



**HAL**  
open science

## Splitting trees with neutral mutations at birth

Mathieu Richard

► **To cite this version:**

Mathieu Richard. Splitting trees with neutral mutations at birth. *Stochastic Processes and their Applications*, 2014, pp. Volume 124, Issue 10, October 2014, Pages 3206-3230. 10.1016/j.spa.2014.05.008 . hal-00834732v2

**HAL Id: hal-00834732**

**<https://hal.science/hal-00834732v2>**

Submitted on 10 Jun 2014

**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.

# Splitting trees with neutral mutations at birth

MATHIEU RICHARD<sup>1</sup>

June 10, 2014

## Abstract

We consider a population model where individuals behave independently from each other and whose genealogy is described by a chronological tree called splitting tree. The individuals have i.i.d. (non-exponential) lifetime durations and give birth at constant rate to clonal or mutant children in an infinitely many alleles model with neutral mutations.

First, to study the allelic partition of the population, we are interested in its frequency spectrum, which, at a fixed time, describes the number of alleles carried by a given number of individuals and with a given age. We compute the expected value of this spectrum and obtain some almost sure convergence results thanks to classical properties of Crump-Mode-Jagers (CMJ) processes counted by random characteristics.

Then, by using multitype CMJ-processes, we get asymptotic properties about the number of alleles that have undergone a fixed number of mutations with respect to the ancestral allele of the population.

*Key words:* Splitting tree; Crump-Mode-Jagers branching processes; neutral mutation; infinitely many alleles model; frequency spectrum; multitype branching processes; almost-sure limit theorem; epidemiology.

*MSC 2010:* Primary 60J80, secondary 60J28, 92D25, 60J85, 60J27, 92D10, 60G55.

## 1 Introduction

We consider a general branching model with neutral mutations occurring at birth. We suppose that individuals carry alleles, have i.i.d. (and not necessarily exponential) life lengths and give birth at constant rate  $b$  during their lives to children who can be mutants with probability  $p$  or clones of their parents with probability  $1 - p$ . We are working with an infinite alleles model, that is, when a mutation occurs, the allele of the mutant child was never encountered before. Moreover, mutations are neutral because they do not imply advantages or disadvantages (all individuals have identically distributed dynamics).

Without mutation, the model is linked to a genealogical tree called *splitting tree* [16, 17, 26]. Moreover, if  $\Xi(t)$  denotes the number of alive individuals at time  $t$ , the process  $\Xi := (\Xi(t), t \geq 0)$  is a *Crump-Mode-Jagers process* (or general branching process) [21] which is binary (births occur singly) and homogeneous (constant birth rate).

We are first interested in the allelic partition of the population and more precisely in properties of the *frequency spectrum*  $(M_t^{i,a}, i \geq 1)$  where  $M_t^{i,a}$  is the number of distinct alleles

---

<sup>1</sup>CMAP, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. mathieu.richard@cmap.polytechnique.fr

carried by exactly  $i$  individuals at time  $t$  and that appeared after time  $t - a$  (or equivalently the alleles whose ages are less than  $a$  at time  $t$ ). Roughly speaking, this spectrum classifies the different alleles depending on their ages and on their sizes.

This kind of question was first studied by Ewens [14] who discovered the well known 'sampling formula' named after him and which describes the law of the allelic partition for a Wright-Fisher model with neutral mutations. In our model, we cannot get a counterpart of the Ewens' sampling formula in the sense that we cannot compute the joint law of  $(M_t^{i,a}, i \geq 1)$  for fixed  $0 < a \leq t$ . However, we obtain two kinds of results concerning the frequency spectrum  $(M_t^{i,a}, i \geq 1)$ . We first compute the expected value of the number of alleles carried by  $i$  individuals at time  $t$  and with age  $a$ . When the process  $\Xi$  is supercritical, we also obtain the asymptotic behaviors of the frequency spectrum and of the relative abundances of alleles as  $t \rightarrow \infty$  on the survival event of  $\Xi$ .

Similar models to ours have been studied in the literature. In [18], Griffiths and Pakes study the case of a Bienaymé-Galton-Watson process where children can independently be mutants with a given probability: the authors obtained asymptotic results about the number of alleles and the frequency spectrum at generation  $n$  as  $n \rightarrow \infty$ . In [30], Pakes gets analogous properties concerning continuous-time Markov branching processes. In particular, his formula of the expected frequency spectrum can be seen as a counterpart of ours, stated in Section 3.1. These two works [18, 30] have recently been used by Kimmel and coworkers in several articles. In [24], the authors are interested in the evolution of parts of DNA called Alu sequences. They model the evolution of these sequences using the infinite-alleles branching process of [18] with a linear-fractional offspring distribution. In [37], the authors consider the model of Pakes [30] and they especially obtain an explicit expression of the limiting mean frequency spectrum in the particular case of birth and death processes and they also study the variance frequency spectrum.

In the two articles [3, 4], Bertoin considers an infinite alleles model with neutral mutations in a subcritical or critical Bienaymé-Galton-Watson process where individuals independently give birth to a random number of clonal and mutant children according to the same joint distribution. In [3], he defines a tree of alleles where all individuals of a common allele are gathered in clusters and specifies the law of the allelic partition of the total population by describing the joint law of the sizes of the clusters and the numbers of their mutant children. In [4], Bertoin obtains the convergence of the sizes of allelic families for a large initial population and a small mutation rate.

Recent results have been obtained about splitting trees with mutations. In [11], Delaporte studies sequences of splitting trees with general and neutral mutations occurring at birth and investigates scaling limits in a regime of large population size and rare mutations. In two related papers [6, 7], Champagnat and Lambert consider a model really close to ours: the authors also work with splitting trees but in their model, individuals independently experience mutations at Poisson times during their lives. In [6], explicit formulas about the expected frequency spectrum at time  $t$  are stated. In [7], the authors are interested in large and old families; they obtain asymptotic results about the sizes of the largest families and about the ages of the oldest ones as  $t \rightarrow \infty$ . Their model with Poissonian mutations and our model with mutations occurring at birth are compared in [8] in the particular case of exponential lifelengths.

Finally, Taïb [36] considers CMJ-processes with a more general kind of mutations at birth (for example, the probability that a mutation occurs can depend on the age of the mother) and thanks to the theory of CMJ-processes counted with characteristics, he obtains several

asymptotic results, especially about the frequency spectrum of the population. However, in our particular case, some of the non-explicit limits he obtained can be computed thanks to a recent paper of Lambert [26] giving for example the one-dimensional marginals of  $\Xi$  and its asymptotic behavior.

We then obtain properties about the number of mutations undergone by alleles. More specifically, for  $i \geq 0$  and  $t \geq 0$ , we study  $L_i(t)$  the number of alleles at time  $t$  that have been affected by  $i$  more mutations than the ancestral allele and  $K_i(t)$  the number of individuals that carry such alleles. Our model can represent the spread of an epidemic: individuals are infected hosts, deaths are recoveries or actual deaths and births are transmissions of the disease, which can mutate to new strains. Then,  $L^i(t)$  is simply the number of strains of the disease that are present at  $t$  and that are at a distance  $i$  from the original strain of the disease. Moreover, the process  $K_i$  describes the number of individuals infected by such strains of the disease.

We compute the expected values of  $L_i(t)$  and  $K_i(t)$  and obtain asymptotic results about  $K_i(t)$  as  $t \rightarrow \infty$  by considering a multitype CMJ-process where the type of an allele is the number of mutations it has undergone. Multitype branching processes are also used in carcinogenesis, that is, in the evolution of cancerous cells. In [10, 13], cancerous cells are modeled by a multitype branching process where a cell is of type  $k$  if it has undergone  $k$  mutations and where the more a cell has undergone mutations, the faster it grows. The object of study is the time  $\tau_k$  of appearance of the first cell of type  $k$ . Branching processes and birth and death processes appear in other works dealing with the evolution of cancerous cells (see for example [29, 20, 34, 35] and references therein). For instance, in [34], the authors study the arrival time of the first resistant cell in a model of cancerous cells undergoing a medical treatment and becoming resistant after having experienced a certain number of mutations.

The paper is organized as follows: in Section 2, we expose the model that we consider and give some of its properties that will be useful to get our main results. Section 3 is devoted to the study of the frequency spectrum and in Section 4, we are interested in properties of the number of mutations undergone by alleles.

## 2 Preliminaries

### 2.1 The model

In this paper, as a population model, we consider genealogical trees satisfying the branching property and called *splitting trees* [16, 17, 26]. They are random trees satisfying the following assumptions.

- At time  $t = 0$ , there is only one progenitor.
- All individuals have i.i.d. lifespans and reproduction behaviors.
- Conditional on her birth date  $\alpha$  and her lifespan  $\zeta$ , each individual gives birth to children at constant rate  $b \in (0, \infty)$  during  $(\alpha, \alpha + \zeta]$ .
- Births occur singly.

We denote the common lifespan distribution by  $\Lambda(\cdot)/b$  where  $\Lambda$  is a finite positive measure on  $(0, +\infty]$  with total mass  $b$  and called *lifespan measure* [26].

To this splitting tree, we add neutral mutations occurring at birth in the following way. We fix  $p \in (0, 1)$ . When a birth event occurs, independently of others individuals, with probability  $1 - p$ , a child is a clone of her mother, that is carries the same allele and a mutant with probability  $p$ . Moreover, when a mutant appears in the population, its allele was never carried before by any other individual. Thus, we consider an *infinitely many alleles* model with *neutral mutations* because all individuals behave in the same way regardless of allele. On Figure 1, one can find an illustration of this model.

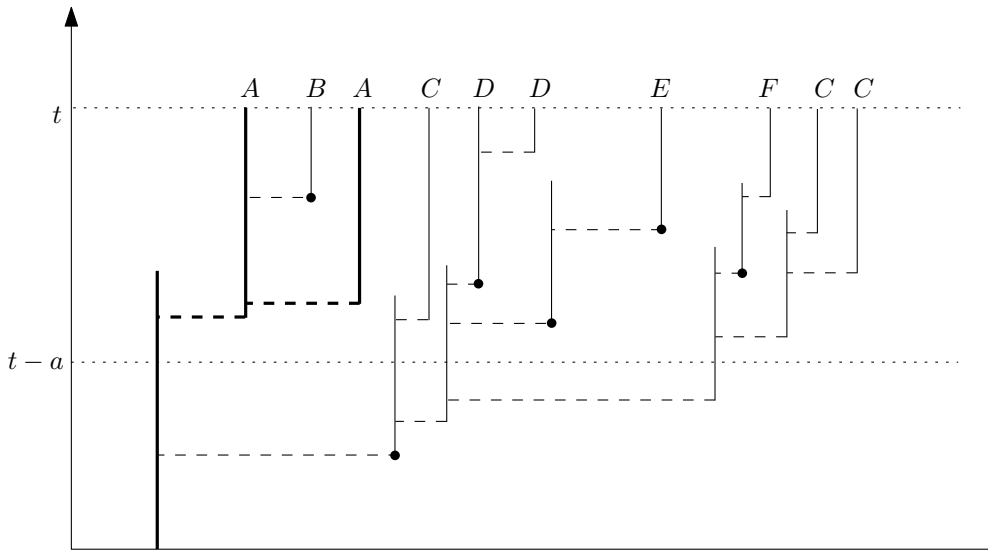


Figure 1: An example of a splitting tree with mutations up to time  $t$ . The vertical axis represents time and the horizontal axis shows filiation (lengths of dashed lined are meaningless). Full circles represent mutations occurring at birth and thick lines, the clonal splitting tree of the ancestor up to time  $t$ . The different letters are the alleles of live individuals at time  $t$ .

