Branching processes with immigration in atypical random environment

Citation for published version:
Foss, S, Korshunov, D & Palmowski, Z 2022, 'Branching processes with immigration in atypical random environment', *Extremes*, vol. 25, pp. 55–77. https://doi.org/10.1007/s10687-021-00427-1

Digital Object Identifier (DOI):
10.1007/s10687-021-00427-1

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Extremes

Extremes
https://doi.org/10.1007/s10687-021-00427-1

Branching processes with immigration in atypical random environment

Sergey Foss[1,2] · **Dmitry Korshunov**[2,3] · **Zbigniew Palmowski**[4]

Received: 20 October 2020 / Revised: 30 March 2021 / Accepted: 6 August 2021 /
Published online: 20 September 2021

Abstract

Motivated by a seminal paper of Kesten et al. (*Ann. Probab.*, *3(1)*, 1–31, 1975) we consider a branching process with a conditional geometric offspring distribution with i.i.d. random environmental parameters A_n, $n \geq 1$ and with one immigrant in each generation. In contrast to above mentioned paper we assume that the environment is long-tailed, that is that the distribution F of $\xi_n := \log((1-A_n)/A_n)$ is long-tailed. We prove that although the offspring distribution is light-tailed, the environment itself can produce extremely heavy tails of the distribution of the population size in the nth generation which becomes even heavier with increase of n. More precisely, we prove that, for all n, the distribution tail $\mathbb{P}(Z_n \geq m)$ of the nth population size Z_n is asymptotically equivalent to $n\overline{F}(\log m)$ as m grows. In this way we generalise Bhattacharya and Palmowski (*Stat. Probab. Lett.*, *154*, 108550, 2019) who proved this result in the case $n = 1$ for regularly varying environment F with parameter $\alpha > 1$. Further, for a subcritical branching process with subexponentially distributed ξ_n, we provide the asymptotics for the distribution tail $\mathbb{P}(Z_n > m)$ which are valid uniformly for all n, and also for the stationary tail distribution. Then we establish the "principle of a single atypical environment" which says that the main cause for the number of particles to be large is the presence of a single very small environmental parameter A_k.

Keywords Branching process · Random environment · Random walk in random environment · Subexponential distribution · Slowly varying distribution

Mathematics Subject Classification (2010) 60J70 · 60G55 · 60J80

The work of S Foss and D Korshunov is partially supported by the RFBR grant 19-51-53010 (2019-2020). The work of Z Palmowski is partially supported by the Polish National Science Centre under the grant 2018/29/B/ST1/00756 (2019-2022).

✉ Zbigniew Palmowski
zbigniew.palmowski@pwr.edu.pl

Extended author information available on the last page of the article.

1 Introduction and main results

Branching processes considered in this paper are motivated by works of (Solomon 1975) and (Kesten et al. 1975), who analysed a neighbourhood random walk in random environment. This is a random walk $(X_t, t \in \mathbb{Z}^+)$ on $\mathbb{Z}$ defined in the following way. Consider a collection $(A_i, i \in \mathbb{Z}^+)$ of i.i.d. (0, 1)-valued random variables. Let $\mathcal{A}$ be the σ-algebra generated by $(A_i, i \in \mathbb{Z}^+)$. Let $(X_k, k \in \mathbb{N})$ be a random walk in random environment, that is a collection of $\mathbb{Z}$-valued random variables such that $X_0 = 0$ and, for $k \geq 0$,

$$\mathbb{P}\big(X_{k+1} = X_k + 1 \mid \mathcal{A}, X_0 = i_0, \dots, X_k = i_k\big) = A_{i_k}$$

and

$$\mathbb{P}\big(X_{k+1} = X_k - 1 \mid \mathcal{A}, X_0 = i_0, \dots, X_k = i_k\big) = 1 - A_{i_k}$$

for all $i_j \in \mathbb{Z}$, $0 \leq j \leq k$. The collection $(A_i, i \in \mathbb{Z}^+)$ is called a random environment.

For this random walk, (Kesten et al. 1975) studied the appropriately scaled limiting distribution of the hitting time $T_n = \inf\{k > 0 : X_k = n\}$ of any state $n \in \mathbb{Z}$. Their analysis is based on the representation of $T_n, n > 0$ in terms of the total number of particles up to the nth generation of a certain branching process in random environment with size-1 immigration at each generation step. In this model the offspring distribution in the nth generation is geometric with a random parameter A_n.

In other words, let $(Z_n, n \geq 0)$ be a branching process in random environment with one immigrant each time that starts from $Z_0 \equiv 0$. Then the following representation holds:

$$Z_{n+1} = \sum_{i=1}^{Z_n+1} B_{n+1,i} \tag{1}$$

where, conditioned on A_n, $(B_{n+1,i}, i \geq 1)$ are independent copies of a geometric random variable B_{n+1} with probability mass function

$$\mathbb{P}(B_{n+1} = k \mid A_n) = A_n(1 - A_n)^k \quad \text{for all } k \geq 0,\ n \geq 0. \tag{2}$$

Denote

$$\xi_n \ := \ \log \frac{1 - A_n}{A_n},$$

so $\xi_n(\omega) > 0$ if and only if $A_n(\omega) < 1/2$, let F be the common distribution of ξ_n.

Following (Kesten et al. 1975), let U_i^n denote the number of transitions of $(X_k, k \geq 0)$ from i to $i - 1$ within time interval $[0, T_n)$, i.e.,

$$U_i^n = Card\{k < T_n : X_k = i, X_{k+1} = i - 1\},$$

where $Card(C)$ is the cardinality of the set C. It is easy to derive that

$$T_n = n + 2 \sum_{i=-\infty}^{\infty} U_i^n. \tag{3}$$

Note that $U_i^n = 0$ for all $i \geq n$ and $U := \sum_{i \leq 0} U_i^n < \infty$ a.s. if $X_k \to \infty$ a.s. as $k \to \infty$. It has been established in (Kesten et al. 1975), that

$$\sum_{i=1}^{n} U_i^n \stackrel{d}{=} \sum_{l=0}^{n-1} Z_l. \tag{4}$$

Then (Kesten et al. 1975) have analysed T_n under the so-called "Kesten assumptions" on the environment:

$$\mathbb{E}\xi < 0 \quad \text{but } \mathbb{E}e^{\xi} \geq 1 \tag{5}$$

and there exists a unique positive solution κ to the equation

$$\mathbb{E}\left(\left(\frac{1-A}{A}\right)^{\kappa}\right) = \mathbb{E}e^{\kappa\xi} = 1. \tag{6}$$

In particular, the assumption Eq. 6 implies that the random variable ξ has an exponentially decaying right tail. It was shown in (Kesten et al. 1975) that, under the assumptions Eq. 5–Eq. 6, the distributions of appropriately scaled random variables T_n and $\sum_{k=0}^{n-1} Z_k$ become close to each other and converge, as $n \to \infty$, to the distribution of a κ-stable random variable.

The tail asymptotics for the branching process Z_n under the assumptions Eq. 5–Eq. 6 were studied by (Dmitruschenkov and Shklyaev 2017) for all three regimes, subcritical, critical, and supercritical.

The aim of our paper is to study the asymptotic behaviour of the branching process Z_n under the complementary assumption that the distribution F of the random variable ξ is *long-tailed*, that is, $\overline{F}(x) > 0$ for all x and

$$\overline{F}(x-y) \sim \overline{F}(x) \quad \text{as } x \to \infty, \tag{7}$$

for some (and therefore for all) fixed y. Here $\overline{F}(x) = 1 - F(x)$ is the tail distribution function and equivalence Eq. 7 means that the ratio of the left- and right-hand sides tends to 1 as x grows, for all y. In particular, Eq. 7 implies that F is *heavy-tailed*, i.e. $\mathbb{E}e^{c\xi} = \infty$ for all $c > 0$. Given Eq. 7, the distribution G defined by its tail as $\overline{G}(x) = \overline{F}(\log x)$, $x \geq 1$, is *slowly varying at infinity* and therefore *subexponential*, that is,

$$\overline{G * G}(x) \sim 2\overline{G}(x) \quad \text{as } x \to \infty, \tag{8}$$

see, e.g. Theorem 3.29 in (Foss et al. 2013).

A distribution F with finite mean is called *strong subexponential* if

$$\int_0^x \overline{F}(x-y)\overline{F}(y)dy \sim 2\overline{F}(x)\int_0^\infty \overline{F}(y)dy \quad \text{as } x \to \infty. \tag{9}$$

Any strong subexponential distribution F is subexponential, and its *integrated tail distribution* F_I with the tail distribution function

$$\overline{F}_I(x) = \min\left(1, \int_x^\infty \overline{F}(y)dy\right).$$

is subexponential too (see e.g. (Foss et al. 2013, Theorem 3.27)). In what follows, we write $F_I(x, y] := \overline{F}_I(x) - \overline{F}_I(y)$.

