Large deviations for Branching Processes in Random Environment

Vincent Bansaye, Julien Berestycki

September 26, 2009

Abstract

A branching process in random environment $(Z_n, n \in \mathbb{N})$ is a generalization of Galton Watson processes where at each generation the reproduction law is picked randomly. In this paper we give several results which belong to the class of *large deviations*. By contrast to the Galton-Watson case, here random environments and the branching process can conspire to achieve atypical events such as $Z_n \leq e^{cn}$ when c is smaller than the typical geometric growth rate $\bar{L}$ and $Z_n \geq e^{cn}$ when $c > \bar{L}$.

One way to obtain such an atypical rate of growth is to have a typical realization of the branching process in an atypical sequence of environments. This gives us a general lower bound for the rate of decrease of their probability.

When each individual leaves at least one offspring in the next generation almost surely, we compute the exact rate function of these events and we show that conditionally on the large deviation event, the trajectory $t \mapsto \frac{1}{n} \log Z_{[nt]}, t \in [0,1]$ converges to a deterministic function $f_c : [0,1] \mapsto \mathbb{R}_+$ in probability in the sense of the uniform norm. The most interesting case is when $c < \bar{L}$ and we authorize individuals to have only one offspring in the next generation. In this situation, conditionally on $Z_n \leq e^{cn}$, the population size stays fixed at 1 until a time $\sim nt_c$. After time nt_c an atypical sequence of environments let Z_n grow with the appropriate rate ($\neq \bar{L}$) to reach c. The corresponding map $f_c(t)$ is piecewise linear and is 0 on $[0, t_c]$ and $f_c(t) = c(t - t_c)/(1 - t_c)$ on $[t_c, 1]$.

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

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

1 Introduction

Let $\mathcal{P}$ be the space of probability measures on the integer, that is

$$\mathcal{P} := \{p : \mathbb{N} \mapsto [0,1] : \sum_{k\geq 0} p(k) = 1\},$$

and denote by $m(p)$ the mean of p :

$$m(p) = \sum_{k\geq 0} kp(k).$$

A branching process in random environment (BPRE for short) $(Z_n, n \in \mathbb{N})$ with environment distribution $\mu \in \mathcal{M}_1(\mathbb{P})$ is a discrete time Markov process which evolves as follows : at time n, we draw $\mathbf{p}$ according to μ independently of the past and then each individual $i = 1, \ldots, Z_n$ reproduces independently according to the same $\mathbf{p}$, i.e. the probability that individual i gives birth to k offsprings in the next generation is $\mathbf{p}(k)$ for each i. We will denote by $\mathbb{P}_{z_0}$ the distribution probability of this process started from z_0 individuals. When we write $\mathbb{P}$ and unless otherwise mentioned, we mean that the initial state is equal to 1.

Thus, we consider an i.i.d. sequence of random environment $(\mathbf{p}_i)_{i\in\mathbb{N}}$ with common distribution μ. Traditionally, the study of BPRE has relied on analytical tools such as generating functions. More precisely, denoting by f_i the probability generating function of $\mathbf{p}_i$, one can note that the BPRE $(Z_n, n \in \mathbb{N})$ is characterized by the relation

$$\mathbb{E}\big(s^{Z_{n+1}} | Z_0, \ldots, Z_n;\ f_0, \ldots, f_n\big) = f_n(s)^{Z_n} \qquad (0 \leq s \leq 1,\ n \geq 0).$$

For classical references on these processes see [1, 2, 3, 6, 15, 23].

A good picture to keep in mind when thinking of a BPRE is the following : consider a population of plants which have a one year life-cycle (so generations are discrete and non-overlapping). Each year the climate or weather conditions (the environment) vary which impacts the reproductive success of the plant. Given the climate, all the plants reproduce according to the same given mechanism. In this context, μ can be thought of as the distribution which controls the successive climates, which are supposed to be iid, and the plant population then obeys a branching process in random environment. By taking a Dirac mass for μ we recover the classical case of Galton Watson processes.

At least intuitively one easily sees that some information on the behavior of the BPRE Z_n can be read from the process $M_n = \Pi_1^n m(\mathbf{p}_i)$ and that their typical behavior should be similar :

$$Z_n \approx M_n, \qquad (n \in \mathbb{N}).$$

Hence the following dichotomy is hardly surprising: A BPRE is supercritical (resp. critical, resp. subcritical) if the expectation of $\log(m(\mathbf{p}))$ with respect to μ the law of the environments :

$$\mathbb{E}(\log(m(\mathbf{p}))),$$

is positive (resp. zero, resp. negative). In the supercrticial case, the BPRE survives with a positive probability, in the critical and subcritical case, it becomes extinct a.s.

Moreover, in the supercritical case, we have the following expected result [3, 16]. Assuming that $\mathbb{E}(\sum_{k\in\mathbb{N}} k^s\mathbf{p}(k)/m(\mathbf{p})) < \infty$ for some $s > 1$, there exists a finite r.v. W such that

$$M_n^{-1}Z_n \stackrel{n\to\infty}{\longrightarrow} W, \qquad \mathbb{P}(W > 0) = \mathbb{P}(\forall n, Z_n > 0).$$

which ensures that conditionally on the non-extinction of $(Z_n)_{n\in\mathbb{N}}$

$$\log(Z_n)/n \to \mathbb{E}(\log(m(\mathbf{p}))) \quad \text{a.s.}$$

This result is a generalization in random environment of the well known Kesten-Stigum Theorem for Galton- Watson processes : let N be the reproduction law of the GW process $(Z_n, n \geq 0)$ and let $m = \mathbb{E}(N)$ be its mean. Assume that $E(N\log_+ N) < \infty$, then

$$W_n := Z_n/(m^n) \stackrel{n\to\infty}{\longrightarrow} W, \qquad \mathbb{P}(W > 0) = \mathbb{P}(\forall n, Z_n > 0).$$

The distribution of W is completely determined by that of N and a natural question concerns the tail behavior of W near 0 and infinity. Results in this direction can be found for instance in [8, 12, 13, 22] for the Galton Watson case and [17] for the BPRE case. In a large deviation context, the tail behavior of W can be related to event where Z_n grows with an atypical rate. Another way to study such events is to consider the asymptotic behavior of Z_{n+1}/Z_n. This is the approach taken in [5] to prove that $|W_n - W|$ decays supergeometrically when $n \to \infty$, assuming that $\mathbb{P}(N = 0) = 0$. Yet another approach is the study of so-called moderate deviations (see [21] for the asymptotic behavior of $\mathbb{P}(Z_n = v_n)$ with $v_n = O(m^n)$).

Finally, we observe that Kesten Stigum Theorem for Galton Watson processes can be reinforced into the following statement:

$$(t \mapsto \frac{1}{n}\log Z_{[nt]}, t \in [0,1]) \Rightarrow (t \mapsto t\log(m), t \in [0,1]).$$

in the sense of the uniform norm almost surely (see for instance [20] for this type of trajectorial results, unconditioned and conditioned on abnormally low growth rates).

In this work we will consider large deviation events for BPREs $A_c(n), c \geq 0$ of the form

$$A_c(n) = \begin{cases} \{0 < \frac{1}{n}\log Z_n \leq c\} \text{ for } c < \mathbb{E}(\log(m(\mathbf{p})) \\ \{\frac{1}{n}\log Z_n \geq c\} \text{ for } c > \mathbb{E}(\log(m(\mathbf{p})) \end{cases},$$

and we are interested in how fast the probability of such events is decaying. More precisely, we are interested in the cases where

$$-\frac{1}{n}\log(\mathbb{P}(A_c(n))) \to \chi(c), \qquad \text{with } \ \chi(c) < \infty.$$

Let us discuss very briefly the Galton-Watson case first (see [14, 20, 22]). Assume first that the Galton Watson process is supercritical ($m := \mathbb{E}(N) > 1$) and and that all the moments of the reproduction law are finite. If we are in the Böttcher case ($\mathbb{P}(N \leq 1) = 0$) then there are no large deviations, i.e.

$$c \neq \log m \ \Rightarrow \ \phi(c) = \infty.$$

If, on the other hand, we are in the Schrödder case ($\mathbb{P}(N = 1) > 0$) then $\phi(c)$ can be non-trivial for $c \leq \log m$. This case is discussed in [20] (see also [14] for finer results for lower deviations) where it is shown that to achieve a lower-than-normal rate of growth $c \leq \log m$ the process first refrains from branching for a long time until it can start to grow at the normal rate $\log m$ and reach its objective. More precisely, it is a consequence of Theorem 2 below that conditionally on $Z_n \leq e^{cn}$,

$$\left(\frac{1}{n}\log(Z_{[nt]}), t \in [0,1]\right) \to (f(t), t \in [0,1])$$

in probability in the sense of uniform norm, where $f(t) = \log(m).(t - (1 - c/\log(m)))_+$. When the reproduction law has infinite moments, the rate function ϕ is non-trivial for $c \geq \log m$. In the critical or subcritical case, there are no large deviations.

We will see that the situation for BPRE differs in many aspects from that of the Galton-Watson case: for instance the rate function is non-trivial as soon as $m(\mathbf{p})$ is not constant and more than 1 with positive probability. This is due to the fact that we can deviate following an atypical sequence of environments, as explained in the next Section, and as already observed by Kozlov for upper values in the supercritical case [18]. When we condition by $Z_n \leq e^{cn}$ and we assume $\mathbb{P}(Z_1 = 1) > 0$ the process $(\frac{1}{n}\log(Z_{[nt]}), t \in [0,1])$ still converges in probability uniformly to a function $f_c(t)$ which has the same shape as f above, that is there exists $t_c \in [0,1]$ such that $f_c(t) = 0$ for $t \leq t_c$ and then f_c is linear and reach c, but the slope of this later piece can now differs from the typical rate $\mathbb{E}(\log m(\mathbf{p}))$.

2 Main results

Denote by $(L_i)_{i\in\mathbb{N}}$ the sequence of iid log-means of the successive environments,

$$L_i := \log(m(\mathbf{p}_i)), \qquad S_n := \sum_{i=0}^{n-1} L_i,$$

and

$$\bar{L} := \mathbb{E}(\log(m(\mathbf{p}))) = \mathbb{E}(L).$$

Define $\phi_L(\lambda) := \log(\mathbb{E}(\exp(\lambda L)))$ the Laplace transform of L and let ψ be the large deviation function associated with $(S_n)_{n\in\mathbb{N}}$:

$$\psi(c) = \sup_{\lambda\in\mathbb{R}}\{c\lambda - \phi_L(\lambda)\}.$$

We briefly recall some well known fact about the rate function ψ (see [11] for a classical reference on the matter). The map $x \mapsto \psi(x)$ is strictly convex and C^∞ in the interior of the set $\{\Lambda'(\lambda), \lambda \in \mathcal{D}_\Lambda^o\}$ where $\mathcal{D}_\Lambda = \{\lambda : \Lambda(\lambda) < \infty\}$. Furthermore, $\psi(\bar{L}) = 0$, and ψ is decreasing (strictly) on the left of $\bar{L}$ and increasing (strictly) on its right.

