SOME PROPERTIES OF STATIONARY CONTINUOUS STATE BRANCHING PROCESSES

ROMAIN ABRAHAM, JEAN-FRANÇOIS DELMAS, AND HUI HE

Abstract. We consider the genealogical tree of a stationary continuous state branching process with immigration. For a sub-critical stable branching mechanism, we consider the genealogical tree of the extant population at some fixed time and prove that, up to a deterministic time-change, it is distributed as a continuous-time Galton-Watson process with immigration. We obtain similar results for a critical stable branching mechanism when only looking at immigrants arriving in some fixed time-interval. For a general sub-critical branching mechanism, we consider the number of individuals that give descendants in the extant population. The associated processes (forward or backward in time) are pure-death or pure-birth Markov processes, for which we compute the transition rates.

1. Introduction

1.1. **State of the art.** Inference of the genealogical tree of some given population (or of a sample of extant individuals) is a central question in evolutionary biology (see for instance [25]) and, to perform this task by the usual maximum likelihood method, the distribution of this genealogical tree must be known.

The most popular model in this context is the Wright-Fisher model where the genealogical tree of a sample of extant individuals is given by the Kingman coalescent [31]. One major feature of this model is to consider a constant size population although many extensions have been proposed to take into account population size change (see e.g. [23]). Other models have also been considered where the distribution of the genealogical tree or a sample of the current population can be explicitly described: linear birth-death process [37], continuous time Galton-Watson trees [24, 28], Brownian tree [3] see also [2], splitting trees [34]. Some recent results on the coalescent process associated with some branching process by time-reversal can be found in [44, 29, 20].

We consider here continuous state branching processes with immigration so that the total population size is stationary. More precisely, let ψ be a branching mechanism of the form

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,+\infty)} \left(\mathrm{e}^{-\lambda r} - 1 + \lambda r\right) \pi(dr), \quad \lambda \geq 0, \tag{1}$$

where $\alpha = \psi'(0) \in \mathbb{R}$, $\beta \geq 0$ and π is a σ-finite measure on $(0, +\infty)$ such that $\int_{(0,+\infty)}(r \wedge r^2)\pi(dr) < +\infty$. We will mainly assume that ψ is sub-critical, that is $\alpha > 0$, and that ψ satisfies

---

Date: February 1, 2021.

2010 *Mathematics Subject Classification.* 60J80, 60J27, 92D25.

Key words and phrases. Continuous state branching process with immigration, quasi-stationary distribution, genealogical tree, ancestral process.

We thank Beijing Normal University, Orléans University and Ecole des Ponts for the invitations during which this paper was written. Hui He also thanks S. Feng for enlightening discussions. This research was partially supported by NSFC (No. 11671041 and 11531001).

the following conditions:

$$\int^{+\infty} \frac{d\lambda}{\psi(\lambda)} < +\infty \quad \text{(Grey condition) and} \quad \int_{0+} \left(\frac{1}{\lambda\alpha} - \frac{1}{\psi(\lambda)} \right) d\lambda < \infty. \tag{2}$$

The Grey condition implies in particular that $\beta > 0$ or $\int_{(0,1)} r\pi(dr) = +\infty$; it also implies that the corresponding branching process dies out in finite time. The latter condition is equivalent to the so-called '$L \log L$'-condition, see [32, Lemma 1]; it also implies that the corresponding branching process renormalized by its expectation is a positive closed martingale.

A continuous state branching process (CB process for short) is a positive real valued Markov process $(Y_t, t \geq 0)$ that satisfies the following branching property: the process Y starting from $Y_0 = x + x'$ is distributed as $Y^{(1)} + Y^{(2)}$ where $Y^{(1)}$ and $Y^{(2)}$ are independent copies of Y starting respectively from $Y_0^{(1)} = x$ and $Y_0^{(2)} = x'$. The distribution of the process Y is then uniquely determined by its branching mechanism, see Section 2.1. As we only consider critical or sub-critical branching mechanisms together with Grey condition (2), the population becomes a.s. extinct in finite time. We denote by $c(t)$ the "probability" of non-extinction at time $t > 0$ under the canonical measure (see Section 2.1 below) which is defined by:

$$\int_{c(t)}^{+\infty} \frac{d\lambda}{\psi(\lambda)} = t, \quad t > 0.$$

The extinction probability is then given by $\mathrm{P}_x(Y_t = 0) = \mathrm{e}^{-xc(t)} > 0$ for all $t > 0$, where P_x denotes the distribution of the CB starting from $x \geq 0$ at time 0. The second condition in (2) insures that $\lim_{t\to+\infty} \mathrm{e}^{\alpha t}\, Y_t$ exist P_x-a.s. and that the following limit is well defined:

$$\kappa = \lim_{t\to+\infty} c(t)\, \mathrm{e}^{\alpha t} \in (0, +\infty) \tag{3}$$

where according to Lemma 1 in [32], κ satisfies $c^{-1}(\kappa) = \int_0^\kappa \left(\frac{1}{\alpha\lambda} - \frac{1}{\psi(\lambda)} \right) d\lambda$.

One way to avoid this extinction is to add an immigration characterized by a function ϕ defined on $\mathbb{R}^+$ which describes the intensity of the immigration and the size of the immigrant population, see for example [36] and references therein:

$$\phi(\lambda) = \beta'\lambda + \int_{(0,+\infty)} \left(1 - \mathrm{e}^{-\lambda r} \right) r\pi'(dr), \quad \lambda \geq 0, \tag{4}$$

where $\beta' \geq 0$ and π' is a σ-finite measure on $(0, +\infty)$ such that $\int_{(0,+\infty)} (1 \wedge r)\, \pi'(dr) < +\infty$. We shall mainly consider the immigration (which, when the Grey condition holds, appears when conditioning the initial CB process on non-extinction, see [33, 14]) given by $\beta' = 2\beta$ and $\pi' = \pi$:

$$\phi(\lambda) = \psi'(\lambda) - \alpha = 2\beta\lambda + \int_{(0,+\infty)} \left(1 - \mathrm{e}^{-\lambda r} \right) r\pi(dr). \tag{5}$$

We can then consider a CB process with immigration (CBI process for short) indexed by $\mathbb{R}$, $Y = (Y_t,\ t \in \mathbb{R})$, whose one-dimensional distributions are the same in time. Some properties of this process have been investigated in [14]. By convention, the stationary case will correspond to a sub-critical branching mechanism ψ and the corresponding immigration ϕ given by (5). We shall denote by $\bar{u}$ the Laplace transform of Y_t, see (16), which is given by:

$$\bar{u}(\lambda) = \frac{\kappa\alpha\, \mathrm{e}^{-\alpha c^{-1}(\lambda)}}{\psi(\lambda)}, \quad \lambda > 0. \tag{6}$$

The description of the genealogy of CB processes is done using Lévy trees (see [16]), and of CBI processes as a real tree with an infinite spine on which some Lévy trees are grafted. As the population size in our CBI processes is stationary, we can look at the extant population at any fixed time, say $t = 0$ in all the paper. We want to describe the distribution of the genealogical tree of this extant population. A complete description of this genealogy is already done in [2] for a quadratic branching mechanism $\psi(\lambda) = \alpha\lambda + \beta\lambda^2$ together with the description of the genealogical tree of a sample of the extant population. We focus in this paper on general branching mechanisms.

1.2. **Main results.** For a general sub-critical branching mechanism ψ, the description of the genealogy of the extant population, in the stationary case, can be seen as a birth process (forward in time) and a death process (backward in time) coming from infinity. Let $1+M_t^0$ be the number of descendants of the extant population forward in time at time $t \in (-\infty, 0)$. Notice that a.s. $\lim_{t\to-\infty} M_t^0 = 0$. The ancestral process $(M_t^0, t < 0)$ describes in some sense the genealogy of the extant population at time 0. Asymptotics of M_t^0 as t increases to 0 are given in [14] (see also references therein for related results on coalescent processes). Recall that for a birth and death Markov process, its distribution is completely characterized by its birth and death rates, that is infinitesimal generator, see [13, Section 1.1]. We have the following result, see Propositions 5.2 and 5.4.

Theorem 1.1. *Assume ψ given by (1) is sub-critical (i.e. $\alpha > 0$) and satisfies conditions (2) and ϕ is given by (5).*

(i) *The forward in time process $(M_t^0, t < 0)$ is a càd-làg inhomogeneous pure birth Markov process starting from* 0 *at time* $-\infty$ *with birth rate given by for $m > n \geq 0$ and $t > 0$:*

$$q^{\mathrm{b}}_{n,m}(-t) = \frac{(m+1)}{(m+1-n)!}\, c(t)^{m-n} \left|\psi^{(m-n+1)}\big(c(t)\big)\right|.$$

(ii) *The backward in time process $(M^0_{(-t)-}, t > 0)$ is a càd-làg inhomogeneous pure death Markov process starting from $+\infty$ at time* 0*, with death rate given by for $n > m \geq 0$ and $t > 0$:*

$$q^{\mathrm{d}}_{n,m}(t) = \binom{n+1}{m} \frac{\left|\bar{u}^{(m)}\big(c(t)\big)\right|}{\left|\bar{u}^{(n)}\big(c(t)\big)\right|} \left|\psi^{(n-m+1)}\big(c(t)\big)\right|.$$

We now consider a stable branching mechanism:

$$\psi(\lambda) = \alpha\lambda + \gamma\lambda^b \tag{7}$$

with $\alpha > 0$, $\gamma > 0$ and $b \in (1, 2]$. The case $b = 2$ corresponds to $\pi = 0$ in (1), and the case $b \in (1, 2)$ corresponds to $\beta = 0$ and $\pi(dr)$ equal (up to a multiplicative constant) to $r^{-b-1}\, dr$. See Remark 5.5 for an explicit computation of the birth rate for $1 < b < 2$. See also Remark 5.1 for an explicit computation of the birth and death rates in the quadratic case $b = 2$, which already appears in Proposition 3.2 and 3.3 in [10].

We now present a deterministic time change for which the genealogy of the extant population (forward in time) in the stationary case is a time homogeneous Galton-Watson process with immigration. The time change relies on the non-extinction "probability" $c(t)$ of the associated CB process under the canonical measure which is given (see Example 3.1 p. 62 in [35] where $\bar{v}_t$ corresponds to $c(t)$ in our setting) for $t > 0$ by:

$$c(t) = \left(\frac{\alpha}{\gamma\left(\mathrm{e}^{(b-1)\alpha t} - 1\right)}\right)^{\frac{1}{b-1}}. \tag{8}$$

We consider the time change $T(t) = -R^{-1}(t)$ where:

$$R(t) = \log\left(\frac{\tilde{\psi}(c(t))}{\tilde{\psi}(0)}\right) \quad \text{with} \quad \tilde{\psi}(\lambda) = \frac{\psi(\lambda)}{\lambda} = \alpha + \gamma\lambda^{b-1} \tag{9}$$

and we consider the process $\tilde{M} = (\tilde{M}_t = M^0_{T(t)}, t > 0)$. The main result of the paper is the following theorem.

Theorem 1.2. *Assume ψ is given by (7) (with $\alpha > 0$ and $b \in (1, 2]$) and ϕ by (5). The time-changed ancestral process $\tilde{M}$ is distributed as a continuous-time Galton-Watson process with immigration.*

The characteristics of the Galton-Watson process with immigration (length of the branches, immigration rate, offspring distribution, immigration size) are precised in Theorem 4.1. This process may also be viewed as a sized-biased continuous-time Galton-Watson process, see Remark 4.2.

As a corollary of this theorem, we study the sizes of the families of the extant population ranked according to their immigration time. The vector of the sizes of these families in the stable case is distributed as the jumps of a time-changed subordinator which yields a Poisson-Kingman distribution (see Remark 4.11). In the quadratic case (see Corollary 4.12) this corresponds to a Poisson-Dirichlet distribution. The computations of Proposition 4.14 prove that for $b \in (1, 2)$, the distribution of the sizes of these families is not a Poisson-Dirichlet distribution since a sized-biased sample of the vector of sizes is not Beta-distributed (except maybe for one very particular case).

In the stable critical case $\psi(\lambda) = \lambda^b$ with $b \in (1, 2]$, the previous results do not make sense since the total population size is always infinite and the ancestral process is trivially infinite at all times. To get a finite extant population, we restrict our attention to the extant individuals whose initial immigrant arrived after some fixed time $-T$. Theorem 1.2 remains valid in this setting with a different change of time, see Theorem 6.1.

If the Grey condition is not satisfied, it is always possible to define the genealogy of a CBI process whose immigration mechanism is given by ϕ in (5), see Corollary 3.3 in [14]. However, the ancestral process is again trivially infinite at all time. This is for example the case for the Neveu's branching mechanism $\psi(\lambda) = \lambda \log(\lambda)$ which appears as the natural limit of the stable branching mechanism $\psi(\lambda) = \lambda^b$ when b goes down to 1. There is a natural link between the CB with Neveu's branching mechanism and the Bolthausen-Sznitman coalescent, see [9]. Inspired by this result, the following result, see Proposition 5.6, gives that looking backward the genealogical tree in the stationary stable case, one recovers, as b decreases to 1, the Bolthausen-Sznitman coalescent. Let $T > 0$ and $n \geq 1$. Conditionally on the number of ancestors at time $-T$ of the extant population being n, that is on $\{M^0_{-T} = n - 1\}$, we label them from 1 to n uniformly at random. Define a continuous time process $(\Pi^{T,[n]}(t), t \geq T)$ taking values in the partitions of $[n] = \{1, 2, \cdots, n\}$, by $\Pi^{T,[n]}(t)$ is the partition of $[n]$ such that i and j are in the same block if and only if the i-th and j-th individuals at level $-T$ have the same ancestor at level $-t$.

Theorem 1.3. *Assume ψ is given by (7) (with $\alpha > 0$ and $b \in (1, 2]$) and ϕ by (5). The law of $(\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}), t \geq 0)$ conditionally on $\{M^0_{-T} = n-1\}$, converges in the sense of finite dimensional distribution to a Bolthausen-Sznitman coalescent as b decreases to 1.*

1.3. **Organisation of the paper.** Sections 2 and 3 are respectively devoted to recall known results on CB process and their genealogy using real trees, and on CBI process and the definition of the number of descendants of the extant population $M = (M_t^0, t < 0)$. For the stable sub-critical setting, we present in Section 4 the proof of Theorem 4.1 (and thus of Theorem 1.2) and the study of the sizes of the families of the extant population ranked according to their immigration time. We compute the birth and death rates of the process M (Theorem 1.1) in Section 5 and apply these expressions in Sub-section 5.3 to prove the convergence of the ancestral process as b goes down to 1 towards the Bolthausen-Sznitman coalescent (Theorem 1.3). We provide some results in Section 6 for the critical stable branching mechanism.

2. Notations

Concerning probability measures and expectations, we shall use P and E for usual real random variables or processes, $\mathbb{P}$ and $\mathbb{E}$ for Lévy trees or Lévy forest, and $\bar{\mathbb{P}}$ and $\bar{\mathbb{E}}$ for the corresponding stationary cases which involve immigration.

The process Y usually refer to a CB or CBI process (under P) and Z usually refer to a CB (under $\mathbb{P}$) or a CBI (under $\bar{\mathbb{P}}$) built on a Lévy tree or a Lévy forest.

We write $\mathbb{N} = \{0, 1, \dots\}$ for the set of integers and $\mathbb{N}^* = \{1, 2, \dots\}$. for the set of positive integers. We denote by δ_a the Dirac mass at a.

