Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 120 (2010) 2064–2077
stochastic processes and their applications
www.elsevier.com/locate/spa

Upper large deviations of branching processes in a random environment—Offspring distributions with geometrically bounded tails

Christian Böinghoff*, Götz Kersting

Fachbereich Informatik und Mathematik, Universität Frankfurt, Fach 187, D-60054 Frankfurt am Main, Germany

Received 30 June 2009; received in revised form 17 February 2010; accepted 24 May 2010
Available online 1 June 2010

Abstract

We generalize a result by Kozlov on large deviations of branching processes (Z_n) in an i.i.d. random environment. Under the assumption that the offspring distributions have geometrically bounded tails and mild regularity of the associated random walk S, the asymptotics of $\mathbf{P}(Z_n \geq e^{\theta n})$ is (on logarithmic scale) completely determined by a convex function Γ depending on properties of S. In many cases Γ is identical with the rate function of (S_n). However, if the branching process is strongly subcritical, there is a phase transition and the asymptotics of $\mathbf{P}(Z_n \geq e^{\theta n})$ and $\mathbf{P}(S_n \geq \theta n)$ differ for small θ.

Keywords: Branching processes; Random environment; Large deviations; Phase transition

1. Introduction and main result

Branching processes in a random environment have been introduced in [16,5]. Initially, they have mainly been studied under the assumption of i.i.d. offspring distributions which are geometric, or more generally, linear fractional (see [14,1]). In recent years, the case of general offspring distributions has attracted attention (compare [7,3,4,11]). In this paper, we focus on upper large deviation probabilities which have been studied in [15] for geometric offspring distributions and in [6] for general offspring distributions. In the latter paper, only a lower bound

* Corresponding author. Tel.: +49 69 798 23409; fax: +49 69 798 28841.
E-mail addresses: boeinghoff@math.uni-frankfurt.de (C. Böinghoff), kersting@math.uni-frankfurt.de (G. Kersting).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2010.05.017

for the (upper) large deviation probabilities is given which is improved here. In the particular case of geometric offspring distributions, direct calculations of generating functions are feasible and the asymptotic of the (upper) large deviation probabilities (including lower order terms) is obtained in [15]. Here, we generalize the rate function found in [15] for offspring distributions that have *geometrically bounded tails* and explain in detail the second order phase transition touched in [15].

Some related results for special cases (e.g. branching processes without extinction) can also be found in [6].

Let us recall the definition of a branching process in an i.i.d. random environment. In each generation, an offspring distribution is chosen at random, independently from one generation to the other. Thus, let Δ be the space of all probability measures on $\mathbb{N}_0$. Equipped with the metric of total variation, Δ is a Polish space. Let Q be a random variable taking values in Δ. An infinite sequence $\Pi = (Q_1, Q_2, \ldots)$ of i.i.d. copies of Q is called a *random environment* and Q_n the offspring distribution in generation $n-1$.

Definition. Let Π be a random environment. Then a process $Z = (Z_0, Z_1, \ldots)$ with values in $\mathbb{N}_0$ is called a *branching process in a random environment* Π, if Z_0 is independent of Π and if, given Π, Z is a Markov chain and for every $n \geq 1$, $z \in \mathbb{N}_0$, $q_1, q_2, \ldots \in \Delta$

$$\mathcal{L}\left(Z_n | Z_{n-1} = z, \Pi = (q_1, q_2, \ldots)\right) = \mathcal{L}(\xi_1 + \cdots + \xi_z) \tag{1}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. random variables with distribution q_n.

Z_n is called the nth generation size. For convenience, we assume $Z_0 = 1$ which means we start with only one individual in the initial generation. We denote the corresponding probability measure on the underlying probability space by $\mathbf{P}$. All results in this paper are under the measure $\mathbf{P}$, which is called the annealed approach.

A first large deviation statement is readily obtained: The limit

$$\gamma = \lim_{n\to\infty} -\frac{1}{n} \ln \mathbf{P}(Z_n > 0) \tag{2}$$

exists and $0 \leq \gamma < \infty$ (except the degenerated case $Z_1 = 0$ which we exclude in the following). Moreover

$$\mathbf{P}(Z_n > 0) \leq \mathrm{e}^{-\gamma n} \tag{3}$$

for all n. This follows from

$$\mathbf{P}(Z_{n+m} > 0) \geq \mathbf{P}(Z_n > 0)\, \mathbf{P}(Z_m > 0).$$

Thus the sequence $(-\ln \mathbf{P}(Z_n > 0))_n$ is subadditive, and the claims follow from properties of non-negative subadditive sequences (see [8]).

Fine properties of Z are mainly determined by an auxiliary process, called associated random walk. It depends on the mean offspring number in each generation.

Definition. Setting

$$X_n = \ln \sum_{y=0}^{\infty} y\, Q_n(\{y\}), \quad n \geq 1$$

the random walk $S = (S_0, S_1, \ldots)$ with initial state $S_0 = 0$ and increments $X_n = S_n - S_{n-1},\ n \geq 1$ is called *associated random walk* for the process $(Z_n)_{n\in\mathbb{N}_0}$.

Notice that the X_n are i.i.d. copies of the logarithmic mean offspring number

$$X = \ln \sum_{y=0}^{\infty} y \; Q(\{y\})$$

which we assume finite a.s. Due to $Z_0 = 1$, we get for the conditioned means of Z_n

$$\mu_n = \mathbf{E}[Z_n | \Pi] = \mathrm{e}^{S_n} \quad \text{a.s.} \tag{4}$$

As we will show in this paper, the asymptotics of the upper large deviation probabilities of $(Z_n)_{n\in\mathbb{N}_0}$ is, on a logarithmic scale, completely determined by a convex function related to the distribution of X resp. to properties of the associated random walk.