The map ψ is called the rate function for the following large deviation principle associated with the random walk S_n. We have for every $c \leq \bar{L}$,

$$\lim_{n\to\infty} -\log(\mathbb{P}(S_n/n \leq c)/n = \psi(c), \tag{1}$$

and for every $c \geq \bar{L}$

$$\lim_{n\to\infty} -\log(\mathbb{P}(S_n/n \geq c)/n = \psi(c). \tag{2}$$

Roughly speaking, one way to get

$$\log(Z_n)/n \in O \qquad (n \to \infty)$$

is to follow environments with a good sequence of environments:

$$\log(\Pi_{i=1}^n m(\mathbf{p}_i))/n = S_n/n \in O.$$

We have then the following upper bound for the rate function for any BPRE under a moment condition analogue to that used in [16]. In ecology applications, this corresponds to explaining a rare event by environmental stochasticity.

Proposition 1. *Assuming that* $\mathbb{E}(\sum_{k\in\mathbb{N}} k^s \mathbf{p}(k)/m(\mathbf{p})) < \infty$ *for some* $s > 1$*, then for every* z_0 *:*

- $\forall c \leq \bar{L}$

$$\limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\log(Z_n)/n \leq c) \leq \ \psi(c).$$

- $\forall c \geq \bar{L}$

$$\limsup_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\log(Z_n)/n \geq c) \leq \ \psi(c).$$

As Theorem 2 below shows, the inequality may be strict. Moreover, this proves that even in the subcritical case, there may be large deviations, contrary to what happens in the Galton Watson case. More precisely, as soon as $\mathbb{P}(m(\mathbf{p}) > 1) > 0$ and $m(\mathbf{p})$ is not constant almost surely, the rate function ψ is non trivial on $(0, \infty)$.

As the proof of this result uses classical probability tilting arguments, we only give the main idea and leave the details to the interested reader. Introduce the probability $\widetilde{\mathbb{P}}$ on $\mathbb{P}$ defined by

$$\widetilde{\mathbb{P}}(\mathbf{p} \in \mathrm{d}p) = \frac{m(p)^{\lambda_c}}{\mathbb{E}(m(\mathbf{p})^{\lambda_c})} \mathbb{P}(\mathbf{p} \in \mathrm{d}p),$$

where λ_c is the point where $\lambda \mapsto \lambda c - \phi_L(\lambda)$ reaches its unique maximum. in $[0, 1]$. Under this new probability

$$\widetilde{\mathbb{E}}(\log m(\mathbf{p})) = c > 0,$$

so $S_n = \sum_{i=1}^n \log m(\mathbf{p}_i)$ is a random walk with drift c and Z_n is a supercritical BPRE with survival probability $\widetilde{p} > 0$. Observe that for a measurable function f $\mathbb{E}_{z_0}(f(Z_n)) = \mathbb{E}(m(\mathbf{p})^{\lambda_c})^n \widetilde{\mathbb{E}}_{z_0}(\exp(-\lambda_c S_n) f(Z_n))$. The result follows with $f(z) = \mathbb{1}_{[c-\epsilon, c+\epsilon]}(\log(z)/n)$.

2.1 Lower deviation in the strongly supercritical case.

We focus here on the so-called *strongly supercritical* case

$$\mathbb{P}(\mathbf{p}(0) = 0) = 1$$

(in which the environments are almost surely supercritical). Let us define for every $c \leq \bar{L}$,

$$\chi(c) \quad := \quad \inf_{t\in[0,1]} \{-t \log(\mathbb{E}(\mathbf{p}(1))) + (1-t)\psi(c/(1-t))\}.$$

It is quite easy to prove that this infimum is reached at a unique point t_c by convexity arguments. Thus

$$\chi(c) = -t_c \log(\mathbb{E}(\mathbf{p}(1))) + (1-t_c)\psi(c/(1-t_c)),\ t_c \in [0, 1 - c/\bar{L}].$$

We can thus define the function $f_c : [0,1] \mapsto \mathbb{R}_+$ for each $c < \bar{L}$ as follows (see figure 1):

$$f_c(t) := \begin{cases} 0, & \text{if } t \le t_c \\ \frac{c}{1-t_c}(t - t_c), & \text{if } t \ge t_c. \end{cases}$$

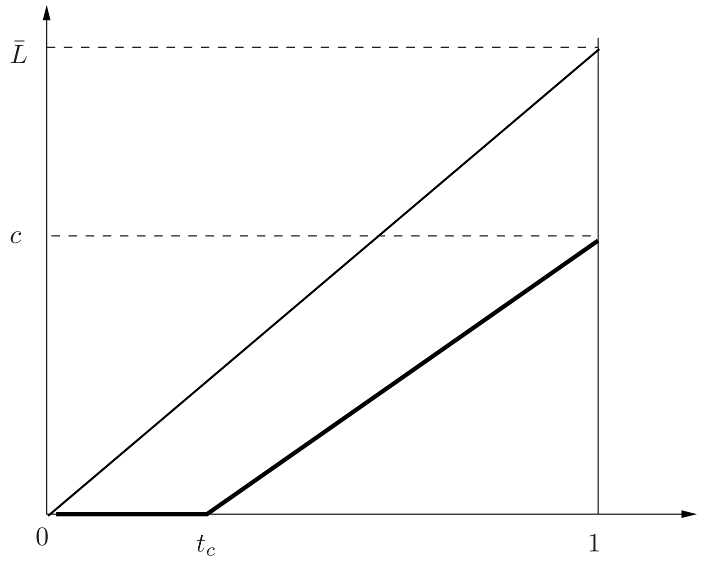

Figure 1: The function $t \mapsto f_c(t)$ for $c \le \bar{L}$.

We will need the following moment assumption $\mathcal{H}$.

$$\left\{ \begin{array}{c} \exists A > 0 \text{ s.t. } \mu(m(\mathbf{p}) > A) = 0, \\ \exists B > 0 \text{ s.t. } \mu(\sum_{k\in\mathbb{N}} k^2 \mathbf{p}(k) > B) = 0 \end{array} \right\} \qquad (\mathcal{H})$$

Observe that the condition in Proposition 1 ($\exists s > 1$ such that $\mathbb{E}(\sum_{k\in\mathbb{N}} k^s \mathbf{p}(k)/m(\mathbf{p})) < \infty$) is included in ($\mathcal{H}$).

The main result is the following theorem which gives the large deviation cost of $Z_n \le \exp(cn)$ and the asymptotic trajectory behavior of Z_n when conditioned on $Z_n \le \exp(cn)$.

Theorem 2. *Assuming that $\mathbb{P}(\mathbf{p}(0) = 0) = 1$ and the hypothesis $\mathcal{H}$ we have*

(a) If $\mu(\mathbf{p}(1) > 0) > 0$, then for every $c < \bar{L}$,

$$-\log(\mathbb{P}(Z_n \le e^{cn}))/n \overset{n\to\infty}{\longrightarrow} \chi(c),$$

and furthermore, conditionally on $Z_n \le e^{cn}$,

$$\sup_{t\in[0,1]} \{\big|\log(Z_{[tn]})/n - f_c(t)\big|\} \overset{n\to\infty}{\longrightarrow} 0, \qquad \textit{in } \mathbb{P}.$$

(b) If $\mu(\mathbf{p}(1) > 0) = 0$, *then for every* $c < \bar{L}$,

$$-\log(\mathbb{P}(Z_n < e^{cn}))/n \stackrel{n\to\infty}{\longrightarrow} \psi(c),$$

and furthermore for every $\inf\{supp \log(m(\mathbf{p}))\} < c < \bar{L}$, *conditionally on* $Z_n \le e^{cn}$,

$$\sup_{t\in[0,1]} \{\big|\log(Z_{[tn]})/n - ct\big|\} \stackrel{n\to\infty}{\longrightarrow} 0, \qquad in\ \mathbb{P}.$$

Let us note that if $\mu(\mathbf{p}(1) > 0) > 0$, then t_c -the take-off point of the trajectory- may either be zero, either be equal to $1 - c/\bar{L}$, or belong to $(0, 1 - c/\bar{L})$ (see Section 3 for examples).

Moreover, when $m := m(\mathbf{p})$ is deterministic, as in the case of a GW process,

- If $\mu(\mathbf{p}(1) > 0) > 0$ (Böttcher case), then $t_c = 1 - c/\log(m)$ and $\chi(c) = t_c \log(\mathbb{E}(\mathbf{p}(1)))$.
- If $\mu(\mathbf{p}(1) > 0) = 0$ (Schrodder case), then $\chi(c) = -\infty$.

Let us first give a heuristic interpretation of the above theorem. Observe that

$$\mathbb{P}(Z_k = 1, k = 1, \dots, tn) = \mathbb{E}(\mathbf{p}(1))^{tn} = \exp(\log(\mathbb{E}(\mathbf{p}(1)))tn)$$

and that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_{(1-t)n}/n \le c) = (1-t)\psi(c/(1-t)).$$

This suggests that

$$\mathbb{P}(Z_k = 1, k = 1, \dots, tn \,; S_n - S_{tn} \le cn) \asymp \exp(n[t\log(\mathbb{E}(\mathbf{p}(1))) + (1-t)\psi(c/(1-t))])$$

and $\chi(c)$ is just the "optimal" cost of such an event with respect to the choice of t. It is not hard to see that the event $\{Z_k = 1, k = 1, \dots, tn \,; S_n - S_{tn} \le cn\}$ is asymptotically included in $\{Z_n \le cn\}$ and hence $\chi(c)$ is an upper bound for the rate function for Z_n. Adding that once $Z_n >> 1$ is large enough it has no choice but to follow the random walk S_n associated to the environment sequence, χ is actually the good candidate to be the rate function.

Thus, roughly speaking, to deviate below c, the process $(\log(Z_{[nt]})/n)_{t\in[0,1]}$ stays bounded until an optimal time t_c and then deviates in straight line to c thanks to a good sequence of environments. The proof in Section 5 and 6 follows this heuristic.

Another heuristic comment concerns the behavior of the environment sequence conditionally on the event $Z_n \le e^{cn}$. Before time $[nt_c]$ we see a sequence of iid environments which are picked according to the original probability law μ biased by $\mathbf{p}(1)$ the probability to have one offspring (think of the case where μ charges only two environments). After time $[nt_c]$ we know that the distribution of the sequence $(L_i)_{i\ge[nt_c]}$ is the law of a sequence of iid L_i conditioned on $\sum_{i=[nt_c]}^{n} L_i \le [nc]$. This implies that the law of the environments is that of an exchangeable sequence with common distribution μ tilted by the log-means.

To conclude this section, we comment on the hypothesis $\mathbb{P}(\mathbf{p}(0)=0)=1$. It is known (see [6]) that for a Galton Watson process Z_n with survival probability p and generating function f, under the $N\log N$ condition, for all $j\in\mathbb{N}$

$$\gamma^{-n}\mathbb{P}(Z_n=j)\to\alpha_j \tag{*}$$

where $\forall j\in\mathbb{N}:\alpha_j\in(0,\infty)$ and $\gamma=f'(p)$. In the case where $\mathbb{P}(Z_1=0)=0$ (no death), $\gamma=f'(p)=f'(0)=\mathbb{P}(Z_1=1)$ which tells us that the cost of staying bounded is the cost of keeping the population size fixed at 1, a fact that we also use for our analysis of BPRE. This suggests that the analogue of γ for BPRE should also play a role in the lower deviations events when $\mathbb{P}(\mathbf{p}(0)=0)<1$. However there is not yet an analogue of $(*)$ for BPRE and the situation is probably more complex.

