



**HAL**  
open science

# A PHASE TRANSITION FOR LARGE VALUES OF BIFURCATING AUTOREGRESSIVE MODELS

Vincent Bansaye, Siméon Valère Bitseki Penda

► **To cite this version:**

Vincent Bansaye, Siméon Valère Bitseki Penda. A PHASE TRANSITION FOR LARGE VALUES OF BIFURCATING AUTOREGRESSIVE MODELS. 2019. hal-02416177

**HAL Id: hal-02416177**

**<https://hal.science/hal-02416177v1>**

Preprint submitted on 17 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A PHASE TRANSITION FOR LARGE VALUES OF BIFURCATING AUTOREGRESSIVE MODELS

VINCENT BANSAYE AND S. VALÈRE BITSEKI PENDA

ABSTRACT. We describe the asymptotic behavior of the number  $Z_n[a_n, \infty)$  of individuals with a large value in a stable bifurcating autoregressive process. The study of the associated first moment  $\mathbb{E}(Z_n[a_n, \infty))$  is equivalent to the annealed large deviation problem  $\mathbb{P}(Y_n \geq a_n)$ , where  $Y$  is an autoregressive process in a random environment and  $a_n \rightarrow \infty$ .

The population with large values and the trajectorial behavior of  $Z_n[a_n, \infty)$  is obtained from the ancestral paths associated to the large deviations of  $Y$  together with its environment. The study of large deviations of autoregressive processes in random environment is of independent interest and achieved first in this paper. The proofs of trajectorial estimates for bifurcating autoregressive process involves then a law of large numbers for non-homogenous trees.

Two regimes appear in the stable case, depending on the fact that one of the autoregressive parameter is greater than one or not. It yields two different asymptotic behaviors for the large local densities and maximal value of the bifurcating autoregressive process.

**Keywords:** Branching process, autoregressive process, random environment, large deviations.

**Mathematics Subject Classification (2010):** 60J80, 60J20, 60K37, 60F10, 60J20, 60C05, 92D25.

## 1. INTRODUCTION

The bifurcating autoregressive (BAR) process  $X = (X_n)_{n \geq 1}$  is a model for affine random transmission of a real value along a binary tree. To define this process, we consider a real value random variable  $X_1$ , independent of a sequence of i.i.d bivariate random variables  $((\eta_{2k}, \eta_{2k+1}), k \geq 1)$  with law  $\mathcal{N}_2(0, \Gamma)$ , where

$$\Gamma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \rho \in (-1, 1).$$

Then  $X$  is defined inductively for  $k \geq 1$  by

$$(1) \quad \begin{cases} X_{2k} = \alpha X_k + \eta_{2k} \\ X_{2k+1} = \beta X_k + \eta_{2k+1}, \end{cases}$$

where  $\alpha, \beta$  are non-negative real numbers. Informally, the value  $X_k$  of individual  $k$  is randomly transmitted to its two offsprings  $2k$  and  $2k + 1$  following an autoregressive process.

This model has been introduced in the symmetric case  $\alpha = \beta$  by Cowan [20] and Cowan and Staudte [21] to analyze cell lineages data. It allows to study the evolution and the transmission of a trait after division, in particular it size or it growth rate. In these works, Cowan and Staudte have focused on the study of the bacteria *Escherichia Coli*. *E. Coli* is a rod-shaped bacterium which reproduces by dividing in the middle, producing two cells. Several extensions of their model have been proposed and we refer e.g. to the works of Basawa and Higgins [6, 7] and Basawa and Zhou [8, 9], where the model of Cowan and Staudte is studied for long memory and non-Gaussian

noise.

In 2005, Steward *et al.* [43] have designed an experimental protocol which brings evidence of aging and asymmetry in *E. Coli*. In order to study the dynamic of evolution of values in a population of cells, taking into account such possible asymmetry, Guyon and *al.* [32] have considered and used the model of Cowan and Staudte with  $(\alpha, \beta) \in (0, 1)^2$ , which means that  $\alpha$  and  $\beta$  may be different. Since then, this model has been extended to more complex settings and studied from a statistical and probabilistic point of view. Bercu *et al.* [11] consider an extension of BAR model with non-Gaussian noise and long memory. They use martingale approach in order to study the asymptotic behavior of least-squares estimators of unknown parameters of their model. In the following of this work, Bitseki and Djellout [13] and Bitseki *et al.* [14] have studied the deviation inequalities and the moderate deviations principle of unknown parameters in the BAR models introduced by Bercu *et al.* [11] and Guyon [31]. Several extension of this model were proposed where the missing data are taken into account in the study of cell lineage data, see de Saporta *et al.* [24, 25, 27] and Delmas and Marsalle [22]. Furthermore Bercu and Blandin [10] and de Saporta *et al.* [26], have considered random coefficients, while Bitseki Penda and Olivier [15] have worked on nonlinear autoregressive models.

More generally, polarisation of cells is a fundamental process in cell biology. Asymmetry at cell division have been observed and studied in different contexts, see e.g. [37] for plasmids, [42] for extra chromosomal DNA, [35] for mitochondrias and [5] for parasites. Asymmetry at division is a key feature of aging, cell variability and differentiation [19].

The BAR models turns out to be an interesting toy model to explore mathematically the effect of randomness and asymmetry in the long time behavior of characteristics of cells in division processes. In this paper, we focus on large values among the cell population. We can, describe here how lineages and stochasticity from the Gaussian additive term interplay to explain this part of the distribution of traits when time goes to infinity.

**1.1. Stability and random cell lineage.** To analyze the long time behavior of the bifurcating autoregressive model, the genealogy of the population is involved. It is given here by the binary tree. Each vertex is seen as a positive integer different from 0 and it denotes one individual of the population. The initial individual is thus denoted by 1 and for  $n \in \mathbb{N}$ ,

$$\mathbb{G}_n = \{2^n, 2^n + 1, \dots, 2^{n+1} - 1\} \quad \text{and} \quad \mathbb{T}_n = \bigcup_{m=0}^n \mathbb{G}_m$$

denote respectively the  $n$ -th generation and the first  $(n + 1)$  generations of the population. The collection of values  $(X_i : i \in \mathbb{G}_n)$  of the individuals in generation  $n$  is represented by the random punctual measure

$$Z_n = \sum_{i \in \mathbb{G}_n} \delta_{X_i}.$$

The process  $Z$  is a multitype branching process where types are real valued. The first moment of measure  $Z_n$  satisfies the following simple many-to-one formula:

$$(2) \quad \mathbb{E}(Z_n([a, b])) = 2^n \mathbb{P}(Y_n \in [a, b]),$$

for any real numbers  $a \leq b$ , where  $(Y_n)_n$  is the autoregressive process in random environment corresponding to a uniform random path in the binary tree. More precisely,  $Y$  is defined by

$$(3) \quad Y_0 = X_1 \quad \text{and} \quad \forall n \geq 1, \quad Y_n = \theta_n Y_{n-1} + \varepsilon_n,$$

where  $\Theta = (\theta_n, n \geq 1)$  is a sequence of i.i.d. random variables such that

$$\mathbb{P}(\theta_1 = \alpha) = \mathbb{P}(\theta_1 = \beta) = 1/2$$

and  $(\varepsilon_n, n \geq 1)$  is a sequence of i.i.d. centered standard Gaussian random variables independent of  $(\theta_n, n \geq 1)$ . We observe that for any  $n \geq 1$ ,

$$(4) \quad Y_n = \left( \prod_{k=1}^n \theta_k \right) Y_0 + \bar{Y}_n, \text{ where } \bar{Y}_n = \sum_{k=1}^n \left( \prod_{\ell=k+1}^n \theta_\ell \right) \varepsilon_k,$$

using the convention  $\prod_{\ell \in \emptyset} = 1$ . Conditionally on  $(\theta_n, n \geq 1)$ ,  $\bar{Y}_n$  is a Gaussian random variable with standard deviation

$$(5) \quad A_n = A(\theta_1, \dots, \theta_n) = \sqrt{\sum_{k=1}^n \prod_{\ell=k+1}^n \theta_\ell^2}.$$

Moreover, for each  $n$ ,  $A_n$  is distributed as  $A_n^* = A(\theta_n, \dots, \theta_1)$  and by monotonicity the latter converges a.s. as  $n \rightarrow \infty$  to

$$A_\infty^* = \sqrt{\sum_{k=1}^{\infty} \prod_{\ell=1}^{k-1} \theta_\ell^2} \in (0, \infty].$$

If  $\alpha\beta \geq 1$  then  $A_\infty^* = \infty$  a.s. and the mean proportion of individuals whose value is in a compact set goes to 0 as  $n$  goes to infinity, i.e.  $Z_n([a, b])/2^n \rightarrow 0$  a.s. We focus here on the stable case :

$$\alpha\beta < 1.$$

Then  $(Y_n)_n$  converges in law to its unique stationary distribution  $\pi$  on  $\mathbb{R}$ , which is a mixed centered gaussian random variable with random standard deviation distributed as  $A_\infty^*$ . Guyon [31] has shown in that case that a strong law of large numbers holds :

$$\frac{1}{2^n} Z_n([a, b]) \xrightarrow{n \rightarrow \infty} \pi([a, b]) \quad \text{a.s.}$$

**1.2. Large deviations and local densities in the stable case : main results.** We are interested here in the number of individuals with large values and the extremal values among the population. More precisely we study

$$Z_n([a_n, \infty)) = \#\{i \in \mathbb{G}_n : X_i \geq a_n\}$$

when  $a_n \rightarrow \infty$ . In particular we expect a law of large number (or concentration) effect, informally

$$Z_n([a_n, \infty)) \approx \mathbb{E}(Z_n([a_n, \infty))) = 2^n \mathbb{P}(Y_n \in [a_n, \infty))$$

at the logarithm scale.

Proving this law of large number effect and determining the asymptotic behavior of  $\mathbb{E}(Z_n([a_n, \infty)))$  leads us to study the large deviation event  $\{Y_n \geq a_n\}$ . More precisely, we describe the past trajectory  $(Y_i : i \leq n)$  conditionally on this event  $\{Y_n \geq a_n\}$  together with the associated environment  $(\theta_i : i \leq n)$ . It allows us both to estimate the mean behavior  $\mathbb{P}(Y_n \geq a_n)$  and to control the subtree associated to the local density  $Z_n([a_n, \infty))$  and thus correlations of the values between individuals. This large deviation issue on the autoregressive process  $Y$  is both crucial here for the analysis of  $Z_n$  and of independent interest. In the case  $\alpha = \beta < 1$ ,  $Y_n$  is a gaussian random variable and  $\mathbb{P}(Y_n \geq a_n)$  is derived analytically. In the general stable case  $\alpha\beta < 1$ , two regimes appear, which correspond to two different behaviors of the past trajectory and associated environments  $\theta$ . We obtain the following classification

- i)* Case  $\alpha = \max\{\alpha, \beta\} < 1$ . Then, the deviation event  $\{Y_n \geq a_n\}$  is achieved by selecting the most favorable environment  $\alpha$  in the last generations. The deviation event  $\{Y_n \geq a_n\}$  relies on extreme events in these last generations, which are quantified by the tail of Gaussian random variables with standard deviation  $\alpha$ . We obtain that for any  $x > 0$ ,

$$\lim_{n \rightarrow +\infty} \frac{1}{xn} \log \mathbb{P}(Y_n \geq \sqrt{xn}) = -\frac{1 - \alpha^2}{2}.$$

Moreover we prove that conditionally on  $\{Y_n \geq a_n\}$ , the past trajectory  $(Y_{n-i} : 0 \leq i \leq n)$  is approximated by a geometric progression  $(a_n \alpha^i : 0 \leq i \leq n)$ .

- ii)* Case  $\beta < 1 < \alpha$  and  $\alpha\beta < 1$ . Then there exists a unique  $\kappa > 0$  such that

$$\alpha^\kappa + \beta^\kappa = 2$$

and now the deviation event  $\{Y_n \geq a_n\}$  comes from the deviation of the environment  $\{A_n \geq a_n\}$ . To estimate the probability of that latter, we use Kesten theorem, which ensures that  $\mathbb{P}(A_\infty \geq a_n) \sim ca_n^\kappa$  as  $n \rightarrow \infty$ . We thus restrict the study here to the case when  $\{A_n^* \geq a_n\}$  is comparable to  $\{A_\infty^* \geq a_n\}$ , which will correspond to the fact that the supremum over all time of the associated random walk drifting to  $-\infty$  is reached before time  $n$  when it is larger than  $a_n$ . That leads us to introduce  $\gamma \in (0, 1)$  defined by

$$\gamma := \inf_{s \geq 0} (\alpha^s + \beta^s)/2.$$

Then for a sequence  $a_n$  tending to infinity not too fast, we prove that  $\mathbb{P}(Y_n \geq a_n)$  is equivalent to  $a_n^{-\kappa}$ . More precisely we need that  $a_n^\kappa$  is negligible compared to  $n^{3/2}(1/\gamma)^n$  as  $n$  tends to infinity. In particular, for any  $\rho \geq 1$  such that  $\rho^\kappa \in (0, 1/\gamma)$ , we get

$$\lim_{n \rightarrow +\infty} \frac{1}{n \log(\rho)} \log \mathbb{P}(Y_n \geq \rho^n) = -\kappa.$$

We expect that the regimes  $\max(\alpha, \beta) = 1$  and  $a_n = \exp(n\rho)$  with  $\rho > 1/\gamma$  could be studied with an approach similar to *ii)*. As far as we seen,  $\{A_n^* \geq a_n\}$  will not be comparable to  $\{A_\infty^* \geq a_n\}$  if  $\rho > 1/\gamma$  and describing the behavior of  $\{Y_n \geq a_n\}$  should require more work. We also mention that non gaussian noise could lead to different behavior, in particular if the tail of  $\varepsilon$  is at the same scale as the deviation probability coming from the environment.