Here we require the rate function $\Lambda : \mathbb{R} \to \bar{\mathbb{R}}^+$ which fulfills for all $\theta \in \mathbb{R}$

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(S_n \geq \theta n) \leq -\inf_{y\geq\theta} \Lambda(y) = \Lambda(\theta)$$
$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(S_n > \theta n) \geq -\inf_{y>\theta} \Lambda(y) = \Lambda(\theta+)$$

and

$$\mathbf{P}(S_n \geq \theta n) \leq \mathrm{e}^{-\Lambda(\theta)n}. \tag{5}$$

Λ is a convex, nondecreasing function, given by

$$\Lambda(\theta) = \sup_{s\geq 0} \{s\theta - \ln \varphi(s)\}$$

with $\varphi(s) = \mathbf{E}[\mathrm{e}^{sX}]$. 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. [9,8]).

Remark. Usually, Λ is defined as the Legendre transform of $\ln \varphi$ and the supremum is taken over all $s \in \mathbb{R}$. As we are here only interested in upper deviations, it is convenient to set $\Lambda(\theta) = 0$ for all $\theta \leq \mathbf{E}[X]$. $\square$

From γ and Λ we obtain another convex function $\Gamma(\theta)$, $\theta \geq 0$ that turns out to determine the asymptotics of upper large deviations of Z.

Let $\zeta \geq 0$. Γ_ζ is defined as the largest convex function fulfilling

$$\Gamma_\zeta(0) \leq \zeta, \qquad \Gamma_\zeta(\theta) \leq \Lambda(\theta)$$

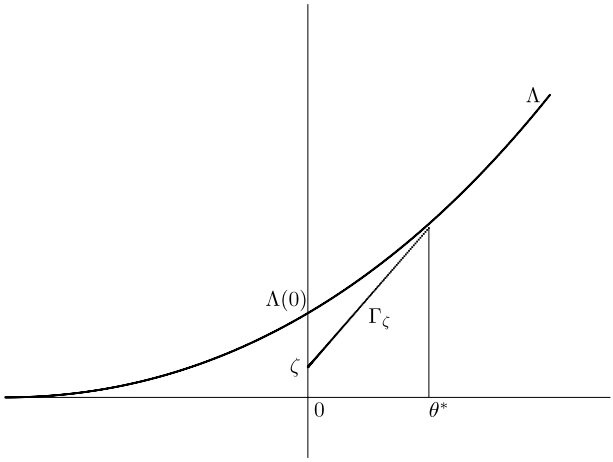
for all $\theta \geq 0$ and set $\Gamma = \Gamma_\gamma$. It is not difficult to see that this function is given by

$$\Gamma_\zeta(\theta) = \begin{cases} \zeta\left(1 - \dfrac{\theta}{\theta^*}\right) + \dfrac{\theta}{\theta^*}\Lambda(\theta^*), & \text{if } \theta < \theta^* \\ \Lambda(\theta), & \text{else} \end{cases}$$

where $0 \leq \theta^* \leq \infty$ is such that

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

The following picture gives Γ_ζ in the case $\zeta < \Lambda(0)$.



If $\zeta \geq \Lambda(0)$ then $\theta^* = 0$ and $\Gamma_\zeta = \Lambda$.

For our result we need two assumptions.

Assumption 1. There is a $s > 0$ such that the moment generating function

$$\varphi(s) = \mathbf{E}\left[\mathrm{e}^{sX}\right] < \infty.$$

In particular $\mathbf{E}[X] \geq -\infty$ exists.

The second assumption concerns the tails of the offspring distributions.

Assumption 2. There are constants $k_0 \in \mathbb{N}_0$, $0 \leq a < b$ and $c > 0$ such that Q a.s. takes values in the set of all probability distributions $q \in \Delta$ with the following property:

If ξ has distribution q and expectation $\mathbb{E}[\xi] = m$, then

$$\mathbb{E}\left[(\xi - j)^+\right] \leq cm\left(\frac{a+m}{b+m}\right)^{j-k_0}, \quad j \geq k_0. \tag{6}$$

Note that $\mathbb{E}[(\xi - j)^+]$ is decreasing in j. We require that this takes place at a geometric rate, where the rate may slow down as $\mathbb{E}[\xi]$ gets larger. The linear factor m excludes degenerated cases where 0 carries most of the mass. Essentially the assumption means that ξ has geometrically bounded tails. This is seen from the following examples:

Examples. • Geometric distributions with success probability p and expectation $m = \frac{1-p}{p}$.

Then $\mathbb{P}(\xi \geq i) = (1-p)^i = \left(\frac{m}{m+1}\right)^i$ and

$$\mathbb{E}\left[(\xi - j)^+\right] = \sum_{i>j} \mathbb{P}(\xi \geq i) = m\left(\frac{m}{m+1}\right)^j$$

which fulfills (6) for $k_0 = 0$, $a = 0$, $b = 1$ and $c = 1$.

• Distributions fulfilling the following condition:

There are constants $c > 0$ and $d \in (0, 1)$ such that

$$\mathbb{P}(\xi = j) \leq cmd^j, \quad j \geq 0. \tag{7}$$

In contrast to (6), (7) implies that the exponential decay rate is uniformly bounded and thus, the expectations are also uniformly bounded by a constant.

