# On branching process with rare neutral mutation

AIRAM BLANCAS[*] and VÍCTOR RIVERO[**]

*Centro de Investigación en Matemáticas (CIMAT A.C.). Calle Jalisco s/n, Col. Valenciana, 36240 Guanajuato, Gto. México. E-mail: [*]airam@cimat.mx; [**]rivero@cimat.mx*

In this paper, we study the genealogical structure of a Galton–Watson process with neutral mutations. Namely, we extend in two directions the asymptotic results obtained in Bertoin [*Stochastic Process. Appl.* **120** (2010) 678–697]. In the critical case, we construct the version of the model in Bertoin [*Stochastic Process. Appl.* **120** (2010) 678–697], conditioned not to be extinct. We establish a version of the limit theorems in Bertoin [*Stochastic Process. Appl.* **120** (2010) 678–697], when the reproduction law has an infinite variance and it is in the domain of attraction of an $\alpha$-stable distribution, both for the unconditioned process and for the process conditioned to nonextinction. In the latter case, we obtain the convergence (after re-normalization) of the allelic sub-populations towards a tree indexed CSBP with immigration.

*Keywords:* branching process; domain of attraction of $\alpha$-stable laws; neutral mutations; Q-processes; regular variation

## 1. Model description and main results

A Galton–Watson process models a population where at every generation each individual reproduces according to the same distribution, independently of the others and then dies. For background about branching processes, we refer to [1,13] and [15]. A number of variants, involving different types of conditioning and limit theorems, are core of branching processes theory. For instance, when the process dies with probability one [18] proved that the distribution of the process conditioned to nonextinction exists, under some assumptions on the moments of the reproduction law. The proof was simplified and the moment assumptions removed in [9] and in [16]. More generally, [14] introduced the $Q$-process.

As a further extension of the Galton–Watson model, [3] studied the so called Galton–Watson process with neutral mutations. This emerges assuming that the mutations modify the genotype of individuals but not the dynamic of the population, which is modeled by a standard Galton–Watson. Since mutations appear in the ancestral lines of the population, each individual begets children that do not necessarily inherit its genetic type (allele). In addition, we suppose that the population has infinity alleles, that is, each mutation event originates a different allele. We denote the size of a typical family by $\xi^{(+)} := \xi^{(c)} + \xi^{(m)}$, where $\xi^{(c)}$, $\xi^{(m)}$ are nonnegative random variables which determine respectively, the number of clones and mutants children of a typical individual. We exclude the degenerate cases $\xi^{(c)} \equiv 0$ or $\xi^{(m)} \equiv 0$.

In [3], established asymptotic features on the genealogy of allelic sub-families in a Galton–Watson process with neutral mutations. In his development, the genealogy of the population is described by a planar rooted tree where the mutations are represented by marks in the edges
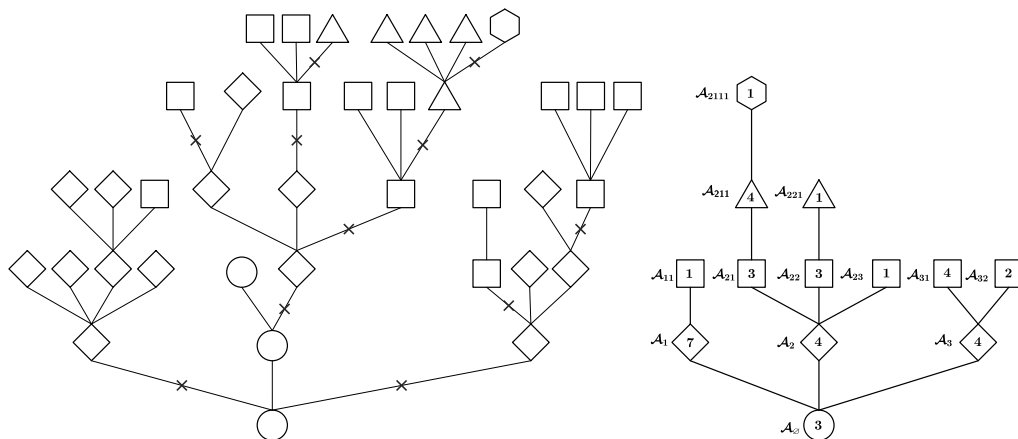
**Figure 1.** The tree on the left, illustrates a Galton Watson process with neutral mutation, the allele type of an individual is represented by the form of its vertex. On the right we have its tree of alleles, from this we can deduce for instance, that $M_0 = 1$, $T_0 = 3$, $M_1 = 3$, $T_1 = 15$.

between parents and mutant children. See Figure 1 (left). The vertices with $n$ marks in their ancestral line are associated with the called *n-type individuals*. Those individuals with the $n$th mark in the edge between them and their parents are known as *mutants of the n-type*. We denote by the $T_n$ the total population of individuals of the $n$th type and by $M_n$ the total number of mutants of $n$th type. By convention, the individual in the generation 0th, the ancestors, are consider as mutants of the 0th type, that is $M_0 = a$, $\mathbb{P}_a$-a.s.

It is well know that the branching property is the most basic property in the analysis of Galton–Watson processes. Then it is natural to expect a branching property for the Galton–Watson process with neutral mutations. Indeed it has the *general branching property*, which states that conditionally on the set of children of a stopping line, the families that those beget are independent copies of the initial tree. The concept of *stopping line* was introduced by [5], where the reader is referred for the formal definition. Roughly, a *line* is a family of edges such that every branch from the root contains at most one edge in that family. A stopping line is a random line such that the event "an edge is in the line", only depends on the marks found on their ancestral line. In particular, the set of edges connecting the mutants of the $n$ type with their parents is a stopping line. Then for every $n$, each mutant of the $n$-type begets a sub-family which is independent of the others and has the same distribution as the original tree. Another important consequence of the general branching property is given in the following lemma.

**Lemma 1.1 ([3], Lemma 1).** *Under $\mathbb{P}_a$, $\{M_n : n \in \mathbb{Z}_+\}$ is a Galton–Watson process with reproduction law $\mathbb{P}_1(M_1 \in \cdot)$. More generally, $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, is a Markov chain, with transition probabilities*

$$\mathbb{P}_a(T_n = k, M_{n+1} = l | T_{n-1} = i, M_n = j) = \mathbb{P}_j(T_0 = k, M_1 = l), \qquad (1.1)$$

*for all $i, j, k, l \in \mathbb{Z}_+$ and $j \leq k$.*

**Remark 1.1.** Since the mutants of the $n$th type are also individuals of the $n$th type, the transition probabilities in (1.1) are zero when $j > k$.

Let $\{P^n_{(i,j),(k,l)} : i, j, k, l \in \mathbb{Z}_+\}$ denotes the $n$-step transition probabilities of the Markov chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, that is

$$P^n_{(i,j),(k,l)} = \mathbb{P}_a(T_{r+n} = k, M_{r+n+1} = l | T_r = i, M_{r+1} = j), \qquad i, j, k, l \in \mathbb{Z}_+, n \in \mathbb{N}. \quad (1.2)$$

Then $P^n_{(i,j),(k,l)}$ depends only on the mutants coordinate. Actually, it is not difficult to prove using induction, that the following identity holds

$$P^n_{(i,j),(k,l)} = \sum_{j_{n-1}=1}^{\infty} \mathbf{P}^{n-1}_{(j,j_{n-1})} \mathbb{P}_{j_{n-1}}(T_0 = k, M_1 = l), \quad (1.3)$$

where $j_0 = j$ and $\{\mathbf{P}^n_{(i,j)} : i, j \in \mathbb{Z}_+\}$ denotes the $n$-step transition probabilities of $\{M_n : n \in \mathbb{Z}_+\}$.

We now introduce the space of finite sequence of integers

$$\mathbb{U} := \bigcup_{k \in \mathbb{Z}_+} \mathbb{N}^k,$$

where $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}^0 = \{\varnothing\}$. We recall that this set gives us the label of the vertices in Ulam–Harris–Neveu tree. More precisely, the root corresponds to $\{\varnothing\}$, one vertex at level $k > 0$ is $u = (u_1, \ldots, u_k)$ and $uj = (u_1, \ldots, u_k, j)$ represents its $j$th children. The level of the vertex $u$ is denoted by $|u|$. We shall consider the *tree of alleles* $\mathcal{A} := \{\mathcal{A}_u : u \in \mathbb{U}\}$ constructed recursively in [3]. Define $\mathcal{A}_\varnothing = T_0$ and $\mathcal{A}_{uj}$ as the size of the $j$th allelic sub-population of the type $|u| + 1$ which descend from the allelic sub-family indexed by the vertex $u$. In the case of ties, sub-families are ordered by convention uniformly at random. See Figure 1 (right). A further consequence of the general branching property is that, the tree of alleles enjoys a branching property. To provide a formal statement, we first define the degree of the tree of alleles $\mathcal{A}$ at some vertex $u \in \mathbb{U}$ as

$$d_u := \max\{j \geq 1 : \mathcal{A}_{uj} > 0\},$$

where we agree that $\max \varnothing = 0$. The notation $(d_u \downarrow)$ means that the $d_u$-tuple has been rearranged in the decreasing order of the first coordinate, by convention, in the case of ties the coordinates are ranked uniformly at random.

**Lemma 1.2 ([3], Lemma 2).** *For any integers $a \geq 1$ and $k \geq 0$, under $\mathbb{P}_a$ conditionally on $\{(\mathcal{A}_u, d_u) : |u| \leq k\}$, for each vertex $u$ at level $k$ with $\mathcal{A}_u > 0$, the family of variables $\{(\mathcal{A}_{uj}, d_{uj}) : 1 \leq j \leq d_u\}$ are independent with distribution $(T_0, M_1)^{(d_u \downarrow)}$ under $\mathbb{P}_1$.*

It is important to observe the following identities

$$T_k = \sum_{|u|=k} \mathcal{A}_u \quad \text{and} \quad M_{k+1} = \sum_{|u|=k} d_u. \quad (1.4)$$

Hence, given a population with neutral mutations, $\{(\mathcal{A}_u, d_u) : |u| \leq k\}$ records the genealogy of allelic sub-families together with their sizes. Also, the size of their generations is a Galton–Watson process.

The first goal in this paper is to construct the version of the chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, conditioned on nonextinction of mutants, hence we are interested in the situation where

$$T = \inf\{n \geq 1 : M_n = 0\} < \infty, \tag{1.5}$$

with a strictly positive probability. According to Corollary 1 of [3], this occurs when $\mathbb{E}(\xi^{(c)}) < 1$ and $\mathbb{E}(\xi^{(+)}) \leq 1$. This implies that the Galton Watson process of mutants $\{M_n : n \in \mathbb{Z}_+\}$ is critical or subcritical, that is $m := \mathbb{E}_1(M_1) \leq 1$.

We can now state our first theorem.

**Theorem 1.3.** *Let $a \in \mathbb{N}$ and $\{\mathcal{F}_n : n \in \mathbb{Z}_+\}$ be the natural filtration of the process $\{(T_{n-1}, M_n) : n \in \mathbb{N}\}$. Then, $\mathbb{P}_a^{\uparrow}$ is locally absolutely continuous with respect to $\mathbb{P}_a$ with Radom–Nikodim martingale density*

$$Y_n = \frac{M_n q^{M_n - a}}{(f'(q))^n} \mathbf{1}_{\{n < T\}},$$

*where $f(y) = \mathbb{E}_1(y^{M_1})$ and $q = \mathbb{P}_1(0 < T < \infty)$, that is*

$$d\mathbb{P}_a^{\uparrow}|_{\mathcal{F}_n} = \frac{Y_n}{a} d\mathbb{P}_a \Big|_{\mathcal{F}_n}, \qquad n \in \mathbb{N}.$$

*Furthermore, $\mathbb{P}_a^{\uparrow}$ is the law of a Markov chain $\{(T_n^{\uparrow}, M_{n+1}^{\uparrow}) : n \in \mathbb{Z}_+\}$ with $n$-step transition probabilities,*

$$Q_{(i,j),(k,l)}^n = \frac{lq^{l-j}}{j(f'(q))^n} P_{(i,j),(k,l)}^n, \qquad j, l \geq 1, \tag{1.6}$$

*where $\{P_{(i,j),(k,l)}^n : i, j, k, l \in \mathbb{Z}_+\}$ denotes the $n$-step transition probabilities of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$.*

We next ensure that the process defined in the above theorem is distributed as $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ conditionally on nonextinction of mutants.

**Theorem 1.4.** *Suppose that $\mathbb{E}(\xi^{(c)}) < 1$ and $\mathbb{E}(\xi^{(+)}) \leq 1$.*

(i) *Let $a, n \in \mathbb{N}$ with $n$ fixed. The conditional law of the process $\{(T_k, M_{k+1}) : 0 \leq k \leq n - 1\}$ under $\mathbb{P}_a(\cdot | n + k < T < \infty)$ converges, as $k \to \infty$, towards the probability measure $\mathbb{P}_a^{\uparrow}$, in the sense that for any $n$*

$$\lim_{k \to \infty} \mathbb{P}_a(A | n + k < T < \infty) = \mathbb{P}_a^{\uparrow}(A), \qquad \forall A \in \mathcal{F}_n. \tag{1.7}$$

(ii) *The Yaglom limit*

$$\lim_{n \to \infty} \mathbb{P}(T_{n-1} = i, M_n = j | n < T < \infty),$$

*exists and has a generating function $\widehat{\varphi}(x, y)$ such that for all $n \in \mathbb{N}$,*

$$m^n \widehat{\varphi}(x, y) = \widehat{f}(\varphi_n(x, y)) - \widehat{f}(\varphi_n(x, 0)), \qquad x, y \in [0, 1]. \tag{1.8}$$

The proof of the above results is based on classical methods and this due to the fact that the generating function of $(T_n, M_{n+1})$ can be written in terms of that of $M_n$, as it is established in Section 2.

We now turn to analyze the asymptotic behavior of the tree of alleles. In this purpose we will consider for ever $n \in \mathbb{N}$, a Galton–Watson process $\{Z_k^{(+n)} : k \in \mathbb{Z}_+\}$ such that the reproduction law

$$\pi_k^+ = \mathbb{P}(\xi^{(c)} + \xi^{(m)} = k), \qquad k \in \mathbb{Z}_+,$$

is critical (with mean one) and has a finite variance $\sigma^2$. We assume that each child is a clone of her mother with probability $1 - p(n)$ and a mutant with probability $p(n)$, so the joint law of $(\xi^{(c)}, \xi^{(m)})$, denoted by $\pi = \{\pi_{k,l} : k, l \in \mathbb{Z}_+\}$, that is,

$$\pi_{k,l} = \mathbb{P}(\xi^{(c)} = k, \xi^{(m)} = l), \qquad k, l \in \mathbb{Z}_+, \tag{1.9}$$

satisfies

$$\pi_{k,l} = \pi_{k+l}^+ \binom{k+l}{k} (1 - p(n))^k p(n)^l, \qquad k, l \in \mathbb{Z}_+. \tag{1.10}$$

As usual, in the remainder of this paper the relation $f \sim g$ refers to $\lim_{x \to \infty} f(x)/g(x) = 1$.