We have thus obtained both the asymptotic behavior of  $\mathbb{E}(Z_n([a_n, \infty))) = 2^n \mathbb{P}(Y_n \in [a_n, \infty))$  and the ancestral path of  $(Y, \Theta)$  on the event  $Y_n \geq a_n$ . Roughly speaking, a trajectorial version of the many-to-one formula (2) allows us to derive the genealogy and process contributing to the local density  $Z_n([a_n, \infty))$  from the ancestral path of  $(Y, \Theta)$  before time  $n$ . We obtain the following long time estimates

- i)* Case  $\alpha = \max\{\alpha, \beta\} < 1$ . Then, for any  $x \in (0, 2 \log(2)/(1 - \alpha^2))$ ,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log Z_n([x\sqrt{n}; +\infty)) = \log(2) - x(1 - \alpha^2)/2 \quad \text{a.s.}$$

In that case, the proof follows ideas of Biggins for branching random walks and we consider a subpopulation of individuals which realizes the deviation and thus the large values  $[x\sqrt{n}, \infty)$ . It is described by a non-homogeneous Galton-Watson process, inherited from the trajectory of values in last generation following the geometric paths obtained above for the deviation  $\{Y_n \geq a_n\}$ . The main difference with branching random walk is this time non-homogeneity since the trajectory associated with the deviation of the value (or value) is not straightline but geometric and limited to the last generations.

ii) Case  $\beta < 1 < \alpha$  and  $\alpha\beta < 1$ . For all  $\rho \geq 1$  such that  $\rho^\kappa < \min(2, 1/\gamma)$ , we have

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log(Z_n([\rho^n, \infty)) = \log(2) - \kappa \log(\rho) \quad \text{in probability.}$$

The subtree associated with the individuals in  $[\rho^n, \infty)$  is very different from the previous case. Indeed, only some particular paths of the binary genealogical tree are involved in local densities, corresponding to the deviation event  $\{A_n^* \geq \rho^n\}$ . The evolution of the chain on these paths does not deviate from its usual Gaussian autoregressive expected behavior. The proof is here more involved since a study of the process restricted to the subtree of  $\mathbb{T}_n$  corresponding to the environment  $\Theta_n = (\theta_i : i \leq n)$  associated to the large deviation of  $Y_n$ . This phenomenon is linked to the random environment structure inherited from the tree and can be compared to the results in the weakly subcritical case in Kimmel's branching model [5]. In that latter, an analogous phase transition occurs but only the mean behavior had been obtained. The technique developed in this paper should help to obtain a convergence in probability. The issue of the a.s. convergence remains open. We could expect to prove it by estimating for instance the speed of convergence in probability.

In words, two sources of randomness are combined here : the gaussian additive term of the autoregressive process and the choice of paths in the binary genealogy. In the first regime, the local densities is inherited from both randomness, in the very last generations. The picture is quite close to branching random walks, but the fact that deviation creating local densities occurs only at the end, not starting from the beginning. In the second regime, only the path in the tree is involved for large traits and the deviation starts much earlier to create these local densities. The picture is very different from branching random walks since a deterministic subtree, inherited from functional  $A_n$  support the local density. The study of the autoregressive model  $Y$  together with its environment  $\theta$  yields a natural way to analyse this issue.

**1.3. Extremal value.** The maximal value in generation  $n$

$$M_n = \max\{X_i : i \in \mathbb{G}_n\}$$

has different behavior and speed. If  $\alpha = \max\{\alpha, \beta\} < 1$ , then

$$\lim_{n \rightarrow +\infty} \frac{1}{\sqrt{n}} M_n = \sqrt{\frac{2 \log(2)}{1 - \alpha^2}} \quad \text{a.s.}$$

If  $\beta < 1 < \alpha$ , then

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log(M_n) = \frac{\log(2)}{\kappa} \quad \text{in probability.}$$

The issue of maximal value has attracted lots of attention and has been extensively studied for branching random walks from the works of Hammersley [33], Kingman [38], Biggins [12] and Bramson [18]. Fine estimates have been obtained recently, which have shed light on the branching structure of the process, see in particular Hu and Shi [34], Faraud et al. [29], Addario-Berry and Reed [1] and Aïdékon [2].

The behavior of the maximal value is different in this model. Technics developed for branching random walk are used and adapted here, in particular the use of many to one formula to exploit the description of the path leading to extremal particules [41] and the construction of an imbedded branching process. The main novelty lies in the weak regime  $\beta < 1 < \alpha$  where a law of large numbers of a well chosen subtree of the binary tree is involved.

**1.4. Notation.** We will use the following notations. Let  $(x_n)$  and  $(y_n)$  be two sequences of real numbers.

- We recall that  $x_n = \mathcal{O}(1)$  means that  $x_n \rightarrow 0$  as  $n \rightarrow +\infty$ .
- We write  $x_n = \mathcal{O}(y_n)$  if  $x_n/y_n = \mathcal{O}(1)$ .
- We write  $x_n \sim y_n$  if  $x_n/y_n \rightarrow 1$  as  $n \rightarrow +\infty$ .
- We write  $x_n \lesssim y_n$  if there is a positive constant  $c$ , independent of  $n$ , such that  $x_n \leq cy_n$ .
- We write  $x_n \asymp y_n$  if there are two positives constants  $c_1$  and  $c_2$  such that  $c_1 y_n \leq x_n \leq c_2 y_n$  as  $n \rightarrow +\infty$ .

## 2. LARGE DEVIATIONS AND ANCESTRAL PATHS IN THE STABLE CASE

This section has two goals. The first is to evaluate the probabilities of large deviations of random coefficients autoregressive process  $(Y_n, n \in \mathbb{N})$ . The second is to understand the asymptotic behavior of the paths of this process conditionally to the large deviations events. For simplicity and owing to the motivation of this paper (see the next section), we focus on the case where the random coefficients take only two values with identical probability. Then, let  $(\theta_n, n \geq 1)$  be a sequence of i.i.d. random variables which take their values in  $\{\alpha, \beta\}$  with  $0 < \beta \leq \alpha$ , and  $\mathbb{P}(\theta_1 = \alpha) = 1/2$ . Let  $(\varepsilon_n, n \geq 1)$  be a sequence of i.i.d. centered standard Gaussian random variables independent of  $(\theta_n, n \geq 1)$ . We assume without loss of generality that  $Y_0 = 0$  and consider the process  $(Y_n, n \in \mathbb{N})$  defined by

$$Y_0 = 0 \quad \text{and} \quad \forall n \geq 1, \quad Y_n = \theta_n Y_{n-1} + \varepsilon_n,$$

in such a way that from (4), we have

$$Y_n = \sum_{k=1}^n \left( \prod_{\ell=k+1}^n \theta_\ell \right) \varepsilon_k.$$

**2.1. Large deviations for strictly stable case.** We consider here the case

$$\beta \leq \alpha < 1$$

and show that the (annealed) large deviation behavior of  $(Y_n, n \in \mathbb{N})$  is given by a geometric growth in environment  $\alpha$  in the last generations.

**Theorem 1.** *Let  $(a_n)_{n \in \mathbb{N}}$  be a sequence of non-decreasing real numbers which tends to infinity. We have*

$$(6) \quad \lim_{n \rightarrow +\infty} \frac{1}{a_n^2} \log \mathbb{P}(Y_n \geq a_n) = -\frac{1 - \alpha^2}{2}.$$

Moreover, for all sequence  $(\ell_n)_n$  of integers such that  $n - \ell_n \rightarrow \infty$  and  $\ell_n = \mathcal{O}(\log a_n)$ ,

$$(7) \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left( \sup_{\ell \in \{0, \dots, \ell_n\}} \left| \frac{Y_{n-\ell}}{a_n} - \alpha^\ell \right| \leq \varepsilon, \quad (\theta_{n-\ell_n}, \dots, \theta_n) = (\alpha, \dots, \alpha) \mid Y_n \geq a_n \right) = 1$$

for any  $\varepsilon > 0$ .

*Remark 1.* More generally, one can see the proof to check that admissible sequences  $(\ell_n)_n$  just need to satisfy  $n - \ell_n \rightarrow \infty$  and  $\lim_{n \rightarrow \infty} 2^{\ell_n} \exp(-c(a_n \alpha^{\ell_n})^2) = 0$  for some finite constant  $c$ .

*Proof.* Let  $\mathcal{F}_n = \sigma(\theta_1, \dots, \theta_n)$  be the filtration defined by  $\theta_1, \dots, \theta_n$ . We recall that  $\mathcal{L}(Y_n | \mathcal{F}_n) = \mathcal{N}(0, A_n^2)$ , the centered normal law with variance  $A_n^2$ , where  $A_n$  is defined in (5). We observe that

$$A_n^2 \leq \sum_{k=0}^{\infty} \alpha^{2k} = \frac{1}{1 - \alpha^2}$$

since  $\beta \leq \alpha$ . First, we have

$$(8) \quad \begin{aligned} \mathbb{P}(Y_n \geq a_n) &= \mathbb{E} \left[ \mathbb{P}(\mathcal{N}(0, A_n^2) \geq a_n | \mathcal{F}_n) \right] \\ &\leq \mathbb{P} \left( \mathcal{N} \left( 0, \frac{1}{1 - \alpha^2} \right) \geq a_n \right) \lesssim \frac{1}{a_n} \exp \left( -\frac{(1 - \alpha^2)a_n^2}{2} \right), \end{aligned}$$

where the last inequality follows from the classical upper tail inequality for standard normal distribution, see for instance [16].

Next, we introduce a sequence  $(\ell_n)_{n \in \mathbb{N}}$  which tends to infinity such that  $\ell_n \leq n$  and  $\ell_n = o(a_n^2)$ . We set

$$B_{\ell_n} = \{\theta_n = \alpha, \dots, \theta_{n - \ell_n} = \alpha\}, \quad u_{\ell_n}^2 = \sum_{k=0}^{\ell_n} \alpha^{2k} = \frac{1 - \alpha^{2\ell_n + 2}}{1 - \alpha^2}$$

and using again  $\mathcal{L}(Y_n | \mathcal{F}_n) = \mathcal{N}(0, A_n^2)$ , we have

$$(9) \quad \begin{aligned} \mathbb{P}(Y_n \geq a_n, B_n) &= \frac{1}{2^{\ell_n + 1}} \times \mathbb{E} \left[ \mathbb{P}(\mathcal{N}(0, \alpha^{2\ell_n + 2} A(\theta_1, \dots, \theta_{n - \ell_n - 1})^2 + u_{\ell_n}^2) \geq a_n | \mathcal{F}_{n - \ell_n - 1}) \right] \\ &\geq \frac{1}{2^{\ell_n + 1}} \times \mathbb{P}(\mathcal{N}(0, u_{\ell_n}^2) \geq a_n) \\ &\gtrsim \frac{1}{2^{\ell_n}} \times \frac{u_{\ell_n}}{a_n} \exp \left( -\frac{a_n^2}{2u_{\ell_n}^2} \right), \end{aligned}$$

where the last inequality follows from the classical lower bounds for the tail of standard normal distribution (see for e.g. [16]).

Now, from (8) and (9) we get

$$\frac{1}{2^{\ell_n}} \frac{u_{\ell_n}}{a_n} \exp \left( -\frac{a_n^2}{2u_{\ell_n}^2} \right) \lesssim \mathbb{P}(Y_n \geq a_n) \lesssim \frac{1}{a_n} \exp \left( -\frac{(1 - \alpha^2)a_n^2}{2} \right).$$

Finally, letting  $n$  go to infinity in the previous inequalities leads to

$$-\frac{1 - \alpha^2}{2} \leq \liminf_{n \rightarrow +\infty} \frac{1}{a_n^2} \log \mathbb{P}(Y_n \geq a_n) \leq \limsup_{n \rightarrow +\infty} \frac{1}{a_n^2} \log \mathbb{P}(Y_n \geq a_n) \leq -\frac{1 - \alpha^2}{2}$$

since  $\ell_n = o(a_n^2)$  and  $u_{\ell_n}^2 \rightarrow 1/(1 - \alpha^2)$ , which ends the proof of (6).