- Distributions with Gaussian tails:
  There are constants $\alpha, c > 0$ such that
  $$\mathbb{P}(\xi = j) \le cm \exp(-\alpha j^2), \quad j \ge 0.$$
- Distributions with support in $\{0, \dots, k_0\}$ trivially fulfill condition (6).

The following theorem, which is our main result, has been obtained by Kozlov in the particular case of geometric offspring distributions.

Theorem 1. *Let Assumptions 1 and 2 be fulfilled. Then for any $\theta \ge 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(Z_n \ge \mathrm{e}^{\theta n}) \le -\Gamma(\theta),$$
$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(Z_n > \mathrm{e}^{\theta n}) \ge -\Gamma(\theta+)$$

where $\Gamma = \Gamma_\gamma$. Note that $\Gamma(0) = \gamma$.

Theorem 1 says that a phase transition occurs if $\gamma < \Lambda(0)$. Then $\theta^* > 0$ and there are different regimes $\theta < \theta^*$ and $\theta \ge \theta^*$. Now under mild assumptions, $\gamma \le \Lambda(0)$ (see [1,7]) and $\gamma = \Lambda(0)$ as well as $\gamma < \Lambda(0)$ occurs. The latter case is known as strongly subcritical (defined by $\mathbf{E}[X\mathrm{e}^X] < 0$), being different in general from just subcriticality (see e.g. [4,7,12]).

To get a feel for Theorem 1 we make the following remarks:

First we mention that the large deviation event $\{Z_n \ge \mathrm{e}^{\theta n}\}$ is essentially realized in an exceptional environment and not by exceptionally big offspring numbers. This would require either exponentially many individuals reproducing in an exceptional manner or one individual having exponentially many offsprings. Both probabilities are (by Assumption 2 for the latter) of lower order than exponential. Next recall that

$$\ln \mathbf{P}(S_n \ge 0) \sim -\Lambda(0)n, \qquad \ln \mathbf{P}(Z_n \ge 1) \sim -\gamma n.$$

In the case $\gamma = \Lambda(0)$, the events $\{S_n \ge 0\}$ and $\{Z_n > 0\}$ have the same exponential decay rate. Thus it is natural to expect that the events $\{S_n \ge \theta n\}$ and $\{Z_n \ge \mathrm{e}^{\theta n}\}$ essentially coincide. From our theorem we see that in the case of $\gamma < \Lambda(0)$ this is also true if $\theta \ge \theta^*$. For $\theta < \theta^*$, however, matters change. There we also have to consider the events $\{Z_{\lfloor \lambda n \rfloor} \ge 1, S_n - S_{\lfloor \lambda n \rfloor} \ge \theta n\}$ with $0 < \lambda < 1$, which in view of (2) have exponentially small probability as well. Surprisingly, for λ properly chosen, this event has exponentially larger probability than $\{S_n \ge \theta n\}$. Thus it is of advantage to keep the population just alive at the beginning and to enforce exponential growth only later. For the environment, this means that S first decreases linearly up to time $\lfloor \lambda n \rfloor$ and then increases linearly. For further details we refer to the proof in Section 3.

To better understand the meaning of Γ, we note that Γ_0 determines the large deviations of $S_n - M_n$, where

$$M_n = \min\{S_0, \dots, S_n\}.$$

That is, under Assumption 1 and for all $\theta \ge 0$

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}\left(S_n - M_n \ge \theta n\right) \le -\Gamma_0(\theta),$$
$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}\left(S_n - M_n > \theta n\right) \ge -\Gamma_0(\theta+).$$

This result immediately follows from the proof of Theorem 1 (see Section 3): As we shall see, for $\{Z_n \geq \mathrm{e}^{\theta n}\}$, the occurrence of $\{S_n \geq \theta n\}$ is not required, but essentially only that of $\{S_n - M_n \geq \theta n\}$ and one has to enforce survival of Z until M_n is attained. The latter event has exponentially small probability, represented by γ.

For related results see [9], Section IX.9 (in the context of polymers) or [6] (for the lower deviations of supercritical branching processes in a random environment without extinction). In the latter publication, it has been proved that the most probable way to observe an exceptionally small population number consists of a period with bounded population size followed by a period of geometric growth (corresponding to S growing linearly).

The paper is organized as follows. In the next section, we discuss two characteristics of the distribution of Z_n, which are of interest on their own. Our main theorem is proved in Section 3.

2. Two characteristics of Z

As we will see in this section, Assumption 2 assures that certain characteristics of the distribution of Z_n, known for the case when the generating functions of the offspring distributions are of linear fractional form, are useful also in the more general case treated here. We derive bounds for the normalized variance and tail probabilities in terms of the associated random walk. Let

$$U_n = \mathrm{e}^{-S_n}$$
$$V_n = \sum_{k=0}^{n} \mathrm{e}^{-S_k}.$$

Then $U_n = \mathbf{E}[Z_n|\Pi]^{-1}$. The following two results shed some light on the significance of V_n.