Without mutation, if  $\Xi(t)$  is the number of extant individuals at time  $t$ , then the process  $(\Xi(t), t \geq 0)$  is a Crump-Mode-Jagers (CMJ) process or general branching process (see [21, 22] and references therein). We use the formalism commonly employed in CMJ-processes. Denote by  $I = \bigcup_{n=0}^{\infty} (\mathbb{N}^*)^n$  the Ulam-Harris family history space with the convention  $\mathbb{N}^0 = \{0\}$ . The interpretation is that an individual  $\mathbf{i} = (i_1, \dots, i_{n-1}, i_n) \in I$  is the  $i_n$ -th child of the  $i_{n-1}$ -th child of ... of the  $i_1$ -th child of the ancestor, who is labeled  $\{0\}$ . For more details, the interested reader can see [21]. With each individual  $\mathbf{i} \in I$ , one associates a non-negative random variable  $\lambda_{\mathbf{i}}$  (its life length), and a point process  $\xi_{\mathbf{i}}$  called birth point process. The sequence  $(\lambda_{\mathbf{i}}, \xi_{\mathbf{i}})_{\mathbf{i} \in I}$  is assumed to be i.i.d. but  $\lambda_{\mathbf{i}}$  and  $\xi_{\mathbf{i}}$  are not necessarily independent.

In our particular case, the common distribution of lifespans is  $\Lambda(\cdot)/b$  and conditional on its lifespan, the birth point process of an individual is distributed as a Poisson point process during its life. If we denote by  $\xi$  the birth point process of the ancestor, its characteristic measure is then given by

$$\mu(dt) := \mathbb{E}[\xi(dt)] = dt\Lambda((t, +\infty]). \quad (1)$$

In the same way, if  $\xi_m$  (resp  $\xi_c$ ) is the birth point process of the mutant (resp. clonal) children of the ancestor, then by the thinning property of Poisson point processes,  $\xi_m$  and  $\xi_c$  are independent,  $\xi = \xi_m + \xi_c$  and with obvious notation,  $\mu_m(dt) = p\mu(dt)$  and  $\mu_c(dt) = (1-p)\mu(dt)$ .

## 2.2 Basic properties about the process $\Xi$

We consider here the model without mutation and recall some known facts about the CMJ-processes  $\Xi$ , which will be useful in the following.

We say that the process  $\Xi$  is subcritical, critical or supercritical if

$$m := \int_{(0,\infty]} r\Lambda(dr)$$

is respectively less than, equal to or greater than 1. We set

$$\psi(\lambda) := \lambda - \int_{(0,\infty]} (1 - e^{-\lambda r}) \Lambda(dr). \quad (2)$$

Since this function is convex and satisfies  $\psi(0^+) = 0$ , we can define  $\eta$  as its largest root. Moreover, since  $\psi'(0^+) = 1 - m$ , when the process  $\Xi$  is subcritical or critical, we have  $\eta = 0$  and when it is supercritical,  $\eta$  is positive. In the last case,  $\eta$  is called the *Malthusian parameter* of the population as explained in the forthcoming Proposition 2.1.

To obtain properties about the splitting trees, Lambert defined in [26] a contour process which characterized them. This contour process is a spectrally positive Lévy process with Laplace exponent  $\psi$ . Let  $W$  be the *scale function* associated with it (see [2, Ch. VII]), defined as the unique increasing continuous function  $(0, \infty) \rightarrow (0, \infty)$  satisfying

$$\int_0^\infty W(x)e^{-\lambda x} dx = \frac{1}{\psi(\lambda)}, \quad \lambda > \eta. \quad (3)$$

In the entire paper, by convenience, we assume that the following hypothesis holds.

$$\text{The measure } \Lambda \text{ has no atom.} \quad (\text{H})$$

According to [25, p. 234], under the hypothesis (H), the function  $W$  is continuously differentiable and thanks to the Lemma 4.1 in [33], we have

$$W'(t) = bW(t) - \int_0^t W(t-x)\Lambda(dx), \quad t \geq 0. \quad (4)$$

In fact, if (H) is not satisfied, most of the results stated in this paper still hold by replacing  $W'(t)$  by  $bW(t) - \int_0^t W(t-x)\Lambda(dx)$ . We mainly choose to assume (H) in order to simplify the results and their proofs.

Thanks to the scale function  $W$ , we deduce the one-dimensional marginals of  $\Xi$  (see [33, p. 293])

$$\mathbb{P}(\Xi(t) = 0) = 1 - \frac{W'(t)}{bW(t)} \quad (5)$$

and for  $n \in \mathbb{N}^*$ ,

$$\mathbb{P}(\Xi(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}. \quad (6)$$

In other words, conditional on being positive,  $\Xi(t)$  is geometric with success probability  $1/W(t)$ . In particular, for  $t \geq 0$ ,

$$\mathbb{E}[\Xi(t)] = \frac{W'(t)}{b}. \quad (7)$$

Let  $\text{Ext} := \left\{ \Xi(t) \xrightarrow[t \rightarrow \infty]{} 0 \right\}$  be the extinction event of the splitting tree. We denote by  $\mathbb{P}^* := \mathbb{P}(\cdot | \text{Ext}^c)$  the law of the process  $\Xi$  conditional on its survival event.

**Proposition 2.1** (Lambert [26]). *We have*

$$\mathbb{P}(\text{Ext}) = 1 - \eta/b.$$

Moreover, if  $\Xi$  is supercritical ( $m > 1$ ), under  $\mathbb{P}^*$ ,

$$e^{-\eta t} \Xi(t) \xrightarrow[t \rightarrow \infty]{} \mathcal{E} \quad \text{a.s.} \quad (8)$$

where  $\mathcal{E}$  is exponential with parameter  $\psi'(\eta)$ .

In fact, Lambert proved the convergence in distribution in [26] and a.s. convergence holds according to [28] (see [33, p.285]). The convergence (8) justifies why we call  $\eta$  the Malthusian parameter of the population since  $\Xi(t)$  grows like  $e^{\eta t}$  on the non-extinction event.

Most of the results stated in Sections 3 and 4 deal with long-time behaviors of several processes. To obtain them, we need the asymptotic behaviors of the scale function  $W$  and of its derivative  $W'$ . Different regimes appear depending on whether  $\Xi$  is subcritical, critical or supercritical. We record them in the following lemma.

**Lemma 2.2.** (i) *If  $m > 1$  (supercritical case), as  $t \rightarrow \infty$ , we have*

$$W(t) \sim \frac{1}{\psi'(\eta)} e^{\eta t} \quad \text{and} \quad W'(t) \sim \frac{\eta}{\psi'(\eta)} e^{\eta t}.$$

(ii) *If  $m = 1$  (critical case) and  $\sigma^2 := \int_0^\infty r^2 \Lambda(dr) < \infty$ , we have  $W(t) \underset{t \rightarrow \infty}{\sim} \frac{2}{\sigma^2} t$ . If we also suppose that*

$$\lim_{t \rightarrow \infty} t^2 \int_t^\infty (x-t) \Lambda(dx) = 0, \quad (9)$$

*we have  $\lim_{t \rightarrow \infty} W'(t) = \frac{2}{\sigma^2}$ .*

(iii) *If  $m < 1$  (subcritical case),  $\lim_{t \rightarrow \infty} W(t) = \frac{1}{1-m}$ . If we also assume that there is a negative  $\tilde{\eta}$  satisfying*

$$\psi(\tilde{\eta}) = 0 \quad \text{and} \quad \int_0^\infty r e^{-\tilde{\eta} r} \Lambda(dr) < \infty, \quad (10)$$

*then  $W'(t) \underset{t \rightarrow \infty}{\sim} \frac{\tilde{\eta}}{\psi'(\tilde{\eta})} e^{\tilde{\eta} t}$ .*

In the critical case, the condition (9) holds if  $\Lambda$  has a finite third moment. Concerning the subcritical case, it is possible to find a negative root of  $\psi$  because this function is convex,  $\psi(0) = 0$  and  $\psi'(0) = 1 - m > 0$ . However, to have this root requires that the exponential moment  $\int_0^\infty e^{\lambda r} \Lambda(dr)$  is finite for  $\lambda$  large enough. The proof of Lemma 2.2 is postponed to the appendix.

### 2.3 The clonal process

In the sequel, an important role will be played by the clonal process  $(\Xi_c(t), t \geq 0)$  ( $c$  for clonal) where  $\Xi_c(t)$  is the number of extant individuals at time  $t$ , carrying the same allele as the ancestor (see Figure 1). Since mutations occur independently of  $(\Xi(t), t \geq 0)$ , by the thinning property of Poisson processes, the process  $\Xi_c$  defines a splitting tree with lifespan measure

$$\Lambda_c := (1 - p)\Lambda.$$

Moreover, conditional on its arrival time in the population, each allele evolves like  $\Xi_c$  and independently of other alleles.

According to the previous section, the process  $\Xi_c$  is subcritical, critical or supercritical if  $(1 - p)m$  is respectively less than, equal to, or greater than 1. In particular, when  $\Xi$  is critical or subcritical,  $\Xi_c$  is obviously subcritical.

Furthermore, we associate with  $\Xi_c$  the function  $\psi_c$ , which satisfies

$$\psi_c(\lambda) := \lambda - (1 - p) \int_{(0, \infty]} (1 - e^{-\lambda r}) \Lambda(dr) = p\lambda + (1 - p)\psi(\lambda). \quad (11)$$

Let  $\eta_c$  be the largest root of  $\psi_c$ . When  $\Xi_c$  is subcritical or critical,  $\eta_c = 0$  while in the supercritical case,  $\eta_c > 0$ . Moreover, in the latter, by the definition of  $\eta_c$  and by using (11), we have

$$\psi(\eta_c) = \frac{p\eta_c}{p - 1} < 0.$$

It implies that  $\eta_c < \eta$  since  $\psi$  is convex and its largest root is  $\eta$ .

Finally, all the properties about  $\Xi$ , stated in the paragraph 2.2, also hold for  $\Xi_c$ . To obtain them, it suffices to respectively replace  $b$ ,  $\psi$ ,  $\eta$  and  $W$  by  $b(1 - p)$ ,  $\psi_c$ ,  $\eta_c$  and  $W_c$  where  $W_c$  is the scale function associated with  $\psi_c$ , solution of

$$\int_0^\infty W_c(x) e^{-\lambda x} dx = \frac{1}{\psi_c(\lambda)}, \quad \lambda > \eta_c.$$

## 3 Frequency spectrum

At a given time, for any allele, we call *family* the set of all individuals that share this allele. To study the allelic partition of the population, we study its associated frequency spectrum, which, roughly speaking, sort the different families according to their sizes and their ages. More precisely, for  $i \in \mathbb{N}^*$  and  $a > 0$ , let  $M_t^{i,a}$  be the number of alleles whose ages are less than  $a$  and carried by  $i$  individuals at  $t$ . Then, for fixed  $0 \leq a < t$ , the sequence  $(M_t^{i,a}, i \geq 1)$  is the frequency spectrum at time  $t$  of families with ages less than  $a$ .