2.2 Upper deviation in the strongly supercritical case

Assume as above that

$$\mathbb{P}(\mathbf{p}(0)=0)=1,$$

and that for every $k\geq 1$,

$$\mathbb{E}(Z_1^k)<\infty,$$

we have the following large deviation result for upper values.

Theorem 3. *For every $c>\bar{L}$,*

$$-\frac{1}{n}\log(\mathbb{P}(Z_n\geq e^{cn}))\overset{n\to\infty}{\longrightarrow}\psi(c),$$

and furthermore for $c<\sup\{supp\log(m(\mathbf{p}))\}$, conditionally on $Z_n\geq\exp(cn)$,

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

To put it in words, this says that the cost of achieving a higher than normal rate of growth is just the cost of seeing an atypical sequence of environments in which this rate is expected. Furthermore, conditionally on $Z_n\geq e^{cn}$, the trajectory $(\log(Z_{[nt]})/n)_{t\in[0,1]}$ is asymptotically a straight line.

Kozlov [18] gives the upper deviations of Z_n in the case where the generating functions f are a.s. linear fractional and verify a.s. $f''(1)=2f'(1)^2$. In the strongly supercritical case and under those hypothesis, he proves that for every $\theta>0$, there exists $I(\theta)>0$ such that

$$\mathbb{P}(\log(Z_n)\geq\theta n)\sim I(\theta)\mathbb{P}(S_n\geq\theta n),\quad(n\to\infty).$$

Thus, Kozlov gets a finer result in the linear fractional case with $f''(1)=2f'(1)^2$ a.s. by proving that the upper deviations of the BPRE Z_n above $\bar{L}$ are exactly given by the large deviations of the random walk S_n.

Proposition 1 shows that the rates of upper and lower deviations are at least those of the environments, but Theorem 2 and the remark below show that the converse is not always true.

Theorem 3 is the symmetric for upper deviations of case (b) of Theorem 2 for lower deviations. It is natural to ask if there is an analogue of case (a) as well. In this direction, we make the following two remarks.

- If there exists $k > 1$ such that
$$\mathbb{E}(Z_1^k) = \infty,$$
then the cost of reaching c can be less that $\psi(c)$, since the BPRE might "explode" to a very large value in the first generation and then follow a geometric growth. This mirrors nicely what happens for lower deviations in the case (a). However we do not have an equivalent of Theorem 2 for upper deviations as such a result seems much harder to obtain for now.

- In the case when
$$\mathbb{P}(m(\mathbf{p}) < 1) > 0,$$
then by Theorem 3 in [16],
$$s_{\max} := \sup_{s \geq 1} \{\mathbb{E}(W^s) < \infty\} < \infty.$$
Thus, the BPRE $(Z_n)_{n\in\mathbb{N}}$ might deviate from the exponential of the random walk of environments :
$$\lim_{n\to\infty} -\log(\mathbb{P}(\exp(-S_n)Z_n \geq \exp(n\epsilon))/n < \infty, \quad (\epsilon > 0),$$
which would yield a more complicated rate function for deviations.

2.3 No large deviation without supercritical environment

Finally, we consider the case when environments are a.s. subcritical or critical :
$$\mathbb{P}(m(\mathbf{p}) \leq 1) = 1,$$
and we assume that for every $j \in \mathbb{N}$, there exists $M_j > 0$ such that
$$\sum_{k=0}^{\infty} k^j \mathbf{p}(k) \leq M_j \quad \text{a.s.} \qquad (\mathcal{M}).$$
Note that the condition $(\mathcal{M})$ implies $(\mathcal{H})$ simply by considering $j = 2$.

In that case, even if $\mathbb{P}_1(Z_1 \geq 2) > 0$, there is no large deviation, as in the case of a Galton Watson process.

Proposition 4. *Suppose $(\mathcal{M})$ and that $\mathbb{P}(m(\mathbf{p}) \leq 1) = 1$, then for every $c > 0$,*
$$\lim_{n\to\infty} -\log(\mathbb{P}(Z_n \geq \exp(cn))/n = \infty.$$

We recall that by Proposition 1, this result does not hold if $\mathbb{P}(m(\mathbf{p}) > 1) > 0$.

The next short section shows a concrete example where t_c is non trivial. Section **??** is devoted to the proof of Proposition 1. Section 4 is devoted to proving two key lemmas which are then used repeatedly. The first gives the cost of keeping the population bounded for a long time. The second tells us that once the population passes a threshold, it grows geometrically following the product of the means of environments. In Section **??**, we start by computing the rate function and then we describe the trajectory. Section 5 is devoted to upper large deviation while Section 6 to case when environments are a.s. subcritical or critical.

3 A motivating example : the case of two environments

Suppose we have two environments P_1 and P_2 with $\mu(\mathbf{p} = P_1) = q$. Call $L_1 = \log m(P_1)$ and $L_2 = \log m(P_2)$ their respective log mean and suppose $L_1 < L_2$. The random walk S_n is thus the sum of iid variables $X : \mathbb{P}(X = L_1) = q, \mathbb{P}(X = L_2) = 1 - q$.

Recall that if X is a Bernoulli variable with parameter p the Fentchel Legendre transform of $\Lambda(\lambda) = \log(\mathbb{E}(e^{\lambda X}))$ is

$$\Lambda^*(x) = x \log(x/p) + (1 - x) \log((1 - x)/(1 - p)).$$

Hence the rate function for the large deviation principle associated to the random walk S_n is defined for $L_1 \leq x \leq L_2$ by

$$\psi(x) = z \log(z/p) + (1 - z) \log((1 - z)/(1 - p)) \text{ where } z = \frac{x - L_1}{L_2 - L_1}.$$

Recall that $\mathbb{E}(\mathbf{p}(1)) = qP_1(1) + (1 - q)P_2(1)$ is the probability that an individual has exactly one descendent in the next generation.

The following figure 2 shows the function $t \mapsto -t \log(\mathbb{E}(\mathbf{p}(1))) + (1 - t)\psi(c/(1 - t))$, so $\chi(c)$ is the minimum of this function and t_c is the t where this minimum is reached. Figure 2 is drawn using the values $L_1 = 1$, $L_2 = 2$, $q = .5$, $\mathbb{E}(\mathbf{p}(1)) = .4$, $c = 1.1$ and $1 - c/\bar{L} \sim .27$. Thus, we ask $Z_n \leq e^{1.1n}$ whereas Z_n behaves normally as $e^{1.5n}$ and this example illustrate Theorem 2 a) with $t_c \in (0, 1 - c/\bar{L})$.

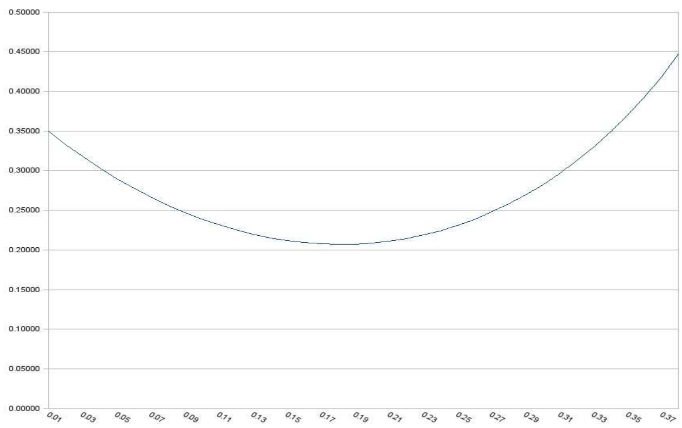

Figure 2: In this example $t_c \sim 0.18$, the slope of the function f_c after t_c is 1.34.

As an illustration and a motivation we propose the following model for parasites infection. In [7], we consider a branching model for parasite infection with cell division. In every generation, the cells give birth to two daughter cells and the cell population is the binary tree. We want to take into account unequal sharing of parasites, following experiments made in Tamara's Laboratory in Hopital Necker (Paris), and we distinguish a first (resp. a second) daughter cell. Then we call $Z^{(1)}$ (resp. $Z^{(2)}$) the number of descendants of a given parasite of the mother cell in the first (resp. the second daughter), where $(Z^{(1)}, Z^{(2)})$ is any couple of random variable (it may be non symmetric, dependent...). A key role for limit theorems is played by the process $(Z_n)_{n\in\mathbb{N}}$ which gives the number of parasites in a random cell line (choosing randomly one of the two daughter cells at each cell division and counting the number of parasites inside). This process follows a Branching process with two equiprobable environment with respective reproduction law $Z^{(1)}$ and $Z^{(2)}$. Thus, here $q = 1/2$, $L_1 = \log(\mathbb{E}(Z^{(1)}))$ and $L_2 = \log(\mathbb{E}(Z^{(2)}))$.

We are interested in determining the number of cells with a large number of parasites and we call $N_n^{\leq c}$ (resp $N_n^{\geq c}$) the number of cells in generation n which contain less (resp. more) than $\exp(cn)$ parasites, for $c > 0$. An easy computation (which follows (17) in [7]) shows that

$$\mathbb{E}(N_n^{\leq c}) = 2^n \mathbb{P}(Z_n \leq \exp(cn)), \quad \mathbb{E}(N_n^{\geq c}) = 2^n \mathbb{P}(Z_n \geq \exp(cn)).$$

If $\mathbb{P}(Z^{(1)} = 0) = \mathbb{P}(Z^{(2)} = 0) = 1$, Section 2.1 ensures that for every $c \geq \sqrt{\mathbb{E}(Z^{(1)})\mathbb{E}(Z^{(2)})}$,

$$\lim_{n\to\infty} \log(\mathbb{E}(N_n^{\leq c}))/n = \log(2) - \chi(c).$$

Moreover Section 2.2 ensures that for every $c \geq \sqrt{\mathbb{E}(Z^{(1)})\mathbb{E}(Z^{(2)})}$,

$$\lim_{n\to\infty} \log(\mathbb{E}(N_n^{\leq c}))/n = \log(2) - \psi(c).$$

4 Lower deviations: Proof of Theorem 2

Let us briefly describe the main steps which compose the proof of Theorem 2. First, we compute the rate of decrease of the probability that the population remains bounded. We then prove that when the population is large enough, its growth is given by the product of mean of the successive environments. This allows us to give both the rate function for large deviations and the point where the trajectory of the process under the conditional event takes off. Finally, we describe the full asymptotic path conditionally on the large deviation event.

4.1 The cost of staying bounded

We start with the following elementary result, which says that staying bounded has the same exponential cost than staying fixed at 1.

Lemma 5. *For every $N \geq 1$,*

$$\lim_{n\to\infty} \log(\mathbb{P}(Z_n \leq N))/n = \log(\mathbb{E}(\mathbf{p}(1))).$$

Moreover, if $\mathbb{E}(\mathbf{p}(1)) > 0$, then for every fixed N there is a constant C such that for every $n \in \mathbb{N}$,

$$\mathbb{P}(Z_n \leq N) \leq C n^N \mathbb{E}(\mathbf{p}(1))^{n+1}.$$

Proof. We call $(N_i)_{i\geq 1}$ the number of offspring of a random lineage. More explicitly, we call N_0 the size of the offspring of the ancestor in generation 0 and pick uniformly one individual among this offspring. We call N_1 the size of the offspring of this individual and so on...