For the second part, we focus on the case  $\beta < \alpha$ , while the case  $\alpha = \beta$  is simpler. We write

$$\tau_n = \sup\{i = 1, \dots, n : \theta_i = \beta\}$$

with convention  $\sup \emptyset = 0$ . We first observe that

$$\mathbb{P}(Y_n \geq a_n; \tau_n = n - i) = \mathbb{P}(\theta_n = \alpha, \dots, \theta_{n - i + 1} = \alpha, \theta_{n - i} = \beta, Y_n \geq a_n)$$

and

$$\mathcal{L}(Y_n | \theta_n = \alpha, \dots, \theta_{n - i + 1} = \alpha, \theta_{n - i} = \beta) = \mathcal{L} \left( \alpha^i \beta Y_{n - i - 1} + \sum_{k=0}^i \alpha^k \varepsilon_{n - k} \right).$$



We get

$$\begin{aligned}
& \mathbb{P}(Y_n \geq a_n; \tau_n = n - i) \\
&= \left(\frac{1}{2}\right)^{i+1} \times \mathbb{P}\left(\alpha^i \beta Y_{n-i-1} + \sum_{k=0}^i \alpha^k \varepsilon_{n-k} \geq a_n\right) \\
&= \left(\frac{1}{2}\right)^{i+1} \times \mathbb{E}\left(\mathbb{P}\left(\mathcal{N}\left(0, \frac{1 - (\alpha^2)^{i+1}}{1 - \alpha^2} + (\alpha^i \beta)^2 A_{n-i-1}^2\right) \geq a_n \mid \mathcal{F}_{n-i}\right)\right) \\
&\leq \left(\frac{1}{2}\right)^{i+1} \times \mathbb{P}\left(\mathcal{N}\left(0, w_i^2\right) \geq a_n\right),
\end{aligned}$$

where using again  $A_{n-i-1} \leq 1/(1 - \alpha^2)$ ,  $w_i$  is a non-negative real number defined by

$$w_i^2 = \frac{1 - (\alpha^2)^{i+1}}{1 - \alpha^2} + \frac{(\alpha^i \beta)^2}{1 - \alpha^2} = \frac{1 - \alpha^{2i}(\alpha^2 - \beta^2)}{1 - \alpha^2}.$$

Then

$$\mathbb{P}(Y_n \geq a_n; \tau_n = n - i) \lesssim \left(\frac{1}{2}\right)^{i+1} \frac{1}{a_n} \exp\left(-\frac{a_n^2}{2w_i^2}\right)$$

and summing over  $i$ , we get by monotonicity of  $w_i$ ,

$$\mathbb{P}(Y_n \geq a_n, B_n^c) \lesssim \sum_{i=0}^{\ell_n} \left(\frac{1}{2}\right)^{i+1} \frac{1}{a_n} \exp\left(-\frac{a_n^2}{2w_i^2}\right) \lesssim \frac{1}{a_n} e^{-a_n^2/w_{\ell_n}^2}.$$

Let  $s \in \mathbb{N}$  such that  $\alpha^{2s} < \alpha^2 - \beta^2$ . Using now  $\mathbb{P}(Y_n \geq a_n) \geq \mathbb{P}(Y_n \geq a_n, B_{\ell_n+s})$  and (9), we get

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{P}(Y_n \geq a_n, B_n^c)}{\mathbb{P}(Y_n \geq a_n)} \lesssim \limsup_{n \rightarrow \infty} 2^{\ell_n+s} e^{-a_n^2(1/w_{\ell_n}^2 - 1/u_{\ell_n+s}^2)} = 0,$$

since  $1/w_{\ell_n}^2 - 1/u_{\ell_n+s}^2 \sim c\alpha^{2\ell_n}$  with  $c > 0$  and  $w_{\ell_n}/u_{\ell_n+s} \rightarrow 1$  and  $\ell_n = \mathcal{O}(\log a_n)$ . It ensures that

$$(10) \quad \lim_{n \rightarrow +\infty} \mathbb{P}(B_{\ell_n} \mid Y_n \geq a_n) = 1.$$

This proves a part of (7).

Besides, recalling that  $n - \ell_n \rightarrow \infty$ , we consider  $k_n \leq n$  such that  $k_n - \ell_n \rightarrow \infty$  and  $k_n = \mathcal{O}(\log a_n)$ .

Then, (10) ensures that

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{\ell=0, \dots, \ell_n} \left| \frac{Y_{n-\ell}}{a_n} - \alpha^\ell \right| > \varepsilon \mid Y_n \geq a_n\right) \leq \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{\ell=0, \dots, \ell_n} \left| \frac{Y_{n-\ell}}{a_n} - \alpha^\ell \right| > \varepsilon \mid Y_n \geq a_n, B_{k_n}\right).$$

Conditionally on  $B_{k_n}$ , we can write

$$\frac{Y_{n-\ell}}{a_n} = \frac{\alpha^{k_n-\ell+1}}{a_n} Y_{n-k_n-1} + \frac{1}{a_n} \sum_{k=0}^{k_n-\ell} \alpha^k \varepsilon_{n-k-\ell}$$

where  $Y_{n-k_n-1}$  and  $(\varepsilon_{n-k-\ell} : k = 0, \dots, k_n - \ell)$  are still independent and  $(\varepsilon_{n-k-\ell} : k = 0, \dots, k_n - \ell)$  are distributed as standard gaussian random variables. Then,

$$(11) \quad \frac{Y_{n-\ell}}{a_n} - \alpha^\ell = \frac{\alpha^{k_n-\ell+1}}{a_n} Y_{n-k_n-1} + \sum_{k=0}^{k_n-\ell} \left( \frac{\alpha^k}{a_n} \varepsilon_{n-k-\ell} - \alpha^\ell (1 - \alpha^2) \alpha^{2k} \right) + \eta_{n,\ell},$$

where

$$(12) \quad \eta_{n,\ell} = \alpha^\ell(1 - \alpha^2) \sum_{k=0}^{k_n - \ell} \alpha^{2k} - \alpha^\ell, \quad \sup_{\ell \in \{0, \dots, \ell_n\}} |\eta_{n,\ell}| \xrightarrow{n \rightarrow \infty} 0.$$

Next, for all  $\ell \in \{0, \dots, \ell_n\}$ , (8) yields

$$\mathbb{P} \left( \left| \frac{\alpha^{k_n - \ell + 1}}{a_n} Y_{n - k_n - 1} \right| \geq \varepsilon \right) \leq \mathbb{P} \left( |Y_{n - k_n - 1}| \geq \frac{a_n}{\alpha^{k_n - \ell_n}} \varepsilon \right) \lesssim \frac{\alpha^{k_n - \ell_n}}{a_n} \exp \left( - \frac{(1 - \alpha^2) \varepsilon^2 \alpha^{-2(k_n - \ell_n + 1)} a_n^2}{2} \right).$$

Dividing now by  $\mathbb{P}(Y_n \geq a_n)$  and using again the lowerbound (9) with  $k_n$  yields

$$\begin{aligned} & \sup_{\ell \in \{0, \dots, \ell_n\}} \mathbb{P} \left( \left| \frac{\alpha^{k_n - \ell + 1}}{a_n} Y_{n - k_n - 1} \right| \geq \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) \\ & \lesssim 2^{k_n} \alpha^{k_n - \ell_n} \exp \left( - \frac{(1 - \alpha^2) a_n^2}{2} \left( \varepsilon^2 \alpha^{-2(k_n - \ell_n + 1)} - \frac{1}{1 - (\alpha^2)^{k_n + 1}} \right) \right) \\ & = \mathcal{O}(1/\ell_n) \end{aligned}$$

since  $k_n - \ell_n \rightarrow \infty$  and  $k_n = \mathcal{O}(\log a_n)$ . Using  $\mathbb{P}(\sup_i X_i \geq a) \leq \sum_i \mathbb{P}(X_i \geq a)$  and the fact that  $Y_{n - k_n - 1}$  is independent of  $B_{k_n}$ , it ensures that

$$(13) \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left( \sup_{\ell \in \{0, \dots, \ell_n\}} \left| \frac{\alpha^{k_n - \ell + 1}}{a_n} Y_{n - k_n - 1} \right| \geq \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) = 0.$$

Finally, to control the remaining term in (11), we prove that

$$(14) \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left( \sup_{\ell \in \{0, \dots, \ell_n\}} \left| \sum_{k=0}^{k_n - \ell} \left( \frac{\alpha^k}{a_n} \varepsilon_{n - k - \ell} - \alpha^\ell (1 - \alpha^2) \alpha^{2k} \right) \right| \geq \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) = 0.$$

It amounts to solve a minimisation problem for the cost of a trajectory  $(\varepsilon_{n - k - \ell} : \ell \in \{0, \dots, \ell_n\})$ , under the constraint  $Y_n \geq a_n$ . This problem can be solved explicitly and we actually prove in Appendix that for any  $k \in \{0, \dots, k_n\}$

$$(15) \quad \mathbb{P} \left( \left| \frac{\alpha^k}{a_n} \varepsilon_{n - k} - (1 - \alpha^2) \alpha^{2k} \right| > \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) \lesssim \exp \left( -C \frac{a_n^2}{2} \right),$$

for some positive constant  $C$ , which yields directly (14).

Combining the last estimates ensures that

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \sup_{\ell \in \{0, \dots, \ell_n\}} \left| \frac{1}{a_n} Y_{n - \ell} - \alpha^\ell \right| \geq \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) = 0,$$

which ends the proof recalling (10).  $\square$

**2.2. Large deviations in the weakly stable case.** In this regime,  $\alpha\beta < 1$  and

$$\beta \leq 1 < \alpha.$$

Then there exists a unique  $\kappa > 0$  such that

$$\alpha^\kappa + \beta^\kappa = 2.$$

We also set  $\kappa_0 > 0$  and  $\gamma \in (0, 1)$ , respectively defined by

$$\log(\alpha) \alpha^{\kappa_0} + \log(\beta) \beta^{\kappa_0} = 0, \quad \gamma := (\alpha^{\kappa_0} + \beta^{\kappa_0})/2 = \inf_{s \geq 0} \{(\alpha^s + \beta^s)/2\}.$$

We introduce now the random walk  $(S_k)_{k \geq 1}$  defined by  $S_1 = 0$  and

$$S_k = \sum_{\ell=1}^{k-1} \log \theta_\ell \quad (k \geq 2).$$

Recalling from the introduction that  $A_n^* = A(\theta_n, \dots, \theta_1)$ , we can write

$$(A_n^*)^2 = \sum_{k=1}^n \exp(2S_k).$$

We restrict here the study to the case when  $\{A_n^* \geq a_n\}$  is comparable to  $\{A_\infty^* \geq a_n\}$ . In the proof below, it corresponds to the fact that the supremum

$$\bar{S} = \sup_{k \geq 1} S_k$$

is reached before time  $n$  when it is larger than  $a_n$ . It forces  $a_n$  to be not too large in the statement below.

**Theorem 2.** *Let  $(a_n, n \in \mathbb{N})$  be a sequence of real numbers such that  $a_n^\kappa \cdot n^{-3/2} \cdot \gamma^n \rightarrow 0$  as  $n \rightarrow \infty$ . Then,*

$$(16) \quad \mathbb{P}(Y_n \geq a_n) \asymp \mathbb{P}(A_n \geq a_n) \asymp a_n^{-\kappa}.$$

When the supremum is reached latter, additional work is needed and we expect that other equivalents can be proved using a change of probability.

*Proof.* First we determine the asymptotic behavior of  $\mathbb{P}(A_n \geq a_n)$  using the result of Kesten [36] which guarantees that  $(A_\infty^*)^2 = \sum_{k=1}^\infty \exp(2S_k)$  is comparable to  $\exp(2\bar{S})$ . More precisely, we first observe that

$$\mathbb{E}[(\theta_1^2)^\kappa] = 1 \quad \text{and} \quad \mathbb{E}[(\theta_1^2)^\kappa \log^+ \theta_1^2] < \infty.$$

Besides the following equalities hold:

$$2S_k = \sum_{\ell=1}^{k-1} \log \theta_\ell^2 \quad \text{and} \quad 2\bar{S} = \sup_k \sum_{\ell=1}^{k-1} \log \theta_\ell^2.$$

Note that for any  $a > 0$ ,  $\mathbb{P}(A_n^* \geq a) \leq \mathbb{P}((A_\infty^*)^2 \geq a^2)$ . Now, using the results of Section XI.6 in [30] and Section 1 in [36] (see also Section 2 in [28]), we get, for some positive constant  $C$ ,

$$\mathbb{P}((A_\infty^*)^2 \geq a^2) \sim_{a \rightarrow \infty} C \times \mathbb{P}(\exp(2\bar{S}) \geq a^2) = C \times \mathbb{P}(\exp(\bar{S}) \geq a).$$

Moreover, we know, see for e.g. Section XI.6 in [30], that

$$\mathbb{P}(\exp(\bar{S}) \geq a) \sim_{a \rightarrow \infty} c_0 a^{-\kappa},$$

where  $c_0$  is positive constant. We conclude that

$$\mathbb{P}(A_n^* \geq a_n) \lesssim \mathbb{P}(\exp(\bar{S}) \geq a_n) \sim c_0 a_n^{-\kappa} \quad \text{when } n \rightarrow \infty.$$

Besides, for any  $x \in (0, 1]$ ,

$$\mathbb{P}(A_n^* \geq a_n x) \geq \mathbb{P}\left(\sum_{k=1}^n \exp(2S_k) \geq a_n^2 x^2, \tau_{\bar{S}} \leq n\right) \geq \mathbb{P}(\exp(\bar{S}) \geq a_n x, \tau_{\bar{S}} \leq n),$$

where  $\tau_{\bar{S}} := \inf\{k > 0; S_k = \bar{S}\}$  is the first hitting time of the maximum of the random walk  $(S_k, k > 0)$ . Adding from Theorem 15, Chapter 4 of [17] that

$$\mathbb{P}(\tau_{\bar{S}} > n) \lesssim n^{-3/2} \gamma^n,$$

and using  $n^{-3/2}\gamma^n = \mathcal{O}(a_n^{-\kappa})$ , we get that  $\mathbb{P}(\tau_{\bar{S}} > n)$  is negligible compared to  $\mathbb{P}(\exp(\bar{S}) \geq a_n)$  and

$$\mathbb{P}(A_n^* \geq a_n x) \gtrsim \mathbb{P}(\exp(\bar{S}) \geq a_n x)$$

uniformly for  $n \geq 0$  and  $x \in (0, 1]$ .

Combining these estimates and recalling that the symmetry in law of  $(\theta_1, \dots, \theta_n)$  ensures that  $A_n$  is distributed as  $A_n^*$ , we get

$$(17) \quad \mathbb{P}(A_n \geq a_n x) = \mathbb{P}(A_n^* \geq a_n x) \asymp \mathbb{P}(\exp(\bar{S}) \geq a_n x)$$

uniformly for  $x \in (0, 1]$  and

$$(18) \quad \forall x \in (0, 1], \quad \mathbb{P}(\exp(\bar{S}) \geq a_n x) \sim_{n \rightarrow \infty} c_0 (x a_n)^{-\kappa}; \quad \sup_{x \in (0, 1], n \geq 0} \frac{\mathbb{P}(\exp(\bar{S}) \geq a_n x)}{(x a_n)^{-\kappa}} < \infty.$$

It ensures in particular that  $\mathbb{P}(A_n \geq a_n) \asymp a_n^{-\kappa}$  and ends the proof of the second part of the statement.

