liminf_{n \rightarrow \infty} \tilde{\mathbb{E}}\left[\frac{\left[1 - 1 \wedge 1/N p_n\right]^2}{1 + N/2}\right] \geq \frac{\delta(1 - 1/2)^2}{1 + N/2} > 0,$$

which proves (8) and ends up 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. [11] pages 10/11). \square

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

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

First, by (1),

$$-\frac{1}{tn} \log(\mathbb{P}(Z_{[tn]} > 0)) \xrightarrow{n \rightarrow \infty} \gamma .$$

Second, using that $\log(\mathbb{P}(Z_1 > z))/\log(z) \xrightarrow{z \rightarrow \infty} -\beta$, we have

$$-\frac{1}{n} \log(\mathbb{P}(Z_1 \geq \exp(sn))) \xrightarrow{n \rightarrow \infty} s\beta .$$

Finally, by Proposition 1, we get that

$$\limsup_{n \rightarrow \infty} -\frac{1}{(1-t)n} \log(\mathbb{P}_{\exp(sn)}(Z_{n-[tn]} \geq \exp(\theta n))) \leq \Lambda((\theta - s)/(1-t)+)$$

since

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

Combining the first inequality and the last three limits ensures that

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(\mathbb{P}(Z_n \geq \exp(\theta n))) \leq \inf_{t \in [0,1], s \in [0,\theta]} \left\{ t\gamma + \beta s + (1-t)\Lambda((\theta - s)/(1-t)+) \right\}.$$

As convex nonnegative function, Λ has at most one jump (to infinity). Thus the above infimum is $\psi(\theta)$. To see this, we only have to consider the jump point. Say, there are $s_\theta \in [0, \theta]$ and $t_\theta \in [0, 1)$ such that

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

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

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

Now letting $\epsilon \rightarrow 0$ proves the result and thereby the upper bound of Theorem 1. \square

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

We introduce the minimum of the associated random walk up to time n :

$$M_n := \min_{0 \leq k \leq n} S_k \quad \text{a.s.}$$

Using that $\mathbb{P}(Z_n > 0 | \Pi) \leq \mathbb{E}[Z_n | \Pi] = \exp(S_n)$ and $\mathbb{P}(Z_n > 0 | \Pi)$ decreasing a.s., we get the following classical inequality (see e.g. [10])

$$\mathbb{P}(Z_n > 0 | \Pi) \leq e^{M_n} \quad \text{a.s.} \tag{10}$$

Actually, the above estimate gives the correct exponential decay rate (see e.g. [10]):

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

In Lemma 1, the above relation is generalized and proved rigorously under assumption $\mathcal{H}(\beta)$.

For the proof of the lower bound of the main theorem, we need the following key bound for the tail probability of Z_n .

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

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

Let us explain briefly this result. The probability to survive until time n evolves as $\exp(M_n)$, nice environment sequences correspond to large values of $(S_n - M_n)$ and high reproduction of the initial individual gives the last term $k^{-\beta}$. Conditionally on the environment sequence and the survival of the process, the growth of the process follows $\exp(S_n - M_n)$: this corresponds to the 'best time period' for the growth of the process. Thus, this theorem essentially says that conditionally on $Z_n > 0$, the tail distribution of $Z_n / e^{S_n - M_n}$ is at most polynomial with exponent $-\beta$.

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

$$f_{0,n}(s) := \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n = k | \Pi) = \mathbb{E}[s^{Z_n} | \Pi], \quad \text{a.s.} \quad (0 \leq s \leq 1). \quad (11)$$

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

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

and

$$g_{0,n}(s) := \sum_{k=0}^{\infty} s^k \mathbb{P}(Z_n > k | \Pi) = \frac{1 - f_{0,n}(s)}{1 - s} \quad \text{a.s.} \quad (0 \leq s \leq 1). \quad (12)$$

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

$$h_{\mu,k}(s) := \frac{1}{(1 - f_k(s))^\mu} - \frac{1}{(f'_k(1)(1 - s))^\mu} = \frac{g_k(1)^\mu - g_k(s)^\mu}{(g_k(1)g_k(s)(1 - s))^\mu} \quad \text{a.s.} \quad (0 \leq s \leq 1). \quad (13)$$

Finally, we define for all $1 \leq k \leq n$,

$$\begin{aligned} U_k &:= (f'_1(1) \cdots f'_k(1))^{-1} = f'_{0,k}(1)^{-1} = e^{-S_k}, \\ f_{k,n} &:= f_{k+1} \circ f_{k+2} \circ \cdots \circ f_n, \quad 0 \leq k < n; \quad f_{n,n} = id \quad \text{a.s.} \end{aligned}$$