2.1. **Continuous state branching processes.** We refer to [11, 22, 36] for a presentation and general results on CB processes. We recall that a CB process with branching mechanism ψ (denoted CB(ψ)) is a càd-làg non-negative real-valued Markov process $Y = (Y_t,\ t \geq 0)$ whose transition kernels are characterized, for every $s, t, \lambda \geq 0$, by

$$\mathrm{E}\left[\mathrm{e}^{-\lambda Y_{s+t}} \,\middle|\, Y_s\right] = \mathrm{e}^{-u(t,\lambda)\, Y_s}, \tag{10}$$

where $(u(\lambda, t); t \geq 0, \lambda \geq 0)$ is the unique non-negative solution of the integral equation

$$u(\lambda, t) + \int_0^t \psi\big(u(\lambda, s)\big)\, ds = \lambda, \tag{11}$$

or equivalently the unique non-negative solution of the integral equation

$$\int_{u(\lambda,t)}^{\lambda} \frac{dr}{\psi(r)} = t. \tag{12}$$

When the Grey condition holds, see (2), we define the finite function c on $(0, +\infty)$ by:

$$c(t) = u(+\infty, t) = \lim_{\lambda \to +\infty} u(\lambda, t). \tag{13}$$

Let P_r denote the distribution of the CB process Y starting from $r \geq 0$ at time 0 (that is conditionally on $\{Y_0 = r\}$) and E_r the corresponding expectation operator. We denote by N the canonical (infinite but σ-finite) measure of the CB process Y on the Skorohod space $D(\mathbb{R}_+)$ of real-valued càd-làg function defined on $\mathbb{R}_+$; so that if $\sum_{i \in I} \delta_{Y_i}$ is a Poisson point measure on $D(\mathbb{R}_+)$ with intensity $r\mathrm{N}(dY)$, with $r \in \mathbb{R}_+$, then the process $(\tilde{Y}_t, t \geq 0)$ defined by

$$\tilde{Y}_t = \sum_{i \in I} Y_t^i$$

is distributed as Y under P_r, see [35, Theorem 8.24] in the more general framework of measure-valued branching processes. (Notice the set of indices I is countable.) In particular, we have for $\lambda, t \geq 0$:

$$\mathrm{N}\left[1-\mathrm{e}^{-\lambda Y_t}\right] = \lim_{r\to 0} \frac{1}{r}\mathrm{E}_r\left[1-\mathrm{e}^{-\lambda Y_t}\right] = u(\lambda, t).$$

When the Grey condition holds, the function $c(t)$ satisfies for $t > 0$:

$$c(t) = \mathrm{N}[Y_t > 0], \quad u(c(t), s) = c(t+s) \quad \text{and} \quad c'(t) = -\psi(c(t)).$$

2.2. **Continuous branching process with immigration.** We define the stationary CBI processes with branching mechanism ψ and immigration ϕ, given respectively by (1) and (4), as follows. Let $\sum_{i\in I} \delta_{(h_i, Y_i)}$ be a Poisson point measure on $\mathbb{R} \times D(\mathbb{R}_+)$ with intensity

$$\nu_0(dh, dY) = dh\left(\beta'\mathrm{N}(dY) + \int_{(0,+\infty)} \mathrm{P}_r(dY)\, \pi'(dr)\right), \tag{14}$$

and define, with the convention that $Y^i_s = 0$ for $s < 0$, the process $(Y_t, t \in \mathbb{R})$ by:

$$Y_t = \sum_{i\in I} Y^i_{t-h_i}.$$

Using the master formula for Poisson point measure, we get that for $\lambda \geq 0$ and $t \in \mathbb{R}$:

$$\mathrm{E}\left[\mathrm{e}^{-\lambda Y_t}\right] = \exp\left(-\int_0^{+\infty} \phi(u(\lambda, r))\, dr\right),$$

where u is still the function given by (12). If for some (and thus for all) $\lambda > 0$, we have:

$$\int_0^{+\infty} \phi(u(\lambda, r))\, dr < +\infty, \tag{15}$$

then we get that Y_0 is a.s. finite. Following [35, Section 9.5], we then obtain that $(Y_t, t \geq 0)$ is a càd-làg Markov process started at Y_0 whose transition kernels are characterized, for $s, t \in \mathbb{R}$, $\lambda \geq 0$, by:

$$\mathrm{E}\left[\mathrm{e}^{-\lambda Y_{s+t}} \,\middle|\, Y_s\right] = \exp\left(-u(\lambda, t)Y_s - \int_0^t \phi\big(u(\lambda, r)\big)\, dr\right).$$

The so called CBI process started at $r \in \mathbb{R}$ with branching mechanism ψ and inmmigration mechanism ϕ is defined as $(Y'_t + \sum_{i\in I} Y^i_{t-t_i}\mathbf{1}_{\{t_i>0\}}, t \geq 0)$, where $(Y'_t, t \geq 0)$ is a CB with branching mechanism ψ started from r at time 0 independent of the Poisson point measure $\sum_{i\in I} \delta_{(h_i, Y_i)}$. We refer to [30] for more results on CBI processes. Notice that at this stage no hypothesis are required on ψ and ϕ. Now, provided that Y_0 is a.s. finite, we get that $(Y_t, t \geq 0)$ is a CBI started at (its stationary measure) Y_0.

By translation invariance of the Lebesgue measure dt, we deduce that if (15) holds, then the processes $(Y_{t+t_0}, t \geq 0)$, for $t_0 \in \mathbb{R}$ have all the same distribution. This gives that $(Y_t, t \in \mathbb{R})$ is a $[0, +\infty)$-valued stationary càd-làg Markov process; and we call it the stationary CBI processes with branching mechanism ψ and immigration ϕ.

When one considers a sub-critical branching mechanism ψ satisfying (2) and the particular immigration mechanism ϕ given by (5), we get according to [14, Corollary 3.13] and the definition of $\bar{u}(\lambda)$ given in (6) that for $\lambda > 0$:

$$\int_0^{+\infty} \phi(u(\lambda, r))\, dr = -\log(\bar{u}(\lambda)) < +\infty,$$

and the process $(Y_t, t \geq 0)$ conditioned on $\{Y_0 = r\}$ with $r > 0$, can be viewed as the CB process with branching mechanism ψ started at r and conditioned on non-extinction; this observation motivates the particular choice for this immigration mechanism. In this particular case, we have for every $\lambda \geq 0$ and $t \in \mathbb{R}$,

$$\mathrm{E}\left[\mathrm{e}^{-\lambda Y_t}\right] = \bar{u}(\lambda). \tag{16}$$

2.3. **Real trees and Lévy trees.** We refer to [15, 18] for general results on real trees and to [17] for Lévy trees. We recall that a metric space $(\mathbf{t}, d)$ is a real tree if the following two properties hold for every $u, v \in \mathbf{t}$.

(i) There is a unique isometric map $f_{u,v}$ from $[0, d(u,u)]$ into $\mathbf{t}$ such that

$$f_{u,v}(0) = u \qquad \text{and} \qquad f_{u,v}\big(d(u,v)\big) = v.$$

(ii) If φ is a continuous injective map from $[0,1]$ into $\mathbf{t}$ such that $\varphi(0) = u$ and $\varphi(1) = v$, then the range of φ is also the range of $f_{u,v}$.

The range of the map $f_{u,v}$ is denoted $[\![u, v]\!]$. It is the unique continuous path that links u to v in the tree. In order to simplify the notations, we often omit the distance d in the notation and say that $\mathbf{t}$ is a real tree.

A rooted real tree is a real tree $(\mathbf{t}, d)$ with a distinguished vertex ∂ called the root. Two real trees (resp. rooted real trees) $\mathbf{t}_1$ and $\mathbf{t}_2$ are called equivalent if there is an isometry (resp. a root-preserving isometry) that maps $\mathbf{t}_1$ onto $\mathbf{t}_2$. We set $\mathbb{T}$ the set of all equivalence classes of rooted compact real trees. We endow the set $\mathbb{T}$ with the pointed Gromov-Hausdorff distance (see [18]) and the associated Borel σ-field. The set $\mathbb{T}$ is then Polish.

Let $\mathbf{t} \in \mathbb{T}$ be a rooted tree. We define a partial order $\prec$ (called the genealogical order) on $\mathbf{t}$ by:

$$u \prec v \iff u \in [\![\partial, v]\!] \setminus \{v\}$$

and we say in this case that u is an ancestor of v. The height of a vertex $u \in \mathbf{t}$ is defined by

$$H(u) = d(\partial, u),$$

and we denote by $H(\mathbf{t}) = \sup\{d(\partial, u),\ u \in \mathbf{t}\}$ the height of the tree $\mathbf{t}$. Let $a > 0$. The truncation of $\mathbf{t}$ at level a is the tree $\mathrm{Tr}_a(\mathbf{t}) = \{u \in \mathbf{t},\ H(u) \leq a\}$, and the population of the tree $\mathbf{t}$ at level a is the sub-set

$$\mathcal{Z}_{\mathbf{t}}(a) = \{u \in \mathbf{t},\ H(u) = a\}. \tag{17}$$

We denote by $(\mathbf{t}^{(i),*}, i \in I)$ the connected components of the open set $\mathbf{t} \setminus \mathrm{Tr}_a(\mathbf{t})$. For every $i \in I$, there exists a unique point $\partial_i \in \mathcal{Z}_{\mathbf{t}}(a)$ such that $\partial_i \in [\![\partial, u]\!]$ for every $u \in \mathbf{t}^{(i),*}$. We then set $\mathbf{t}^{(i)} = \mathbf{t}^{(i),*} \cup \{\partial_i\}$ so that $\mathbf{t}^{(i)}$ is a compact rooted real tree with root ∂_i and we consider the point measure on $\mathcal{Z}_{\mathbf{t}}(a) \times \mathbb{T}$:

$$\mathcal{N}_a^{\mathbf{t}} = \sum_{i \in I} \delta_{(\partial_i, \mathbf{t}^{(i)})}.$$

We now recall the definition of the excursion measure associated with a ψ-Lévy tree from [17] which is critical or sub-critical ($\alpha \geq 0$) and satisfies the Grey condition (first condition in (2)). Let ψ be a branching mechanism defined by (1). Then, there exists a measure $\mathbb{N}$ on $\mathbb{T}$ such that:

(i) **Existence of a local time.** For every $a \geq 0$ and for $\mathbb{N}(d\mathcal{T})$-a.e. $\mathcal{T} \in \mathbb{T}$, there exists a finite measure ℓ^a on $\mathcal{T}$ such that
  (a) $\ell^0 = 0$ and, for every $a > 0$, ℓ^a is supported on $\mathcal{Z}_{\mathcal{T}}(a)$.
  (b) For every $a > 0$, $\{\ell^a \neq 0\} = \{H(\mathcal{T}) > a\}$, $\mathbb{N}(d\mathcal{T})$-a.e.

(c) For every $a>0$, we have $\mathbb{N}(d\mathcal{T})$-a.e. for every bounded continuous function φ on $\mathcal{T}$,

$$\begin{aligned}\langle \ell^a, \varphi\rangle &= \lim_{\varepsilon\to 0+} \frac{1}{c(\varepsilon)} \int \mathcal{N}_a^{\mathcal{T}}(du\; d\mathcal{T}')\varphi(u)\mathbf{1}_{\{H(\mathcal{T}')\geq\varepsilon\}}\\ &= \lim_{\varepsilon\to 0+} \frac{1}{c(\varepsilon)} \int \mathcal{N}_{a-\varepsilon}^{\mathcal{T}}(du\; d\mathcal{T}')\varphi(u)\mathbf{1}_{\{H(\mathcal{T}')\geq\varepsilon\}}.\end{aligned}$$

(ii) **Branching property.** For every $a>0$, the conditional distribution of the point measure $\mathcal{N}_a^{\mathcal{T}}(du\; d\mathcal{T}')$, under the probability measure $\mathbb{N}(d\mathcal{T}\,|\,H(\mathcal{T})>a)$ and given $\mathrm{Tr_a}(\mathcal{T})$, is that of a Poisson point measure on $\mathcal{Z}_{\mathcal{T}}(a)\times\mathbb{T}$ with intensity $\ell^a(du)\mathbb{N}(d\mathcal{T}')$.
(iii) **Regularity of the local time process.** We can choose a modification of the process $(\ell^a, a\geq 0)$ in such a way that the mapping $a\longmapsto \ell^a$ is $\mathbb{N}(d\mathcal{T})$-a.e. càd-làg for the weak topology on finite measures on $\mathcal{T}$.
(iv) **Link with CB processes.** Under $\mathbb{N}(d\mathcal{T})$, the process $(\langle \ell^a, 1\rangle, a\geq 0)$ is distributed as a $\mathrm{CB}(\psi)$ process under N.

If necessary, we shall write $\ell^a(\mathcal{T})$ for ℓ^a in order to stress the dependence in the Lévy tree $\mathcal{T}$. We define the population size process as $Z=(Z_a, a\geq 0)$, where the "size" of the population at level a is given by:

$$Z_a = \langle \ell^a(\mathcal{T}), 1\rangle. \tag{18}$$

We recall that under $\mathbb{N}$, the process Z is distributed as Y under the canonical measure N. Let us mention that the super-critical case, which we do not study here, could have been handled using a Girsanov transformation, as in [1].

2.4. **Forests.**

Definition 2.1 (Forest and leveled forest)**.** *A forest is a family $\mathbf{f}=(\mathbf{t}_i)_{i\in I}$, at most countable, of elements of $\mathbb{T}$. A leveled forest is a family $\bar{\mathbf{f}}=(h_i,\mathbf{t}_i)_{i\in I}$, at most countable, of elements of $\mathbb{R}\times\mathbb{T}$. We denote by $\mathbb{F}$ (resp. $\bar{\mathbb{F}}$) the set of (resp. leveled) forests.*

If $\bar{\mathbf{f}}=(h_i,\mathbf{t}_i)_{i\in I}$ is a leveled forest, denoting by d_i the distance in the tree $\mathbf{t}_i$ and ∂_i the root of $\mathbf{t}_i$, we can associate with it a tree $(\mathbf{t}(\bar{\mathbf{f}}),\bar{d})$ by

$$\mathbf{t}(\bar{\mathbf{f}}) = \mathbb{R}\sqcup\left(\bigsqcup_{i\in I}\mathbf{t}_i^*\right)$$

where $\sqcup$ denotes the disjoint union of sets, $\mathbf{t}_i^*=\mathbf{t}_i\setminus\{\partial_i\}$, and, for every $u,v\in\mathbf{t}(\bar{\mathbf{f}})$,

$$\bar{d}(u,v)=\begin{cases}|u-v| & \text{if } u,v\in\mathbb{R},\\ d_i(u,v) & \text{if } u,v\in\mathbf{t}_i^*,\\ |u-h_i|+d_i(\partial_i,v) & \text{if } u\in\mathbb{R} \text{ and } v\in\mathbf{t}_i^*,\\ d_i(\partial_i,u)+|h_i-h_j|+d_j(\partial_j,v) & \text{if } u\in\mathbf{t}_i^*,\ v\in\mathbf{t}_j^* \text{ with } i\neq j.\end{cases}$$

Remark 2.2. It is easy to check that $\mathbf{t}(\bar{\mathbf{f}})$ is indeed a real tree. It is neither rooted nor compact, and can be seen as a tree with a two-sided infinite spine (the set $\mathbb{R}$).

Remark 2.3. If $\mathbf{f}=(h_i,\mathbf{t}_i)_{i\in I}$ and $(h_i,\tilde{\mathbf{t}}_i)_{i\in I}$ are two families of real numbers and real trees such that, for every $i\in I$, the trees $\mathbf{t}_i$ and $\tilde{\mathbf{t}}_i$ are equivalent, then the trees constructed by the above procedure are also equivalent, so the construction is valid for families of elements of $\mathbb{R}\times\mathbb{T}$.

We extend the notion of ancestor in the tree $\mathbf{t}(\bar{\mathbf{f}})$ by

$$u \prec v \iff \begin{cases} u < v & \text{if } u, v \in \mathbb{R}, \\ u \leq h_i & \text{if } u \in \mathbb{R} \text{ and } v \in \mathbf{t}_i^*, \\ u \prec_i v & \text{if } u, v \in \mathbf{t}_i^*, \end{cases}$$

where $\prec_i$ denotes the genealogical order in the tree $\mathbf{t}_i$. We also extend the notion of height of a vertex $u \in \mathbf{t}(\mathbf{f})$ by

$$H(u) = \begin{cases} u & \text{if } u \in \mathbb{R}, \\ h_i + H_i(u) & \text{if } u \in \mathbf{t}_i^*, \end{cases}$$

where H_i denotes the height of a vertex in the tree $\mathbf{t}_i$.

Definition 2.4. *(Ancestral tree) Let $\bar{\mathbf{f}}$ be a leveled forest and let $\bar{\mathbf{t}} = \mathbf{t}(\bar{\mathbf{f}})$ be its associated tree. For every $a \in \mathbb{R}$, we define $\mathcal{Z}_{\bar{\mathbf{t}}}(a)$ the population at height a, by (17) with $\mathbf{t}$ replaced by $\bar{\mathbf{t}}$ and the ancestral tree $\mathcal{A}_{\bar{\mathbf{t}}}(a)$ of the population at level a by*

$$\mathcal{A}_{\bar{\mathbf{t}}}(a) = \mathcal{Z}_{\bar{\mathbf{t}}}(a) \cup \mathrm{Anc}(\mathcal{Z}_{\bar{\mathbf{t}}}(a)),$$

where $\mathrm{Anc}(\mathcal{Z}_{\bar{\mathbf{t}}}(a)) = \cup_{v \in \mathcal{Z}_{\bar{\mathbf{t}}}(a)} \{u \in \bar{\mathbf{t}},\, u \prec v\}$ is the set of all the ancestors in $\bar{\mathbf{t}}$ of the vertices of $\mathcal{Z}_{\bar{\mathbf{t}}}(a)$.

When there is no confusion we write $\mathcal{A}$ for $\mathcal{A}_{\bar{\mathbf{t}}}$.