We start now with our first main result.

Theorem 1.1 *Under the assumption Eq. 7,*

$$\mathbb{P}(Z_1 > m) \sim \overline{F}(\log m) \quad \text{as } m \to \infty.$$

If, in addition, the distribution F is subexponential, then, for any fixed $n \geq 2$,

$$\mathbb{P}(Z_n > m) \sim n\overline{F}(\log m) \quad \text{as } m \to \infty.$$

Theorem 1.1 shows that the tail of Z_1 is surprisingly heavy and is getting heavier in each next generation. What should be underlined, this type of behaviour is a consequence of the environment only, and not of the branching mechanism which is of geometric type. In contrast to a series of papers (Seneta 1973), (Darling 1970), (Schuh and Barbour 1977), (Hong and Zhang 2019), we do not analyse the convergence results for $n \to \infty$, with focusing on the tail behaviour of the distribution of Z_n for each n.

Consider now a branching process with state-independent immigration satisfying the stability condition

$$-a := \mathbb{E}\xi < 0 \quad \text{where } \mathbb{E}|\xi| < \infty. \tag{10}$$

The classical Foster criterion implies that the distribution of Z_n stabilises in time, i.e. the distribution of the Markov chain Z_n converges to a unique limiting/stationary distribution as n grows. It follows from Theorem 1.1 that, for any n, the tail of the stationary distribution must be asymptotically heavier than $n\overline{F}(\log m)$, i.e. $\mathbb{P}(Z > m)/\overline{F}(\log m) \to \infty$ as $m \to \infty$, where Z is sampled from the stationary distribution. The distribution tail asymptotics of Z_n and Z are specified in the following two results. The first result provides two asymptotic lower bounds, for finite and infinite time horizons, where the first bound is uniform for all generations.

Theorem 1.2 *Let the stability condition Eq. 10 hold and*

$$A \leq \widehat{A} \quad \text{a.s. for some constant } \widehat{A} < 1, \tag{11}$$

or, equivalently, ξ be bounded below by $\log(1/\widehat{A} - 1)$. Then the following lower bounds hold.

(i) If the distribution F is long-tailed, then

$$\mathbb{P}(Z_n > m) \geq (a^{-1} + o(1))F_I(\log m, \log m + na] \\ \text{as } m \to \infty \text{ uniformly for all } n \geq 1. \tag{12}$$

(ii) If the integrated tail distribution F_I is long-tailed, then

$$\mathbb{P}(Z > m) \geq (a^{-1} + o(1))\overline{F}_I(\log m) \quad \text{as } m \to \infty. \tag{13}$$

The next result presents conditions for the existence of upper bounds that match the lower bounds of Theorem 1.2.

Theorem 1.3 *Let the stability condition Eq. 10 hold and the distribution F be such that*

$$\overline{F}(m-\sqrt{m}) \sim \overline{F}(m) \quad \textit{and} \quad \overline{F}(m)e^{\sqrt{m}} \to \infty \quad \textit{as } m \to \infty. \tag{14}$$

Then the following upper bounds hold.

(i) If the distribution F is strong subexponential, then

$$\begin{aligned}\mathbb{P}(Z_n > m) \;\le\; (a^{-1}+o(1))F_I(\log m, \log m + na] \\ \textit{as } m \to \infty \textit{ uniformly for all } n \ge 1.\end{aligned} \tag{15}$$

(ii) If the integrated tail distribution F_I is subexponential, then

$$\mathbb{P}(Z > m) \;\le\; (a^{-1}+o(1))\overline{F}_I(\log m) \quad \textit{as } m \to \infty. \tag{16}$$

Distributions satisfying the first condition in Eq. 14 are called *square-root insensitive*, see e.g. (Foss et al. 2013, Sect. 2.8). Typical examples of distributions satisfying Eq. 14 are: any regularly varying distribution, the log-normal distribution and a Weibull distribution with parameter less than $1/2$.

We do not know, how essential is the square-root insensitivity condition for the upper bounds in Theorem 1.3 to hold. In the literature, there are various scenarios where extra randomness leads to appearance of further terms in the tail asymptotics due to the effects caused by the central limit theorem. Namely, for the Weibull distribution $\overline{F}(x) = \exp(-x^{\beta})$ with parameter $\beta \in [1/2, 1)$, the number of extra terms appearing in the tail asymptotics depends on the interval $[n/(n+1), (n+1)/(n+2))$, $n = 1, 2, \ldots$ the parameter β belongs to – see e.g. (Asmussen et al. 1998) and (Foss and Korshunov 2000) for the distributional tail asymptotics of the stationary queue length in a single-server queue or (Denisov et al. To appear) for the tail asymptotics of the stationary distribution in a Markov chain with asymptotically zero drift. However, we are not certain that similar arguments may be relevant to the model considered in the present paper.

If the distribution F satisfies all the conditions of Theorems 1.2 and 1.3, then the corresponding lower and upper bounds match each other and we conclude the following tail asymptotics.

Theorem 1.4 *Let the stability condition Eq. 10 hold, and the distribution F satisfy Eq. 14 and be bounded below in the sense of Theorem 1.2. Then the following tail asymptotics hold.*

(i) If the distribution F is strong subexponential, then

$$\mathbb{P}(Z_n > m) \sim a^{-1}F_I(\log m, \log m + na] \textit{ as } m \to \infty \textit{ uniformly for all } n \ge 1. \tag{17}$$

(ii) If the integrated tail distribution F_I is subexponential, then

$$\mathbb{P}(Z > m) \sim a^{-1}\overline{F}_I(\log m) \quad \textit{as } m \to \infty. \tag{18}$$

These asymptotics may be intuitively interpreted as follows: Z_n takes a large value if one of the ξ's is sufficiently large, i.e. one of the success probabilities A's is small. This phenomenon may be named as *the principle of a single atypical environment* and formulated as follows.

For any $c > 1$ and $\varepsilon > 0$ let us introduce events

$$E_n^{(k)}(m, c, \varepsilon) = \big\{Z_k \le c,\ \xi_k > \log m + (a+\varepsilon)(n-k),$$
$$|S_{j,n-1} - (n-j)\mathbb{E}\xi| \le c + \varepsilon(n-j) \text{ for all } j \in [k+1, n-1]\big\}, \quad k \le n-1,$$

where $S_{j,n} := \xi_j + \ldots + \xi_n$. The event $E_n^{(k)}(m, c, \varepsilon)$ describes all trajectories such that the value of Z_k is relatively small, then the success probability A_k is close to zero and, as a result, a single atypical environment occurs, and after time k the environment follows the strong law of large numbers with drift $-a$. As stated in the next theorem, the union of all these events provides the most probable way for the large deviations of Z_n to occur.

Theorem 1.5 *Assume that conditions of Theorems 1.2 and 1.3 hold. Then, for any fixed* $\varepsilon > 0$*,*

$$\lim_{c\to\infty} \lim_{m\to\infty} \inf_{n\ge 1} \mathbb{P}\left(\bigcup_{k=0}^{n-1} E_n^{(k)}(m, c, \varepsilon) \,\Big|\, Z_n > m\right) = 1. \tag{19}$$

A similar phenomenon has been observed by (Vatutin and Zheng 2012) for the survival probability of a subcritical branching process in random environment without immigration where the increments of the associated random walk have a regularly varying at infinity distribution.

Let us highlight a natural link of branching processes in random environment to stochastic difference equations. It follows from the recurrence equation

$$\begin{aligned}\mathbb{E}(Z_n \mid \mathcal{A},\ Z_{n-1}) &= (Z_{n-1}+1)\mathbb{E}(B_n \mid \mathcal{A}) \\ &= (Z_{n-1}+1)\left(\frac{1}{A_{n-1}} - 1\right) = (Z_{n-1}+1)e^{\xi_{n-1}}\end{aligned}$$

that, for each n, the conditional expectation of Z_n,

$$\mathbb{E}(Z_n \mid \mathcal{A}) = \sum_{k=0}^{n-1} e^{\sum_{l=k}^{n-1}\xi_l} = \sum_{k=0}^{n-1} e^{S_{k,n-1}}, \tag{20}$$

is distributed as a finite time horizon perpetuity, and its limit $\mathbb{E}(Z \mid \mathcal{A})$ as the solution to the stochastic fixed point equation. Their tail asymptotic behaviour in the subexponential case is the same as given in Eq. 17–Eq. 18, that is,

$$\mathbb{P}\big[\mathbb{E}(Z_n \mid \mathcal{A}) > m\big] \sim a^{-1} F_I(\log m, \log m + na]$$
$$\text{as } m \to \infty \text{ uniformly for all } n \ge 1, \tag{21}$$
$$\mathbb{P}\big[\mathbb{E}(Z \mid \mathcal{A}) > m\big] \sim a^{-1}\overline{F}_I(\log m) \quad \text{as } m \to \infty, \tag{22}$$

see (Dyszewski 2016) for Eq. 22 and (Korshunov 2021) for general case.