By a telescope summation argument similar to [13], we have

$$\begin{aligned} \frac{1}{(1 - f_{0,n}(s))^\mu} &= \frac{U_0^\mu}{(1 - f_{0,n}(s))^\mu} \\ &= \frac{U_n^\mu}{(1 - f_{n,n}(s))^\mu} + \sum_{k=0}^{n-1} \left(\frac{U_k^\mu}{(1 - f_{k,n}(s))^\mu} - \frac{U_{k+1}^\mu}{(1 - f_{k+1,n}(s))^\mu} \right) \\ &= \frac{U_n^\mu}{(1 - s)^\mu} + \sum_{k=0}^{n-1} U_k^\mu \left(\frac{1}{(1 - f_{k+1}(f_{k+1,n}(s)))^\mu} - \frac{1}{(f'_{k+1}(1)(1 - f_{k+1,n}(s)))^\mu} \right) \\ &= \frac{U_n^\mu}{(1 - s)^\mu} + \sum_{k=0}^{n-1} U_k^\mu h_{\mu,k+1}(f_{k+1,n}(s)), \quad s \geq 0. \end{aligned} \quad (14)$$

Proof of Theorem 2. In the same vein as [11], we are obtaining 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 $s \rightarrow 1$. In that purpose, we use (14) for $\mu = \beta - 1$ and get

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

Then we calculate the first derivative of $g_{0,n}$:

$$\begin{aligned} & g'_{0,n}(s) \\ = & -(\beta-1)^{-1} \left(U_n^{\beta-1} + (1-s)^{\beta-1} \sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s)) \right)^{-1-1/(\beta-1)} \\ & \times \left(-(\beta-1)(1-s)^{\beta-2} \sum_{k=0}^{n-1} U_k^{\beta-1} h_{\beta-1,k+1}(f_{k+1,n}(s)) \right. \\ & \quad \left. + (1-s)^{\beta-1} \sum_{k=0}^{n-1} U_k^{\beta-1} h'_{\beta-1,k+1}(f_{k+1,n}(s)) f'_{k+1,n}(s) \right) \\ \leq & \frac{\sum_{k=0}^{n-1} U_k^{\beta-1} \left(h_{\beta-1,k+1}(f_{k+1,n}(s)) - (\beta-1)^{-1} h'_{\beta-1,k+1}(f_{k+1,n}(s)) f'_{k+1,n}(s) (1-s) \right)}{U_n^{\beta} (1-s)^{2-\beta}} \quad (15) \end{aligned}$$

Now Lemma 4 in the appendix ensures that there exists $c > 0$ such that for every $s \in [0, 1)$,

$$h_{\beta-1,k}(s) \leq c\Upsilon(1/(1-s)), \quad (16)$$

$$-h'_{\beta-1,k}(s) \leq c\Upsilon(1/(1-s))/(1-s) \quad \text{a.s.} \quad (17)$$

Moreover, using (14), Lemma 4 in the appendix for $0 < \mu < \beta - 1$ and $U_k \leq \exp(-M_n)$ for every $0 \leq k \leq n$, there exists a $c \geq 1$ such that for every $s \in [0, 1)$,

$$\frac{1}{(1-f_{k+1,n}(s))^\mu} \leq \frac{e^{-\mu M_n}}{(1-s)^\mu} + n c e^{-\mu M_n} \leq c e^{-\mu M_n} (n+1)/(1-s)^\mu \quad \text{a.s.}$$

Combining this inequality with (16) ensures that there exists $c > 0$ such that

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

Moreover, $f_{k+1,n}(s) \leq 1 - f'_{k+1,n}(s)(1-s)$ by convexity of $f_{k+1,n}$ and (17) ensures that

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

Using the two last estimates with $\mu = (\beta - 1)/2$ together in (15) yields

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

Moreover for all $k \geq 1$ and $s \in [0, 1]$,

$$\begin{aligned} g'_{0,n}(s) & \geq \sum_{j=k/2}^k j \mathbb{P}(Z_n > j|\Pi) s^{j-1} \\ & \geq s^k \frac{k^2}{2} \mathbb{P}(Z_n > k|\Pi). \end{aligned} \quad (18)$$

By letting $s = 1 - 1/k$ in the two last inequalities, we get

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

which ends up the proof since $U_n = \exp(-S_n)$. \square

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

Lemma 1. *Under assumption $\mathcal{H}(\beta)$, for all $\theta \geq 0$, $b > 0$ and Υ positive nondecreasing and slowly varying at infinity,*

$$\gamma = - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\Upsilon(n^b e^{\theta n} e^{-M_n}) e^{M_n}].$$

Proof of Lemma 1. First let $\Upsilon = 1$. We use (14) with some $0 < \mu < \beta - 1$ and (38) ensures that

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

For the upper bound, we use (10) and get

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

As Υ is nondecreasing,

$$\gamma \geq \limsup_{n \rightarrow \infty} - \frac{1}{n} \log \mathbb{E}[\Upsilon(n^b e^{\theta n} e^{-M_n}) e^{M_n}].$$

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

$$\xi(t) := - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{tM_n}].$$

We note that $\xi(t) \geq 0$ and by Lemma V.4 in [18], $\xi(t)$ is finite and convex. So ξ is continuous.