3. The stationary Lévy tree

3.1. **Random forests, CB and CBI processes.** Let ψ be a branching mechanism defined by (1) which is critical or sub-critical ($\alpha \geq 0$) and satisfies the Grey condition (first condition in (2)). For $r > 0$, we denote by $\mathbb{P}_r(d\mathbf{f})$ the probability distribution on $\mathbb{F}$ of the random forest $\mathcal{F} = (\mathcal{T}_i)_{i \in I}$ given by the atoms of a poisson point measure on $\mathbb{T}$ with intensity $r\mathbb{N}(d\mathbf{t})$. Under $\mathbb{P}_r$, the family $(\ell^a(\mathcal{T}_i))_{i \in I}$ of the corresponding local times at level $a \geq 0$ is well defined, and we define the local time at level a of the forest $\mathcal{F}$ by

$$\ell^a(\mathcal{F}) = \sum_{i \in I} \ell^a(\mathcal{T}_i). \tag{19}$$

Let the size-population process $Z = (Z_a, a \geq 0)$ be defined (under $\mathbb{P}_r$) by (18) with the local time $\ell^a(\mathcal{T})$ replaced by $\ell^a(\mathcal{F})$. By property (iv) of the Lévy tree excursion measure, and the definition of the probability measure $\mathbb{P}_r$, we get that under $\mathbb{P}_r$, the process Z is a CB started at time 0 from r.

If $\mathbf{f} = (\mathbf{t}_i)_{i \in I}$ is a forest and $h \in \mathbb{R}$, the pair $(h, \mathbf{f})$ can be viewed as the leveled forest $(h, \mathbf{t}_i)_{i \in I}$. Eventually, a family of leveled forests $(\bar{\mathbf{f}}_i)_{i \in I}$ can be viewed as a leveled forest since a countable disjoint union of countable sets remains countable. Conversely a tree is a forest, thus the measure $\mathbb{N}(d\mathbf{t})$ on $\mathbb{T}$ can be viewed as a measure $\mathbb{N}(d\mathbf{f})$ on $\mathbb{F}$.

We denote by $\bar{\mathbb{P}}(d\bar{\mathbf{f}})$ the probability distribution on $\bar{\mathbb{F}}$ of the random leveled forest $\bar{\mathcal{F}} = (h_i, \mathcal{F}_i)_{i \in I}$ given by the atoms of a Poisson point measure on $\mathbb{R} \times \mathbb{F}$ with intensity

$$\nu(dh, d\mathbf{f}) = dh \left(\beta \mathbb{N}[d\mathbf{f}] + \int_0^{+\infty} \pi(dr)\, \mathbb{P}_r(d\mathbf{f}) \right), \tag{20}$$

and let $\bar{\mathcal{T}} = \mathbf{t}(\bar{\mathcal{F}})$ be the random tree associated with this leveled forest. The random tree $\bar{\mathcal{T}}$ under $\bar{\mathbb{P}}$ can be viewed as stationary version of the Lévy tree with branching mechanism ψ conditioned on non-extinction, see [14], Section 3. We call the random tree $\bar{\mathcal{T}}$ the stationary Lévy tree.

For every $i \in I$, the local time measure $\ell^a(\mathcal{F}_i)$ at level a of the leveled forest $(h_i, \mathcal{F}_i)$ is a.s. well-defined by (19). We then define, for every $a \in \mathbb{R}$, the local time measure at level a for the tree $\bar{\mathcal{T}}$ by

$$\ell^a(\bar{\mathcal{T}}) = \sum_{i\in I} \ell^{a-h_i}(\mathcal{F}_i)\mathbf{1}_{\{h_i \leq a\}}. \tag{21}$$

By standard property of Poisson point measures, we have the following result, where $Z = (Z_a, a \in \mathbb{R})$ is defined (under $\bar{\mathbb{P}}$) by (18) with the local time $\ell^a(\mathcal{T})$ replaced by $\ell^a(\bar{\mathcal{T}})$.

Proposition 3.1. *Assume that (2) holds. Under $\bar{\mathbb{P}}$, the process Z is a stationary CBI process with branching mechanism ψ and immigration mechanism ϕ given by (5).*

Proof. Since Z under $\mathbb{N}$ is distributed as Y under N, we deduce that the push-forward of the measure $\nu(dh, d\mathbf{f})$ defined in (20) by the function $(h, \mathbf{f}) \mapsto (h, \ell(\mathbf{f}))$, where $\ell(\mathbf{f}) = (\ell^a(\mathbf{f}), a \geq 0)$, is the measure $\nu_0(dh, dY)$ defined in (20). Then use results recalled in Section 2.2, to deduce that under $\bar{\mathbb{P}}$, the process Z is a stationary CBI process with branching mechanism ψ and immigration mechanism ϕ. □

3.2. **Branching points of the ancestral tree.** Recall $\bar{\mathcal{T}}$ is defined under $\bar{\mathbb{P}}$ in the previous section. For $t \in \mathbb{R}$, we write $\mathcal{A}(t)$ the ancestral tree $\mathcal{A}_{\bar{\mathcal{T}}}(t)$ of the population at level t defined by Definition 2.4. Notice that $\bar{\mathbb{P}}$-a.s. $\mathcal{A}(t)$ has only a finite number of vertices at any level $s < t$ and we set for $s < t$:

$$M_s^t = \text{Card}\,\{u \in \mathcal{A}_{\bar{\mathcal{T}}}(t),\ H(u) = s\} - 1. \tag{22}$$

The number M_s^t is exactly the number of individuals of the tree $\bar{\mathcal{T}}$ at level s that have descendants at level t, the immortal (or two-sided infinite) spine being excluded (which explains the -1 in the definition of M_s^t).

Under $\bar{\mathbb{P}}$, since the intensity $\nu(dh, d\mathbf{f})$ is invariant by translation in h, we get that the distribution of the ancestral tree $\mathcal{A}(t)$ does not depend on $t \in \mathbb{R}$. Therefore, we can fix the level at which the current population is considered, say $t = 0$, and look at the ancestral process $M^0 = (M_s^0, s < 0)$ which is a pure-birth process starting at time $s = -\infty$ from 0.

We define the jumping times of the process M^0 inductively by setting

$$\tau_0 = \sup\{t > 0,\ M_{-t}^0 \neq 0\} \tag{23}$$

and for $n \geq 1$,

$$\tau_n = \sup\{t < \tau_{n-1},\ M_{-t}^0 \neq M_{(-t)-}^0\}, \tag{24}$$

and we define the size of the n-th jump of the process M^0, $n \geq 0$, by

$$\xi_n = M_{-\tau_n}^0 - M_{-(\tau_n)-}^0 = M_{-\tau_n}^0 - M_{-\tau_{n-1}}^0. \tag{25}$$

In the sequel, we will distinguish between the jumps that are due to a new immigration (i.e. a branching point on the infinite spine) and those coming from a reproduction of an individual of the ancestral tree. For that purpose, recall that the tree is constructed from a random leveled forest $(h_i, \mathcal{T}_i)_{i\in I}$.

We then define for every $n \geq 1$,

$$\tau_n^I = -\inf\{h_i > -\tau_{n-1},\ H(\mathcal{T}_i) \geq -h_i\} \quad \text{and} \tag{26}$$

$$\tau_n^B = -\inf\{t > -\tau_{n-1},\ M_t^0 \neq M_{t-}^0 \text{ and } t \neq h_i\ \forall i \in I\}, \tag{27}$$

so that

$$\tau_n = \tau_n^I \vee \tau_n^B. \tag{28}$$

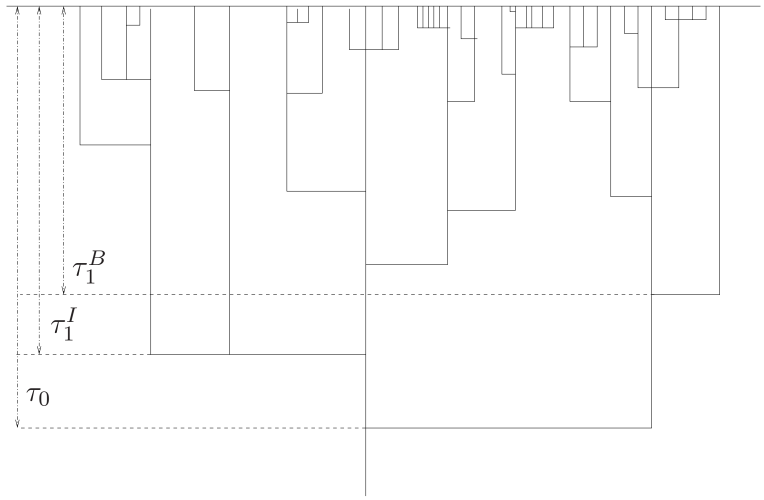

FIGURE 1. The ancestral tree of the extant population at time $t = 0$ and the first jumping times (time $t = 0$ corresponds to the top horizontal line).

Thanks to Theorem 2.7.1 of [16], we have, for every $r \in [0, 1]$, every $t > u > 0$ and every $n \in \mathbb{N}^*$,

$$\bar{\mathbb{E}}[r^{\xi_1} \mid \tau_0 = t, \xi_0 = n, \tau_1^B = u, \tau_1^I < u] = g_t(t-u, r), \tag{29}$$

where

$$g_t(s, r) = r\, \frac{\psi'(c(t-s)) - \gamma_\psi(c(t-s), (1-r)c(t-s))}{\psi'(c(t-s)) - \gamma_\psi(c(t-s), 0)}, \tag{30}$$

with

$$\forall a, b \geq 0, \qquad \gamma_\psi(a, b) = \begin{cases} \frac{\psi(a)-\psi(b)}{a-b} & \text{if } a \neq b, \\ \psi'(a) & \text{if } a = b. \end{cases}$$

On the other hand, by standard properties of Poisson point measures (see also Proposition 5.2 in [14]), we have:

$$\bar{\mathbb{E}}[r^{\xi_1} \mid \tau_0 = t, \xi_0 = n, \tau_1^I = u, \tau_1^B < u] = 1 - \frac{\phi((1-r)c(u))}{\phi(c(u))}\cdot \tag{31}$$

4. Properties of the ancestral process in the sub-critical stable case

In this section, the branching mechanism ψ, the immigration mechanism ϕ, and the function $\tilde{\psi}$ are given by (7), (5) and (9), that is, for $\lambda \geq 0$:

$$\psi(\lambda) = \alpha\lambda + \gamma\lambda^b, \quad \phi(\lambda) = b\gamma\lambda^{b-1}, \quad \tilde{\psi}(\lambda) = \alpha + \gamma\lambda^{b-1}, \tag{32}$$

with $\alpha > 0$, $\gamma > 0$ and $b \in (1, 2]$. We recall[1] the extinction probability $c(t)$ defined by (13) and given by (8), and elucidate the Laplace transform of the CBI $\bar{u}$ as well as the constant κ defined in (3) and (6):

$$c(t) = \left(\frac{\alpha}{\gamma\left(\mathrm{e}^{(b-1)\alpha t} - 1\right)}\right)^{\frac{1}{b-1}}, \quad \bar{u}(\lambda) = \left(1 + \frac{\gamma}{\alpha}\lambda^{b-1}\right)^{-\frac{b}{b-1}} \quad \text{and} \quad \kappa = \left(\frac{\alpha}{\gamma}\right)^{\frac{1}{b-1}}. \tag{33}$$

[1]According to Example 3.1 p. 62 in [35] (where $v_t(\lambda)$ corresponds to $u(\lambda, t)$ in our setting), we also have $u(\lambda, t) = \mathrm{e}^{-\alpha t}\, \lambda \left[1 + \gamma\, \alpha^{-1}\, (1 - \mathrm{e}^{-\alpha(b-1)t})\lambda^{b-1}\right]^{-1/(b-1)}$.

The expression of the function $g_t(s,r)$ of (30) does not depend on s and t. We have for $r \in [0,1]$:

$$g_t(s,r) = g_B(r) \quad \text{with} \quad g_B(r) = \frac{br - 1 + (1-r)^b}{b-1}. \tag{34}$$

We also define the generating function g_I by, for $r \in [0,1]$:

$$g_I(r) = \frac{(b-1)}{b} g_B'(r) = 1 - (1-r)^{b-1}. \tag{35}$$

4.1. **Distribution of the time-changed ancestral process.** We make precise the time change given in (9): for $t > 0$

$$R(t) = \log\left(\frac{\tilde{\psi}(c(t))}{\tilde{\psi}(0)}\right) = \log\left(\frac{\mathrm{e}^{(b-1)\alpha t}}{\mathrm{e}^{(b-1)\alpha t} - 1}\right).$$

The function R is continuous and strictly decreasing; we also have that $\lim_{t\to 0} R(t) = +\infty$ and $\lim_{t\to+\infty} R(t) = 0$. Thus the function R is one-to-one from $(0,+\infty)$ to $(0,+\infty)$. We consider the time-changed ancestral process $\tilde{M} = (\tilde{M}_t, t \geq 0)$ defined by $\tilde{M}_0 = 0$, and for $t > 0$:

$$\tilde{M}_t = M^0_{T(t)} \quad \text{with} \quad T(t) = -R^{-1}(t).$$

The next theorem, whose proof is given in Section 4.2, is the main result of this section. It states that the ancestral process is a continuous-time Galton-Watson process with immigration (GWI process).

Theorem 4.1. *Consider the sub-critical stable branching mechanism with immigration* (32)*. The time-changed ancestral process $\tilde{M}$ is distributed under $\bar{\mathbb{P}}$ as a GWI process, $X = (X_t, t \geq 0)$, with:*

(i) $X_0 = 0$ *a.s.;*
(ii) *the branching rate of X is 1;*
(iii) *the offspring distribution has generating function g_B defined in (34);*
(iv) *the immigration rate is $\frac{b}{b-1}$;*
(v) *the number of immigrants has generating function g_I defined in (35).*

Recall that the distribution of the process X is characterized by the Markov property and its infinitesimal transition probabilities. Let us denote by $p = (p_n, n \geq 0)$ (resp. $q = (q_n, n \geq 0)$) the distribution on $\mathbb{N}$ associated with the generating function g_B (resp. g_I).

We compute the infinitesimal generator of X at time t. First, since $p_0 = g_B(0) = 0$, we have, for every $t \geq 0$, $h > 0$ and every $0 \leq k < n$,

$$\mathrm{P}(X_{t+h} = k | X_t = n) = 0.$$

Furthermore, by Equation (35), we have, for every $n \geq 1$,

$$np_n = \frac{b}{b-1} q_{n-1}.$$

Therefore, as $h \to 0+$, we have for every $0 \leq n < k$,

$$\mathrm{P}(X_{t+h} = k | X_t = n) = \left(np_{k-n+1} + \frac{b}{b-1} q_{k-n}\right) h + o(h) = (k+1) p_{k-n+1} h + o(h).$$

Eventually, since $p_1 = g_B'(0) = 0$, we have, for every $n \geq 0$, as $h \to 0+$,

$$\mathrm{P}(X_{t+h} = n | X_t = n) = 1 - \sum_{k=n+1}^{+\infty} (k+1) p_{k-n+1} h + o(h) = 1 - \left(\frac{b}{b-1} + n\right) h + o(h).$$

To sum up, we have the following transition rates for the GWI process X as $h \to 0+$,

$$
\mathrm{P}(X_{t+h} = k | X_t = n) = \begin{cases} (k+1)p_{k-n+1}h + o(h) & \text{if } k \geq n+1, \\ 1 - \left(\frac{b}{b-1} + n\right) h + o(h) & \text{if } k = n, \\ o(h) & \text{otherwise.} \end{cases} \tag{36}
$$

Taking the limit of $h^{-1}(\mathrm{P}(X_{t+h} = k|X_t = n) - \mathbf{1}_{\{k=n\}})$, as h goes down to 0, gives the infinitesimal generator of X at time t. In particular, if $(\tau'_n, n \geq 0)$ is the sequence of jumping times of X (with $\tau'_0 = 0$), we have for $r \in [0,1]$, $n, k \geq 0$,

$$
\mathrm{E}\left[r^{X_{\tau'_{n+1}} - X_{\tau'_n}} \mid X_{\tau'_n} = k \right] = g_{[k]}(r), \tag{37}
$$

where for $r \in [0,1]$,

$$
\begin{aligned}
g_{[k]}(r) &= \frac{k}{k(b-1)+b}\left(br - 1 + (1-r)^b\right) + \frac{b}{k(b-1)+b}\left(1 - (1-r)^{b-1}\right) \\
&= \frac{k(b-1)}{k(b-1)+b} g_B(r) + \frac{b}{k(b-1)+b} g_I(r).
\end{aligned} \tag{38}
$$

Remark 4.2. Let $\chi = (\chi_t, t \geq 0)$ be a continuous-time Galton-Watson process (GW process) with branching rate 1, offspring distribution p and starting at $\chi_0 = 1$. Recall that the size-biased version of χ is the process $\hat{\chi} = (\hat{\chi}_t, t \geq 0)$ such that for every $T > 0$ and every bounded measurable functional φ, we have:

$$
\mathrm{E}\left[\varphi(\hat{\chi}_t,\, t \in [0,T])\right] = \frac{1}{\mathrm{E}[\chi_T]} \mathrm{E}\left[\chi_T\, \varphi(\chi_t,\, t \in [0,T])\right]. \tag{39}
$$

Then, the GWI process X of Theorem 4.1 is distributed as $\hat{\chi} - 1$.