Note that $(N_i)_{i\geq 1}$ are iid with common ditribution $\mathbb{P}(N=k)=\mathbb{E}(\mathbf{p}(k))$. Hence, for every $n\geq N$, recalling that $\mathbb{P}(p(0)=0)=1$,

$$\begin{aligned}\mathbb{P}(Z_n\leq N)&\leq \mathbb{P}(\text{less than } N \text{ of the } (N_i)_{0\leq i\leq n-1} \text{ are} > 1)\\ &\leq \sum_{k=0}^{N}\binom{n}{k}(1-\mathbb{E}(\mathbf{p}(1)))^k\mathbb{E}(\mathbf{p}(1))^{n-k}\\ &\leq (N+1)n^N\mathbb{E}(\mathbf{p}(1))^{n-N}.\end{aligned}$$

Adding that

$$\mathbb{P}(Z_n\leq N)\geq \mathbb{P}(Z_n=1)=\mathbb{E}(\mathbf{p}(1))^n,$$

allows us to conclude. □

Our proof actually shows the stronger

$$\lim_{n\to\infty}\log(\mathbb{P}(Z_n\leq n^a))/n=\log(\mathbb{E}(\mathbf{p}(1))),$$

for $a\in(0,1)$.

4.2 The cost of deviating from the environments

The aim of this section is to show that once the process "takes off" (i.e. once the population passes a certain threshold), it has to follow the products of the means of the environments sequence.

Lemma 6. *Assuming $\mathcal{H}$, for all $\epsilon>0$ and $\eta>0$, there exist $N,D\in\mathbb{N}$ such that for all $z_0\geq N$ and $n\in\mathbb{N}$,*

$$\mathbb{P}_{z_0}(Z_n\leq z_0\exp(S_n-n\epsilon)\mid (\mathbf{p}_i)_{i=0}^{n-1})\leq D\eta^n\quad a.s.$$

so that

$$\mathbb{P}_{z_0}(Z_n\leq \exp(S_n-n\epsilon))\leq D\eta^n.$$

Define for every $0\leq i\leq n-1$,

$$R_i:=Z_{i+1}/Z_i,$$

so that

$$Z_n=Z_0\Pi_{i=0}^{n-1}R_i.$$

For all $\lambda\geq 0$, $q\in\mathbb{N}$ and $0\leq i\leq n-1$ define the function

$$\Lambda_q(\lambda,p):=\mathbb{E}\big(\exp(\lambda[L_i-\epsilon-\log(R_i)])\mid \mathbf{p}_i=p,\ Z_i=q\big),$$

(this quantity does not depend on i by Markov property) and

$$M_N(\lambda,p):=\sup_{q\geq N}\Lambda_q(\lambda,p).$$

The proof will use the following Lemma, the proof of which is given at the end of this section.

Lemma 7. *Fix $\epsilon > 0$, there exist $\alpha \in (0,1), \lambda_0 \in (0,1)$ and $N \in \mathbb{N}$ such that*

$$M_N(\lambda_0, \mathbf{p}) \leq 1 - \alpha \qquad a.s.$$

where $\mathbf{p}$ is a random probability with law μ.

We proceed with the proof of Lemma 6 assuming that the above result holds.

Proof. For every $\lambda > 0, \epsilon > 0, k : in\mathbb{N}$, by Markov inequality

$$\begin{aligned}
&\mathbb{P}_{z_0}\big(Z_n \leq kz_0 \exp(S_n - n\epsilon) \mid (\mathbf{p}_i)_{i=0}^{n-1}\big) \\
&\quad = \mathbb{P}_{z_0}\big(z_0 \Pi_{i=0}^{n-1} R_i \leq kz_0 \exp(\sum_{i=0}^{n-1}[L_i - \epsilon]) \mid (\mathbf{p}_i)_{i=0}^{n-1}\big) \\
&\quad \leq k^\lambda \; \mathbb{E}_{z_0}\big(\exp\{\lambda \sum_{i=0}^{n-1}[L_i - \epsilon - \log R_i]\} \mid (\mathbf{p}_i)_{i=0}^{n-1}\big).
\end{aligned}$$

Observe that conditionally on $\mathbf{p}_j, R_j$ depends on $(\mathbf{p}_i)_{i=0}^{j}$ and $(Z_0, R_0, R_1, \ldots, R_{j-1})$ only through Z_j. Furthermore, under $\mathbb{P}_{z_0}$ we have that almost surely $\forall n \in \mathbb{N} : Z_n \geq z_0$ since $\mathbb{P}(\mathbf{p}(0) > 0) = 0$. Hence we get for every $\lambda \geq 0$,

$$\begin{aligned}
&\mathbb{P}_{z_0}\big(Z_n \leq kz_0 \exp(S_n - n\epsilon) \mid (\mathbf{p}_i)_{i=0}^{n-1}\big) \\
&\leq k^\lambda \; \mathbb{E}_{z_0}\Big\{ \exp(\lambda \sum_{i=0}^{n-2}[L_i - \epsilon - \log(R_i)]) \\
&\qquad \times \mathbb{E}_{z_0}\big[\exp(\lambda[L_{n-1} - \epsilon - \log(R_{n-1})]) \mid \mathbf{p}_{n-1}, Z_{n-1}\big] \;\Big|\; (\mathbf{p}_i)_{i=0}^{n-1}\Big\} \\
&\leq k^\lambda \; \mathbb{E}_{z_0}\big(\exp(\lambda \sum_{i=0}^{n-2}[L_i - \epsilon - \log(R_i)]) \mid (\mathbf{p}_i)_{i=0}^{n-2}\big) M_{z_0}(\lambda, \mathbf{p}_{n-1}) \\
&\leq \ldots \\
&\leq k^\lambda \; \Pi_{i=0}^{n-1} M_{z_0}(\lambda, \mathbf{p}_i).
\end{aligned}$$

Fix $\epsilon > 0$, by Lemma 7 we can find $\alpha \in (0,1), \lambda_0 \in (0,1)$ and $\exists N \in \mathbb{N}$ such that almost surely $\forall i \in \mathbb{N}$, $M_N(\lambda_0, \mathbf{p}_i) \leq 1 - \alpha$. Hence, for all $z_0 \geq N, k \in \mathbb{N}$ we have,

$$\mathbb{P}_{z_0}(Z_n \leq kz_0 \exp(S_n - n\epsilon) \mid (\mathbf{p}_i)_{i=1}^{n}) \leq k^{\lambda_0} \prod_{i=1}^{n} M_{z_0}(\lambda_0, \mathbf{p}_i) \leq k^{\lambda_0}(1-\alpha)^n \quad \text{a.s.} \tag{3}$$

Fix now $\eta > 0$ and chose $k \in \mathbb{N}$ such that $(1-\alpha)^k \leq \eta$. By starting with an initial population of size larger than kN and dividing it up in subgroups of size at least N which then evolve independently, we see that Lemma 6 follows easily from(3).

□

We now prove Lemma 7.

Proof. Observe that the $(\Lambda_q(\lambda, \mathbf{p}_i))_{i\in\mathbb{N}}$ are iid with common distribution

$$\Lambda_q(\lambda) = \Lambda_q(\lambda, \mathbf{p}_0) = \mathbb{E}(\exp(\lambda[L_0 - \epsilon - \log R_0]) \,|\mathbf{p}_0,\ Z_0 = q).$$

By Taylor's formula, for every $\lambda \geq 0$, there exists $c_\lambda \in [0, \lambda]$ such that

$$\Lambda_q(\lambda) = 1 + \lambda\mathbb{E}\big(L_0 - \epsilon - \log(R_0) \mid \mathbf{p}_0,\ Z_0 = q\big) + \lambda^2\Lambda_q''(c_\lambda). \tag{4}$$

Conditionally on $\mathbf{p}_0 = p$ and $Z_0 = q$, $R_0 = q^{-1}\sum_{i=1}^q X_j$ where the X_j are iid with distribution p. We then remark that since $R_0 \geq 1$ and $\exp L_0 \geq 1$ a.s., using $|\log(R_0) - L_0| < |R_0 - \exp L_0|$ and Cauchy-Schwartz inequality

$$\big|\mathbb{E}\big[\log(R_0) - L_0 \mid \mathbf{p}_0, Z_0 = q\big]\big| \leq \left(\frac{1}{q}\mathrm{Var}_{\mathbf{p}_0}\right)^{1/2}.$$

By hypothesis $\mathcal{H}, \mathrm{Var}_{\mathbf{p}_0}$ is bounded so there exists $N \in \mathbb{N}$ such that for every $q \geq N$,

$$\big|\mathbb{E}\big[\log(R_0) - L_0|\mathbf{p}_0, Z_0 = q\big]\big| \leq \epsilon/2 \qquad \text{a.s.}$$

Next, in order to bound the second order term $\Lambda_q''(\lambda)$, using $\log x \leq x - 1$ when $x > 0$, we note that

$$\begin{aligned}\mathbb{E}\big(\log(R_0)^2 \mid \mathbf{p}_0,\ Z_0 = q\big) &\leq 1 + \frac{2}{q^2}\mathbb{E}\big(\sum_{j=1}^q X_j^2 \mid \mathbf{p}_0,\ Z_0 = q\big)\\ &\leq 1 + \frac{2}{q}B \qquad \text{a.s.,}\end{aligned}$$

where B is the constant from $\mathcal{H}$. Hence, for any $\lambda \in [0, 1]$,

$$\begin{aligned}\Lambda_q''(\lambda) &= \mathbb{E}\Big[(L_0 - \epsilon - \log R_0)^2 e^{\lambda(L_0-\epsilon-\log R_0)}\big|\ \mathbf{p}_0,\ Z_0 = q\Big]\\ &\leq e^{L_0}\mathbb{E}\Big[(L_0 - \epsilon - \log(R_0))^2 \mid \mathbf{p}_0,\ Z_0 = q\Big]\\ &\leq 4A\Big[(\log A)^2 + \epsilon^2 + \mathbb{E}\big(\log(R_0)^2 \mid \mathbf{p}_0,\ Z_0 = q\big)\Big] \qquad \text{a.s.,}\end{aligned}$$

where A is the constant from $\mathcal{H}$. Thus we conclude that for all $\lambda \in [0, 1]$ and $q \in \mathbb{N}$

$$\Lambda_q''(\lambda) \leq M \qquad \text{a.s.,}$$

where M is a finite constant. Then, for all $q \geq N$ and $\lambda \in [0, 1]$,

$$\Lambda_q(\lambda) \leq 1 - \lambda\epsilon/2 + \lambda^2 M \qquad \text{a.s.,}$$

and thus

$$M_N(\lambda, \mathbf{p}_0) \leq 1 - \lambda\epsilon/2 + \lambda^2 M \qquad \text{a.s. .}$$

Choose now $\lambda_0 \in (0, 1]$ small enough such that $\lambda_0\epsilon/2 - \lambda_0^2 M = \alpha > 0$, then $M_N(\lambda_0, \mathbf{p}_0) \leq 1 - \alpha$ *a.s.* This ends up the proof of Lemma 7. □

4.3 Deviation cost and take-off point

For each $c < \bar{L}$, we start by giving the rate function for lower deviations and we prove that $(Z_{[nt]})_{t\in[0,1]}$ begins to take large values at time t_c. We then show that no jump occur at time t_c and that $(\log(Z_{[nt]})/n)_{t\in[t_c,1]}$ grows linearly to complete the proof of Theorem 2.