Notice that  $M_t^{i,t}$  is simply the number of alleles carried by  $i$  particles at time  $t$  (regardless of their ages) and that  $M_t := \sum_{i \geq 1} M_t^{i,t}$  is the number of different alleles at time  $t$ .

In the example on Figure 1, taking no account of ages of alleles, the frequency spectrum  $(M_t^{i,t}, i \geq 1)$  is  $(3, 2, 1, 0, \dots)$  because three alleles ( $B, E, F$ ) are carried by one individual,  $A$  and  $D$  are carried by two individuals and  $C$  is the only allele carried by three individuals. If we are only interested in alleles younger than  $a$  at time  $t$ , we have  $(M_t^{i,a}, i \geq 1) = (3, 1, 0, \dots)$



since alleles  $A$  and  $D$  appeared in the population before  $t - a$ .

Although it is not possible to obtain the joint distribution of  $(M_t^{i,a}, i \geq 1)$  for fixed  $t$  and  $a$  as in the 'Ewens sampling formula', we are able to get some properties of this frequency spectrum.

### 3.1 Expected frequency spectrum

We first give an exact expression of the expected frequency spectrum at any time  $t$ . For  $0 < a < t$  and  $i \geq 1$ , we denote by  $M_t^{i,da}$  the number of alleles carried by  $i$  individuals at time  $t$  and with ages in  $[a - da, a]$ . The following proposition yields its expected value.

**Proposition 3.1.** *For  $0 < a < t$  and  $i \geq 1$ , we have*

$$\mathbb{E} \left[ M_t^{i,da} \right] = \frac{p}{b(1-p)} W'(t-a) \left( 1 - \frac{1}{W_c(a)} \right)^{i-1} \frac{W'_c(a)}{W_c^2(a)} da. \quad (12)$$

*Proof.* Conditional on  $\Xi(t-a)$ , thanks to the branching property and classical properties of Poisson point processes,  $M_t^{i,da}$  is the sum of  $\Xi(t-a)$  i.i.d. random variables distributed as the number of atoms in the interval  $[t-a, t-a+da]$  of a Poisson point process with parameter  $bp\mathbb{P}(\Xi_c(a) = i)$ . Hence,

$$\mathbb{E} \left[ M_t^{i,da} \right] = \mathbb{E}[\Xi(t-a)] bp \mathbb{P}(\Xi_c(a) = i) da.$$

According to (6) and (7),  $\mathbb{E}[\Xi(t-a)] = W'(t-a)/b$  and

$$\mathbb{P}(\Xi_c(a) = i) = \left( 1 - \frac{1}{W_c(a)} \right)^{i-1} \frac{W'_c(a)}{b(1-p)W_c(a)^2}$$

and we obtain the desired result.  $\square$

As a consequence of this proposition, we deduce the expected value of the frequency spectrum. In particular, when  $\Xi$  is supercritical, we see that  $M_t^{i,a}$  grows as  $e^{\eta t}$  when  $t \rightarrow \infty$ , that is, with the same growth rate as  $\Xi(t)$ .

**Corollary 3.2.** *For  $a \leq t$  and  $i \geq 1$ ,*

$$\begin{aligned} \mathbb{E} \left[ M_t^{i,a} \right] &= \frac{p}{b(1-p)} \int_0^a W'(t-x) \left( 1 - \frac{1}{W_c(x)} \right)^{i-1} \frac{W'_c(x)}{W_c^2(x)} dx \\ &\quad + \frac{1}{b(1-p)} \left( 1 - \frac{1}{W_c(t)} \right)^{i-1} \frac{W'_c(t)}{W_c^2(t)} \mathbf{1}_{\{t=a\}}. \end{aligned} \quad (13)$$

Moreover, if  $m > 1$ , for  $a \geq 0$ , as  $t \rightarrow \infty$ ,

$$e^{-\eta t} \mathbb{E} \left[ M_t^{i,a} \right] \rightarrow \frac{\eta}{b} \frac{p}{1-p} \frac{J^{i,a}}{\psi'(\eta)} \quad (14)$$

where

$$J^{i,a} := \int_0^a e^{-\eta u} \left( 1 - \frac{1}{W_c(u)} \right)^{i-1} \frac{W'_c(u)}{W_c^2(u)} du.$$

In [30], Pakes obtained a similar result: in Lemma 3.1.2, he computed the expected frequency spectrum for a Markov branching process where children can independently be mutants of their mothers with probability  $p$ . In [32, Chap 3], the expected frequency spectrum formula (13) is used to obtain the orders of magnitude of the sizes (resp. of the ages) of the largest families (resp. of the oldest families) at time  $t$  when  $t \rightarrow \infty$  (see also [8]).

*Proof of Corollary 3.2.* We obtain (13) by integrating (12) on  $(0, a)$ . The second term of the r.h.s. corresponds to  $\mathbb{P}(\Xi_c(t) = i)$ , that is, the probability that the progenitor has  $i$  alive clonal descendants at time  $t$ .

We get the convergence result (14) by using the dominated convergence theorem. First, as  $t \rightarrow \infty$ , using the Lemma 2.2,  $W'(t-x)$  is equivalent to  $\frac{\eta}{\psi'(\eta)} e^{\eta(t-x)}$ . Then, we obtain the convergence of the integral because  $\left(1 - \frac{1}{W_c(x)}\right)^{i-1} \frac{W'_c(x)}{W_c^2(x)} \leq b(1-p)$  and because there exists  $C > 0$  such that  $e^{-\eta t} W'(t-x) \leq C$  for  $t$  large enough.  $\square$

### 3.2 Convergence results

In this paragraph, we suppose that the process  $\Xi$  is supercritical ( $m > 1$ ) and are interested in improvements of the convergence result (14). We recall that  $\mathbb{P}^*$  is the law of the process  $\Xi$  conditional on its survival event.

The following proposition yields the asymptotic behavior, as  $t \rightarrow \infty$  and under  $\mathbb{P}^*$ , of the frequency spectrum  $(M_t^{i,a}, i \geq 1)$ . We also obtain the convergence of the relative abundances  $M_t^{i,a}/M_t$  of families of sizes  $i$  and ages less than  $a$ .

**Proposition 3.3.** *Under  $\mathbb{P}^*$ , with probability 1,*

$$e^{-\eta t} M_t \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J \mathcal{E}, \quad (15)$$

$$e^{-\eta t} M_t^{i,a} \xrightarrow[t \rightarrow \infty]{} \frac{p}{1-p} J^{i,a} \mathcal{E}, \quad (16)$$

and

$$\frac{M_t^{i,a}}{M_t} \xrightarrow[t \rightarrow \infty]{} \frac{J^{i,a}}{J} \quad (17)$$

where

$$J := \int_0^\infty e^{-\eta u} \frac{W'_c(u)}{W_c(u)} du$$

and where  $\mathcal{E}$  is the exponential random variable with parameter  $\psi'(\eta)$  defined by (8).

Notice that (16) is consistent with (14) since  $\mathbb{P}(\text{Ext}^c) = \eta/b$  and  $\mathbb{E}[\mathcal{E}] = 1/\psi'(\eta)$ . Moreover, we point out that the relative abundances  $M_t^{i,a}/M_t$  converges to a deterministic limit.

Similar convergence results were obtained by Z. Taïb in [36] who considered a similar model with more general birth point processes and mutations. However, his convergence theorems give non-explicit limits. In our particular case, we can find the distribution of the limits because the one-dimensional marginals and the asymptotic behavior of  $\Xi$  are known (see above-mentioned Proposition 2.1 and formulas (5) and (6)).

*Proof of Proposition 3.3.* To prove the three convergence results, we use general properties about CMJ-processes counted by random characteristics developed by Jagers and Nerman (see [21, 22, 23]). We mainly follow ideas of Taïb and all the properties we use are recorded in the appendix of [36]. We first check that the general assumptions (C.1-4) of this appendix hold in our case: recalling that  $\mu$  is the measure defined by (1), we must have

- (a)  $\mu(\mathbb{R}^+) > 1$  (supercritical case).
- (b) There is a number  $\lambda_0 > 0$  such that  $\int_0^\infty e^{-\lambda_0 u} \mu(du) = 1$ .
- (c)  $\beta := \int_0^\infty u e^{-\lambda_0 u} \mu(du) < \infty$ .
- (d)  $\mu$  is not supported by any lattice  $k\mathbb{Z}$  for  $k > 0$ .

First, by the definition of  $\mu$ , (a) is satisfied since we have supposed that  $m > 1$  and point (d) is straightforward. Moreover, the largest root  $\eta$  of  $\psi$  satisfies (b) seeing that

$$\int_0^\infty e^{-\eta u} \mu(du) = \int_0^\infty e^{-\eta u} \Lambda((u, +\infty]) du = \int_{(0, \infty]} \Lambda(dr) \frac{1 - e^{-\eta r}}{\eta} = \frac{\eta - \psi(\eta)}{\eta} = 1. \quad (18)$$

Finally, let us check that (c) holds. Since  $\Lambda$  is a finite measure with mass  $b$ , we have

$$\beta = \int_0^\infty u e^{-\eta u} \Lambda((u, +\infty]) du \leq b \int_0^\infty u e^{-\eta u} du < \infty.$$

We are now able to prove (16). We follow the proof of Theorem 3.3 in [36] using a random characteristic. We denote by  $\tau_1 < \tau_2 < \dots$  the successive birth times of the children of the ancestor, by  $\rho_1, \rho_2, \dots$  the independent sequence of i.i.d. Bernoulli random variables with parameter  $1 - p$  saying if a child is a clone or a mutant and by  $\Xi_c^1, \Xi_c^2, \dots$  the i.i.d. clonal population processes of the mutant children. Then, for  $u \geq a$ , we set the characteristic

$$\chi(u) := \sum_{j \geq 1} (1 - \rho_j) \mathbf{1}_{\{u-a \leq \tau_j < u\}} \mathbf{1}_{\{\Xi_c^j(u-\tau_j)=i\}},$$

that is,  $\chi(u)$  is the number of mutant children of the ancestor born between  $u - a$  and  $u$  and whose alleles are carried by  $i$  individuals at  $u$ . For  $u \leq a$ , we set  $\chi(u) = 0$ .

In the same way, for any individual  $k$  and  $u > a$ , denote by  $\chi_k(u)$  the number of mutant children of  $k$ , born between  $u - a$  and  $u$  units of time after the birth of  $k$  and whose alleles are carried by  $i$  individuals  $u$  units of time after the birth of  $k$ . Hence,  $(M_t^{i,a}, t \geq 0)$  can be counted by the characteristic  $\chi$ , meaning that for  $t \geq 0$ ,

$$M_t^{i,a} = \sum_{k=1}^{y_t} \chi_k(t - \alpha_k) + \mathbf{1}_{\{t \leq a\}} \mathbf{1}_{\{\Xi_c(t)=i\}} \quad (19)$$

where  $\alpha_k$  denotes the birth time of  $k$  and  $y_t$  is the number of individuals born before  $t$ . The second term in the r.h.s. of (19) pertains to the ancestor which is not counted by the characteristic. However, this term vanishes when  $t > a$  and so disappears at the limit  $t \rightarrow \infty$ .