Indeed, the process $\hat{\chi}$ is a Markov process as a Doob h-transform of a Markov process (the process $(\chi_t/\mathbb{E}[\chi_t], t \geq 0)$ is a martingale). Its transition rates are given by the following computations. For every $t \geq 0$, $\varepsilon > 0$ and every integers $1 \leq n < k$, we have:

$$
\begin{aligned}
\frac{1}{\varepsilon}\mathrm{P}(\hat{\chi}_{t+\varepsilon} = k | \hat{\chi}_t = n) &= \frac{1}{\varepsilon} \frac{\mathrm{E}[\mathbf{1}_{\{\hat{\chi}_{t+\varepsilon} = k,\, \hat{\chi}_t = n\}}]}{\mathrm{E}[\mathbf{1}_{\{\hat{\chi}_t = n\}}]} \\
&= \frac{1}{\varepsilon} \frac{\mathrm{E}[\chi_{t+\varepsilon} \mathbf{1}_{\{\chi_{t+\varepsilon} = k,\, \chi_t = n\}}]}{\mathrm{E}[\chi_t \mathbf{1}_{\{\chi_t = n\}}]} \frac{\mathrm{E}[\chi_t]}{\mathrm{E}[\chi_{t+\varepsilon}]} \\
&= \frac{1}{\varepsilon} \frac{k}{n} \mathrm{P}(\chi_{t+\varepsilon} = k | \chi_t = n) \frac{\mathrm{E}[\chi_t]}{\mathrm{E}[\chi_{t+\varepsilon}]}.
\end{aligned}
$$

We deduce that for $0 \leq n < k$:

$$
\lim_{\varepsilon \to 0+} \frac{1}{\varepsilon} \mathrm{P}(\hat{\chi}_{t+\varepsilon} - 1 = k | \hat{\chi}_t - 1 = n) = \frac{k+1}{n+1}(n+1)p_{k-n+1}.
$$

According to the transition rates given in (36), we deduce that X is distributed as $\hat{\chi} - 1$.

The following result is an application of Theorem 4.1. Recall κ defined in (33) and Z in Proposition 3.1.

Corollary 4.3. *Let X be the GWI process defined in Theorem 4.1. Then there exists a random variable W distributed as κZ_0 under $\bar{\mathbb{P}}$, such that*

$$
\lim_{t \to \infty} \mathrm{e}^{-\frac{t}{b-1}}\, X_t \overset{a.s.}{=} W. \tag{40}
$$

Proof. It is known from Corollary 6.5 in [14] that a.s. $\lim_{s\downarrow 0} \frac{M_{-s}}{c(s)} = Z_0$. Using the expressions of R and c, we have:

$$c\big(R^{-1}(t)\big) = \left(\frac{\alpha}{\gamma}(\mathrm{e}^t - 1)\right)^{\frac{1}{b-1}}, \tag{41}$$

and thus $\lim_{t\to\infty} \mathrm{e}^{-\frac{t}{b-1}} c\big(R^{-1}(t)\big) = \kappa$. Then (40) follows readily from Theorem 4.1. □

Remark 4.4. If a GW process or a GWI process has finite offspring mean and finite immigration mean, then limits such as (40) are well-known, see for example Section III.7 in [5]. However, as the immigration mean if infinite since $g_I'(1-) = +\infty$, we deduce that in our setting $\mathrm{E}[X_t] = \infty$. We have not found results such as (40) in the literature.

Remark 4.5. According to (16) and (33), one can check that

$$\mathrm{E}[\mathrm{e}^{-\lambda W}] = \left(1 + \lambda^{b-1}\right)^{-\frac{b}{b-1}} = \mathrm{E}[\mathrm{e}^{-\lambda^{b-1} G}], \tag{42}$$

where G has the $\Gamma(\frac{b}{b-1}, 1)$ distribution. For $b = 2$, one get that W is $\Gamma(2,1)$. For $b \in (1,2)$, according to Proposition 1.5 in [7], using notations from Propositions 4.2 and 4.3 in [27], we get that W is distributed as $\chi_{b-1,b}$ and thus has a generalized positive Linnik distribution with parameter $(b-1, b)$ see the first paragraph of Section 2.3 in [27] and the references therein. Remark 2.2 and (2.25) in [27] give that W has intensity $f_{b-1,b}$ on $(0,+\infty)$, where for $\mathrm{a} \in (0,1)$, $\mathrm{b} > 0$ and $z > 0$:

$$f_{\mathrm{a},\mathrm{b}}(z) = \frac{1}{\pi} \int_0^\infty \frac{\mathrm{e}^{-zy} \sin(\pi \mathrm{b} F_{\mathrm{a}}(y))}{[y^{2\mathrm{a}} + 2y^{\mathrm{a}} \cos(\mathrm{a}\pi) + 1]^{\frac{\mathrm{b}}{2\mathrm{a}}}}\, dy, \tag{43}$$

with

$$F_{\mathrm{a}}(y) = 1 - \frac{1}{\pi \mathrm{a}} \cot^{-1}\left(\cot(\pi \mathrm{a}) + \frac{y^{\mathrm{a}}}{\sin(\pi \mathrm{a})}\right).$$

We also give another representation of the density of W using the fact that in our case $\mathrm{b} = \mathrm{a} + 1 = b$. Indeed, according to (4.7), Proposition 4.3 (iii) in [27] we have that:

$$f_{\mathrm{a},\mathrm{a}+1}(z) = -z f'_{\mathrm{a},1}(z) \tag{44}$$

using the representation of $f_{\mathrm{a},1}$ from Proposition 2.8 and (2.22):

$$f_{\mathrm{a},1}(z) = \int_0^\infty \frac{\mathrm{e}^{z/y}}{y} \Delta_{\mathrm{a},1}(y)\, dy \quad \text{with} \quad \Delta_{\mathrm{a},1}(y) = \frac{1}{\pi} \frac{\sin(\pi(1 - F_{\mathrm{a}}(y)))}{[y^{2\mathrm{a}} + 2y^{\mathrm{a}} \cos(\mathrm{a}\pi) + 1]^{\frac{1}{2\mathrm{a}}}}.$$

Remark 4.6. Let χ be the GW process introduced in Remark 4.2. Recall from [5] Formula (4) p. 108 that $\mathbb{E}[\chi_t] = \mathrm{e}^{\frac{t}{b-1}}$. Let $W' = \lim_{t\to+\infty} \mathrm{e}^{-\frac{t}{b-1}} \chi_t$. By [27] Proposition 4.1 and Proposition 4.3 (iii), the distribution of W' has density $-f'_{b-1,1}$. Then, Equation (44) readily implies that the distribution of W is the size-biased distribution of W' i.e., for every bounded continuous function φ, we have:

$$\mathrm{E}[\varphi(W)] = \mathrm{E}[W' \varphi(W')]. \tag{45}$$

Another way of getting this identity is to use the relationship between the processes X and χ. For every bounded continuous function φ, we have:

$$\mathrm{E}\left[\varphi(\mathrm{e}^{-\frac{t}{b-1}} X_t)\right] = \mathrm{E}\left[\frac{\chi_t}{\mathrm{E}[\chi_t]} \varphi(\mathrm{e}^{-\frac{t}{b-1}} (\chi_t - 1))\right].$$

Moreover, the expression of $f_{b-1,1}$ implies that the variable W' admits every moment of order $\theta < b$. Then the martingale $(\chi_t/\mathrm{E}[\chi_t], t \geq 0)$ is uniformly integrable (see [4] for this result for a dicrete time GW process) and taking the limit in the previous equation gives (45).

4.2. **Proof of Theorem 4.1.** We first prove two intermediate lemmas, the first one on the Markov property for the ancestral process M^0 and the second one on the distribution of the jumping times of $\tilde{M}$. Recall the notations of Subsection 3.2 for the jumping times $(\tau_n, n \geq 0)$ and the jumping sizes $(\xi_n, n \geq 0)$ of the ancestral process M^0, see (23), (24) and (25).

Lemma 4.7. *We set $(\mathcal{G}_n,\, n \geq 0)$ the filtration generated by the process $((M^0_{-\tau_n}, \tau_n), n \geq 0)$. Under $\bar{\mathbb{P}}$, for every $n > 0$, conditionally given $\mathcal{G}_{n-1}$, the random variables ξ_n and τ_n are independent. Moreover, the conditional distribution of the random variable ξ_n given $\mathcal{G}_{n-1}$ has generating function $g_{[M^0_{-\tau_{n-1}}]}$, with $g_{[\cdot]}$ defined in* (38).

Proof. According to Remark 5.6 in [14], we have for $t > 0$:

$$\bar{\mathbb{E}}[r^{\xi_0} \mid \tau_0 = t] = 1 - (1-r)^{b-1} = g_I(r). \tag{46}$$

Thus ξ_0 and τ_0 are independent. Then by the branching property, it suffices to study the case $n = 1$. Let us first compute the conditional distribution of τ_1.

Recall that $\tau_1 = \tau_1^I \vee \tau_1^B$. By standard properties of Poisson point measures, we have

$$\begin{aligned}
\bar{\mathbb{P}}(\tau_1^I < u | \tau_0 = t, M_{-\tau_0} = n) &= \exp\left\{-\int_u^t ds \int_{(0,+\infty)} r\pi(dr)\mathbb{P}_r(H(\mathcal{T}) > s)\right\} \\
&= \exp\left\{-\int_u^t ds \int_{(0,+\infty)} r\pi(dr)\left(1 - \mathrm{e}^{-r\mathbb{N}[H(\mathcal{T})>s]}\right)\right\} \\
&= \exp\left\{-\int_u^t ds\, \phi(c(s))\right\} \\
&= \exp\left\{-b\int_u^t \frac{\alpha\, ds}{\mathrm{e}^{(b-1)\alpha s} - 1}\right\} \\
&= \left(\frac{\mathrm{e}^{\alpha t}\, c(t)}{\mathrm{e}^{\alpha u}\, c(u)}\right)^b.
\end{aligned} \tag{47}$$

Moreover, by Theorem 2.7.1 of [16], we have, using $\tilde{\psi}(\lambda) = \psi(\lambda)/\lambda$:

$$\bar{\mathbb{P}}(\tau_1^B < u | \tau_0 = t, M_{-\tau_0} = n) = \left(\frac{\tilde{\psi}(c(t))}{\tilde{\psi}(c(u))}\right)^n. \tag{48}$$

Recall that $-c'(u) = \psi(c(u)) = \alpha c(u) + \gamma c(u)^b$. We deduce the conditional distribution of τ_1:

$$\begin{aligned}
&\bar{\mathbb{P}}(\tau_1 \in du | \tau_0 = t, M_{-\tau_0} = n) \\
&\quad = \bar{\mathbb{P}}(\tau_1^B \in du,\, \tau_1^I < u | \tau_0 = t, M_{-\tau_0} = n) + \bar{\mathbb{P}}(\tau_1^I \in du,\, \tau_1^B < u | \tau_0 = t, M_{-\tau_0} = n) \\
&\quad = \bar{\mathbb{P}}(\tau_1^B \in du | \tau_0 = t, M_{-\tau_0} = n) \bar{\mathbb{P}}(\tau_1^I < u | \tau_0 = t, M_{-\tau_0} = n) \\
&\qquad\qquad\qquad\qquad + \bar{\mathbb{P}}(\tau_1^I \in du | \tau_0 = t, M_{-\tau_0} = n) \bar{\mathbb{P}}(\tau_1^B < u | \tau_0 = t, M_{-\tau_0} = n) \\
&\quad = \left(\frac{\mathrm{e}^{\alpha t}\, c(t)}{\mathrm{e}^{\alpha u}\, c(u)} \right)^b \left(\frac{\alpha + \gamma c(t)^{b-1}}{\alpha + \gamma c(u)^{b-1}} \right)^n \left[\frac{n\gamma(b-1)(-c'(u))c(u)^{b-2}}{\alpha + \gamma c(u)^{b-1}} + b\left(-\alpha + \frac{(-c'(u))}{c(u)} \right) \right] du \\
&\quad = \left(\frac{\mathrm{e}^{\alpha t}\, c(t)}{\mathrm{e}^{\alpha u}\, c(u)} \right)^b \left(\frac{\alpha + \gamma c(t)^{b-1}}{\alpha + \gamma c(u)^{b-1}} \right)^n \left[n\gamma(b-1)c(u)^{b-1} + b\gamma c(u)^{b-1} \right] du \\
&\quad = \left(\frac{\mathrm{e}^{\alpha t}\, c(t)}{\mathrm{e}^{\alpha u}\, c(u)} \right)^b \left(\frac{\alpha + \gamma c(t)^{b-1}}{\alpha + \gamma c(u)^{b-1}} \right)^n \gamma(nb + b - n) c(u)^{b-1} du.
\end{aligned}$$

We deduce that:

$$\begin{aligned}
\bar{\mathbb{P}}(\tau_1^B \in du,\, \tau_1^I < u | \tau_1 \in du, \tau_0 = t, M_{-\tau_0} = n) &= \frac{n(b-1)}{nb + b - n}, \\
\bar{\mathbb{P}}(\tau_1^I \in du,\, \tau_1^B < u | \tau_1 \in du, \tau_0 = t, M_{-\tau_0} = n) &= \frac{b}{nb + b - n}.
\end{aligned}$$

Using formulas (29), (31), (34), the expression of ϕ, and the definition (38) of $g_{[n]}$, we get that

$$\begin{aligned}
\bar{\mathbb{E}}[r^{\xi_1} | \tau_0 = t, \xi_0 = n, \tau_1 = u] &= g_t(t-u, r) \frac{n(b-1)}{nb + b - n} + \left(1 - \frac{\phi((1-r)c(u))}{\phi(c(u))} \right) \frac{b}{nb + b - n} \\
&= g_B(r) \frac{n(b-1)}{nb + b - n} + g_I(r) \frac{b}{nb + b - n} \\
&= g_{[n]}(r).
\end{aligned}$$

Since the latter expression does not depend on u, this proves the conditional independence between ξ_1 and τ_1. Moreover, we indeed recover the expression of (37) for the conditional generating function of ξ_1. $\square$

Remark 4.8. Lemma 4.7 implies in particular the independence between the first jumping time τ_0 and the states of M^0, i.e. the sequence $(M^0_{-\tau_n}, n \geq 0)$.

We denote, for every $n \geq 0$, the scaled jumping time $\tilde{\tau}_n = R(\tau_n)$ and the corresponding time intervals $\Delta_n = \tilde{\tau}_n - \tilde{\tau}_{n-1}$ with the convention $\tilde{\tau}_{-1} = 0$. The following lemma gives the distribution of the time intervals given the states of the process $\tilde{M}$.

Lemma 4.9. *Conditionally given $(\tilde{M}_{\tilde{\tau}_n}, n \geq 0)$, the random variables $(\Delta_n, n \geq 0)$ are independant with for all $u \geq 0$ and $n \geq 0$:*

$$\bar{\mathbb{P}}\big(\Delta_n > u \mid (\tilde{M}_{\tilde{\tau}_k}, k \geq 0)\big) = \exp\left(-\left(\tilde{M}_{\tilde{\tau}_n} + \frac{b}{b-1} \right) u \right). \tag{49}$$

Proof. Let us first compute the distribution of $\tilde{\tau}_0 = \Delta_0$. For every $u \geq 0$, we have:

$$\bar{\mathbb{P}}\left(\tilde{\tau}_0 > u \mid (\tilde{M}_{\tilde{\tau}_k}, k \geq 0) \right) = \bar{\mathbb{P}}\left(R(\tau_0) > u \mid (M^0_{-\tau_k}, k \geq 0) \right) = \bar{\mathbb{P}}\big(\tau_0 < R^{-1}(u)\big),$$

using the independance between τ_0 and the states of M^0, see Remark 4.8, and that R is non-increasing.

By the branching property, for every $r > 0$, conditionally on Z_{-r}, the random variable M^0_{-r} is distributed under $\bar{\mathbb{P}}$ according to a Poisson distribution with parameter $c(r)Z_{-r}$. We get:

$$\bar{\mathbb{P}}(\tau_0 < r) = \bar{\mathbb{P}}(M^0_{-r} = 0) = \bar{\mathbb{E}}\left[\mathrm{e}^{-c(r)Z_{-r}}\right].$$

Note that Z is a CBI, so that (16) holds (with Y distributed as Z). Thus, using (33) as well as (41), we deduce that:

$$\bar{\mathbb{P}}(\tilde{\tau}_0 > u) = \bar{u}\left(c\left(R^{-1}(u)\right)\right) = \mathrm{e}^{-\frac{bu}{b-1}},$$

which is the sought after expression since $\tilde{M}_0 = 0$.

Let us now compute the distribution of $\Delta_1 = \tilde{\tau}_1 - \tilde{\tau}_0$. First, using that $\tau_n = \tau_n^I \vee \tau_n^B$ (see (28)) and Equations (47) and (48), we have

$$\begin{aligned}
\bar{\mathbb{P}}\left(\tilde{\tau}_1 > u \,\middle|\, \tilde{\tau}_0 = t, \tilde{M}_{\tilde{\tau}_0} = k\right) &= \bar{\mathbb{P}}\left(\tau_1 < R^{-1}(u) \,\middle|\, \tau_0 = R^{-1}(t),\ M^0_{-\tau_0} = k\right) \\
&= \bar{\mathbb{P}}\left(\tau_1^I < R^{-1}(u) \,\middle|\, \tau_0 = R^{-1}(t),\ M^0_{-\tau_0} = k\right) \\
&\qquad \times \bar{\mathbb{P}}\left(\tau_1^B < R^{-1}(u) \,\middle|\, \tau_0 = R^{-1}(t),\ M^0_{-\tau_0} = k\right) \\
&= \frac{\mathrm{e}^{\alpha b R^{-1}(t)}\, c\big(R^{-1}(t)\big)^b}{\mathrm{e}^{\alpha b R^{-1}(u)}\, c\big(R^{-1}(u)\big)^b} \frac{\tilde{\psi}\Big(c\big(R^{-1}(t)\big)\Big)^k}{\tilde{\psi}\Big(c\big(R^{-1}(u)\big)\Big)^k}.
\end{aligned}$$

Using the expressions of R and (41), we have

$$\tilde{\psi}\Big(c\big(R^{-1}(t)\big)\Big) = \alpha\,\mathrm{e}^t \quad \text{and} \quad \mathrm{e}^{\alpha R^{-1}(t)} = \left(\frac{\mathrm{e}^t}{\mathrm{e}^t - 1}\right)^{\frac{1}{b-1}}.$$

This and (41) again give:

$$\bar{\mathbb{P}}\left(\Delta_1 > u \,\middle|\, \tilde{\tau}_0 = t, \tilde{M}_{\tilde{\tau}_0} = k\right) = \bar{\mathbb{P}}\left(\tilde{\tau}_1 > u + t \,\middle|\, \tilde{\tau}_0 = t, \tilde{M}_{\tilde{\tau}_0} = k\right) = \mathrm{e}^{-\left(k + \frac{b}{b-1}\right)u}.$$

By an easy induction, Lemma 4.7 implies that, conditionally given $\mathcal{G}_0$, the random variable τ_1 is independent of the states $(\tilde{M}_{\tilde{\tau}_k}, k \geq 1)$. Therefore, we get

$$\bar{\mathbb{P}}\left(\Delta_1 > u \,\middle|\, \tilde{\tau}_0 = t, (\tilde{M}_{\tilde{\tau}_n}, n \geq 0)\right) = \mathrm{e}^{-\left(\tilde{M}_{\tilde{\tau}_0} + \frac{b}{b-1}\right)u}.$$

The proof then follows by induction and by the Markov property. □

Proof of Theorem 4.1. Lemmas 4.7 and 4.9 imply the Markov property for the process $\tilde{M}$, Lemma 4.9 gives the transition rates and Lemma 4.7 gives the distribution of the jumps. This and (36), (37) and (38) give the result. □

4.3. **Distribution of the sizes of the families of the current population.** Recall the forest $\bar{\mathcal{F}} = (h_i, \mathcal{F}_i)_{i \in I}$ from Section 3.1 and the process Z from Proposition 3.1. Let us denote by

$$I_0 = \{i \in I,\ h_i < 0 \text{ and } \ell^{-h_i}(\mathcal{F}_i) \neq 0\}$$

the immigrants that have descendants at time 0. We order the set I_0 by the date of arrival of the immigrant: $I_0 = \{i_k, k \geq 0\}$ with $-\tau_0 = h_{i_0} < h_{i_1} < h_{i_2} < \cdots < 0$. For every $k \geq 0$, we set ζ_k the size of the population at time 0 generated by the k-th immigrant, that is:

$$\zeta_k = \langle \ell^{-h_{i_k}}(\mathcal{F}_{i_k}), 1 \rangle.$$

Notice that $\sum_{k=0}^{+\infty} \zeta_k = Z_0$.

Let $\{\sigma_t : t \geq 0\}$ be a $(b-1)$-stable subordinator: $\mathrm{E}[\mathrm{e}^{-x\sigma_t}] = \mathrm{e}^{-tx^{b-1}}$. Recall κ defined in (33).

Proposition 4.10. *Consider the sub-critical stable branching mechanism with immigration* (32)*. The random point measure $\sum_{i\in\mathbb{N}} \delta_{\kappa\zeta_i}(dx)$ is a Poisson point measure on $[0,\infty)$ with intensity $g(x)\,dx$ where for $x>0$:*

$$g(x) = \frac{b}{x}\mathrm{E}\left[\mathrm{e}^{-(x/\sigma_1)^{b-1}}\right]. \tag{50}$$

We also have that for all $\lambda \geq 0$:

$$\int_0^\infty (1-\mathrm{e}^{-\lambda x})\, g(x)\, dx = \frac{b}{b-1}\log\left(1+\lambda^{b-1}\right) = -\log\left(\bar{u}(\kappa\lambda)\right). \tag{51}$$

Proof. Recall the GWI process X from Theorem 4.1 and the random variable W from (40) in Corollary 4.3. Let $\{0 = T_0 < T_1 < T_2 < \cdots\}$ be the immigration times of X which forms a Poisson process with rate $b/(b-1)$. Recall g_B and g_I defined in (34) and (35). Let $\{X^i, i\geq 0\}$ be a sequence of independent continuous time Galton-Watson processes with branching rate 1 such that the offspring law has generating function g_B and the law of X_0^i has generating function g_I. For $i\geq 0$, we set $W_i \overset{a.s.}{=} \lim_{t\to\infty} \mathrm{e}^{-\frac{t}{b-1}} X_t^i$ so that $\{W_i : i\geq 0\}$ are independent random variables with the same distribution. By Theorem 3 on Page 116 of [5], we get that for $x\geq 0$, $\mathrm{E}[\mathrm{e}^{-xW_i}] = g_I(\varphi(x))$, where φ is a one-to-one map form $[0,\infty)$ to $(0,1]$ such that for $x\in(0,1]$:

$$\varphi^{-1}(x) = (1-x)\exp\left\{\int_1^x \left(\frac{g_B'(1)-1}{g_B(r)-r} + \frac{1}{1-r}\right) dr\right\}.$$

We get that for $x\in(0,1]$:

$$\varphi^{-1}(x) = (1-x)\exp\left\{\int_0^{1-x} \frac{u^{b-2}}{1-u^{b-1}} du\right\} = \left(\frac{(1-x)^{b-1}}{1-(1-x)^{b-1}}\right)^{1/(b-1)}.$$

This gives that for $x\geq 0$:

$$\varphi(x) = 1 - \left(\frac{x^{b-1}}{1+x^{b-1}}\right)^{1/(b-1)}.$$

We then deduce that:

$$\mathrm{E}[\mathrm{e}^{-xW_i}] = g_I(\varphi(x)) = 1-(1-\varphi(x))^{b-1} = \frac{1}{1+x^{b-1}}. \tag{52}$$

Let $\{(\sigma_s^i, s\geq 0), i\in\mathbb{N}\}$ be a sequence of independent $(b-1)$-stable subordinators and $\{E^i, i\in\mathbb{N}\}$ be a sequence of independent exponentially distributed random variables with parameter 1. Then it is easy to see from (52) that

$$(W_i, i\in\mathbb{N}) \overset{d}{=} (\sigma^i_{E^i}, i\in\mathbb{N}) \overset{d}{=} \left((E^i)^{\frac{1}{b-1}}\sigma_1^i, i\in\mathbb{N}\right),$$

where the last equality follows from scale invariant property of stable subordinator. Thus we have:

$$\left(\mathrm{e}^{-\frac{T_i}{(b-1)}} W_i, i\in\mathbb{N}\right) \overset{d}{=} \left(\mathrm{e}^{-\frac{T_i}{(b-1)}}(E^i)^{\frac{1}{b-1}}\sigma_1^i, i\in\mathbb{N}\right). \tag{53}$$

On the other hand, notice that $\sum_i \delta_{T_i}(dt)\delta_{E^i}(dx)$ is a Poisson random measure on $[0,\infty)^2$ with intensity $\frac{b}{b-1}dt\,\mathrm{e}^{-x}\,dx$. Thus $\sum_i \delta_{\{\mathrm{e}^{-T_i}E^i\}}(ds)$ is a Poisson random measure on $[0,\infty)$ with intensity $\frac{b}{b-1}s^{-1}\,\mathrm{e}^{-s}\,ds$. Indeed, for any bounded positive measurable function f on $[0,\infty)$, one has

$$\mathrm{E}\left[\mathrm{e}^{-\sum_i f(\mathrm{e}^{-T_i E^i})}\right] = \exp\left\{-\int_0^\infty \int_0^\infty (1-\mathrm{e}^{-f(\mathrm{e}^{-t}\,x)})\frac{b}{b-1}dt\ \mathrm{e}^{-x}\,dx\right\}$$
$$= \exp\left\{-\int_0^\infty \int_0^\infty (1-\mathrm{e}^{-f(s)})\frac{b}{b-1}dt\,\mathrm{e}^t\,\mathrm{e}^{-s\,\mathrm{e}^t}\,ds\right\}$$
$$= \exp\left\{-\int_0^\infty (1-\mathrm{e}^{-f(s)})\frac{b}{b-1}s^{-1}\,\mathrm{e}^{-s}\,ds\right\},$$

where the last equality follows from $\int_0^\infty \mathrm{e}^{t-s\,\mathrm{e}^t}\,dt = \int_1^\infty \mathrm{e}^{-st}\,dt = s^{-1}\,\mathrm{e}^{-s}$. Hence, we deduce that

$$F(ds\,dx) := \sum_i \delta_{\{\mathrm{e}^{-T_i}\,E^i\}}(ds)\delta_{\sigma_1^i}(dx)$$

is a Poisson point measure on $[0,\infty)^2$ with intensity $\frac{b}{b-1}s^{-1}\,\mathrm{e}^{-s}\,ds\,\mathrm{P}(\sigma_1\in dx)$. Define

$$G(ds) = \sum_i \delta_{\{\mathrm{e}^{-\frac{T_i}{b-1}}(E^i)^{\frac{1}{b-1}}\sigma_1^i\}}(ds).$$

We shall prove that G is a Poisson point measure on $[0,\infty)$ with intensity $g(s)ds$. We only need to identify the intensity measure. For any positive measurable function f on $[0,\infty)$, we have:

$$\log \mathrm{E}\left[\exp\left\{-\int_0^\infty f(s)G(ds)\right\}\right] = \log \mathrm{E}\left[\exp\left\{-\int_{[0,\infty)^2} f(s^{1/(b-1)}x)\,F(dsdx)\right\}\right]$$
$$= -\int_{[0,\infty)^2}\left(1-\mathrm{e}^{-f(s^{1/(b-1)}x)}\right)\frac{b}{b-1}\,\mathrm{e}^{-s}\,\frac{ds}{s}\,\mathrm{P}(\sigma_1\in dx)$$
$$= -\int_{[0,\infty)^2}\left(1-\mathrm{e}^{-f(t)}\right)\frac{b}{t}\,\mathrm{e}^{-(t/x)^{b-1}}\,dt\,\mathrm{P}(\sigma_1\in dx)$$
$$= -\int_{[0,\infty)}(1-\mathrm{e}^{-f(t)})g(t)\,dt.$$

Then the desired result follows. We now prove the last part of the proposition. We have:

$$\int_{[0,\infty)}(1-\mathrm{e}^{-\lambda t})\,g(t)\,dt = \frac{b}{b-1}\int_0^\infty s^{-1}\,\mathrm{e}^{-s}\,ds\int_0^\infty \mathrm{P}(\sigma_1\in dx)(1-\mathrm{e}^{-\lambda s^{1/(b-1)}x})$$
$$= \frac{b}{b-1}\int_0^\infty ds\,s^{-1}\,\mathrm{e}^{-s}\left(1-\mathrm{e}^{-s\lambda^{b-1}}\right)$$
$$= \frac{b}{b-1}\log\left(1+\lambda^{b-1}\right).$$

To conclude, thanks to (53), we only need to prove that

$$(54)\qquad (\kappa\zeta_i, i\in\mathbb{N}) \overset{d}{=} \left(\mathrm{e}^{-\frac{T_i}{(b-1)}}\,W_i, i\in\mathbb{N}\right).$$

Since for $t\geq 0$, we have $X_t = \sum_{T_i\leq t} X^i_{t-T_i}$, we deduce from Fatou lemma that:

$$(55)\qquad W \geq \sum_{i\geq 0}\mathrm{e}^{-T_i/(b-1)}\,W_i.$$

We deduce from (53) that a.s.:

$$\left(\sum_{i=0}^{k} \mathrm{e}^{-\frac{T_i}{b-1}} W_i, k \in \mathbb{N}\right) \stackrel{d}{=} \left(\sum_{i=0}^{k} \mathrm{e}^{-\frac{T_i}{b-1}} (E^i)^{\frac{1}{b-1}} \sigma_1^i, k \in \mathbb{N}\right) \stackrel{d}{=} (\sigma_{S_k}, k \in \mathbb{N}),$$

where $S_k = \sum_{i=0}^k \mathrm{e}^{-T_i} E^i$. Notice that $(S_k, k \geq 0)$ is independent of $(\sigma_s, s \geq 0)$. Let $N(dr, ds)$ be a Poisson point measure on $\mathbb{R}_+ \times \mathbb{R}_+^*$ with intensity $s^{-1} \mathrm{e}^{-s} \, drds$. We define for all $t \geq 0$ $\Gamma_t = \int_{\mathbb{R}_+ \times \mathbb{R}_+^*} s \mathbf{1}_{\{r \leq t\}} N(dr, ds)$. Notice that $(\Gamma_t, t \geq 0)$ is a Gamma subordinator with Lévy measure $s^{-1} \mathrm{e}^{-s} \, ds$. Since $\sum_i \delta_{\{\mathrm{e}^{-T_i} E^i\}}(ds)$ is a Poisson point measure on $[0, \infty)$ with intensity $\frac{b}{b-1} s^{-1} \mathrm{e}^{-s} \, ds$, we get that $\{\mathrm{e}^{-T_i} E^i, i \in \mathbb{N}\}$ are the jump sizes of $\left(\Gamma_t, \, t \in [0, \frac{b}{b-1}]\right)$. We deduce that S_∞, and thus $\sum_{i \geq 0} \mathrm{e}^{-\frac{T_i}{b-1}} W_i$, are distributed as $\Gamma_{\frac{b}{b-1}}$, and then are $\Gamma\left(\frac{b}{b-1}, 1\right)$-distributed. Thanks to Remark 4.5, we deduce that $\sum_{i \geq 0} \mathrm{e}^{-\frac{T_i}{b-1}} W_i$ and W have the same distribution. Then we deduce from (55) that a.s.:

$$W = \sum_{i \geq 0} \mathrm{e}^{-T_i/(b-1)} W_i.$$

Then, Theorem 4.1 and Corollary 4.3 imply that (54) holds. □

Remark 4.11. From the proof of Proposition 4.10, we get that $\{\kappa \zeta_i, i \in \mathbb{N}\}$ are the jump sizes of $\{\sigma_{\Gamma_t}, 0 \leq t \leq \frac{b}{b-1}\}$, where $\left(\Gamma_t, \, t \in [0, \frac{b}{b-1}]\right)$ is a Gamma subordinator with Lévy measure $s^{-1} \mathrm{e}^{-s} \, ds$ independent of $\{\sigma_t : t \geq 0\}$, which is a $(b-1)$-stable subordinator $(\mathrm{E}[\mathrm{e}^{-x\sigma_t}] = \mathrm{e}^{-tx^{b-1}})$. This induces a Poisson-Kingman partition; see [40].

The distribution of $(\zeta_k, k \geq 0)$ is related to the Poisson-Dirichlet distribution in the quadratic case. Recall that $Z_0 = \sum_{k \in \mathbb{N}} \zeta_k$.

Corollary 4.12. *Consider the sub-critical quadratic branching mechanism with immigration* (7) *with* $b = 2$. *Let* $(\zeta_{(k)}, k \in \mathbb{N})$ *be the decreasing order statistics of* $(\zeta_k, k \in \mathbb{N})$. *Then, the random sequence* $\left(\zeta_{(k)}/Z_0, k \in \mathbb{N}\right)$ *has a Poisson-Dirichlet distribution with parameter* 2.

Proof. When $b = 2$, we have $\sigma(t) = t$. Then $\{\kappa \zeta_k, k \in \mathbb{N}\}$ are jump sizes of $\{\Gamma_t : 0 \leq t \leq 2\}$. The result follows from Proposition 5 in [41], see also [31]. □

Remark 4.13. Assume that $\psi(\lambda) = \alpha\lambda + \gamma\lambda^2$. According to (53) above and Theorem 2.21 in [19], we have the following so-called GEM representation: the sequence $(\zeta_k/Z_0, k \in \mathbb{N})$ is distributed as

$$(U_0, (1 - U_0)U_1, \cdots, (1 - U_0) \cdots (1 - U_{k-1})U_k, \cdots,), \tag{56}$$

where $\{U_i, i \geq 0\}$ are independent random variable with the same Beta-$(1, 2)$ distribution; see [19] and references therein. Moreover, Corollary 4.12 above and Theorem 2.7 in [19] give that the size-biased permutation of $\left(\zeta_{(k)}/Z_0, k \in \mathbb{N}\right)$ also has the same law as the family of age-ordered in (56).