We consider the first time at which the population reaches the threshold N

$$\tau(N) := \inf\{k : Z_k > N\}, \quad \tau_n(N) = \min(\tau(N), n).$$

Recalling that

$$\chi(c) = \inf_{t\in[0,1-c/\bar{L}]} \{-t\log(\mathbb{E}(\mathbf{p}(1))) + (1-t)\psi(c/(1-t))\}$$

and that t_c is the unique minimizer,we have the following statement.

Proposition 8. *For each $c < \bar{L}$,*

$$\lim_{n\to\infty} -\frac{1}{n}\log\mathbb{P}(\log(Z_n)/n \leq c) = \chi(c).$$

Furthermore, for N large enough, conditionally on $Z_n \leq e^{cn}$,

$$\tau_n(N)/n \overset{n\to\infty}{\longrightarrow} t_c \quad \textit{in } \mathbb{P}.$$

For the proof, we need the following lemma, which tells us that once the population is above N, the cost of a deviation for $(Z_n)_{n\geq 0}$ is simply the cost of the necessary sequence of environments, i.e. the deviation cost for the random walk $(S_n)_{n\geq 0}$.
By decomposing the total probability cost of reaching nc in two pieces (staying bounded until time nt and then having $(S_n - S_{t_c n}) \simeq nc$) and then minimizing over t gives us the correct rate function. The unicity of this minimizer t_c ensures then that the take-off point $\tau_n(N)/n$ converges to t_c.

Lemma 9. *Assume $\mathcal{H}$.*

(i) For each $\eta > 0, \epsilon > 0$, there exists $D, N \in \mathbb{N}$ such that for all $c \geq 0$, $z_0 \geq N$ and $n \in \mathbb{N}$,

$$\mathbb{P}_{z_0}\big(Z_n \leq z_0\exp(cn)\big) \leq D(\eta^n + \exp(-n\psi^*(c+\epsilon))),$$

where $\psi^*(x) = \psi(x)$ for $x \leq \bar{L}$ and $\psi^*(x) = 0$ for $x \geq \bar{L}$.

(ii) For every $\epsilon > 0$ and for every $c_0 \leq \bar{L} - \epsilon$ such that $\psi(c_0) < \infty$, there exists N such that for all $z_0 \geq N$ and $c \in [c_0, \bar{L} - \epsilon]$,

$$\liminf_{n\to\infty} -\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n \leq z_0 e^{cn}) \geq \psi(c+\epsilon)$$

and

$$\limsup_{n\to\infty} -\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n \leq z_0 e^{cn}) \leq \psi(c).$$

Proof. For each $z_0 \in \mathbb{N}, c \le \bar{L}, n \in \mathbb{N}$ and $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}_{z_0}\big(Z_n \le z_0 \exp(cn)\big) \\
&\quad \le \mathbb{P}_{z_0}\big(Z_n \le z_0 \exp(cn),\ S_n - n\epsilon \ge cn\big) + \mathbb{P}_{z_0}\big(S_n - n\epsilon \le cn\big) \\
&\quad \le \mathbb{P}_{z_0}\big(Z_n \le z_0 \exp(S_n - n\epsilon)\big) + \mathbb{P}_{z_0}\big(S_n \le (c+\epsilon)n\big).
\end{aligned}$$

We bound this using Lemma 6 and the following classical fact (see [11]): If $c \le \bar{L}$,

$$\forall n \in \mathbb{N} : \mathbb{P}(S_n \le nc) \le \exp(-n\psi(c)) \tag{5}$$

Thus, there exist $D, N := D(\epsilon, \eta), N(\epsilon, \eta)$ such that for all $c \le \bar{L} - \epsilon$, $z_0 \ge N$,

$$\mathbb{P}_{z_0}\big(Z_n \le z_0 \exp(cn)\big) \le D\eta^n + \exp(-n\psi^*(c+\epsilon)).$$

which yields (i).

The first part of (ii) is an easy consequence of (i) by taking $\eta < \inf\{\exp(-\psi(c)), c \in [c_0, \bar{L} - \epsilon]\}$. The second part comes directly from Proposition 1. □

Proof of Proposition 8. If $\mathbb{E}(\mathbf{p}(1)) = 0$ then $\chi(c) = \psi(c)$ and $t_c = 0$. Noting that $Z_n \ge 2^n$ *a.s.* gives directly the second part of the lemma, while the first part follows essentially from Lemma 9 (ii).

We suppose now that $\mathbb{E}(\mathbf{p}(1)) > 0$. For each $c \le \bar{L}$ and $i = 1, \ldots, n$, we have for every $z_0 \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(\tau_n(N) = i,\ \ Z_n \le \exp(cn)) &\le \mathbb{P}(Z_{i-1} \le N)\mathbb{P}_N(Z_{n-i} \le \exp(cn)) \\
&\le \mathbb{P}(Z_{i-1} \le N)\mathbb{P}_N(Z_{n-i} \le N\exp(cn)).
\end{aligned}$$

Using Lemma 9 and Lemma 5, for all $\eta > 0$ and $\epsilon > 0$, there exists $N, M \in \mathbb{N}$ such that for all $z_0 \ge N$,

$$\mathbb{P}(\tau_n(N) = i,\ \ Z_n \le \exp(cn)) \le Mn^N\mathbb{E}(\mathbf{p}(1))^i[\eta^{n-i} + \exp(-(n-i)\psi^*(cn/(n-i)+\epsilon)].$$

Summing over i leads to

$$\begin{aligned}
\mathbb{P}(\log(Z_n)/n \le c) &= \sum_{i=1}^{n} \mathbb{P}(\tau_n(N) = i,\ \log(Z_n)/n \le c) \\
&\le \sum_{i=1}^{n} Mn^N\mathbb{E}(\mathbf{p}(1))^i[\eta^{n-i} + \exp(-(n-i)\psi^*(cn/(n-i)+\epsilon)].
\end{aligned}$$

Thus, using standard inequalities and that $a^n + b^n \le (a+b)^n$ when $a, b \ge 0$, we see that

$$\begin{aligned}
&\liminf_{n\to\infty} -\frac{1}{n}\log\mathbb{P}(\log(Z_n)/n \le c) \\
&\quad \ge \inf_{t\in[0,1]} \{-t\log(\mathbb{E}(\mathbf{p}(1))) - (1-t)\log(\eta + \exp(-\psi^*(c/(1-t)+\epsilon)))\}.
\end{aligned}$$

This yields the results by letting $\eta, \epsilon \to 0$, and noting that the infimum is necessarily reached on $[0, 1 - c/\bar{L}]$ (since $\psi^*(c/(1-t)) = 0$ as soon as $t \geq 1 - c/\bar{L}$).

More generally, given $0 \leq a < b \leq 1$ it is an easy adaptation of the above argument to show that

$$\begin{aligned} \liminf -\frac{1}{n} \log \mathbb{P}(\log(Z_n)/n \leq c, \tau_n(N)/n \in [a, b]) \\ \geq \inf_{t \in [a,b] \cap [0, 1-c/\bar{L}]} \{-t \log \mathbb{E}(\mathbf{p}(1)) + (1-t)\psi(c/(1-t))\} . \end{aligned} \tag{6}$$

The upper bound is much easier since it is enough to exhibit a trajectory having $\chi(c)$ as it asymptotic cost. By construction it should be clear that

$$\mathbb{P}(Z_{[t_c n]} = 1, Z_n \leq e^{cn}) = \mathbb{P}(Z_{[t_c n]} = 1)\mathbb{P}(Z_{n-[t_c n]} \leq e^{cn})$$

By Lemma 5 and Proposition 1,

$$\begin{aligned} \limsup_n -\frac{1}{n} \log \mathbb{P}(Z_{[t_c n]} = 1, Z_n \leq e^{cn}) &\leq -t_c \log \mathbb{E}(\mathbf{p}(1)) + (1-t_c)\psi(c/(1-t_c)) \\ &= \chi(c). \end{aligned}$$

Combining this inequality with the lower bound given by (**??**), this concludes the proof of the first point of Proposition 8.

For the convergence of $\tau_n(N)/n \to t_c$, observe that as t_c is the unique minimizer of $t \in [0, 1] \mapsto \{-t \log \mathbb{E}(\mathbf{p}(1)) + (1-t)\psi(c/(1-t))\}$, if $t_c \notin (a, b)$ we have

$$\inf_{t \in [a,b] \cap [0, 1-c/\bar{L}]} \{-t \log \mathbb{E}(\mathbf{p}(1)) + (1-t)\psi(c/(1-t))\} > \chi(c).$$

This means by (**??**) and (6) that conditionally on $Z_n \leq e^{cn}$ the event $\tau_n(N)/n \in (a, b)$ becomes negligible with respect to the event $\tau_n(N)/n \in [t_c - \epsilon, t_c + \epsilon]$ for any $\epsilon > 0$. This proves that $\tau_n(N)/n \to_p t_c$. □

Proposition 8 already proves half of Theorem 2. We now proceed to the proof of the path behavior. Define a process $t \mapsto Y^{(n)}(t)$ for $t \in [0, 1]$ by

$$Y_t^{(n)} = \frac{1}{n} \log(Z_{[nt]}).$$

The second part of Theorem 2 tells us that $Y^{(n)}(t)$ converges to f_c in probability in the sense of the uniform norm. To prove this we need two more ingredients, first we need to show that after time $\tau_n(N)/n \simeq t_c$ the trajectory of $Y^{(n)}(t)$ converges to a straight line (this is the object of the following section 4.4) and then that $Y^{(n)}$ does not jump at time $\tau_n(N)/n$ (in section 4.5).

4.4 Trajectories in large populations

The following proposition shows that for a large enough initial population and conditionally on $Y^{(n)}(1) < c$ the process $Y^{(n)}$ converges to the deterministic function $t \mapsto ct$.

Proposition 10. *For all $c < \bar{L}$ and $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that for $z_0 \geq N$,*

$$\lim_{n\to\infty} \mathbb{P}_{z_0}\big(\sup_{x\in[0,1]} \{|Y^{(n)}(x) - cx| \geq \epsilon \mid Z_n \leq z_0 \exp(cn)\big) = 0.$$

Before the proof, let us give a little heuristic of this result. Informally, for all $t \in (0,1)$ and $\epsilon > 0$,

$$\mathbb{P}_{z_0}(Y_t^{(n)} = c+\epsilon,\ Z_n \leq \exp(cn)) = \mathbb{P}_{z_0}(Z_{[nt]} = \exp(tn(c+\epsilon)))\mathbb{P}_{\exp(tn(c+\epsilon))}(Z_{n-[nt]} \leq \exp(cn)).$$

Then, for z_0 large enough, Lemma 9 ensures that

$$\begin{aligned} \lim_{n\to\infty} -\log(\mathbb{P}_{z_0}(Y_t^{(n)} &= c+\epsilon,\ Z_n \leq \exp(cn)))/n \\ &= t\psi(c+\epsilon) + (1-t)\psi(c - \epsilon t/(1-t)) \qquad (7)\\ &> \psi(c), \end{aligned}$$

by strict convexity of ψ. Adding that $\limsup_n -\frac{1}{n}\log \mathbb{P}_{z_0}(Z_n \leq z_0 e^{cn}) \leq \psi(c)$ by Proposition 1 entails that the probability of this event becomes negligible as $n \to \infty$.