We prove now that  $\mathbb{P}(Y_n \geq a_n) \asymp a_n^{-\kappa}$ . Setting  $W_n = A_n/a_n$  and recalling that  $\mathcal{L}(Y_n | \mathcal{F}_{n-1}) = \mathcal{N}(0, a_n^2 W_n^2)$ , we first consider

$$\begin{aligned} \mathbb{P}(Y_n \geq a_n, W_n < 1) &\sim \mathbb{E} \left[ \mathbf{1}_{\{W_n < 1\}} W_n \exp\left(-\frac{1}{2W_n^2}\right) \right] \\ &= \int_0^1 \left(1 + \frac{1}{x^2}\right) \exp\left(-\frac{1}{2x^2}\right) \mathbb{P}(A_n \geq a_n x) dx. \end{aligned}$$

Besides, (18) and the fact that  $(1 + \frac{1}{x^2}) \exp(-\frac{1}{2x^2}) x^{-\kappa}$  is integrable on  $(0, 1]$  ensure by bounded convergence that

$$a_n^\kappa \int_0^1 \left(1 + \frac{1}{x^2}\right) \exp\left(-\frac{1}{2x^2}\right) \mathbb{P}(\exp(\bar{S}) \geq a_n x) dx \xrightarrow{n \rightarrow \infty} c \in (0, \infty).$$

Using the uniform estimate in (17) we obtain from the two previous displays that

$$\mathbb{P}(Y_n \geq a_n, W_n < 1) \asymp a_n^{-\kappa}.$$

Next, using again  $\mathcal{L}(Y_n | \mathcal{F}_{n-1}) = \mathcal{N}(0, a_n^2 W_n^2)$  and the fact that  $\mathbb{P}(\mathcal{N}(0, a_n^2 x^2) \geq a_n) \in [\mathbb{P}(\mathcal{N}(0, 1) \geq 1), 1]$  for  $x \geq 1$ , we get

$$\mathbb{P}(Y_n \geq a_n, W_n \geq 1) \asymp \mathbb{P}(W_n \geq 1) = \mathbb{P}(A_n \geq a_n) \asymp a_n^{-\kappa},$$

which ends the proof of the theorem.  $\square$

### 3. LOCAL DENSITIES FOR BIFURCATING AUTOREGRESSIVE PROCESSES

We can now study the autoregressive process

$$Z_n = \sum_{i \in \mathbb{G}_n} \delta_{X_i}$$

defined in Introduction. We recall that  $\mathbb{T}$  is the regular binary tree describing the underlying population and  $\mathbb{G}_n = \{2^n, 2^n + 1, \dots, 2^{n+1} - 1\}$  is the generation  $n$ .

### 3.1. The strictly stable case.

**Theorem 3.** *If  $\beta \leq \alpha < 1$ , then for any  $x \in [0, \sqrt{2 \log(2)/(1 - \alpha^2)})$ ,*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log Z_n([x\sqrt{n}; +\infty)) = \log 2 - x^2(1 - \alpha^2)/2 \quad a.s.$$

The proof relies on the large deviation results of the previous section and a law of large number principle, which is in the same spirit as proofs for branching random walk. The results of the previous section also tells us that the bulk of  $Z_n([x\sqrt{n}; +\infty))$  corresponds to individuals coming from the first daughter in the last divisions and that their value has undergone a geometric deviation from their stable distribution in these last divisions.

The proof of the upper bound is a classical Borel Cantelli argument using the first moment.

*Proof of the upper bound of Theorem 3.* Let  $\varepsilon > 0$  and write  $\alpha(x) = \log 2 - x^2(1 - \alpha^2)/2$ . By Markov inequality we have

$$\mathbb{P} \left( \frac{1}{n} \log Z_n([x\sqrt{n}; +\infty)) \geq \alpha(x) + \varepsilon \right) \leq \exp(-n(\alpha(x) + \varepsilon)) \times \mathbb{E} [Z_n([x\sqrt{n}; +\infty))].$$

Besides, the many-to-one formula (2) and (8) yield

$$\mathbb{E} [Z_n([x\sqrt{n}; +\infty))] = 2^n \mathbb{P}(Y_n \geq x\sqrt{n}) \lesssim 2^n \exp \left( -\frac{(1 - \alpha^2)x^2 n}{2} \right).$$

We obtain

$$\mathbb{P} \left( \frac{1}{n} \log Z_n([x\sqrt{n}; +\infty)) \geq \alpha(x) + \varepsilon \right) \leq \exp(-n\varepsilon)$$

and Borel Cantelli Lemma allows us to conclude that

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log Z_n([x\sqrt{n}; +\infty)) \leq \alpha(x) \quad a.s.$$

which ends the proof of the upper bound of Theorem 3.  $\square$

For the lower bound, we need now to prove a law of large number result on the subpopulation producing the local density  $[x\sqrt{n}, \infty[$  in generation  $n$ . Using the results of the previous section, it is achieved by following individuals who undergo a deviation in the last generation in environment  $\alpha$ . We first derive from the previous section the following result on the large deviations of  $Y$  and then proceed with the proof of the lower bound.

**Lemma 1.** *For any  $x \geq 0$*

$$(19) \quad \liminf_{p \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n+p} \log \mathbb{P}(Y_p \geq x\sqrt{n+p}) = -x^2(1 - \alpha^2)/2.$$

*Proof.* We first observe that by monotonicity of  $Y$  with respect to the initial condition and Markov property,  $\mathbb{P}(Y_{n+p} \geq x\sqrt{n+p}) \geq \mathbb{P}(Y_n \geq 0)\mathbb{P}(Y_p \geq x\sqrt{n+p}) = \mathbb{P}(Y_p \geq x\sqrt{n+p})/2$ . Using (6), we get the upper bound

$$\liminf_{p \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n+p} \log \mathbb{P}(Y_p \geq x\sqrt{n+p}) \leq -x^2(1 - \alpha^2)/2.$$

We prove now the converse inequality. We know from Theorem 1 that conditionally on  $\{Y_n \geq a_n\}$ , the process favors the best environment, that is the coefficient  $\alpha$ , at least in the last time. We then have

$$\begin{aligned} \mathbb{P}(Y_p \geq x\sqrt{n+p}) &\geq \mathbb{P}(Y_p \geq x\sqrt{n+p}, \theta_p = \alpha, \dots, \theta_1 = \alpha) \\ &= \frac{1}{2^p} \times \mathbb{P}\left(\sum_{\ell=0}^{p-1} \alpha^\ell \varepsilon_{p-\ell} \geq x\sqrt{n+p}\right) \\ &= \frac{1}{2^p} \times \mathbb{P}\left(\mathcal{N}(0, 1) \geq \sqrt{\frac{1-\alpha^2}{1-\alpha^{2p}}} x\sqrt{n+p}\right) \\ &\gtrsim \frac{1}{2^p} \times \frac{1}{x\sqrt{n+p}} \exp\left(-\frac{(1-\alpha^2)x^2(n+p)}{2(1-\alpha^{2p})}\right). \end{aligned}$$

Now, applying the log function in both sides of the last inequality, dividing by  $n+p$  and letting  $n$  and  $p$  go to infinity gives the expected lower bound and ends the proof.  $\square$

*Proof of the the lower bound of Theorem 3.* We first observe that for any  $n, p, a \geq 0$ ,

$$Z_{n+p}([a, \infty)) \geq \sum_{u \in \mathbb{G}_n: X_u \geq 0} \#\{v \in \mathbb{G}_{n+p} : u \preceq v, X_v \geq a\} \quad \text{a.s.}$$

Using the monotonicity of the autoregressive process with respect to its initial value and the branching property,

$$Z_{n+p}([a, \infty)) \geq \sum_{u \in \mathbb{G}_n: X_u \geq 0} X_{p,a}^{(u)} \quad \text{a.s.}$$

where  $(X_{p,a}^{(u)})_{u \in \mathbb{G}_n}$  are i.i.d. r.v. with the same law  $\mu_{p,a}$  defined by

$$\mu_{p,a} = \mathbb{P}_{\delta_0}(Z_p([a; +\infty)) \in \cdot).$$

Besides, Proposition 28 in [31] ensures that

$$\frac{1}{2^n} \#\{u \in \mathbb{G}_n : X_u \geq 0\} \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}_+} \mu(dx) \quad \text{a.s.},$$

where  $\mu$  is the law of  $Y_\infty^* = \sum_{k=1}^{\infty} (\prod_{\ell=1}^{k-1} \theta_\ell) \varepsilon_k$ . Actually, note that

$$\int_{\mathbb{R}_+} \mu(dx) = \lim_{n \rightarrow \infty} \mathbb{P}(Y_n \geq 0) = \frac{1}{2}.$$

Setting by now  $a = x\sqrt{n+p}$ , the many-to-one formula yields

$$u_{n,p} = \overline{\mu_{p,a}} = \mathbb{E}[X_{p,a}^{(u)}] = \mathbb{E}[Z_p[x\sqrt{n+p}, \infty)] = 2^p \mathbb{P}(Y_p \geq x\sqrt{n+p}).$$

Besides, the large deviation estimate proved above in Lemma 1 ensures that

$$(20) \quad \liminf_{p \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n+p} \log u_{n,p} = -x^2(1-\alpha^2)/2.$$

Recalling that  $x \in [0, \sqrt{2 \log(2)/(1-\alpha^2)})$ , we consider by now  $p$  large enough such that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log(\#\{u \in \mathbb{G}_n : X_u \geq 0\}) + \liminf_{n \rightarrow \infty} \frac{1}{n} \log u_{n,p} = \log(2) + \liminf_{n \rightarrow \infty} \frac{1}{n} \log u_{n,p} > 0.$$

Writing  $(X_{n,i} : i = 1 \dots, N_n) = (X_{p,a}^{(u)} : u \in \mathbb{G}_n, X_u \geq 0)$ , we are now in position to apply the law of large number given in Proposition 1 in Appendix 4.2 and get

$$\liminf_{n \rightarrow \infty} \frac{1}{\#\{u \in \mathbb{G}_n : X_u \geq 0\} u_{n,p}} \sum_{u \in \mathbb{G}_n : X_u \geq 0} X_{p,a}^{(u)} = 1 \quad \text{a.s.}$$

Then,

$$\liminf_{n \rightarrow \infty} \frac{1}{2^{n-1} u_{n,p}} Z_{n+p}([x\sqrt{n+p}, \infty)) \geq 1 \quad \text{a.s.}$$

and using (20) we obtain

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log Z_n([x\sqrt{n}, +\infty)) \geq \log(2) - x^2(1 - \alpha^2)/2 \quad \text{a.s.}$$

which ends the proof of the lower bound and Theorem 3.  $\square$

We derive now the asymptotic behavior of the highest value  $M_n = \max\{X_i : i \in \mathbb{G}_n\}$ .

**Corollary 1.**

$$\lim_{n \rightarrow +\infty} \frac{1}{\sqrt{n}} M_n = \sqrt{\frac{2 \log(2)}{1 - \alpha^2}} \quad \text{a.s.}$$

*Proof.* Setting  $v_+^2 = 2 \log(2)/(1 - \alpha^2) + \varepsilon$  for some  $\varepsilon > 0$ , we use the classical estimate

$$\mathbb{P}(M_n/\sqrt{n} \geq v_+) \leq \mathbb{E}(Z_n[\sqrt{n}v_+, \infty)) = 2^n \mathbb{P}(Y_n \geq \sqrt{n}v_+)$$

and recalling (6), we get  $\sum_n \mathbb{P}(M_n/\sqrt{n} \geq v_+) < \infty$ . Then  $\limsup_{n \rightarrow \infty} M_n/\sqrt{n} \leq v_+$  a.s. by Borel Cantelli lemma, which proves the upperbound letting  $\varepsilon \rightarrow 0$ .

Moreover, setting  $v_- = 2 \log(2)/(1 - \alpha^2) - \varepsilon$ , we observe that  $\{Z([\sqrt{n}v_-, \infty)) > 0\} \subset \{M_n \geq \sqrt{n}v_-\}$ . So

$$\liminf_{n \rightarrow \infty} M_n/\sqrt{n} \geq v_- \quad \text{a.s.}$$

is a direct consequence of Theorem 3. It ends the proof.  $\square$

**3.2. The weakly stable case.** We assume now that  $\alpha > 1$ . We recall that  $\alpha\beta < 1$ , so  $\beta < 1$ . Moreover  $\kappa \in (0, \infty)$  is defined by  $\alpha^\kappa + \beta^\kappa = 2$ .

**Theorem 4.** *For all  $c \in [0, \infty)$  such that  $c^\kappa < \min(2, 1/\gamma)$  and  $c > 1$ , we have*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log \#\{i \in \mathbb{G}_n : X_i \geq c^n\} = \log(2/c^\kappa) \quad \text{in probability.}$$

The proof of the upperbound is achieved by a classical Borel Cantelli argument using the first moment evaluated in the previous section, as for the strictly stable case. The proof of the lowerbound is different and more involved. We need to focus on a subtree producing the large values  $[c^n, \infty)$  at time  $n$  and characterized by  $A(\theta_1, \dots, \theta_n) \geq c^n$ . We then control correlations through the common ancestor of nodes and perform  $L^2$  estimates.