In the paper [3], it has been assumed that the number of ancestors and mutation rate respectively have the following behavior

$$a(n) \sim nx \quad \text{and} \quad p(n) \sim cn^{-1}, \qquad \text{as } n \to \infty; \tag{1.11}$$

where $c, x$ are some positive constants. In this setting, it has been proved that

$$\mathcal{L}\big(\{(n^{-2}T_k, n^{-1}M_{k+1}) : k \in \mathbb{Z}_+\}, \mathbb{P}_{a(n)}^{p(n)}\big) \implies \{(Z_{k+1}, cZ_{k+1}) : k \in \mathbb{Z}_+\}, \tag{1.12}$$

where $\{Z_k : k \in \mathbb{Z}_+\}$ is a (discrete time) continuous state branching process, in short CSBP, with reproduction measure

$$\nu(dy) = \frac{c}{\sqrt{2\pi\sigma^2 y^3}} \exp\left(-\frac{c^2 y}{2\sigma^2}\right) dy, \qquad y > 0, \tag{1.13}$$

and initial population of size $x/c$. Here and all through the paper the symbol $\implies$ will denote the weak convergence of finite dimensional distributions.

From [13], we know that the transition probabilities of any CSBP process $\{Y_k : k \in \mathbb{Z}_+\}$ with reproduction measure $\vartheta$ are characterized as follows:

$$\mathbb{E}(e^{-\lambda Y_{k+1}} | Y_k = y) = e^{-y\kappa(\lambda)}, \qquad k \in \mathbb{Z}_+, \lambda, y \geq 0, \tag{1.14}$$

where $\kappa$ is the cumulant of a subordinator with Lévy measure $\vartheta$, so that $\int_{(0,\infty)}(1 \wedge y)\vartheta(dy) < \infty$. In this work, we will consider subordinators without drift, thus

$$\kappa(\lambda) = \int_{(0,\infty)} \left(1 - e^{-\lambda y}\right)\vartheta(dy), \qquad \lambda > 0. \tag{1.15}$$

Applying successively the property (1.14), we obtain

$$\mathbb{E}_x\left(e^{-s_1 Y_1 \cdots - s_k Y_k}\right) = e^{-x\kappa(l_{k-1}(s_1))}, \qquad s_i \geq 0, i = 1, 2, \ldots, k, \tag{1.16}$$

where $l$ is defined by induction as follows: $l_0(s) = s$,

$$l_i(s_{n-i}) = s_{n-i} + \kappa\left(l_{i-1}(s_{n-i+1})\right), \qquad i \in \mathbb{N}. \tag{1.17}$$

Combining the convergence (1.12) and the identity (1.14), together with the Lévy–Itô decomposition of a subordinator, one could infer that conditionally on $n^{-2}T_k \sim y$ the sequence of the sizes of the sub-population carrying a same allele of the $(k+1)$-type and normalized by a factor $n^{-2}$ should converge in distribution to the sequence of atoms of a Poisson random measure on $\mathbb{R}_+$ with intensity given in (1.13). Thus, the limit of a sequence of tree of alleles can be defined as follows.

**Definition 1.5 ([3], Definition 1).** Fix $x > 0$ and $\vartheta$ a measure on $(0,\infty)$ with $\int_{(0,\infty)}(1 \wedge y)\vartheta(dy) < \infty$. A tree-indexed CSBP with reproduction measure $\vartheta$ and initial population of size $x$, is a process $\{\mathcal{Y}_u : u \in \mathbb{U}\}$ with values in $\mathbb{R}_+$ and indexed by the universal tree, whose distribution is characterized by induction on the levels as follows:

  (i) $\mathcal{Y}_\varnothing = x$ a.s.;
  (ii) for every $k \in \mathbb{Z}_+$ conditionally on $\{\mathcal{Y}_v : v \in \mathbb{U}, |v| \leq k\}$, the sequences $\{\mathcal{Y}_{uj} : j \in \mathbb{N}\}$ for the vertices $u \in \mathbb{U}$ at generation $|u| = k$ are independent, and each sequence is distributed as the family of the atoms of a Poisson random measure on $(0,\infty)$ with intensity $\mathcal{Y}_u\vartheta$, where the atoms are repeated according to their multiplicity, ranked in the decreasing order, and completed by an infinite sequence of 0 if the Poisson measure is finite.

Roughly, the tree-indexed CSBP is a process indexed according to the Ulam–Harris–Neveu tree such that the vertices $u \in \mathbb{U}$ at level $|u| = k > 1$ represent the sizes of the sub-populations at generation $k$ in the CSBP $\{Y_k : k \in \mathbb{Z}_+\}$, which descent from the same parent at generation $k - 1$.

Now it can be seen that the convergence (1.12) can be written as follows

$$\mathcal{L}\left(\left\{\left(n^{-2}\mathcal{A}_u, n^{-1}d_u\right) : k \in \mathbb{Z}_+\right\}, \mathbb{P}_{a(n)}^{p(n)}\right) \implies \left\{(\mathcal{Z}_u, c\mathcal{Z}_u) : u \in \mathbb{U}\right\}, \tag{1.18}$$

where $\{\mathcal{Z}_u : u \in \mathbb{U}\}$ is a tree-indexed CSBP with reproduction $\nu$ given in (1.13) and random initial population of size $x/c$. We recall under some assumptions. It is $d_u$ denotes the outer degree at the vertex $u \in \mathbb{U}$ in the tree of alleles. This latter convergence is the main result of [3]. It uses an argument on convergence of triangular arrays, described in page 690 therein, that can be extended to a more general context, see, for example, the forthcoming Lemma 4.6.

One main goal of this paper is to investigate the asymptotic behavior of the population in the same sense of [3] but on a complimentary class of reproduction laws. Instead of assuming that it has finite variance as in Bertoin's paper, we suppose that there exists $\alpha \in (1, 2)$ such that,

$$\bar{\pi}^+(j) := \mathbb{P}\big(\xi^{(+)} > j\big) \in RV_\infty^{-\alpha}, \qquad j \in \mathbb{Z}_+, \tag{1.19}$$

where $RV_\infty^{-\alpha}$ denotes the class of functions which are regularly varying at $\infty$ with index $-\alpha$, see Chapter I in [4] for background. Note that the case $\alpha \in (0, 1)$ is excluded because it contradicts the assumption that $\pi^+$ is critical.

In order to extend the main result of [3] to our setting, we prove that there exists a regularly varying function $r$ with index $\alpha$ such that

$$r(n)\mathbb{P}\big(\xi^+ > ny\big) \xrightarrow[n\to\infty]{} c_\alpha y^{-\alpha}, \qquad \forall y > 0, \tag{1.20}$$

where $c_\alpha = 1/\Gamma(3-\alpha)$. The proof of this fact is given in Proposition 4.1. Moreover, the following behavior will be assumed instead of the hypothesis (1.11),

$$a(n) \sim xr(n)p(n) \quad \text{and} \quad p(n) \sim cn^{-1} \qquad \text{as } n \to \infty. \tag{1.21}$$

The result below extends to our setting the main result in [3].

**Theorem 1.6.** *If* (1.19) *and* (1.21) *holds, then the following convergence holds in the sense of finite dimensional distributions*

$$\mathcal{L}\big(\big\{\big((r(n))^{-1}\mathcal{A}_u, \big(r(n)p(n)\big)^{-1}d_u\big) : u \in \mathbb{U}\big\}, \mathbb{P}_{a(n)}^{p(n)}\big) \implies \big\{\big(\mathcal{Z}_u^{1/\alpha}, \mathcal{Z}_u^{1/\alpha}\big) : u \in \mathbb{U}\big\},$$

*where* $\{\mathcal{Z}_u^{1/\alpha} : u \in \mathbb{U}\}$ *is a tree-indexed CSBP with reproduction measure*

$$\nu^\alpha(dy) = c_\alpha' y^{-1-1/\alpha}\, dy, \qquad y > 0, \alpha \in (1, 2), \tag{1.22}$$

*where* $c_\alpha' = \alpha^{-1}/\Gamma(1-\alpha^{-1})$.

Finally, we establish the convergence of the finite dimensional distributions of the rescaled chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, conditioned to nonextinction of mutants, towards a continuous state branching process with immigration in discrete time.

**Theorem 1.7.** *If the reproduction law is critical, there exist sequences $b_1(n)$ and $b_2(n)$ such that the following joint convergence in the sense of finite dimensional distributions holds*:

$$\mathcal{L}\big(\big\{\big(b_1(n)T_{k-1}, b_2(n)M_k\big) : k \in \mathbb{Z}_+\big\}, \mathbb{P}_{a(n)}^{p(n)\uparrow}\big) \implies \big\{(Y_k, \beta Y_k) : k \in \mathbb{Z}_+\big\},$$

*where* $\{Y_k : k \in \mathbb{Z}_+\}$ *is a CSBP with immigration, which is characterized by the following conditions*:

  (i) *if the reproduction law has finite variance $\sigma^2$ and* (1.11) *holds, then its reproduction measure is given by* (1.13) *and the immigration measure is $z\nu(dz)$ and $\beta = c$; moreover $b_1(n) = n^{-2}$ and $b_2(n) = n^{-1}$;*

(ii) *if the assumptions* (1.19) *and* (1.21) *hold, the reproduction measure is* $\nu^\alpha(dz)$ *as defined in* (1.22), *the immigration measure is* $z\nu^\alpha(dz)$ *and* $\beta = 1$; *the normalizing constants are given by* $b_1(n) = (r(n)p(n))^{-1}$ *and* $b_2(n) = (r(n))^{-1}$.

The remainder of the paper is organized as follows. In Section 2, we provide some preliminary facts. Section 3 is devoted to construct and to interpret the process $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ conditioned to non extinction. In Section 4, we investigate the asymptotic behavior of the population in the framework where the reproduction law is in the domain of attraction of an $\alpha$-stable distribution. In the last section, we prove Theorem 1.7, which is the result that explains the asymptotic behavior of the process conditioned to nonextinction.

## 2. Preliminaries

In this section, we obtain some useful formulas for the generating function of $(T_n, M_{n+1})$, denoted for $n \in \mathbb{Z}_+$ by

$$\varphi_n(x, y) := \mathbb{E}_1\left(x^{T_{n-1}} y^{M_n}\right), \qquad x, y \in [0, 1],$$

where for notational convenience $\varphi_1(x, y) := \varphi(x, y)$. Observe that the generating function of $M_n$ is

$$f_n(y) := \varphi_n(1, y), \qquad y \in [0, 1], \tag{2.1}$$

and as before we denote $f_1(y) =: f(y)$.

According with the classical theory of branching processes, the extinction probability of the Galton–Watson process $\{M_n : n \in \mathbb{Z}_+\}$, that we denote by $q$, is the smallest root of $f(y) = y$, which is less or equal than one depending on whether the mean of the reproduction law, $m := \mathbb{E}_1(M_1)$ is $> 1$ or $\leq 1$, respectively. In order to avoid trivial cases, we assume throughout that

(H1) $\mathbb{P}(M_1 = 1) > 0$,
(H2) $\mathbb{P}(M_1 = 0) + \mathbb{P}(M_1 = 1) < 1$, and $\mathbb{P}(M_1 = j) \neq 1$, for any $j$.

We also know that the $n$-step transition probabilities $\{\mathbf{P}^n_{(i,j)} : i, j \in \mathbb{Z}_+\}$ of the Galton–Watson process $\{M_n : n \in \mathbb{Z}_+\}$ satisfy

$$\sum_{j=0}^{\infty} \mathbf{P}^n_{(i,j)} y^j = \left(f_n(y)\right)^i, \qquad i \geq 1. \tag{2.2}$$

For a Galton–Watson process with neutral mutations, let $g$ be the generating function of the reproduction law of a typical individual, that is

$$g(x, y) := \mathbb{E}\left(x^{\xi^{(c)}} y^{\xi^{(m)}}\right), \qquad x, y \in [0, 1].$$

Proposition 1 of [3] ensures that the law of $(T_0, M_1)$ can be obtained applying the Lagrange inversion formula to the equation

$$\varphi(x, y) = xg\left(\varphi(x, y), y\right), \qquad x, y \in [0, 1]. \tag{2.3}$$

Thanks to this latter equality it is possible to deduce that $T \leq \infty$ if $\mathbb{E}(\xi^{(c)}) < 1$ and $\mathbb{E}(\xi^{(+)}) \leq 1$. Similarly, we have that $\mathbb{E}(\xi^{(+)2}) < \infty$ if and only if $\mathbb{E}(M_1^2) < \infty$. This equality is also a key tool to establish the following identity

$$\mathbb{P}_a(T_0 = k, M_1 = l) = \frac{a}{k}\pi_{k-a,l}^{*k}, \qquad k \geq a \geq 1 \text{ and } l \geq 0, \tag{2.4}$$

where $\pi^{*k}$ denotes the $k$th convolution of $\pi$, as defined in (1.9). Using the previous display, we can write the hypothesis (H1) and (H2) in terms of the reproduction distribution of a typical individual.

Moreover, letting $P_{(i,j),(k,l)}^n$ be the $n$-step transition probabilities of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, for this process we have a equality similar to (2.2), that is,

$$\sum_{k,l=0}^{\infty} P_{(i,j),(k,l)}^n x^k y^l = (\varphi_n(x, y))^j, \qquad i, j \geq 1. \tag{2.5}$$

We get the latter equality by induction. Namely, we apply the Chapman–Kolmogorov equation to express the $(n + 1)$-step transition probabilities in terms of the transitions in one step and use (1.1).

A simple but key relation for our analysis is

$$\varphi_n(x, y) = f_{n-1}(\varphi(x, y)), \qquad x, y \in [0, 1]. \tag{2.6}$$

Due to (2.1), the proof of the latter identity is equivalent to establish the following equality

$$\varphi_n(x, y) = \varphi_{n-1}(1, \varphi(x, y)), \qquad x, y \in [0, 1], \tag{2.7}$$

which follows from the standard calculations:

$$\varphi_n(x, y) = \mathbb{E}_1\big(\mathbb{E}_1\big(x^{T_{n-1}} y^{M_n}|T_{n-2}, M_{n-1}\big)\big)$$

$$= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j)$$

$$\times \sum_{k=j}^{\infty}\sum_{l=0}^{\infty} x^k y^l \mathbb{P}_1(T_{n-1} = k, M_n = l|T_{n-2} = i, M_{n-1} = j)$$

$$= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j)\sum_{k=j}^{\infty}\sum_{l=0}^{\infty} x^k y^l \mathbb{P}_j(T_0 = k, M_1 = l)$$

$$= \sum_{i,j=0}^{\infty} \mathbb{P}_1(T_{n-2} = i, M_{n-1} = j)(\varphi(x, y))^j$$

$$= \varphi_{n-1}(1, \varphi(x, y));$$

where we used the Markov property of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, Lemma 1.1 and the branching property.