Proof. Observe that $\{\exists x \in [x_0,x_1] : Y^{(n)}(x) > cx+\epsilon\} = \{\exists x \in [x_0,x_1] : Y^{(n)}(x) \in (cx+\epsilon, \bar{L}x]\}$, because a.s. $t \mapsto Y^{(n)}(t)$ is an increasing function so that the only way $Y^{(n)}$ can cross $x \mapsto \bar{L}x$ downward is continuously. Hence we can divide the proof in the following steps :

- (i) There exists $0 < x_0 < x_1 < 1$ such that for every $\epsilon > 0$ and for z_0 large enough $\lim_{n\to\infty} \mathbb{P}_{z_0}\big(\sup_{x\notin[x_0,x_1]}\{|Y^{(n)}(x) - cx| \geq \epsilon \mid Z_n \leq z_0\exp(cn)\big) = 0$ (see Figure 4.4).
- (ii) We show that for z_0 large enough $\lim_{n\to\infty} \mathbb{P}_{z_0}\big(\exists x \in [x_0,x_1] : cx+\epsilon \leq Y^{(n)}(x) \leq \bar{L}x \mid Z_n \leq z_0\exp(cn)\big) = 0$.
- (iii) The fact that for z_0 large enough $\lim_{n\to\infty} \mathbb{P}_{z_0}\big(\exists x \in [x_0,x_1] : Y^{(n)}(x) \leq cx-\epsilon \mid Z_n \leq z_0\exp(cn)\big) = 0$ then follows from the same arguments as in (ii).

We start by proving (ii) which is the key point. We can assume $\epsilon < (\bar{L}-c)x_0$ and $\epsilon < (\bar{L}-c)(1-x_1)$ and we define

$$R_c := \{(x,y) : x \in [x_0,x_1], y \in [cx+\epsilon, \bar{L}x]\}.$$

We know from Lemma 9 that $\limsup_{n\to\infty} -\frac{1}{n}\log \mathbb{P}_{z_0}(Z_n \leq z_0\exp(cn)) \leq \psi(c)$ (for z_0 large enough). Hence, we will have proved the result if we show that for z_0 large enough

$$\liminf_{n\to\infty} -\frac{1}{n}\log \mathbb{P}_{z_0}\big(\exists x \in [x_0,x_1] : (x, Y^{(n)}(x)) \in R_c, Z_n \leq z_0 e^{cn}\big) > \psi(c). \qquad (8)$$

Lemma 9 or heuristic (7) suggest that the asymptotic cost of the event $\{Y^{(n)}(x) = y, Y^{(n)}(1) < c\}$ is given by the map

$$x, y \in [0,1] \mapsto x\psi(y/x) + (1-x)\psi((c-y)/(1-x)).$$

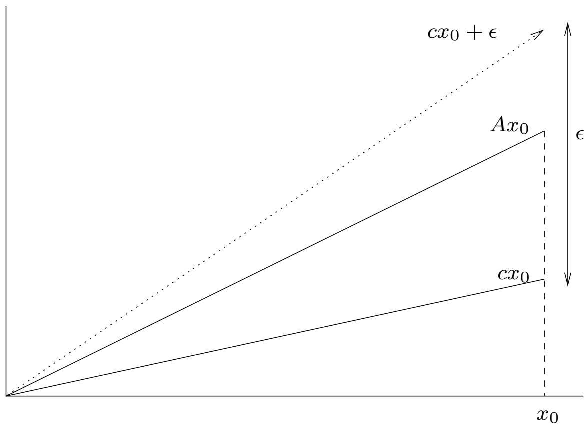

Figure 3: Proof of point (i). By chosing x_0 small enough, reaching $cx_0+\epsilon$ requires the population to deviate from the environments since the environments alone can only reach Ax_0.

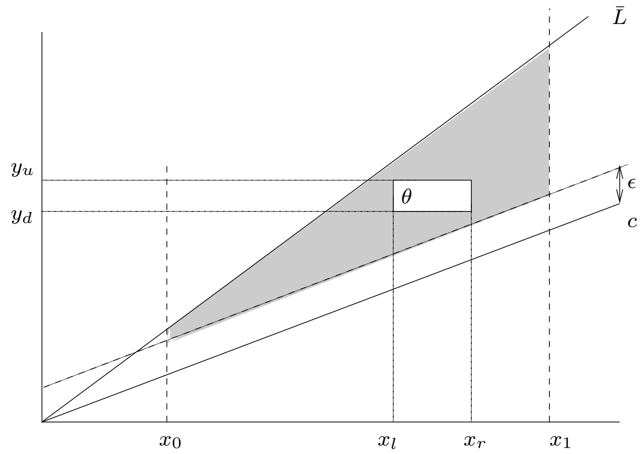

Figure 4: Proof of point (ii). R_c (the grey aera) and the cell θ.

More precisely, consider a cell $\theta = [x_l, x_r] \times [y_d, y_u] \subset R_c$ and define for every $\eta \geq 0$,

$$C_{c,\eta}(\theta) := x_l \psi(y_d/x_l + \eta) + (1 - x_r)\psi((c - y_d)/(1 - x_r) + \eta).$$

Observe that

$$\{\exists x : (x, Y^{(n)}(x)) \in \theta\} \subset \{Y^{(n)}(x_l) \leq y_u\} \cap \{Y^{(n)}(x_d) \geq y_d\},$$

so using the Markov property and the fact that $z_0 \mapsto \mathbb{P}_{z_0}(Y^{(n)}(1) \leq c)$ is decreasing

$$\begin{aligned}
&\mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in \theta,\ Y^{(n)}(1) \leq c) \\
&\quad\leq \mathbb{P}_{z_0}(Y^{(n)}(x_l) \leq y_u, Y^{(n)}(x_r) \geq y_d, Y^{(n)}(1) - Y^{(n)}(x_r) \leq c - Y^{(n)}(x_r)) \\
&\quad\leq \mathbb{P}_{z_0}(Y^{(n)}(x_l) \leq y_u) \sup_{y \geq y_d} \mathbb{P}_{[\exp ny]}(Y^{(n)}(1 - x_r) \leq (c - y)/(1 - x_r)) \\
&\quad\leq \mathbb{P}_{z_0}(Y^{(n)}(x_l) \leq y_u)\mathbb{P}_{[\exp ny_d]}(Y^{(n)}(1 - x_r) \leq (c - y_d)/(1 - x_r))
\end{aligned}$$

Hence, using Lemma 9 (ii), we see that for every $\eta > 0$ small enough, there exists $N(\eta, \theta)$ large enough such that for every $z_0 \geq N(\eta, \theta)$,

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\exists x \in [0,1] : (x, Y^{(n)}(x)) \in \theta, Y^{(n)}(1) < c) \geq C_{c,\eta}(\theta).$$

By continuity of $\eta, \theta \to C_{\eta,c}(\theta)$,

$$\inf_{\substack{\theta \subset R_c,\\ \mathrm{diam}(\theta)\leq\delta}} \big\{C_{\eta,c}(\theta)\big\} \overset{\delta,\eta\to 0}{\longrightarrow} \inf_{z\in R_c} \{C_{0,c}(\{z\})\}.$$

Moreover for every $z = (x, y) \in R_c$, $x \in [x_0, x_1]$ and $y/x > c$, so by strict convexity of ψ,

$$C_{0,c}(\{z\}) = x\psi(y/x) + (1-x)\psi((c-y)/(1-x)) > \psi(c).$$

Then $\inf_{z\in R_c}\{C_{0,c}(\{z\})\} > \psi(c)$, and there exists $\delta_0 > 0$ and $\eta > 0$ such that for every cell θ whose diameter is less than δ_0, for every $z_0 \geq N(\eta, \theta)$,

$$\liminf_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\exists x \in [0,1] : (x, Y^{(n)}(x)) \in \theta, Y^{(n)}(1) < c) > \psi(c). \tag{9}$$

Fix an arbitrary region $R \subset R_c^\circ$ included in the interior of R_c. We can chose $0 < \delta \leq \delta_0$ such that there is a cover of R by the union of a finite collection $\mathcal{K}$ of rectangular regions $[x(i), x(i+1)] \times [y(j), y(j+1)]$ with $i \in \{1, \ldots, N_\delta\}$ and $j \in \{1, \ldots, N(i)\}$ such that their diameter is never more than δ.
Observe that for every $z_0 \geq 1$,

$$\begin{aligned}\mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in R', Y^{(n)} \leq c) &\leq \sum_{\theta\in\mathcal{K}} \mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in \theta, Y^{(n)} \leq c)\\ &\leq |\mathcal{K}| \sup_{\theta\in\mathcal{K}} \mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in \theta, Y^{(n)} \leq c).\end{aligned}$$

Then using (9) simultaneously for each cell $\theta \in \mathcal{K}$, we conclude that for every $z_0 \geq N = \max\{N(\theta, \eta) : \theta \in \mathcal{K}\}$,

$$\begin{aligned}&\liminf_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in R', Y^{(n)} \leq c)\\ &\quad= \min_{\theta\in\mathcal{K}} \liminf_{n\to\infty} -\frac{1}{n} \log \mathbb{P}_{z_0}(\exists x : (x, Y^{(n)}(x)) \in \theta, Y^{(n)} \leq c)\\ &\quad> \psi(c).\end{aligned}$$

As R' is arbitrary in the interior of R_c this concludes the proof of (8) and (ii).

Let us now proceed with the proof of (i). Recall that under hypothesis $\mathcal{H}, \mathbb{P}(L > \log A) = 0$ (i.e. the support of L is bounded by $\log A$.) Fix $\zeta > 0$ and take x_0, x_1 such that $\epsilon/x_0 > A + \zeta, x_0 c < \epsilon$ and $c + \epsilon/(1-x_1) > A + \zeta, \epsilon > c(1-x_1)$.

$$\begin{aligned}&\mathbb{P}_{z_0}\big(\exists x \leq x_0 : |Y^{(n)}(x) - cx| > \epsilon, Y^{(n)}(1) \leq c\big)\\ &\leq \mathbb{P}_{z_0}\big(\exists x \leq x_0 : Y^{(n)}(x) - cx > \epsilon\big)\\ &\leq \mathbb{P}_{z_0}\big(Y^{(n)}(x_0) > \epsilon\big)\\ &\leq \mathbb{P}_{z_0}\big(Y^{(n)}(x_0) > x_0(A+\zeta)\big)\\ &\leq \mathbb{P}_{z_0}\big(\log(Z_{[nx_0]}) > S_{[nx_0]} + \zeta n x_0\big)\end{aligned}$$

since $nx_0(A+\zeta)-S_{nx_0}>\zeta nx_0$. Hence this requires a "deviation from the environments" and by Lemma 6 for η fixed, there exists $D\geq 0$ such that for z_0 large enough,

$$P_{z_0}\big(\exists x\leq x_0:|Y^{(n)}(x)-cx|>\epsilon, Y^{(n)}(1)<c+\log z_0/n\big)\leq D\eta^{nx_0}.$$

Picking η small enough ensures that this is in $o(\exp(-n\psi(c)))$. The argument for the $[x_1,1]$ part of the interval is similar. Thus, recalling that $\limsup_{n\to\infty}-\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n\leq z_0\exp(cn))\leq\psi(c)$ for z_0 large enough, we get (i). $\square$

We can also prove the following stronger result. For every $c<\bar{L}$, for every $\epsilon>0$, there exists $N\in\mathbb{N}$ and $\alpha>0$, such that for $z_0\geq N$,

$$\lim_{n\to\infty}\sup_{c'\in[c-\alpha,c+\alpha]}\mathbb{P}_{z_0}\big(\sup_{x\in[0,1]}\{|Y^{(n)}(x)-c'x|\geq\epsilon\mid Z_n\leq z_0\exp(c'n)\big)=0. \qquad (10)$$

Indeed the proof of Lemma 9 (ii) also ensures that for every $\epsilon>0$ and for every $c_0\leq\bar{L}-\epsilon$ such that $\psi(c_0)<\infty$ there exists N such that for $z_0\geq N$,

$$\liminf_{n\to\infty}\inf_{c\in[c_0,\bar{L}]}\{-\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n\leq z_0e^{cn})-\psi(c+\epsilon)\}\geq 0.$$

Then, following the proof of (ii) above with now

$$\inf_{c\in[c_0,\bar{L}]}\{\inf_{\substack{\theta\subset R_c,\\ \mathrm{diam}(\theta)\leq\delta}}\{C_{\eta,c}(\theta)\}-\inf_{z\in R_c}\{C_{0,c}(\{z\})\}\}\overset{\delta,\eta\to 0}{\longrightarrow}0,$$

there exists $\delta_0>0$ and $\eta>0$ such that for every cell θ whose diameter is less than δ_0, for every $z_0\geq N(\eta,\theta)$, (9) becomes

$$\beta=\liminf_{n\to\infty}\inf_{c\in[c_0,\bar{L}]}\{-\frac{1}{n}\log\mathbb{P}_{z_0}(\exists x\in[0,1]:(x,Y^{(n)}(x))\in\theta, Y^{(n)}(1)<c)-\psi(c)\}>0.$$

Moreover for every $\epsilon>0$,

$$\begin{aligned}\limsup_{n\to\infty}\sup_{c'\in[c-\alpha,c+\alpha]}-\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n\leq\exp(c'n)) &\leq \limsup_{n\to\infty}-\frac{1}{n}\log\mathbb{P}_{z_0}(Z_n\leq\exp((c-\alpha)n))\\ &= \psi(c-\alpha).\end{aligned}$$

Putting the two last inequalities together with $\alpha>0$ such that $\psi(c-\alpha)\leq\psi(c+\alpha)+\beta$ and $[c-\alpha,c+\alpha]\subset[c_0,\bar{L}-\epsilon]$ gives (10).

4.5 End of the proof of Theorem 2

We have proved the part of Theorem 2 concerned with the rate function in section 4.3. Now we tackle the part which gives the full trajectory convergence. For convenience we write $\mathbb{P}^{n,c}(\cdot)$ the conditional probability $\mathbb{P}(\cdot|Z_n\leq e^{cn})$.

We begin to prove that $(Z_n)_{n\in\mathbb{N}}$ does not make a big jump when it goes up to N in the following sense.

Lemma 11. *For every $c < \bar{L}$ and $N \in \mathbb{N}$,*

$$\sup_{n\in\mathbb{N}} \mathbb{P}^{n,c}(Z_{\tau_n(N)} \geq N + M) \overset{M\to\infty}{\longrightarrow} 0.$$

Proof. By the Markov property, for any b and $a \leq N$ fixed,

$$\begin{aligned}
&\mathbb{P}^{n,c}(Z_{\tau_n(N)} \geq N + M \,|\tau_n(N) = b, Z_{\tau_n(N)-1} = a) \\
&\quad = \mathbb{P}_a(Z_1 \geq N + M \mid Z_{n-b} \leq e^{cn}) \\
&\quad \leq \mathbb{P}_N(Z_1 \geq N + M \mid Z_{n-b} \leq e^{cn}) \\
&\quad = \frac{\mathbb{P}_N(Z_{n-b} \leq e^{cn} \mid Z_1 \geq N + M))\mathbb{P}_N(Z_1 \geq N + M)}{\mathbb{P}_N(Z_{n-b} \leq e^{cn})}
\end{aligned}$$

by Bayes' formula. Observe that

$$\mathbb{P}_N(Z_{n-b} \leq e^{cn} \mid Z_1 \geq N + M)) \leq \mathbb{P}_N(Z_{n-b} \leq e^{cn}),$$

so that

$$\mathbb{P}(Z_{\tau_n(N)} \geq N + M \,|Z_n \leq e^{cn}, \tau_n(N) = b, Z_{\tau_n(N)-1} = a) \leq \mathbb{P}_N(Z_1 \geq N + M).$$

This is uniform with respect to a and b so that summing over them yields

$$\forall n \in \mathbb{N}, \quad \mathbb{P}^{n,c}(Z_{\tau_n(N)} \geq N + M) \leq \mathbb{P}_N(Z_1 \geq M + N),$$

which completes the proof letting $M \to \infty$. □

We can now prove the second part of Theorem 2 in the case $\mathbb{P}(\mathbf{p}(1) > 0) > 0$ (case a). Let $\epsilon, \eta > 0$ and $M, N \geq 1$ and note that

$$\mathbb{P}^{n,c}\big(\sup_{t\in[0,1]} \{|Y^{(n)}(t) - f_c(t)|\} \geq \eta\big) \leq A_n + B_n + C_n \tag{11}$$

where

$$\begin{cases} A_n = \mathbb{P}^{n,c}\big(\sup_{t\in[0,1]}\{|Y^{(n)}(t) - f_c(t)|\} \geq \eta,\ \tau(N)/n \in [t_c - \epsilon, t_c + \epsilon],\ Z_{\tau(N)} \leq N + M\big) \\ B_n = \mathbb{P}^{n,c}\big(\tau(N)/n \notin [t_c - \epsilon, t_c + \epsilon]\big) \\ C_n = \mathbb{P}^{n,c}\big(Z_{\tau(N)} \geq N + M\big). \end{cases} \tag{12}$$

We need to show that those three quantities go to zero. The term C_n is small since there is no jump at time $\tau(N)$, and more precisely by Lemma 11, we can find M such that for n large enough

$$C_n \leq \epsilon.$$

The B_n bit deals with the event that $\tau(N)$ is outside of its normal window. Thanks to Lemma 9 (ii), there exists N large enough so that

$$B_n \overset{n\to\infty}{\longrightarrow} 0.$$

Thus it remains only to show that $A_n \to 0$. This is essentially a consequence of Proposition 10 since A_n is related to the trajectory after $\tau(N)$.. Let us now give the details.

We start by observing that for every $\epsilon < \eta/2c$, for n large enough,

$$\sup_{t\in[0,t_c+\epsilon]} \{|\log(N)/n| + |f_c(t)|\} \leq \eta/2,$$

so that conditionally on the event $\{\tau(N)/n \in [t_c-\epsilon, t_c+\epsilon]\}$,

$$\sup_{t\in[0,\tau(N)/n[} \{|Y^{(n)}(t) - f_c(t)|\} < \eta.$$

Then, fixing $\epsilon > 0$ such that

$$\sup_{t_c-\epsilon\leq\alpha\leq t_c+\epsilon,\ t\in[0,1]} \{f_c(\alpha+t) - ct/(1-\alpha)\} \leq \eta/2,$$

we have for every $n \in \mathbb{N}$,

$$\begin{aligned}
&A_n \\
\leq\ & \mathbb{P}^{n,c}\big(\sup_{\tau(N)/n\leq t\leq 1} \{|\log(Z_{[nt]})/n - f_c(t)|\} \geq \eta,\ \tau(N)/n \in [t_c-\epsilon, t_c+\epsilon],\ Z_{\tau(N)} \leq N+M\big) \\
\leq\ & \sup_{\substack{z_0\in[N,N+M] \\ t_c-\epsilon\leq\alpha\leq t_c+\epsilon}} \mathbb{P}_{z_0}\big(\sup_{t\leq 1-\alpha} \{|\log(Z_{[nt]})/n - f_c(\alpha+t)|\} \geq \eta \mid Z_{[(1-\alpha)n]} \leq \exp(cn)\big), \\
\leq\ & \sup_{\substack{z_0\in[N,N+M] \\ t_c-\epsilon\leq\alpha\leq t_c+\epsilon}} \mathbb{P}_{z_0}\big(\sup_{t\leq 1-\alpha} \{|\log(Z_{[nt]})/n - \frac{ct}{1-\alpha}|\} \geq \eta/2 \mid Z_{[n(1-\alpha)]} \leq \exp(cn)\big) \\
\leq\ & \sup_{\substack{z_0\in[N,N+M] \\ c/(1-t_c+\epsilon)\leq x\leq c/(1-t_c-\epsilon)}} \mathbb{P}_{z_0}\big(\sup_{t\leq c/x} \{|\log(Z_{[nt]})/n - xt|\} \geq \eta/2|\ Z_{[nc/x]} \leq \exp(nc/x.x)\big)
\end{aligned}$$

By (10), there exists $\epsilon > 0$ such that $A_n \overset{n\to\infty}{\longrightarrow} 0$. Then using (11),

$$\mathbb{P}^{n,c}\big(\sup_{t\in[0,1]} \{|\log(Z_{[nt]})/n - f_c(t)|\} \geq \eta\big) \overset{n\to\infty}{\longrightarrow} 0.$$

Thus in the case $\mathbb{P}(\mathbf{p}(1) > 0) > 0$, we get that conditionally on $Z_n \leq e^{cn}$,

$$\sup_{t\in[0,1]} \{\big|\log(Z_{[tn]})/n - f_c(t)\big|\} \overset{n\to\infty}{\longrightarrow} 0, \qquad \text{in } \mathbb{P}.$$

The case $\mathbb{P}(\mathbf{p}(1) > 0) = 0$ is easier (and amounts to make $t_c = 0$ in the proof above).