*Proof of the upperbound of Theorem 4.* Let  $\varepsilon > 0$ . By Markov inequality we have

$$\mathbb{P}\left(\frac{1}{n} \log \#\{i \in \mathbb{G}_n : X_i \geq c^n\} \geq \log(2/c^\kappa) + \varepsilon\right) \leq (c^\kappa/2)^n \exp(-n\varepsilon) \times \mathbb{E}[\#\{i \in \mathbb{G}_n : X_i \geq c^n\}].$$

Besides, using the many-to-one formula and (16),

$$\mathbb{E}[\#\{i \in \mathbb{G}_n : X_i \geq c^n\}] = 2^n \mathbb{P}(Y_n \geq c^n) \lesssim 2^n c^{-n\kappa}.$$

From the foregoing we are led to

$$\mathbb{P}\left(\frac{1}{n} \log \#\{i \in \mathbb{G}_n : X_i \geq c^n\} \geq \log(2/c^\kappa) + \varepsilon\right) \lesssim \exp(-n\varepsilon).$$

Finally, the Borel Cantelli Lemma allows us to conclude that

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log \#\{i \in \mathbb{G}_n : X_i \geq c^n\} \leq \log(2/c^\kappa) \quad \text{a.s.}$$

which ends the proof of the upper bound.  $\square$

Let us turn to the the proof of the lower bound and first give the outline. Recall that  $(\theta_n, n \geq 1)$  is a sequence of i.i.d. random variables having the same law that  $\theta$  and  $A_n = A(\theta_1, \dots, \theta_n)$ . For any  $i \in \mathbb{G}_n$ , we write  $(i_1, \dots, i_n) \in \{0, 1\}^n$  its binary decomposition. This decompositions yields the unique path in the tree from the root 1 to the vertex  $i$ . Setting  $\bar{0} = \alpha$  and  $\bar{1} = \beta$ , we define

$$A[i] = \sqrt{\sum_{k=1}^n \prod_{\ell=k+1}^n \bar{i}_\ell^2}$$

and we consider the subset of  $\mathbb{G}_n$  defined by

$$T_n = \{i \in \mathbb{G}_n : A[i] \geq c^n\}.$$

Note that we have

$$(21) \quad \mathbb{P}(A_n \geq c^n) = \frac{\#T_n}{2^n} \quad \text{and} \quad \#\{i \in \mathbb{G}_n : X_i \geq c^n\} \geq \#\{i \in T_n : X_i \geq c^n\} \quad \text{a.s.}$$

We set

$$F_n = \frac{\#\{i \in T_n : X_i \geq c^n\}}{\#T_n}.$$

The subtree  $T_n$  provides the bulk of  $Z_n[c^n, \infty)$ . Roughly the proof will follow from the fact that  $\liminf_{n \rightarrow +\infty} F_n > 0$  in probability. Indeed, using Theorem 2, it will ensure that

$$(22) \quad \frac{\#\{i \in \mathbb{G}_n : X_i \geq c^n\}}{2^n} \geq \frac{\#\{i \in T_n : X_i \geq c^n\}}{2^n} = F_n \times \mathbb{P}(A_n \geq c^n) \gtrsim (c^n)^{-\kappa}.$$

We actually prove using  $L^2$  computations that  $F_n$  is close to the non degenerated sequence

$$(23) \quad \begin{aligned} f_n &= \mathbb{E}[F_n] = \frac{1}{\#T_n} \mathbb{E}\left[\sum_{i \in T_n} \mathbf{1}_{\{X_i \geq c^n\}}\right] = \frac{1}{\#T_n} \mathbb{E}\left[\sum_{i \in \mathbb{G}_n} \mathbf{1}_{\{X_i \geq c^n\}} \mathbf{1}_{\{i \in T_n\}}\right] \\ &= \frac{2^n}{\#T_n} \times \mathbb{P}(Y_n \geq c^n, A_n \geq c^n) = \mathbb{P}(Y_n \geq c^n | A_n \geq c^n) \geq \mathbb{P}(\mathcal{N}(0, 1) \geq 1) > 0, \end{aligned}$$

where  $(Y_n, n \geq 0)$  is the autoregressive process with random coefficients defined in (3) and we use the many to one formula coupling the process and its environment.

For that purpose, for each  $i \in \mathbb{G}_n$ , we set

$$Z_n(i) = \mathbf{1}_{\{X_i \geq c^n\}} - \mathbb{P}(X_i \geq c^n).$$

Thus  $F_n - f_n = \frac{1}{\#T_n} \sum_{i \in T_n} Z_n(i)$  and

$$(24) \quad \mathbb{E}[(F_n - f_n)^2] = \frac{1}{\#T_n^2} \sum_{i \in T_n} \mathbb{E}[Z_n(i)^2] + \frac{1}{\#T_n^2} \sum_{\substack{(i,j) \in T_n^2 \\ i \neq j}} \mathbb{E}[Z_n(i)Z_n(j)].$$



For the first term of the right hand side of (24), we have

$$\frac{1}{\#T_n^2} \sum_{i \in T_n} \mathbb{E} [Z_n(i)^2] \leq \frac{1}{\#T_n^2} \sum_{i \in T_n} \mathbb{E} [\mathbf{1}_{\{X_i \geq c^n\}}] = \frac{\mathbb{P}(Y_n \geq c^n | A_n \geq c^n)}{\#T_n} \xrightarrow{n \rightarrow +\infty} 0,$$

since  $\#T_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Let us deal with the second term of (24). For all  $p \in \{0, \dots, n-1\}$ , we denote by  $\mathcal{T}_p^{(n)}$  the set of all the individuals of the generation  $p$  who are ancestors of at least one individual in the sub-population  $T_n$ . For each  $i \in \mathbb{G}_p$ , we denote by  $T_n(i)$  the set of individuals belonging to  $T_n$  who are descendants of  $i$ . Writing  $i \preceq j$  when  $i$  is an ancestor of  $j$ , it means that

$$\mathcal{T}_p^{(n)} = \{i \in \mathbb{G}_p : \exists j \in T_n \text{ such that } i \preceq j\} \quad \text{and} \quad T_n(i) = \{j \in T_n : i \preceq j\}.$$

Writing  $i \wedge j$  the most recent common ancestor of two individuals  $i$  and  $j$  and gathering the couples in function of their most recent common ancestor, we obtain

$$\begin{aligned} & \frac{1}{\#T_n^2} \sum_{\substack{(i,j) \in T_n^2 \\ i \neq j}} \mathbb{E} [Z_n(i)Z_n(j)] \\ &= \frac{1}{\#T_n^2} \sum_{p=0}^{n-1} \sum_{u \in \mathcal{T}_p^{(n)}} \sum_{\substack{(i,j) \in T_n^2 \\ i \wedge j = u}} \mathbb{E} [Z_n(i)Z_n(j)] \\ (25) \quad &= \frac{1}{\#T_n^2} \sum_{p=0}^{n-1} \sum_{u \in \mathcal{T}_p^{(n)}} 2\mathbb{E} \left[ \mathbb{E} \left[ \sum_{i \in T_n(2u)} Z_n(i) \middle| X_{2u} \right] \times \mathbb{E} \left[ \sum_{j \in T_n(2u+1)} Z_n(j) \middle| X_{2u+1} \right] \right]. \end{aligned}$$

For all  $p$  fixed in  $\{0, \dots, n-1\}$  and  $u \in \mathbb{G}_p$ , we set  $A_n^{p,u} = A(\bar{u}_1, \dots, \bar{u}_p, \theta_{p+1}, \dots, \theta_n)$  and for any  $x \in \mathbb{R}$ ,

$$P_n^p(x, u) = \mathbb{P} \left( Y_n \geq c^n \middle| Y_p = x, A_n^{p,u} \geq c^n \right)$$

and by convention  $P_n^p(x, u) = 0$  if  $\mathbb{P}(A_n^{p,u} \geq c^n) = 0$ . Besides the many-to-one formula ensures that

$$(26) \quad \mathbb{E} \left[ \sum_{i \in T_n(u)} Z_n(i) \middle| X_u \right] = \#T_n(u) (P_n^p(X_u, u) - \mathbb{E} [P_n^p(X_u, u)]).$$

The rest of the proof of the theorem relies on the two following lemmas. The first one ensures that an ergodic property holds on lineages, which allows to forget the beginning of the trajectory.

**Lemma 2.** *For any  $p \in \mathbb{N}$  and any  $u \in \mathbb{G}_p$ , we have*

$$P_n^p(X_u, u) - \mathbb{E} [P_n^p(X_u, u)] \xrightarrow{n \rightarrow +\infty} 0 \quad \text{in probability.}$$

*Proof of Lemma 2.* Let  $p$  be a fixed natural integer and let  $u \in \mathbb{G}_p$ . First we will prove that conditionally on  $\{A_n \geq c^n\}$ ,  $\prod_{k=1}^n \theta_k / c^n$  converges to 0 in probability. For that purpose we will show that for all  $\varepsilon > 0$ , we have

$$(27) \quad \mathbb{P} \left( \prod_{k=1}^n \theta_k \geq \varepsilon c^n \right) = o(\mathbb{P}(A_n \geq c^n)).$$

We set

$$S_k = \sum_{\ell=1}^k \log(\theta_\ell); \quad \bar{S} = \sup \{S_k, k \geq 1\} \quad \text{and} \quad \tau_{\bar{S}} = \inf \{k \geq 1, S_k = \bar{S}\}.$$

Let  $\xi \in (0, 1)$  and  $\delta > 0$  such that  $\log(\beta) \leq -\delta < \mathbb{E}[\log(\theta_1)] < 0$ . We have

$$\begin{aligned} \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n\right) &= \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n, \tau_{\bar{S}} > \xi n\right) + \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n, \tau_{\bar{S}} \leq \xi n\right) \\ &\leq \mathbb{P}(\tau_{\bar{S}} > \xi n) + \mathbb{P}(\bar{S} > n(\log(c) + \delta)) \\ (28) \quad &+ \sum_{j=1}^{\xi n} \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n, \tau_{\bar{S}} = j, S_j \leq n(\log(c) + \delta)\right). \end{aligned}$$

Using the asymptotic of  $\tau_{\bar{S}}$  given in Theorem 15, Chapter 4 of [17], we obtain

$$(29) \quad \mathbb{P}(\tau_{\bar{S}} > \xi n) \lesssim (\xi n)^{-3/2} \gamma^{\xi n}.$$

Next, from Kesten's results, see e.g. [36], Section 1 or [28] Section 2, we have

$$(30) \quad \mathbb{P}(\bar{S} \geq n(\log(c) + \delta)) \sim (c^n)^{-\kappa} \exp(-\kappa \delta n).$$

Finally, for the last term of the inequality (28) we have

$$\begin{aligned} &\sum_{j=1}^{\xi n} \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n, \tau_{\bar{S}} = j, S_j \leq n(\log(c) + \delta)\right) \\ &\leq \sum_{j=1}^{\xi n} \mathbb{P}(\exp(S_n - S_j) \geq \varepsilon \exp(-n\delta), \tau_{\bar{S}} = j, S_j \geq S_n \geq n \log(c) + \log(\varepsilon)) \\ &= \sum_{j=1}^{\xi n} \mathbb{P}\left(\tau_{\bar{S}} = j, S_j > n \log(c) + \log(\varepsilon)\right) \mathbb{P}\left(\exp(S_n - S_j) \geq \varepsilon \exp(-n\delta) \mid \tau_{\bar{S}} = j\right). \end{aligned}$$

We first observe that conditioning on  $\tau_{\bar{S}} = j$  make stochastically decrease the random after time  $j$  and

$$\begin{aligned} \mathbb{P}\left(\exp(S_n - S_j) \geq \varepsilon \exp(-n\delta) \mid \tau_{\bar{S}} = j\right) &\leq \mathbb{P}(S_n - S_j \geq -n\delta + \log(\varepsilon)) \\ &\lesssim \mathbb{P}(S_{n-j} \geq -n\delta) \\ &\leq \exp\left(-(n-j)\psi\left(-\frac{n}{n-j}\delta\right)\right), \end{aligned}$$

where  $\psi$  is the rate function associated to  $(S_k, k \geq 1)$ , see e.g. [23] Chapter 2 for more details. Since  $\psi$  is nonincreasing in  $(-\infty, \mathbb{E}[\log(\theta_1)]]$ , we have for  $j \in \{1, \dots, \xi n\}$ ,

$$\exp\left(-(n-j)\psi\left(-\frac{n}{n-j}\delta\right)\right) \lesssim \exp(-(1-\xi)n\psi(-\delta)).$$

From the foregoing we obtain

$$\begin{aligned} (31) \quad &\sum_{j=1}^{\xi n} \mathbb{P}\left(\prod_{k=1}^n \theta_k \geq \varepsilon c^n, \tau_{\bar{S}} = j, S_j \leq n(\log(c) + \delta)\right) \\ &\lesssim \mathbb{P}(S_{\tau_{\bar{S}}} > n \log(c)) \times \exp(-(1-\xi)n\psi(-\delta)) \lesssim (c^n)^{-\kappa} \times \exp(-(1-\xi)n\psi(-\delta)). \end{aligned}$$

From (29)-(31) we conclude that (27) holds and then that

$$\mathbb{P} \left( \prod_{k=1}^n \theta_k \geq \varepsilon c^n \mid A_n \geq c^n \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Similarly, it is easy to see that for  $p \in \mathbb{N}$  and  $u \in \mathbb{G}_p$  and  $x \in \mathbb{R}$ , we also have

$$(32) \quad \mathbb{P} \left( \left| x \frac{\prod_{k=p+1}^n \theta_k}{c^n} \right| \geq \varepsilon \mid A_n^{p,u} \geq c^n \right) \xrightarrow{n \rightarrow +\infty} 0.$$

We observe now that

$$(33) \quad P_n^p(x, u) = \mathbb{P} \left( x \frac{\prod_{k=p+1}^n \theta_k}{c^n} + \frac{\tilde{Y}_n}{c^n} \geq 1 \mid A_n^{p,u} \geq c^n \right),$$

where  $\tilde{Y}_n = \sum_{k=p+1}^n (\prod_{\ell=k+1}^n \theta_\ell) \varepsilon_k$ . Combining (33), (32) and the fact that  $\tilde{Y}_n/c^n$  is a mixed of Gaussian random variable and the fact that  $X_u$  is stochastically bounded (since gaussian), we get

$$P_n^p(X_u, u) - \mathbb{P} \left( \tilde{Y}_n \geq c^n \mid A_n^{p,u} \geq c^n \right) \rightarrow 0 \quad \text{in probability.}$$

By dominated convergence theorem, we obtain that  $\mathbb{E}[P_n^p(X_u, u)] - \mathbb{P} \left( \tilde{Y}_n \geq c^n \mid A_n^{p,u} \geq c^n \right)$  goes to 0 as  $n$  tends to infinity. The proof follows by combining these two limits.  $\square$

The next lemma proves that the common ancestor of lineages of the bulk  $Z_n[c^n, \infty)$  are at the beginning of the tree, thus providing the decorrelation needed for a law of large number.