When $b \in (1, 2)$, it does not seem possible to get a result similar to Corollary 4.12 or Remark 4.13, see Remark 4.16 below.

We consider the size-biased sample V of $(\zeta_k/Z_0, k \in \mathbb{N})$ under $\bar{\mathbb{P}}$. Let K be a $\mathbb{N}$-valued random variable such that, conditionally on $(\zeta_k/Z_0, k \in \mathbb{N})$, K is equal to k with probability ζ_k/Z_0. Then,

V is distributed as ζ_K/Z_0 under $\bar{\mathbb{P}}$:

$$\mathrm{P}(V \in dx) = \sum_{k\geq 0} x\, \bar{\mathbb{P}}(\zeta_k/Z_0 \in dx).$$

We shall also consider the size-biased sample ζ^* of $(\zeta_k, k \in \mathbb{N})$, which is distributed as ζ_K.

Recall function g from (50) and that $f_{b-1,b}$ defined in (43) is the density of κZ_0. Then with Proposition 4.10 and Remark 4.11 in hand, Theorem 2.1 of [38] implies that the distribution of V and $\kappa\zeta^*$ have densities given by:

$$f_V(x) = x \int_0^\infty t g(xt) f_{b-1,b}((1-x)t)\, dt \quad \text{for } x \in (0,1),$$

and

$$f_{\kappa\zeta^*}(x) = x g(x) \int_x^\infty f_{b-1,b}(t-x)\, \frac{dt}{t} \quad \text{for } x > 0.$$

See also (25) and (19) in Section 3 of [40]. In the following proposition we characterize the law of ζ^* via its Laplace transform and compute the moments of V. Recall $\bar{u}$ from (33). We set:

$$G = -\frac{\bar{u}'}{\bar{u}}\cdot$$

Proposition 4.14. *Consider the sub-critical stable branching mechanism with immigration* (32)*. We have for $\lambda \geq 0$,*

$$\mathrm{E}[\mathrm{e}^{-\lambda\zeta^*}] = \int_0^\infty G(\lambda+\mu)\bar{u}(\mu)\, d\mu, \tag{57}$$

and for $n \geq 1$:

$$\mathrm{E}[V^n] = \int_0^\infty v_n(t)\, dt \quad \textit{with} \quad v_n(t) = (-1)^n \frac{t^n}{n!} G^{(n)}(t)\bar{u}(t). \tag{58}$$

Proof. First, by property of Poisson point measure and (51), we get:

$$\begin{aligned}
\bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty} \zeta_i\, \mathrm{e}^{-\lambda\zeta_i - \mu Z_0}\right] &= \partial_\lambda \left(\partial_\rho\, \bar{\mathbb{E}}\left[\mathrm{e}^{-\rho \sum_{i=0}^{+\infty} \mathrm{e}^{-\lambda\zeta_i} - \mu Z_0}\right]\right)_{|\rho=0} \\
&= -\exp\left\{-\int_0^\infty (1-\mathrm{e}^{-\mu x})\kappa g(\kappa x)\, dx\right\} \partial_\lambda \int_0^\infty \mathrm{e}^{-(\mu+\lambda)x}\, \kappa g(\kappa x)\, dx \\
&= \exp\left\{-\int_0^\infty (1-\mathrm{e}^{-\mu x})\kappa g(\kappa x)\, dx\right\} \int_0^\infty x\, \mathrm{e}^{-(\mu+\lambda)x}\, \kappa g(\kappa x)\, dx \\
&= \bar{u}(\mu) G(\lambda+\mu).
\end{aligned}$$

Then (57) follows from

$$\mathrm{E}[\mathrm{e}^{-\lambda\zeta^*}] = \bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty} \frac{\zeta_i}{Z_0}\, \mathrm{e}^{-\lambda\zeta_i}\right] = \int_0^\infty d\mu\, \bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty} \zeta_i\, \mathrm{e}^{-\lambda\zeta_i - \mu Z_0}\right].$$

Next, observe that for $n \geq 1$,

$$\bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty} \zeta_i^n\, \mathrm{e}^{-\mu Z_0}\right] = (-1)^{n-1} \left(\partial_\lambda^{n-1}\, \bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty} \zeta_i\, \mathrm{e}^{-\lambda\zeta_i - \mu Z_0}\right]\right)_{|\lambda=0} = (-1)^{n-1} G^{(n-1)}(\mu)\bar{u}(\mu).$$

We deduce that:

$$\begin{aligned}
\mathrm{E}[V^n] &= \bar{\mathbb{E}}\left[\sum_{i=0}^{+\infty}\left(\frac{\zeta_i}{Z_0}\right)^{n+1}\right] \\
&= \int_{(0,+\infty)^n} dt_1 \dots dt_n\, \mathbf{1}_{\{0<t_1<t_2\dots<t_n\}} \int_{t_n}^{\infty} (-1)^n G^{(n)}(t)\bar{u}(t)\, dt \\
&= \int_0^{\infty} (-1)^n \frac{t^n}{n!} G^{(n)}(t)\, \bar{u}(t)\, dt.
\end{aligned}$$

This finishes the proof. □

Remark 4.15. The moment of V in (58) can be computed explicitly. Set $\eta = b-1$ and $a = \gamma/\alpha$. Recall from (33) that $\bar{u}(t) = (1+at^\eta)^{-(\frac{1}{\eta}+1)}$. As $G^{(n-1)}(t)$ is a linear combination of functions $t^{k\eta-n}(1+at^\eta)^{-k}$ for $k \in \{1,\dots,n\}$, one gets that for $n \geq 1$:

$$\lim_{t\to 0+} t^n G^{(n-1)}(t)\, \bar{u}(t) = \lim_{t\to+\infty} t^n G^{(n-1)}(t)\, \bar{u}(t) = 0.$$

Recall v_n defined in (58), so that for $n \geq 0$ and $k \geq 0$:

$$\frac{v_n(t)}{(1+at^\eta)^k} = (-1)^n \frac{t^n}{n!} \frac{G^{(n)}(t)}{(1+at^\eta)^{k+1+1/\eta}}\cdot$$

Then, for $n \geq 1$ and $k \geq 0$, by integration by parts, we have

$$\begin{aligned}
(59)\quad & \int_0^\infty \frac{v_n(t)}{(1+at^\eta)^k}\, dt \\
&= \left(1 - \frac{\eta(k+1)+1}{n}\right)\int_0^\infty \frac{v_{n-1}(t)}{(1+at^\eta)^k}\, dt + \frac{\eta(k+1)+1}{n}\int_0^\infty \frac{v_{n-1}(t)}{(1+at^\eta)^{k+1}} dt,
\end{aligned}$$

and

$$\int_0^\infty \frac{v_0(t)}{(1+at^\eta)^k}\, dt = \int_0^\infty \frac{(-\bar{u}'(t))}{(1+at^\eta)^k}\, dt = \int_0^\infty \frac{a(1+\frac{1}{\eta})\eta t^{\eta-1}}{(1+at^\eta)^{k+2+\frac{1}{\eta}}} dt = \frac{\eta+1}{\eta(k+1)+1}\cdot$$

The previous recursion formula gives the value of $\int_0^\infty v_n(t)(1+at^\eta)^{-k}\, dt$ for all $n \geq 0$ and $k \geq 0$. Using (59) with $k=0$, we get:

$$(60)\qquad \mathrm{E}[V^n] = \left(1 - \frac{(\eta+1)}{n}\right)\int_0^\infty v_{n-1}(t)\, dt + \frac{\eta+1}{n}\int_0^\infty \frac{v_{n-1}(t)}{1+at^\eta}\, dt.$$

In particular, one has:

$$\begin{aligned}
\mathrm{E}[V] &= -\eta + (\eta+1)\frac{\eta+1}{2\eta+1} = \frac{-\eta^2+\eta+1}{2\eta+1}, \\
\mathrm{E}[V^2] &= (1 - \frac{\eta+1}{2})E[X] - \frac{\eta(\eta+1)^2}{2\eta+1} + \frac{(\eta+1)^2(2\eta+1)}{2(3\eta+1)} \\
&= \frac{\eta^4 - 7\eta^3 + \eta^2 + 7\eta + 2}{2(2\eta+1)(3\eta+1)}, \\
\mathrm{E}[V^3] &= \frac{23\eta^5 - 80\eta^4 - 30\eta^3 + 74\eta^2 + 43\eta + 6}{6(2\eta+1)(3\eta+1)(4\eta+1)}\cdot
\end{aligned}$$

Remark 4.16. Based on the moment formulas above, one can check that V is not Beta-distributed if $b < 2$ (except maybe for one particuler value of η ($\eta \simeq 0.428$) where the three first moments of V coincide with those of a Beta distribution). If $b = 2$, then according to Remark 4.13, V is Beta(1,2)-distributed.

5. Birth and death rates of the ancestral process

In this section, we assume that ψ and ϕ are defined as (1) and (5), respectively. We assume that Conditions (2) hold. Recall the function $\bar{u}$ defined in (16) and (6) which is the Laplace transform of Z_0 under $\bar{\mathbb{P}}$. Similar to the arguments on page 1330 in [10], see also Proposition 3.12 in [14], we have for $r \geq s > 0$ and $x, y \in [0,1]$:

$$\bar{\mathbb{E}}[x^{M^0_{-r}} y^{M^0_{-s}}] = \bar{u}(\lambda_0)\, \mathrm{e}^{\alpha(r-s)}\, \frac{\psi\Big(u\big(c(s)(1-y), r-s\big)\Big)}{\psi\big(c(s)(1-y)\big)}, \tag{61}$$

with

$$\lambda_0 = \lambda_0(x, y) = c(r)(1-x) + xu\big(c(s)(1-y), r-s\big). \tag{62}$$

We first summarize the results of the next two sections concerning the quadratic case, see also [10].

Remark 5.1. In the quadratic case, $\psi(\lambda) = \alpha\lambda + \gamma\lambda^2$, we have:

$$c(t) = \frac{\alpha}{\gamma(\mathrm{e}^{\alpha t} - 1)} \quad \text{and} \quad \bar{u}(\lambda) = \Big(1 + \frac{\gamma}{\alpha}\lambda\Big)^{-2}.$$

We get thanks to (61) and (62) (taking $r = s = t$ and $x = y$) that for $t > 0$ and $x \in [0,1]$:

$$\bar{\mathbb{E}}\left[x^{M^0_{-t}}\right] = \bar{u}(c(t)(1-x)) = \left(\frac{\mathrm{e}^{\alpha t} - 1}{\mathrm{e}^{\alpha t} - x}\right)^2.$$

We get that for $n \geq 0$:

$$\bar{\mathbb{P}}(M^0_{-t} = n) = (n+1)\, \mathrm{e}^{-\alpha t n} \left(1 - \mathrm{e}^{-\alpha t}\right)^2.$$

For the death rate, we deduce from (63) that for $n \geq 1$: $q^{\mathrm{d}}_{n,m}(t) = 0$ if $n - 2 \geq m \geq 0$ and if $m = n - 1$

$$q^{\mathrm{d}}_{n,n-1}(t) = n(\alpha + \gamma c(t)).$$

For the birth rate, we deduce from (79) that for $n \geq 0$: $q^{\mathrm{b}}_{n,m}(-t) = 0$ if $m \geq n + 2$ and if $m = n + 1$

$$q^{\mathrm{b}}_{n,n+1}(-t) = (n+2)\gamma c(t).$$

5.1. **Death process.** Recall the ancestral process $M^0 = (M^0_t, t < 0)$ defined in Section 3.2. Notice that the branching property gives that the ancestral process is a Markov process. We first study the death rate of the time reversed ancestral process $\hat{M}^0 = (M^0_{-t}, t > 0)$. Notice that $\hat{M}^0$ is a Markov process as the time reversal of a Markov process. Recall that for a birth and death Markov process, its distribution is completely characterized by its birth and death rates, that is its infinitesimal generator, see [13, Section 1.1].

Proposition 5.2. *Let ψ and ϕ be defined by (1) and (5) such that conditions (2) hold. The process $\hat{M}^0$ is a càd-làg death process starting at time 0 from $+\infty$ and with death rate given for $n > m \geq 0$ and $t > 0$ by:*

$$q^{\mathrm{d}}_{n,m}(t) = \lim_{\varepsilon \to 0+} \frac{1}{\varepsilon}\, \bar{\mathbb{P}}\left(M^0_{-(t+\varepsilon)} = m \middle| M^0_{-t} = n\right) = \binom{n+1}{m} \frac{\left|\bar{u}^{(m)}\big(c(t)\big)\right|}{\left|\bar{u}^{(n)}\big(c(t)\big)\right|} \left|\psi^{(n-m+1)}\big(c(t)\big)\right|. \tag{63}$$

There is no closed formula for the stable case unless it is quadratic. However, the next lemma gives an explicit asymptotic for the birth rate when the stable index b goes down to 1.

Lemma 5.3. *Consider the sub-critical stable branching mechanism with immigration* (32)*, that is* $\psi(\lambda) = \alpha\lambda + \gamma\lambda^b$ *with* $\alpha > 0$ *and* $b \in (1, 2]$*. Then we have for* $n > m \geq 0$ *and* $t > 0$*:*

$$\lim_{b\to 1+} q^{\mathrm{d}}_{n,m}(t) = \frac{n+1}{(n+1-m)(n-m)}\,\frac{1}{\gamma t}.$$

Proof. Recall that in the stable case (see (33)):

$$\bar{u}(\lambda) = \left(1 + \frac{\gamma}{\alpha}\lambda^{b-1}\right)^{-\frac{b}{b-1}}, \quad c(t) = \left(\frac{\alpha}{\gamma\left(\mathrm{e}^{(b-1)\alpha t} - 1\right)}\right)^{\frac{1}{b-1}}.$$

One can check that $\bar{u}^{(n)}(\lambda)$, for $n \geq 1$, has the form:

$$\bar{u}^{(n)}(\lambda) = \sum_{k=1}^{n} C_{b,k,n} \left(\frac{\gamma}{\alpha}\right)^k \left(1 + \frac{\gamma}{\alpha}\lambda^{b-1}\right)^{-\frac{b}{b-1}-k} \lambda^{-(2-b)k-(n-k)},$$

where $C_{b,k,n}$ are constants depending only on b, k and n and such that $(-1)^n C_{b,k,n} \geq 0$. On the other hand, writing λ^{-b} as $(\lambda^{b-1})^{-\frac{b}{b-1}}$, one sees that, with the same constants $C_{b,k,n}$:

$$\left(\frac{\alpha}{\gamma}\right)^{\frac{b}{b-1}} \left(\lambda^{-b}\right)^{(n)} = \left(\left(\frac{\gamma}{\alpha}\lambda^{b-1}\right)^{-\frac{b}{b-1}}\right)^{(n)} = \sum_{k=1}^{n} C_{b,k,n} \left(\frac{\gamma}{\alpha}\right)^k \left(\frac{\gamma}{\alpha}\lambda^{b-1}\right)^{-\frac{b}{b-1}-k} \lambda^{-(2-b)k-(n-k)}.$$

Note that:

$$\left(1 + \frac{\gamma}{\alpha}c(t)^{b-1}\right)^{-\frac{b}{b-1}} = \left(\frac{\gamma}{\alpha}c(t)^{b-1}\right)^{-\frac{b}{b-1}} \mathrm{e}^{-\alpha b t}.$$

Since the constants $C_{b,k,n}$ have all the same sign for given n, we deduce that for $n \geq 1$, as $b \to 1+$:

$$\begin{aligned}
\bar{u}^{(n)}(c(t)) &\sim \left(\frac{\alpha}{\gamma}\right)^{\frac{b}{b-1}} (\lambda^{-b})^{(n)}_{|\lambda=c(t)}\, \mathrm{e}^{-\alpha b t} \\
&= (-b)(-b-1)\cdots(-b-n+1)c(t)^{-b-n} \left(\frac{\alpha}{\gamma}\right)^{\frac{b}{b-1}} \mathrm{e}^{-\alpha b t} \\
&\sim (-1)^n n!\, c(t)^{-b-n} \left(\frac{\alpha}{\gamma}\right)^{\frac{b}{b-1}} \mathrm{e}^{-\alpha b t}.
\end{aligned}$$

We deduce that:

$$\begin{aligned}
\lim_{b\to 1+} q^{\mathrm{d}}_{n,m}(t) &= \lim_{b\to 1+} \binom{n+1}{m} \frac{\left|\bar{u}^{(m)}\big(c(t)\big)\right|}{\left|\bar{u}^{(n)}\big(c(t)\big)\right|} \left|\psi^{(n-m+1)}\big(c(t)\big)\right| \\
&= \lim_{b\to 1+} \binom{n+1}{m} \frac{m!}{n!} c(t)^{n-m} \left| b(b-1)\cdots(b-n+m)c(t)^{b-n+m-1} \right| \\
&= \lim_{b\to 1+} \binom{n+1}{m} \frac{(n-1-m)!\,m!}{n!} (b-1)c(t)^{b-1} \\
&= \frac{n+1}{(n+1-m)(n-m)}\,\frac{1}{\gamma t}.
\end{aligned}$$

□

Proof of Proposition 5.2. The proof is divided in three steps.