# 3. The process conditioned to nonextinction

This section is devoted to study the process $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ conditioned to nonextinction.

## 3.1. Construction

Here our aim is to prove Theorem 1.3, which ensures the existence of the law of a Markovian process that we understand as the chain $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, conditioned to nonextinction of mutants in the population.

**Proof of Theorem 1.3.** An application of the Monotone Convergence theorem, along with an elementary computation, shows that

$$\frac{d}{ds} \mathbb{E}_a\left(s^{M_n}\right)\Big|_{s=q} = \mathbb{E}_a\left(M_n q^{M_n-1}\right).$$

Moreover, the following identity is deduced from the branching property of the Galton–Watson process $\{M_n : n \in \mathbb{Z}_+\}$ and the properties of its generating function

$$\frac{d}{ds} \mathbb{E}_a\left(s^{M_n}\right)\Big|_{s=q} = a q^{a-1} f_n'(q).$$

The latter and former identities imply in turn that

$$\mathbb{E}_a\left(M_n q^{M_n-1}\right) = a q^{a-1} f_n'(q).$$

Then by the Markov property,

$$\mathbb{E}_a\left(M_{n+k} q^{M_{n+k}-1} | \mathcal{F}_n\right) = M_n q^{M_n-1} f_k'(q).$$

Combining the latter with the fact that $f_k'(q) = [f'(q)]^k$ (see [1], Lemma 3.3), we have that

$$Y_n = \frac{M_n q^{M_n-a}}{(f'(q))^n}, \qquad n \geq 0,$$

is a martingale. Now from the theory of $h$-transforms (see Chapter 11 in [6]), there exists a Markovian process, that we denote by $\{(T_n^\uparrow, M_{n+1}^\uparrow) : n \in \mathbb{Z}_+\}$ whose law satisfies

$$\mathbb{P}_a^\uparrow\left(T_0^\uparrow = i_0, M_1^\uparrow = j_1, \ldots, T_{n-1}^\uparrow = i_{n-1}, M_n^\uparrow = j_n\right) := \mathbb{P}_a(A_n) \frac{j_n q^{j_n-a}}{a(f'(q))^n}, \qquad (3.1)$$

where for every $n \in \mathbb{N}$

$$A_n = \{T_0 = i_0, M_1 = j_1, \ldots, T_{n-1} = i_{n-1}, M_n = j_n\},$$

$$i_0, j_1, \ldots, i_{n-1}, j_n \in \mathbb{N}. \tag{3.2}$$

$\square$

## 3.2. Conditional laws

Here we will prove Theorem 1.4. The building block in this aim will be the generating function of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$, hence some of the results given in Section 2 will be necessary.

**Proof of Theorem 1.4.** (i) Let $A_n$ be an event of the form given in (3.2). It thus follows from the Markov and branching properties of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ that

$$\mathbb{P}_a(n + k < T < \infty) = \mathbb{E}_a\left(\mathbf{1}_{\{M_{n+k} > 0\}} q^{M_{n+k}}\right),$$

where as before $q = \mathbb{P}_1(0 < T < \infty)$. We also have that

$$\mathbb{P}_a(A_n, n + k < T < \infty) = \mathbb{E}_a\left(\mathbf{1}_{A_n} \mathbb{E}_{j_n}\left(\mathbf{1}_{\{M_k > 0\}} q^{M_k}\right)\right).$$

Then using (2.2), we get

$$\mathbb{P}_a(A_n | n + k < T < \infty) = \mathbb{P}_a(A_n) \frac{\sum_{j=1}^{\infty} \mathbf{P}^k_{(j_n, j)} q^j}{\sum_{j=1}^{\infty} \mathbf{P}^{n+k}_{(1, j)} q^j},$$

we recall that $\mathbf{P}^n_{(i,j)}$ denotes the $n$-step transition probabilities of $\{M_n : n \in \mathbb{Z}_+\}$. Besides, Theorem 7.4 of [1] establishes that the following limit holds

$$\lim_{k \to \infty} \frac{\mathbf{P}^{n+k}_{(i_1, j)}}{\mathbf{P}^k_{(i_2, j)}} = i_1 i_2^{-1} \left(f'(q)\right)^k q^{i_1 - i_2}.$$

Finally, thanks to the hypothesis (H2) we can use the previous identity to obtain

$$\lim_{k \to \infty} \mathbb{P}_a(A_n | n + k < T < \infty) = \mathbb{P}_a(A_n) \frac{j_n q^{j_n - a}}{a (f'(q))^n}, \qquad a \in \mathbb{N},$$

which finishes the first part of the proof.

(ii) We will first ensure the convergence of the generating function. Since $\{M_n > 0\}$ on the event $\{n < T < \infty\}$, we deduce that for all $x, y \in [0, 1]$:

$$\widehat{\varphi}_n(x, y) := \mathbb{E}_1\left(x^{T_{n-1}} y^{M_n} | n < T < \infty\right)$$

$$= \frac{\varphi_n(x, y) - \varphi_n(x, 0)}{1 - \mathbb{P}(M_n = 0)}.$$

From the identity (2.6) and the fact that $f_n(0) = \mathbb{P}(M_n = 0)$, the previous expression can be written as follows

$$\widehat{\varphi}_n(x, y) = \frac{1 - f_{n-1}(0)}{1 - f_n(0)} \left( \frac{f_{n-1}(\varphi(x, y)) - f_{n-1}(0)}{1 - f_{n-1}(0)} - \frac{f_{n-1}(\varphi(x, 0)) - f_{n-1}(0)}{1 - f_{n-1}(0)} \right). \qquad (3.3)$$

We now take $u = f_{n-1}(0)$ and use $m = f'(1)$ to obtain

$$\lim_{n \to \infty} \frac{1 - f_{n-1}(0)}{1 - f_n(0)} = \lim_{u \to 1} \frac{1 - u}{1 - f(u)} = \frac{1}{m}.$$

Observe that the function

$$n \longmapsto \frac{1 - f_{n-1}(s)}{1 - f_{n-1}(0)},$$

is decreasing for each $s$. Therefore, as $n$ tends to infinity the expression

$$\frac{f_{n-1}(s) - f_{n-1}(0)}{1 - f_{n-1}(0)} = 1 - \frac{1 - f_{n-1}(s)}{1 - f_{n-1}(0)},$$

has a limit, say $1 - \widehat{f}(s)$. According to Theorem 1.8.1 in [1], we know that the generating function, $\widehat{f}(s)$, of the Yaglom distribution of $\{M_n : n \in \mathbb{Z}_+\}$ given by the following limit

$$\rho_k = \lim_{n \to \infty} \mathbb{P}(M_n = k | n < T < \infty) \qquad \text{for all } k \in \mathbb{N}.$$

The above cited theorem also ensures that

$$1 - \widehat{f}(f(s)) = m(1 - \widehat{f}(s)), \qquad s \in [0, 1]. \qquad (3.4)$$

Putting all the pieces together in (3.3), we obtain

$$\widehat{\varphi}(x, y) := \lim_{n \to \infty} \widehat{\varphi}_n(x, y) = \frac{\widehat{f}(\varphi(x, y)) - \widehat{f}(\varphi(x, 0))}{m}.$$

We now prove by induction (1.8). If $n = 1$, it is the just proved equality. Then suppose (1.8) holds for $n = k$. In order to get the identity for $n = k + 1$ note that by the induction hypothesis

$$m^{k+1} \widehat{\varphi}(x, y) = m[1 - \widehat{f}(\varphi_k(x, 0))] - m[1 - \widehat{f}(\varphi_k(x, y))], \qquad x, y \in [0, 1].$$

From the above, we deduce the claim using first (3.4) and then (2.6). □

**Remark 3.1.** In the previous proof, we established the existence of a Yaglom limit when $m \leq 1$, however as in the classical case similar arguments can be used to show existence of a Yaglom limit in the supercritical case.

## 3.3. Interpretation

Motivated by the interpretation of a Galton–Watson process conditioned to nonextinction given in [12], our objective in the present subsection is to describe the chain $\{(T_n^\uparrow, M_{n+1}^\uparrow) : n \in \mathbb{Z}_+\}$ in terms of immigration of mutants. We start calculating the generating function of the $n$-step transition probabilities of this process.

**Proposition 3.1.** *Letting $Q_{(i,j),(k,l)}^n$ be the n-step transition probabilities of the process*

$$\{(T_n^\uparrow, M_{n+1}^\uparrow) : n \in \mathbb{Z}_+\},$$

*we have, for $x, y \in [0, 1]$*

$$\sum_{k,l=1}^\infty Q_{(i,j),(k,l)}^n x^k y^l = \frac{yq^{1-j}}{[f'(q)]^n} [\varphi_n(x, qy)]^{j-1} \frac{\partial}{\partial y} \varphi(x, qy) \prod_{i=1}^{n-1} f'(\varphi_i(x, qy)). \tag{3.5}$$

**Proof.** Let us start by pointing out the following formula

$$\sum_{k,l=1}^\infty Q_{(i,j),(k,l)}^n x^k y^l = \frac{yq^{1-j}}{[f'(q)]^n} \left[ \varphi_n^{j-1}(x, u) \frac{\partial}{\partial u} \varphi_n(x, u) \right]_{u=qy}, \qquad x, y \le 1.$$

This is a consequence of the fact that for each $n \in \mathbb{Z}_+$ the generating function of $(T_n, M_{n+1})$ is infinitely differentiable in $(x, y) \in [0, 1]^2$, the identity (2.5) and some elementary computation. The claimed formula is obtained by applying repeatedly (2.6) and the recursion $f_n(y) = f(f_{n-1}(y))$.  □

Taking $x = 1$ in (3.5) and recalling the fact that the transition probabilities of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ depend only on the second coordinate, we can identify a Galton–Watson process with immigration (see [11] for background).

**Corollary 3.2.** *If $\{M_n : n \in \mathbb{Z}_+\}$ is critical or subcritical, then $\{M_n^\uparrow - 1 : n \in \mathbb{Z}_+\}$ is a Galton–Watson process with immigration $[f, \frac{f'}{m}]$.*

Note that $\{M_n^\uparrow : n \in \mathbb{Z}_+\}$ is the $Q$-process associated to the Galton–Watson process $\{M_n : n \in \mathbb{Z}_+\}$ (see for instance [1] or [12]). The following corollary is analogous to Proposition 1 in [3].

**Corollary 3.3.** *If $\{M_n : n \in \mathbb{Z}_+\}$ is critical or subcritical, then the generating function of $(T_0^\uparrow, M_1^\uparrow)$ is determined by the equation*

$$\mathbb{E}_1\left(x^{T_0^\uparrow} y^{M_1^\uparrow}\right) = \frac{xy}{m} \frac{\partial}{\partial y} g(\varphi(x, y), y), \qquad x, y \in [0, 1].$$

*Moreover, the distribution of $(T_0^\uparrow, M_1^\uparrow)$ is given by*

$$\mathbb{P}_a^\uparrow\big(T_0^\uparrow = k, M_1^\uparrow = l\big) = \frac{l}{mk}\pi_{k-a,l}^{*k}, \qquad k \geq a \geq 1 \text{ and } l \geq 0,$$

*we recall that $\pi^{*k}$ denotes the kth convolution of the law $\pi$, defined in (1.9).*

**Proof.** Taking $n = 1$ in the equality (3.5),

$$\mathbb{E}_1\big(x^{T_0^\uparrow} y^{M_1^\uparrow}\big) = \frac{y}{m}\frac{\partial}{\partial y}\varphi(x, y).$$

Then the first identity is obtained using the identity (2.3). To get the second one, recall the definition of $\mathbb{P}^\uparrow$ given in (3.1) and (2.4). □

We can now give an interpretation to the process $\{(T_n^\uparrow, M_{n+1}^\uparrow) : n \in \mathbb{Z}_+\}$, in terms of a tree of alleles with immigration $\mathcal{A}^\uparrow = \{\mathcal{A}_u^\uparrow : u \in \mathbb{U}\}$. This tree will provide a description of the genealogical structure in a population conditioned to non extinction.

We start defining $\mathcal{A}_\varnothing^\uparrow = T_0^\uparrow$ that is, the total number of individuals without mutations into the population, then according to a distribution with generating function $f'/m$, a random number of individuals of the same genetic type arrive. We enumerate the $M_1^\uparrow$ allelic sub-populations of the first type beget by $T_0^\uparrow$ in decreasing order, with the convention that in the case of ties, sub-populations of the same size are ranked uniformly at random. Using Corollary 3.2, we choose uniformly at random one of the first type sub-families in the tree of alleles, removing it and replace it by a population of size $T_0^\uparrow$ which begets allelic subpopulation according to $M_1^\uparrow$, where $(T_0^\uparrow, M_1^\uparrow)$ is given by Corollary 3.3. We continue with the construction by iteration, $\mathcal{A}_{uj}^\uparrow$ is the size of the $j$th sub-population allelic of type $|u| + 1$ which descend from the allelic sub-family indexed by the vertex $u$. Then we choose one of the sub-families of type $|u| + 1$ to replace it by one of size $T_0^\uparrow$, which begets allelic subpopulation according to $M_1^\uparrow$.

## 4. Asymptotic behavior: The $\alpha$-stable case

Our goal in this section is to prove Theorem 1.6. For that end, until further notice we will consider a sequence of Galton–Watson processes $\{Z_k^{(+n)} : k \in \mathbb{Z}_+\}$ such that the reproduction law $\pi^{(+)}$ is critical with heavy tails; the mutations appear in the population according to (1.10); and the mutation rate, together with the ancestors behavior is given by (1.21).

### 4.1. Approximations for the reproduction law

We start by describing the normalizing sequence appearing in Theorem 1.6.

**Lemma 4.1.** *If the condition* (1.19) *holds, then there exists a sequence* $\{r(n) : n \geq 0\}$ *that is regularly varying at infinity with index* $\alpha$ *such that*

$$r(n)\pi^+(n\,dy) \xrightarrow[n\to\infty]{} c_\alpha \frac{dy}{y^{1+\alpha}},$$

*in the sense of vague convergence on* $(0, \infty)$, *where* $c_\alpha = 1/\Gamma(3 - \alpha)$. *In particular,*

$$\exp\left\{-t\int_{(0,\infty)}\left(1 - e^{-\lambda y} - \lambda y\right)r(n)\pi^+(n\,dy)\right\} \xrightarrow[n\to\infty]{} e^{-t\lambda^\alpha}.$$

The proof is an elementary application of standard results from the theory of Regular Variation (see, e.g., [4] for background). We include a proof in Appendix A for sake of completeness.

In order to link the asymptotic behaviour of the reproduction law of a typical individual with that of the joint distribution of clones and mutants, we first link their Laplace transform. Although in the present setting we use some ideas of the standard Tauberian–Abelian theorem, we remark that it is not straightforward application of this theorem because we consider sequences of measures indexed by the positive integers changing, unlike to the standard case, where only the normalizing constants change.