Now by properties of slowly varying sequences (see [9], proposition 1.3.6, page 16), for any $\delta > 0$, $x^{-\delta} \Upsilon(x) \rightarrow 0$ as $x \rightarrow \infty$ (see appendix) and

$$- \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\Upsilon(n^b e^{\theta n} e^{-M_n}) e^{M_n}] \geq -\delta\theta - \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{(1+\delta)M_n}].$$

Letting $\delta \rightarrow 0$ and using continuity of χ , this ends up the proof. \square

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

$$\mathbb{P}(S_n \geq \theta n) \leq e^{-\Lambda(\theta)n} \tag{19}$$

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

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

We decompose the probability of having an extraordinarily large population according to $S_n - M_n$.

$$\mathbb{P}(Z_n \geq e^{\theta n}) = \mathbb{P}(Z_n \geq e^{\theta n}, S_n - M_n \geq \theta n) + \mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n} | \Pi); S_n - M_n < \theta n]. \quad (20)$$

The asymptotic of the first term can be found using (19) (see [11]):

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

This ensures that

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}, S_n - M_n \geq \theta n) \geq \chi(\theta), \quad (21)$$

where

$$\chi(\theta) = \inf_{0 < t \leq 1} \{t\gamma + (1-t)\Lambda(\theta/(1-t))\}.$$

For the second term, we use Theorem 2 and the Markov property for $(S_n : n \geq 0)$:

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

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

$$\mathbb{E}[\Upsilon(n^{2/(\beta-1)} e^{-S_k} e^{\theta n}) e^{S_k}; \tau_k = k] = \mathbb{E}[\Upsilon(n^{2/(\beta-1)} e^{-M_k} e^{\theta n}) e^{M_k}, \tau_k = k] \leq \mathbb{E}[\Upsilon(n^{2/(\beta-1)} e^{-M_k} e^{\theta n}) e^{M_k}]$$

and we deduce from (19) that

$$\begin{aligned} &\mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n} | \Pi); S_n - M_n < \theta n] \\ &\leq c n^{[\beta]} \sum_{k=1}^n \mathbb{E}[\Upsilon(n^{2/(\beta-1)} e^{-M_k} e^{\theta n}) e^{M_k}] \sum_{j=0}^{m_\epsilon} e^{-\beta(\theta - (j+1)\epsilon)n} \mathbb{P}(S_{n-k} \in [nj\epsilon, n(j+1)\epsilon], M_{n-k} \geq 0) \\ &\leq c n^{[\beta]} \sum_{k=1}^n \mathbb{E}[\Upsilon(n^{2/(\beta-1)} e^{-M_k} e^{\theta n}) e^{M_k}] \sum_{j=0}^{m_\epsilon} e^{-\beta(\theta - (j+1)\epsilon)n} e^{-\Lambda(j\epsilon n/(n-k))(n-k)} \\ &\leq c \theta n^5 \sup_{0 < t \leq 1, 0 \leq s \leq \theta} \left\{ \mathbb{E} \left[\Upsilon(n^{2/(\beta-1)} e^{-M_{\lfloor tn \rfloor}} e^{\theta n}) e^{M_{\lfloor tn \rfloor}} \right] \cdot e^{-(\beta s + (1-t)\Lambda((\theta-s)/(1-t)))n} \right\}. \end{aligned}$$

Together with Lemma 1, this yields

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{E}[\mathbb{P}(Z_n \geq e^{\theta n} | \Pi); S_n - M_n < \theta n] \geq \psi(\theta),$$

where

$$\psi(\theta) = \inf_{0 < t \leq 1, 0 \leq s \leq \theta} \left\{ \gamma t + \beta s + (1-t)\Lambda((\theta-s)/(1-t)) \right\}.$$

Combining this inequality with (20) and (21) gives

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \geq \min\{\chi(\theta); \psi(\theta)\}.$$

Adding that $\psi(\theta) \leq \chi(\theta)$ since the infimum is considered on a larger set for ψ than for χ , we get

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n \geq e^{\theta n}) \geq \psi(\theta),$$

which proves the lower bound of Theorem 1. \square

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

First, Lemma 1 still holds for $\beta > 2$ by following the same proof. Indeed, using (14) for $\mu = 1$ together with Lemma 4 given in the appendix ensures that

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

The main difficulty is to obtain an equivalent of Theorem 2. For this, we need to calculate higher order derivatives of $g_{0,n}$ and the upper bound on the tail probability of Z_n contains an additional term:

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

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

For the proof, we use the functions

$$h_k(s) = \frac{1}{(1 - f_k(s))} - \frac{1}{f_k'(1)(1 - s)} = \frac{g_k(1) - g_k(s)}{g_k(1)g_k(s)(1 - s)} \quad a.s. \quad (0 \leq s < 1)$$

and

$$H(s) = \sum_{k=0}^{n-1} U_k h_{k+1}(f_{k+1,n}(s)) \quad a.s. \quad (0 \leq s < 1). \quad (22)$$

Then (14) with $\mu = 1$ gives

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

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

$$\frac{d^l}{ds^l} g_{0,n}(s)^{-1} = (1 - s)H^{(l)}(s) - lH^{(l-1)}(s) \quad a.s. \quad (0 \leq s < 1). \quad (23)$$

The rest of the section is organized as follows. First, we prove the following technical lemma which gives useful bounds for power generating series. Then we derive Theorem 3. Finally the main lines of the proof of the lower bound of Theorem 1 for $\beta > 2$ are explained (following the proof for $\beta \in (1, 2]$). For simplicity of notation, we introduce \leq_c which means that the inequality is fulfilled up to a multiplicative constant c which does not depend on s, k, l or ω .