The remainder of the paper is dedicated to the proofs of the results above. We close our paper by Section 6 which contains some discussion and possible extensions.

2 Finite time horizon tail asymptotics, proof of Theorem 1.1

Let B have, conditionally on A, a geometric distribution with probability mass function

$$\mathbb{P}(B = k \mid A) = A(1 - A)^k \quad \text{for all } k \geq 0.$$

Then,

$$\mathbb{P}(B > m) = \mathbb{E}\big((1 - A)^{m+1}\big). \tag{23}$$

Next, for k conditionally independent copies $B^{(1)}, \ldots, B^{(k)}$ of B, the event $B^{(1)} + \ldots + B^{(k)} > m$ may be described as the number of successes in the first of corresponding $m + k$ Bernoulli trials is smaller than k, which yields the following binomial representation that is convenient for further analysis,

$$\mathbb{P}(B^{(1)} + \ldots + B^{(k)} > m \mid A) = \sum_{j=0}^{k-1} \binom{m+k}{j} A^j (1 - A)^{m+k-j}. \tag{24}$$

The above representations call for the following two auxiliary results.

Lemma 2.1 *Under the assumption Eq. 7,*

$$\mathbb{E}\big((1 - A)^m\big) \sim \overline{F}(\log m) \quad \text{as } m \to \infty. \tag{25}$$

Lemma 2.2 *Under the assumption Eq. 7, there exist $\gamma < \infty$ and $\varepsilon > 0$ such that*

$$\mathbb{E}A^j (1 - A)^m \leq \gamma \frac{j^j m^m}{(m + j)^{m+j}} \overline{F}(\log m - \log j) \quad \text{for all } m > 1 \text{ and } j \leq \varepsilon m.$$

In particular, for any fixed $j \geq 1$,

$$\mathbb{E}A^j (1 - A)^m = o(\overline{F}(\log m)) \quad \text{as } m \to \infty. \tag{26}$$

Proof of Lemma 2.1 Since, for any fixed $\varepsilon > 0$,

$$\mathbb{E}\big((1 - A)^{m+1};\ A > \varepsilon\big) \leq (1 - \varepsilon)^{m+1}$$

is exponentially decreasing as $m \to \infty$, the asymptotic behaviour of the right-hand side in Eq. 25 is determined by the tail behavior of A near 0. Notice that, for $0 < a < b < 1$,

$$\begin{aligned}\mathbb{P}\left(A \in (a, b]\right) &= \mathbb{P}\left(\log \frac{1 - A}{A} \in \left[\log \frac{1 - b}{b},\ \log \frac{1 - a}{a}\right)\right) \\ &= \mathbb{P}\big(\xi \in [\log(1/b - 1),\ \log(1/a - 1))\big).\end{aligned} \tag{27}$$

Hence, for any fixed $c > 0$, we have

$$\begin{aligned}\mathbb{E}(1 - A)^m &\geq \mathbb{E}[(1 - A)^m;\ A \leq c/m] \\ &\geq (1 - c/m)^m \mathbb{P}(A \leq c/m) \\ &= (1 - c/m)^m \overline{F}(\log(m/c - 1)).\end{aligned}$$

It follows from the long-tailedness of the distribution F of ξ that the right-hand side of above equation is asymptotically equivalent to $e^{-c}\overline{F}(\log m)$ as $m \to \infty$. Letting $c \downarrow 0$ we complete the proof of the lower bound

$$\mathbb{E}(1-A)^m \ \geq\ (1+o(1))\overline{F}(\log m) \quad \text{as } m \to \infty.$$

To obtain the matching upper bound, let us consider the following decomposition which is valid for all integer $K \in [1, [m/2]-1]$:

$$\begin{aligned}
&\mathbb{E}(1-A)^m \\
&= \mathbb{E}\left[(1-A)^m;\ A \leq \frac{K}{m}\right] + \sum_{k=K}^{[m/2]-1} \mathbb{E}\left[(1-A)^m;\ A \in \left(\frac{k}{m}, \frac{k+1}{m}\right]\right] \\
&\quad + \mathbb{E}\left[(1-A)^m;\ A > \frac{[m/2]}{m}\right] \\
&\leq \mathbb{P}\left(A \leq \frac{K}{m}\right) + \sum_{k=K}^{[m/2]-1} \left(1-\frac{k}{m}\right)^m \mathbb{P}\left(A \leq \frac{k+1}{m}\right) + \left(1-\frac{[m/2]}{m}\right)^m \\
&\leq \overline{F}\left(\log\left(\frac{m}{K}-1\right)\right) + \sum_{k=K}^{[m/2]-1} e^{-k}\overline{F}\left(\log\left(\frac{m}{k+1}-1\right)\right) + \left(1-\frac{[m/2]}{m}\right)^m.
\end{aligned}$$

Let us show that the series in the middle term in the last line is negligible for large values of K. Indeed, firstly,

$$\frac{m}{k+1} - 1 \ \geq\ \frac{1}{2}\frac{m}{k+1} \quad \text{for all } k \leq \frac{m}{2} - 1$$

and hence

$$\sum_{k=K}^{[m/2]-1} e^{-k}\overline{F}\left(\log\left(\frac{m}{k+1}-1\right)\right) \ \leq\ \sum_{k=K}^{[m/2]-1} e^{-k}\overline{F}(\log m - \log(k+1) - \log 2).$$

Since the distribution F is assumed long-tailed, $\overline{F}(x-1) \leq e\overline{F}(x)$ for all sufficiently large x. Hence, there exists a constant $\gamma < \infty$ such that $\overline{F}(x-y) \leq \gamma e^y \overline{F}(x)$ for all $x, y > 0$. Therefore,

$$\begin{aligned}
\sum_{k=K}^{[m/2]-1} e^{-k}\overline{F}\left(\log\left(\frac{m}{k+1}-1\right)\right) &\leq \gamma \overline{F}(\log m) \sum_{k=K}^{\infty} e^{-k} e^{\log(k+1)+\log 2} \\
&\leq \varepsilon(K)\overline{F}(\log m)
\end{aligned} \tag{28}$$

where

$$\varepsilon(K) := \gamma \sum_{k=K}^{\infty} e^{-k} e^{\log(k+1)+\log 2} \ \to\ 0 \quad \text{as } K \to \infty.$$

Hence we conclude that

$$\mathbb{E}(1-A)^m \ \leq\ \overline{F}(\log(m/K-1)) + \varepsilon(K)\overline{F}(\log m) + O(1/2^m) \quad \text{as } m \to \infty.$$

Due to the long-tailedness of F this implies that, for any fixed K,

$$\mathbb{E}(1-A)^m \ \le\ (1+o(1))\overline{F}(\log m)+\varepsilon(K)\overline{F}(\log m)+O(1/2^m) \quad \text{as } m\to\infty.$$

The long-tailedness of F also implies that $2^m\overline{F}(\log m)\to\infty$. Thus

$$\mathbb{E}(1-A)^m \ \le\ (1+o(1))\overline{F}(\log m)+\varepsilon(K)\overline{F}(\log m) \quad \text{as } m\to\infty,$$

and since $\varepsilon(K)\to 0$ as $K\to\infty$, the proof is complete. □

Proof of Lemma 2.2 There exist $K\in\mathbb{N}$ and $\varepsilon_1>0$ such that the following inequalities hold

$$\log(k+1)\ \le\ k/6 \quad \text{for all } k\ge K \tag{29}$$

and

$$\left(1-\frac{j}{m}\right)^m\ \ge\ \frac{1}{3^j} \quad \text{for all } m>K \text{ and } j\le\varepsilon_1 m. \tag{30}$$

Similar to the case $j=0$ considered in the proof of Lemma 2.1, we make use of the following decomposition:

$$\begin{aligned}\mathbb{E}A^j(1-A)^m\ &=\ \mathbb{E}\left[A^j(1-A)^m;\ A\le\frac{Kj}{3m}\right]\\ &\quad+\sum_{k=K}^{[3m/j]}\mathbb{E}\left[A^j(1-A)^m;\ A\in\left(k\frac{j}{3m},(k+1)\frac{j}{3m}\right]\right]\\ &=:\ E_1+E_2.\end{aligned} \tag{31}$$