**Lemma 4.2.** *For every positive integer* $n$, *let* $\phi_n$ *be the Laplace transform of* $\xi^{(n)} = (\xi^{(cn)}, \xi^{(mn)})$ *under the measure* $\mathbb{P}_1^{p(n)}$. *Assume that* $\{\lambda(n) : n \in \mathbb{Z}_+\}$ *is a positive sequence such that* $\lambda(n) \to 0$, *as* $n \to \infty$. *Then, as* $n \to \infty$

$$\phi_n\big(\lambda(n), \theta\big) \sim \phi^+\big((1 - p(n))(1 - e^{-\lambda(n)}) + p(n)(1 - e^{-\theta})\big), \qquad \forall\theta \geq 0, \qquad (4.1)$$

*where* $\phi^+$ *is the Laplace transform of* $\xi^{(+)}$. *In particular* $\phi_n^m$, *respectively* $\phi_n^c$, *the Laplace transform of* $\xi^{(mn)}$, *respectively* $\xi^{(cn)}$, *satisfies*

$$\phi_n^m(\theta) \sim \phi^+\big(p(n)(1 - e^{-\theta})\big), \qquad \forall\theta \geq 0, \qquad (4.2)$$

$$\phi_n^c\big(\lambda(n)\big) \sim \phi^+\big((1 - p(n))(1 - e^{-\lambda(n)})\big) \qquad as\ n \to \infty. \qquad (4.3)$$

**Proof.** According to (1.10), conditionally to $\xi^{(+)} = k$ the distribution of $\xi^{(m)}$ is Binomial with parameter $(k, p)$. This fact implies the following equality in law

$$\big(\xi^{(c)}, \xi^{(m)}\big) \overset{\mathcal{L}}{=} \sum_{i=1}^{\xi^{(+)}}(\mathbf{1}_{\{U_i > p\}}, \mathbf{1}_{\{U_i \leq p\}}), \qquad (4.4)$$

where $\{U_i : i \in \mathbb{N}\}$ are independent random variables with common distribution that of an uniform random variable in $(0, 1)$. Therefore,

$$\phi_n\big(\lambda(n), \theta\big) = \sum_{k=0}^{\infty}\mathbb{P}_1^{p(n)}\big(\xi^{(+)} = k\big)\big[(1 - p(n))e^{-\lambda(n)} + p(n)e^{-\theta}\big]^k$$

$$= \phi^+\big(-\log\big(1 - \big(1 - (1 - p(n))e^{-\lambda(n)} - p(n)e^{-\theta}\big)\big)\big).$$

We conclude the proof using (1.11) and the elementary asymptotic estimate

$$\frac{\log(1-y)}{y} \xrightarrow[y \to 0]{} -1. \tag{4.5}$$

$\square$

In the same way, it is possible to establish the following estimate.

**Corollary 4.3.** *For every positive integer $n$, let $\psi_n$ be the characteristic function of $\xi^{(n)} = (\xi^{(cn)}, \xi^{(mn)})$ under the measure $\mathbb{P}_1^{p(n)}$. Then*

$$\psi_n(\lambda(n), \theta) \sim \phi^+\big((1 - p(n))(1 - e^{i\lambda(n)}) + p(n)(1 - e^{i\theta})\big),$$

$$\text{as } \lambda(n) \xrightarrow[n \to \infty]{} 0, \forall \theta \geq 0. \tag{4.6}$$

*In particular $\psi_n^m$, respectively $\psi_n^c$, the characteristic function of $\xi^{(mn)}$, respectively of $\xi^{(cn)}$, satisfies*

$$\psi_n^m(\theta) \sim \phi^+\big(p(n)(1 - e^{i\theta})\big),$$

$$\psi_n^c(\lambda(n)) \sim \phi^+\big((1 - p(n))(1 - e^{i\lambda(n)})\big),$$

*as $n \to \infty$, $\lambda(n) \to 0$ and for all $\theta \geq 0$.*

**Proof.** Similarly to previous lemma, using (4.4) we have

$$\psi_n(\lambda(n), \theta) = \sum_{k=0}^{\infty} \mathbb{P}_1^{p(n)}(\xi^{(+)} = k)\big[(1 - p(n))e^{i\lambda(n)} + p(n)e^{i\theta}\big]^k$$

$$= \sum_{k=0}^{\infty} \mathbb{P}_1^{p(n)}(\xi^{(+)} = k) \exp\big\{k \log\big(1 - \big((1 - p(n))(1 - e^{i\lambda(n)}) - p(n)(1 - e^{i\theta})\big)\big)\big\}.$$

To conclude, we apply the asymptotic estimate (4.5). $\square$

We can now use the results above to give an estimate for the reproduction measure.

**Proposition 4.4.** *For every positive integer $n$, let $\pi^{(cn)}$, $\pi^{(mn)}$ be the reproduction laws of $\xi^{(cn)}$ and $\xi^{(mn)}$, respectively. Assume that $\{y(n) : n \geq 0\}$ is any sequence such that $y(n) \to \infty$ as $n \to \infty$. In the regime (1.19), the asymptotic behavior of the tail distribution of $\xi^{(\cdot)}$ is given by*

$$\bar{\pi}^{(\cdot)}(y(n)) \sim c_\alpha \bar{\pi}^+(y(n)/\mathbb{E}(\xi^{(\cdot)})) \qquad \text{as } n \to \infty, \tag{4.7}$$

*where $(\cdot) = cn, mn$ and $c_\alpha = 1/\Gamma(3 - \alpha)$.*

The proof of this proposition is deferred to the Appendix B because it use some elements of the proof of Lemma 4.1, which is included in Appendix A.

## 4.2. Proof of Theorem 1.6

The aim of this section is to prove Theorem 1.6. For that end, we required two auxiliaries results that we will next state.

Let $n \in \mathbb{N}$ fixed. According with the construction of alleles trees, given a vertex $u$, at level $k \geq 1$, in $\{\mathcal{A}_u : u \in \mathbb{U}\}$, a vertex $uj$ represents the size of the $j$th allelic sub-populations of type $k + 1$ begot by $u$, and this holds for every $j \in \mathbb{N}$. Thus, the labels of the vertices at level $k + 1$ determine the variable $T_{k+1}$. Moreover, for all $k \in \mathbb{N}$, the total number of vertices at level $k$ correspond to $M_k$. Hence, a first step to establish the convergence in Theorem 1.6 will be describe the scaling limit of the process $\{(T_k, M_{k+1}) : k \in \mathbb{Z}_+\}$, towards a CSBP $\{Z_k^{1/\alpha} : k \in \mathbb{Z}_+\}$. That is the purpose of the Proposition 4.5 below, whose proof its deferred to Section 4.3.

**Proposition 4.5.** *Assuming* (1.19) *and* (1.21), *we have*

$$\mathcal{L}\left(\left\{\left(\frac{T_k}{r(n)}, \frac{M_{k+1}}{r(n)\,p(n)}\right) : k \in \mathbb{Z}_+\right\}, \mathbb{P}_{a(n)}^{p(n)}\right) \implies \left\{\left(Z_{k+1}^{1/\alpha}, Z_{k+1}^{1/\alpha}\right) : k \in \mathbb{Z}_+\right\}, \qquad (4.8)$$

*where* $\{Z_k^{1/\alpha} : k \in \mathbb{Z}_+\}$ *is a CSBP process with reproduction measure* $\nu^\alpha$ *given in Theorem 1.6.*

To obtain from this result the convergence claimed in Theorem 1.6, we will need the Lemma 4.6 below, whose proof is given in Section 4.4.

**Lemma 4.6.** *Let* $b(n)$ *be a sequence of integers such that* $b(n) \sim b\, r(n)\, p(n)$ *for some* $b > 0$.

(i) *For every* $n \in \mathbb{N}$, *let* $\{\chi_j^{(n)} : 1 \leq j \leq b(n)\}$ *be a sequence of independent identically distributed random variables with distribution* $(\frac{T_0}{r(n)}, \frac{M_1}{r(n)\,p(n)})$. *Defining for every* $n \in \mathbb{N}$, $\gamma_j^{(n)} := \delta_{\chi_j^{(n)}}$ *and* $\gamma_n = \sum_{j=1}^\infty \gamma_j^{(n)}$, *the following weak convergence of measures holds*

$$\gamma_n \xrightarrow[n \to \infty]{} \gamma, \qquad (4.9)$$

*where* $\gamma$ *is a Poisson point measure with intesity* $b\eta$, *with* $\eta$ *the image of the measure* $\nu^\alpha$ *(given in Theorem 1.6) by the action of the map* $x \mapsto (x, x)$.

(ii) *We have the following convergence, under the measure* $\mathbb{P}_1^{p(n)}$

$$\left(\frac{T_0}{r(n)}, \frac{M_1}{r(n)\,p(n)}\right)^{(b(n)^\downarrow)} \implies (\mathbf{a}_1, \mathbf{a}_2, \ldots), \qquad (4.10)$$

*where for all* $k \in \mathbb{N}$, $\mathbf{a}_k = (a_k, a_k)$ *with* $\{a_k : k \in \mathbb{N}\}$, *the atoms of a Poisson random measure on* $(0, \infty)$ *with intensity* $b\nu^\alpha$ *ranked in decreasing order; the measure* $\nu^\alpha$ *is given in Theorem 1.6.*

Taking for granted the above results, we can provide a proof to Theorem 1.6.

### 4.2.1. *Proof of Theorem* 1.6

We will establish

$$\mathcal{L}\left(\left(\left(\left(\frac{\mathcal{A}_u}{r(n)}, \frac{d_u}{r(n)p(n)}\right) : |u| \leq k\right) : k \in \mathbb{Z}_+\right), \mathbb{P}_{a(n)}^{p(n)}\right) \tag{4.11}$$
$$\implies \quad \left(\left(\left(\mathcal{Z}_u^{1/\alpha}, \mathcal{Z}_u^{1/\alpha}\right) : |u| \leq k\right) : k \in \mathbb{Z}_+\right).$$

Actually by the monotone class theorem, it is enough to show for nonnegative measurable continuous functions $f_1, \ldots, f_k$

$$\mathbb{E}_{a(n)}^{p(n)}\left[\prod_{i=1}^{k} f_i\left((r(n))^{-1}\mathcal{A}_u, (r(n)p(n))^{-1}d_u : |u| \leq i\right)\right]$$
$$\xrightarrow[n\to\infty]{} \mathbb{E}^{\mathbb{Q}_x}\left[\prod_{i=1}^{k} f_i\left(\mathcal{Z}_u^{1/\alpha}, \mathcal{Z}_u^{1/\alpha} : |u| \leq i\right)\right],$$

where $\mathbb{Q}_x$ is the law of a tree-indexed CSBP started with an initial population of size $x$ constructed from the subordinator $\{\tau_t^\alpha : t \geq 0\}$. This will be done by induction on $k$. The case $k = 1$ is given in the convergence (4.10). Assuming the result holds for $k$, we will prove the convergence for $k + 1$. Conditioning with respect to $\mathcal{G}_k = \sigma(\mathcal{A}_u^{(n)}, d_u^{(n)} : |u| \leq k)$ and using Lemma 1.2, we have

$$\mathbb{E}_{a(n)}^{p(n)}\left[\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k+1} f_i\left((r(n))^{-1}\mathcal{A}_u, (r(n)p(n))^{-1}d_u : |u| = i\right)|\mathcal{G}_k\right)\right]$$
$$= \mathbb{E}_{a(n)}^{p(n)}\left[\prod_{i=1}^{k} f_i\left((r(n))^{-1}\mathcal{A}_u, (r(n)p(n))^{-1}d_u : |u| \leq i\right)\right.$$
$$\left. \times \mathbb{E}_1^{p(n)}\left(f_{k+1}\left(((r(n))^{-1}T_0, (r(n)p(n))^{-1}M_1)^{d_u\downarrow} : |u| = k\right)\right)\right].$$

Besides, by the induction hypothesis $d_u \sim r(n)p(n)\mathcal{Z}_u^{1/\alpha}$ with $|u| = k$, therefore when $n \to \infty$ in the previous equality we obtain

$$\mathbb{E}^{\mathbb{Q}_x}\left[\prod_{i=1}^{k} f_i\left(\mathcal{Z}_u^{1/\alpha}, \mathcal{Z}_u^{1/\alpha} : |u| \leq i\right)\mathbb{E}_1\left(f_{k+1}\left((\mathbf{a}_1', \mathbf{a}_2', \ldots)\right)\right)\right],$$

where $\mathbf{a}_k' = (a_k', a_k')$ are the atoms of a Poisson random measure on $(0, \infty)$ with intensity $b\mathcal{Z}_u^{1/\alpha}\nu^\alpha$, repeated according to their multiplicity and ranked in the decreasing order. Due to the definition of a tree-indexed CSBP this concludes the proof.

### 4.3. Proof of Proposition 4.5

Observe that is enough to show the convergence of the Laplace transforms associated with the finite dimensional distributions of each processes. Since $\{Z_k^{1/\alpha} : k \in \mathbb{Z}_+\}$ is a CSBP with transition probabilities characterized by the subordinator $\tau^\alpha$, we use (1.16) to get its Laplace transform. Therefore we will establish the following convergence:

$$\mathbb{E}_{a(n)}^{p(n)}\left(\prod_{i=1}^{k} e^{-\frac{s_{i-1}}{r(n)}T_{i-1} - \frac{t_i}{r(n)p(n)}M_i}\right) \xrightarrow[n\to\infty]{} \exp\{-x\kappa\left(l_{k-1}(s_0 + t_1)\right)\}, \tag{4.12}$$

for all $s_i, t_i \geq 0$, $i = 1, 2, \dots, k$. According to (1.15), $\kappa$ denotes the cumulant of the subordinator that characterizes the transition probabilities. In this case, we have the $1/\alpha$-stable subordinator $\tau^\alpha$ with the Lévy measure $\nu^\alpha$. The function $l$ is given in (1.17).

We will prove the converge (4.12) by induction on $k$. The aim of the following lemma is to ensure the above claimed result holds for $k = 0$.

**Lemma 4.7.** *For $\alpha \in (1, 2)$, let $\tau^\alpha$ be an $1/\alpha$-stable subordinator with no drift and Lévy measure $\nu^\alpha$. Assuming (1.19) and (1.21)*

(i) *the following convergences hold*:

$$\mathcal{L}\left(\left(\frac{T_0}{r(n)}, \frac{M_1}{r(n)p(n)}\right), \mathbb{P}_{a(n)}^{p(n)}\right) \implies \left(\tau_x^\alpha, \tau_x^\alpha\right); \tag{4.13}$$

(ii) *under the measure $\mathbb{P}_1^{p(n)}$, the behavior of the joint tail distribution of $T_0$ and $M_1$ is given by*

$$\lim_{n\to\infty} r(n)p(n)\mathbb{P}_1^{p(n)}\left(\frac{T_0}{r(n)} > s, \frac{M_1}{r(n)p(n)} > t\right) = \bar{\nu}^\alpha(s \wedge t), \tag{4.14}$$

*where $\bar{\nu}^\alpha$ denotes the tail function of the Lévy measure $\nu^\alpha$.*

A key tool to establish Lemma 4.7 is the following lemma.