Proposition 1. *Under Assumption 2, there is an $\alpha < \infty$ such that*

$$\frac{\mathbf{E}[Z_n^2|\Pi]}{\mathbf{E}[Z_n|\Pi]^2} \leq \alpha V_n \quad a.s.$$

For the tail probabilities, the following estimate holds:

Theorem 2. *Under Assumption 2, there is a $\beta > 0$ such that for all $z \geq 0$:*

$$\mathbf{P}(Z_n \geq z|\Pi) \leq 2\exp\left(-\frac{\beta U_n}{V_n} z\right) \quad a.s.$$

or likewise

$$\mathbf{P}\left(\frac{Z_n}{\mathbf{E}[Z_n|\Pi]} \geq z \middle| \Pi\right) \leq 2\exp\left(-\frac{\beta}{V_n}\, z\right) \quad a.s.$$

For the proof, we use a formula derived in [11]. Let

$$f_n(s) = \sum_{k=0}^{\infty} s^k Q_n(\{k\})$$

be the probability generating function of the offspring distribution of an individual in generation $n-1$. Note that $X_n = \ln f_n'(1)$. Then, as is well known

$$\mathbf{E}[s^{Z_n}|\Pi] = f_1(f_2(\cdots f_n(s)\cdots)) = f_{0,n}(s), \quad s \geq 0. \tag{8}$$

We like to use an alternative expression for this generating function. Let

$$f_{k,n} = f_{k+1} \circ f_{k+2} \circ \cdots \circ f_n, \quad 0 \le k < n; \qquad f_{n,n} = \mathrm{id}$$
$$g_k(s) = \frac{1}{1 - f_k(s)} - \frac{1}{f_k'(1)(1-s)}, \quad s \ge 0. \tag{9}$$

Below, we shall see that the singularity of $g_k(s)$ at $s = 1$ is removable under Assumption 2. By chain rule for differentiation (recall $f_k(1) = 1$),

$$U_k = \left(f_1'(1) \cdots f_k'(1)\right)^{-1} = f_{0,k}'(1)^{-1}.$$

By a telescope summation argument,

$$\begin{aligned}
\frac{1}{1 - f_{0,n}(s)} &= \frac{U_0}{1 - f_{0,n}(s)} \\
&= \frac{U_n}{1 - f_{n,n}(s)} + \sum_{k=0}^{n-1} \left(\frac{U_k}{1 - f_{k,n}(s)} - \frac{U_{k+1}}{1 - f_{k+1,n}(s)} \right) \\
&= \frac{U_n}{1-s} + \sum_{k=0}^{n-1} U_k \left(\frac{1}{1 - f_{k+1}(f_{k+1,n}(s))} - \frac{1}{f_{k+1}'(1)(1 - f_{k+1,n}(s))} \right) \\
&= \frac{U_n}{1-s} + \sum_{k=0}^{n-1} U_k g_{k+1}(f_{k+1,n}(s)), \quad s \ge 0.
\end{aligned} \tag{10}$$

Note that (10) does not only hold for those s in the domain of convergence of $f_{0,n}(s)$, but for all $s \ge 0$. The following lemma is similar to a statement that has been proved in [11] for every $s \in [0, 1)$:

Lemma 1. *Let $f(s) = \sum_{k=0}^{\infty} s^k p_k$ be the generating function of ξ with distribution $q = (p_k)$ and expectation $\mathbb{E}[\xi] = m$. Then under Assumption 2, the function*

$$g(s) = \frac{1}{1 - f(s)} - \frac{1}{m(1-s)}$$

is continuous for $s \ge 0$ (that means the singularity in $s = 1$ is removable), and there is a number $d < \infty$ such that for all $s \ge 0$

$$0 \le g(s) \le d\left(\frac{1}{m} + 1\right).$$

In particular $\mathbb{E}[\xi(\xi - 1)] \le d(m + m^2)$.

Proof. Defining $r_j = \sum_{k>j} p_k$, we rewrite $f(s) - 1$ to extract the factor $s - 1$ (see [10], Chapter XI):

$$f(s) - 1 = (s-1) \sum_{j=0}^{\infty} s^j r_j = (s-1)h(s),$$

where

$$h(s) = \sum_{j=0}^{\infty} s^j r_j. \tag{11}$$

In the same manner

$$\begin{aligned} h(s) - h(1) &= (s-1)\sum_{j=0}^{\infty} s^j (r_{j+1} + r_{j+2} + \cdots) \\ &= (s-1)k(s), \end{aligned}$$

with

$$k(s) = \sum_{j=0}^{\infty} s^j (r_{j+1} + r_{j+2} + \cdots).$$

Note that by (11), $h(1) = m$ and thus

$$\begin{aligned} \frac{1}{1-f(s)} - \frac{1}{m(1-s)} &= \frac{1}{m(s-1)} - \frac{1}{(s-1)h(s)} \\ &= \frac{k(s)}{mh(s)}. \end{aligned}$$

Since $\mathbb{E}[(\xi - j)^+] = r_{j+1} + r_{j+2} + \cdots$, by Assumption 2 the functions $f(s)$, $h(s)$ and $k(s)$ are finite for $s < \frac{b+m}{a+m}$. Therefore g is continuous in $s = 1$ and thus everywhere.

Next let $1 \le s < \frac{1}{2} + \frac{b+m}{2(a+m)}$, thus $s \le \frac{b}{a}$. As h is nondecreasing, $h(s) \ge m$ for all $s \ge 1$. Thus (also recall $m = r_0 + r_1 + \cdots$),

$$\begin{aligned} \frac{1}{1-f(s)} - \frac{1}{m(1-s)} &\le \frac{k(s)}{m^2} \\ &= \frac{1}{m}\sum_{j=0}^{\infty} s^j \frac{r_{j+1} + r_{j+2} + \cdots}{r_0 + r_1 + \cdots} \\ &\le \frac{1}{m}\left(\sum_{j=0}^{k_0-1} s^j + c s^{k_0} \sum_{j=0}^{\infty} s^j \left(\frac{a+m}{b+m}\right)^j\right) \\ &\le \frac{k_0 \left(\frac{b}{a}\right)^{k_0}}{m} + \frac{2\left(\frac{b}{a}\right)^{k_0} c(b+m)}{(b-a)m}, \end{aligned} \tag{12}$$

where we used (6) in the prelast step. The last inequality follows from $s \le \frac{b}{a}$ in the first sum and $s < \frac{1}{2}\frac{b+m}{2(a+m)}$ in the second geometric summation.