The maximum of the function $x^j(1-x)^m$ over the interval $[0,1]$ is attained at point $j/(m+j)$ and is equal to $j^jm^m/(m+j)^{m+j}$. Therefore, for some $\varepsilon=\varepsilon(K)\le\varepsilon_1$,

$$\begin{aligned}E_1&\le\frac{j^jm^m}{(m+j)^{m+j}}\mathbb{P}\left(A\le\frac{Kj}{3m}\right)\\ &=\frac{j^jm^m}{(m+j)^{m+j}}\overline{F}\left(\log\left(\frac{3m}{Kj}-1\right)\right)\\ &\le\gamma_1\frac{j^jm^m}{(m+j)^{m+j}}\overline{F}(\log m-\log j) \quad \text{for some } \gamma_1<\infty \text{ and all } j\le\varepsilon m,\end{aligned} \tag{32}$$

owing to the long-tailedness of F. Further, the series on the right hand side of Eq. 31 possesses the following upper bound

$$\begin{aligned}E_2\ &\le\ \sum_{k=K}^{[3m/j]}(k+1)^j\left(\frac{j}{3m}\right)^j\left(1-\frac{kj}{3m}\right)^m\mathbb{P}\left(A\le(k+1)\frac{j}{3m}\right)\\ &\le\ \left(\frac{j}{3m}\right)^j\sum_{k=K}^{[3m/j]}(k+1)^je^{-kj/3}\overline{F}(\log(3m/(k+1)j-1))\end{aligned}$$

because $(1 - kj/3m)^m \le e^{-kj/3}$. Let us now bound the latter series. It follows from the inequality Eq. 29 that

$$(k+1)^j e^{-kj/3} = e^{j(\log(k+1)-k/3)} \le e^{-jk/6} \quad \text{for all } k \ge K.$$

Then, using arguments similar to those in Eq. 28,

$$\begin{aligned} E_2 &\le \left(\frac{j}{3m}\right)^j \sum_{k=K}^{[3m/j]} e^{-jk/6} \overline{F}(\log(3m/(k+1)j - 1)) \\ &\le \gamma_2 \left(\frac{j}{3m}\right)^j \overline{F}(\log m - \log j) \quad \text{for some } \gamma_2 < \infty, \end{aligned} \tag{33}$$

which implies the result due to the inequalities Eq. 32 and

$$\frac{j^j m^m}{(m+j)^{m+j}} = \left(\frac{j}{m}\right)^j \left(1 - \frac{j}{m+j}\right)^{m+j} \ge \left(\frac{j}{3m}\right)^j$$

which is guarantied by Eq. 30. □

Proof of Theorem 1.1 We prove the statement by induction in $n \ge 1$. The assertion for $n = 1$ follows from the equality

$$\mathbb{P}(Z_1 > m) = \mathbb{P}(B_1 > m) = \mathbb{E}\big((1 - A_0)^{m+1}\big),$$

the representation Eq. 23 and Lemma 2.1. Assume that the assertion of Theorem 1.1 is valid for some $n \ge 1$. Let us show that then it follows for $n + 1 \ge 2$. Our aim is to obtain the tail asymptotics of the distribution of

$$Z_{n+1} = \sum_{i=1}^{Z_n+1} B_{n+1,i},$$

where $(B_{n+1,i}, i \ge 1)$ are independent copies of a geometric random variable B_{n+1} with success probability A_n (its probability mass function is specified in Eq. 2) and independent of Z_n conditioned on $\mathcal{A}$. Then the following representation holds

$$\begin{aligned} \mathbb{P}(Z_{n+1} > m) &= \sum_{k=0}^{\infty} \mathbb{P}\left(\sum_{j=1}^{k+1} B_{n+1,j} > m, Z_n = k\right) \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left[\mathbb{P}\left(\sum_{j=1}^{k+1} B_{n+1,j} > m \,\middle|\, \mathcal{A}\right)\right] \mathbb{P}(Z_n = k), \end{aligned} \tag{34}$$

where we have conditioned on $\mathcal{A}$ and used the fact that Z_n and $(B_{n+1,i}, i \ge 1)$ are independent conditioned on $\mathcal{A}$.

We start with the proof of the upper bound. For that, let us split the summation in Eq. 34 into three parts, from 0 to K, from $K + 1$ to $\varepsilon m - 1$ and from εm to ∞

where integer K is chosen large enough and real $\varepsilon > 0$ small enough. This splitting together with non-negativity of the B's implies that

$$\begin{aligned}
&\mathbb{P}(Z_{n+1} > m) \\
&\quad \le \mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{K} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big]\mathbb{P}(Z_n < K) \\
&\qquad + \sum_{k=K}^{\varepsilon m} \mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{k+1} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big]\mathbb{P}(Z_n = k) + \mathbb{P}(Z_n > \varepsilon m) \\
&\quad \le \mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{K} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big] + \sum_{k=K}^{\varepsilon m} \mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{k+1} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big]\mathbb{P}(Z_n = k) \\
&\qquad + \mathbb{P}(Z_n > \varepsilon m).
\end{aligned}$$

By the induction hypothesis and long-tailedness of F, for any fixed ε,

$$\mathbb{P}(Z_n > \varepsilon m) \sim n\overline{F}(\log(\varepsilon m)) \sim n\overline{F}(\log m) \quad \text{as } m \to \infty.$$

So it is left to show that, for any fixed K,

$$\mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{K} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big] \sim \overline{F}(\log m) \quad \text{as } m \to \infty, \tag{35}$$

and that, for any $\delta > 0$, there exist a sufficiently large K and a sufficiently small $\varepsilon > 0$ such that

$$\begin{aligned}
&\sum_{k=K}^{\varepsilon m} \mathbb{E}\Big[\mathbb{P}\Big(\sum_{j=1}^{k+1} B_{n+1,j} > m\Big|\mathcal{A}\Big)\Big]\mathbb{P}(Z_n = k) \\
&\quad \le \delta\overline{F}(\log m) \quad \text{for all sufficiently large } m.
\end{aligned} \tag{36}$$

We start with proving Eq. 36.

Let $\xi(A)$ be a Bernoulli random variable with success probability A and $S_{m+k}(A)$ be the sum of $m+k$ independent copies of $\xi(A)$. It follows from the representation Eq. 24 that

$$\begin{aligned}
\mathbb{P}(B^{(1)} + \ldots + B^{(k)} > m \mid A) &= \mathbb{P}(S_{m+k}(A) \le k-1) \\
&\le (\mathbb{E}(e^{-\beta\xi(A)}))^{m+k} e^{\beta(k-1)} \\
&= (1 - A + e^{-\beta}A)^{m+k} e^{\beta(k-1)}, \quad \text{for all } \beta > 0.
\end{aligned}$$

The minimal value of the right hand side is attained for β such that $e^{-\beta} = \frac{(1-A)(k-1)}{A(m+1)}$, hence

$$\mathbb{P}(B^{(1)} + \ldots + B^{(k)} > m \mid A) \le \frac{(m+k)^{m+k}}{(m+1)^{m+1}(k-1)^{k-1}} A^{k-1}(1-A)^{m+1}.$$

This allows us to conclude from Lemma 2.2 that, for $k \le \varepsilon m$,

$$\begin{aligned}\mathbb{P}(B^{(1)} + \ldots + B^{(k)} > m) &\le \frac{(m+k)^{m+k}}{(m+1)^{m+1}(k-1)^{k-1}} \mathbb{E} A^{k-1}(1-A)^{m+1} \\ &\le \gamma \overline{F}(\log(m+1) - \log(k-1)).\end{aligned}$$

Therefore,

$$\sum_{k=K}^{\varepsilon m} \mathbb{P}(B_{n+1,1} + \ldots + B_{n+1,k+1} > m)\mathbb{P}(Z_n = k) \le \gamma \sum_{k=K}^{\varepsilon m} \overline{F}(\log(m+1) - \log k)\mathbb{P}(Z_n = k).$$

Representing $\mathbb{P}(Z_n = k)$ as the difference $\mathbb{P}(Z_n > k-1) - \mathbb{P}(Z_n > k)$ and rearranging the sum on the right hand side we conclude that this sum is not greater than

$$\begin{aligned}&\overline{F}(\log(m+1) - \log K)\mathbb{P}(Z_n > K-1) \\ &\quad + \sum_{k=K}^{\varepsilon m-1} \big(\overline{F}(\log(m+1) - \log(k+1)) - \overline{F}(\log(m+1) - \log k)\big)\mathbb{P}(Z_n > k).\end{aligned}$$

Then the induction hypothesis yields an upper bound, for some $\gamma_1 < \infty$,

$$\begin{aligned}&\sum_{k=K}^{\varepsilon m} \mathbb{P}(B_{n+1,1} + \ldots + B_{n+1,k+1} > m)\mathbb{P}(Z_n = k) \\ &\le \gamma \overline{F}(\log(m+1) - \log K)\mathbb{P}(Z_n > K-1) \\ &\quad + \gamma_1 \sum_{k=K}^{\epsilon m-1} \big(\overline{F}(\log(m+1) - \log(k+1)) - \overline{F}(\log(m+1) - \log k)\big)\overline{F}(\log k).\end{aligned}$$