Lemma 2. *Under assumption $\mathcal{H}(\beta)$, for every $l \leq \lceil \beta \rceil - 1$,*

$$f_{0,n}^{(l)}(1) \leq_c n^{l-1} e^{S_n} e^{(l-1)(S_n - M_n)} \quad a.s. \quad (24)$$

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 $l = \lceil \beta \rceil - 1$

$$|H^{(l)}(s)| \leq_c n^l e^{l(S_n - M_n)} \quad (25)$$

$$|H^{(l)}(s)| \leq_c n^l e^{\lceil \beta \rceil - 2(S_n - M_n)} + n \Upsilon(n^2 e^{-M_n} (1-s)^{-1}) (1-s)^{-\lceil \beta \rceil - \beta} e^{-S_n} e^{(\beta-1)(S_n - M_n)} \quad (26)$$

$$|H^{(l)}(s)| \leq_c n^l e^{\lceil \beta \rceil - 1(S_n - M_n)} + n^2 \Upsilon(n^2 e^{-M_n} (1-s)^{-1}) e^{-S_n} e^{\beta(S_n - M_n)} (1-s)^{-\lceil \beta \rceil - \beta} + n \Upsilon(n^2 e^{-M_n} (1-s)^{-1}) e^{-S_n} e^{(\beta-1)(S_n - M_n)} (1-s)^{-1 - \lceil \beta \rceil - \beta}. \quad (27)$$

Proof. We prove the Lemma by induction with respect to l and all the following relations hold a.s. for every $s \in [0, 1]$. For $l = 1$, (24) is trivially fulfilled since $f'_{0,n}(1) = e^{S_n}$. First, we consider $l < \lceil \beta \rceil - 2$ and we assume that (24) holds for every $i \leq l$. We are first proving that (25) holds for l and then that (24) holds for $l + 1$.

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

$$f_{k+1,n}^{(i)}(s) \leq f_{k+1,n}^{(i)}(1) \leq_c n^{i-1} e^{S_n} e^{(i-1)(S_n - S_k - \min_{j=k,\dots,n} \{S_j - S_k\})} \leq_c n^{i-1} e^{S_n} e^{(i-1)(S_n - M_n)}. \quad (28)$$

Lemma 6 given in the appendix ensures that (see Lemma 6) for the definition of $u_{j,l}$

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

and using (28)

$$u_{j,l}(s) \leq_c n^{l-j} e^{jS_n} e^{(l-j)(S_n - M_n)} \leq_c n^{l-1} e^{S_n} e^{(l-1)(S_n - M_n)}.$$

By Lemma 5 also given in the appendix, for $j < \lceil \beta \rceil - 2$, the derivatives $h_k^{(j)}$ are bounded by a constant that does not depend on ω . Thus

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

Then recalling (22), we have

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

which gives (25) for $l < \lceil \beta \rceil - 2$.

We can now prove that (24) is fulfilled for $l + 1 < \lceil \beta \rceil - 1$. Using Lemma 6 again (see (45)) with $f = g_{0,n}$ and $h(x) = 1/x$, we get

$$\begin{aligned} \frac{d^l}{ds^l} g_{0,n}(s)^{-1} &= \sum_{j=1}^l (-1)(-2) \cdots (-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s) \\ &= -g_{0,n}(s)^{-2} g_{0,n}^{(l)}(s) + \sum_{j=2}^l (-1)(-2) \cdots (-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s), \end{aligned} \quad (29)$$

where

$$u_{j,l}(s) = \sum_{i=(i_1, \dots, i_{2j}) \in \mathcal{C}(j,l)} c_i (g_{0,n}^{(i_1)}(s))^{i_2} \cdots (g_{0,n}^{(i_{2j-1})}(s))^{i_{2j}},$$

and $\mathcal{C}(j,l) = \{(i_1, \dots, i_{2j}) \in \mathbb{N}^{2j} \mid i_1 i_2 + i_3 i_4 + \dots = l \text{ and } i_2 + i_4 + \dots = j\}$.

Moreover, the induction assumption (24) and (35) give for every $i \leq l-1$,

$$g_{0,n}^{(i)}(1) \leq_c n^i e^{S_n} e^{i(S_n - M_n)}.$$

Thus

$$u_{j,l}(1) \leq_c n^l e^{j S_n} e^{l(S_n - M_n)}.$$

By (22), the left hand-side of (29) is equal to $(1-s)H^{(l)}(s) - lH^{(l-1)}(s)$. By (25), for $l < \lceil \beta \rceil - 2$, $(1-s)H^{(l)}(s)$ vanishes for $s=1$. Thus letting $s=1$ and noting that $g_{0,n}(1) = e^{S_n}$ yields

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

As we have already proved (25) for $l < \lceil \beta \rceil - 2$, we get

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

Using (35), we get (24) for $l+1$, which completes the induction and proves (24) for $l < \lceil \beta \rceil - 1$.