According to Theorem 5 in the appendix of [36], if there exists  $\eta' < \eta$  such that

$$\int_0^\infty e^{-\eta' t} \mu(dt) < \infty \quad (20)$$

and such that

$$\mathbb{E} \left[ \sup_{u \geq 0} e^{-\eta' u} \chi(u) \right] < \infty, \quad (21)$$

then, under  $\mathbb{P}^*$ ,

$$\lim_{t \rightarrow \infty} \frac{M_t^{i,a}}{y_t} = \int_0^\infty \eta e^{-\eta u} \mathbb{E} [\chi(u)] du \quad \text{a.s.}$$

However, conditioning by the point process  $\xi_m$  which is independent from the  $\Xi_c^j$ 's, we have

$$\begin{aligned} \int_0^\infty \eta e^{-\eta u} \mathbb{E} [\chi(u)] du &= \int_0^\infty \eta e^{-\eta u} du \mathbb{E} \left[ \sum_{j \geq 1} (1 - \rho_j) \mathbf{1}_{\{u-a \leq \tau_j < u\}} \mathbf{1}_{\{\Xi_c^j(u-\tau_j)=i\}} \right] \\ &= \int_0^\infty \eta e^{-\eta u} du \mathbb{E} \left[ \int_{u-a}^u \mathbb{P}(\Xi_c(u-t) = i) \xi_m(dt) \right] \\ &= \int_0^\infty \eta e^{-\eta u} du \int_{u-a}^u \mathbb{P}(\Xi_c(u-t) = i) \mu_m(dt) \\ &= \int_0^\infty \mu_m(dt) e^{-\eta t} \int_0^a \eta e^{-\eta u} \mathbb{P}(\Xi_c(u) = i) du \end{aligned}$$

using Fubini-Tonelli's theorem and a change of variables. Since  $\mu_m = p\mu$  and  $\eta$  satisfies (18),  $\int_0^\infty e^{-\eta t} \mu_m(dt) = p$ . Then, thanks to (6),

$$\lim_{t \rightarrow \infty} \frac{M_t^{i,a}}{y_t} = \frac{p\eta}{b(1-p)} \int_0^a e^{-\eta u} \left( 1 - \frac{1}{W_c(u)} \right)^{i-1} \frac{W_c'(u)}{W_c^2(u)} du \quad \text{a.s.} \quad (22)$$

We still have to find  $\eta' < \eta$  such that conditions (20) and (21) are satisfied. First, as in (18), for  $\eta' \in (0, \eta)$ ,  $\int_0^\infty e^{-\eta' u} \mu(du) = 1 - \psi(\eta')/\eta' < \infty$ . Second, the characteristic  $\chi$  is stochastically dominated by a Poisson process with parameter  $b$ , say  $(N_t, t \geq 0)$ . Then, to prove (21), it is sufficient to show that  $\mathbb{E} \left[ \sup_{t \geq 0} e^{-\eta' t} N_t \right]$  is finite. At the end of the proof of Proposition 5.1 in [6, p.1028], the authors show that for  $\kappa$  large enough, the process  $A := \left( \left( e^{-\eta' t} N_t + \kappa \right)^2, t \geq 0 \right)$  is a supermartingale, which implies that for  $y \geq 0$ ,

$$\mathbb{P} \left( \sup_{t \geq 0} e^{-\eta' t} N_t \geq y \right) = \mathbb{P} \left( \sup_{t \geq 0} A_t \geq (y + \kappa)^2 \right) \leq \frac{\mathbb{E}[A_0]}{(y + \kappa)^2} = \frac{\kappa}{(y + \kappa)^2},$$

where the inequality is due to the maximal inequality [31, p. 58]. Then,

$$\mathbb{E} \left[ \sup_{t \geq 0} e^{-\eta' t} N_t \right] = \int_0^\infty dy \mathbb{P} \left( \sup_{t \geq 0} e^{-\eta' t} N_t \geq y \right) < \infty.$$

Furthermore, using again [36, Thm 5] with the characteristic  $\chi'(u) = \mathbf{1}_{\{0 \leq u \leq \lambda\}}$ , on  $\mathbb{P}^*$

$$\frac{\Xi_t}{y_t} = \frac{\sum_{k=1}^{y_t} \chi_k'(t - \alpha_k)}{y_t} \xrightarrow{t \rightarrow \infty} \int_0^\infty \eta e^{-\eta x} \mathbb{E} [\chi'(x)] dx = \int_0^\infty \eta e^{-\eta x} dx \int_{(x, \infty)} \frac{\Lambda(dr)}{b} = \frac{\eta}{b} \quad (23)$$

because  $\psi(\eta) = 0$ . Notice that in that case, condition (21) is easily satisfied because  $\chi'(u) \leq 1$  for  $u \geq 0$ . Finally, using together (8), (22) and (23) we get the convergence result (16).

We now prove (15). According to [36, Thm 3.3], on  $\mathbb{P}^*$ ,

$$\frac{M_t}{y_t} \xrightarrow{t \rightarrow \infty} \int_0^\infty e^{-\eta x} \mu_m(dx) \left( 1 - \eta \int_0^\infty e^{-\eta u} \mathbb{P}(\Xi_c(u) = 0) \right) du = p\eta \int_0^\infty e^{-\eta u} \frac{W_c'(u)}{b(1-p)W_c(u)} du.$$

By using the last display, (8) and (23), on  $\mathbb{P}^*$ , we have

$$e^{-\eta t} M_t \xrightarrow{t \rightarrow \infty} \mathcal{E} \frac{p}{1-p} \int_0^\infty e^{-\eta u} \frac{W_c'(u)}{W_c(u)} du.$$

Finally, (17) is straightforward from (15) and (16).  $\square$

The following result deals with the asymptotic behavior of  $M_t^{i,t}$ , the number of alleles carried by  $i$  individuals at time  $t$ .

**Proposition 3.4.** *Under  $\mathbb{P}^*$ , with probability 1 as  $t \rightarrow \infty$ ,*

$$\frac{M_t^{i,t}}{M_t} \rightarrow \frac{1}{i} \frac{\int_0^\infty e^{-\eta u} \left( 1 - \frac{1}{W_c(u)} \right)^i du}{\int_0^\infty e^{-\eta u} \ln(W_c(u)) du}.$$

Notice that the limit can be considered as a weighted Fisher log-series distribution, which is commonly used to describe species abundances [15]. More precisely, this limit can be written

$$\frac{\frac{1}{i} \int_0^1 (1-\omega)^i d\nu(\omega)}{\int_0^1 \log(1/\omega) d\nu(\omega)}$$

where  $d\nu(\omega) = e^{-\eta u(\omega)} du(\omega)$  and  $u(\omega) = W^{-1}(1/\omega)$ .

*Proof.* Since  $M_t^{i,t}$  can be counted by the random characteristic

$$\bar{\chi}(u) := \sum_{j \geq 1} (1 - \rho_j) \mathbf{1}_{\{\tau_j < u\}} \mathbf{1}_{\{\Xi_c^j(u - \tau_j) = i\}},$$

similar computations as in the proof of Proposition 3.3 lead to the convergence of  $M_t^{i,t}/M_t$  to  $J^{i,\infty}/J$  as  $t \rightarrow \infty$ . We then get the result by integrating by parts  $J$  and  $J^{i,\infty}$  by noticing that  $W_c(0) = 1$  according to [33, Lemma 4.1].  $\square$

**Remark 3.5.** *In most cases, it is not possible to simplify the expressions of  $J$  and  $J^{i,a}$  appearing in the limits in Propositions 3.3 and 3.4 since the scale functions  $W$  and  $W_c$  are generally unknown.*

*However, when the life lengths are exponentially distributed with parameter  $d < b$  (in that case,  $\Xi$  is simply a Markovian linear birth and death process with birth rate  $b$  and death rate  $d$ ), we know from [26, p393] that*

$$W_c(x) = \begin{cases} \frac{d - (1-p)be^{((1-p)b-d)x}}{d - (1-p)b} & \text{if } d \neq (1-p)b \\ 1 + dx & \text{if } d = (1-p)b \end{cases}. \quad (24)$$

*In that case, as it was recently done in [37], thanks to simple changes of variables, the integrals  $J$  and  $J^{i,\infty}$  can be expressed as Gauss hypergeometric functions or as confluent hypergeometric*

functions [1] whose parameters only depend on  $b$ ,  $d$ ,  $p$  and  $i$ . Then,  $J$  and  $J^{i,\infty}$  can be numerically computed.

Finally, when  $\Xi$  is a pure birth process, that is, when  $\Lambda(\cdot)/b = \delta_\infty$ , there are simple expressions of these integrals. Indeed, in that case, we have  $\psi(\lambda) = (\lambda - b)\mathbf{1}_{\{\lambda > 0\}}$ ,  $\eta = b$ ,  $W_c(x) = e^{(1-p)bx}$  from (24) and obviously,  $\mathbb{P}(\text{Ext}) = 0$  and  $\mathbb{P}^* = \mathbb{P}$ . Moreover, easy computations lead to simple expressions of  $J$  and  $J^{i,a}$  and we obtain convergence results such as

$$\frac{M_t^{i,t}}{M_t} \xrightarrow{t \rightarrow \infty} \frac{1}{i(1-p)} \sum_{k=0}^i (-1)^k \frac{1}{1+(1-p)k} \quad \text{a.s.}$$

## 4 Number of undergone mutations

Let us say that the allele of the progenitor is of *type 0*. Then, recursively, for  $k \geq 1$ , we say that an allele is of *type k* if it is carried by a mutant child of an individual carrying an allele of type  $k-1$ . Equivalently, they are alleles that have been affected by  $k$  more mutations than the ancestral allele. For  $t \geq 0$  and  $i \geq 0$ , we denote by  $L_i(t)$  the number of alleles of type  $i$  at time  $t$  and by  $K_i(t)$  the number of individuals that carry such alleles at time  $t$ .

In this section, we obtain similar results as those of Section 3, that is, we compute the expected values  $\mathbb{E}[K_i(t)]$  and  $\mathbb{E}[L_i(t)]$  for fixed  $i$  and  $t$ . We then study the asymptotic behavior of  $K_i(t)$  as  $t \rightarrow \infty$  when the clonal process is supercritical.

The main argument we use to get these results is to consider  $(K_i(t), i \geq 0, t \geq 0)$  as a multitype CMJ-process. More precisely, regardless of its type, every individual has a lifespan distributed as  $\Lambda(\cdot)/b$  and an individual with type  $i$  gives birth to individuals of the  $i$ -th type at rate  $b(1-p)$  and to individuals of type  $i+1$  at rate  $bp$ . Moreover, the process  $(K_i, i \geq 1)$  belongs to the class of *reducible* multitype processes since an individual of type  $i$  cannot give birth to an individual of type  $j$  with  $j < i$ . Then, if at some time, there is no more individuals of type  $i$ , no such individual will reappear in the population.