For $s \ge \frac{1}{2} + \frac{b+m}{2(a+m)}$ (and therefore $s \ge 1$), we simply drop the negative term and then use $s \ge \frac{1}{2} + \frac{b+m}{2(a+m)}$:

$$\begin{aligned} -\frac{1}{f(s)-1} + \frac{1}{m(s-1)} &\le \frac{1}{m(s-1)} \\ &\le \frac{2(a+m)}{m(b-a)}. \end{aligned} \tag{13}$$

Note that the last estimate does not require the assumption $f(s) < \infty$. Choosing a sufficiently large constant d, (12) and (13) prove the claim for every $s \ge 1$. For $s < 1$, we use that $g(s) \le 2g(1)$ for $0 \le s \le 1$ (see [11], Lemma 2.1 therein). The last claim follows since $g(1) = m^{-2}\mathbb{E}[\xi(\xi-1)]$. $\square$

Proof of Proposition 1. For the proof of the proposition, we need an expression for $f''_{0,n}(1)$, which is well-known (see e.g. [2]). From $f_{0,n} = f_{0,n-1} \circ f_n$, by chain rule for differentiation $f'_{0,n}(1) = f'_{0,n}(1) f'_n(1)$ and $f''_{0,n}(1) = f''_{0,n-1}(1)(f'_n(1))^2 + f'_{0,n-1}(1) f''_n(1)$, thus

$$\frac{f''_{0,n}(1)}{(f'_{0,n}(1))^2} = \frac{f''_{0,n-1}(1)}{(f'_{0,n-1}(1))^2} + \frac{f''_n(1)}{f'_{0,n-1}(1)(f'_n(1))^2}.$$

Recalling $U_k = \mathrm{e}^{-S_k}$ and $V_n = \sum_{k=0}^n \mathrm{e}^{-S_k}$, the last estimate in Lemma 1 implies

$$\frac{f''_n(1)}{f'_{0,n-1}(1)(f'_n(1))^2} \le d(U_{n-1} + U_n)$$

and thereby

$$\frac{\mathbf{E}[Z_n(Z_n-1)|\Pi]}{\mathbf{E}[Z_n|\Pi]^2} = \frac{f''_{0,n}(1)}{(f'_{0,n}(1))^2} \le 2dV_n \quad \text{a.s.}$$

Now $V_n \ge U_n = (\mathbf{E}[Z_n|\Pi])^{-1}$, thus choosing $\alpha = 2d + 1$ yields the claim. □

Proof of Theorem 2. We obtain from (10) and Lemma 1

$$\frac{1}{1 - f_{0,n}(s)} \le \frac{U_n}{1-s} + d\sum_{k=0}^{n-1}(U_k + U_{k+1}) \le \frac{U_n}{1-s} + 2dV_n. \tag{14}$$

As $f_{0,n}(s) > 1$ for $s > 1$, we only have a useful bound if the right-hand side of (14) is negative. Thus for $s < 1 + \frac{U_n}{2dV_n}$, we get

$$f_{0,n}(s) \le \frac{U_n - (2dV_n - 1)(s-1)}{U_n - 2dV_n(s-1)}. \tag{15}$$

For $s \le 1 + \frac{U_n}{4dV_n}$, since the right-hand side in (15) is nondecreasing for $s \ge 1$, and as $V_n \ge 1$,

$$f_{0,n}(s) \le 1 + \frac{1}{2d}.$$

Therefore (without loss of generality $d \ge \frac{1}{2}$), for every $1 \le s \le 1 + \frac{U_n}{4dV_n}$,

$$\mathbf{P}(Z_n \ge z|\Pi) \le s^{-z} f_{0,n}(s) \le 2s^{-z} \quad \text{a.s.} \tag{16}$$

Note that $\frac{U_n}{V_n} \le 1$, so there is a $\beta > 0$ such that $\mathrm{e}^{\beta\frac{U_n}{V_n}} \le 1 + \frac{U_n}{4dV_n}$ (β is defined by $\mathrm{e}^{\beta} = 1 + \frac{1}{4d}$). Taking $s = \mathrm{e}^{\beta\frac{U_n}{V_n}}$ in (16) yields Theorem 2. □

3. Proof of Theorem 1

We shall use the following representation of Γ.

Lemma 2. *For any* $\theta \ge 0$

$$\Gamma(\theta) = \inf_{0<\lambda\le 1}\{(1-\lambda)\gamma + \lambda\Lambda(\theta/\lambda)\}.$$

Proof. Let us denote this infimum by $\iota(\theta)$. We show that it fulfills the properties defining $\Gamma(\theta)$. First for any $\theta', \theta'' \ge 0, t \in (0,1)$ and $\epsilon > 0$ there are $\lambda', \lambda'' \in (0,1]$ such that in view of

convexity of Λ

$$\begin{aligned}
&t\iota(\theta') + (1-t)\iota(\theta'') \\
&\quad \geq t(1-\lambda')\gamma + t\lambda'\Lambda(\theta'/\lambda') + (1-t)(1-\lambda'')\gamma + (1-t)\lambda''\Lambda(\theta''/\lambda'') - \epsilon \\
&\quad = \left(t\lambda' + (1-t)\lambda''\right)\gamma + (t\lambda' + (1-t)\lambda'')\frac{t\lambda'}{t\lambda' + (1-t)\lambda''}\Lambda(\theta'/\lambda') \\
&\qquad + (t\lambda' + (1-t)\lambda'')\frac{(1-t)\lambda''}{t\lambda' + (1-t)\lambda''}\Lambda(\theta''/\lambda'') - \epsilon \\
&\quad \geq \left(1 - (t\lambda' + (1-t)\lambda'')\right)\gamma + (t\lambda' + (1-t)\lambda'')\Lambda\left(\frac{t\theta' + (1-t)\theta''}{t\lambda' + (1-t)\lambda''}\right) - \epsilon \\
&\quad \geq \iota(t\theta' + (1-t)\theta'') - \epsilon.
\end{aligned}$$