Let us prove the bound on $H^{(l)}(s)$ for $l = \lceil \beta \rceil - 2$. Using again Lemmas 5 and 6 and (24) yields

$$\begin{aligned} &\left| \frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s)) \right| \\ &= \left| \sum_{j=1}^{l-1} h_{k+1}^{(j)}(f_{k+1,n}(s)) u_{j,l}(s) + h_{k+1}^{(l)}(f_{k+1,n}(s)) (f'_{k+1,n}(s))^l \right| \\ &\leq_c n^{l-1} e^{S_n} e^{(\lceil \beta \rceil - 2)(S_n - M_n)} + \Upsilon(1/(1 - f_{k+1,n}(s))) (1 - f_{k+1,n}(s))^{-(\lceil \beta \rceil - \beta)} (f'_{k+1,n}(s))^l. \end{aligned} \quad (30)$$

Now by the same arguments as in the proof of Theorem 2, $\Upsilon(1/(1 - f_{k+1,n}(s))) \leq \Upsilon(n^2 e^{-M_n} (1-s)^{-1})$ and by convexity,

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

Using also $f'_{k+1,n}(s) \leq e^{S_n - M_n}$, by (30) follows

$$\left| \frac{d^l}{ds^l} h_{k+1}(f_{k+1,n}(s)) \right| \leq_c n^{l-1} e^{S_n} e^{(\lceil \beta \rceil - 2)(S_n - M_n)} + \Upsilon(n^2 e^{-M_n} (1-s)^{-1}) (1-s)^{-(\lceil \beta \rceil - \beta)} e^{(\beta-2)(S_n - M_n)}.$$

Combining this inequality with (22) proves (26).

This implies that $(1-s)H^{(l)}(s) \rightarrow 0$ as $s \rightarrow 1$ for $l = \lceil \beta \rceil - 2$. Thus we can apply the same arguments to get an upper bound for $g_{0,n}^{(l)}(1)$ and prove (24) for $l = \lceil \beta \rceil - 1$.

Finally, let $l = \lceil \beta \rceil - 1$. We apply just the same arguments as before. Then Lemmas 5 and 6 yield

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

Using again (22), this proves (27). \square

Proof of Theorem 3 for $\beta > 2$. Let $l = \lceil \beta \rceil - 1$. Without loss of generality, we assume $\Upsilon \geq 1$. The following relations hold a.s. Using (29) and (23),

$$g_{0,n}^{(l)}(s) = g_{0,n}(1)^2 \left(- (1-s) H^{(l)}(s) + l H^{(l-1)}(s) + \sum_{j=2}^l (-1)(-2) \cdots (-j) g_{0,n}(s)^{-(j+1)} u_{j,l}(s) \right)$$

Now using (26), (27), (29) as well as $\exp(S_n) \leq \exp(S_n - M_n)$ for the first terms and (24) together with (35) for the last term yields

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

Analogously to (18), we get the following estimate for every $1/2 \leq s < 1$,

$$g_{0,n}^{(l)}(s) \geq s^k \frac{k^{l+1}}{2} \mathbb{P}(Z_n > k | \Pi).$$

Choosing $s = 1 - 1/k$ yields

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

Using that for all $a \geq 1$ and $b \geq 0$, the function $x \rightarrow a^{-x} \exp((x-1)b)$ is monotone and that $\beta \leq \lceil \beta \rceil < \beta + 1 \leq \lceil \beta \rceil + 1$, we have for all $k \geq 1$,

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

Combining the two last inequalities leads to

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

which completes the proof. \square

Proof of the lower bound of Theorem 1 for $\beta > 2$. The proof now follows the proof for $\beta \in (1, 2]$. Theorem 3 yields

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \geq \min \{ \psi_{\gamma, \beta, \Lambda}(\theta), \psi_{\gamma, \lceil \beta \rceil + 1, \Lambda}(\theta) \},$$

where ψ is defined in (4). Now using the characterization of ψ given in forthcoming Lemma 3, we deduce that for any $\theta \geq 0$,

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

Thus

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

and we get the expected lower bound. \square

6 Proof of the corollary

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

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

Then we can apply the lower bound in Theorem 1 for every $\beta > 0$. This yields for all $\beta > 0$ and $\theta \geq 0$,

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \geq \psi_{\gamma, \beta, \Lambda}(\theta).$$

Now taking the limit $\beta \rightarrow \infty$, the monotone convergence of $\psi_{\gamma, \beta, \Lambda}$ yields

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Z_n > e^{\theta n}) \geq \psi_{\gamma, \infty, \Lambda}(\theta),$$

where

$$\begin{aligned} \psi_{\gamma, \infty, \Lambda}(\theta) &:= \lim_{\beta \rightarrow \infty} \inf_{t \in [0, 1], s \in [0, \theta]} \left\{ t\gamma + \beta s + (1-t)\Lambda((\theta-s)/(1-t)) \right\} \\ &= \inf_{t \in [0, 1]} \left\{ t\gamma + (1-t)\Lambda(\theta/(1-t)) \right\}. \end{aligned}$$

This gives the upper bound and the lower bound follows readily the proof given in Section 3 where we consider the natural associated path (or see [11]). \square

7 Appendix

We give in this section several technical results useful for the proofs.