Step 1: Preliminary computations.
We set for $\lambda, \mu \in [0,1]$ and $t, \varepsilon > 0$:

$$\begin{aligned} g_{t,\varepsilon}(\mu) &= \mathrm{e}^{\alpha\varepsilon} \frac{\psi\Big(u\big(c(t)(1-\mu), \varepsilon\big)\Big)}{\psi\big(c(t)(1-\mu)\big)}, \\ \lambda^*_{t,\varepsilon} = \lambda^*_{t,\varepsilon}(\lambda,\mu) &= c(t+\varepsilon)(1-\lambda) + \lambda u\big(c(t)(1-\mu),\varepsilon\big), \\ f^{\mathrm{d}}_{t,\varepsilon}(\lambda,\mu) &= \bar{u}(\lambda^*_{t,\varepsilon}) g_{t,\varepsilon}(\mu), \\ f_0(\mu) &= \bar{u}\big(c(t)(1-\mu)\big). \end{aligned}$$

Thanks to (61) and (62), we deduce that:

$$f^{\mathrm{d}}_{t,\varepsilon}(\lambda,\mu) = \bar{\mathbb{E}}\left[\lambda^{M^0_{-(t+\varepsilon)}} \mu^{M^0_{-t}}\right] \quad \text{and} \quad f_0(\mu) = f^{\mathrm{d}}(1,\mu) = \bar{\mathbb{E}}\left[\mu^{M^0_{-t}}\right]. \tag{64}$$

We get for $n > m \geq 0$:

$$q^{\mathrm{d}}_{n,m}(t) = \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon} \frac{\bar{\mathbb{P}}(M^0_{-(t+\varepsilon)} = m, M^0_{-t} = n)}{\bar{\mathbb{P}}(M^0_{-t} = n)} = \lim_{\varepsilon\to 0+} \frac{1}{\varepsilon} \frac{\partial^n_\mu \partial^m_\lambda\, f^{\mathrm{d}}_{t,\varepsilon}(0,0)}{m!\, f^{(n)}_0(0)}. \tag{65}$$

First notice that for $n \geq 1$:

$$f^{(n)}_0(0) = (-1)^n \bar{u}^{(n)}(c(t))\, c(t)^n. \tag{66}$$

We now study $\partial^n_\mu \partial^m_\lambda\, f^{\mathrm{d}}_{t,\varepsilon}(0,0)$. We set:

$$I_{t,\varepsilon}(\mu) = \partial_\lambda\, \lambda^*_{t,\varepsilon}(\lambda,\mu) = u\big(c(t)(1-\mu),\varepsilon\big) - c(t+\varepsilon).$$

Notice that $g_{t,\varepsilon}(\mu)$ and $I_{t,\varepsilon}(\mu)$ are independent of λ. We deduce that for $m \geq 0$:

$$\partial^m_\lambda\, f^{\mathrm{d}}_{t,\varepsilon}(\lambda,\mu) = \bar{u}^{(m)}(\lambda^*_{t,\varepsilon}) I_{t,\varepsilon}(\mu)^m\, g_{t,\varepsilon}(\mu). \tag{67}$$

We also note that for $k \geq 1$,

$$I^{(k)}_{t,\varepsilon}(\mu) = (-1)^k c(t)^k\, \partial^k_\lambda u\big(c(t)(1-\mu),\varepsilon\big) \quad \text{and} \quad \partial^k_\mu\, \lambda^*_{t,\varepsilon}(\lambda,\mu) = \lambda\, I^{(k)}_{t,\varepsilon}(\mu). \tag{68}$$

We deduce that for $k \geq 1$:

$$\partial^k_\mu\, \lambda^*_{t,\varepsilon}(0,0) = 0 \quad \text{and} \quad \partial^k_\mu\, \bar{u}\big(\lambda^*_{t,\varepsilon}(\lambda,\mu)\big)_{|(\lambda,\mu)=(0,0)} = 0. \tag{69}$$

We end this first step by a remark. We deduce from

$$\partial_\lambda u(\lambda,t) = \frac{\psi(u(\lambda,t))}{\psi(\lambda)} \tag{70}$$

(see (12)), Equation (11) and elementary computations that:

$$\partial^k_\lambda\, u(\lambda,\varepsilon) = \begin{cases} 1 + o(1) & \text{if } k = 1, \\ -\psi^{(k)}(\lambda)\,\varepsilon + o(\varepsilon) & \text{if } k \geq 2, \end{cases} \tag{71}$$

as ε goes down to 0, where $o(1)$ means a quantity which goes down to 0 with ε.

Step 2: Study of $\partial^n_\mu\, (I^m_{t,\varepsilon})(0)$.
We now study the value of $\partial^n_\mu\, (I^m_{t,\varepsilon})(0)$ for $n \geq 0$ and $m \geq 0$ and $(n,m) \neq (0,0)$. The case $n > m = 0$ is trivial as $\partial^n_\mu\, (I^m_{t,\varepsilon})(0) = 0$. We have for all $m > n \geq 0$:

$$I_{t,\varepsilon}(0) = 0 \quad \text{and thus} \quad \partial^n_\mu\, (I^m_{t,\varepsilon})(0) = 0. \tag{72}$$

For the case $m = 1$, we deduce from (68) and (71) that for $n \geq 1$:

$$I_{t,\varepsilon}^{(n)}(0) = \begin{cases} -c(t) + o(1) & \text{if } n = 1, \\ (-1)^{n+1} c(t)^n \, \psi^{(n)}\big(c(t)\big)\varepsilon + o(\varepsilon) & \text{if } n \geq 2. \end{cases} \tag{73}$$

For $n = m$, Faa di Bruno's formula, $I_{t,\varepsilon}(0) = 0$ (see (72)) and (73) give that:

$$\partial_\mu^m \, (I_{t,\varepsilon}^m)(0) = m!(-1)^m c(t)^m + o(1). \tag{74}$$

We shall prove by induction over $m \geq 1$ that for all $n > m \geq 1$:

$$\partial_\mu^n \, (I_{t,\varepsilon}^m)(0) = \binom{n}{m-1} m!(-1)^{n+1} c(t)^n \, \psi^{(n-m+1)}\big(c(t)\big)\, \varepsilon + o(\varepsilon). \tag{75}$$

Thanks to (73), we get that (75) holds for $m = 1$ and all $n > m$. Let us assume that (75) holds for $m - 1$ (and all $n > m - 1$), and let us prove it holds for m (and all $n > m$). We have for $n > m$:

$$\begin{aligned}
\partial_\mu^n \, (I_{t,\varepsilon}^m)(0) &= \sum_{k=0}^n \binom{n}{k} I_{t,\varepsilon}^{(k)}(0) \; \partial_\mu^{n-k} (I_{t,\varepsilon}^{m-1})(0) \\
&= n I_{t,\varepsilon}^{(1)}(0) \, \partial_\mu^{n-1} (I_{t,\varepsilon}^{m-1})(0) + \binom{n}{m-1} I_{t,\varepsilon}^{(n-m+1)}(0) \, \partial_\mu^{m-1} \, (I_{t,\varepsilon}^{m-1})(0) + O(\varepsilon^2) \\
&= \left[n(m-1)! \binom{n-1}{m-2} + (m-1)! \binom{n}{m-1} \right] (-1)^{n+1} c(t)^n \, \psi^{(n-m+1)}\big(c(t)\big)\varepsilon + o(\varepsilon) \\
&= \binom{n}{m-1} m!(-1)^{n+1} c(t)^n \, \psi^{(n-m+1)}\big(c(t)\big)\, \varepsilon + o(\varepsilon),
\end{aligned}$$

where, for the second equality we used that $I_{t,\varepsilon}(0) = 0$ (see (72)) for the term $k = 0$, then $\partial_\mu^{n-k} \, (I_{t,\varepsilon}^{m-1})(0) = 0$ (see (72)) for the terms $k > n - m + 1$, and then $I_{t,\varepsilon}^{(k)}(0) \; \partial_\mu^{n-k} \, (I_{t,\varepsilon}^{m-1})(0) = O(\varepsilon^2)$ (see (73) and the induction hypothesis) for $n - m + 1 > k \geq 2$; and for the third equality (73) (for $k = 1$ and $k = n - m + 1$), the induction hypothesis and (74). Thus (75) holds for all $n > m \geq 1$.

Step 3: Computation of $q_{n,m}^{\mathrm{d}}(t)$.

If we derive (67) n times with respect to μ and evaluate the derivative at $(0, 0)$, we get for $n > m \geq 0$:

$$\begin{aligned}
\partial_\mu^n \partial_\lambda^m \, f_{t,\varepsilon}^{\mathrm{d}}(0,0) &= \bar{u}^{(m)}\big(c(t+\varepsilon)\big) \partial_\mu^n \, (I_{t,\varepsilon}^m \, g_{t,\varepsilon})(0) \\
&= \bar{u}^{(m)}\big(c(t+\varepsilon)\big) \sum_{k=m}^n \binom{n}{k} g_{t,\varepsilon}^{(n-k)}(0) \, \partial_\mu^k \, (I_{t,\varepsilon}^m)(0),
\end{aligned} \tag{76}$$

where for the first equality we used that all the terms in Leibniz' formula are 0 except one thanks to (69), and for the second Leibniz' formula again with (72).

Since $g_{t,\varepsilon}(\mu) = \mathrm{e}^{\alpha\varepsilon} \, \partial_\lambda u\big(c(t)(1-\mu), \varepsilon\big)$, see (70), we deduce from (71) that, for $k \geq 1$:

$$g_{t,\varepsilon}^{(k)}(0) = (-1)^{k+1} c(t)^k \, \psi^{(k+1)}\big(c(t)\big)\, \varepsilon + o(\varepsilon). \tag{77}$$

This and (73) imply that for $n - 1 > m \geq 0$:

$$\sum_{k=m+1}^{n-1} \binom{n}{k} g_{t,\varepsilon}^{(n-k)}(0) \, \partial_\mu^k \, (I_{t,\varepsilon}^m)(0) = O(\varepsilon^2) = o(\varepsilon).$$

Then, we deduce from (76) and (75) that for $t>0$ and $n>m\geq 0$ (with the convention that $\binom{n}{m-1}=0$ if $m=0$):

$$\begin{aligned}\partial_\mu^n\partial_\lambda^m\, f^{\rm d}_{t,\varepsilon}(0,0) &= \bar u^{(m)}\big(c(t+\varepsilon)\big)\left[\binom{n}{m}\partial_\mu^m\,(I^m_{t,\varepsilon})(0)\, g^{(n-m)}_{t,\varepsilon}(0)+\partial_\mu^n\,(I^m_{t,\varepsilon})(0)\, g_{t,\varepsilon}(0)\right]+o(\varepsilon)\\ &= \bar u^{(m)}\big(c(t)\big)\left[\binom{n}{m}+\binom{n}{m-1}\right] m!(-1)^{n+1}c(t)^n\,\psi^{(n-m+1)}\big(c(t)\big)\,\varepsilon+o(\varepsilon)\\ &= \binom{n+1}{m}\bar u^{(m)}\big(c(t)\big)m!(-1)^{n+1}c(t)^n\,\psi^{(n-m+1)}\big(c(t)\big)\,\varepsilon+o(\varepsilon). \end{aligned} \tag{78}$$

Notice that $o(\varepsilon)$ in the last equality is uniform on $t\in[a,b]$ far any given $0<a<b<+\infty$. We then deduce from the latter equality, (65) and (66) that for $n>m\geq 0$:

$$q^{\rm d}_{n,m}(t)=-\binom{n+1}{m}\frac{\bar u^{(m)}\big(c(t)\big)\,\psi^{(n-m+1)}\big(c(t)\big)}{\bar u^{(n)}\big(c(t)\big)}.$$

This finishes the proof. □

5.2. **Birth process.** Recall that the ancestral process $M^0=(M^0_t,t<0)$ defined in Section 3.2 is a Markov process thanks to the branching property.

Proposition 5.4. *Let ψ and ϕ be defined by (1) and (5) such that Conditions (2) hold. The process M^0 is a càd-làg birth process starting at time $-\infty$ from* 0 *and with birth rate given for $n>m\geq 0$ and $t>0$ by:*

$$q^{\rm b}_{n,m}(-t)=\lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}\bar{\mathbb P}\left(M^0_{-(t-\varepsilon)}=m|M^0_{-t}=n\right)=\frac{(m+1)}{(m+1-n)!}\,c(t)^{m-n}\,\left|\psi^{(m-n+1)}\big(c(t)\big)\right|. \tag{79}$$

Concerning the birth rate, it is possible to have an explicit formula in the stable case.

Remark 5.5. Consider the sub-critical stable branching mechanism with immigration (32). Using (33), we deduce that for $m>n\geq 0$ and $t>0$:

$$q^{\rm b}_{n,m}(-t)=\frac{(m+1)}{(m-n+1)!}\,|b(b-1)\cdots(b-m+n)|\,\frac{\alpha}{\mathrm{e}^{(b-1)\alpha t}-1}.$$

Proof of Proposition 5.4. We keep notations from the proof of Proposition 5.2 for $f^{\rm d}_{t,\varepsilon}$ and f_0. We set for $\lambda,\mu\in[0,1]$ and $t>\varepsilon>0$ $f^{\rm b}_{t,\varepsilon}(\lambda,\mu)=f^{\rm d}_{t-\varepsilon,\varepsilon}(\mu,\lambda)$ for $\lambda,\mu\in[0,1]$ and $t>\varepsilon>0$. Thanks to (64), we have that:

$$f^{\rm b}_{t,\varepsilon}(\lambda,\mu)=f^{\rm d}_{t-\varepsilon,\varepsilon}(\mu,\lambda)=\bar{\mathbb E}\left[\lambda^{M^0_{-t+\varepsilon}}\mu^{M^0_{-t}}\right].$$

Recall f_0 defined in (64) and its derivative given by (66). We get for $m>n\geq 0$:

$$q^{\rm b}_{n,m}(-t)=\lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}\frac{\bar{\mathbb P}(M^0_{-t+\varepsilon}=m,M^0_{-t}=n)}{\bar{\mathbb P}(M^0_{-t}=n)}=\lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}\frac{\partial_\mu^n\partial_\lambda^m\, f^{\rm b}_{t,\varepsilon}(0,0)}{m!\, f_0^{(n)}(0)}. \tag{80}$$

Since $\partial_\mu^n\partial_\lambda^m\, f^{\rm b}_{t,\varepsilon}(0,0)=\partial_\mu^m\partial_\lambda^n\, f^{\rm d}_{t-\varepsilon,\varepsilon}(0,0)$, using the continuity in ε of the function c, we deduce from (78), and the fact that $o(\varepsilon)$ in (78) is uniform in t on any closed interval of $(0,+\infty)$, that for $m>n\geq 0$ and $t>\varepsilon>0$:

$$\begin{aligned}\partial_\mu^n\partial_\lambda^m\, f^{\rm b}_{t,\varepsilon}(0,0) &= \binom{m+1}{n}\bar u^{(n)}\big(c(t-\varepsilon)\big)n!(-1)^{m+1}c(t-\varepsilon)^m\,\psi^{(m-n+1)}\big(c(t-\varepsilon)\big)\,\varepsilon+o(\varepsilon)\\ &= \binom{m+1}{n}\bar u^{(n)}\big(c(t)\big)n!(-1)^{m+1}c(t)^m\,\psi^{(m-n+1)}\big(c(t)\big)\,\varepsilon+o(\varepsilon).\end{aligned}$$

We then deduce from the latter equality, (80) and (66) that for $n > m \geq 0$:

$$q^{\mathrm{b}}_{n,m}(-t) = (-1)^{m-n+1} \frac{m+1}{(m+1-n)!} \, c(t)^{m-n} \, \psi^{(m-n+1)}\big(c(t)\big).$$

This finishes the proof. □

5.3. **Bolthausen-Sznitman coalescent as limit of the ancestral process.** The Bolthausen-Sznitman coalescent, $(\Pi(t), t \geq 0)$, is a continuous-time Markov chain taking values in the set of patitions of $\mathbb{N}^*$. It can be easily defined by considering its restriction $\Pi^{[n]} = (\Pi^{[n]}(t), t \geq 0)$ to the set $[n] := \{1, 2, \cdots, n\}$, for $n \geq 1$. Denote by $\mathcal{P}_n$ be the set of partitions of $[n]$. Then, the process $\Pi^{[n]}$ is a continuous-time $\mathcal{P}_n$-valued Markov chain whose transition rates are as follows: if $\#\Pi^{[n]}(t) = k$, then any m of the present blocks coalesce at rate

$$\frac{(m-2)!(k-m)!}{(k-1)!}, \quad 2 \leq m \leq k \leq n,$$

where $\#\Pi^{[n]}(t)$ denotes the number of blocks of $\Pi^{[n]}(t)$. The Bolthausen-Sznitman coalescent was first introduced in [12]. It is also a member of the class of coalescents with multiple collisions introduced in [39] and [42]. We refer to the survey [8] for further results on coalescent processes.