**Lemma 3.**

$$\sup_{n > p} \left\{ \frac{\sum_{k=p}^{n-1} \sum_{u \in \mathcal{T}_k^{(n)}} \#T_n(2u) \#T_n(2u+1)}{\#T_n^2} \right\} \xrightarrow{p \rightarrow +\infty} 0.$$

*Proof of Lemma 3.* Let  $I_n$  and  $J_n$  be two independent indices uniformly drawn from  $T_n$  and independent of  $(X_i, i \in \mathbb{T})$ . Let  $p < n$ . We have

$$\mathbb{P}(p \leq |I_n \wedge J_n| < n) = \frac{1}{\#T_n^2} \sum_{(i,j) \in T_n^2, i \neq j} \mathbb{1}_{\{|i \wedge j| \geq p\}}.$$

Next, gathering pairs in function of their most recent common ancestor we obtain

$$\mathbb{P}(p \leq |I_n \wedge J_n| < n) = \frac{1}{\#T_n^2} \sum_{(i,j) \in T_n^2, i \neq j} \mathbb{1}_{\{|i \wedge j| \geq p\}} = \frac{1}{\#T_n^2} \sum_{k=p}^{n-1} \sum_{u \in \mathcal{T}_k^{(n)}} \#T_n(2u) \cdot \#T_n(2u+1).$$

Since for any  $u \in \mathcal{T}_p^{(n)}$ ,

$$\mathbb{P}(p \leq |I_n \wedge J_n| < n \mid u \preceq I_n) = \frac{\#T_n(u)}{\#T_n},$$

then

$$\mathbb{P}(p \leq |I_n \wedge J_n| < n) \leq \frac{\max \left\{ \#T_n(u) : u \in \mathcal{T}_p^{(n)} \right\}}{\#T_n}.$$

Moreover, since  $\alpha > \beta$ , we easily observe that

$$\max \left\{ \#T_n(u) : u \in \mathcal{T}_p^{(n)} \right\} = \#T_n(2^p).$$

Thus we need to check that

$$(34) \quad \frac{\sup_{n>p} \{ \#T_n(2^p) \}}{\#T_n} \xrightarrow{p \rightarrow +\infty} 0.$$

For that purpose, we set for  $p < n$ ,

$$N_1(i) = \# \{k : i_k = 1\}, \quad B_n^p = \{i \in T_n(2^p) : N_1(i) < \varepsilon(n-p)\},$$

where  $\varepsilon > 0$  is fixed. We recall that for any  $i \in \mathbb{G}_n$ ,  $(i_1, \dots, i_n) \in \{0, 1\}^n$  is the binary decomposition providing the unique path in the binary tree from the root to  $i$ . In word,  $B_n^p$  is the set of individuals of  $T_n(2^p)$  whose binary decomposition has less than  $\varepsilon(n-p)$  “1”. We first check that there exists  $\varepsilon > 0$  such that

$$(35) \quad \frac{\#B_n^p}{\#T_n} \xrightarrow{p \rightarrow +\infty} 0.$$

It means that the number of extremal individuals created by the prolific individual “ $2^p$ ” is negligible with respect to the total number of extremal individuals when  $p$  becomes large. Now let  $\psi$  be rate function associated to the simple random walk (with step 0 and 1 with probability  $1/2$ ). For  $n > p$  large enough, we have

$$\frac{\#B_n^p}{2^{n-p}} = \frac{\# \{i \in \mathbb{G}_{n-p} : N_1(i) < \varepsilon(n-p)\}}{2^{n-p}} \leq \exp(-(n-p)\psi(\varepsilon)).$$

Recalling that  $\#T_n \sim (2/c^\kappa)^n$ , there exists  $c_0 > 0$  such that

$$\frac{\#B_n^p}{\#T_n} \leq c_0 \left( \frac{\exp(\psi(\varepsilon))}{2} \right)^p \left( \frac{c^\kappa}{\exp(\psi(\varepsilon))} \right)^n.$$

Now we recall that  $c^\kappa < 2$  and  $\psi(0) = \log(2)$ . So we can choose  $\varepsilon > 0$  such that  $c^\kappa < \exp(\psi(\varepsilon))$ . Then

$$\frac{\#B_n^p}{\#T_n} \leq c_0 \left( \frac{\exp(\psi(\varepsilon))}{2} \right)^p \cdot \left( \frac{c^\kappa}{\exp(\psi(\varepsilon))} \right)^p \leq \left( \frac{c^\kappa}{2} \right)^p$$

for  $n > p$ . The last term tends to 0 as  $p$  tends to infinity, which yields (35).

Now using (35), a sufficient condition for (34) is

$$(36) \quad \frac{\#T_n(2^p) \setminus B_n^p}{\#T_n} \xrightarrow{p \rightarrow +\infty} 0,$$

where  $T_n(2^p) \setminus B_n^p$  is the set of individuals of  $T_n(2^p)$  whose the number of components which are equal to 1 is greater than  $\varepsilon(n-p)$ . We prove (36) as follows. For  $i \in \mathbb{G}_n$  and  $1 \leq a < b \leq n$ , we define  $\tau_{ab}i \in \mathbb{G}_n$  as the label obtained from  $i$  by permutation of  $i_a$  and  $i_b$  in the binary decomposition (or path in the binary tree). We also introduce

$$\mathcal{J}_n^k = \{u \in T_n(2^p) : N_1(u) = k\}$$

the set of individuals in  $T_n$ , whose ancestor in generation  $p$  is  $2^p$ , and which contains exactly  $k$  components equal to 1.

Let us first fix  $a$  with  $1 \leq a \leq p$  and consider the set

$$\mathcal{J}_n^k(a) = \{\tau_{ab}u : u \in \mathcal{J}_n^k, p+1 \leq b \leq n, u_b = 1\}.$$

First  $\mathfrak{J}_n^k(a) \cap T_n(2^p) = \emptyset$  and  $\mathfrak{J}_n^k(a) \subset T_n$  since by definition of  $A_n$  and using  $\beta < \alpha$ ,  $A_n(\tau_{ab}u) \geq A_n(u)$  for any  $u$  such that  $u_a = 0$  and  $u_b = 1$ . Moreover for any  $a \neq a'$  and  $k \neq k'$ ,  $\mathfrak{J}_n^k(a) \cap \mathfrak{J}_n^{k'}(a') = \emptyset$ , so

$$\#T_n \setminus T_n(2^p) \geq \# \bigcup_{\substack{k \geq (n-p)\varepsilon \\ 1 \leq a \leq p}} \mathfrak{J}_n^k(a) = \sum_{\substack{k \geq (n-p)\varepsilon, \\ 1 \leq a \leq p}} \#\mathfrak{J}_n^k(a).$$

Besides,

$$\#\mathfrak{J}_n^k(a) = \frac{k}{n-p-k} \#\mathfrak{J}_n^k,$$

since

- for any  $u \in \mathfrak{J}_n^k$ , we have  $k$  choices for  $b$  such that  $u_b = 1$ ;
- for any  $v \in \mathfrak{J}_n^k(a)$ , we have  $n-p-k$  choices for  $0$  in the  $n-p$  last components of  $v$  that you can change for  $1$ :  $\#\{(u, b) : u \in \mathfrak{J}_n^k, p+1 \leq b \leq n, u_b = 1, \tau_{ab}u = v\} = (n-p-k)$ .

Putting the three last facts together, we get

$$\#T_n \setminus T_n(2^p) \geq \sum_{\substack{(n-p)\varepsilon \leq k < n-p, \\ 1 \leq a \leq p}} \frac{k}{n-p-k} \#\mathfrak{J}_n^k \geq \varepsilon \sum_{\substack{(n-p)\varepsilon \leq k < n-p, \\ 1 \leq a \leq p}} \#\mathfrak{J}_n^k.$$

Excluding the path  $(0, \dots, 0, 1, \dots, 1)$  corresponding to  $k = n-p$ ,  $T_n(2^p) \setminus B_n^p$  is the union of the disjoint sets  $\mathfrak{J}_n^k$  for  $(n-p)\varepsilon \leq k < n-p$ . Thus, we get

$$\#T_n \setminus T_n(2^p) \geq \varepsilon p \times \#T_n(2^p) \setminus B_n^p$$

and

$$\limsup_{p \rightarrow \infty} \sup_{n \geq p} \frac{\#T_n(2^p) \setminus B_n^p}{\#T_n \setminus T_n(2^p)} = 0.$$

This proves (36) and ends the proof of Lemma 3.  $\square$

We can now proceed with the proof of Theorem 4.

*Proof of the lower bound of Theorem 4.* Now let  $\varepsilon > 0$ . From Lemma 3 let  $p_\varepsilon$  such that

$$\frac{1}{\#T_n^2} \sum_{k=p_\varepsilon}^{n-1} \sum_{u \in \mathcal{T}_k^{(n)}} \#T_n(2u) \#T_n(2u+1) \leq \varepsilon.$$

Adding that  $|Z_n(i)| \leq 1$  and splitting the sum for  $p$  larger than  $p_\varepsilon$ , we get

$$\frac{1}{\#T_n^2} \sum_{\substack{(i,j) \in T_n^2 \\ i \neq j}} \mathbb{E}[Z_n(i)Z_n(j)] \leq R_n + \varepsilon,$$

where

$$R_n = \frac{1}{\#T_n^2} \sum_{p=0}^{p_\varepsilon-1} \sum_{u \in \mathcal{T}_p^{(n)}} \sum_{\substack{(i,j) \in T_n^2 \\ i \wedge j = u}} \mathbb{E}[Z_n(i)Z_n(j)].$$

Besides, recalling (25)-(26),

$$R_n = \frac{1}{\#T_n^2} \sum_{p=0}^{p_\varepsilon-1} \sum_{u \in \mathcal{T}_p^{(n)}} 2 \cdot \#T_n(2u) \cdot \#T_n(2u+1) \\ \times [P_n^{p+1}(X_{2u}, 2u) - \mathbb{E}(P_n^{p+1}(X_{2u}, 2u))][P_n^{p+1}(X_{2u+1}, 2u+1) - \mathbb{E}(P_n^{p+1}(X_{2u+1}, 2u+1))].$$

Next,  $P_n^p$  is bounded and by dominated convergence theorem, Lemma 2 yields  $R_n \xrightarrow[n \rightarrow +\infty]{} 0$ . We conclude from the foregoing that

$$\frac{1}{\#T_n^2} \sum_{\substack{(i,j) \in T_n^2 \\ i \neq j}} \mathbb{E}[Z_n(i)Z_n(j)] \xrightarrow[n \rightarrow +\infty]{} 0$$

and then that

$$\mathbb{E}[(F_n - f_n)^2] \xrightarrow[n \rightarrow +\infty]{} 0.$$

Combining this limit with (22) and (23) ends the proof of the lowerbound and Theorem 4.  $\square$

Finally, we derive the asymptotic value of the maximal value  $M_n$  in the weakly stable case.

**Corollary 2.**

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log(M_n) = \frac{\log(2)}{\kappa} \quad \text{in probability.}$$

The proof follows the proof of Corollary 1 in the strictly stable case and is left to the reader. One need now to use the estimate of Theorem 2 for the upper bound and check inequalities in probability for the lower bound.

#### 4. APPENDIX

##### 4.1. Proof of (15) : minimization problem for large values in the strictly stable case.

We first focus on the minimization problem and then use it to get the estimates of probabilities given in (15). We set  $\vartheta = 1/\alpha^2 > 1$  for convenience and consider the space

$$H = \{y \in \mathbb{R}^{\mathbb{N}} : \sum_{k \geq 0} y_k^2 \vartheta^k < \infty\}.$$

It is easy to see that  $H$  is an Hilbert space endowed with the inner product  $\langle x, y \rangle = \sum_{k \geq 0} x_k y_k \vartheta^k$ . We write  $\|y\|^2 = \sum_{k \geq 0} y_k^2 \vartheta^k$  the associated norm. Thanks to Cauchy-Schwarz inequality, one can observe that  $y \rightarrow \sum_{k \geq 0} y_k$  is well defined and continuous on  $H$ . For  $D \subset H$ , we define

$$(37) \quad I(D) = \inf \left\{ \|y\|^2 : \sum_{j \geq 0} y_j \geq 1, y \in D \right\}.$$

and we have the following minimization result for a quadratic form with a linear constraint.