**Lemma 4.8.** *In the regime (1.19) and (1.21), the normalized random walk defined by*

$$\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} = \left(a(n)/r(n)p(n), 0\right) + \sum_{i=1}^{\lfloor r(n)t \rfloor} \left((\xi_i^{(cn)} - 1)/n, \xi_i^{(mn)}/r(n)p(n)\right), \qquad t \geq 0,$$

*converges weakly*

$$\{\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} : t \geq 0\} \implies \{(x + X_t, t) : t \geq 0\},$$

*where $\{X_t : t \geq 0\}$ is an $\alpha$-stable process with no-negative jumps with and characteristic exponent $c_\alpha|\lambda|^\alpha$.*

Given that this result is similar to other existing results in the literature we prefer to postpone its proof to the Appendix C and focus in the proof of Lemma 4.7.

**Proof of Lemma 4.7(i).** From Lemma 3 of [3], we know that the first passage time below 0 for the centered random walk $\mathbf{S}_{1,k}^{(n)} = a(n) + \sum_{i=1}^{k} (\xi_k^{(cn)} - 1)$ has the same distribution as $T_0$. Let $(\varsigma^{(n)}(0), \Sigma^{(n)}(0))$ be the random variables

$$\varsigma^{(n)}(0) = \inf\{k \in \mathbb{Z}_+ : \mathbf{S}_{1,k}^{(n)} = 0\} \quad \text{and} \quad \Sigma^{(n)}(0) := \sum_{i=1}^{\varsigma^{(n)}(0)} \xi_i^{(mn)}.$$

According to Lemma 3 of [3], we have that $(\varsigma^{(n)}(0), \Sigma^{(n)}(0))$ has the same distribution as $(T_0, M_1)$ under $\mathbb{P}_{a(n)}^{p(n)}$. On the other hand, we also have the following two identities

$$\frac{\varsigma^{(n)}(0)}{r(n)} = \frac{1}{r(n)} \inf\{k \in \mathbb{Z}_+ : \mathbf{S}_{1,k}^{(n)} = 0\} = \inf\{t \geq 0 : \bar{\mathbf{S}}_{1,\lfloor r(n)t \rfloor}^{(n)} = 0\}$$

and

$$\left( \frac{1}{r(n)} \varsigma^{(n)}(0), \bar{\mathbf{S}}_{\varsigma^{(n)}(0)}^{(n)} \right) = \left( \frac{1}{r(n)} \varsigma^{(n)}(0), \left( \bar{\mathbf{S}}_{1,\lfloor \varsigma^{(n)}(0) \rfloor}^{(n)}, \frac{1}{r(n)p(n)} \Sigma^{(n)}(0) \right) \right). \tag{4.15}$$

From the Lemma 4.8, we have the weak convergence

$$\{\bar{\mathbf{S}}_{\lfloor r(n)t \rfloor}^{(n)} : t \geq 0\} \quad \Longrightarrow \quad \{(x + X_t, t) : t \geq 0\},$$

and in fact the convergence holds in the sense of Skorohod's topology, see Chapter IV of [17]. Since $X$ is an $\alpha$-stable process, Theorem 1 in Chapter VII of [2] ensures that the first passage time below $-x$ for the process $X$

$$\tau_x^\alpha = \inf\{t \geq 0 : X_t \leq -x\}, \qquad x \geq 0.$$

is a stable subordinator of parameter $1/\alpha$. We will conclude from these facts that the claimed convergence holds as soon as

$$\left( \frac{1}{r(n)} \varsigma^{(n)}(0), \bar{\mathbf{S}}_{\varsigma^{(n)}(0)}^{(n)} \right) \quad \Longrightarrow \quad \left( \tau_x^\alpha, (x + X_t, t)|_{t=\tau_x^\alpha} \right). \tag{4.16}$$

But according to Theorem 13.6.5 of [17] about weak convergence of first passage times and undershoots and overshoots, when there is convergence in Skorohod's topology, we have

$$\left( \frac{1}{r(n)} \varsigma^{(n)}(0), \bar{\mathbf{S}}_{1,\varsigma^{(n)}(0)}^{(n)} \right) \quad \Longrightarrow \quad \left( \tau_x^\alpha, X_{\tau_x^\alpha} + x \right).$$

Moreover, since we have the joint convergence

$$\left\{ \left( \bar{\mathbf{S}}_{1,\lfloor \varsigma^{(n)}(0)t \rfloor}^{(n)}, \bar{\mathbf{S}}_{2,\lfloor \varsigma^{(n)}(0)t \rfloor}^{(n)} \right) : t \geq 0 \right\} \quad \Longrightarrow \quad \{(x + X_t, t) : t \geq 0\}$$

in the sense of Skorohod's topology, and the second coordinate is a determinist linear function, it is an elementary exercise to extend the above mentioned result of [17] to get that the convergence in (4.16) holds. $\qquad\square$

**Proof of Lemma 4.7(ii).** We will apply the same techniques used in the proof of statement (ii) in Lemma 4 of [3]. Let us start observing that for every $x, y \in \mathbb{R}$

$$e^{-sx-ty} = st \int_0^\infty \int_0^\infty e^{-sx-ty} \mathbf{1}_{\{x<u, y<v\}} \, du \, dv, \qquad s, t \geq 0.$$

Thus Fubini's theorem implies that for any random vector $(X, Y)$ the following identity holds.

$$1 - \mathbb{E}\big(e^{-sX-tY}\big) = st \int_0^\infty \int_0^\infty e^{-sx-ty} \mathbb{P}(X \geq u \text{ or } Y \geq v) \, du \, dv, \qquad s, t \geq 0.$$

In particular,

$$1 - \mathbb{E}_1^{p(n)}\Big(e^{-\frac{s}{r(n)} T_0 - \frac{t}{r(n)p(n)} M_1}\Big) = st \int_0^\infty \int_0^\infty e^{-su-tv} \bar{\mu}_n\big(r(n)u, r(n)p(n)v\big) \, du \, dv, \qquad s, t \geq 0,$$

where $\bar{\mu}_n(x, y) := \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. Hence by the branching property,

$$
\begin{aligned}
\mathbb{E}_{a(n)}^{p(n)}&\Big(e^{-\frac{s}{r(n)} T_0 - \frac{t}{r(n)p(n)} M_1}\Big) \\
&= \bigg(1 - st \int_0^\infty \int_0^\infty e^{-su-tv} \bar{\mu}_n\big(r(n)u, r(n)p(n)v\big) \, du \, dv\bigg)^{a(n)}.
\end{aligned}
\tag{4.17}
$$

According to the first part of this lemma together with (1.21), the previous display converges as $n \to \infty$ towards

$$\mathbb{E}\big(e^{-(s+t)\tau_x^\alpha}\big) = \exp\bigg(-x \int_0^\infty \big(1 - e^{-(s+t)y}\big) v^\alpha(dy)\bigg).$$

Taking logarithms in the last two identities we obtain that

$$\lim_{n\to\infty} sta(n) \int_0^\infty \int_0^\infty e^{-su-tv} \bar{\mu}_n\big(r(n)u, r(n)p(n)v\big) \, du \, dv = x \int_0^\infty \big(1 - e^{-(s+t)y}\big) v^\alpha(dy).$$

Hence, it only remains to see that the line above is equal to

$$xst \int_0^\infty \int_0^\infty e^{-sy-tz} \bar{v}^\alpha(y \wedge z) \, dy \, dz.$$

For that end, we observe the equality

$$\int_0^\infty \int_0^\infty e^{-sy-tz} \bar{v}^\alpha(y \wedge z) \, dy \, dz = \int_0^\infty v^\alpha(du) \int_0^\infty \int_0^\infty e^{-sy-tz} \mathbf{1}_{\{u>y \text{ or } u>z\}} \, dy \, dz,$$

and we obtain the claimed identity by uniqueness of the Laplace transform. $\qquad\square$

Continuing with the proof of Proposition 4.5, we assume that (4.12) holds for $n = k$ and prove the convergence for $n = k + 1$. To this end, we use the Markov property of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ and the fact that, conditionally to $M_n = j$, the pair $(T_n, M_{n+1})$ has the same distribution as $(T_0, M_1)$ under $\mathbb{P}_j$, to obtain

$$\mathbb{E}^{p(n)}_{a(n)}\left(\prod_{i=1}^{k+1} e^{-\frac{s_{i-1}}{r(n)} T_{i-1} - \frac{t_i}{r(n)p(n)} M_i}\right)$$

$$= \mathbb{E}^{p(n)}_{a(n)}\left(e^{-\frac{s_0}{r(n)} T_0 - \frac{t_1}{r(n)p(n)} M_1} \cdots\right.$$

$$\left.\times\, e^{-\frac{s_{k-1}}{r(n)} T_{k-1} - \left(\frac{t_k}{r(n)p(n)} - \frac{1}{r(n)p(n)} \log \mathbb{E}^{p(n)}_{r(n)p(n)} (e^{-\frac{s_k}{r(n)} T_0 - \frac{t_{k+1}}{r(n)p(n)} M_1})\right) M_k}\right).$$

Due to the assumption $r(n)p(n) \sim a(n)x$ in hypothesis (1.21), we obtain as a consequence of Lemma 4.7(i), that

$$\mathbb{E}^{p(n)}_{a(n)}\left(\prod_{i=1}^{k+1} e^{-\frac{s_{i-1}}{r(n)} T_{i-1} - \frac{t_i}{r(n)p(n)} M_i}\right)$$

$$\sim \mathbb{E}^{p(n)}_{a(n)}\left(e^{-\frac{s_0}{r(n)} T_0 - \frac{t_1}{r(n)p(n)} M_1} \cdots e^{-\frac{s_{k-1}}{r(n)} T_{k-1} - \frac{1}{r(n)p(n)}(t_k + \kappa(s_k + t_{k+1})) M_k}\right).$$

Then using the induction hypothesis with

$$s'_{i-1} + t'_i = \begin{cases} rls_{i-1} + t_i, & i < k, \\ l(s_{i-1} + t_i), & i = k, \end{cases} \tag{4.18}$$

we get

$$\mathbb{E}^{p(n)}_{a(n)}\left(e^{-\frac{s_0}{r(n)} T_0 - \frac{t_1}{r(n)p(n)} M_1} \cdots e^{-\frac{s_{k-1}}{r(n)} T_{k-1} - \frac{1}{r(n)p(n)}(t_k + \kappa(s_k + t_{k+1})) M_k}\right)$$

$$\xrightarrow[n\to\infty]{} \exp\left\{-x\kappa\left(l_{k-1}(s'_0 + t'_1)\right)\right\}.$$

This concludes the proof because of the recursive definition of $l_i$ given in (1.17) together with the choice of $s'_{i-1} + t'_i$,

$$\kappa\left(l_{k-2}(s'_1 + t'_2)\right) = \kappa\left(l_{k-1}(s_1 + t_2)\right),$$

as consequence of $l_{k-1}(s'_0 + ct'_1) = l_k(s_0 + t_1)$.

## 4.4. Proof of Lemma 4.6

By the construction $\{\gamma_j^{(n)} : 1 \le j \le b(n)\}$ is a sequence of independent random variables. Besides, the convergence (4.14) in Lemma 4.7 implies $\gamma_j^{(n)} \xrightarrow{P} 0$ as $n \to \infty$, uniformly in $j$. Then

$$\sup_j \mathbb{E}\left(\left|\gamma_j^{(n)} \wedge 1\right|\right) \to 0.$$

Hence, according with the definition given in Chapter 4 of [10], we have that $\{\gamma_j^{(n)} : 1 \leq j \leq b(n)\}$ is a null array. Thus, we will get the convergence (4.9) as an application of Theorem 16.18 of [10], once we verify the following conditions:

(a) $\sum_j \mathbb{P}(\gamma_j^{(n)}(B) > 0) \to \eta(B)$, as $n \to \infty$, for all $B \in \hat{\mathcal{B}}$, where $\eta$ is as defined in the statement of the lemma,

(b) $\sum_j \mathbb{P}(\gamma_j^{(n)}(B) > 1) \to 0$, as $n \to \infty$, for all $B \in \mathcal{B}$.

Here $\mathcal{B}$ is the Borel $\sigma$-algebra of $[0, \infty)^2$, $\hat{\mathcal{B}} =: \{B \in \mathcal{B} : \gamma(\partial B) = 0 \text{ c.s.}\}$, with $\gamma$ the measure defined in the statement of this lemma and the symbol $\partial$ denotes the boundary of $B$. Observe that the class of sets $B = ((b, \infty) \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times (b', \infty))$ is a $\pi$-system which generates a $\lambda$-system that coincides with $\mathcal{B}$. Then, by Dynkin's theorem it is enough to establish the conditions above for sets of the latter form $B$. In this setting, the condition (b) holds because $\gamma_j^{(n)}(\cdot)$ takes only the values 0 or 1, for any $j$ and $n$. To establish (a), observe the following identities

$$\sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) > 0) = \sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) = 1)$$

$$= \sum_{j=1}^{b(n)} \bar{\mu}_n\big((r(n))^{-1}s, (r(n)p(n))^{-1}t\big)$$

$$= b(n)\bar{\mu}_n\big((r(n))^{-1}s, (r(n)p(n))^{-1}t\big),$$

here we recall the identity $\bar{\mu}_n(x, y) = \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. Assuming that $b(n) \sim b r(n)p(n)$ for some $b > 0$, from Lemma 4.7 and the last equality we have

$$\sum_{j=1}^{b(n)} \mathbb{P}(\gamma_j^{(n)}(B) > 0) \xrightarrow[n \to \infty]{} b \bar{\nu}^\alpha(s \wedge t).$$

To get the first convergence in the lemma, it remains to observe that the equality $b\bar{\nu}^\alpha(s \wedge t) = \eta(B)$ holds. But this follows from the equalities,

$$\int_B \eta(dx, dy) = b \int_{(x,x) \in B} \nu^\alpha(dx) = b \int_{(s \wedge t, \infty)} \nu^\alpha(dx).$$

We will now prove the convergence (ii). For $i = 1, 2$, $\chi_{ij}^{(n)}$ denotes the $i$th coordinate of the sequence $\chi_j^{(n)}$ that appears in the statement (i). Assuming that $\chi_i^{(n)} := (\chi_{1i}^{(n)}, \chi_{2i}^{(n)}) \leq \chi_j^{(n)} := (\chi_{1j}^{(n)}, \chi_{2j}^{(n)})$ if and only if $\chi_{1i}^{(n)} \leq \chi_{1j}^{(n)}$ or $\chi_{2j}^{(n)} \leq \chi_{2j}^{(n)}$, let us define $j_1$ as the index where the maximum of the sequence $\{\chi_i^{(n)} : 1 \leq i \leq b(n)\}$ is reached.

$$\chi_{j_1}^{(n)} = \max_{1 \leq i \leq b(n)} \chi_i.$$

Similarly for $k = 2, \ldots, b(n)$, let $j_k$ be the index of the $k$-order statistic

$$\chi_{j_k}^{(n)} = \max_{i \in J_k} \chi_i,$$

where $J_k = \{1, \ldots, b(n)\} \setminus \{j_1, \ldots, j_{k-1}\}$. Then observe that

$$\mathbb{P}\big(\chi_{j_1}^{(n)} \geq \mathbf{c}_1, \chi_{j_2}^{(n)} \geq \mathbf{c}_2, \ldots, \chi_{j_k}^{(n)} \geq \mathbf{c}_k\big) = \mathbb{P}\big(\gamma_n(C_1) \geq 1, \gamma_n(C_2) \geq 2, \ldots, \gamma_n(C_k) \geq k\big),$$

if $\mathbf{c}_i = (c_i, c_i)$, $C_i = (0, 1) \times (0, 1) \setminus (0, c_i) \times (0, c_i)$ and $c_1 > \cdots > c_k$. Taking now the limit as $n \to \infty$ in the equality below and using the convergence in (4.14), we have

$$\mathbb{P}(\chi_{j_1} \geq \mathbf{c}_1, \chi_{j_2} \geq \mathbf{c}_2, \ldots, \chi_{j_k} \geq \mathbf{c}_k) \xrightarrow[n \to \infty]{} \mathbb{P}\big(\gamma(C_1) \geq 1, \gamma(C_2) \geq 2, \ldots, \gamma(C_k) \geq k\big).$$

This implies the desired convergence because

$$\mathbb{P}\big(\gamma(C_1) \geq 1, \ldots, \gamma(C_k) \geq k\big) = \mathbb{P}(\mathbf{a}_{j_1} \geq \mathbf{c}_1, \mathbf{a}_{j_2} \geq \mathbf{c}_2, \ldots, \mathbf{a}_{j_k} \geq \mathbf{c}_k),$$

where $\mathbf{a}_k = (a_k, a_k)$ with $\{a_k : k \in \mathbb{N}\}$, the atoms of a Poisson random measure on $(0, \infty)$ with intensity $b\nu^\alpha$ ranked in decreasing order; the measure $\nu^\alpha$ is given in Theorem 1.6. As before we used the indices $j_k$ to rank in decreasing order the sequence $\mathbf{a}_k$.