7.1 Characterization of the rate function ψ

Lemma 3. *Let $0 \leq \gamma \leq \Lambda(0)$ and $\beta > 0$. The function ψ defined for $\theta \geq 0$ by*

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

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

$$\psi(0) = \gamma, \quad \psi(\theta) \leq \Lambda(\theta), \quad \psi(\theta + x) \leq \psi(\theta) + \beta x. \quad (31)$$

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

$$\begin{aligned}
& \lambda\psi(\theta') + (1 - \lambda)\psi(\theta'') \\
& \geq \lambda[t'\gamma + \beta s' + (1 - t')\Lambda((\theta' - s')/(1 - t'))] \\
& \quad + (1 - \lambda)[t''\gamma + \beta s'' + (1 - t'')\Lambda((\theta'' - s'')/(1 - t''))] - \epsilon \\
& \geq [\lambda t' + (1 - \lambda)t'']\gamma + [\lambda s' + (1 - \lambda)s'']\beta \\
& \quad + \left(1 - [\lambda t' + (1 - \lambda)t'']\right)\Lambda\left(\frac{\lambda\theta' + (1 - \lambda)\theta'' - (\lambda s' + (1 - \lambda)s'')}{1 - [\lambda t' + (1 - \lambda)t'']}\right) - \epsilon \\
& \geq \psi(\lambda\theta' + (1 - \lambda)\theta'') - \epsilon.
\end{aligned}$$

Letting $\epsilon \rightarrow 0$ entails that ψ is convex.

Second, following the previous computation, we verify that ψ fulfills (31). For any $\theta \geq 0$ and $\epsilon > 0$, there exist $t' \in [0, 1)$ and $s' \in [0, \theta]$ such that

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

Taking the limit $\epsilon \rightarrow 0$ yields the second property in (31). Furthermore, letting $t = 0$ and $s = 0$ implies $\psi(\theta) \leq \Lambda(\theta)$ and $t \rightarrow 1$ entails that $\psi(0) \leq \gamma$. This completes the proof of (31).

Finally, let κ be any convex function which satisfies (31). Using these assumptions ensures that for all $t \in [0, 1)$ and $0 \leq s \leq \theta$,

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

Taking the infimum over s and t , we get $\psi(\theta) \geq \kappa(\theta)$ and the proof is complete. \square

We give now describe a last characterization of ψ that results from Lemma 3 (see Figure 3). Let θ^* and θ^\dagger be defined as in (5) and (6) and assume $0 < \theta^* < \theta^\dagger < \infty$. As convex and monotone function, Λ has at most one jump (to infinity). Let this jump be in $0 < \theta_j \leq \infty$ and $\Lambda(\theta)$ is differentiable for $\theta < \theta_j$. As Λ is also continuous from below, $\Lambda(\theta_j) < \infty$. Now, by the preceding characterization, ψ is the largest convex function, starting in $\psi(0) = \gamma$, being at most as large as Λ and having at most slope β .

The largest convex function through the point $(0, \gamma)$ being smaller/equal than Λ has to be linear and has to be a tangent of Λ . By definition of θ^* , the tangent at Λ in θ^* goes through the point $(0, \gamma)$. Thus ψ is linear for $\theta < \theta^*$ and follows this tangent. For $\theta > \theta^*$, ψ is identical with Λ until the slope of Λ is exactly β (or until Λ jumps to infinity). At this point θ^\dagger , the last condition

becomes important and ψ is linear with slope β for $\theta > \theta^\dagger$. Summing up,

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

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

7.2 Slowly varying functions

In this section, we recall some properties of regularly varying functions and we refer to [9] for details. The function $\Upsilon : (0, \infty) \rightarrow (0, \infty)$ is a slowly varying function if for every $a > 0$,

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

We need a Tauberian result from [16], p. 423. See also [9], Theorem 1.5.11, page 28. For any $\alpha > -1$, the function $g(s) := \sum_{k=0}^{\infty} s^k k^\alpha$ satisfies

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

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

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

For $\alpha = -1$, $\sum_{k=1}^{\infty} s^k/k = -\log(1 - s)$. As the logarithm is a slowly varying function, we rewrite the previous results in the following way, which will be convenient in the proofs.

There exists a nondecreasing positive slowly varying function Υ such that for all $\alpha \geq -1$ and $s \in [0, 1)$

$$\sum_{k=1}^{\infty} s^k k^\alpha \leq \Upsilon(1/(1 - s))(1 - s)^{-1-\alpha}. \quad (32)$$

7.3 Bounds for generating functions

Let L be a random variable with values in $\{0, 1, 2, \dots\}$ 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 the following function associated to f ,

$$g(s) := \sum_{k=0}^{\infty} s^k q_k = \frac{1 - f(s)}{1 - s}, \quad (33)$$

where the last identity comes from Cauchy product of power series (see also [11]). We recall that the l -th derivative of a function f is denoted by $f^{(l)}$ and that $f^{(l)}(s)$ and $g^{(l)}(s)$ exist for every $s \in [0, 1)$. As

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

all derivatives of f and g are nonnegative, nondecreasing functions. We are using g instead of f in the proofs since the associated sequence $(q_k)_{k \in \mathbb{N}}$ is monotone, which is more convenient. Calculating the l -th derivative of $f(s) = 1 - (1-s)g(s)$ gives

$$f^{(l)}(s) = lg^{(l-1)}(s) - (1-s)g^{(l)}(s). \quad (34)$$

Thus $g^{(l-1)}(1)$ and $f^{(l)}(1)$ both essentially describe the l -th moment of the corresponding probability distribution. More precisely, if $g^{(l-1)}(1)$ is finite, then $f^{(l)}(1)$ is finite and

$$f^{(l)}(1) = lg^{(l-1)}(1). \quad (35)$$

Conversely if $f^{(l)}(1) < \infty$, then

$$g^{(l-1)}(1) = \sum_{k=1}^{\infty} k(k-1)\cdots(k-l+1)q_k \leq \sum_{k=1}^{\infty} k^l p_k < \infty$$

and $g^{(l-1)}(1) < \infty$.