Due to the long-tailedness of F, for any $\delta > 0$ there exists a sufficiently large K such that the first term on the right hand side is not greater than $\delta \overline{F}(\log m)$, for all sufficiently large m. After rearranging we conclude that the sum on the right hand side is not greater than

$$\begin{aligned}&\overline{F}(\log(m+1) - \log(\varepsilon m))\overline{F}(\log(\varepsilon m - 1)) \\ &\quad + \sum_{k=K+1}^{\varepsilon m-1} \overline{F}(\log(m+1) - \log k)\big(\overline{F}(\log(k-1)) - \overline{F}(\log k)\big).\end{aligned} \tag{37}$$

Since F is long-tailed, the first term here is asymptotically equivalent to

$$\overline{F}(\log(1/\varepsilon))\overline{F}(\log m) \quad \text{as } m \to \infty,$$

so it is not greater than $\delta \overline{F}(\log m)$ for all sufficiently large m provided $\overline{F}(\log(1/\varepsilon)) \le \delta/2$. The sum in Eq. 37 equals

$$\sum_{k=K+1}^{\varepsilon m-1} \overline{G}\left(\frac{m+1}{k}\right) G(k-1, k],$$

where the distribution G is defined via its tail as $\overline{G}(x) = \overline{F}(\log x)$, and can be bounded by the integral

$$\begin{aligned}\int_K^{\varepsilon m} \overline{G}(m/z)G(dz) &= \mathbb{P}(e^{\xi_1+\xi_2} > m;\ e^{\xi_2} \in (K, \varepsilon m]) \\ &= \mathbb{P}(\xi_1 + \xi_2 > \log m;\ \xi_2 \in (\log K,\ \log m - \log(1/\varepsilon)]).\end{aligned}$$

Since the distribution F is assumed to be subexponential, we can choose a sufficiently large K and a sufficiently small $\varepsilon > 0$ such that the latter probability is not greater than $\delta\overline{F}(\log m)$ for all sufficiently large m, see (Foss et al. 2013, Theorem 3.6), which completes the proof of Eq. 36.

To complete the proof of the upper bound it now suffices to show Eq. 35. This follows immediately from the representation Eq. 24, the asymptotics Eq. 26 and Lemma 2.1.

We will obtain now the matching lower bound. For that, let us split the sum in Eq. 34 into two parts, from 0 to cm and from $cm + 1$ to ∞ where c is a large number sent to infinity later on. This splitting implies that

$$\begin{aligned}&\mathbb{P}(Z_{n+1} > m) \\ &\geq \sum_{k=0}^{cm} \mathbb{E}\big[\mathbb{P}(B_{n+1,1} > m \mid \mathcal{A})\big]\mathbb{P}(Z_n = k) \\ &\quad + \sum_{k=cm+1}^{\infty} \mathbb{E}\left[\mathbb{P}\left(\sum_{j=1}^{k+1} B_{n+1,j} > m \middle| \mathcal{A}\right)\right]\mathbb{P}(Z_n = k) \\ &\geq \mathbb{E}\big[\mathbb{P}(B_{n+1,1} > m \mid \mathcal{A})\big]\mathbb{P}(Z_n \leq cm) \\ &\quad + \mathbb{E}\left[\mathbb{P}\left(\sum_{j=1}^{cm} B_{n+1,j} > m \middle| \mathcal{A}\right)\right]\mathbb{P}(Z_n > cm), \end{aligned} \tag{38}$$

since all the B's are non-negative. By Lemma 2.1,

$$\mathbb{E}\big[\mathbb{P}(B_{n+1,1} > m \mid \mathcal{A})\big]\mathbb{P}(Z_n \leq cm) \sim \overline{F}(\log m) \quad \text{as } m \to \infty. \tag{39}$$

Further, by the law of large numbers,

$$\mathbb{P}\left(\sum_{j=1}^{cm} B_{n+1,j} > m \middle| \mathcal{A}\right) \to 1 \quad \text{a.s. as } c \to \infty.$$

Hence, the dominated convergence theorem allows us to conclude that

$$\mathbb{E}\left[\mathbb{P}\left(\sum_{j=1}^{cm} B_{n+1,j} > m \middle| \mathcal{A}\right)\right] \to 1 \quad \text{as } c \to \infty. \tag{40}$$

Finally, by the induction hypothesis and long-tailedness of F, for any fixed c,

$$\mathbb{P}(Z_n > cm) \sim n\overline{F}(\log(cm)) \sim n\overline{F}(\log m) \quad \text{as } m \to \infty. \tag{41}$$

Substituting Eq. 39–Eq. 41 into Eq. 38 and letting $c \to \infty$ we conclude the induction step for the lower bound. □

3 Proof of the lower bound, Theorem 1.2

Note that, by the strong law of large numbers, for any fixed $\varepsilon > 0$,

$$\inf_{n\geq 1} \mathbb{P}(C_S(c,\varepsilon,k,n) \text{ for all } k \leq n) \to 1 \quad \text{as } c \to \infty, \tag{42}$$

where

$$C_S(c,\varepsilon,k,n) := \{|S_{k,n} - (n-k+1)\mathbb{E}\xi| \leq c + \varepsilon(n-k+1)\}$$

and $S_{k,n} = \xi_k + \ldots + \xi_n$.

We show that, under the long-tailedness condition Eq. 7, the most probable way for a big value of Z_n to occur is due to atypical random environment when one of the following events occurs, $k \leq n-1$:

$$C_A(k,n) := \left\{A_k \leq \frac{c_1}{M(m,k,n)},\ C_S(c_2,\varepsilon,j,n-1) \text{ for all } j \in [k+1,n-1]\right\},$$

where

$$M(m,k,n) := me^{\varepsilon(n-1-k)+c_2}\prod_{j=k+1}^{n-1}\frac{1}{a_{A_j}} = me^{\varepsilon(n-1-k)+c_2-S_{k+1,n-1}},$$

$a_A := \mathbb{E}\{B \mid A\} = 1/A - 1 = e^{\xi}$, $c_1, c_2, \varepsilon > 0$ are fixed, c_2 will be sent to infinity later on, while c_1 and ε will be sent to 0. Since A is bounded by $\widehat{A} < 1$, a_A is bounded away from 0 by $1/\widehat{A} - 1$.

Let us bound from below the probability of the union of events $C_A(k,n)$. We start with the following lower bound

$$\mathbb{P}\left(\bigcup_{k=0}^{n-1} C_A(k,n)\right) \geq \sum_{k=0}^{n-1}\mathbb{P}(C_A(k,n)) - \sum_{k\neq l}\mathbb{P}(C_A(k,n)\cap C_A(l,n)). \tag{43}$$

On the event $C_S(c_2,\varepsilon,k+1,n-1)$ we have

$$a(n-1-k) \leq \varepsilon(n-1-k) + c_2 - S_{k+1,n-1} \leq 2c_2 + (2\varepsilon + a)(n-1-k) \tag{44}$$

and hence

$$\begin{aligned}
&\sum_{k=0}^{n-1}\mathbb{P}(C_A(k,n))\\
&\geq \sum_{k=0}^{n-1}\mathbb{P}\left(A_k \leq \frac{c_1}{me^{2c_2+(2\varepsilon+a)(n-1-k)}},\ C_S(c_2,\varepsilon,j,n-1) \text{ for all } j \in [k+1,n-1]\right)\\
&= \sum_{k=0}^{n-1}\mathbb{P}\left(A_k \leq \frac{c_1}{me^{2c_2+(2\varepsilon+a)(n-1-k)}}\right)\mathbb{P}\big(C_S(c_2,\varepsilon,j,n-1) \text{ for all } j \in [k+1,n-1]\big)\\
&\geq \mathbb{P}\big(C_S(c_2,\varepsilon,j,n-1) \text{ for all } j \in [1,n-1]\big)\sum_{k=0}^{n-1}\mathbb{P}\left(A_k \leq \frac{c_1}{me^{2c_2+(2\varepsilon+a)(n-1-k)}}\right),
\end{aligned}$$

and

$$\begin{aligned}\sum_{k\neq l}\mathbb{P}(C_A(k,n)\cap C_A(l,n)) &\le \sum_{k\neq l}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{a(n-1-k)}},\ A_l\le \frac{c_1}{me^{a(n-1-l)}}\Big)\\ &= \sum_{k\neq l}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{a(n-1-k)}}\Big)\mathbb{P}\Big(A_l\le \frac{c_1}{me^{a(n-1-l)}}\Big)\\ &\le \Big(\sum_{k=0}^{n-1}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{a(n-1-k)}}\Big)\Big)^2.\end{aligned}$$