Even though the reducible branching processes are less studied than the irreducible processes (which enable the use of Perron-Frobenius theory of positive matrices), they are useful in the model we consider. Indeed, if we are interested in the study of the process  $K^i$  for a given  $i$ , since the types cannot decrease, we can transform a problem with an infinite and countable number of types to a problem with a finite number of possible types by studying the multitype process  $(K_j, j = 1, \dots, i)$ .

The reader interested in multitype CMJ-processes can refer to the third and fourth chapters of Mode's book [27].

### 4.1 Expected number of mutations

For  $f, g$  two continuous functions on  $[0, +\infty)$ , we denote by  $f \star g$  the standard convolution product

$$f \star g(t) = \int_0^t f(t-x)g(x)dx, \quad t \geq 0.$$

We use the following notation for the consecutive convolutions of a function. For a continuous function  $f$ , let  $f^{\star(1)} := f$  and for  $i \geq 2$ ,  $f^{\star(i)} := f \star f^{\star(i-1)}$ . In the following result, we give simple expressions of the mean values of  $K_i(t)$  and  $L_i(t)$  that are entirely determined by the functions  $W_c$  and  $W'_c$ .

**Proposition 4.1.** *We fix  $t > 0$ . We then have*

$$\mathbb{E}[K_i(t)] = \frac{1}{b(1-p)} \left( \frac{p}{1-p} \right)^i (W'_c)^{\star(i+1)}(t), \quad i \geq 0 \quad (25)$$

$$\mathbb{E}[L_0(t)] = \frac{1}{b(1-p)} \frac{W'_c(t)}{W_c(t)} \quad (26)$$

$$\mathbb{E}[L_i(t)] = \frac{1}{b(1-p)} \left( \frac{p}{1-p} \right)^i (W'_c)^{\star(i)} \star \frac{W'_c}{W_c}(t), \quad i \geq 1. \quad (27)$$

In [27], C. Mode considers general multitype branching processes. In paragraph 4.4, he gives an equation satisfied by the Laplace transform of  $\mathbb{E}[K_i(\cdot)]$ . In most cases, this equation cannot be solved but it enables to obtain the asymptotic behavior of  $\mathbb{E}[K_i(t)]$  as  $t \rightarrow \infty$ . In our particular case, one can solve the equation of Mode to obtain the Laplace transform of  $\mathbb{E}[K_i(\cdot)]$  and to compute  $\mathbb{E}[K_i(t)]$  thanks to an inverse Laplace transform. Nevertheless, in the following proof, we prefer to use a more direct approach which avoids long computations.

*Proof of Proposition 4.1.* We prove (25) by induction on  $i \geq 0$ . For  $i = 0$ , according to (7),  $\mathbb{E}[K_0(t)] = \mathbb{E}[\Xi_c(t)] = \frac{W'_c(t)}{b(1-p)}$ . We now suppose that the result is true at rank  $i - 1$ . Denote by  $K_i^{\text{da}}(t)$  the number of individuals of type  $i$  that are alive at time  $t$  and whose alleles appeared in the population between times  $t - a$  and  $t - a + da$ . Then, in the same manner as in the proof of Proposition 3.1, using the branching property, conditional on  $K_{i-1}(t - a) = n$ ,  $K_i^{\text{da}}(t)$  is the sum of  $n$  independent random variables. These random variables are distributed as the number at time  $a$ , of all the clonal descendants of mutant individuals born in a time interval  $[0, da]$ . Then,

$$\mathbb{E}[K_i^{\text{da}}(t)] = \mathbb{E}[K_{i-1}(t - a)] \mathbb{E}[\Xi_c(a)] b p da.$$

Finally, using the induction hypothesis and integrating over  $a \in (0, t)$ , we have

$$\mathbb{E}[K_i^a(t)] = \frac{1}{b(1-p)} \left( \frac{p}{1-p} \right)^{i-1} \int_0^t (W'_c)^{\star(i)}(t-a) \frac{W'_c(a)}{b(1-p)} b p da,$$

which is what we wanted to prove.

Since  $L_0(t) = \mathbf{1}_{\{\Xi_c(t) > 0\}}$ , (26) is straightforward from (5). We prove (27) with similar techniques as we proved (25). The mean number of alleles of type  $i$  with age  $a$  at time  $t$  is

$$\mathbb{E}[L_i^{\text{da}}(t)] = \mathbb{E}[K_{i-1}(t - a)] b p \mathbb{P}(\xi_c(a) > 0) da.$$

We then get the result thanks to (5), (25) and by integrating the last display on  $[0, t]$ .  $\square$

From Proposition 4.1, we obtain the asymptotic behaviors of  $\mathbb{E}[K_i(t)]$  and  $\mathbb{E}[L_i(t)]$  as  $t \rightarrow \infty$ . Different regimes appear depending on the class of criticality of the clonal process  $\Xi_c$ .

**Corollary 4.2.** *We suppose that one of the following hypotheses holds*

- $(1-p)m > 1$ ,
- $(1-p)m = 1$ ,  $\sigma^2 < \infty$  and  $\lim_{t \rightarrow \infty} t^2 \int_t^\infty (x-t) \Lambda(dx) = 0$ ,

- $(1-p)m < 1$  and  $\psi_c$  has a negative root  $\tilde{\eta}_c$  satisfying  $\int_0^\infty r e^{-\tilde{\eta}_c r} \Lambda(dr) < \infty$ .

Then, for  $i \geq 0$ , we have

$$\mathbb{E}[K_i(t)] \underset{t \rightarrow \infty}{\sim} \frac{1}{b(1-p)i!} \left(\frac{p}{1-p}\right)^i C_p^{i+1} \cdot t^i e^{\eta_p t}$$

where

$$\eta_p := \begin{cases} \eta_c & \text{if } (1-p)m > 1 \\ 0 & \text{if } (1-p)m = 1 \\ \tilde{\eta}_c & \text{if } (1-p)m < 1 \end{cases} \quad \text{and } C_p := \begin{cases} \frac{\eta_c}{\psi'_c(\eta_c)} & \text{if } (1-p)m > 1 \\ \frac{1}{(1-p)\sigma^2} & \text{if } (1-p)m = 1 \\ \frac{\eta_c}{\psi'_c(\tilde{\eta}_c)} & \text{if } (1-p)m < 1 \end{cases}.$$

*Proof.* This result is a consequence of the following lemma thanks to Proposition 4.1 and Lemma 2.2 which gives the asymptotic behaviors of  $W'_c(t)$  as  $t \rightarrow \infty$  in all cases.  $\square$

**Lemma 4.3.** *Let  $f$  be a non-negative continuous function on  $[0, +\infty)$  such that for some  $a \in \mathbb{R}$  and  $l \geq 0$ ,*

$$e^{at} f(t) \xrightarrow[t \rightarrow \infty]{} l. \quad (28)$$

Then, for  $i \geq 1$

$$\frac{e^{at}}{t^{i-1}} f^{*(i)}(t) \xrightarrow[t \rightarrow \infty]{} \frac{l^i}{(i-1)!}.$$

*Proof.* We use induction on  $i \geq 1$  to prove the two properties

$$S_i := \sup_{t \in (0, +\infty)} \frac{e^{at}}{t^{i-1}} f^{*(i)}(t) < \infty \quad \text{and} \quad \frac{e^{at}}{t^{i-1}} f^{*(i)}(t) \xrightarrow[t \rightarrow \infty]{} \frac{l^i}{(i-1)!}. \quad (29)$$

This is obvious when  $i = 1$  since  $f$  is continuous and since we assume (28). We now suppose that (29) holds for an integer  $i \geq 1$ . By a change of variables, we have

$$f^{*(i+1)}(t) = \int_0^t f(t-x) f^{*(i)}(x) dx = t \int_0^1 f(t(1-x)) f^{*(i)}(tx) dx.$$

Hence,

$$t^{-i} e^{at} f^{*(i+1)}(t) = \int_0^1 \left( e^{at(1-x)} f(t(1-x)) \right) \left( \frac{e^{atx}}{(tx)^{i-1}} f^{*(i)}(tx) \right) x^{i-1} dx. \quad (30)$$

Thus, we first have  $S_{i+1} \leq S_1 S_i \int_0^1 x^{i-1} dx < \infty$ . Moreover, as  $t \rightarrow \infty$ , the r.h.s. of (30) converges to  $l \frac{l^i}{(i-1)!} \int_0^1 x^{i-1} dx = \frac{l^{i+1}}{i!}$  by using (28), the recurrence hypothesis and the dominated convergence theorem (the integrand is upper-bounded by  $S_1 S_i x^{i-1}$  which is integrable on  $(0, 1)$ ).  $\square$

We next determine the asymptotic behavior of  $\mathbb{E}[L_i(t)]$  as  $t \rightarrow \infty$ .

**Corollary 4.4.** *Let  $i$  be a positive integer.*



(i) If  $(1-p)m > 1$ , we have

$$\mathbb{E}[L_i(t)] \underset{t \rightarrow \infty}{\sim} \frac{J_c}{b(1-p)(i-1)!} \left(\frac{p}{1-p}\right)^i \left(\frac{\eta_c}{\psi'_c(\eta_c)}\right)^i \cdot t^{i-1} e^{\eta_c t}$$

where  $J_c := \int_0^\infty e^{-\eta_c u} \frac{W'_c(u)}{W_c(u)} du$ .

(ii) If  $(1-p)m = 1$ , if  $\sigma^2 < \infty$  and if (9) holds,

$$\mathbb{E}[L_i(t)] \underset{t \rightarrow \infty}{\sim} C_i t^i \ln t$$

where  $C_i = \frac{1}{b(1-p)(i-1)!} \left(\frac{2p}{\sigma^2(1-p)^2}\right)^i$ .

(iii) If  $(1-p)m < 1$  and if there is  $\tilde{\eta}_c$  satisfying (10),

$$\mathbb{E}[L_i(t)] \underset{t \rightarrow \infty}{\sim} \frac{1 - (1-p)m}{b(1-p)i!} \left(\frac{p}{1-p}\right)^i \left(\frac{\tilde{\eta}_c}{\psi'(\tilde{\eta}_c)}\right)^{i+1} \cdot t^i e^{\tilde{\eta}_c t}.$$

*Proof.* We begin by proving the subcritical case. According to Lemma 2.2(iii),  $W'_c(t)/W_c(t)$  behaves as  $(1 - (1-p)m) \frac{\tilde{\eta}_c}{\psi'(\tilde{\eta}_c)} e^{\tilde{\eta}_c t}$  as  $t \rightarrow \infty$ . Then point (iii) stems from a slight modification of Lemma 4.3.

Concerning the critical case, according to Lemmas 2.2 and 4.3, for any  $\varepsilon > 0$ , there is  $A > 0$  such that, if  $t - x \geq A$ , then

$$\left| \frac{(W'_c)^{\star(i)}(t-x)}{t^i} - \underbrace{\frac{1}{(i-1)!} \left(\frac{2}{(1-p)\sigma^2}\right)^i}_{:=A_i} \right| \leq \varepsilon.$$

Then, for  $\varepsilon > 0$  and  $t \geq A$ , we have

$$t^{-i} (W'_c)^{\star(i)} \star \frac{W'_c}{W_c}(t) = I_1(t) + I_2(t)$$

where

$$I_1(t) := \int_0^{t-A} t^{-i} (W'_c)^{\star(i)}(t-x) \frac{W'_c(x)}{W_c(x)} dx \in \left[ (A_i - \varepsilon) \int_0^{t-A} \frac{W'_c(x)}{W_c(x)} dx, (A_i + \varepsilon) \int_0^{t-A} \frac{W'_c(x)}{W_c(x)} dx \right]$$

and

$$I_2(t) := \int_0^A t^{-i} (W'_c)^{\star(i)}(x) \frac{W'_c(t-x)}{W_c(t-x)} dx \leq t^{-i} b(1-p) \int_0^A (W'_c)^{\star(i)}(x) dx$$

where we have used that  $\frac{W'_c(x)}{W_c(x)} = b(1-p)\mathbb{P}(\Xi_c(t) > 0)$ . Hence, as  $t \rightarrow \infty$ ,  $I_2(t)$  vanishes and  $I_1(t)$  is equivalent to  $A_i \ln(t)$  since we know from Lemma 2.2(ii) that  $W'_c(x)/W_c(x) \sim 1/x$  as  $x \rightarrow \infty$ . Thus,

$$\mathbb{E}[L^i(t)] = \frac{1}{b(1-p)} \left(\frac{p}{1-p}\right)^i (W'_c)^{\star(i)} \star \frac{W'_c}{W_c}(t) \underset{t \rightarrow \infty}{\sim} \frac{A_i}{b(1-p)} \left(\frac{p}{1-p}\right)^i t^i \ln t.$$

Finally, we get the asymptotic behavior in the supercritical case by a direct application of the following Lemma 4.5. Indeed, according to the proof of Lemma 4.4,  $(W'_c)^{\star(i)}(t)$  grows as  $t^i e^{\eta_c t}$  as  $t \rightarrow \infty$  and from Lemma 2.2 (i), we have  $\frac{W'_c(t)}{W_c(t)} \rightarrow \eta_c$  as  $t \rightarrow \infty$ .  $\square$

**Lemma 4.5.** *Let  $f, g$  be two continuous and positive functions such that  $f(t)$  converges as  $t \rightarrow \infty$  and such that for some  $a > 0$ ,*

$$\lim_{t \rightarrow \infty} \frac{e^{-at}}{t^i} g(t) = l'. \quad (31)$$

Then, as  $t \rightarrow \infty$ ,

$$f \star g(t) \sim l' e^{at} t^i \int_0^\infty e^{-ax} f(x) dx.$$

*Proof.* We have

$$\frac{e^{-at}}{t^i} f \star g(t) = \int_0^\infty e^{-ax} f(x) h(t, x) dx$$

where  $h(t, x) := \frac{e^{-a(t-x)} g(t-x)}{t^i} \mathbf{1}_{\{0 \leq x \leq t\}}$ . Moreover,  $h(t, x) \xrightarrow[t \rightarrow \infty]{} l' \mathbf{1}_{\{x \geq 0\}}$  and according to (31)

$$h(t, x) \leq \frac{e^{-a(t-x)} g(t-x)}{(t-x)^i} \mathbf{1}_{\{x \leq t\}} \leq M$$

where  $M$  is a constant that does not depend on  $t$  and  $x$ . Finally, since  $f(t)$  converges as  $t \rightarrow \infty$ ,  $\int_0^\infty f(x) e^{-ax} dx$  is finite and the result stems from the dominated convergence theorem.  $\square$

## 4.2 Asymptotic behavior of $K_i(t)$ when $\Xi_c$ is supercritical

In Corollary 4.2, we proved that  $\mathbb{E}[K_i(t)]$  grows as  $t^i e^{\eta_c t}$  as  $t \rightarrow \infty$  when the clonal process is supercritical. We now improve that convergence result by studying the asymptotic behavior of  $K_i(t)$ . Notice that in the subcritical and critical cases, the process  $K_i$  goes extinct a.s.

**Theorem 4.6.** *We suppose that  $(1-p)m > 1$  and that  $\sigma^2 := \int_0^\infty \Lambda(dz) z^2 < \infty$ . Then, for  $i \geq 0$ , almost surely and in quadratic mean, we have*

$$\frac{e^{-\eta_c t}}{t^i} K_i(t) \xrightarrow[t \rightarrow \infty]{} \kappa_i$$

where  $\mathbb{P}(\kappa_i = 0) = 1 - \frac{\eta_c}{b(1-p)}$  and conditional on being non zero,  $\kappa_i$  is exponential with mean

$$\frac{1}{i!} \left( \frac{p}{1-p} \right)^i \frac{\eta_c^i}{\psi'_c(\eta_c)^{i+1}}.$$

*Proof.* The case  $i = 0$  is straightforward since  $K_0 = \Xi_c$  and the result holds according to Proposition 2.1. Then, from now on, we suppose that  $i \geq 1$ .

The proof is divided in three steps. In the first one, we prove that the a.s. and  $L^2$  convergences hold. We then identify the law of the limit: we find conditions characterizing its Laplace transform (step 2) and then exhibit the random variable which satisfies these conditions (step 3).

### Step 1: Proof of the convergence

To prove the a.s. and quadratic mean convergences toward  $\kappa_i$  as  $t \rightarrow \infty$ , we use Theorem 4.3 (iv) in [27, p.173]. Actually, in this book, C. Mode only proved that  $t^{-1}e^{-\eta_c t}K_1(t)$  converges as  $t \rightarrow \infty$  but the same result holds about  $K_i$  for  $i \geq 2$  by using similar techniques.

Let us check the hypotheses of Theorem 4.3 of [27]. First, we have to show that  $N$ , the number of children of the ancestor, is square integrable. Conditional on having a life length  $z$ ,  $N$  is a Poisson random variable with parameter  $bz$ . Then,

$$\mathbb{E}[N^2] = b^{-1} \int_0^\infty \Lambda(dz)((bz)^2 + bz) = b\sigma^2 + m < \infty.$$

Second, let us consider an individual of type  $i$  whose life length is infinite. Denote by  $\mu_i(t)$  her mean number of children that are born before she reaches age  $t$ . We need to prove that  $\max_i \mu_i(t)$  is zero when  $t = 0$  and finite when  $t > 0$ . In our case, it is obvious since for any  $i \geq 0$ ,  $\mu_i(t) = bt$ .

Finally, if  $\zeta$  is the life length of an individual, we have to check that  $\int_0^\infty dt \mathbb{P}(\zeta \geq t)^p$  is finite for  $p = 1$  and also for some  $p > 1$ . We have

$$\int_0^\infty dt \mathbb{P}(\zeta \geq t)^p \leq \int_0^\infty dt \mathbb{P}(\zeta \geq t) = \mathbb{E}[\zeta] = \frac{m}{b} < \infty.$$

### Step 2: Characterization of the Laplace transform of $\kappa_i$

For  $i \geq 1$ , denote by  $\varphi_i$  the Laplace transform of  $\kappa_i$

$$\varphi_i(a) := \mathbb{E}[e^{-a\kappa_i}], \quad a > 0.$$

$\mathbb{E}[\kappa_i] = \frac{1}{b(1-p)} \frac{1}{i!} \left(\frac{p}{1-p}\right)^i \left(\frac{\eta_c}{\psi'_c(\eta_c)}\right)^{i+1}$ . In particular,  $\kappa_i$  is almost surely finite and

$$\varphi_i(0) = 1 \text{ and } \varphi_i \text{ is differentiable at } 0 \text{ with } \varphi'_i(0) = \frac{1}{b(1-p)} \frac{1}{i!} \left(\frac{p}{1-p}\right)^i \left(\frac{\eta_c}{\psi'_c(\eta_c)}\right)^{i+1}. \quad (32)$$

We now want to prove that for  $i \geq 1$  and  $a > 0$ , we have

$$\varphi_i(a) = \int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ b(1-p) \left( \int_0^z \varphi_i(ae^{-\eta_c u}) du - z \right) \right\}. \quad (33)$$

To show that the latter holds, we look at the children of the ancestor. More precisely, conditional on having a lifespan  $z$ , before time  $t$ , the ancestor has a number of mutant (resp. clonal) children distributed as a Poisson random variable with parameter  $bp(z \wedge t)$  (resp.  $b(1-p)(z \wedge t)$ ). Moreover, these two random variables are independent.

For  $z > 0$ ,  $t \geq 0$  and  $k, l \in \mathbb{N}$ , by the branching property and classical properties about Poisson processes, conditional on the event “during its lifespan  $z$  and before age  $t$ , the ancestor has  $k$  mutant and  $l$  clonal children”, we have

$$K_i(t) \stackrel{(d)}{=} \sum_{q=1}^k \tilde{K}_i^q(t - U_q) + \sum_{r=1}^l \hat{K}_i^r(t - V_r)$$

where

- the processes  $\tilde{K}_i^1, \dots, \tilde{K}_i^k$  are i.i.d. and distributed as  $K_{i-1}$ ,
- the processes  $\hat{K}_i^1, \dots, \hat{K}_i^k$  are i.i.d. and distributed as  $K_i$ ,
- the random variables  $U_1, \dots, U_p, V_1, \dots, V_q$  are i.i.d. and uniform in  $[0, z \wedge t]$ .

Moreover, all these quantities are mutually independent. Hence, for  $s \in [0, 1]$ , we have

$$\begin{aligned} \mathbb{E} \left[ s^{K_i(t)} \right] &= \int_0^\infty \frac{\Lambda(dz)}{b} \sum_{k,l \geq 0} e^{-bz \wedge t} \frac{(bp(z \wedge t))^k}{k!} \frac{(b(1-p)(z \wedge t))^l}{l!} \mathbb{E} \left[ s^{K_{i-1}(t-U_1)} \right]^k \mathbb{E} \left[ s^{K_i(t-V_1)} \right]^l \\ &= \int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ -bz \wedge t + pb(z \wedge t) \mathbb{E} \left[ s^{K_{i-1}(t-U_1)} \right] + (1-p)b(z \wedge t) \mathbb{E} \left[ s^{K_i(t-V_1)} \right] \right\}. \end{aligned}$$

Thus, if we denote by  $G_{i,t}$  the probability generating function of  $K_i(t)$ , we have for  $s \in [0, 1]$

$$G_{i,t}(s) = \int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ -b(z \wedge t) + pb \int_0^{z \wedge t} G_{i-1,t-u}(s) du + (1-p)b \int_0^{z \wedge t} G_{i,t-v}(s) dv \right\}. \quad (34)$$

According to the first step of the proof, as  $t \rightarrow \infty$ ,  $t^{-i}e^{-\eta ct}K_i(t)$  converges a.s. and then in distribution to  $\kappa_i$ . Then,

$$\begin{aligned} G_{i,t} \left( e^{-at^{-i}e^{-\eta ct}} \right) &= \mathbb{E} \left[ e^{-aK_i(t)t^{-i}e^{-\eta ct}} \right] \xrightarrow{t \rightarrow \infty} \varphi_i(a), \\ G_{i-1,t-u} \left( e^{-at^{-i}e^{-\eta ct}} \right) &\xrightarrow{t \rightarrow \infty} 1, \end{aligned} \quad (35)$$

and

$$G_{i,t-v} \left( e^{-at^{-i}e^{-\eta ct}} \right) \xrightarrow{t \rightarrow \infty} \varphi_i(ae^{-\eta cv}). \quad (36)$$

We set

$$C(t, z) := -b(z \wedge t) + pb \int_0^{z \wedge t} G_{i-1,t-u} \left( e^{-at^{-i}e^{-\eta ct}} \right) du + (1-p)b \int_0^{z \wedge t} G_{i,t-v} \left( e^{-at^{-i}e^{-\eta ct}} \right) dv.$$

Then, since for all  $t \geq 0$  and  $i \in \mathbb{N}$ ,  $G_{i,t}$  and  $G_{i-1,t}$  are less than 1, using the dominated convergence theorem and equations (35) and (36), we get

$$\lim_{t \rightarrow \infty} C(t, z) = -b(1-p)z + (1-p)b \int_0^z \varphi_i(ae^{-\eta cv}) dv$$

Finally, if we apply (34) with  $s = e^{-at^{-i}e^{-\eta ct}}$  and let  $t$  go to  $\infty$ , we obtain (33). Indeed, we can again use the dominated convergence theorem since  $C(t, z) \leq 0$  for  $t \geq 0$  and  $z > 0$  and since  $\Lambda$  is a finite measure.

In summary, we have proved that  $\varphi_i$  satisfies (32) and (33). In fact, according to the following Lemma 4.7 whose proof is postponed to the end of the section,  $\varphi_i$  is the only function that fulfills these two conditions, that is,  $\kappa_i$  is characterized by (32) and (33).

**Lemma 4.7.** *Suppose that  $\Lambda$  satisfies the hypotheses of Theorem 4.6 and let  $A \in (0, +\infty)$ . Then, there is a unique function  $f : [0, +\infty) \rightarrow [0, 1]$  satisfying the following properties*

- $f$  is continuous on  $[0, +\infty)$  and  $f(0) = 1$ .*
- $f$  is differentiable at 0 and  $f'(0) = A$ .*
- For any  $a > 0$ ,  $f(a) = \int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ b(1-p) \int_0^z (f(ae^{-\eta cu}) - 1) du \right\}$ .*

### Step 3: Identification of the solution

We denote  $P := \frac{\eta_c}{b(1-p)}$  and for  $i \geq 1$ , let  $E_i$  be a non-negative random variable such that  $\mathbb{P}(E_i = 0) = 1 - P$  and conditional on being positive,  $E_i$  is exponential with parameter  $\theta_i := \left( \frac{1}{i!} \left( \frac{p}{1-p} \right)^i \frac{\eta_c^i}{\psi_c'(\eta_c)^{i+1}} \right)^{-1}$ . To complete the proof of Theorem 4.6, it remains to check that the Laplace transform of  $E_i$ , that is  $\varphi_{E_i}(a) = 1 - P + P \frac{\theta_i}{a + \theta_i}$ , satisfies the conditions (32) and (33) of the Step 2.

Condition (32) is trivially satisfied since  $\mathbb{E}[E_i] = P/\theta_i$ . Furthermore, we have

$$\begin{aligned} \int_0^z \varphi_{E_i}(e^{-\eta_c u}) du &= z(1-P) + P \int_0^z \frac{\theta_i e^{\eta_c u}}{a + \theta_i e^{\eta_c u}} du \\ &= z(1-P) + \frac{P}{\eta_c} (\ln(a + \theta_i e^{\eta_c z}) - \ln(a + \theta_i)). \end{aligned}$$

Then,

$$\begin{aligned} \int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ b(1-p) \left( \int_0^z \varphi_{E_i}(ae^{-\eta_c u}) du - z \right) \right\} &= \int_0^\infty \frac{\Lambda(dz)}{b} e^{-\eta_c z} \frac{a + \theta_i e^{\eta_c z}}{a + \theta_i} \\ &= \frac{\theta_i}{a + \theta_i} + \frac{a}{a + \theta_i} \int_0^\infty \frac{\Lambda(dz)}{b} e^{-\eta_c z}. \end{aligned}$$

Since  $\eta_c$  is a root of the function  $\psi_c$  defined by (11), we have  $(1-p) \int_0^\infty \Lambda(dz) e^{-\eta_c z} = b(1-p) - \eta_c$ . Then,

$$\int_0^\infty \frac{\Lambda(dz)}{b} \exp \left\{ b(1-p) \left( \int_0^z \varphi_{E_i}(ae^{-\eta_c u}) du - z \right) \right\} = \frac{\theta_i}{a + \theta_i} + \frac{a}{a + \theta_i} (1-P) = \varphi_{E_i}(a).$$

Thus,  $\varphi_{E_i}$  satisfies condition (33), which ends the proof of Theorem 4.6.  $\square$

*Proof of Lemma 4.7.* The proof of this result is mostly inspired from that of Theorem 7.2 in [12]. Let  $\nu$  be the positive measure on  $[0, +\infty)$  defined by  $\nu(dx) := (1-p)e^{-\eta_c x} dx \int_x^{+\infty} \Lambda(dz)$ . This is a probability measure since by the Fubini theorem,

$$\int_0^\infty \nu(dx) = (1-p) \int_0^\infty \Lambda(dz) \int_0^z e^{-\eta_c x} dx = (1-p) \int_0^\infty \Lambda(dz) \frac{1 - e^{-\eta_c z}}{\eta_c} = \frac{\eta_c - \psi_c(\eta_c)}{\eta_c} = 1$$

where we have used that  $\eta_c$  is the largest root of  $\psi_c$ .

Let  $f$  and  $g$  be two functions satisfying the three conditions of Lemma 4.7. For  $a > 0$ , we define  $\Delta(a) := \frac{f(a) - g(a)}{a}$ . Since  $f$  and  $g$  fulfill the conditions (i) and (ii), we have

$$\frac{f(a) - g(a)}{a} = \frac{f(a) - 1}{a} + \frac{1 - g(a)}{a} \xrightarrow{a \rightarrow 0} f'(0) - g'(0) = 0.$$

Hence,  $\Delta(0^+) = 0$  and since  $f, g$  are continuous and bounded on  $[0, +\infty)$ , the function  $\Delta$  also is. Moreover, since  $f$  and  $g$  satisfy (iii) and take values in  $[0, 1]$ , using the inequality  $|e^{-x} - e^{-y}| \leq |x - y|$  for  $x, y \geq 0$ , we have

$$\begin{aligned} \Delta(a) &\leq (1-p) \int_0^\infty \Lambda(dz) \int_0^z \Delta(ae^{-\eta_c u}) e^{-\eta_c u} du \\ &= (1-p) \int_0^\infty du \Delta(ae^{-\eta_c u}) e^{-\eta_c u} \int_u^\infty \Lambda(dz) \\ &= \mathbb{E} [\Delta(ae^{-\eta_c Z})] \end{aligned}$$

where  $Z$  is a random variable distributed as the probability measure  $\nu$ . Thus, iterating this argument,

$$\Delta(a) \leq \mathbb{E} \left[ \Delta \left( a e^{-\eta c(Z_1 + \dots + Z_n)} \right) \right] \quad (37)$$

where the random variables  $Z_i$ 's are i.i.d. and distributed as  $\nu$ . Using the definition of  $\nu$ , these random variables are a.s. positive. Thus, the sum  $Z_1 + \dots + Z_n$  a.s. goes to  $+\infty$  as  $n \rightarrow \infty$  and the r.h.s of (37) goes to  $\Delta(0^+) = 0$  as  $n \rightarrow \infty$  by the dominated convergence theorem (recall that  $\Delta$  is bounded). Finally,  $\Delta(a) = 0$  for any positive  $a$  and the proof is complete.  $\square$

**Remark 4.8.** We can also view  $(L^i(t), i \geq 0, t \geq 0)$  as a multitype CMJ-process where the individuals of this branching process are the different alleles; roughly speaking, all the individuals sharing the same alleles are gathered into clusters. Moreover, an allele of type  $i$  has a lifespan distributed as the extinction time of  $\Xi_c$  and gives birth to alleles of type  $i + 1$  according to the jump times of  $(Y_{\int_0^t \Xi(s) ds}, t \geq 0)$  where  $Y$  is a homogeneous Poisson process with rate  $bp$ , independent of  $\Xi_c$ .

In the supercritical case, it seems not possible to study the asymptotic behavior of  $L_i(t)$  with the techniques we used to obtain that of  $K_i(t)$ . The problem is that in this multitype process, the number of “children” of an allele may be infinite since it survives forever with positive probability  $\frac{\eta c}{b(1-p)}$ . Then, some conditions of [27, Thm 4.3] are not fulfilled. As an example, the second moment property we checked in the first step of the proof of Theorem 4.6 is false here.

## A Asymptotic behaviors of $W$ and $W'$

### A.1 Proof of Lemma 2.2(i)

We suppose here that  $m > 1$  which implies that  $\eta > 0$ . To obtain the asymptotic behavior of  $W$ , we use a Tauberian theorem about Laplace transforms. Indeed, we have

$$\int_0^\infty e^{-\lambda x} (e^{-\eta x} W(t)) dx = \frac{1}{\psi(\eta + \lambda)} \underset{\lambda \rightarrow 0}{\sim} \frac{1}{\lambda \psi'(\eta)}.$$

Then, since  $t \mapsto e^{-\eta t} W(t)$  is non-decreasing according to equation (4) in [9] (applied with  $q = 0$ ), a Tauberian theorem entails the desired result (see for instance [2, p.10]).

To get the behavior of  $W'$ , since  $t \mapsto e^{-\eta t} W'(t)$  is not necessarily non-decreasing, we cannot use the same method. However, according to (4),

$$e^{-\eta t} W'(t) = b e^{-\eta t} W(t) - \int_0^t e^{-\eta(t-x)} W(t-x) e^{-\eta x} \Lambda(dx).$$

Since  $t \mapsto e^{-\eta t} W(t)$  is non-decreasing and converges to  $1/\psi'(\eta)$ , the monotone convergence theorem implies that

$$\lim_{t \rightarrow \infty} e^{-\eta t} W'(t) = \frac{b}{\psi'(\eta)} - \frac{1}{\psi'(\eta)} \int_0^\infty e^{-\eta x} \Lambda(dx).$$

The proof is then complete since from (2), we have  $0 = \psi(\eta) = \eta - b + \int_0^\infty e^{-\eta x} \Lambda(dx)$ .

## A.2 Proof of Lemma 2.2(ii)

We assume here that  $m = 1$  and  $\sigma^2 = \int_0^\infty z^2 \Lambda(dz) < \infty$ . First, notice that in that case  $\psi'(0) = 1 - m = 0$  and that  $\psi''(0) = \sigma^2$ . So

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)} \underset{\lambda \rightarrow 0}{\sim} \frac{2}{\sigma^2 \lambda^2}.$$

Hence, a Tauberian theorem yields the asymptotic behavior of  $W(t)$  as  $t \rightarrow \infty$ .