For $\mu \in (0, 1]$, we also define the function

$$h_\mu(s) := \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu}. \quad (36)$$

The following useful lemmas give 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 can rewrite assumption $\mathcal{H}(\beta)$ in the following way

$$q_k \leq d g(0) (g(1) \wedge 1) k^{-\beta} \quad (k \geq 1). \quad (37)$$

Lemma 4. *Let $\beta > 1$ and assume that (37) 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 $s \in [0, 1]$,*

$$h_\mu(s) \leq c. \quad (38)$$

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 $s \in [0, 1]$,

$$h_{\beta-1}(s) \leq \Upsilon(1/(1-s)) \quad (39)$$

$$-h'_{\beta-1}(s) \leq \Upsilon(1/(1-s))/(1-s). \quad (40)$$

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

In the proofs, we use again the notation \leq_c which means that the inequality is fulfilled up to a multiplicative constant which depends on β and μ but is independent of s and the order of the differentiation.

Proof. Using $g(s) \geq g(0)$, we have

$$\begin{aligned} h_\mu(s) &= \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(s)(1-s))^\mu} \\ &\leq \frac{g(1)^\mu - g(s)^\mu}{(g(1)g(0)(1-s))^\mu} \\ &\leq (g(1) \wedge 1)^{-1} \frac{(\sum_{k=0}^{\infty} g(0)^{-1} q_k)^\mu - (\sum_{k=0}^{\infty} s^k q_k g(0)^{-1})^\mu}{(1-s)^\mu}. \end{aligned} \quad (41)$$

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

$$1 = q_0/g(0) \leq x := \sum_{k=0}^{\infty} s^k q_k g(0)^{-1} \leq a := \sum_{k=0}^{\infty} q_k g(0)^{-1}. \quad (42)$$

Then $x^{\mu-1} \leq 1$ and using the inequality of concavity in (41) with $q_k \leq dg(0) \cdot (g(1) \wedge 1) \cdot k^{-\beta}$ leads to

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

The estimates (38) and (39) on h_μ for $0 < \mu < (\beta - 1) \wedge 1$ and $\mu = \beta - 1$ now follow directly from (32). For $\mu = 1$, $\beta > 2$ and $s = 1$, the sum is finite and (38) also holds in this case.

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

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

Differentiating both sides yields

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

and thus

$$-h'_{\beta-1}(s) \leq (\beta-1) \left(\frac{h_{\beta-1}(s)g'(s)}{g(s)} + \frac{g'(s)}{g(s)g(1)^{\beta-1}(1-s)^{\beta-1}} - \frac{g(1)^{\beta-1} - g(s)^{\beta-1}}{g(s)^{\beta-1}g(1)^{\beta-1}(1-s)^\beta} \right)$$

As g is nondecreasing, we can skip the the last term which is negative. Using (37) and (39), we get

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

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

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

The result now follows from (32) and the fact that the product of two slowly varying functions is still slowly varying. \square

We consider now

$$h(s) = h_1(s) = \frac{g(1) - g(s)}{g(1)g(s)(1-s)}.$$

Lemma 5. *We assume that (37) holds for some $\beta > 1$. Then there exists a finite constant $c = c(\beta, d) < \infty$ such that for every $s \in [0, 1)$,*

$$\begin{aligned} |h^{(l)}(s)| &\leq c && \text{if } 0 \leq l < \beta - 2 \\ |h^{(\lceil \beta \rceil - 2)}(s)| &\leq c \Upsilon(1/(1-s)) (1-s)^{-(\lceil \beta \rceil - \beta)} && \text{if } \beta \geq 2 \\ |h^{(\lceil \beta \rceil - 1)}(s)| &\leq c \Upsilon(1/(1-s)) (1-s)^{-1 - (\lceil \beta \rceil - \beta)}. \end{aligned} \quad (43)$$

Proof. By (36) and Cauchy product of power series, for every $s \in [0, 1)$,

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

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

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

Moreover, (37) ensures that for all $s \in [0, 1)$ and $j < \beta - 2$,

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

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

$$\begin{aligned} |h^{(l)}(s)| &\leq c g(s)^{-1} \left(g(1)^{-1} g(0) \cdot (g(1) \wedge 1) \cdot \sum_{k=0}^{\infty} k^l s^{k-l} \sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^l \binom{l}{j} g^{(j)}(1) |h^{(l-j)}(s)| \right) \\ &\leq c \sum_{k=0}^{\infty} k^l s^{k-l} \sum_{j=k+1}^{\infty} j^{-\beta} + \sum_{j=1}^l \binom{l}{j} |h^{(l-j)}(s)| \end{aligned} \quad (44)$$

We can prove the first statement of the lemma by induction on l . For $l = 0$, it is given by Lemma 4. Assuming that the bounds holds for $l' < l < \beta - 2$, the previous inequality ensures that

$$|h^{(l)}(s)| \leq c \left(1 + \sum_{j=0}^{l-1} |h^{(j)}(s)| \right)$$

since $\sum_{k=0}^{\infty} k^l \sum_{j=k+1}^{\infty} j^{-\beta} < \infty$. This ends up the induction and proves the first estimate in (43).

We consider now $l = \lceil \beta \rceil - 2$ and 'continue the induction'. Using the bound of $h^{(l)}$ for $l < \beta - 2$ and (44) 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 (32).

Finally, we prove the bound for $l = \lceil \beta \rceil - 1$ in the same way. By (44):

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

and Lemma 32 allows us to conclude. \square

7.4 Successive Differentiation for composition of functions

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

Lemma 6. *Let f and h be real-valued, l -times differentiable functions. Then*

$$\frac{d^l}{ds^l} h(f(s)) = \sum_{j=1}^l h^{(j)}(f(s)) u_{j,l}(s), \quad (45)$$

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

$$u_{j,l}(s) = \sum_{i=(i_1, \dots, i_{2j}) \in \mathcal{C}(j,l)} c_i (f^{(i_1)}(s))^{i_2} \dots (f^{(i_{2j-1})}(s))^{i_{2j}}, \quad (46)$$

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

$$\mathcal{C}(j, l) := \{(i_1, \dots, i_{2j}) \in \mathbb{N}^{2j} \mid i_1 i_2 + i_3 i_4 + \dots = l \text{ and } i_2 + i_4 + \dots = j\}.$$

Proof. We prove the formula by induction with respect to l . For $l = 1$, by chain rule of differentiation, (45) is fulfilled. Assume that (45) and (46) hold for l . Then by product rule for differentiation,

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

Now

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

with new constants given by

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

Furthermore,

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

with some new constants $0 \leq \hat{c}_i < \infty$. This ends up the induction. \square

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. A limit theorem for a critical branching process in random environment. *Discrete Math. Appl.* **5** (1993) 45-58.
- [3] V.I. Afanasyev and J. Geiger and G. Kersting and V.A. Vatutin. Functional limit theorems for strongly subcritical branching processes in random environment. *Stochastic Process. Appl.* **115** (2005) 1658-1676.
- [4] K.B. Athreya and S. Karlin. On branching processes with random environments: I, II. *Ann. Math. Stat.* **42** (1971) 1499-1520, 1843-1858.
- [5] K.B. Athreya and P.E. Ney. *Branching Processes*. Dover Publications, INC, New York (2004), *unabridged republication 1972*
- [6] V.I. Afanasyev and J. Geiger and G. Kersting and V.A. Vatutin. Criticality for branching processes in random environment. *Ann. Probab.* **33** (2005) 645-673.
- [7] V. Bansaye and J. Berestycki. Large deviations for Branching Processes in Random Environment. *Markov Processes related Fields* **15**. 493-524. (2009).
- [8] J. D. Biggins, N. H. Bingham (1993). Large deviations in the supercritical branching process. *Adv. in Appl. Probab.* 25, no. 4, 757-772.
- [9] N. H. Bingham, C. M. Goldie, J. L. Teugels. *Regular Variation*. Cambridge University Press Cambridge (1987).
- [10] M. Birkner and J. Geiger and G. Kersting. *Branching processes in random environment - a view on critical and subcritical cases*. Springer. Berlin (2005) 265-291.
- [11] C. Böinghoff and G. Kersting. On large deviations of branching in a random environment - Offspring distributions having at most geometric tails. To appear in *Stoch. Proc. Appl.* (2009).
- [12] A. Dembo and O. Zeitoni. *Large Deviations Techniques and Applications*. Jones and Barlett Publishers International. London (1993).
- [13] J. Geiger and G. Kersting. The survival probability of a critical branching process in a random environment. *Theory Probab. Appl.* **45** (2000) 517-525.

- [14] J. Geiger and 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.
- [15] W. Feller. An Introduction to Probability Theory and Its Applications- Volume I. John Wiley & Sons, Inc. New York (1968). 3. edition.
- [16] W. Feller. An Introduction to Probability Theory and Its Applications- Volume II. John Wiley & Sons, Inc. New York (1966). 1. edition.
- [17] Y. Guivarc'h, Q. Liu (2001). Asymptotic properties of branching processes in random environment. *C.R. Acad. Sci. Paris*, t.332, Serie I. 339-344.
- [18] F. den Hollander. Large Deviations. American Mathematical Society. Providence, RI (2000).
- [19] O. Kallenberg. Foundations of Modern Probability. Springer. London (2001), 2. edition.
- [20] 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.
- [21] M. V. Kozlov. On large deviations of branching processes in a random environment: geometric distribution of descendants. *Discrete Math. Appl.* **16** (2006) 155-174.
- [22] A. Rouault (2000). 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, 15–38.
- [23] W. L. Smith and W.E. Wilkinson. On branching processes in random environments. *Ann. Math. Stat.* **40** (1969) 814-824.
- [24] Daniel Tokarev. Galton-Watson Processes and Extinction in Population Systems. *PhD-Thesis*, Monash University (2007).