As follows from Eq. 27,

$$\begin{aligned}\sum_{k=0}^{n-1}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{2c_2+(2\varepsilon+a)(n-1-k)}}\Big) &= \sum_{k=0}^{n-1}\mathbb{P}\Big(\xi\ge \log\Big(\frac{me^{2c_2+(2\varepsilon+a)k}}{c_1}-1\Big)\Big)\\ &\ge \sum_{k=0}^{n-1}\overline{F}(\log m+2c_2+(2\varepsilon+a)k-\log c_1)\\ &\ge \frac{1}{2\varepsilon+a}\int_{\log m+2c_2-\log c_1}^{\log m+2c_2-\log c_1+(2\varepsilon+a)n}\overline{F}(x)dx\end{aligned}$$

since the tail function $\overline{F}(x)$ is decreasing. Therefore,

$$\sum_{k=0}^{n-1}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{2c_2+(2\varepsilon+a)(n-1-k)}}\Big)\ge \frac{1+o(1)}{2\varepsilon+a}\int_{\log m}^{\log m+(2\varepsilon+a)n}\overline{F}(x)dx$$

as $m\to\infty$ uniformly for all $n\ge 1$ because the distribution F is long-tailed. Similarly,

$$\sum_{k=0}^{n-1}\mathbb{P}\Big(A_k\le \frac{c_1}{me^{a(n-1-k)}}\Big)\le \frac{1+o(1)}{a}\int_{\log m}^{\log m+na}\overline{F}(x)dx.$$

Therefore,

$$\begin{aligned}\sum_{k=0}^{n-1}\mathbb{P}(C_A(k,n)) \ge \frac{1+o(1)}{2\varepsilon+a}\int_{\log m}^{\log m+na}\overline{F}(x)dx\,\mathbb{P}\big(C_S(c_2,\varepsilon,j,n-1)\\ \text{for all } j\in[1,n-1]\big),\end{aligned}$$

and, as $m\to\infty$ uniformly for all $n\ge 1$,

$$\begin{aligned}\sum_{k\neq l}\mathbb{P}(C_A(k,n)\cap C_A(l,n)) &= O\Big(\int_{\log m}^{\log m+na}\overline{F}(x)dx\Big)^2\\ &= o\Big(\int_{\log m}^{\log m+na}\overline{F}(x)dx\Big),\end{aligned}$$

because the integral tends to 0 due to the integrability of the tail of F. Substituting these bounds into Eq. 43 and applying Eq. 42, for any fixed $\varepsilon > 0$ we can conclude the following lower bound,

$$\mathbb{P}\left(\bigcup_{k=0}^{n-1} C_A(k,n)\right) \geq \frac{g(c_2)+o(1)}{2\varepsilon + a} \int_{\log m}^{\log m + na} \overline{F}(x)dx \tag{45}$$

as $m \to \infty$ uniformly for all $n \geq 1$, where $g(c_2) \to 1$ as $c_2 \to \infty$.

As above, conditioning on $\mathcal{A}$ yields

$$\begin{aligned} \mathbb{P}(Z_n > m) &= \mathbb{E}[\mathbb{P}(Z_n > m \mid \mathcal{A})] \\ &\geq \mathbb{E}[\mathbb{P}(Z_n > m \mid \mathcal{A});\ C_A(n)], \end{aligned} \tag{46}$$

where $C_A(n) := \bigcup_{k=0}^{n-1} C_A(k,n)$. Then, owing to Eq. 45, for the proof of Eq. 12 it suffices to show that

$$\liminf_{m\to\infty} \inf_{C_A(n)} \mathbb{P}(Z_n > m \mid \mathcal{A}) \geq e^{-c_1} \quad \text{uniformly for all } n \geq 1. \tag{47}$$

Hence we are left with the proof of Eq. 47. Since the event $C_A(n)$ is the union of events $C_A(k,n)$, $k \leq n-1$, the probability of the event

$$C_B(k,n) := \left\{ B_{k+1,1} > me^{c_2+\varepsilon(n-1-k)} \prod_{j=k+1}^{n-1} \frac{1}{a_{A_j}} \right\},$$

conditionally on $C_A(n)$, possesses the following asymptotic lower bound

$$\begin{aligned} \mathbb{P}(C_B(k,n) \mid C_A(n)) &\geq (1-A_k)^{me^{c_2+\varepsilon(n-1-k)} \prod_{j=k+1}^{n-1} \frac{1}{a_{A_j}}} \mid C_A(n) \\ &\geq \left(1 - \frac{c_1}{me^{c_2+\varepsilon(n-1-k)} \prod_{j=k+1}^{n-1} \frac{1}{a_{A_j}}}\right)^{me^{c_2+\varepsilon(n-1-k)} \prod_{j=k+1}^{n-1} \frac{1}{a_{A_j}}} \\ &\to e^{-c_1} \quad \text{as } m \to \infty. \end{aligned}$$

Therefore, it only remains to show that

$$\inf_{C_A(k,n)} \mathbb{P}(Z_n > m \mid C_B(k,n),\ \mathcal{A}) \to 1 \tag{48}$$

as $m \to \infty$ uniformly for all $k \leq n-1$ and $n \geq 1$.

To prove this convergence, let us note that, conditioned on $\mathcal{A}$,

$$\begin{aligned} &\mathbb{P}\big[Z_j \leq l a_{A_{j-1}} e^{-\varepsilon} \big| Z_{j-1} = l, \mathcal{A}\big] \\ &= \mathbb{P}\big[B_{j,1} + \ldots + B_{j,l+1} \leq l a_{A_{j-1}} e^{-\varepsilon} \big| \mathcal{A}\big] \\ &\leq \mathbb{P}\left[\frac{B_{j,1}}{a_{A_{j-1}}} + \ldots + \frac{B_{j,l}}{a_{A_{j-1}}} \leq l e^{-\varepsilon} \middle| \mathcal{A}\right] \\ &= \mathbb{P}\left[\left(e^{-\varepsilon/2} - \frac{B_{j,1}}{a_{A_{j-1}}}\right) + \ldots + \left(e^{-\varepsilon/2} - \frac{B_{j,l}}{a_{A_{j-1}}}\right) \geq l(e^{-\varepsilon/2} - e^{-\varepsilon}) \middle| \mathcal{A}\right] \\ &\leq \mathbb{P}\left[\left(e^{-\varepsilon/2} - \frac{B_{j,1}}{a_{A_{j-1}}}\right) + \ldots + \left(e^{-\varepsilon/2} - \frac{B_{j,l}}{a_{A_{j-1}}}\right) \geq l e^{-\varepsilon} \varepsilon/2 \middle| \mathcal{A}\right]. \end{aligned}$$

Applying the exponential Markov inequality, we obtain the following upper bound, for all $\lambda > 0$,

$$\mathbb{P}\big[Z_j \le l a_{A_{j-1}} e^{-\varepsilon} \big| Z_{j-1} = l, \mathcal{A}\big]$$
$$\le e^{-l\lambda e^{-\varepsilon}\varepsilon/2} \mathbb{E}\Big[e^{\lambda\left(\left(e^{-\varepsilon/2} - \frac{B_{j,1}}{a_{A_{j-1}}}\right) + \ldots + \left(e^{-\varepsilon/2} - \frac{B_{j,l}}{a_{A_{j-1}}}\right)\right)} \Big| \mathcal{A}\Big].$$

Since

$$\begin{aligned}\mathbb{E}\Big[e^{\lambda\left(e^{-\varepsilon/2} - \frac{B}{a_A}\right)} \Big| A\Big] &= e^{\lambda e^{-\varepsilon/2}} \frac{A}{1-(1-A)e^{-\lambda \frac{A}{1-A}}} \\ &= e^{\frac{\lambda}{1-A} + \lambda(e^{-\varepsilon/2}-1)} \frac{A}{e^{\lambda \frac{A}{1-A}} - (1-A)} \\ &\le e^{\lambda(e^{-\varepsilon/2}-1)} \frac{e^{\frac{\lambda}{1-A}}}{\frac{\lambda}{1-A} + 1}\end{aligned}$$

and since A is bounded away from 1, there exists a sufficiently small $\lambda_0 > 0$ such that

$$\mathbb{E}\Big[e^{\lambda_0\left(e^{-\varepsilon/2} - \frac{B}{a_A}\right)} \Big| A\Big] \le 1 \quad \text{for all } A \in (0, \widehat{A}),$$

Therefore,

$$\mathbb{P}\big[Z_j \le l a_{A_{j-1}} e^{-\varepsilon} \big| Z_{j-1} = l,\ \mathcal{A}\big] \le e^{-l\delta} \quad \text{where } \delta = \lambda_0 e^{-\varepsilon}\varepsilon/2 > 0.$$

which, due to monotonicity property of the branching process Z_n, implies that

$$\mathbb{P}\big[Z_j \le l a_{A_{j-1}} e^{-\varepsilon} \big| Z_{j-1} \ge l,\ \mathcal{A}\big] \le e^{-l\delta}.$$