# 5. Asymptotic behavior: The conditioned to nonextinction case

This section is devoted to establish Theorem 1.7. Following the same strategy of Proposition 4.5, we shall establish by induction the convergence of Laplace transforms of the finite dimensional distributions associated with the processes involved. With this aim, we first deduce the Laplace transform of the finite dimensional distributions of a CSBP with immigration $\{Z_n^I : n \in \mathbb{N}\}$, with mechanism $(\vartheta, \iota)$. We recall that it is defined for every $n \in \mathbb{N}$ as follows

$$Z_{n+1}^I = \tau_n\big(Z_n^I\big) + I_n,$$

where $\{I_n : n \in \mathbb{N}\}$ is a sequence of nonnegative random variables with common probability measure $\omega$, which determine the distribution of individual immigrants arriving in the population. Let us denote its Laplace transform of $\omega$ by $\iota$, i.e.

$$\iota(\lambda) = \int_0^\infty e^{-\lambda x} \omega(dx), \qquad \lambda \geq 0. \tag{5.1}$$

Let $\{T^{(n)}(t) : t \geq 0\}_{n \geq 0}$ be a sequence of independent subordinators (without drift) and also independents of $I_n$, with the same distribution and Laplace transform given in (1.15). Thereby

$$\mathbb{E}\big(e^{-\lambda T^{(n)}(Z_n^I)} | Z_n^I\big) = e^{-Z_n^I \kappa(\lambda)}.$$

This previous equality together with the Markov property imply that

$$\mathbb{E}_x\left(e^{-s_1 Z_1^I - \cdots - s_k Z_k^I}\right) = \prod_{i=0}^{k-1} \iota\left(l_i\left(s_{k-i}\right)\right) e^{-x\kappa(l_{k-1}(s_1))} \qquad \text{for all } s_i \geq 0, i = 1, 2, \ldots, k;$$

where the functions $l.(s_{k-.})$ are defined in (1.17). Thus, the proof of statement (i) of Theorem 1.7 require to establish that for all $s_i, t_i \geq 0$, $i = 1, 2, \ldots, k$, the convergence below holds

$$\mathbb{E}_{a_n}^{p(n)\uparrow}\left(\prod_{i=1}^{k} e^{-\frac{s_{i-1}}{n^2} T_{i-1} - \frac{t_i}{n} M_i}\right) \xrightarrow[n\to\infty]{} \exp\left\{-x\kappa\left(l_{k-1}(s_0+c_\vartheta t_1)\right)\right\} \prod_{i=0}^{k-1} \iota\left(l_{k-i}(s_{i-1}+\tilde{c}_\vartheta t_i)\right), \quad (5.2)$$

where $\kappa$ and $l$ are respectively defined in (1.15) and (1.17), taking in particular $\vartheta = \nu$ given in (1.13). To obtain (ii) of Theorem 1.7, the previous convergence is proved with $\vartheta = \nu^\alpha$ defined in (1.22). The following lemma establishes the above convergence in the case $k = 1$. In its proof we use the reference [3] to justify the statement corresponding to $\vartheta = \nu$ and previous results here obtained to establish the case where $\vartheta = \nu^\alpha$.

**Lemma 5.1.** *If* (1.11) *holds, then we have the following convergence*

$$\mathcal{L}\left(\left(b_1(n)T_0, b_2(n)M_1\right), \mathbb{P}_{a(n)}^{p(n)\uparrow}\right) \implies (\tau, c_\vartheta \tau),$$

*where $\tau$ is a random variable with Laplace transform $e^{-\kappa(s)}\iota(s)$, where $\kappa(s)$, $\iota(s)$ are given in* (1.15) *and* (5.1), *according to*

  (i) *if the reproduction law has finite variance $\sigma^2$, $\vartheta = c^{-1}\nu$, where $\nu$ is the measure in* (1.13). *Moreover, $b_1(n) = n^{-2}$, $b_2(n) = n^{-1}$ and $c_\vartheta = c$;*

  (ii) *otherwise, under the assumptions* (1.19) *and* (1.21), *$b_1(n) = (r(n))^{-1}$, $b_2(n) = (r(n)p(n))^{-1}$, $\vartheta = \nu^\alpha$ is given by* (1.22) *and $c_\vartheta = 1$.*

**Proof.** We will only prove the claim in the setting in (i). The proof in the other case is fully analogue. To simplify the notation, we just write $b_1$ and $b_2$. We prove the convergence of Laplace transform $(b_1 T_0, b_2 M_1)$ under the measure $\mathbb{P}_{a(n)}^{p(n)\uparrow}$. First, recalling the definition of the conditional measure given in (3.1), an elementary calculation using the branching property shows

$$\mathbb{E}_{a(n)}^{p(n)\uparrow}\left(e^{-sT_0 - tM_1}\right) = \mathbb{E}_{a(n)-1}^{p(n)}\left(e^{-sT_0 - tM_1}\right)\mathbb{E}_1^{p(n)}\left(e^{-sT_0 - tM_1} M_1\right). \tag{5.3}$$

Thanks to Lemma 4 of [3] and Lemma 4.7.

$$\mathbb{E}_{a(n)}^{p(n)}\left(\exp(-sb_1 T_0 - tb_2 M_1)\right) \xrightarrow[n\to\infty]{} \exp\left(-x\int_0^\infty \left(1 - e^{-(s+c_\vartheta t)y}\right)c_\vartheta^{-1}\vartheta\,(dy)\right), \tag{5.4}$$

so it remains to calculate the limit of the second factor in (5.3). Using again [3] together with the equality (4.17), we have

$$\mathbb{E}_1^{p(n)}\left(\exp(-sb_1 T_0 - tb_2 M_1)\right) = 1 - st\int_0^\infty \int_0^\infty e^{-sx-ty}\bar{\mu}_n\left(b_1^{-1}x, b_2^{-1}y\right)dx\,dy,$$

where as before $\bar{\mu}_n(x, y) := \mathbb{P}_1^{p(n)}(T_0 > x \text{ or } M_1 > y)$. Due to Lemma 4(ii) in [3] and Lemma 4.7(ii) the latter implies

$$\mathbb{E}_1^{p(n)}\big(\exp(-sb_1 T_0 - tb_2 M_1) M_1\big) \xrightarrow[n \to \infty]{} s \int_0^\infty \int_0^\infty e^{-sx - ty} \bar{\vartheta}\left(x \wedge \frac{y}{c_\vartheta}\right) dx\, dy$$

$$- st \int_0^\infty \int_0^\infty y e^{-sx - ty} \bar{\vartheta}\left(x \wedge \frac{y}{c_\vartheta}\right) dx\, dy.$$

Computing the integrals, we get

$$\mathbb{E}_1^{p(n)}\big(\exp(-sb_1 T_0 - tb_2 M_1) M_1\big) \xrightarrow[n \to \infty]{} \int_0^\infty e^{-(s + c_\vartheta t)z} z \vartheta\,(dz). \tag{5.5}$$

This finishes the proof. $\qquad\square$

We can now continue with the proof of Theorem 1.7. We assume that (5.2) holds for $k$ and verify it also holds for $k + 1$. Let $\mathcal{F}_k = \sigma((M_{j-1}, T_j), j \le k)$ and $(T_0', M_1')$ be an independent copy of $(T_0, M_1)$. Recalling the definition of the measure $\mathbb{P}_a^\uparrow$ in Theorem 1.3 we have for any $a \in \mathbb{N}$ and $p > 0$ that,

$$\mathbb{E}_a^{p\uparrow}\big(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_k T_k - \theta_{k+1} M_{k+1}}\big)$$

$$= \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_{k-1} T_{k-1} - \theta_k M_k} \frac{1}{a} \mathbb{E}_{M_k}^p\big(e^{-\lambda_k T_0' - \theta_{k+1} M_1'} M_1'\big)\right).$$

Then applying the identity (5.3) we get

$$\mathbb{E}_a^{p\uparrow}\big(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_k T_k - \theta_{k+1} M_{k+1}}\big)$$

$$= \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_{k-1} T_{k-1} - \theta_k M_k} \frac{M_k}{a} \big(\mathbb{E}_1^p\big(e^{-\lambda_k T_0' - \theta_{k+1} M_1'}\big)\big)^{M_k - 1}\right.$$

$$\left. \times \mathbb{E}_1^p\big(e^{-\lambda_k T_0' - \theta_{k+1} M_1'} M_1'\big)\right).$$

Using the Markov property of $\{(T_n, M_{n+1}) : n \in \mathbb{Z}_+\}$ and writing the terms suitably, we get

$$\mathbb{E}_a^{p\uparrow}\big(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_k T_k - \theta_{k+1} M_{k+1}}\big)$$

$$= e^{-\frac{1}{n} \log \mathbb{E}_n^p (e^{-\lambda_k T_0' - \theta_{k+1} M_1'})} \mathbb{E}_1^p\big(e^{-\lambda_k T_0' - \theta_{k+1} M_1'} M_1'\big)$$

$$\times \mathbb{E}_a^p\left(e^{-\lambda_0 T_0 - \theta_1 M_1} \cdots e^{-\lambda_{k-1} T_{k-1}} e^{-(\theta_k - b_2 \log \mathbb{E}_{b_2^{-1}}^{p(n)} (e^{-\lambda_k T_0' - \theta_{k+1} M_1'})) M_k} \frac{M_k}{a}\right).$$

In the previous equality, we have $(\lambda_{i-1}, \theta_i) = (b_1 s_{i-1}, b_2 t_i)$, $i = 1, \ldots, k+1$ and $(a, p) = (a(n), p(n))$, and the we use hypotheses (1.11) and (1.21) to obtain

$$
\mathbb{E}_{a(n)}^{p(n)\uparrow} \left( \prod_{i=1}^{k+1} e^{-b_1 s_{i-1} T_{i-1} - b_2 t_i M_i} \right)
$$

$$
\sim e^{-b_2 \log \mathbb{E}_{b_2-1}^{p(n)} (e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'})} \mathbb{E}_1^{p(n)} \left( e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'} M_1' \right)
$$

$$
\times \mathbb{E}_{a(n)}^{p(n)\uparrow} \left( e^{-b_1 s_0 T_0 - b_2 t_1 M_1} \cdots e^{-b_1 s_{k-1} T_{k-1} - b_2 (t_k + \kappa (s_k + c t_{k+1})) M_k} \right).
$$

Now we have to calculate the limit of each factor. The first one converges towards to 1 thanks to (5.4). Besides, to get

$$
\mathbb{E}_1^{p(n)} \left( e^{-b_1 s_k T_0' - b_2 t_{k+1} M_1'} M_1' \right) \xrightarrow[n\to\infty]{} \iota\big(\kappa_0(s_k + c_\vartheta t_{k+1})\big), \tag{5.6}
$$

we use the convergence (5.5) together with the convention $\kappa_0(s) = s$. As in Proposition 4.5, in order to conclude we use the induction hypothesis with $s_{i-1}' + t_i'$, $1 \le i \le k$ as defined in (4.18).

# Appendix A:  Proof of Lemma 4.1

A consequence of (1.19) is that the measure defined on $[0, \infty)$ by

$$
\mu(x) := \int_0^x z \bar{\pi}^+(z) \, dz, \qquad x \ge 0. \tag{A.1}
$$

is such that $x \mapsto \mu(x)$ is $RV_\infty^{2-\alpha}$. Then from the Tauberian–Abelian theorem (see Theorem 1.7.1 in [4]), its Laplace transform $\mathcal{L}_\mu \in RV_0^{-(2-\alpha)}$ and

$$
\mu(x) \sim \frac{1}{\Gamma(3-\alpha)} \mathcal{L}_\mu(1/x), \qquad x \to \infty.
$$

Observe that

$$
\lambda^2 \mathcal{L}_\mu(\lambda) = \mathbb{E}\big(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+}\big), \qquad \lambda \to 0. \tag{A.2}
$$

As consequence of the definition of the measure $\mu$ and the approximations above,

$$
\bar{\pi}^+(1/\lambda) \sim c_\alpha \mathbb{E}\big(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+}\big), \qquad \lambda \to 0, \tag{A.3}
$$

where $c_\alpha = 1/\Gamma(3-\alpha)$. Hence for all $x > 0$,

$$
\frac{1}{\mathbb{E}(1 - e^{-\lambda \xi^+} - \lambda \xi^+ e^{-\lambda \xi^+})} \bar{\pi}^+(x/\lambda) \sim \frac{\bar{\pi}^+(x/\lambda)}{c_\alpha \bar{\pi}^+(1/\lambda)} \xrightarrow[\lambda \to 0]{} c_\alpha x^{-\alpha}. \tag{A.4}
$$

We set $r(n) = (\mathbb{E}[1 - e^{-\xi^+/n} - \xi^+ e^{-\xi^+/n}/n])^{-1}$ and define the measure on $(0, \infty)$, $m_n(dy) = r(n)\bar\pi^+(n\,dy)$. The convergence in (A.4) implies

$$m_n(x, \infty) \xrightarrow[n\to\infty]{} \int_x^\infty \frac{c_\alpha}{\alpha} \frac{dy}{y^{1+\alpha}} \qquad \text{for all } x > 0.$$

Therefore, for all $0 < x \le y \le \infty$

$$m_n(x, y] \xrightarrow[n\to\infty]{} \int_x^y \frac{c_\alpha}{\alpha} \frac{dz}{z^{1+\alpha}}.$$

This implies that the measure on $(0, \infty)$ defined by $m_n(dy) = r(n)\bar\pi^+(n\,dy)$ converges vaguely towards $c_\alpha \frac{dy}{y^{1+\alpha}}$. We also have

$$\int y^2 \mathbb{1}_{\{y \le x\}} r(n)\pi^+(n\,dy) \xrightarrow[n\to\infty]{} c_\alpha \int y^2 \mathbb{1}_{\{y \le x\}} \frac{dy}{y^{1+\alpha}}.$$

Using an argument of monote class to deduce the above convergence over intervals $I \subset (0, \infty)$. Thus, we obtain the convergence of the Laplace transform of the measure $\mu$. This complete the proof because $\mu$ is regularly varying at infinity with indices $2 - \alpha$ and its Laplace transform satisfies the identity (A.2).