Letting $\epsilon \to 0$ gives the convexity of ι. Next choosing $\lambda = 1$ implies $\iota(\theta) \leq \Lambda(\theta)$ and letting $\lambda \to 0$ entails $\iota(0) \leq \gamma$. Finally let $\kappa(\theta)$ be any convex function below $\Gamma(\theta)$ and γ. Then for any $\lambda \in (0, 1]$, $\theta \geq 0$

$$\begin{aligned}
(1-\lambda)\gamma + \lambda\Lambda(\theta/\lambda) &\geq (1-\lambda)\kappa(0) + \lambda\kappa(\theta/\lambda) \\
&\geq \kappa\left((1-\lambda)0 + \lambda(\theta/\lambda)\right) \\
&= \kappa(\theta).
\end{aligned}$$

It follows $\iota(\theta) \geq \kappa(\theta)$, and the proof is complete. □

3.1. The upper bound

We follow ideas of Kozlov.

Lemma 3. *Under Assumptions 1 and 2, for any $\theta \geq 0$:*

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(Z_n \geq \mathrm{e}^{\theta n}) \leq -\inf_{y \geq \theta} \Gamma(y).$$

Proof. We restrict ourselves to $\theta > 0$ as the case $\theta = 0$ is covered by (2). Let

$$M_n = \min_{0 \leq k \leq n} S_k$$

and $\epsilon > 0$ such that $\theta - \epsilon > 0$. We have

$$\begin{aligned}
\mathbf{P}(Z_n \geq \mathrm{e}^{\theta n}) &= \mathbf{P}(Z_n \geq \mathrm{e}^{\theta n}, S_n - M_n \geq (\theta - \epsilon)n) + \mathbf{P}(Z_n \geq \mathrm{e}^{\theta n}, S_n - M_n < (\theta - \epsilon)n) \\
&= p_{1n} + p_{2n} \quad \text{(say)}.
\end{aligned}$$

Now, as Z_k is independent of $S_n - S_k$ and by (3) and (5),

$$\begin{aligned}
&\mathbf{P}(Z_n \geq \mathrm{e}^{\theta n}, S_n - M_n \geq (\theta - \epsilon)n) \leq \sum_{k=0}^{n-1} \mathbf{P}(Z_k > 0, S_n - S_k \geq (\theta - \epsilon)n) \\
&\quad \leq \sum_{k=0}^{n-1} \mathrm{e}^{-\gamma k} \mathrm{e}^{-\Lambda\left((\theta-\epsilon)\frac{n}{n-k}\right)\cdot(n-k)} \\
&\quad = \sum_{k=0}^{n-1} \exp\left(-n\left(\gamma\frac{k}{n} + \Lambda\left((\theta-\epsilon)\frac{n}{n-k}\right)\frac{n-k}{n}\right)\right).
\end{aligned}$$

In view of Lemma 2

$$p_{1n} \le n\mathrm{e}^{-\Gamma(\theta-\epsilon)n}.$$

As to p_{2n}, by means of Theorem 2

$$\mathbf{P}(Z_n \ge \mathrm{e}^{\theta n}|\Pi) \le 2\exp\left(-\frac{\beta U_n}{V_n}\mathrm{e}^{\theta n}\right).$$

Now $V_n \le (n+1)\mathrm{e}^{-M_n}$, thus

$$\begin{aligned} p_{2n} &\le \mathbf{E}\left[2\exp\left(-\frac{\beta U_n}{V_n}\mathrm{e}^{\theta n}\right); S_n - M_n < (\theta-\epsilon)n\right] \\ &\le \mathbf{E}\left[2\exp\left(-\beta(n+1)^{-1}\mathrm{e}^{\theta n-(S_n-M_n)}\right); S_n - M_n < (\theta-\epsilon)n\right] \\ &\le 2\exp\left(-\beta(n+1)^{-1}\mathrm{e}^{\epsilon n}\right). \end{aligned}$$

By standard arguments from large deviation theory (see [9]), we have

$$\begin{aligned} \limsup_{n\to\infty}\frac{1}{n}\ln\mathbf{P}(Z_n \ge \mathrm{e}^{\theta n}) &\le \limsup_{n\to\infty}\frac{1}{n}\ln\left[p_{1n}+p_{2n}\right] \\ &= \max\left[\limsup_{n\to\infty}\frac{1}{n}\ln p_{1n}, \limsup_{n\to\infty}\frac{1}{n}\ln p_{2n}\right] \\ &\le \max\left[-\Gamma(\theta-\epsilon), -\infty\right]. \end{aligned}$$

As Γ is left-continuous, taking the limit $\epsilon \to 0$ yields the result. □

3.2. The lower bound

Lemma 4. *Under Assumptions 1 and 2, for any $\theta \ge 0$:*

$$\liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}(Z_n > \mathrm{e}^{\theta n}) \ge -\inf_{y>\theta}\Gamma(y).$$

A weaker result has been proved in [6] under different assumptions.

Proof. Without loss of generality, we restrict ourselves to the case $\Gamma(\theta+) < \infty$. For every $0 < \lambda \le 1$, by Markov property,

$$\begin{aligned} \mathbf{P}(Z_n > \mathrm{e}^{\theta n}) &= \mathbf{P}(Z_{\lceil(1-\lambda)n\rceil} > 0)\mathbf{P}(Z_n > \mathrm{e}^{\theta n}|Z_{\lceil(1-\lambda)n\rceil} > 0) \\ &\ge \mathbf{P}(Z_{\lceil(1-\lambda)n\rceil} > 0)\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}). \end{aligned}$$

We fix θ', θ'' with $\theta < \theta' < \theta''$. Then

$$\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}) \ge \mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n).$$

An inequality due to Paley and Zygmund (see e.g. [13], p. 63) gives for any $\mathbb{R}^+$-valued random variable ξ with $0 < \mathbb{E}[\xi] < \infty$ for any $0 < r < 1$

$$\mathbb{P}(\xi > r\mathbb{E}[\xi]) \ge (1-r)^2\frac{\mathbb{E}[\xi]^2}{\mathbb{E}[\xi^2]}.$$

Thus for any $0 < r < 1$

$$\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > r\mathbf{E}[Z_{\lfloor\lambda n\rfloor}|\Pi]|\Pi) \ge (1-r)^2\frac{\mathbf{E}[Z_{\lfloor\lambda n\rfloor}|\Pi]^2}{\mathbf{E}\left[Z^2_{\lfloor\lambda n\rfloor}|\Pi\right]}.$$

From Proposition 1

$$\frac{\mathbf{E}[Z_{\lfloor\lambda n\rfloor}|\Pi]^2}{\mathbf{E}\left[Z^2_{\lfloor\lambda n\rfloor}|\Pi\right]} \geq \frac{1}{\alpha V_{\lfloor\lambda n\rfloor}} \geq \frac{\mathrm{e}^{M_{\lfloor\lambda n\rfloor}}}{\alpha(n+1)}.$$

Thus with $r = \mathrm{e}^{-(\theta'-\theta)} \geq \mathrm{e}^{-(\theta'-\theta)n}$,

$$\begin{aligned}\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n) &= \mathbf{E}\left[\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}|\Pi); \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n\right] \\ &\geq (\alpha(n+1))^{-1}(1-r)^2\mathbf{E}[\mathrm{e}^{M_{\lfloor\lambda n\rfloor}}; \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n] \\ &\geq (\alpha(n+1))^{-1}(1-r)^2\mathbf{P}(M_{\lfloor\lambda n\rfloor} \geq 0, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n).\end{aligned}$$

Let $\tilde{\theta} = \frac{1}{2\lambda}(\theta' + \theta'')$. First we assume $\varphi(s) < \infty$ for all $s \in \mathbb{R}^+$. Then $\Lambda(\tilde{\theta}) = \tau\tilde{\theta} - \ln\varphi(\tau)$ defines τ (see e.g. [8] for properties of the rate function) and we can change measure according to

$$\tilde{\mathbf{E}}[\Phi(S_1, \ldots, S_n)] = \rho^{-n}\mathbf{E}\left[\Phi(S_1, \ldots, S_n)\mathrm{e}^{\tau S_n}\right]$$

where $\rho = \varphi(\tau)$. Thus

$$\begin{aligned}&\mathbf{P}(M_{\lfloor\lambda n\rfloor} \geq 0, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n) \\ &\quad\geq \rho^{\lfloor\lambda n\rfloor}\tilde{\mathbf{E}}\left[\mathrm{e}^{-\tau S_{\lfloor\lambda n\rfloor}}; \theta'\lambda^{-1}\lceil\lambda n\rceil < S_{\lfloor\lambda n\rfloor} < \theta''\lambda^{-1}\lfloor\lambda n\rfloor, M_{\lfloor\lambda n\rfloor} \geq 0\right] \\ &\quad\geq \rho^{\lfloor\lambda n\rfloor}\mathrm{e}^{-\tau\theta''\lambda^{-1}\lfloor\lambda n\rfloor}\tilde{\mathbf{P}}(\theta'\lambda^{-1}\lceil\lambda n\rceil < S_{\lfloor\lambda n\rfloor} < \theta''\lambda^{-1}\lfloor\lambda n\rfloor, M_{\lfloor\lambda n\rfloor} \geq 0).\end{aligned}$$

Now S is under $\tilde{\mathbf{P}}$ a random walk with $\tilde{\mathbf{E}}[S_n] = \tilde{\theta}n$. Therefore $\tilde{\mathbf{P}}(\theta'\lambda^{-1}n < S_n < \theta''\lambda^{-1}n) \to 1$ and $\tilde{\mathbf{P}}(M_n \geq 0) \to \tilde{\mathbf{P}}(M_\infty \geq 0) = p > 0$. Thus for n large enough,

$$\tilde{\mathbf{P}}(\theta'\lambda^{-1}\lceil\lambda n\rceil < S_{\lfloor\lambda n\rfloor} < \theta''\lambda^{-1}\lfloor\lambda n\rfloor, M_{\lfloor\lambda n\rfloor} \geq 0) \geq \frac{p}{2}$$

and therefore, for any $0 < \lambda \leq 1$

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}(Z_n > \mathrm{e}^{\theta n}) &\geq \liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}(Z_{\lceil(1-\lambda)n\rceil} > 0) \\ &\quad + \liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}(Z_{\lfloor\lambda n\rfloor} > \mathrm{e}^{\theta n}, \theta'\lambda^{-1}\lceil\lambda n\rceil < S_{\lfloor\lambda n\rfloor} < \theta''\lambda^{-1}\lfloor\lambda n\rfloor) \\ &= -\Big((1-\lambda)\gamma + \lambda(\theta''\lambda^{-1}\tau - \ln\varphi(\tau))\Big).\end{aligned}$$

Letting $\theta', \theta'' \to \theta$ (as $\Lambda(\theta+) < \infty$, Λ is continuous in θ)

$$\theta''\lambda^{-1}\tau - \ln\varphi(\tau) \to \Lambda(\lambda^{-1}\theta+).$$

For the general case of $\varphi(s) = \infty$, we condition on $\{\max_{i=1,\ldots,\lfloor\lambda n\rfloor} X_i < x\}$. For the conditioned random walk, the moment generating function is finite on $\mathbb{R}^+$ and we can find a $\tilde{\tau}_x$ such that the above calculation holds. Now

$$\begin{aligned}&\mathbf{P}(M_{\lfloor\lambda n\rfloor} \geq 0, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n) \\ &\quad\geq \mathbf{P}\left(M_{\lfloor\lambda n\rfloor} \geq 0, \theta' n < S_{\lfloor\lambda n\rfloor} < \theta'' n\middle|\max_{i=1,\ldots,\lfloor\lambda n\rfloor} X_i < x\right)\mathbf{P}(X < x)^{\lfloor\lambda n\rfloor}\end{aligned}$$

and the moment generating function of the conditioned random walk is $\mathbf{P}(X < x)^{-n}\varphi_x^n(s)$, where $\varphi_x(s) = \mathbf{E}[\mathrm{e}^{sx}; X < x]$. Thus, letting $\theta', \theta'' \to \theta$ (and thereby $\tilde{\tau}_x \to \tau_x$),

$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(Z_n > \mathrm{e}^{\theta n}) \geq -\Big((1-\lambda)\gamma + \lambda(\theta\lambda^{-1}\tau_x - \ln\varphi_x(\tau_x))\Big),$$

where

$$\begin{aligned}\theta\lambda^{-1}\tau_x - \ln\varphi_x(\tau_x) &= \theta\lambda^{-1}\tau_x - \ln\mathbf{E}[\mathrm{e}^{\tau_x X}|X < x] + \ln\mathbf{P}(X < x)\\ &= \sup_{s\geq 0}\Big\{\theta\lambda^{-1}s - \ln\mathbf{E}[\mathrm{e}^{sX}; X < x]\Big\}.\end{aligned}$$

The right-hand side is non-increasing in x. Thus, by monotone convergence, we may interchange the limit $x \to \infty$ and the supremum and letting $x \to \infty$,

$$\begin{aligned}\theta\lambda^{-1}\tau_x - \ln\varphi_x(\tau_x) &\to \sup_{s\geq 0}\Big\{\theta\lambda^{-1}s - \ln\mathbf{E}[\mathrm{e}^{sX}]\Big\}\\ &= \Lambda(\theta\lambda^{-1}).\end{aligned}$$

By Lemma 2,

$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(Z_n > \mathrm{e}^{\theta n}) \geq -\Gamma(\theta+)$$

which entails the result. □

Acknowledgements

We are indebted to Mikhail Kozlov who pointed out a substantial error in an earlier version of this paper and thank the anonymous referees for their useful remarks. The research was supported in part by the German Research Foundation (DFG) and the Russian Foundation of Basic Research, Grant 31120121.

References

[1] V.I. Afanasyev, Limit theorems for a conditional random walk and some applications, Diss. Cand. Sci. Moscow, MSU, 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, J. Geiger, G. Kersting, V.A. Vatutin, Criticality for branching processes in random environment, Ann. Probab. 33 (2005) 645–673.

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

[5] K.B. Athreya, S. Karlin, On branching processes with random environments: I, II, Ann. Math. Statist. 42 (1971) 1499–1520, 1843–1858.

[6] V. Bansaye, J. Berestycki, Large deviations for branching processes in random environment, Preprint, 2008. arXiv:math.PR/0810.4991v1.

[7] M. Birkner, J. Geiger, G. Kersting, Branching Processes in Random Environment—A View on Critical and Subcritical Cases, Springer, Berlin, 2005, pp. 269–291.

[8] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Jones and Barlett Publishers International, London, 1993.

[9] F. den Hollander, Large Deviations, American Mathematical Society, Providence, RI, 2000.

[10] W. Feller, An Introduction to Probability Theory and its Applications—Volume I, 3rd ed., John Wiley & Sons, Inc., New York, 1968.

[11] J. Geiger, G. Kersting, The survival probability of a critical branching process in a random environment, Theory Probab. Appl. 45 (2000) 517–525.
[12] J. Geiger, G. Kersting, V.A. Vatutin, Limit theorems for subcritical branching processes in random environment, Ann. Inst. H. Poincaré Probab. Statist. 39 (2003) 593–620.
[13] O. Kallenberg, Foundations of Modern Probability, 2nd ed., Springer, London, 2001.
[14] 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.
[15] M.V. Kozlov, On large deviations of branching processes in a random environment: geometric distribution of descendants, Discrete Math. Appl. 16 (2006) 155–174.
[16] W.L. Smith, W.E. Wilkinson, On branching processes in random environments, Ann. Math. Statist. 40 (1969) 814–824.