Then the induction arguments lead to the following upper bound

$$\mathbb{P}\Bigg[Z_n \le l e^{-\varepsilon(n-1-k)} \prod_{i=k+1}^{n-1} a_{A_i} \Bigg| Z_{k+1} \ge l,\ \mathcal{A}\Bigg] \le \sum_{j=k+1}^{n-1} e^{-l\delta e^{-\varepsilon(j-1-k)} \prod_{i=k+1}^{j-1} a_{A_i}}.$$

We take

$$l = m e^{\varepsilon(n-1-k)} \prod_{i=k+1}^{n-1} \frac{1}{a_{A_i}}$$

to conclude that

$$\mathbb{P}(Z_n > m \mid C_B(k,n),\ \mathcal{A}) \ge 1 - \sum_{j=k+1}^{n-1} e^{-m\delta e^{\varepsilon(n-1-j)} \prod_{i=j}^{n-1} \frac{1}{a_{A_i}}}.$$

Due to the representation

$$\log e^{c_2} \prod_{i=j}^{n-1} \frac{A_i}{1-A_i} = c_2 + \sum_{i=j}^{n-1} \log \frac{A_i}{1-A_i} = c_2 - \sum_{i=j}^{n-1} \xi_i,$$

we get

$$\mathbb{P}(Z_n > m \mid C_B(k,n),\ \mathcal{A}) \geq 1 - \sum_{j=k+1}^{n-1} e^{-m\delta e^{\varepsilon(n-1-j)-c_2}},$$

for any sequence of ξ's such that

$$c_2 - \sum_{i=j}^{n-1} \xi_i \geq 0 \quad \text{for all } j \in [k, n-1],$$

which is the case on $C_S(c_2, \varepsilon, k, n-1)$ and hence on $C_A(k,n)$, as follows from the first inequality in Eq. 44 for all $\varepsilon \in (0, -\mathbb{E}\xi)$. So, we have shown Eq. 48, and the proof of the first lower bound in Theorem 1.2 is complete.

The lower limit for the stationary distribution follows similar arguments if we start with an analogue of Eq. 46,

$$\begin{aligned}\mathbb{P}(Z > m) &= \lim_{n\to\infty} \mathbb{P}(Z_n > m) \\ &\geq \lim_{n\to\infty} \mathbb{E}[\mathbb{P}(Z_n > m \mid \mathcal{A});\ C_A(n)].\end{aligned} \tag{49}$$

Then, similar to Eq. 45, we may use the fact that F_I is long-tailed to conclude that

$$\lim_{n\to\infty} \mathbb{P}(C_A(n)) \geq \frac{g(c_2) + o(1)}{2\varepsilon + a} \overline{F}_I(\log m) \quad \text{as } m \to \infty, \tag{50}$$

which together with Eq. 47 justifies the lower bound for the stationary tail distribution.

4 Proof of the upper bound, Theorem 1.3

Let W_n be a branching process without immigration, that is, $W_0 = 1$ and

$$W_{n+1} = \sum_{i=1}^{W_n} B_{n+1,i} \quad \text{for } n \geq 0.$$

Let $W_n^{(0)}$ be the number of particles in Z_n generated by the immigrant arriving at time 0, $W_n^{(1)}$ be the number of particles in Z_n generated by the immigrant arriving at time 1 and so on. All these processes extinct in a finite time and are independent being conditioned on the environment $\mathcal{A}$. In addition, $W_n^{(k)}$ has the same distribution with W_{n-k} given the same success probabilities. By the definition of Z_n,

$$Z_n = W_n^{(0)} + W_n^{(1)} + \ldots + W_n^{(n-1)},$$

and hence, for any fixed $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(Z_n > m) &\leq \mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \text{ for some } k \in [0, n-1]\big) \\ &= \mathbb{E}\big[\mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \text{ for some } k \in [0, n-1] \mid \mathcal{A}\big)\big].\end{aligned}$$

Splitting the area of integration into two parts, we get the following upper bound

$$\begin{aligned}\mathbb{P}(Z_n > m) \le\ & \mathbb{P}(S_{k,n-1} > \log m - \sqrt{\log m} - 2\varepsilon(n-k) \text{ for some } k \in [0, n-1]) \\ & + \mathbb{E}\big[\mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \text{ for some } k \in [0, n-1] \mid \mathcal{A}\big); \\ & S_{k,n-1} \le \log m - \sqrt{\log m} - 2\varepsilon(n-k) \text{ for all } k \in [0, n-1]\big]. \end{aligned} \tag{51}$$

Using Eq. 10 and strong subexponentiality of F we conclude that

$$\begin{aligned} & \mathbb{P}(S_{k,n-1} + 2\varepsilon(n-k) > \log m - \sqrt{\log m} \text{ for some } k \in [0, n-1]) \\ & \sim \frac{1}{a-2\varepsilon} \int_{\log m - \sqrt{\log m}}^{\log m - \sqrt{\log m} + n(a-2\varepsilon)} \overline{F}(x)dx \end{aligned} \tag{52}$$

as $m \to \infty$ uniformly for all n, see (Korshunov 2002) and also (Foss et al. 2013), Theorem 5.3.

Further, by the Markov inequality,

$$\begin{aligned}\mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \mid \mathcal{A}\big) & \le \frac{\mathbb{E}(W_n^{(k)} \mid \mathcal{A})}{me^{-\varepsilon(n-k)}(1-e^{-\varepsilon})} \\ & = \frac{e^{S_{k,n-1}}}{me^{-\varepsilon(n-k)}(1-e^{-\varepsilon})}. \end{aligned}$$

Hence, on the event $\{S_{k,n-1} \le \log m - \sqrt{\log m} - 2\varepsilon(n-k) \text{ for all } k \in [0, n-1]\}$ we have

$$\mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \mid \mathcal{A}\big) \le \frac{e^{-\varepsilon(n-k)}}{e^{\sqrt{\log m}}(1-e^{-\varepsilon})},$$

which implies that

$$\begin{aligned} & \mathbb{E}\big[\mathbb{P}\big(W_n^{(k)} > me^{-\varepsilon(n-k)}(1-e^{-\varepsilon}) \text{ for some } k \in [0, n-1] \mid \mathcal{A}\big); \\ & \quad S_{k,n-1} \le \log m - \sqrt{\log m} - 2\varepsilon(n-k) \text{ for all } k \in [0, n-1]\big] \\ & \le \frac{1}{e^{\sqrt{\log m}}(1-e^{-\varepsilon})} \sum_{k=0}^{n-1} e^{-\varepsilon(n-k)} \\ & \le \frac{1}{e^{\sqrt{\log m}}(1-e^{-\varepsilon})^2}. \end{aligned} \tag{53}$$

Substituting Eqs. 52 and 53 into Eq. 51, we deduce that, uniformly for all $n \ge 1$,

$$\mathbb{P}(Z_n > m) \le \frac{1+o(1)}{a-2\varepsilon} \int_{\log m - \sqrt{\log m}}^{\log m - \sqrt{\log m} + na} \overline{F}(x)dx + \frac{1}{e^{\sqrt{\log m}}(1-e^{-\varepsilon})^2}.$$

By the condition Eq. 14, $\overline{F}(\log m - \sqrt{\log m}) \sim \overline{F}(\log m)$ and $\overline{F}(\log m)e^{\sqrt{\log m}} \to \infty$ as $m \to \infty$, hence

$$\mathbb{P}(Z_n > m) \le \frac{1+o(1)}{a-2\varepsilon} \int_{\log m}^{\log m + na} \overline{F}(x)dx,$$

uniformly for all $n \geq 1$. Due to the arbitrary choice of $\varepsilon > 0$, the proof of the upper bound Eq. 15 is complete.

The above arguments can be streamlined if we made use of the link Eq. 20 to stochastic difference equations. Indeed, conditioning on the environment leads to

$$\begin{aligned}\mathbb{P}(Z_n > m) &= \mathbb{E}\big[\mathbb{P}(Z_n > m \mid \mathcal{A})\big] \\ &\leq \mathbb{P}\big[\mathbb{E}(Z_n \mid \mathcal{A}) > me^{-\sqrt{\log m}}\big] \\ &\quad + \mathbb{E}\big[\mathbb{P}\big(Z_n > m \mid \mathcal{A}\big);\ \mathbb{E}(Z_n \mid \mathcal{A}) \leq me^{-\sqrt{\log m}}\big].\end{aligned}$$

For the first term on the right hand side we apply the asymptotics Eq. 21. To estimate the second term, we can apply the Markov inequality to get

$$\begin{aligned}\mathbb{P}\big(Z_n > m \mid \mathcal{A}\big) &\leq \frac{\mathbb{E}(Z_n \mid \mathcal{A})}{m} \\ &\leq \frac{me^{-\sqrt{\log m}}}{m} = e^{-\sqrt{\log m}}\end{aligned}$$

on the event $\mathbb{E}(Z_n \mid \mathcal{A}) \leq me^{-\sqrt{\log m}}$ which completes the proof.