# Appendix B: Proof of Proposition 4.4

We prove the statement for clones, the mutants case is similar. First, note that in the same way as in the proof above,

$$\mu^{cn}(x) = \int_0^x s\bar\pi^{cn}(s)\,ds, \qquad x \ge 0, \tag{B.1}$$

is a measure on $[0, \infty)$ with Laplace transform $\mathcal{L}_{\mu^{cn}}$ such that

$$\lambda^2 \mathcal{L}_{\mu^{cn}}(\lambda) = \mathbb{E}\big(1 - e^{-\lambda\xi^{(cn)}} - \lambda\xi^{(cn)}e^{-\lambda\xi^{(cn)}}\big), \qquad \lambda \ge 0.$$

We now replace $\lambda$ by a sequence $\{\lambda(n) : n \ge 0\}$ such that $\lambda(n) \to 0$ as $n \to \infty$,

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}\big(\lambda(n)\big) = 1 - \phi_n^c\big(\lambda(n)\big) + \lambda(n)\big(\phi_n^c\big)'\big(\lambda(n)\big).$$

From (4.2) we have an estimate of the term $\phi_n^c$. In order to estimate $(\phi_n^c)'$ we use the fact that for every fixed $n$, conditionally to $\xi^{(+)} = k$ the distribution of $\xi^{(cn)}$ is Binomial with parameter $(k, 1 - p(n))$. Then we apply the same techniques as in Lemma 4.2 to get the following estimate

$$\big(\phi_n^c\big)'\big(\lambda(n)\big) \sim \big(1 - p(n)\big)e^{-\lambda(n)}\phi^{+\prime}\big(\big(1 - p(n)\big)\big(1 - e^{-\lambda(n)}\big)\big), \qquad n \to \infty.$$

Putting both estimates together we infer:

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}\big(\lambda(n)\big) \sim 1 - \phi^+\big((1-p(n))(1-e^{-\lambda(n)})\big)$$
$$+ \big(1-p(n)\big)\big(1-e^{-\lambda(n)}\big)\phi_n^{+\prime}\big((1-p(n))(1-e^{-\lambda(n)})\big)$$
$$- \big(1-p(n)\big)\big(1-e^{-\lambda(n)}-\lambda(n)e^{-\lambda(n)}\big)\phi_n^{+\prime}\big((1-p(n))(1-e^{-\lambda(n)})\big),$$

as $n \to \infty$. Furthermore,

$$\big(1-p(n)\big)\big(1-e^{-\lambda(n)}-\lambda(n)e^{-\lambda(n)}\big)\phi_n^{+\prime}\big((1-p(n))(1-e^{-\lambda(n)})\big) \xrightarrow[n\to\infty]{} 0,$$

and from (A.2) we have

$$\lambda(n)^2 \mathcal{L}_\mu\big(\lambda(n)\big) = 1 - \phi^+\big(\lambda(n)\big) + \lambda(n)\phi^{+\prime}\big(\lambda(n)\big), \qquad n \to \infty,$$

where $\mathcal{L}_\mu$ is the Laplace transform of the measure $\mu$ defined in (A.1). From these last two displays, we obtain

$$\lambda(n)^2 \mathcal{L}_{\mu^{cn}}\big(\lambda(n)\big) \sim \big((1-p(n))(1-e^{-\lambda(n)})\big)^2 \mathcal{L}_\mu\big((1-p(n))(1-e^{-\lambda(n)})\big) + O\big(\lambda(n)^2\big),$$

as $n \to \infty$. Due to the estimate $\lambda(n) \sim 1 - e^{-\lambda(n)}$ as $n \to \infty$, the approximation of $\mathcal{L}_\mu$ given in (A.3) implies

$$c_\alpha \lambda(n)^2 \mathcal{L}_{\mu^{cn}}\big(\lambda(n)\big) \sim \bar{\pi}^+\left(\frac{1}{\lambda(n)(1-p(n))}\right) + O\big(\lambda(n)^2\big), \qquad n \to \infty. \qquad \text{(B.2)}$$

Hence, it is remains to prove

$$\lim_{n\to\infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} = c_\alpha. \qquad \text{(B.3)}$$

In this aim, define for every $y \geq 0$, the following measure

$$m_{\lambda(n)}^{cn}(0,y] := m_{\lambda(n)}^{cn}(y) = \frac{\mu^{cn}(y/\lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))}, \qquad \forall y > 0.$$

Observe that

$$\int_{[0,\infty)} e^{-\theta s} d_y\left(\frac{\mu^{cn}(y/\lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))}\right) = \frac{1}{\mathcal{L}_{\mu^{cn}}(\lambda(n))} \int_{[0,\infty)} e^{-\theta\lambda(n)y}\mu^{cn}(dy)$$
$$= \frac{\mathcal{L}_{\mu^{cn}}(\theta\lambda(n))}{\mathcal{L}_{\mu^{cn}}(\lambda(n))}, \qquad \forall \theta > 0.$$

From the previous display and the estimate in (B.2), we get

$$\mathcal{L}_{m_{\lambda(n)}^{cn}}(\theta) \xrightarrow[n\to\infty]{} \theta^{-(2-\alpha)}, \qquad \forall \theta > 0.$$

Writing now $\theta^{-(2-\alpha)}$ in terms of the gamma function we have

$$\theta^{-(2-\alpha)} = \frac{1}{\Gamma(2-\alpha)} \int_0^\infty s^{(2-\alpha)-1} e^{-\theta s} \, ds.$$

Since the convergence of the Laplace transform implies the weak convergence of measures (see Theorem 13.1.2 in [7]), we have

$$m_{\lambda(n)}^{cn}(y) \xrightarrow[n\to\infty]{} \frac{1}{\Gamma(2-\alpha)} \int_0^y s^{(2-\alpha)-1} \, ds = \frac{y^{2-\alpha}}{\Gamma(3-\alpha)}. \tag{B.4}$$

Moreover, by the definition of the measure $\mu^{(cn)}$ we can obtain the following inequality for any $y < 1$,

$$\left(1/\lambda(n)\right)^2 y(1-y)\bar{\pi}^{cn}\left(y/\lambda(n)\right) \leq \mu^{cn}\left(1/\lambda(n)\right) - \mu^{cn}\left(y/\lambda(n)\right) \leq \left(1/\lambda(n)\right)^2 (1-y)\bar{\pi}^{cn}\left(1/\lambda(n)\right).$$

Due to (B.4) this implies

$$c_\alpha \frac{1-y^{2-\alpha}}{1-y} \leq \liminf_{n\to\infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} \leq \limsup_{n\to\infty} \frac{\bar{\pi}^{cn}(1/\lambda(n))}{(\lambda(n))^2 \mathcal{L}_{\mu^{cn}}(\lambda(n))} \leq c_\alpha \frac{1-y^{2-\alpha}}{y(1-y)},$$

for all $y < 1$. To conclude, we make $y \uparrow 1$.

## Appendix C: Proof of Lemma 4.8

Before proving Lemma 4.8, we would like to present some basic aspect of functional convergence of stochastic process, further details can be found in [8]. It is well known that the law of a Lévy process $\{X_t : t \geq 0\}$ on $\mathbb{R}^d$ is determined by that of random variable $X_1$, which is infinitely divisible random variable, and according to the Lévy–Khintchine formula has characteristic exponent

$$\Psi(\mathbf{u}) = i\mathbf{u} \cdot \mathbf{b} - \frac{1}{2}\mathbf{u} \cdot c\mathbf{u}^T + \int \left(e^{i\mathbf{u}\cdot\mathbf{x}} - 1 - i\mathbf{u}\cdot\mathbf{h}(\mathbf{x})\right)\pi(d\mathbf{x}),$$

where $\mathbf{b} \in \mathbb{R}^d$, $c$ is a $d \times d$ symmetric nonnegative matrix, $\pi$ is a positive measure on $\mathbb{R}^d$ with $\pi(\{\mathbf{0}\}) = 0$ and $\int (1 \wedge |\mathbf{x}|^2)\pi(d\mathbf{x}) < \infty$, $\mathbf{h}$ is a truncation function from $\mathbb{R}^d$ to $\mathbb{R}^d$, that is, bounded measurable satisfying

$$\mathbf{h}(\mathbf{x}) = o(|\mathbf{x}|), \qquad |\mathbf{x}| \to 0.$$

Hence an infinitely divisible distribution, and therefore a Lévy process, is uniquely characterized by the triple $(\mathbf{b}, c, \pi)$. Another useful related quantity is a $d \times d$ symmetric nonnegative matrix, called the modified second characteristic, and defined as follows

$$\tilde{c}^{ij} = c^{ij} + \int h^i(\mathbf{x})h^j(\mathbf{x})\pi(d\mathbf{x}), \qquad i, j = 1, 2, \ldots, d.$$

1606 A. Blancas and V. Rivero

According to Theorem VII.2.9 of [8], if $\{\pi_n : n \in \mathbb{N}\}$ is a sequence of infinitely divisible distributions on $\mathbb{R}^d$. Then $\pi_n$ converges weakly to $\pi$ if and only if

$$\mathbf{b}_n \to \mathbf{b},$$
$$\tilde{c}_n \to \tilde{c},$$
$$\pi_n(\mathbf{g}) \to \pi(\mathbf{g}) \qquad \text{for all } \mathbf{g} \in C_1(\mathbb{R}^d),$$

where $C_1(\mathbb{R}^d)$ is a convergence-determining class for the weak convergence induced by all continuous bounded nonnegative functions $\mathbb{R}^d \to \mathbb{R}$, vanishing at the origin and with limit at infinity. We will assume furthermore that $\mathbf{h}$ is a continuous function and that it is the same function for all the independent distributions considered here.

In a more general sense, a $d$-dimensional semimartingale $W$, has associated a characteristic triplet $(B, C, \nu)$ consisting in:

-   $B = (B^i)_{i \leq d}$ a predictable process with components of finite variation over each interval $[0, t]$.
-   $C = (C^{ij})_{i,j \leq d}$ a continuous process, namely

    $$C^{ij} = \langle W^{i,c}, W^{j,c} \rangle,$$

    where $W^c$ is the continuous martingale part of $W$.
-   $\nu$ a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$.

A second modified characteristic $\tilde{C}$ is also defined,

$$\tilde{C}_t^{ij} = C_t^{ij} + (h^i h^j) * \nu_t - \sum_{s \leq t} \left( \int h^i(\mathbf{x})\nu(\{s\} \times d\mathbf{x}) \right) \left( \int h^j(\mathbf{w})\nu(\{s\} \times d\mathbf{w}) \right).$$

If $W$ has no fixed times of discontinuity, in which case $B$ is continuous, and $|h(x)|^2 * \nu_t < \infty$, it reduces to

$$\tilde{C}_t^{ij} = C_t^{ij} + (h^i h^j) * \nu_t.$$

According to Theorem VII.3.4 of [8], the necessary and sufficient conditions to assure the functional convergence of a sequence of semimartingales $W^n$ towards $W$ are given also in terms of their characteristics:

$$\sup_{s \leq t} |B_s^n - B_s| \to 0 \qquad \text{for all } t \geq 0,$$
$$\tilde{C}^n \to \tilde{C} \qquad \text{for all } t \in D, \tag{C.1}$$
$$\mathbf{g} * \nu_t^n \to \mathbf{g} * \nu_t \qquad \text{for all } t \in D, \mathbf{g} \in C_1(\mathbb{R}^d),$$

where $D$ is a dense subset of $\mathbb{R}_+$.

We now turn to prove the convergence claimed in Proposition 4.8. In order to apply the Theorem VII.3.4 of [8], first we will prove the convergence of the characteristics of the process

$$\tilde{\mathbf{S}}^n_{N_{r(n)t}} - (r(n)t/n, 0), \qquad t \geq 0, \tag{C.2}$$

where

$$\tilde{\mathbf{S}}^n_k = \sum_{i=1}^{k} (\xi_i^{(cn)}/n, \xi_i^{(mn)}/r(n)p(n)), \qquad k \in \mathbb{N},$$

and $\{N_t : t \geq 0\}$ is a Poisson process with parameter one, independent of the sequence

$$\xi^{(n)} = \left\{ (\xi_k^{(cn)}, \xi_k^{(mn)}) : k \in \mathbb{Z}_+ \right\}.$$

The following lemma establishes the previous statement. We will use this result as a device to study the characteristics of $\bar{\mathbf{S}}^n_{\lfloor r(n)t \rfloor}$, which are closely related to those of $\tilde{\mathbf{S}}^n_{N_{r(n)t}} - (r(n)t/n, 0)$.

**Lemma C.1.** *The process defined in* (C.2) *is a semimartingale with characteristics relatives to a continuous truncation function* **h** *given by*

$$\mathbf{b}_t^n = r(n)t\mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] - \big(r(n)t/n, 0\big),$$

$$c_t^{n,ij} = 0, \tilde{c}_t^{n,ij} = r(n)t\mathbb{E}\big[h_i\big(b(n)\xi^{(n)}\big)h_j\big(b(n)\xi^{(n)}\big)\big], \qquad i, j = 1, 2, \tag{C.3}$$

$$F_t^n(d\mathbf{x}) = r(n)t\pi(d\mathbf{x}),$$

*where* $b(n)\xi^{(n)} = (\xi^{(cn)}/n, \xi^{(mn)}/r(n)p(n))$, $\pi^{(n)}(d\mathbf{x}) = \mathbb{P}(\xi^{(cn)} \in dx_1, \xi^{(mn)} \in dx_2)$. *Moreover, in the regime determined by* (1.19) *and* (1.21), *we have the following weak convergence in the sense of Skorohod topology*

$$\big(\{\tilde{\mathbf{S}}^n_{N_{r(n)t}} - \big(r(n)t/n, 0\big) : t \geq 0\}, \mathbb{P}^{p(n)}_{a(n)}\big) \implies \{(X_t, t) : t \geq 0\}, \tag{C.4}$$

*where* $X_t$ *is a spectrally positive* $\alpha$-*stable process with parameter* $\alpha \in (1, 2)$. *In particular, we obtain the convergence of the characteristics in* (C.3) *towards those relatives to* $\{(X_t, t) : t \geq 0\}$ *and characteristic exponent* $c_\alpha |\lambda|^\alpha$, *that is*

$$\mathbf{b}_t = \left( t\left( \int_{(0,\infty)} \lambda\big(h(y) - y\big)c_\alpha y^{-(\alpha+1)} \, dy \right), t \right),$$

$$c_t^{ij} = 0, \qquad \tilde{c}_t^{ij} = \mathbb{E}\big[h_i(X_t)h_j(X_t)\big], \qquad i, j = 1, 2, \tag{C.5}$$

$$F_t(d\mathbf{x}) = t c_\alpha x_1^{-(\alpha+1)} \, dx_1 \delta_0(dx_2).$$

**Proof of Lemma C.1.** Note that for $\mathbf{u} = (\lambda, \theta) \in \mathbb{R}^2$,

$$\mathbb{E}\big(e^{i\mathbf{u}\cdot\tilde{\mathbf{S}}^n_{N_{r(n)t}}}\big) = e^{r(n)t(\psi_n(\frac{\lambda}{n}, \frac{\theta}{r(n)p(n)})-1)}, \qquad t \geq 0. \tag{C.6}$$

Then the exponent in the right-hand side of the previous equality can be written as follows

$$
t\left[ i\int_{(0,\infty)}\int_{(0,\infty)} \mathbf{u}\cdot\mathbf{h}\big(b(n)\mathbf{x}\big)r(n)\pi^{(n)}(d\mathbf{x}) \right.
$$
$$
\left. + \int_{(0,\infty)}\int_{(0,\infty)} \big(e^{i\mathbf{u}\cdot b(n)\mathbf{x}}-1-i\mathbf{u}\cdot\mathbf{h}\big(b(n)\mathbf{x}\big)\big)r(n)\pi^{(n)}(d\mathbf{x}) \right], \tag{C.7}
$$

where $b(n)\mathbf{x}=(x_1/n, x_2/r(n)p(n))$, $\pi^{(n)}(d\mathbf{x})=\mathbb{P}(\xi^{(cn)}\in dx_1, \xi^{(mn)}\in dx_2)$. From here $\tilde{\mathbf{S}}_{N_{r(n)t}}$ is infinitely divisible, also we can deduce that the characteristics of the process $\{\tilde{\mathbf{S}}^n_{N_{r(n)t}} - (r(n)t/n, 0): t\geq 0\}$ are given by (C.3). Thanks to Theorem II.3.11 of [8] this process is a Lévy process and a semimartingale.

Besides, to get the convergence in (C.4) we shall prove the convergence of the characteristic functions. This fact is verified using Corollary 4.3, together with the fact that conditionally to $\xi^{(+)}=k$ the distribution of $\xi^{(cn)}$ is Binomial with parameter $(k, 1-p(n))$, as well as the assumption that $\xi^+$ has mean 1. Indeed, the expression can be written (C.7) as

$$
t\int_{(0,\infty)}\left( e^{-((1-p(n))(1-e^{i\lambda/n})-p(n)(1-e^{i\theta/r(n)p(n)}))y}-1 \right.
$$
$$
-i\mathbf{u}\cdot\mathbf{h}\left(\frac{y}{n}(1-p(n)), \frac{y}{r(n)p(n)}p(n)\right)\Bigg)r(n)\pi^+(dy)
$$
$$
+ti\int_{(0,\infty)}\left( \mathbf{u}\cdot\mathbf{h}\left(\frac{y}{n}(1-p(n)), \frac{y}{r(n)p(n)}p(n)\right) \right.
$$
$$
\left. -\left(\frac{\lambda}{n}(1-p(n))+\frac{\theta}{r(n)p(n)}p(n)\right)y\right)r(n)\pi^+(dy)
$$
$$
+i\left(\frac{\lambda}{n}(1-p(n))+\frac{\theta}{r(n)p(n)}p(n)\right)r(n)t.
$$

Then by the continuity of $\mathbf{h}$ and the assumptions (1.11) and (1.20), the previous display becomes

$$
\sim t\int_{(0,\infty)}\left( e^{-((1-p(n))(1-e^{i\lambda/n})-p(n)(1-e^{i\theta/r(n)p(n)}))y}-1-i\lambda h\left(\frac{y}{n}\right)\right)r(n)\pi^+(dy)
$$
$$
+ti\int_{(0,\infty)}\left( \lambda h\left(\frac{y}{n}\right)-\left(\frac{\lambda}{n}(1-p(n))+\frac{\theta}{r(n)p(n)}p(n)\right)y\right)r(n)\pi^+(dy)
$$
$$
+i\left(\frac{\lambda}{n}(1-p(n))+\frac{\theta}{r(n)p(n)}p(n)\right)r(n)t,
$$

where $h$ is the truncation function from $\mathbb{R}$ to $\mathbb{R}$ obtained as projection of $\mathbf{h}$ in the second coordinate. Making a change of variables $z = y/n$.

$$
\sim t \int_{(0,\infty)} \left( e^{-((1-p(n))(1-e^{i\lambda/n})-p(n)(1-e^{i\theta/r(n)p(n)}))nz} - 1 - i\lambda h(z) \right) r(n)\pi^+(n\, dz)
$$

$$
+ t i \int_{(0,\infty)} \left( \lambda h(z) - \left( \lambda(1-p(n)) + \frac{\theta}{r(n)p(n)} np(n) \right) z \right) r(n)\pi^+(n\, dz) \tag{C.8}
$$

$$
+ i \left( \frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)} p(n) \right) r(n) t.
$$

Finally, we have the convergence

$$
\mathbb{E}\left( e^{i\mathbf{u}\cdot(\tilde{\mathbf{S}}^n_{N_{r(n)t}} - (r(n)t/n,0))} \right) \xrightarrow[n\to\infty]{} e^{t\left( \int_{(0,\infty)} (e^{i\lambda y} - 1 - i\lambda y)c_\alpha y^{-(\alpha+1)}\, dy \right) + it\theta}, \tag{C.9}
$$

where $c_\alpha$ is a constant depending on $\alpha$ that appears in Lemma 4.1. Indeed, the result in Lemma 4.1 implies the following convergence

$$
\int_{(0,\infty)} \left( e^{-((1-p(n))(1-e^{i\lambda/n})-p(n)(1-e^{i\theta/r(n)p(n)}))ny} - 1 - i\lambda h(y) \right) r(n)\pi^+(n\, dy)
$$

$$
\xrightarrow[n\to\infty]{} t \int_{(0,\infty)} \left( e^{i\lambda y} - 1 - i\lambda h(y) \right) c_\alpha y^{-(\alpha+1)}\, dy,
$$

while the second adding in (C.8) converges towards

$$
it \int_{(0,\infty)} \lambda \left( h(y) - y \right) c_\alpha y^{-(\alpha+1)}\, dy.
$$

To finish, we observe that the assumption that $r(n) \in RV^\alpha_\infty$ with $\alpha \in (1,2)$ and $p(n) \sim cn^{-1}$ implies that

$$
\frac{r(n)p(n)}{n} \xrightarrow[n\to\infty]{} 0.
$$

The final term in (C.9) is such that

$$
i \left( \frac{\lambda}{n}(1-p(n)) + \frac{\theta}{r(n)p(n)} p(n) \right) r(n) t - i \frac{\lambda}{n} r(n) t \xrightarrow[n\to\infty]{} i\theta t,
$$

for all $t \geq 0$. From (C.9) the characteristic of $(X_t, t)$ are given by (C.5). As a consequence of Theorem VII.2.9 of [8], we have the convergence of the characteristics. Finally, the characteristic exponent of $X_t$ is $c_\alpha |\lambda|^\alpha$. $\qquad\square$

We have now all the elements to prove Lemma 4.8.

**Proof of Lemma 4.8.** Thanks to Theorem 2.3.11 of [8], $\tilde{\mathbf{S}}^{(n)}_{\lfloor r(n)t \rfloor}$ is a semimartingale with characteristics relatives to $\mathbf{h} : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$B^n_t = \lfloor r(n)t \rfloor \mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] - \big(\lfloor r(n)t \rfloor/n, 0\big),$$

$$C^{n,ij}_t = 0,$$

$$\tilde{C}^{n,ij}_t = \lfloor r(n)t \rfloor \big(\mathbb{E}\big[h_i\big(b(n)\xi^{(n)}\big)h_j\big(b(n)\xi^{(n)}\big)\big] - \mathbb{E}\big[h_i\big(b(n)\xi^{(n)}\big)\big]\mathbb{E}\big[h_j\big(b(n)\xi^{(n)}\big)\big]\big),$$

$$\mathbf{g} * \nu^n_t = \lfloor r(n)t \rfloor \mathbb{E}\big(\mathbf{g}\big(b(n)\xi^{(n)}\big)\big) \qquad \text{for } \mathbf{g} \geq 0 \text{ measurable and positive,}$$

for $i, j = 1, 2$. As usual $\lfloor \cdot \rfloor$ denotes the floor function.

Now only remains to verify the conditions (C.1) where the limit characteristics agree with these in (C.5). In this direction, we recall that

$$\mathbf{b}^n_t = r(n)t\mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] - \big(r(n)t/n, 0\big),$$

and observe

$$\big|B^n_s - \mathbf{b}_s\big| \leq \big|\lfloor r(n)s \rfloor - r(n)s\big| \mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] + \big|(r(n)s/n, 0) - (\lfloor r(n)s \rfloor/n, 0)\big| + \big|\mathbf{b}^n_s - \mathbf{b}_s\big|.$$

Then using the properties of the floor function, we obtain

$$\big|B^n_s - \mathbf{b}_s\big| \leq \mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] + \big|(s/n, 0)\big| + \big|\mathbf{b}^n_s - \mathbf{b}_s\big|.$$

Thus, by the convergence of $\mathbf{b}^n_t$ established in the previous lemma, together with the fact that $r(n) \to \infty$, we get

$$\sup_{s \leq t}\big|B^n_s - \mathbf{b}_s\big| \leq \mathbb{E}\big[\mathbf{h}\big(b(n)\xi^{(n)}\big)\big] + \big|(t/n, 0)\big| + \big|\mathbf{b}^n_t - \mathbf{b}_t\big| \xrightarrow[n \to \infty]{} 0,$$

hence we have the first condition in (C.1). In order to determine the second one, let $b^{n,i}_t$ be the $i$th coordinate of $\mathbf{b}^n_t$, $i = 1, 2$. Once more, applying the properties of the floor function we have

$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,ij}_t - \frac{1}{r(n)t}b^{n,i}_t b^{n,j}_t - \frac{r(n)t}{n}\mathbb{E}\big[h_2\big(b(n)\xi^{(n)}\big)\big]$$

$$\leq \tilde{C}^{n,ij}_t \leq \tilde{c}^{n,ij}_t + \frac{1 - r(n)t}{(r(n)t)^2}b^{n,i}_t b^{n,j}_t + \frac{1 - r(n)t}{n}\mathbb{E}\big[h_2\big(b(n)\xi^{(n)}\big)\big].$$

Also

$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,22}_t - \frac{1}{r(n)t}\big(b^{n,2}_t\big)^2 \leq \tilde{C}^{n,22}_t \leq \tilde{c}^{n,22}_t + \frac{1 - r(n)t}{(r(n)t)^2}\big(b^{n,2}_t\big)^2$$

and

$$\left(1 - \frac{1}{r(n)t}\right)\tilde{c}^{n,11}_t - \frac{1}{r(n)t}\big(b^{n,1}_t\big)^2 \leq \tilde{C}^{n,11}_t \leq \tilde{c}^{n,11}_t + \frac{1 - r(n)t}{(r(n)t)^2}\big(b^{n,1}_t\big)^2 + b^{n,1}_t + \frac{1}{n^2}\big[r(n) - 1\big].$$

As consequence of the convergence, $\mathbf{b}_t^n \to \mathbf{b}_t$, $b_t^{n,i}$ converges for $i = 1, 2$. Then the above inequalities imply $\tilde{C}_t^{n,ij} \to \tilde{c}_t^{ij}$, because $r(n) \in RV_\infty^\alpha$ and $\alpha \in (1,2)$. It is easily proved that also $\mathbf{g} * \nu_t^n \to \mathbf{g} * F_t$ for all $g$ using that $F_t^n = r(n)t\pi(d\mathbf{x})$ converges to $F_t$, as we proved in the previous lemma, together with properties of the floor function. $\qquad\square$

# Acknowledgements

# References

[1] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes. Die Grundlehren der Mathematischen Wissenschaften* **196**. New York: Springer.

[2] Bertoin, J. (1996). *Lévy Processes. Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press.

[3] Bertoin, J. (2010). A limit theorem for trees of alleles in branching processes with rare neutral mutations. *Stochastic Process. Appl.* **120** 678–697. MR2603059

[4] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR1015093

[5] Chauvin, B. (1986). Arbres et processus de Bellman–Harris. *Ann. Inst. Henri Poincaré B*, *Probab. Stat.* **22** 209–232.

[6] Chung, K.L. and Walsh, J.B. (2005). *Markov Processes*, *Brownian Motion*, *and Time Symmetry*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* **249**. New York: Springer.

[7] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. *Vol. II*, 2nd ed. New York: Wiley. MR0270403

[8] Jacod, J. and Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes. Grundlehren der Mathematischen Wissenschaften* **288**. Berlin: Springer. MR0959133

[9] Joffe, A. (1967). On the Galton–Watson branching process with mean less than one. *Ann. Math. Stat.* **38** 264–266. MR0205337

[10] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer.

[11] Kawazu, K. and Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen.* **16** 34–51.

[12] Lambert, A. (2007). Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.* **12** 420–446.

[13] Lambert, A. (2008). Population dynamics and random genealogies. *Stoch. Models* **24** 45–163. MR2466449

[14] Lamperti, J. and Ney, P. (1968). Conditioned branching processes and their limiting diffusions. *Teor. Verojatnost. i Primenen.* **13** 126–137.

[15] Li, Z. (2011). *Measure-Valued Branching Markov Processes. Probability and Its Applications* (*New York*). Heidelberg: Springer. MR2760602

[16] Seneta, E. and Vere-Jones, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3** 403–434. MR0207047

[17] Whitt, W. (2002). *Stochastic-Process Limits*: *An Introduction to Stochastic-Process Limits and Their Application to Queues*. *Springer Series in Operations Research*. New York: Springer. MR1876437

[18] Yaglom, A.M. (1947). Certain limit theorems of the theory of branching random processes. *Dokl. Akad. Nauk SSSR* **56** 795–798.