5 Proof for upper deviation

Here, we assume that for every $k \geq 1$,

$$\mathbb{E}(Z_1^k) < \infty.$$

Lemma 12. *For every $c \geq \bar{L}$, denoting by*

$$s_{max} := \sup\{s > 1 : \mathbb{E}(m(\mathbf{p})^{1-s}) < 1\},$$

we have for every $z_0 \geq 1$,

$$\liminf_{n\to\infty} \inf_{z_0\geq 1} \{-\frac{1}{n}\log\big(\mathbb{P}_{z_0}(Z_n \geq z_0\exp(cn))\big)\} \geq \sup_{0\leq\eta\leq c-\bar{L}} \min(s_{max}\eta,\ \psi(c-\eta)).$$

The first part of Theorem 3 is a direct consequence of this lemma. Indeed, in the case when Z_n is strongly supercritical, $s_{max} = \infty$, then letting $\eta \downarrow 0$, we get, for every $c \geq \bar{L}$,

$$-\log(\mathbb{P}_1(Z_n \leq e^{cn}))/n \overset{n\to\infty}{\longrightarrow} \psi(c).$$

Proof of Lemma 12. For every $\eta > 0$, $\mathbb{P}_{z_0}(Z_n \geq z_0 \exp(cn))$ is smaller than

$$\mathbb{P}_{z_0}(Z_n \geq z_0 \exp(cn)),\ S_n \leq n[c-\eta]) + \mathbb{P}_{z_0}(Z_n \geq z_0 \exp(cn)),\ S_n \geq n[c-\eta]). \tag{13}$$

First, as for every $k \geq 1$, $\mathbb{E}(Z_1^k) < \infty$, by Theorem 3 in [16], for every $s > 1$ such that

$$\mathbb{E}(m(\mathbf{p})^{1-s}) < 1,$$

there exists $C_s > 0$ such that for every $n \in \mathbb{N}$,

$$\mathbb{E}_1(W_n^s) \leq C_s,$$

where $W_n = \exp(-S_n)Z_n$. Note that conditionally on the environments $(\mathbf{p}_i)_{i=0}^{n-1}$, W_n starting from z_0 is the sum of z_0 iid random variable distributed as W_n starting from 1. Thus, there exists C'_s such that for all $n, z_0 \in \mathbb{N}$,

$$\mathbb{E}_{z_0}(W_n^s) \leq z_0^s C'_s.$$

Then, by Markov inequality,

$$\begin{aligned}
\mathbb{P}_{z_0}(Z_n \geq z_0 \exp(cn),\ S_n \leq n[c-\eta]) &\leq \mathbb{P}_{z_0}(Z_n \exp(-S_n) \geq z_0 \exp(n\eta)) \\
&= \mathbb{P}_{z_0}(W_n \geq z_0 \exp(n\eta)) \\
&\leq \frac{\mathbb{E}_{z_0}(W_n^s)}{z_0^s \exp(ns\eta)} \\
&\leq C'_s \exp(-sn\eta). \qquad (14)
\end{aligned}$$

Second, by standard large deviation upper bound, we have

$$\mathbb{P}_{z_0}(Z_n \geq \exp(cn),\ S_n \geq n[c-\eta]) \leq \mathbb{P}(S_n \geq n[c-\eta]) \leq \exp(-n\psi(c-\eta)). \tag{15}$$

Combining (13),(14), and (15) we get

$$\liminf_{n\to\infty} \inf_{z_0 \geq 1} \{-\log\big(\mathbb{P}_{z_0}(Z_n \geq z_0 \exp(cn))\big\}/n \geq \min(s\eta,\ \psi(c-\eta)).$$

Thus,

$$\liminf_{n\to\infty} \inf_{z_0 \geq 1} -\log\big(\mathbb{P}(\log(Z_n)/n \geq c)\big)\}/n \geq \sup_{0 \leq \eta \leq c-\bar{L}} \min(s\eta,\ \psi(c-\eta)).$$

Letting $s \uparrow s_{max} = \sup\{s > 1 : \mathbb{E}(m(\mathbf{p})^{1-s}) < 1\}$ yields the result. □

The proof of the second part of Theorem 3 follows the proof of Proposition 10. Roughly speaking, for all $t \in (0,1)$ and $\epsilon > 0$,

$$\mathbb{P}(Z_{[nt]} = \exp(tn(c+\epsilon)),\ Z_n \geq \exp(cn)) = \mathbb{P}(Z_{[nt]} = \exp(tn(c+\epsilon)))\mathbb{P}_{\exp(tn(c+\epsilon))}(Z_{n-[nt]} \geq \exp(cn)).$$

Then the first part of Theorem 3 ensures that

$$\begin{aligned}\lim_{n\to\infty} -\log(\mathbb{P}(Z_{[nt]} = \exp(tn(c+\epsilon)),\ Z_n \geq \exp(cn)))/n \\ = t\psi(c+\epsilon) + (1-t)\psi(c - t/(1-t)\epsilon) \\ > \psi(c),\end{aligned}$$

by strict convexity of ψ. This entails that $\log(Z_{[nt]})/n \to ct$ as $n \to \infty$.

6 Proof without supercritical environments

We assume here that $\mathbb{P}(m(\mathbf{p}) \leq 1) = 1$. Recall that f_i is the probability generating function of $\mathbf{p}_i$ and that, denoting by

$$F_n := f_0 \circ \cdots \circ f_{n-1},$$

we have for every $k \in \mathbb{N}$,

$$\mathbb{E}_k(s^{Z_{n+1}} \mid f_0, ..., f_n) = F_{n+1}(s)^k \qquad (0 \leq s \leq 1).$$

We assume also that for every $j \geq 1$, there exists $M_j > 0$ such that

$$\sum_{k=0}^{\infty} k^j \mathbf{p}(k) \leq M_j \quad \text{a.s.}$$

Then,

$$f^{(j)}(1) \leq M_j \quad \text{a.s.}$$

We use that for ever $c > 1$ and $k \geq 1$, by Markov inequality,

$$\begin{aligned}\mathbb{P}(Z_n \geq c^n) &= \mathbb{P}(Z_n(Z_n-1)...(Z_n-k+1) \geq c^n(c^n-1)...(c^n-k+1)) \\ &\leq \frac{\mathbb{E}(Z_n(Z_n-1)...(Z_n-k+1))}{c^n(c^n-1)...(c^n-k+1)} \\ &= \frac{\mathbb{E}(F_n^{(k)}(1))}{c^n(c^n-1)...(c^n-k+1)}.\end{aligned}$$

Thus, to get Proposition 4, it is enough to prove that for every $k > 1$,

$$\mathbb{E}(F_n^{(k)}(1)) \leq C_k n^{k^k}$$

and let $k \to \infty$. The last inequality can be directly derived from the following lemma, since here $f_i'(1) \leq 1$ a.s. and there exists $M_j > 0$ such that for every $j \in \mathbb{N}$, $f^{(j)}(1) \leq M_j$ a.s.

Lemma 13. *Let $(g_i)_{1\leq i\leq n}$ be power series with positive coefficients such that*

$$\forall 2 \leq i \leq n, \quad g_i(1) = 1$$

and denote by

$$G_i = g_i \circ ... \circ g_n, \qquad (1 \leq i \leq n).$$

Then, for every $k \geq 0$,

$$\sup_{x\in[0,1]} G_1^{(k)}(x) \leq \max_{\substack{0\leq j\leq k \\ 1\leq i\leq n}} (1, [g_i^{(j)}(1)]^{k^k}). \max_{2\leq i\leq n} (1, g_i'(1))^{nk}.n^{k^k}$$

Proof. This result can be proved by induction. Indeed,

$$\begin{aligned} G_1^{(k+1)} &= [\Pi_{i=1}^n g_i' \circ G_{i+1}]^{(k)} \\ &= \sum_{k_1+...+k_n=k} \Pi_{i=1}^n [g_i' \circ G_{i+1}]^{(k_i)}. \end{aligned}$$

Then, noting that $\#\{i \in [1,n] : k_i > 0\} \leq k$ and $\#\{k_i : k_1 + ... + k_n = k\} \leq n^k$, for every $x \in [0,1]$,

$$G_1^{(k+1)}(x) \leq n^k \max_{\substack{1\leq i\leq n \\ 0\leq k_i\leq k}} \{1, [g_i' \circ G_{i+1}]^{(k_i)}(x)\}^k. \max(1, g_1'(G_2(x))). \max_{2\leq i\leq n} (1, g_i'(1))^n.$$

So,

$$\sup_{x\in[0,1]} G_1^{(k+1)}(x) \leq n^k \max_{\substack{1\leq i\leq n \\ 0\leq k_i\leq k}} \{1, [g_i' \circ G_{i+1}]^{(k_i)}(x)\}^{k+1}. \max_{2\leq i\leq n} (1, g_i'(1))^n.$$

One can complete the induction noting that $k + k^k(k+1) \leq (k+1)^{k+1}$. □

Acknowledgements: The authors wish to thank Amaury Lambert for many interesting discussions at the start of this project.

References

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

[2] K. B. Athreya, S. Karlin (1971). On branching processes with random environments, I : Extinction probability. *Ann. Math. Stat.* 42. 1499-1520.

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

[4] K. B. Athreya; A. Vidyashankar (1993). Large deviation results for branching processes. *Stochastic processes*, 7-12, Springer, New York.

[5] K. B. Athreya (1994). Large deviation rates for branching processes. I Single type case. *Ann. Appl. Probab.* 4, no. 3, 779-790.

[6] K. B. Athreya, P. E. Ney (2004). *Branching processes.* Dover Publications, Inc., Mineola, NY.

[7] V. Bansaye (2008). Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Ann. Appl. Prob.* Vol. 18, No. 3, 967-996.

[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 (1988). On the limit of a supercritical branching process. *J. Appl. Prob.* 25A, 215-228.

[10] F.M. Dekking (1988). On the survival probability of a branching process in a finite state iid environment. *Stochastic Processes and their Applications* 27 p151-157.

[11] A. Dembo, O. Zeitouni (1998). *Large deviations techniques and applications.* Second edition. Applications of Mathematics (New York), 38. Springer-Verlag, New York.

[12] S. Dubuc (1971). La densite de la loi limite d'un processus en casacade expansif. *Z. Wahr.* 19, 281-290.

[13] S. Dubuc (1971). Problemes relatifs a l'iteration des fonctions suggeres par les processus en cascade. *Ann. Inst. Fourier* 21, 171-251.

[14] K. Fleischmann, V. Wachtel (2007). Lower deviation probabilities for supercritical Galton-Watson processes. *Ann. I. H. Poincar Probab. Statist.* 43 , no. 2, 233–255

[15] J. Geiger, G. Kersting, V A. Vatutin (2003). Limit Theorems for subcritical branching processes in random environment. *Ann. I. H. Poincar* 39, no. 4. 593-620.

[16] 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.

[17] B. Hambly (1992). On the limiting distribution of a supercritical branching process in random environment. *J. Appl. Probab.* 29, No 3, 499-518.

[18] M. V. Kozlov (2006). On large deviations of branching processes in a random environment: a geometric distribution of the number of descendants. *Diskret. Mat. 18* , no. 2, 29–47.

[19] Q. Liu (2000). On generalized multiplicative cascades. *Stochastic Process. Appl.* 86, no. 2, 263-286.

[20] P. Mörters, M. Ortgiese (2008). Small value probabilities via the branching tree heuristic. *Bernoulli* 14, no. 1, 277–299.

[21] P. E. Ney; A. N. Vidyashankar (2004). Local limit theory and large deviations for supercritical branching processes. *Ann. Appl. Probab.* 14, no. 3, 1135–1166

[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, W. Wilkinson (1969). On branching processes in random environments. *Ann. Math. Stat.* Vol40, No 3, p814-827.