The proof of the stationary upper bound Eq. 16 follows similar arguments with initial upper bound

$$\begin{aligned}\mathbb{P}(Z > m) &= \lim_{n\to\infty} \mathbb{P}(Z_n > m) \\ &\leq \lim_{n\to\infty} \mathbb{P}\big[\mathbb{E}(Z_n \mid \mathcal{A}) > me^{-\sqrt{\log m}}\big] \\ &\quad + \lim_{n\to\infty} \mathbb{E}\big[\mathbb{P}\big(Z_n > m \mid \mathcal{A}\big);\ \mathbb{E}(Z_n \mid \mathcal{A}) \leq me^{-\sqrt{\log m}}\big].\end{aligned}$$

and further use of the asymptotics Eq. 22 instead of Eq. 21 which is valid due to subexponentiality of the integrated tail distribution F_I. The proof of Theorem 1.3 is complete.

5 Proof of the principle of a single atypical environment, Theorem 1.5

As follows from the arguments presented in Section 3, for any fixed c and $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\bigcup_{k=0}^{n-1} E_n^{(k)}(m, c, \varepsilon)\right) &\sim \frac{1}{a+\varepsilon} \int_{\log m}^{\log m + (a+\varepsilon)n} \overline{F}(x)dx \\ &\geq \frac{1}{a+\varepsilon} \int_{\log m}^{\log m + an} \overline{F}(x)dx\end{aligned}$$

and the event presented on the left hand side implies $Z_n > m$ with high probability, that is,

$$\mathbb{P}\left(Z_n > m \;\middle|\; \bigcup_{k=0}^{n-1} E_n^{(k)}(m, c, \varepsilon)\right) \to 1 \quad \text{as } m \to \infty \text{ uniformly for all } n.$$

Then the equality

$$\mathbb{P}\Big(\bigcup_{k=0}^{n-1} E_n^{(k)}(m,c,\varepsilon) \,\Big|\, Z_n > m\Big)$$
$$= \mathbb{P}\Big(Z_n > m \,\Big|\, \bigcup_{k=0}^{n-1} E_n^{(k)}(m,c,\varepsilon)\Big) \frac{\mathbb{P}\big(\bigcup_{k=0}^{n-1} E_n^{(k)}(m,c,\varepsilon)\big)}{\mathbb{P}(Z_n > m)}$$

and Theorem 1.3 imply that

$$\lim_{m\to\infty} \inf_n \mathbb{P}\Big(\bigcup_{k=0}^{n-1} E_n^{(k)}(m,c,\varepsilon) \,\Big|\, Z_n > m\Big) \geq \frac{a}{a+\varepsilon}.$$

Letting $\varepsilon \downarrow 0$ concludes the proof.

6 Related models

The techniques developed in this paper may be applied to analysing a variety of similar models. We mention here a few of them.

Non-geometric offspring distribution Our analysis of the tail asymptotics of Z_n is particularly based on the representations Eqs. 23 and 24 available for conditionally geometric distribution of the number of offsprings B. In the case of light-tailed ξ this assumption on B is not essential as recent contributions by (Buraczewski and Dyszewski 2018) or (Basrak and Kevei 2020) demonstrate. It would be of interest to develop a technique needed for non-geometric setting in the case of heavy-tailed ξ too.

Random-size immigration One may replace size-1 immigration by a *random-size-immigration* where random sizes are i.i.d. and independent of everything else, with a common light-tailed distribution – or, more generally, the sizes may be stochastically bounded by a random variable with a light-tailed distribution.

A branching process $\{\widehat{Z}_n, n \geq 0\}$ with *state-dependent size-1 immigration* is a particular case here: an immigrant arrives only when the previous generation produces no offspring:

$$\widehat{Z}_{n+1} = \sum_{i=1}^{\max(1,\widehat{Z}_n)} B_{n+1,i}, \quad n \geq 0.$$

Clearly, $\widehat{Z}_n \leq Z_n$ a.s., for any n. Moreover, one can show that, for each n, the low bounds for $\mathbb{P}(Z_n > m)$ and $\mathbb{P}(\widehat{Z}_n > m)$ are asymptotically equivalent. Then, in particular, the statement of Theorem 1.1 stays valid with $\widehat{Z}_n$ in place of Z_n.

Continuous-space analogue Instead of the recursion Eq. 1, one may consider a "continuous-space" recursion of the form

$$Z_{n+1} = Y_{n+1} + \int_0^{Z_n} dB_{n+1}(t)$$

where B_n are subordinators with a light-tailed distribution of the Levy measure (that depends on random parameters) and $\{Y_n\}$ are i.i.d. "innovations" with a light-tailed distribution. A similar problem for a branching process with immigration, but without random environment has been studied in a recent paper by (Foss and Miyazawa 2020).

Acknowledgements The authors are very thankful to two referees for their really valuable comments and suggestions.

References

Asmussen, S., Klüppelberg, C., Sigman, K.: Sampling at subexponential times, with queueing applications. Stoch Process. Appl. **79**, 265–286 (1998)

Basrak, B., Kevei, P.: Limit theorems for branching processes with immigration in a random environment. arXiv:2002.00634v1 (2020)

Bhattacharya, A., Palmowski, Z.: Slower variation of the generation sizes induced by heavy-tailed environment for geometric branching. Stat. Probab. Lett. **154**, 108550 (2019)

Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation, Volume 27 of Encyclopedia of Mathematics and Its Applications. Cambridge University Press, Cambridge (1987)

Buraczewski, D., Dyszewski, P.: Precise large deviation estimates for branching process in random environment. Elec. J Probab. **23**, 1–16 (2018)

Darling, D.A.: The Galton-Watson process with infinite mean. J. Appl. Prob. **7**, 455–456 (1970)

Denisov, D., Korshunov, D., Wachtel, V.: At the Edge of Criticality: Markov Chains with Asymptotically Zero Drift. Cramér–Doob's Approach to Lamperti's Problem. Springer (To appear)

Dmitruschenkov, D.V., Shklyaev, A.V.: Large deviations of branching processes with immigration in random environment. Discrete Math. Appl. **27**, 361–376 (2017)

Dyszewski, P.: Iterated random functions and slowly varying tails. Stochastic Process Appl. **126**, 392–413 (2016)

Foss, S., Korshunov, D.: Sampling at a random time with a heavy-tailed distribution. Markov Proc. Rel Fields **6**(4), 543–568 (2000)

Foss, S., Korshunov, D., Zachary, S. An Introduction to Heavy-tailed and Subexponential Distributions, 2nd Edition. Springer, New York (2013)

Foss, S., Miyazawa, M.: Tails in a fixed-point problem for a branching process with state-independent immigration. Marc. Proc. Rel Fields **26**, 619–636 (2020)

Hong, W., Zhang, X.: Asymptotic behaviour of heavy-tailed branching processes in random environments. Electron. J Probab. **24**, 56 (2019)

Kesten, H., Kozlov, M.V., Spitzer, F.: A limit law for random walk in a random environment. Compositio Mathematica **30**(2), 145–168 (1975)

Korshunov, D.: Large-deviation probabilities for maxima of sums of independent random variables with negative mean and subexponential distribution. Theory Probab. Appl. **46**, 355–366 (2002)

Korshunov, D.: In: A Lifetime of Excursions Through Random Walks and Lévy Processes: In: A Volume in Honour of Ron Doney's 80th Birthday. Kyprianou, A. E. & Chaumon, L. (eds.). Birkhauser, (Progress in Probability; vol. 78), (2021)

Seneta, E.: The simple branching process with infinite mean I. J. Appl. Prob. **10**, 206–212 (1973)

Schuh, H.J., Barbour, A.D.: On the asymptotic behaviour of branching processes with infinite mean. Adv. Appl. Prob. **9**, 681–723 (1977)

Solomon, F.: Random walks in a random environment. Ann. Probab. **3**(1), 1–31 (1975)

Vatutin, V., Zheng, X.: Subcritical branching processes in a random environment without the Cramer condition. Stochastic Process Appl. **122**, 2594–2609 (2012)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Sergey Foss[1,2] · **Dmitry Korshunov**[2,3] · **Zbigniew Palmowski**[4]

Sergey Foss
S.Foss@hw.ac.uk

Dmitry Korshunov
d.korshunov@lancaster.ac.uk

1 Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, Scotland, UK

2 MCA, Novosibirsk State University, Novosibirsk, Russia

3 Lancaster University, Lancaster, UK

4 Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland