Research Article

Birth and Death Processes with Neutral Mutations

Nicolas Champagnat,[1,2] Amaury Lambert,[3] and Mathieu Richard[3,4]

[1] IECN, Université de Lorraine, Campus Scientifique, B.P. 70239,
  54506 Vandœuvre-lès-Nancy Cedex, France
[2] Inria, 54600 Villers-lès-Nancy, France
[3] Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599 CNRS and UPMC Université Paris 06,
  Case courrier 188, 4 Place Jussieu, 75252 Paris Cedex 05, France
[4] CMAP, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France

Correspondence should be addressed to
Mathieu Richard, mathieu.richard@cmap.polytechnique.fr

We review recent results of ours concerning branching processes with general lifetimes and neutral mutations, under the infinitely many alleles model, where mutations can occur either at the birth of particles or at a constant rate during their lives. In both models, we study the allelic partition of the population at time t. We give closed-form formulae for the expected frequency spectrum at t and prove a pathwise convergence to an explicit limit, as $t \to +\infty$, of the relative numbers of types younger than some given age and carried by a given number of particles (small families). We also provide the convergences in distribution of the sizes or ages of the largest families and of the oldest families. In the case of exponential lifetimes, population dynamics are given by linear birth and death processes, and we can most of the time provide general formulations of our results unifying both models.

1. Introduction

We consider a general branching model, where particles have i.i.d. (not necessarily exponential) life lengths and give birth at a constant rate b during their lives to independent copies of themselves. The genealogical tree thus produced is called *splitting tree* [1–3]. The process that counts the number of the alive particles through time is a *Crump-Mode-Jagers process* (or general branching process) [4] which is binary (births occur singly) and homogeneous (constant birth rate).

We enrich this genealogical model with mutations. In Model I, each child is a clone of her mother with probability $1-p$ and a mutant with probability p. In Model II, independently of other particles, each particle undergoes mutations during her life at constant rate θ (and births are always clonal). For both models, we are working under the infinitely many alleles model; that is, a mutation yields a type, also called *allele*, which was never encountered before. Moreover, mutations are supposed to be neutral; that is, they do not modify the way particles die and reproduce. For any type and any time t, we call *family* the set of all particles that share this type at time t.

Branching processes (and especially birth and death processes) with mutations have many applications in biology. In carcinogenesis [5–10], they can model the evolution of cancerous cells. In [11], Kendall modeled carcinogenesis by a birth and death process where mutations occur during life according to an inhomogeneous Poisson process. In [8, 10], 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 this study is the time τ_k of appearance of the first cell of type k. In [7], the authors study the arrival time of the first resistant cell and the number of resistant cells, in a model of cancerous cells undergoing a medical treatment and becoming resistant after having experienced a certain number of mutations.

Branching processes with mutations are also used in epidemiology. Epidemics, and especially their onset, can be modeled by birth and death processes, where particles are infected hosts, births are disease transmissions, and deaths are recoveries or actual deaths. In [12], Stadler provides a statistical method for the inference of transmission rates and of the reproductive value of epidemics in a birth and death model with mutations. In [13], Lambert and Trapman enriched the transmission tree with Poissonian marks modeling detection events of hospital patients infected by an antibiotic-resistant pathogen. They provided an inference method based on the knowledge of times spent by patients at the hospital at the detection of the outbreak.

Let us also mention the existence of models, for example, [14], of phage reproduction within a bacterium by a (possibly time inhomogeneous) birth and death process with Poissonian mutations, where particles model phage in the vegetative phase (DNA strands in the bacterium without protein coating) and death is interpreted as phage maturing (reception of protein coating).

In ecology, the neutral theory of biodiversity [15] gives a prediction of the diversity patterns, in terms of species abundance distributions, that are generated by individual-based models where speciation is caused by mutation or by immigration from mainland. Usually, the underlying genealogical models are assumed to keep the population size constant through time, as in the Moran or Wright-Fisher models, and so have the same well-known properties as models in mathematical population genetics (e.g., Ewens sampling formula), with a different interpretation. See [16, 17] for cases where this assumption is relaxed in favor of the branching property.

In this paper, we are first interested in the allelic partition of the population and more precisely in properties about the *frequency spectrum* $(M_t^{i,a}, i \geq 1)$, where $M_t^{i,a}$ is the number of distinct types younger than a (i.e., whose original mutation appeared after $t - a$) carried by exactly i particles at time t. This kind of question was first studied by Ewens [18] 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 models, it is not possible to obtain a counterpart of Ewens sampling formula but we obtain different kinds of results concerning the frequency spectrum $(M_t^{i,a}, i \geq 1)$. First, we

get a closed-form formula for the expected frequency spectrum, even in the non-Markovian cases. Second, we get pathwise convergence results as $t \to +\infty$ on the survival event, of the relative abundances of types. Third, we investigate the order of magnitude of the sizes of the largest families at time t and of the ages of oldest types at time t, as $t \to +\infty$, and show the convergence in distribution of these quantities properly rescaled. Several regimes appear, depending on whether the *clonal process*, which is the process counting particles of a same type, is subcritical, critical, or supercritical.

We do not know any previous mathematical studies, other than ours, on branching processes with Poissonian mutations, but there are several existing mathematical results on branching models with mutations at birth that we now briefly review.

In discrete time, Griffiths and Pakes [19] studied the case of a Bienaymé-Galton-Watson (BGW) process where at each generation, all particles mutate independently with some probability u. The authors obtained properties about the number of alleles/types in the population, about the time of last mutation in the (sub)critical case and about the expected frequency spectrum. In [20, 21], Bertoin considers an infinite alleles model with neutral mutations in a subcritical or critical BGW process where particles independently give birth to a random number of clonal and mutant children according to the same joint distribution. In [20], the tree of alleles is studied, where all particles of a common type are gathered in clusters and the law of the allelic partition of the total population is given by describing the joint law of the sizes of the clusters and of the numbers of their mutant children. In [21], Bertoin obtains the joint convergence of the sizes of allelic families in the limit of a large initial population size and a small mutation rate.

In continuous time, Pakes [22] studied Markovian branching processes and gave the counterpart in the time-continuous setting, of properties found in the previously cited paper [19]. In particular, his results about the frequency spectrum and the "limiting frequency spectrum" are similar to ours, stated in Section 3. Recently, Maruvka et al. [23, 24] have considered the linear birth and death process with Poissonian mutations. Actually, they rather studied a PDE satisfied by a concentration $n(x, t)$ which can be seen as (but is not proved to be) a deterministic approximation to the number of families of size x at time t. It is remarkable that this PDE has a steady concentration $\bar{n}(x)$, whose behavior as $x \to \infty$ is comparable to the asymptotic behavior of the relative numbers of families of size m as $m \to +\infty$ in the discrete model studied here and in [19]. In the monography [25], Taïb is interested in general branching processes known as Crump-Mode-Jagers processes (see [4, 26] and references therein) where mutations still occur at birth but with a probability that may depend, for example, on the age of the mother. He obtained limited theorems about the frequency spectrum by using random characteristics techniques but in most cases, limits cannot be explicitly computed. Some of our results in Model I are applications of Taïb's, but use techniques specific to splitting trees to yield explicit formulae. We have refrained to apply results of Taïb on the convergence in distribution of properly rescaled sizes of the largest families, on the validity of which we have doubts in the case of supercritical clonal processes (see last section).

The paper is organized as follows. In Section 2, we define the models and give some of their properties that will be useful to state the main results. Section 3 is devoted to the study of the frequency spectrum (small families). Finally, in Section 4, we give the results about ages of the oldest families and about sizes of the largest ones.

Notice that in this paper, most of the results are stated for linear birth and death processes in order to simplify the notation. Most of them are also true with general life length distributions and are proved in Chapter 3 of the Ph.D. thesis [27] for Model I and in [28, 29]

for Model II. Specific effort has been put on finding a unifying formulation for our results as soon as it seemed possible.

2. The Models

2.1. Model without Mutations

We first define the model without mutations and give some of its properties. Afterwards, we will explain the two mutation mechanisms that we consider in this paper.

   As a population model, we consider *splitting trees* [1–3]; that is,

   (i) at time $t = 0$, the population starts with one progenitor;

   (ii) all particles have i.i.d. reproduction behaviors;

   (iii) conditional on her birth date α and her life length ζ, each particle gives birth at a
        constant rate $b \in (0, \infty)$ during $(\alpha, \alpha + \zeta)$, to a single particle at each birth event.

It is important to notice that the common law of life lengths can be as general as possible. Let $Z = (Z(t), t \geq 0)$ be the process counting the number of extant particles through time. We denote the lifespan distribution by $\Lambda(\cdot)/b$, where Λ is a finite positive measure on $(0, +\infty]$ with total mass b and is called a *lifespan measure* [3].

   The total population process Z belongs to a large class of branching processes called *Crump-Mode-Jagers* or *CMJ processes*. In these processes, also called general branching processes [4, 26], one associates with each particle x in the population a non-negative r.v. λ_x (her life length) and a point process ξ_x called birth point process. One assumes that the sequence $(\lambda_x, \xi_x)_x$ is i.i.d. but λ_x and ξ_x are not necessarily independent. Then, the CMJ process is defined as

$$Z(t) = \sum_x \mathbf{1}_{\{\sigma_x \leq t < \sigma_x + \lambda_x\}}, \quad t \geq 0, \tag{2.1}$$

where for any particle x in the population, σ_x is her birth time.

   In our particular case, the common distribution of lifespans is $\Lambda(\cdot)/b$ and conditional on her lifespan, the birth point process of a particle is distributed as a Poisson point process during her life. We can say that the CMJ process Z is *homogeneous* (constant birth rate) and *binary* (births occur singly). We will say that Z is *subcritical, critical, or supercritical* according to whether the mean number of children per particle

$$m := \int_0^\infty u \Lambda(\mathrm{d}u), \tag{2.2}$$

is less than, equal to, or greater than 1.

   The advantage of homogeneous, binary CMJ processes is that they allow for explicit computations, for example, about one-dimensional marginals of Z (see forthcoming Proposition 2.1). More precisely, for $\lambda \geq 0$, define

$$\psi(\lambda) := \lambda - \int_{(0,\infty]} \left(1 - e^{-\lambda u}\right) \Lambda(\mathrm{d}u), \tag{2.3}$$

and let r be the greatest root of ψ. Notice that ψ is convex, $\psi(0) = 0$, and $\psi'(0) = 1 - m$. As a consequence,

$$r = 0 \quad \text{if } Z \text{ is subcritical or critical,}$$
$$r > 0 \quad \text{if } Z \text{ is supercritical.} \tag{2.4}$$

Let W be the so-called *scale function* [30, page 194] associated with ψ, that is, the unique increasing continuous function $(0, \infty) \to (0, \infty)$ satisfying

$$\int_0^\infty W(x)e^{-\lambda x}\mathrm{d}x = \frac{1}{\psi(\lambda)}, \quad \lambda > r. \tag{2.5}$$

Proposition 2.1 (Lambert [3, 17]). *The one-dimensional marginals of Z are given by*

$$\mathbb{P}(Z(t) = 0) = 1 - \frac{W'(t)}{bW(t)} \tag{2.6}$$

and for $n \geq 1$,

$$\mathbb{P}(Z(t) = n) = \left(1 - \frac{1}{W(t)}\right)^{n-1} \frac{W'(t)}{bW(t)^2}. \tag{2.7}$$

In other words, conditional on being nonzero, $Z(t)$ is distributed as a geometric r.v. with success probability $1/W(t)$.

If $\text{Ext} := \{Z(t) \xrightarrow[t \to \infty]{} 0\}$ denotes the extinction event of Z, according to [3], as a consequence of the last proposition,

$$\mathbb{P}(\text{Ext}) = 1 - \frac{r}{b}. \tag{2.8}$$

Thus, thanks to (2.4), extinction occurs a.s. when Z is (sub)critical and $\mathbb{P}(\text{Ext}^c) > 0$ when it is supercritical.

The following proposition justifies the fact that r is called the *Malthusian parameter* of the population in the supercritical case.

Proposition 2.2 (Lambert [3]). *If $m > 1$, conditional on the survival event Ext^c,*

$$e^{-rt}Z(t) \xrightarrow[t \to \infty]{} \mathcal{E} \quad a.s., \tag{2.9}$$

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

In fact, convergence in distribution is proved in [3] and a.s. convergence holds according to [31] (see [32, page 285]).
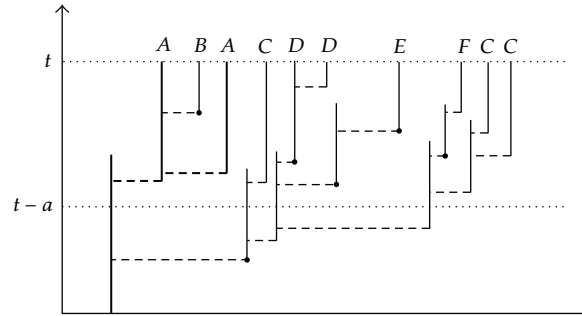
Figure 1: An example of a splitting tree in Model I and of the allelic partition of the whole extant population at time t. Vertical axis is time and horizontal axis shows filiation (horizontal lines have zero length). Full circles represent mutations at birth and thick lines, the clonal splitting tree of the ancestor up to time t. The different letters are the alleles of alive particles at time t.
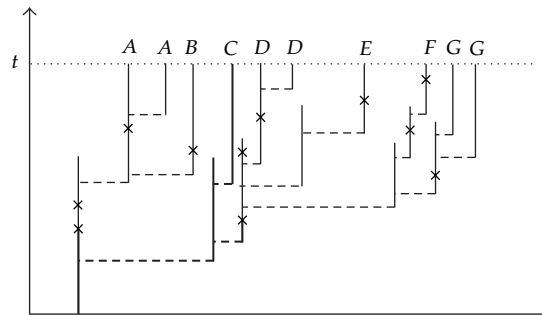

Figure 2: An example of a splitting tree with mutations in Model II and of the allelic partition of the whole extant population at time t. Crosses represent mutations and thick lines represent the clonal splitting tree of the ancestor up to time t. The different letters are the alleles of alive particles at time t.

2.2. Two Mutation Models I and II

We now assume that particles in the population carry types, also called *alleles*. We consider two population models where mutations appear in different ways. In each case, we will make the assumption of *infinitely many alleles*; that is, to every mutation event is associated a different type, so that every type appears only once. We will also assume that mutations are *neutral*; that is, they do not change the way particles die and reproduce.

In Model I, mutations occur at birth. More precisely, there is some $p \in (0, 1)$ such that at each birth event, independently of all other particles, the newborn is a clone of her mother with probability p and a mutant with probability $1 - p$. An illustration is given in Figure 1.

In Model II, particles independently experience mutations during their lives at constant rate $\theta > 0$. In particular, in contrast with Model I, particles can change type several times during their lifetime, but always bear at birth the same type as their mother at this very time. An illustration is given in Figure 2.

In what follows, an important role will be played by the *clonal process*, generically denoted $Z_\star$, counting, as time passes, the number of particles bearing the same type as the progenitor of the population at time 0. It can easily be seen that the genealogy of a clonal population is again a splitting tree, so that $Z_\star$ is also a homogeneous, binary CMJ process. We denote by $b_\star$ its birth rate, by $\psi_\star$ the associated convex function as in (2.3), and by $W_\star$ the

nonnegative function with Laplace transform $1/\psi_\star$. Furthermore, when the clonal population is supercritical, that is, when $\psi'_\star(0+) < 0$, we denote by $r_\star$ its Malthusian parameter, which is the only nonzero root of $\psi_\star$. We will sometimes need to have this generic notation depending on the model considered: Z_p, ψ_p, W_p, and r_p for Model I and Z_θ, ψ_θ, W_θ, and r_θ for Model II.

Concerning Model I, it can be seen [27] that the clonal splitting tree has the same life lengths as the original splitting tree and birth rate $b_p = b(1-p)$, so that its lifespan measure is $(1-p)\Lambda$ and

$$\psi_p(\lambda) = p\lambda + (1-p)\psi(\lambda), \quad \lambda > 0. \tag{2.10}$$

In particular, as in (2.2), the clonal population is subcritical, critical, or supercritical according to whether $m(1-p)$ is less than, equal to, or greater than 1. It should be noted that there is no closed-form formula for W_p.

Concerning Model II, it can be seen [28] that the clonal splitting tree has birth rate $b_\theta = b$ and life lengths distributed as $\min(X, Y)$ where X has the probability distribution $\Lambda(\cdot)/b$ and Y is an independent exponential r.v. with parameter θ. Then we get

$$\psi_\theta(\lambda) = \frac{\lambda\psi(\lambda + \theta)}{\lambda + \theta}, \quad \lambda > 0. \tag{2.11}$$

In particular, $r_\theta = r - \theta$ and the clonal population is subcritical, critical, or supercritical according to whether r is less than, equal to, or greater than θ. It can also be proved that W and W_θ are differentiable and that their derivatives are related via

$$W'_\theta(x) = e^{-\theta x} W'(x), \quad x \geq 0, \tag{2.12}$$

with the requirement that $W_\theta(0) = 1$.

2.3. Exponential Case

An interesting case that we will focus on is the *exponential (or Markovian) case*, when the common distribution of life lengths is exponential with parameter d (with the convention that lifespans are a.s. infinite if $d = 0$), that is, $\Lambda(du) = bde^{-du}du$ or $\Lambda(du) = b\delta_\infty(du)$. In that case, Z is, respectively, a linear birth and death process with birth rate b and death rate d or a pure birth process (or Yule process) with parameter b.

In this case, Z and $Z_\star$ are Markov processes and the quantities defined in Section 2.1 are computable. Indeed, we have

$$\psi(\lambda) = \frac{\lambda(\lambda - b + d)}{\lambda + d}, \quad r = b - d,$$

$$m = 1 - \psi'(0) = \frac{b}{d}, \quad \psi'(r) = 1 - \frac{d}{b}. \tag{2.13}$$

It is also possible to compute the function W, defined by (2.5), while it is generally unknown. From [3, page 393], we have

$$W(x) = \begin{cases} \dfrac{be^{rx} - d}{r}, & \text{if } b \neq d, \\ 1 + bx, & \text{if } b = d, \end{cases} \quad x \geq 0, \tag{2.14}$$

and in all cases

$$W'(x) = be^{rx}, \quad x \geq 0. \tag{2.15}$$

The same results hold for $W_\star$, by, respectively, replacing b, d, and r by

$$b_\star := b(1 - p), \qquad d_\star := d, \qquad r_\star := r - bp, \tag{2.16a}$$

in Model I and by

$$b_\star := b, \qquad d_\star := d + \theta, \qquad r_\star := r - \theta, \tag{2.16b}$$

in Model II.

We will sometimes state results in the total generality of splitting trees, in which case an equation numbered (-a) (resp. (-b)) refers to Model I (resp. Model II), as done previously. However, we will most of the time focus on the exponential case, in which we will as soon as possible use the unified notation using $\star$'s. We will notify when the results can be generalized and will give precise references.

Remark 2.3. In the exponential case, notice that Models I and II are two (incompatible) cases of a more general class of linear birth and death processes with mutations, where particles mutate spontaneously at rate θ, die at rate d, and give birth at rate b and at each birth event: with probability p_2, the mother and the daughter both mutate (and bear either the same new type or two different new types); with probability p_1, the daughter (only) mutates; with probability $p_0 = 1 - p_1 - p_2$, none of them mutates. Then Model I corresponds to the case when $\theta = p_2 = 0$ and Model II to the case when $p_1 = p_2 = 0$. The case studied by Pakes in [22] corresponds to $\theta = 0$, $p_0 = u^2$, $p_1 = 2u(1 - u)$, and $p_2 = (1 - u)^2$. It is still an open question to check whether, when our results hold for both Models I and II with the unified notation, they hold for all linear birth and death processes with mutations.

3. Small Families

Recall that a *family* is a maximal set of particles bearing the same type at the same given time. In this section, we are interested in results about *small families*; that is, families whose sizes and ages are fixed, in opposition to those of Section 4 which concern asymptotic properties of the largest and oldest ones.

More precisely, we give properties of the allelic partition of the entire population by studying the *frequency spectrum* $(M_t^{i,a}, i \geq 1)$, where $M_t^{i,a}$ denotes the number of distinct types,

whose ages are less than a at time t, carried by exactly i particles at time t. Notice that $M_t^{i,t}$ is simply the number of alleles carried by i particles at time t (regardless of their ages).

For instance, in Figure 1, the frequency spectrum $(M_t^{i,t}, i \geq 1)$ is $(3, 2, 1, 0, \dots)$ because three alleles (B, E, and F) are carried by one particle, A and D are carried by two particles, and C is the only allele carried by three particles. Moreover, if we only consider families with ages less than a, $(M_t^{i,a}, i \geq 1)$ equals $(3, 1, 0, \dots)$ because alleles A and C appear in the population before time $t - a$. Similarly, in Figure 2, the frequency spectrum in Model II is $(4, 3, 0, \dots)$.

In the case of branching processes, there is no closed-form formula available for the law of the frequency spectrum as it is the case for the Wright-Fisher model thanks to Ewens sampling formula [18]. Nevertheless, we obtained for both mutation models an exact computation of the expected frequency spectrum and an almost sure asymptotic behavior of this frequency spectrum as $t \to +\infty$.

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 types carried by i particles 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 $i \geq 1$, one has*

$$\mathbb{E}\left[M_t^{i,da}\right] = \frac{p}{b(1-p)} W'(t-a) \left(1 - \frac{1}{W_p(a)}\right)^{i-1} \frac{W_p'(a)}{W_p^2(a)} \, da, \tag{3.1a}$$

$$\mathbb{E}\left[M_t^{i,da}\right] = \frac{\theta W'(t)}{b} \left(1 - \frac{1}{W_\theta(a)}\right)^{i-1} \frac{e^{-\theta a}}{W_\theta^2(a)} \, da. \tag{3.1b}$$

In the exponential case, both expressions are read as

$$\mathbb{E}\left[M_t^{i,da}\right] = (r - r_\star) e^{rt} \left(1 - \frac{1}{W_\star(a)}\right)^{i-1} \frac{e^{-(r-r_\star)a}}{W_\star^2(a)} \, da. \tag{3.2}$$

In [27], (3.1a) is proved in the general case. Its proof uses the branching property and basic properties about Poisson processes. The main argument is that conditional on $Z(t-a)$, $M_t^{i,da}$ is the sum of $Z(t-a)$ independent r.v. distributed as the number of mutants that appear in the population in a time interval da and with i clonal alive descendants at time a. The proof of the general case of (3.1b) in [28] is based on coalescent point processes.

The expected frequency spectrums $\mathbb{E}[M_t^{i,a}]$ can be obtained by integrating (3.1a) and (3.1b) over ages. Taking into account the contribution of the type of the progenitor, we can prove the following result.

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

$$
\mathbb{E}\left[M_t^{i,a}\right] = \frac{p}{b(1-p)} \int_0^a W'(t-x)\left(1 - \frac{1}{W_p(x)}\right)^{i-1} \frac{W_p'(x)}{W_p^2(x)} \mathrm{d}x
$$

$$
+ \frac{1}{b(1-p)}\left(1 - \frac{1}{W_p(t)}\right)^{i-1} \frac{W_p'(t)}{W_p^2(t)} \mathbf{1}_{\{a=t\}}, \tag{3.3a}
$$

$$
\mathbb{E}\left[M_t^{i,a}\right] = \frac{\theta W'(t)}{b} \int_0^a \left(1 - \frac{1}{W_\theta(x)}\right)^{i-1} \frac{e^{-\theta x}}{W_\theta^2(x)} \mathrm{d}x
$$

$$
+ W(t)\left(1 - \frac{1}{W_\theta(t)}\right)^{i-1} \frac{e^{-\theta t}}{W_\theta^2(t)} \mathbf{1}_{\{a=t\}}. \tag{3.3b}
$$

In the exponential case,

$$
\mathbb{E}\left[M_t^{i,a}\right] = (r - r_\star)e^{rt} \int_0^a \left(1 - \frac{1}{W_\star(x)}\right)^{i-1} \frac{e^{-(r-r_\star)x}}{W_\star^2(x)} \mathrm{d}x + \mathbb{P}(Z_\star(t) = i)\mathbf{1}_{\{a=t\}}. \tag{3.4}
$$

The second terms that appear in the r.h.s. correspond to the probabilities that the progenitor has i alive clonal descendants at time t. In the exponential case, we left this probability as such, since its expression depends on the model. It is also possible to get similar equations for the number of families with ages less than a (resp. with size i) by summing over i (resp. by taking $a = t$) in the last expressions.

Remark 3.3. In the exponential case, when the process Z is critical, that is, when $r = b - d = 0$, for $a < t$,

$$
\mathbb{E}\left[M_t^{i,a}\right] = \frac{|r_\star|}{b_\star} \frac{1}{i}\left(1 - \frac{1}{W_\star(a)}\right)^i, \tag{3.5}
$$

which is reminiscent of Fisher log-series of species abundances [17]. Surprisingly, this expression is independent of $t \in (a, \infty)$.

From Corollary 3.2, we deduce the asymptotic behavior of $\mathbb{E}[M_t^{i,a}]$ in the supercritical case.

Proposition 3.4. *We suppose that $m > 1$. In the general case,*

$$
\lim_{t \to +\infty} e^{-rt}\mathbb{E}\left[M_t^{i,a}\right] = \frac{r}{b} \frac{1}{\psi'(r)} J^{i,a}, \tag{3.6}
$$

where, for Model I,

$$
J^{i,a} := \frac{p}{(1-p)} \int_0^a \left(1 - \frac{1}{W_p(u)}\right)^{i-1} \frac{e^{-ru} W_p'(u)}{W_p^2(u)} \mathrm{d}u \tag{3.7a}
$$

and, for Model II,

$$J^{i,a} := \theta \int_0^a \left(1 - \frac{1}{W_\theta(u)}\right)^{i-1} \frac{e^{-\theta u}}{W_\theta^2(u)} du. \tag{3.7b}$$

In the exponential case, one gets the simpler formula:

$$J^{i,a} = (r - r_\star) \int_0^a \left(1 - \frac{1}{W_\star(u)}\right)^{i-1} \frac{e^{-(r-r_\star)u}}{W_\star^2(u)} du. \tag{3.8}$$

Notice that $\mathbb{E}[M_t^{i,a}]$ grows exponentially with parameter r, as does Z on its survival event.

3.2. Convergence Results

In this section and in all following ones, we are interested in long-time behaviors in the two models we consider. Then, from now on, *we assume that the process Z is supercritical.*

This paragraph deals with the improvements of the convergence results (3.6) regarding the expected frequency spectrum. The following results yield the asymptotic behavior as $t \to +\infty$ of the frequency spectrum $(M_t^{i,a}, i \geq 1)$, conditional on the survival event.

The main technique we use to prove them is CMJ processes counted with random characteristics (see [4] and Appendix A in [25]). It enables us to obtain several pathwise convergence results regarding some processes embedded in the supercritical splitting tree.

A characteristic is a random nonnegative function on $[0, +\infty)$. To each particle x in the population is associated a characteristic χ_x, which can be viewed as a score or a weight. It must satisfy that $(\lambda_x, \zeta_x, \chi_x)_x$ is an i.i.d. sequence, where we recall that λ_x is the life length of x and ζ_x its birth process. Then, the process counted with the characteristic χ is defined as

$$Z^\chi(t) := \sum_x \chi_x(t - \sigma_x) \mathbf{1}_{\{\sigma_x \leq t\}}. \tag{3.9}$$

For instance, if $\chi(t) = \mathbf{1}_{\{t \leq \lambda_x\}}$, Z^χ equals Z and if $\chi(t) = \mathbf{1}_{\{t \leq \lambda_x \wedge a\}}$, $Z^\chi(t)$ is the number of extant particles at time t with ages less than a, then, provided technical conditions about χ are satisfied, the convergences of $e^{-rt} Z^\chi(t)$ and of $Z^\chi(t)/Z(t)$ as $t \to +\infty$ hold a.s. on the survival event. In our case, when χ is appropriately chosen, we can use this result to obtain the following statements.

Proposition 3.5. *Let M_t be the number of extant types at time t. Almost surely, on the survival event of Z,*

$$\lim_{t \to +\infty} e^{-rt} M_t = J\mathcal{E},$$

$$\lim_{t \to +\infty} e^{-rt} M_t^{i,a} = J^{i,a} \mathcal{E}, \tag{3.10}$$

where in Model I,

$$J := \frac{rp}{1-p} \int_0^\infty e^{-ru} \ln(W_p(u)) du,$$
(3.11a)

while in Model II,

$$J := \theta \int_0^\infty \frac{e^{-\theta x}}{W_\theta(x)} dx$$
(3.11b)

and where $\mathcal{E}$ is the r.v. defined by (2.9).

In the exponential case, one has

$$J = (r - r_\star) \int_0^\infty \frac{e^{-(r-r_\star)u}}{W_\star(u)} du.$$
(3.12)

Notice that (3.10) is consistent with (3.6) since $\mathbb{P}(\text{Ext}^c) = r/b$ and $\mathbb{E}[\mathcal{E}] = 1/\psi'(r)$. Moreover, (3.10) still holds after $M_t^{i,a}$ is replaced by $M_t^{i,t}$ and $J^{i,a}$ by $J^{i,\infty}$.

3.3. Asymptotic Behavior of the Limiting Frequency Spectrum

Thanks to Proposition 3.5, the proportion $M_t^{i,a}/M_t$ of types carried by i particles and with ages less than a converges a.s. to $J^{i,a}/J$ as $t \to +\infty$. This limit is called "the limiting frequency spectrum" by Pakes in [22]. This paragraph is devoted to the asymptotic behavior, as $i \to +\infty$, of $J^i := J^{i,\infty}$, obtained by taking $a = \infty$ in (3.7a) and (3.7b). In the exponential case,

$$J^i = (r - r_\star) \int_0^\infty \left(1 - \frac{1}{W_\star(u)}\right)^{i-1} \frac{e^{-(r-r_\star)u}}{W_\star^2(u)} du.$$
(3.13)

3.3.1. Supercritical Case

In this paragraph, we only treat the exponential case. Let us assume that the clonal process is supercritical, that is, $r_\star > 0$. Define

$$\nu := \frac{r}{r_\star}, \qquad \mu := \frac{b_\star}{r_\star}, \qquad \gamma := \frac{r - r_\star}{b_\star}.$$
(3.14)

We have $\gamma = p/(1-p)$ in Model I and $\gamma = \theta/b$ in Model II. Recall that J^i is the proportion of types carried by i particles in the large time asymptotic.

Proposition 3.6. *In the exponential case, one has for both models*

$$J^i \underset{i \to +\infty}{\sim} i^{-1-\nu} \gamma \Gamma(\nu + 1) \mu^\nu.$$
(3.15)

Notice that this result is consistent with [24] where Maruvka et al. use an approximation of the frequency spectrum by a concentration driven by a PDE and with [22] where Pakes considers Markov branching processes with multiple simultaneous births, binomial mutations at birth, and no Poissonian mutations.

Remark 3.7. The following proof of Proposition 3.6 easily extends to any life length distributions since it is based on Proposition 2.2 which holds in the general case.

Proof of Proposition 3.6. Since $W_\star(t) \geq 1$ for $t \geq 0$, the sequence $(J^i)_{i \geq 1}$ is positive and non-increasing. Then, according to a Tauberian theorem about series, to prove Proposition 3.6, it is sufficient to prove that $\sum_{i \geq j} J^i$ is equivalent to $i^{-\nu} \gamma \Gamma(\nu) \mu^\nu$ as $j \rightarrow +\infty$.

Recalling (3.13), we have

$$\sum_{i \geq j} J^i = (r - r_\star) \int_0^\infty e^{-rt} \mathbb{P}(Z_\star(t) \geq j) \mathrm{d}t \tag{3.16}$$

and from now on, we follow the proof of [22, Theorem 3.2.1]. Let $s > 0$ be such that $j = e^{r_\star s}$. Then $j^\nu = e^{rs}$ and

$$j^\nu \int_0^\infty e^{-rt} \mathbb{P}(Z_\star(t) \geq j) \mathrm{d}t = \int_{-s}^\infty e^{-rt} \mathbb{P}(Z_\star(t+s) \geq e^{r_\star s}) \mathrm{d}t$$
$$= \int_{-s}^\infty e^{-rt} \mathbb{P}(e^{-r_\star(t+s)} Z_\star(t+s) \geq e^{-r_\star t}) \mathrm{d}t. \tag{3.17}$$

Using Proposition 2.2 and (2.13), $\mathbb{P}(e^{-r_\star(t+s)} Z_\star(t+s) \geq e^{-r_\star t}) \xrightarrow[s \rightarrow +\infty]{} \mathbb{P}(E_\star \geq e^{-r_\star t})$, where $E_\star$ is a nonnegative r.v. such that

$$\mathbb{P}(E_\star = 0) = 1 - \frac{r_\star}{b_\star}, \tag{3.18}$$

and conditional on $\{E_\star > 0\}$, $E_\star$ is an exponential r.v. with parameter $\psi'_\star(r_\star) = 1 - d_\star/b_\star = 1/\mu$. Moreover, using Markov inequality, for $\varepsilon > 0$,

$$\mathbb{P}(e^{-r_\star(t+s)} Z_\star(t+s) \geq e^{-r_\star t}) \leq e^{r_\star(\nu+\varepsilon)t} \mathbb{E}\left[\left(Z_\star(s+t) e^{-r_\star(s+t)}\right)^{\nu+\varepsilon}\right] \leq C e^{r_\star(\nu+\varepsilon)t}, \tag{3.19}$$

using again the a.s. convergence in Proposition 2.2. Then, for $s > 0$ and $t \in \mathbb{R}$, we have

$$e^{-rt} \mathbb{P}(e^{-r_\star(t+s)} Z_\star(t+s) \geq e^{-r_\star t}) \leq e^{-rt} \mathbf{1}_{\{t>0\}} + C e^{r_\star \varepsilon t} \mathbf{1}_{\{t<0\}}, \tag{3.20}$$

and thanks to the dominated convergence theorem,

$$j^\nu \int_0^\infty e^{-rt} \mathbb{P}(Z_\star(t) \geq j) \mathrm{d}t \xrightarrow[j \rightarrow +\infty]{} \int_{\mathbb{R}} e^{-rt} \mathbb{P}(E_\star \geq e^{-r_\star t}) \mathrm{d}t = \frac{1}{\mu} \int_{\mathbb{R}} e^{-rt} e^{-(1/\mu)e^{-r_\star t}} \mathrm{d}t. \tag{3.21}$$

The change of variables $x = e^{-r_\star t}$ in the last integral leads to

$$j^\nu \int_0^\infty e^{-rt} \mathbb{P}(Z_\star(t) \geq j) dt \xrightarrow[j \to +\infty]{} \frac{1}{b_\star} \int_0^\infty x^{\nu-1} e^{-x/\mu} dx = \frac{\Gamma(\nu)}{b_\star} \mu^\nu, \tag{3.22}$$

which terminates the proof. □

3.3.2. Critical Case

We want to obtain a similar result to Proposition 3.6 when the clonal population is critical. It seems that this is not possible in a general setting due to the non explicit expression of the functions W_p and W_θ. However, in the exponential and critical case, we have the simpler expression $W_\star(t) = 1 + b_\star t$ and $r_\star = 0$. Then, we have

$$J^i = r \int_0^\infty e^{-rt} \left(1 - \frac{1}{1 + b_\star t}\right)^{i-1} \frac{1}{(1 + b_\star t)^2} dt. \tag{3.23}$$

Proposition 3.8. *In the exponential case, one has*

$$J_\star^i \underset{i \to +\infty}{\sim} C(\gamma) i^{-3/4} e^{-2\sqrt{\gamma i}}, \tag{3.24}$$

where one recalls that here $\gamma = r/b_\star = r/d_\star$ and one has set $C(x) = \sqrt{\pi} e^{x/2} x^{5/4}$ for $x > 0$.

Proof. By a change of variables, we have set

$$J^i = \frac{r}{b_\star} \int_0^\infty e^{rt/b_\star} \frac{t^{i-1}}{(1 + t)^{i+1}} dt = \gamma \Gamma(i) U(i, 0, \gamma), \tag{3.25}$$

where U is known as a confluent hypergeometric function (see [33, Chapter 13]). Then, using [22, Theorem 3.3.2] with $B = -2$, we have the result. □

4. Asymptotic Results about Large and Old Families

We now state results about ages of the oldest families and about sizes of the largest ones. We mainly focus on the case when clonal populations are subcritical. Then, in Section 4.3, we explain which results hold in the critical and supercritical cases.

We need some notation. For $t \geq 0$,

(i) for $a \geq 0$, let $O_t(a)$ be the number of extant families at time t, with ages greater than a (O for "old"); for convenience, we set $O_t(a) = 0$ if $a < 0$,

(ii) for $x \in \mathbb{R}$, let $L_t(x)$ be the number of families with sizes greater than x at time t (L for "large").

In this section, we are interested in finding the orders of magnitudes of the ages and of the sizes of the families; that is, in finding numbers c_t and x_t such that $\mathbb{E}[O_t(c_t)]$ and $\mathbb{E}[L_t(x_t)]$ converge to positive and finite real numbers as $t \to +\infty$.

4.1. Ages of Old Families in the Subcritical Case

In this section, we suppose that the clonal processes are subcritical and we are interested in ages of old families. Although we only state the results in the exponential case, they also hold in the general case and are proved in [27, Chapter 3] and [29]. However, to obtain the general results in Model I, additional assumptions about the lifespan measure Λ are required, which are easily satisfied in the exponential case (for instance, we need the existence of a negative root of ψ_p, which, with easy computations, is $b(1-p) - d$ in the exponential case).

In the first result, which is a result in expectation, we show that in both models, the ages are of order of magnitude

$$c_t := \frac{r}{r - r_\star} t. \tag{4.1}$$

Proposition 4.1 (see [27, 29]). *One supposes that $Z_\star$ is subcritical. For $a \in \mathbb{R}$, one has*

$$\mathbb{E}[O_t(a + c_t)] \xrightarrow[t \to \infty]{} \frac{|r_\star|}{d_\star} e^{-(r-r_\star)a}. \tag{4.2}$$

This result is a consequence of the expected spectrum formula (3.2), summed over $i \geq 1$ and integrated on $(a + c_t, t)$. We also obtain a more precise result about the convergence in distribution of $O_t(a + c_t)$ as $t \to +\infty$.

Proposition 4.2 (see [27, 29]). *With the same assumptions as in Proposition 4.1, for $a \in \mathbb{R}$, conditional on the survival event, as $t \to \infty$, $O_t(a + c_t)$ converges in distribution to an r.v. $\mathcal{O}$, distributed as a mixed Poisson r.v. whose parameter of mixture is*

$$\frac{b}{r} \frac{|r_\star|}{d_\star} e^{-(r-r_\star)a} E, \tag{4.3}$$

where E is an exponential r.v. with mean 1. Equivalently, $\mathcal{O}$ is geometric on $\{0, 1, \ldots\}$ with success probability

$$\frac{1}{1 + (b/r)(|r_\star|/d_\star)e^{-(r-r_\star)a}}. \tag{4.4}$$

The proof of this proposition in the general case and for Model I, given in [27], follows arguments of Taïb in [25] and uses the notion of CMJ processes counted with *time-dependent* random characteristics developed by Jagers and Nerman in [26, 34]. The difference with (3.9) is that here the characteristics are allowed to depend on time. This theory provides convergences in distribution, as $t \to +\infty$, of quantities of the form

$$Z_t^{\chi^t} := \sum_x \chi_x^t(t - \sigma_x) \mathbf{1}_{\{\sigma_x \leq t\}}, \tag{4.5}$$

under technical conditions about the family of characteristics $(\chi^t(\cdot), t \geq 0)$. The proof of Proposition 4.2 for Model II is given in [29] and does not make use of random characteristics.

The last result deals with the convergence in distribution of the sequence of the ranked ages of extant families. Let $\mathcal{M}(\mathbb{R})$ be the set of nonnegative σ-finite measures on $\mathbb{R}$ and finite on $\mathbb{R}^+$, equipped with the *left-vague topology* induced by the maps $\nu \mapsto \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x)$ for all bounded continuous functions f such that there exists $x_0 \in \mathbb{R}$ satisfying for all $x \leq x_0$, $f(x) = 0$.

Theorem 4.3 (see [27, 29]). *With the same assumptions as previously, let X_t be the point process defined by*

$$X_t(\mathrm{d}x) := \sum_{k \geq 1} \delta_{A_t^k - c_t}(\mathrm{d}x), \tag{4.6}$$

where $A_t^1 \geq A_t^2 \geq \cdots$ is the decreasing sequence of ages of alive families at t. Then, conditional on the survival event, X_t converges as $t \to \infty$ in $\mathcal{M}(\mathbb{R})$ equipped with the left-vague topology to a mixed Poisson point process with an intensity measure

$$\frac{b}{r} \frac{|r_\star|}{d_\star} E\,(r - r_\star)e^{-(r-r_\star)x}\mathrm{d}x, \tag{4.7}$$

where E is an exponential r.v. with mean 1.

4.2. Sizes of the Largest Families in the Subcritical Case of Model II

In this paragraph, we still suppose that the clonal process is subcritical and we are interested in similar results as those of Section 4.1 about the sizes of the largest families. The aim is to find a number x_t such that $L_t(x_t)$ converges to a finite and positive limit as $t \to +\infty$.

Concerning Model I, this problem is still open. On the contrary, it is possible to obtain in Model II the sizes of the largest families. In [29], they are given for any life length distribution but to simplify the results, we only state them in the exponential case. The following result is a consequence of (3.3b) applied with $a = t$ and summed over $i \geq x_t + c$. Recall that the clonal process is assumed to be subcritical, so that $\theta > r$.

Proposition 4.4 (see [29]). *One sets*

$$x_t := \frac{rt - (\theta/(\theta - r))\log t}{-\log(b/(\theta + d))}. \tag{4.8}$$

Then, for $c \in \mathbb{R}$,

$$\mathbb{E}[L_t(x_t + c)] \underset{t \to +\infty}{\sim} A(b, d, \theta)\left(\frac{b}{\theta + d}\right)^{c - 1 + \{-x_t - c\}}, \tag{4.9}$$

where $\{x\}$ denotes the fractional part of a real number x and where $A(b, d, \theta)$ is an explicit constant that only depends on b, d, and θ.

For $t \geq 0$ and $k \geq 1$, we denote by S_t^k the size of the kth largest family in the whole population at time t. Let

$$X_t := \sum_{k \geq 1} \delta_{S_t^k - x_t}, \tag{4.10}$$

be the point measure of the renormalized sizes of the population. To get rid of fractional parts, the following theorem gives the convergence in distribution of $L_t(x_t + c)$ and X_t along a subsequence. More precisely, for $n \geq 1$, let t_n be such that $x_{t_n} = n$; this equation has a unique solution for any n greater than some integer n_0. It satisfies

$$t_n \underset{n \to +\infty}{\sim} \frac{\theta - r}{\theta} \log\left(\frac{\theta + d}{b}\right) n. \tag{4.11}$$

We now state the convergence of the sequence $(X_{t_n}, n \geq n_0)$.

Theorem 4.5 (see [29]). *Conditional on the survival event, the sequence $(X_{t_n}, n \geq n_0)$ of point processes on $\mathbb{Z}$ converges as $n \to +\infty$ on the set $\mathcal{M}(\mathbb{R})$ equipped with the left-vague topology to a mixed Poisson point measure on $\mathbb{Z}$ with an intensity measure*

$$A(b, d, \theta) \frac{b - d}{b}\left(1 - \frac{b}{\theta + d}\right) E \sum_{k \in \mathbb{Z}} \left(\frac{b}{\theta + d}\right)^{k-1} \delta_k, \tag{4.12}$$

where the mixture coefficient E is an exponential r.v. with mean 1.

4.3. Other Results

4.3.1. Critical Case in Model I

The case of a critical clonal process Z_p for a general supercritical splitting tree is treated in Section 3.5.1 of [27] where the counterparts of Propositions 4.1 and 4.2 and Theorem 4.3 are proved.

If $(1 - p)m = 1$, provided that the second moment $\sigma^2 := \int_0^\infty \Lambda(du)u^2$ is finite and that a condition about the tail distribution of Λ holds, ages of oldest families are of order

$$c_t = t - \frac{\log t}{r}. \tag{4.13}$$

Notice that these conditions about Λ are trivially satisfied in the exponential case. These results were also proved in [25, Chapter 4] for any CMJ process Z, that is, with a birth point process as general as possible, but in that case, limits were not explicit.

Similarly to the subcritical case, the problem of sizes of the largest families is still open. Nevertheless, we can state the following conjecture about their order of magnitude.

Conjecture 4.6. *If*

$$x_t := \frac{2}{\sigma^2}\left(rt^2 - t\log t\right), \tag{4.14}$$

as $t \to \infty$, on the survival event, $L_t(x_t)$ converges in distribution to a nondegenerate geometric r.v.

4.3.2. Critical Case in Model II

The general case when Z_θ is critical ($\theta = r$) can be found in Sections 3.4 and 5 in [29]. For both ages and sizes, the counterparts of the results of Sections 4.1 and 4.2 hold.

As in Model I, ages of the oldest families are of order $c_t = t - \log t/r$. Moreover, sizes of the largest ones are of order

$$x_t = \frac{r^2}{4\psi'(r)}\left(t - \frac{\log t}{2r}\right)^2, \tag{4.15}$$

and the point measure,

$$\sum_{k \geq 1} \delta_{\sqrt{S_t^k} - \sqrt{x_t}}, \tag{4.16}$$

converges to a mixed Poisson measure as $t \to +\infty$ but contrary to Theorem 4.5, it does not only hold along a subsequence.

4.3.3. Sizes of the Largest Families in Supercritical Cases

In [27, Chapter 3], general splitting trees in Model I are considered. When the clonal process Z_p is supercritical, that is, when $(1-p)m > 1$, a result about the sizes of the largest families is proved. First notice that, as in (2.9),

$$\mathbb{P}(Z_p(t) \longrightarrow 0) = 1 - \frac{r_p}{(1-p)b}, \tag{4.17}$$

and on $\{Z_p(t) \to 0\}^c$, $e^{-r_p t}Z_p(t)$ a.s. converges as $t \to \infty$ to an exponential random variable. Hence, the sizes of alive families at time t must be of order $e^{r_p t}$ as $t \to +\infty$. We proved this in [27] by showing that $\mathbb{E}[L_t(e^{r_p t})]$ converges as $t \to +\infty$ to an explicit limit.

Notice that we cannot obtain similar results to Proposition 4.2 and Theorem 4.3 concerning the convergence in distribution of $L_t(e^{(b(1-p)-d)t})$ and the convergence of the associated point measure of the decreasing sequence of family sizes.

In [25], Taïb considers a more general model than our Model I; mutation mechanism is the same but Z_p can be any supercritical CMJ process. In his Theorem 4.6, by using a time-dependent characteristic argument, he proved the convergence in distribution of $L_t(e^{r_p t})$ (to a nonexplicit random variable). However, we have doubts about the application of Theorem A.7, since the technical requirements of this theorem do not seem to hold in his case. These technical requirements are neither proved to hold in [25] nor in [34].

In Model II, for a general supercritical splitting tree, if Z_θ is supercritical, that is, $r > \theta$, $Z_\theta(t)$ asymptotically grows like $e^{(r-\theta)t}$. In [29, Proposition 3.2], it is proved that $\mathbb{E}[L_t(e^{(r-\theta)t})]$ converges as $t \to +\infty$, but we were unable to obtain any convergence in distribution in that case.

Acknowledgments

References

[1] J. Geiger, "Size-biased and conditioned random splitting trees," *Stochastic Processes and their Applications*, vol. 65, no. 2, pp. 187–207, 1996.

[2] J. Geiger and G. Kersting, "Depth-first search of random trees, and Poisson point processes," in *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*, vol. 84 of *IMA Volumes in Mathematics and its Applications*, pp. 111–126, Springer, New York, NY, USA, 1997.

[3] A. Lambert, "The contour of splitting trees is a Lévy process," *The Annals of Probability*, vol. 38, no. 1, pp. 348–395, 2010.

[4] P. Jagers, *Branching Processes with Biological Applications*, Wiley-Interscience, London, UK, 1975, Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistic.

[5] M. A. Nowak, F. Michor, and Y. Iwasa, "The linear process of somatic evolution," *Proceedings of the National Academy of Sciences*, vol. 100, no. 25, pp. 14966–14969, 2003.

[6] Y. Iwasa, M. A. Nowak, and F. Michor, "Evolution of resistance during clonal expansion," *Genetics*, vol. 172, no. 4, pp. 2557–2566, 2006.

[7] S. Sagitov and M. C. Serra, "Multitype Bienaymé-Galton-Watson processes escaping extinction," *Advances in Applied Probability*, vol. 41, no. 1, pp. 225–246, 2009.

[8] R. Durrett and S. Moseley, "Evolution of resistance and progression to disease during clonal expansion of cancer," *Theoretical Population Biology*, vol. 77, no. 1, pp. 42–48, 2010.

[9] R. Durrett and J. Mayberry, "Traveling waves of selective sweeps," *The Annals of Applied Probability*, vol. 21, no. 2, pp. 699–744, 2011.

[10] K. Danesh, R. Durrett, L. J. Havrilesky, and E. Myers, "A branching process model of ovarian cancer," *Journal of Theoretical Biology*, vol. 314, pp. 10–15, 2012.

[11] D. G. Kendall, "Birth-and-death processes, and the theory of carcinogenesis," *Biometrika*, vol. 47, pp. 13–21, 1960.

[12] T. Stadler, "Inferring epidemiological parameters based on allele frequencies," *Genetics*, vol. 188, no. 3, pp. 663–672, 2011.

[13] A. Lambert and P. Trapman, "Splitting trees stopped when the first clock rings and Vervaat's transformation," *Journal of Applied Probability*. In press.

[14] J. Gani and G. F. Yeo, "Some birth-death and mutation models for phage reproduction," *Journal of Applied Probability*, vol. 2, pp. 150–161, 1965.

[15] S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press, Princeton, NJ, USA, 2001.

[16] B. Haegeman and R. S. Etienne, "Relaxing the zero-sum assumption in neutral biodiversity theory," *Journal of Theoretical Biology*, vol. 252, no. 2, pp. 288–294, 2008.

[17] A. Lambert, "Species abundance distributions in neutral models with immigration or mutation and general lifetimes," *Journal of Mathematical Biology*, vol. 63, no. 1, pp. 57–72, 2011.

[18] W. J. Ewens, "The sampling theory of selectively neutral alleles," *Theoretical Population Biology*, vol. 3, pp. 87–112, 1972, erratum, ibid. vol. 3, p. 240, 1972; erratum, ibid. vol. 3, p. 376, 1972.

[19] R. C. Griffiths and A. G. Pakes, "An infinite-alleles version of the simple branching process," *Advances in Applied Probability*, vol. 20, no. 3, pp. 489–524, 1988.

[20] J. Bertoin, "The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations," *The Annals of Probability*, vol. 37, no. 4, pp. 1502–1523, 2009.

[21] J. Bertoin, "A limit theorem for trees of alleles in branching processes with rare neutral mutations," *Stochastic Processes and their Applications*, vol. 120, no. 5, pp. 678–697, 2010.

[22] A. G. Pakes, "An infinite alleles version of the Markov branching process," *Australian Mathematical Society A*, vol. 46, no. 1, pp. 146–169, 1989.

[23] Y. E. Maruvka, N. M. Shnerb, and D. A. Kessler, "Universal features of surname distribution in a subsample of a growing population," *Journal of Theoretical Biology*, vol. 262, no. 2, pp. 245–256, 2010.

[24] Y. E. Maruvka, D. A. Kessler, and N. M. Shnerb, "The birth-death-mutation process: a new paradigm for fat tailed distributions," *PLoS ONE*, vol. 6, no. 11, article e26480, 2011.

[25] Z. Taïb, *Branching Processes and Neutral Evolution*, vol. 93 of *Lecture Notes in Biomathematics*, Springer, Berlin, Germany, 1992.

[26] P. Jagers and O. Nerman, "The growth and composition of branching populations," *Advances in Applied Probability*, vol. 16, no. 2, pp. 221–259, 1984.

[27] M. Richard, *Arbres, Processus de branchement non markoviens et Processus de Lvy [Ph.D. thesis]*, UPMC, Paris, France, 2011.

[28] N. Champagnat and A. Lambert, "Splitting trees with neutral Poissonian mutations I: small families," *Stochastic Processes and their Applications*, vol. 122, no. 3, pp. 1003–1033, 2012.

[29] N. Champagnat and A. Lambert, "Splitting trees with neutral Poissonian mutations II:," *Largest and Oldest families*. In press, http://arxiv.org/abs/1108.4812.

[30] J. Bertoin, *Lévy Processes*, vol. 121 of *Cambridge Tracts in Mathematics*, Cambridge University Press, Cambridge, UK, 1996.

[31] O. Nerman, "On the convergence of supercritical general (C-M-J) branching processes," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, no. 3, pp. 365–395, 1981.

[32] M. Richard, "Limit theorems for supercritical age-dependent branching processes with neutral immigration," *Advances in Applied Probability*, vol. 43, no. 1, pp. 276–300, 2011.

[33] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, vol. 55 of *National Bureau of Standards Applied Mathematics Series*, Superintendent of Documents, U.S. Government Printing Office, Washington, DC, USA, 1964.

[34] P. Jagers and O. Nerman, "Limit theorems for sums determined by branching and other exponentially growing processes," *Stochastic Processes and their Applications*, vol. 17, no. 1, pp. 47–71, 1984.