**Lemma 4.** *Writing  $v = ((1 - \alpha^2)\alpha^{2j})_{j \geq 0}$ , for all  $k \leq n$  and for all  $\varepsilon > 0$ , we have*

- i)  $I(H) = \|v\|^2 = 1 - \alpha^2$ ;
- ii)  $I(\{y \in H : |y_k - v_k| \geq \varepsilon\}) > I(H)$ ;
- iii)  $I(\{y \in H : |y_k - v_k| \geq \varepsilon, y_i = 0 \text{ for } i > n\}) \geq I(\{y \in H : |y_k - v_k| \geq \varepsilon\})$ .

*Proof.* First, since  $\|y\|^2$  is invariant by change of sign of a coordinate, we get

$$I(H) = \inf \left\{ \|y\|^2 : \sum_{j \geq 0} y_j = 1, y \in H \right\}.$$

We observe that for all  $y \in H$  such that  $\sum_{j \geq 0} y_j = 1$ , the vector  $z = y - v$  belongs to  $H$  and satisfies  $\sum_{j \geq 0} z_j = 0$ . Then  $y = v + z$  and  $\langle v, z \rangle = 0$  and

$$\|y\|^2 = \|v\|^2 + 2\langle v, z \rangle + \|z\|^2 = \|v\|^2 + \|z\|^2$$

which proves *i*) and *ii*). Next, *iii*) is a direct consequence of

$$\{y \in H : |y - v_k| \geq \varepsilon, \text{ and } y_i = 0 \text{ for } i > n\} \subset \{y \in H : |y - v_k| \geq \varepsilon\}.$$

which ends the proof.  $\square$

We can now deal with the proof of (15). We recall that  $(k_n, n \in \mathbb{N})$  is a sequence of integers smaller than  $n$  and tending to infinity such that  $k_n = \mathcal{O}(\log(a_n))$  and  $B_{k_n} = \{\theta_n = \alpha, \dots, \theta_{n-k_n} = \alpha\}$ . For each  $n \geq 1$ , we subdivide  $[-1, 1]$  as a collection of  $N_n$  successive disjoint intervals  $(I_i^n : 1 \leq i \leq N_n)$  of length (at most)  $1/n^{3/2}$ :

$$[-1, 1] = \cup_{i=1}^{N_n} I_i^n, \quad |I_i^n| \leq 1/n^{3/2}, \quad I_i^n \cap I_j^n = \emptyset \text{ for } i \neq j, \quad N_n \leq 2(n^{3/2} + 1)$$

and either  $I_i^n \subset \mathbb{R}_+$  or  $I_i^n \subset \mathbb{R}_-$ . We set  $I_1^n = [0, 1/n^{3/2}]$ ,  $I_2^n = [-1/n^{3/2}, 0)$  and  $I_0^n = (-\infty, -1) \cup (1, +\infty)$ . Then we can write  $\mathbb{R} = \cup_{i=0}^{N_n} I_i^n$ . For  $k = 0, \dots, k_n$  and  $\varepsilon > 0$ , we write  $\mathcal{I}^n(k)$  the set of paths  $\mathbf{i} \in \{0, \dots, N_n\}^n$  such that the  $k^{\text{th}}$  interval,  $I_{i_k}^n$ , does not intersect  $[v_k - \varepsilon, v_k + \varepsilon]$ :

$$\mathcal{I}^n(k) = \{\mathbf{i} \in \{0, \dots, N_n\}^n : I_{i_k}^n \cap [v_k - \varepsilon, v_k + \varepsilon] = \emptyset\},$$

where we recall that  $v_k = (1 - \alpha^2)\alpha^{2k}$ . We define

$$\Xi_{j,n} = \frac{\alpha^{k_n+1}}{a_n} \theta_{n-k_n-1} \dots \theta_{n-j+1} \varepsilon_{n-j},$$

for  $j \in \{k_n + 2, \dots, n\}$  and  $\Xi_{j,n} = \frac{\alpha^{k_n+1}}{a_n} \varepsilon_{n-j}$ , for  $j \in \{0, \dots, k_n + 1\}$ . We set

$$E_n(\mathbf{i}) = \bigcap_{j=0}^n \left\{ \Xi_{j,n} \in I_{i_j}^n \right\}.$$

For all  $\varepsilon > 0$ , we have

$$(38) \quad \limsup_{n \rightarrow \infty} \mathbb{P} \left( \left| \frac{\alpha^k}{a_n} \varepsilon_{n-k} - v_k \right| > \varepsilon \mid Y_n \geq a_n, B_{k_n} \right) \leq \limsup_{n \rightarrow \infty} \sum_{\mathbf{i} \in \mathcal{I}^n(k)} \mathbb{P}(E_n(\mathbf{i}) \mid Y_n \geq a_n, B_{k_n}).$$

Let  $r \in (0, 1)$  fixed. We set  $\mathcal{I}_1^n(k) = \{\mathbf{i} \in \mathcal{I}^n(k) : (\mathbf{i}_{n^r}, \dots, \mathbf{i}_{n-1}) = \{1, 2\}^{n-n^r}\}$  and

$$\delta_n = 1 - \frac{(n - n^r)}{n^{3/2}} \quad \text{and} \quad D_{n,k} = \left\{ y \in \mathbb{R}^n; \sum_{\ell=0}^{n^r-1} y_\ell \geq 1 - \delta_n; |y_k - v_k| \geq \varepsilon \right\}.$$

Then, we have the following bounds.

**Lemma 5.** For all  $\mathbf{i} \in \mathcal{I}_1^n(k)^c$ ,

$$(39) \quad \mathbb{P}(E_n(\mathbf{i}) \mid Y_n \geq a_n, B_n) \lesssim \alpha^{n^r} n^{-3/2} u_n^{-1} \exp \left( -\frac{a_n^2}{2} \left( (\alpha^{-n^r} n^{-3/2})^2 - u_{k_n}^{-2} \right) \right).$$

For  $\mathbf{i} \in \mathcal{I}_1^n(k)$ ,

$$(40) \quad \mathbb{P}(E_n(\mathbf{i}) | Y_n \geq a_n, B_{k_n}) \lesssim (1/2)^{n-n^r} (n^{3/2})^{n^r} R_n,$$

where

$$R_n = \exp \left( -\frac{a_n^2}{2} \left( \inf_{y \in D_{n,k}} \sum_{j=0}^{n^r-1} \alpha^{-2j} y_j^2 - \frac{1-\alpha^2}{1-\alpha^{2k_n+2}} \right) \right).$$

*Proof of (39).* First, note that for all  $j \geq n^r$ , we have

$$\begin{aligned} \mathbb{P} \left( \Xi_{j,n} \geq \frac{1}{n^{3/2}} \mid Y_n \geq a_n, B_{k_n} \right) &\leq \frac{\mathbb{P}(\Xi_{n^r,n} \geq \frac{1}{n^{3/2}}) \mathbb{P}(B_{k_n})}{\mathbb{P}(Y_n \geq a_n, B_{k_n})} \\ &\leq \frac{\mathbb{P}(\varepsilon_{n-n^r} \geq a_n \alpha^{-n^r} n^{-3/2}) \mathbb{P}(B_{k_n})}{\mathbb{P}(Y_n \geq a_n, B_{k_n})} \\ &\lesssim \alpha^{n^r} n^{-3/2} u_{k_n}^{-1} \exp \left( -\frac{a_n^2}{2} \left( (\alpha^{-n^r} n^{-3/2})^2 - u_{k_n}^{-2} \right) \right). \end{aligned}$$

In the same way, we have

$$\mathbb{P} \left( \Xi_{j,n} \leq -\frac{1}{n^{3/2}} \mid Y_n \geq a_n, B_{k_n} \right) \lesssim \alpha^{n^r} n^{-3/2} u_{k_n}^{-1} \exp \left( -\frac{a_n^2}{2} \left( (\alpha^{-n^r} n^{-3/2})^2 - u_{k_n}^{-2} \right) \right),$$

which implies that

$$(41) \quad \forall j \geq n^r, \quad \mathbb{P} \left( |\Xi_{j,n}| \geq \frac{1}{n^{3/2}} \mid Y_n \geq a_n, B_{k_n} \right) \lesssim \alpha^{n^r} n^{-3/2} u_{k_n}^{-1} \exp \left( -\frac{a_n^2}{2} \left( (\alpha^{-n^r} n^{-3/2})^2 - u_{k_n}^{-2} \right) \right).$$

Now, for  $\mathbf{i} \in \mathcal{I}_1^n(k)^c$ , let  $j_n \geq n^r$  such that  $\mathbf{i}_{j_n} \neq -1$ . Using (41), we get

$$\begin{aligned} \mathbb{P}(E_n(\mathbf{i}) | Y_n \geq a_n, B_{k_n}) &\leq \mathbb{P} \left( \bigcap_{j=0}^{n^r-1} \{ \Xi_{j,n} \in I_{i_j}^n \} \cap \{ \Xi_{j_n,n} \in I_{i_{j_n}}^n \} \mid Y_n \geq a_n, B_{k_n} \right) \\ &\leq \mathbb{P} \left( |\Xi_{j_n,n}| \geq \frac{1}{n^{3/2}} \mid Y_n \geq a_n, B_{k_n} \right) \\ &\lesssim \alpha^{n^r} n^{-3/2} u_{k_n}^{-1} \exp \left( -\frac{a_n^2}{2} \left( (\alpha^{-n^r} n^{-3/2})^2 - u_{k_n}^{-2} \right) \right) \end{aligned}$$

and this ends the proof of (39).  $\square$

*Proof of (40).* We set  $y_i$  the lower bound of  $I_i$  if  $I_i \subset \mathbb{R}_+$  and the upper bound of  $I_i$  if  $I_i \subset \mathbb{R}_-$ . We note that for all  $j \in \{n^r, \dots, n-1\}$ , we have  $\{\Xi_{j,n} \in I_1^n\} \subset \{\varepsilon_{n-j} \geq 0\}$  and  $\{\Xi_{j,n} \in I_2^n\} \subset \{\varepsilon_{n-j} < 0\}$ . We also note that, for all  $j \in \{n^r, \dots, n-1\}$ , we have  $-1/n^{3/2} \leq \Xi_{j,n} \leq 1/n^{3/2}$  and then that

$$1 - \frac{(n-n^r)}{n^{3/2}} \leq 1 - \sum_{j=n^r}^{n-1} \Xi_{j,n} \leq 1 + \frac{(n-n^r)}{n^{3/2}}.$$

Now since  $Y_n/a_n = \sum_{j=0}^{n^r-1} \Xi_{j,n} + \sum_{j=n^r}^{n-1} \Xi_{j,n}$ , we get

$$(42) \quad \mathbb{P}(E_n(\mathbf{i}), Y_n \geq a_n, B_{k_n}) \leq (1/2)^{n-n^r} \mathbb{P}(D_n(\mathbf{i})),$$



where

$$D_n(\mathbf{i}) = \bigcap_{j=0}^{n^r-1} \left\{ \Xi_{j,n} \in I_{i_j}^n \right\} \cap \left\{ \sum_{j=0}^{n^r-1} \Xi_{j,n} \geq 1 - \delta_n \right\} \cap B_{k_n}.$$

Using the convention  $0/0 = 0$  and that

$$D_n(\mathbf{i}) \subset \bigcap_{j=0}^{n^r-1} \left\{ \Xi_{j,n} \in I_{i_j}^n \right\} \cap \left\{ \sum_{j=0}^{n^r-1} (y_{i_j} + n^{-3/2}) \geq 1 - \delta_n \right\} \cap B_{k_n},$$

we get

$$\begin{aligned} \mathbb{P}(D_n(\mathbf{i})) &\leq \mathbb{P}(B_{k_n}) \left( \prod_{j=0}^{n^r-1} \left( \frac{\alpha^j}{a_n |y_{i_j}|} \mathbf{1}_{\{y_{i_j} \neq 0\}} + \frac{1}{2} \mathbf{1}_{\{y_{i_j} = 0\}} \right) \right) \exp \left( -\frac{a_n^2}{2} \sum_{j=0}^{n^r-1} \alpha^{-2j} y_{i_j}^2 \right) \\ &\leq (n^{3/2})^{n^r} \exp \left( -\frac{a_n^2}{2} \inf_{y \in D_{n,k}} \sum_{j=0}^{n^r-1} \alpha^{-2j} y_j^2 \right) \mathbb{P}(B_{k_n}), \end{aligned}$$

where we used the fact that  $1/|y_{i_j}| \leq n^{3/2}$  if  $y_{i_j} \neq 0$  for the last inequality. From (42) we obtain

$$\begin{aligned} \mathbb{P}(E_n(\mathbf{i}) | Y_n \geq a_n, B_{k_n}) &\leq (1/2)^{n-n^r} (n^{3/2})^{n^r} \exp \left( -\frac{a_n^2}{2} \inf_{y \in D_{n^r,k}} \sum_{j=0}^{n^r-1} \alpha^{-2j} y_j^2 \right) \\ &\quad \times \mathbb{P}(B_{k_n}) \mathbb{P}(Y_n \geq a_n, B_{k_n})^{-1}. \end{aligned}$$

Finally, (40) follows using (9).  $\square$

As a direct consequence of (39) and (40), we have, for some positive constant  $C$ ,

$$(43) \quad \sum_{\mathbf{i} \in \mathcal{I}^n(k)} \mathbb{P}(E_n(\mathbf{i}) | Y_n \geq a_n, B_{k_n}) \lesssim (2n^{3/2})^{n^r} e^{-C \frac{a_n^2}{2} (\alpha^{-n^r} n^{-3/2})^2} + (2n^{3/2})^{n^r} (1/2)^{n-n^r} R_n.$$

Recalling (37), we obtain from Lemma 4 that

$$\inf_{y \in D_{n,k}} \sum_{j=0}^{n^r-1} \alpha^{-2j} y_j^2 \geq I(\{y \in \mathbb{R}^n; |y_k - v_k| \geq \varepsilon\}) > I(H) = 1 - \alpha^2.$$

Since  $1 - \alpha^{2k_n+2} \rightarrow 1$  when  $n \rightarrow \infty$ , we get

$$R_n \leq \exp \left( -C \frac{a_n^2}{2} \right),$$

for some positive constant  $C$ . From the latter inequality and (43), we get, since  $n \rightarrow \infty$ ,

$$\sum_{\mathbf{i} \in \mathcal{I}^n(k)} \mathbb{P}(E_n(\mathbf{i}) | Y_n \geq a_n, B_{k_n}) \lesssim e^{-C \frac{a_n^2}{2} (\alpha^{-n^r} n^{-3/2})^2} + e^{-C \frac{a_n^2}{2}} \lesssim e^{-C \frac{a_n^2}{2}}$$

and using (38) this ends the proof of (15).