As in the supercritical case, it is not possible to obtain the asymptotic behavior of  $W'$  with the help of a Tauberian theorem because this function is not monotone. We use known results about the extinction time of a critical CMJ-process. According to the main result of [19], recalling that the measure  $\mu$  is defined by (1) and that  $\zeta$  is the lifespan of the ancestor, if

$$\lim_{t \rightarrow \infty} t^2(1 - \mu([0, t]) = 0 \tag{38}$$

$$\lim_{t \rightarrow \infty} t^2 \mathbb{P}(\zeta > t) = 0, \tag{39}$$

then

$$\lim_{t \rightarrow \infty} t \mathbb{P}(\Xi(t) > 0) = \frac{2 \int_0^\infty r \mu(dr)}{V}$$

where  $V$  is the variance of the number of children of the ancestor. Conditional on “the ancestor has a lifespan  $z$ ”, this number is Poisson with parameter  $bz$ . Then, we have  $V = b \int_0^\infty r^2 \Lambda(dr)$ . Moreover, according to (1),

$$\int_0^\infty r \mu(dr) = \int_0^\infty r dr \int_r^\infty \Lambda(dx) = \frac{\sigma^2}{2}.$$

Thus,

$$\mathbb{P}(\Xi(t) > 0) \underset{t \rightarrow \infty}{\sim} \frac{1}{tb}. \tag{40}$$

It remains to prove that the conditions (38) and (39) hold. By using (1), (38) is equivalent to the condition (9) that we have assumed here. Moreover,  $\mathbb{P}(\zeta > t) = b^{-1} \Lambda((t, +\infty])$ , which goes to 0 faster than  $t^{-2}$  as  $t \rightarrow \infty$  since the second moment of  $\Lambda$  is finite.

On the other hand, according to (5), we know that  $\mathbb{P}(\Xi(t) > 0) = \frac{W'(t)}{bW(t)}$ . Hence,  $W'(t)$  converges to  $2/\sigma^2$  as a consequence of (40) and of the fact that  $W(t)$  behaves as  $2/(\sigma^2 t)$  as  $t \rightarrow \infty$ .

## A.3 Proof of Lemma 2.2(iii)

As in the critical case, the asymptotic behavior of  $W$  can be obtained via a Tauberian theorem because  $\psi(\lambda) \sim (1 - m)\lambda$  as  $\lambda \rightarrow 0$ .

To study  $W'$  in that case, we use again known results about the time of extinction of CMJ-processes. The hypotheses (10) that we make about  $\tilde{\eta}_c$  enables us to use Theorem 6.7.10 in [21], that is,

$$\tilde{C} := \lim_{t \rightarrow \infty} e^{-\tilde{\eta}t} \mathbb{P}(\Xi(t) > 0)$$

exists and is positive if and only if

$$\mathbb{E} \left[ \int_0^\infty e^{-\tilde{\eta}t} \xi(dt) \log\{\xi((0, \infty))\} \right] < \infty$$

where we recall that  $\xi$  is the birth point process of the ancestor. By conditioning on the lifespan of the ancestor and by using classical properties about Poisson point processes, we have

$$\mathbb{E} \left[ \int_0^\infty e^{-\tilde{\eta}t} \xi(dt) \log\{\xi((0, \infty))\} \right] \leq \int_{(0, \infty)} \Lambda(dz) \frac{1 - e^{-\tilde{\eta}z}}{\tilde{\eta}} ((1-p) + b(1-p)^2 z)$$

which is finite thanks to (10). Then, we have proved that there exists  $\tilde{C} > 0$  such that

$$\mathbb{P}(\Xi(t) > 0) \underset{t \rightarrow \infty}{\sim} \tilde{C} e^{\tilde{\eta}t}. \quad (41)$$

It remains to compute  $\tilde{C}$  (which is unknown in most CMJ-processes). Since  $W(t)$  converges to  $(1-m)^{-1}$  as  $t \rightarrow \infty$ , using (5) and (41),

$$W'(t) \underset{t \rightarrow \infty}{\sim} \frac{b}{1-m} \tilde{C} e^{\tilde{\eta}t}. \quad (42)$$

Moreover, by integrating (3) by parts, since  $W(0) = 1$  according to [33, Lem 4.1], we have for  $\lambda > \tilde{\eta}$

$$\int_0^\infty W'(u) e^{-\lambda u} du = \frac{\lambda}{\psi(\lambda)} - 1$$

with the convention  $0/\psi(0) = 1/(1-m) < \infty$ . Then,

$$\int_0^\infty (W'(u) e^{-\tilde{\eta}u}) e^{-\lambda u} du = \frac{\lambda + \tilde{\eta}}{\psi(\lambda + \tilde{\eta})} \underset{\lambda \rightarrow 0}{\sim} \frac{\tilde{\eta}}{\lambda \psi'(\tilde{\eta})}.$$

On the other hand, using [5, Thm. 1.7.6] with (42), we get

$$\int_0^\infty (W'(u) e^{-\tilde{\eta}u}) e^{-\lambda u} du \underset{\lambda \rightarrow 0}{\sim} \frac{b\tilde{C}}{1-m} \frac{1}{\lambda},$$

which implies that  $\tilde{C} = \frac{\tilde{\eta}}{\psi(\tilde{\eta})}$ .

## Acknowledgments

This work was supported by project MANEGE ANR-09-BLAN-0215 (French national research agency). I want to thank Amaury Lambert for his very helpful remarks and comments and Guillaume Achatz who raised a question about the number of mutations undergone by an allele during a talk given at Collège de France. My thanks also to two anonymous referees for their careful check of this paper and for helpful remarks.

## References

- [1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.



- [2] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.
- [3] J. Bertoin. The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations. *Ann. Probab.*, 37(4):1502–1523, 2009.
- [4] J. Bertoin. A limit theorem for trees of alleles in branching processes with rare neutral mutations. *Stochastic Process. Appl.*, 120(5):678–697, 2010.
- [5] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [6] N. Champagnat and A. Lambert. Splitting trees with neutral Poissonian mutations I: Small families. *Stochastic Process. Appl.*, 122(3):1003–1033, 2012.
- [7] N. Champagnat and A. Lambert. Splitting trees with neutral Poissonian mutations II: Largest and oldest families. *Stochastic Processes and their Applications*, 123(4):1368 – 1414, 2013.
- [8] N. Champagnat, A. Lambert, and M. Richard. Birth and death processes with neutral mutations. *Int. J. Stoch. Anal.*, vol. 2012: Art. ID 569081, 20 pages, 2012.
- [9] T. Chan, A. E. Kyprianou, and M. Savov. Smoothness of scale functions for spectrally negative Lévy processes. *Probab. Theory Related Fields*, 150(3-4):691–708, 2011.
- [10] K. Danesh, R. Durrett, L. J. Havrilesky, and E. Myers. A branching process model of ovarian cancer. *Journal of Theoretical Biology*, 314(0):10 – 15, 2012.
- [11] C. Delaporte. Lévy processes with marked jumps II : Application to a population model with mutations at birth. *Preprint available at <http://fr.arxiv.org/abs/1305.6491>*, 2013.
- [12] R. A. Doney. Age-dependent birth and death processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 22:69–90, 1972.
- [13] R. Durrett and S. Moseley. Evolution of resistance and progression to disease during clonal expansion of cancer. *Theoretical Population Biology*, 77(1):42 – 48, 2010.
- [14] W. J. Ewens. The sampling theory of selectively neutral alleles. *Theoret. Population Biology*, 3:87–112; erratum, *ibid.* 3 (1972), 240; erratum, *ibid.* 3 (1972), 376, 1972.
- [15] R. A. Fisher, A. S. Corbet, and C. B. Williams. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, 12:42–58, 1943.
- [16] J. Geiger. Size-biased and conditioned random splitting trees. *Stochastic Process. Appl.*, 65(2):187–207, 1996.
- [17] J. Geiger and G. Kersting. Depth-first search of random trees, and Poisson point processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 111–126. Springer, New York, 1997.

- [18] R. C. Griffiths and A. G. Pakes. An infinite-alleles version of the simple branching process. *Adv. in Appl. Probab.*, 20(3):489–524, 1988.
- [19] J. M. Holte. Extinction probability for a critical general branching process. *Stochastic Processes and their Applications*, 2(3):303 – 309, 1974.
- [20] Y. Iwasa, M. A. Nowak, and F. Michor. Evolution of resistance during clonal expansion. *Genetics*, 172(4):2557–2566, April 2006.
- [21] P. Jagers. *Branching processes with biological applications*. Wiley-Interscience, London, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.
- [22] P. Jagers and O. Nerman. The growth and composition of branching populations. *Adv. in Appl. Probab.*, 16(2):221–259, 1984.
- [23] P. Jagers and O. Nerman. Limit theorems for sums determined by branching and other exponentially growing processes. *Stochastic Process. Appl.*, 17(1):47–71, 1984.
- [24] M. Kimmel and M. Mathaes. Modeling neutral evolution of Alu elements using a branching process. *BMC Genomics*, 11(Suppl 1):S11, 2010.
- [25] A. E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer-Verlag, Berlin, 2006.
- [26] A. Lambert. The contour of splitting trees is a Lévy process. *Ann. Probab.*, 38(1):348–395, 2010.
- [27] C. J. Mode. *Multitype branching processes. Theory and applications*. Modern Analytic and Computational Methods in Science and Mathematics, No. 34. American Elsevier Publishing Co., Inc., New York, 1971.
- [28] O. Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete*, 57(3):365–395, 1981.
- [29] M. A. Nowak, F. Michor, and Y. Iwasa. The linear process of somatic evolution. *Proceedings of the National Academy of Sciences*, 100(25):14966–14969, 2003.
- [30] A. G. Pakes. An infinite alleles version of the Markov branching process. *J. Austral. Math. Soc. Ser. A*, 46(1):146–169, 1989.
- [31] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [32] M. Richard. *Arbres, Processus de branchement non markoviens et Processus de Lévy*. PhD thesis, UPMC, Paris 6, 2011. Available at <http://tel.archives-ouvertes.fr/tel-00649235/fr/>.
- [33] M. Richard. Limit theorems for supercritical age-dependent branching processes with neutral immigration. *Advances in Appl. Probability*, 43-1:276–300, 2011.

- [34] S. Sagitov and M. C. Serra. Multitype Bienaymé-Galton-Watson processes escaping extinction. *Adv. in Appl. Probab.*, 41(1):225–246, 2009.
- [35] S. Sagitov and M. C. Serra. Skeletons of near-critical Bienaymé-Galton-Watson branching processes. *To appear in Adv. Appl. Probab.*, 2013.
- [36] Z. Taïb. *Branching processes and neutral evolution*. Lecture Notes in Biomathematics. 93. Berlin: Springer-Verlag. viii, 112 p. , 1992.
- [37] X. Wu and M. Kimmel. Modeling neutral evolution using an infinite-allele Markov branching process. *Int. J. Stoch. Anal.*, pages Art. ID 963831, 10, 2013.