**4.2. A law of large numbers.** We state here a strong law of large numbers useful for the proof of local densities in the strictly stable case.

**Proposition 1.** *Let  $\{\mathcal{F}_n\}_0^\infty$  be a filtration. Let  $\{X_{n,i} : n, i \geq 1\}$  be non-negative real valued random variables such that for each  $n$ , conditionally on  $\mathcal{F}_n$ ,  $\{X_{n,i} : i \geq 1\}$  are independent and identically r.v. distributed as  $X_n$ .*

*Let  $\{N_n : n \geq 1\}$  be non-negative integer valued r.v. such that for each  $n$ ,  $N_n$  is  $\mathcal{F}_n$  measurable and  $\liminf_{n \rightarrow \infty} N_{n+1}/N_n > 1$ .*

*We assume that  $X_{n,i}$  are uniformly bounded and  $\limsup_{n \rightarrow \infty} \mathbb{E}(X_n)/\mathbb{E}(X_{n+1}) < \liminf_{n \rightarrow \infty} N_{n+1}/N_n$ . Then*

$$\frac{1}{N_n \mathbb{E}(X_n)} \sum_{i=1}^{N_n} X_{n,i} \xrightarrow{n \rightarrow \infty} 1 \quad \text{a.s..}$$

This result is a consequence of a classical law of large numbers where the variables  $X_{n,i}$  depend both on  $n$  and  $i$ , see Kurtz [39] and Athreya Kang [3]. We prove here that gathering terms in a suitable way allows to deal with the case when  $\mathbb{E}(X_n) \rightarrow 0$ .

*Proof.* We write  $u_n = \lfloor 1/\mathbb{E}(X_n) \rfloor \vee 1$  and

$$\sum_{i=1}^{N_n} X_{n,i} = \sum_{k=0}^{V_n-1} Y_{n,k} + R_n,$$

where  $V_n = \lfloor N_n/u_n \rfloor$  goes to infinity and

$$Y_{n,k} = \sum_{i=ku_n}^{(k+1)u_n-1} X_{n,i}, \quad R_n = \sum_{i=V_n u_n}^{N_n} X_{n,i}.$$

Moreover, Markov inequality and boundedness of  $X_n$  yields for any  $t \geq 0$ ,  $n, k \geq 0$ ,

$$\mathbb{P}(Y_{n,k} > t | \mathcal{F}_n) \leq e^{-t} \mathbb{E}(e^{X_n})^{u_n} \leq e^{-t} (1 + C \mathbb{E}(X_n))^{u_n} \leq C' e^{-t}$$

for some constants  $C, C' \geq 0$ , since  $\mathbb{E}(X_n)$  is bounded. The right hand side is integrable on  $[0, \infty)$  and  $\mathbb{E}(Y_{n,k})$  is bounded and

$$\liminf_{n \rightarrow \infty} V_{n+1}/V_n \geq \liminf_{n \rightarrow \infty} N_{n+1}/N_n \cdot \liminf_{n \rightarrow \infty} \mathbb{E}(X_{n+1})/\mathbb{E}(X_n) > 1.$$

We can apply a law of large numbers of Athreya and Kang (Lemma 1 in [3]) to the family of centered and independent variables  $\{Y_{n,k} - \mathbb{E}(Y_{n,0}) : n \geq 1, k \geq 0\}$  which are stochastically dominated and get

$$\frac{1}{V_n} \sum_{k=0}^{V_n-1} (Y_{n,k} - \mathbb{E}(Y_{n,0})) \xrightarrow{n \rightarrow \infty} 0 \quad \text{a.s.}$$

and conclude using that  $\mathbb{E}(Y_{n,0}) = u_n \mathbb{E}(X_n)$  and  $V_n \sim N_n/u_n$  as  $n \rightarrow \infty$  and

$$(44) \quad \frac{R_n}{N_n} = \frac{\sum_{i=V_n u_n}^{N_n} X_{n,i}}{N_n} \xrightarrow{n \rightarrow \infty} 0 \quad \text{a.s.}$$

To justify this last limit, one can consider only the case when  $u_n$  goes geometrically to infinity, since otherwise the boundedness of the r.v.  $X_{n,i}$  allows to use a direct domination and Lemma 1 in [3]. When  $u_n$  grows geometrically, using again the law of large numbers in [3] yields

$$\frac{\sum_{i=V_n u_n}^{N_n} X_{n,i}}{u_n} \xrightarrow{n \rightarrow \infty} 1 \quad \text{a.s.}$$

and the fact that  $u_n/N_n \rightarrow 0$  a.s. ensures (44) and ends the proof.  $\square$

**Acknowledgment.** This work have been supported by the Chair “Modélisation Mathématique et Biodiversité” of VEOLIA Environnement-École Polytechnique-MNHN-F.X and ANR ABIM (ANR-16-CE40-0001).

## REFERENCES

- [1] Addario-Berry, L., and Reed, B. (2009). *Minima in branching random walks*, The Annals of Probability, 37(3), 1044–1079.
- [2] Aidékon, E. (2013). *Convergence in law of the minimum of a branching random walk*, The Annals of Probability, 41(3A), 1362–1426.
- [3] Athreya, K. B., and Kang, H. J. (1998). *Some limit theorems for positive recurrent branching Markov chains: I.*, Advances in Applied Probability, 30, 3, 693–710.
- [4] Athreya, K. B., and Ney, E. (1972). *Branching Process*, Springer, Berlin.
- [5] V. Bansaye (2008). *Proliferating parasites in dividing cells : Kimmel’s branching model revisited*, Ann. Appl. Probab., 18, 3, 967–996.
- [6] Basawa, I. V., and Huggins, R. M. (1999). *Extensions of the bifurcating autoregressive model for cell lineage studies*, Journal of Applied Probability, 36, 1225–1233.
- [7] Basawa, I. V., and Huggins, R. M. (2000). *Inference for the extended bifurcating autoregressive model for cell lineage studies*, Australian & New Zealand Journal of Statistics, 42, 423–432.
- [8] Basawa, I. V., and Zhou, J. (2004). *Non-Gaussian bifurcating models and quasi-likelihood estimation*, Journal of Applied Probability, 41, 55–64.
- [9] Basawa, I. V., and Zhou, J. (2005). *Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors*, Journal of Time Series Analysis, 26, 825–842.
- [10] Bercu, B., Blandin, V. (2015). *A Rademacher-Menchov approach for random coefficient bifurcating autoregressive processes*, Stochastic Processes and their Applications, 125, 1218–1243.
- [11] Bercu, B., de Saporta, B., and Gégout-Petit, A. (2009). *Asymptotic analysis for bifurcating autoregressive processes via a martingale approach*, Electronic Journal of Probability, 87, 2492–2526.
- [12] Biggins, J. D. (1976). *The first-and last-birth problems for a multitype age-dependent branching process*, Advances in Applied Probability, 8(3), 446–459.
- [13] Bitseki Penda, S. V., and Djellout, H. (2014). *Deviation inequalities and moderate deviations for estimators of parameters in bifurcating autoregressive models*, Annales de l’IHP-PS, 50, 806–844.
- [14] Bitseki Penda, S. V., Djellout, H., and Guillin, A. (2014). *Deviation inequalities, Moderate deviations and some limit theorems for bifurcating Markov chains with application*, Annals of Applied Probability, 24, 1, 235–291.
- [15] Bitseki Penda, S. V., and Olivier, A. (2017). *Autoregressive Functions Estimation in Nonlinear Bifurcating Autoregressive Models*, Statistical Inference for Stochastic Processes, 20, 179–210.
- [16] Birnbaum, Z. W. (1942). *An inequality for Mill’s ratio*, Annals of Mathematical Statistics, 13, 2, 245–246.
- [17] Borovkov, A. A. (1976). *Stochastic processes in queueing theory*, Springer, New York.
- [18] Bramson, M. D. (1978). *Maximal displacement of branching Brownian motion*, Communications on Pure and Applied Mathematics, 31(5), 531–581.
- [19] Chartier T. N., Hyenne V., Labbé J.C. (2010). *Mechanisms of asymmetric cell division: from model organisms to tumorigenesis*, Med Sci (Paris), 26(3), 25–258.
- [20] Cowan, R. (1984). *Statistical concepts in the analysis of cell lineage data*, In Proceeding of the 1983 Workshop on Cell Growth and Division, 18–22, LaTrobe University.

- [21] Cowan, R., and Staudte, R. G. (1986). *The bifurcating autoregression model in cell lineage studies*, Biometrics, 42, 769–783.
- [22] Delmas, J. F., Marsalle, L. (2010). *Detection of cellular aging in Galton-Watson process*, Stochastic Processes and their Applications, 12, 2495–2519.
- [23] Dembo, A., and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, Springer-Verlag, New York.
- [24] de Saporta, B., Gégout-Petit, A., and Marsalle, L. (2011). *Parameters estimation for asymmetric bifurcating autoregressive processes with missing data*, Electronic Journal of Statistics, 5, 1313–1353.
- [25] de Saporta, B., Gégout-Petit, A., and Marsalle, L. (2012). *Asymmetry tests for bifurcating autoregressive processes with missing data*, Statistics & Probability Letters, 82 (2012), 1439–1444.
- [26] de Saporta, B., Gégout-Petit, A., and Marsalle, L. (2014). *Random coefficients bifurcating autoregressive processes*, ESAIM: Probability and Statistics, 18, 365–399.
- [27] de Saporta, B., Gégout-Petit, A., and Marsalle, L. (2014). *Statistical study of asymmetry in cell lineage data*, Computational Statistics & Data Analysis, 69 (2014), 15–39.
- [28] Enriquez, N., Sabot, C., and Zindy, O. (2009). *A probabilistic representation of constants in Kesten’s renewal theorem*, Probability Theory Related Fields, 144 , 3-4, 581–613.
- [29] Faraud, G., Hu, Y, and Shi, Z. (2012). *Almost sure convergence for stochastically biased random walks on trees*, Probability Theory Related Fields, 154 , 3-4, 621–660.
- [30] Feller, W. (2008). *An introduction to probability theory and its applications*, Vol. 2, 2nd ed, John Wiley & Sons.
- [31] Guyon, J. (2007). *Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging*, Ann. Appl. Probab., 17, 1538–1569.
- [32] Guyon, J., Bize, A., Paul, G., Stewart, E. J., Delmas, J. F., and Taddéi, F. (2005). *Statistical study of cellular aging*, CEMRACS 2004 Proceedings, ESAIM Proceedings, 14, 100–114.
- [33] Hammersley, J. M. (1974). *Postulates for subadditive processes*, Ann. Probab., 2(4), 652–680.
- [34] Hu, Y., and Shi, Z. (2009). *Minimal value and critical martingale convergence in branching branching random walks and directed polymers on disordered trees*, Ann. Probab., 37, 742–781.
- [35] Katajisto and al. (2015). *Stem cells. Asymmetric apportioning of aged mitochondria between daughter cells is required for stemness*, Science. 2015 Apr 17; 348(6232):340-3.
- [36] Kesten, H. (1973). *Random difference equations and renewal theory for products of random matrices*, Acta Math., 131, 207–248.
- [37] Kimmel, M. (1997). *Quasistationarity in a branching model of division-within-division*, Classical and Modern Branching Processes (K. B. Athreya and P. Jagers, eds.) 157–164. Springer, New York.
- [38] Kingman, J. F. C. (1975). *The first birth problem for an age-dependent branching process*, Ann. Probab., 3(5), 790–801.
- [39] Kurtz, T. (1972). *Inequalities for law of large numbers*, Ann. Math. Statist., 43,1874–1883.
- [40] Nakashima, M. (2013). *Minimal value of Branching Random Walks in Random Environment*, Journal of Theoretical Probability, 26, 1181–1217.
- [41] Roberts, M. I. (2013). *A simple path to asymptotics for the frontier of a branching Brownian motion*. Ann. Probab., Vol. 41, No. 5, 3518–3541
- [42] D. A. Sinclair and L. Guarente. (1997). *Extrachromosomal rDNA circles– a cause of aging in yeast*, Cell, 91(7):1033–1042.
- [43] Stewart, E. J., Madden, R., Paul, G., and Taddéi, F. (2005). *Aging and death in an organism that reproduces by morphologically symmetric division*, PLoS Biol, 3, 2: e45.

VINCENT BANSAYE, CMAP, ÉCOLE POLYTECHNIQUE, ROUTE DE SACLAY, 91128 PALAISEAU, FRANCE.  
*E-mail address:* `vincent.bansaye@polytechnique.edu`

S. VALÈRE BITSEKI PENDA, INSTITUT DE MATHÉMATIQUES DE BOURGOGNE, UMR 5584, CNRS, UNIVERSITÉ DE BOURGOGNE FRANCHE- COMTÉ, F-21000 DIJON, FRANCE.  
*E-mail address:* `Simeon-Valere.Bitseki-Penda@u-bourgogne.fr`