Other constructions of the Bolthausen-Sznitman appear in the literature. See [9] using the genealogy of a continuous state branching process (the corresponding branching mechanism corresponds in some sense to the limit in (7) as b goes down to 1), [21] using a uniform pruning of the branches of a random recursive tree, and [43] using limit of ancestral processes obtained from super-critical Galton-Watson processes; see also references therein for other related results.

Let us consider the ancestral tree $\mathcal{A}(0)$ from Definition 2.4 associated with the stable Lévy forest under $\bar{\mathbb{P}}$ (that is for the stationary regime). Let $T > 0$. Conditionally on $\{M^0_{-T} = n-1\}$, that is the number of individuals of $\mathcal{A}(0)$ at level $-T$ is n, we label all the n individuals from 1 to n uniformly at random. Define a continuous time $\mathcal{P}_n$-valued process $(\Pi^{T,[n]}(t), t \geq T)$, where $\Pi^{T,[n]}(t)$ is the partition of $[n]$ such that i and j are in the same block if and only if the i-th and j-th individuals at level $-T$ have the same ancestor at level $-t$ of the ancestral tree $\mathcal{A}(0)$. By construction, as $\lim_{t\to+\infty} M^0_{-t} = 0$, we have that a.s. $\lim_{t\to+\infty} \Pi^{T,[n]}(t) = [n]$.

Proposition 5.6. *Consider the sub-critical stable branching mechanism with immigration* (32). *The law of* $(\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}), t \geq 0)$*, under* $\bar{\mathbb{P}}(\cdot | M^0_{-T} = n-1)$*, converges in the sense of finite dimensional distribution to a Bolthausen-Sznitman coalescent* $\Pi^{[n]}$*, as* b *decreases to 1.*

Proof. Let $(\chi_t, t \geq 0)$ be the GW process introduced in Remark 4.2 with branching rate 1 and offspring distribution with generating function g_B defined in (34). It is well-known that, conditioned on $\{\chi_T = n\}$, $(\chi_t, 0 \leq t \leq T)$ is an in-homogeneous Markov process. But, by Remark 4.2 and Theorem 1.1, we get that the process $(\chi_t, 0 \leq t \leq T)$ conditionally on $\{\chi_T = n\}$ is distributed as the process $(\tilde{M}_t + 1, 0 \leq t \leq T)$ conditionally on $\{\tilde{M}_T + 1 = n\}$. Thus, the latter is a Markov process. Similar arguments on the genealogical tree imply that the process $(\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}), t \geq 0)$ is Markov (but inhomogeneous in time).

According to [26, Theorem 4.1], it is sufficient to prove that the transition rates of the process $(\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}), t \geq 0)$ converge to those of the process $\Pi^{[n]}$, as $b \to 1+$. We also notice that

$$(\#\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}) - 1, t \geq 0) = (M^0_{-T\,\mathrm{e}^{\gamma t}}, t \geq 0)$$

and that the generations do not overlap. Thus if $\#\Pi^{T,[n]}(T\,\mathrm{e}^{\gamma t}) = k$, then any m of the present blocks coalesce at rate

$$\frac{T\gamma\,\mathrm{e}^{\gamma t}}{\binom{k}{m}} q^{\mathrm{d}}_{k-1,k-m}(T\,\mathrm{e}^{\gamma t})$$

where the death rates $q^{\mathrm{d}}_{n,m}(t) = \lim_{\varepsilon\to 0+} \varepsilon^{-1}\bar{\mathbb{P}}(M^0_{-(t+\varepsilon)} = m | M^0_{-t} = n)$ are computed in Section 5.1 for general branching mechanism. Using Lemma 5.3, we deduce that for $2 \leq m \leq k \leq n$:

$$\lim_{b\to 1+} \frac{T\gamma\,\mathrm{e}^{\gamma t}}{\binom{k}{m}} q^{\mathrm{d}}_{k-1,k-m}(T\,\mathrm{e}^{\gamma t}) = \frac{(m-2)!(k-m)!}{(k-1)!}\cdot$$

This proves the result. □

6. Critical stable case

In this section only, we shall consider the critical stable case with branching mechanism ψ and immigration ϕ given by:

$$\psi(\lambda) = \gamma\lambda^b \quad \text{and} \quad \phi(\lambda) = b\gamma\lambda^{b-1}, \tag{81}$$

with $\gamma > 0$ and $b \in (1, 2]$. We also have (see Example 3.1 p. 62 in [35]) for $\lambda \geq 0$ and $t > 0$:

$$u(\lambda, t) = \frac{\lambda}{(1 + \gamma(b-1)\lambda^{b-1}t)^{1/(b-1)}} \quad \text{and} \quad c(t) = (\gamma(b-1)t)^{-\frac{1}{b-1}}. \tag{82}$$

In this setting, both M^0_{-t} and Z_0 are infinite. For this reason, we only consider the families migrating to the system after some time $-T$.

Let $T > 0$. Recall π is the Lévy measure in (1), $\mathbb{N}$ is the corresponding excursion measure on $\mathbb{T}$ of the Lévy tree, and $\mathbb{P}_r(d\bar{\mathbf{f}})$ is the probability distribution on $\mathbb{F}$ of the random forest $\mathcal{F} = (\mathcal{T}_i)_{i\in I}$ given by the atoms of a poisson point measure on $\mathbb{T}$ with intensity $r\mathbb{N}(d\mathbf{t})$. Similarly to Section 3.1, we consider under $\bar{\mathbb{P}}$ a random leveled forest $\bar{\mathcal{F}}^{(T)} = (h_i, \mathcal{F}_i)_{i\in I^{(T)}}$ given by the atoms of a Poisson point measure on $[-T, +\infty] \times \mathbb{F}$ with intensity

$$\nu(dh, d\mathbf{f}) = \mathbf{1}_{[-T,+\infty)}(h)\, dh \left(\beta_1 \mathbb{N}[d\mathbf{f}] + \int_0^{+\infty} \pi(dr)\, \mathbb{P}_r(d\mathbf{f})\right),$$

and let $\bar{\mathcal{T}}^{(T)} = \mathbf{t}(\bar{\mathcal{F}}^{(T)})$ be the random tree associated with this leveled forest. Set for $a > -T$:

$$\ell^a(\bar{\mathcal{T}}^{(T)}) = \sum_{i\in I^{(T)}} \ell^{a-h_i}(\mathcal{F}_i)\mathbf{1}_{\{h_i\leq a\}} \quad \text{and} \quad Z^{(T)}_a = \langle \ell^a(\bar{\mathcal{T}}^{(T)}), 1\rangle.$$

Thanks to the properties of Poisson point measures, we have, for $\lambda \geq 0$, $t \in [-T, +\infty)$:

$$\begin{aligned}
\bar{\mathbb{E}}\left[\mathrm{e}^{-\lambda Z^{(T)}_t}\right] &= \exp\left\{-\int_{-T}^t \gamma b\, u(\lambda, t-s)^{b-1} ds\right\} \\
&= \exp\left\{-\int_{-T}^t \frac{\gamma b\lambda^{b-1}}{1 + \gamma(b-1)\lambda^{b-1}(t-s)} ds\right\} \\
&= \Big(1 + \gamma(b-1)\lambda^{b-1}(t+T)\Big)^{-\frac{b}{b-1}}.
\end{aligned} \tag{83}$$

We write $\mathcal{A}^{(T)}(0)$ for the genealogical tree $\mathcal{A}_{\bar{\mathcal{T}}^{(T)}}(0)$ of the extant population. We define the ancestral process $M^{(T)} = (M^{(T)}_t, t \in (-T, 0))$, where $1 + M^{(T)}_t$ is the number of ancestors of the extant population living at time t by:

$$M^{(T)}_t = \mathrm{Card}\, \{u \in \mathcal{A}^{(T)}(0),\ H(u) = t\} - 1.$$

Theorem 6.1. *Consider the critical stable branching mechanism with immigration* (81)*. Then the time-changed ancestral process* $(M^{(T)}_{-T\,\mathrm{e}^{-t}}, t \geq 0)$ *is distributed under* $\bar{\mathbb{P}}$ *as the GWI process* $(X_t, t \geq 0)$ *from Theorem 4.1.*

Remark 6.2. In Proposition 19 in [6], it is shown that a reduced tree of a critical stable tree, after a deterministic time change, is a continuous-time Galton-Watson tree with birth rate 1 and offspring distribution given by $g_B(r)$. We get here a similar result with an additional immigration mechanism.

Proof. We only give an outline of the proof as we follow the ideas of the proof of Theorem 4.1. In the critical case the function c is given by (82) and the function g_t of (30) is still given by formula (34). Similarly to (23), (24), we define the jumping times $\{\tau_n^{(T)}, n \geq 0\}$ and jumpings sizes $\{\xi_n^{(T)}, n \geq 0\}$ of the ancestral process $M^{(T)}$. We also define $\tau_n^{(T,B)}$ and $\tau_n^{(T,I)}$ as in (26) with obvious change. Note that in this setting, $\tau_0^{(T)}$ is the immigration time of the first family after $-T$ which survives up to time 0. So, we have for $t \in (0, T)$:

$$\bar{\mathbb{P}}\left(\tau_0^{(T)} < t\right) = \exp\left\{-\int_{-T}^{-t} ds \int_{(0,+\infty)} r\pi(dr)\mathbb{P}_r(H(\mathcal{T}) > s)\right\} = \left(\frac{t}{T}\right)^{b/(b-1)}.$$

Recall g_I defined in (35). We also have, see (46), that for $t \in (0, T)$ and $r \in (0, 1)$:

$$\bar{\mathbb{P}}\left[r^{\xi_0^{(T)}} \mid \tau_0^{(T)} = t\right] = g_I(r).$$

Following the proof of Lemma 4.9, we get for $T > t > u > 0$:

$$\bar{\mathbb{P}}\left(\tau_1^{(T,I)} < u \mid \tau_0^{(T)} = t, M^{(T)}_{-\tau_0^{(T)}} = k\right) = \left(\frac{u}{t}\right)^{\frac{b}{b-1}},$$

$$\bar{\mathbb{P}}\left(\tau_1^{(T,B)} < u \mid \tau_0^{(T)} = t, M^{(T)}_{-\tau_0^{(T)}} = k\right) = \left(\frac{u}{t}\right)^{k}.$$

This further implies that

$$\bar{\mathbb{P}}\left(\tau_1^{(T)} < u \mid \tau_0^{(T)} = t, M^{(T)}_{-\tau_0^{(T)}} = k\right) = \left(\frac{u}{t}\right)^{k+\frac{b}{b-1}}.$$

We deduce that for $u > t > 0$:

$$\bar{\mathbb{P}}\left(\tau_1^{(T)} < T\,\mathrm{e}^{-u} \mid \tau_0^{(T)} = T\,\mathrm{e}^{-t}, M^{(T)}_{-\tau_0^{(T)}} = n\right) = \mathrm{e}^{-\left(n+\frac{b}{b-1}\right)(u-t)}.$$

Arguing as in the proof of Lemma 4.7, we obtain that given $\tau_0^{(T)}$ and $\xi_0^{(T)} = n$, $\tau_1^{(T)}$ and $\xi_1^{(T)}$ are independent and the conditional generating function of $\xi_1^{(T)}$ is given by $g_{[n]}$ defined in (38). The end of the proof is then similar. □

The following proposition, whose proof is left to the reader, is parallel to Corollary 6.5 in [14].

Proposition 6.3. *Consider the critical stable branching mechanism with immigration* (81)*. Then, we have:*

$$\lim_{t\to 0+} \frac{M^{(T)}_{-t}}{c(t)} \overset{a.s.}{=} Z_0^{(T)}$$

We order the set $\{i \in I^{(T)},\ h_i < 0 \text{ and } \ell^{-h_i}(\mathcal{F}_i) \neq 0\}$ of the immigrants that have descendants at time 0, by the date of arrival of the immigrant: $I_0^{(T)} = \{i_k, k \geq 0\}$ with $-\tau_0^{(T)} = h_{i_0} < h_{i_1} < h_{i_2} < \cdots < 0$. For every $k \geq 0$, we set $\zeta_k^{(T)}$ the size of the population at time 0 generated by the

k-th immigrant, that is $\zeta_k^{(T)} = \langle \ell^{-k_{i_k}}(\mathcal{F}_{i_k}), 1\rangle$. Notice that $\sum_{k=0}^{+\infty} \zeta_k^{(T)} = Z_0^{(T)}$. With Theorem 6.1 and Proposition 6.3 in hand, the next two results follow by the same arguments as Proposition 4.10 and Corollary 4.12, respectively.

Proposition 6.4. *Consider the critical stable branching mechanism with immigration* (81)*. The random point measure* $\sum_{k\in\mathbb{N}} \delta_{c(T)\zeta_k^{(T)}}(dx)$ *is a Poisson point measure on* $[0,\infty)$ *with intensity* $g(x)\,dx$*, with* g *defined by (50).*

Proof. Recall W from Corollary 4.3. According to (83) and (42), we deduce that $c(T)Z_0^{(T)}$ and W have the same distribution. Recall X^i and W_i from the proof of Proposition 4.10. Arguing as in the proof of Proposition 4.10, see the proof of (54), and using Theorem 6.1 and Proposition 6.3, we get that:

$$\left(c(T)\zeta_i^{(T)}, i\in\mathbb{N}\right) \stackrel{d}{=} \left(\mathrm{e}^{-\frac{T_i}{(b-1)}}\, W_i : i \in \mathbb{N}\right).$$

Then, follow the proof of Proposition 4.10 to conclude. □

Using Proposition 6.4, we obtain directly the following results, which is the analogue to Corollary 4.12.

Corollary 6.5. *Consider the critical quadratic branching mechanism with immigration* (81) *with* $b = 2$*. Let* $(\zeta_{(k)}^{(T)}, k \in \mathbb{N})$ *be the decreasing order statistics of* $(\zeta_k^{(T)}, k\in\mathbb{N})$*. Then, the random sequence* $\left(\zeta_{(k)}^{(T)}/Z_0^{(T)}, k\in\mathbb{N}\right)$ *has a Poisson-Dirichlet distribution with parameter* 2*.*

One can also consider the critical CBI associated as the limit of the sub-critical CBI when α in (1) goes down to 0. For the stable case, consider the birth rates $q^{\mathrm{b}}_{n,m}(-t)$ defined in Remark 5.5 for $\psi(\lambda) = \alpha\lambda + \gamma\lambda^{\mathrm{b}}$ with $b \in (1,2]$, $\gamma > 0$ and $\alpha > 0$. Letting α goes down to 0, wet get $\lim_{\alpha\to 0} q^{\mathrm{b}}_{n,m}(-t) = q^{\mathrm{b},0}_{n,m}(-t)$ with:

$$q^{\mathrm{b},0}_{n,m}(-t) = \begin{cases} \frac{(m+1)|b(b-2)\cdots(b-m+n)|}{(m-n+1)!}\frac{1}{t} & \text{for } b\in(1,2) \text{ and } m>n,\\ \frac{n+2}{t} & \text{for } b=2 \text{ and } m=n+1,\\ 0 & \text{for } b=2 \text{ and } m>n+1.\end{cases} \tag{84}$$

Then, Theorem 6.1 and (36) implies the following proposition that shows that both constructions for the critical case coincide.

Proposition 6.6. *Assume* $\psi(\lambda) = \gamma\lambda^b$ *with* $b\in(1,2]$*. Then, the ancestral process* $(M_t^{(T)}, -T \le t < 0)$ *is a Markov chain with birth rate* $q^{\mathrm{b},0}_{n,m}(-t)$ *given by (84) for* $m > n \ge 0$ *and* $t > 0$*.*

Acknowledgments

The authors would like to thank the Referees and the Associate Editor for their useful comments which helped to improve the presentation and the results.

References

[1] R. Abraham and J.-F. Delmas. A continuum-tree-valued Markov process. *Ann. Probab.*, 40(3):1167–1211, 2012.

[2] R. Abraham and J.-F. Delmas. Exact simulation of the genealogical tree for a stationary branching population and application to the asymptotics of its total length. arXiv:1612.03715, 2018.

[3] D. Aldous and L. Popovic. A critical branching process model for biodiversity. *Adv. Appl. Probab.*, 37(4):1094–1115, 2005.

[4] G. Alsmeyer and U. Rösler. On the existence of φ-moments of the limit of a normalized supercritical Galton-Watson process. *J. Theor. Probab.*, 17(4):905–928, 2004.

[5] K. B. Athreya and P. Ney. *Branching processes*. Springer-Verlag Berlin, New York, 1972.

[6] J. Berestycki, N. Berestycki, and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5):1835–1887, 2007.

[7] J. Berestycki, N. Berestycki, and J. Schweinsberg. Small-time behavior of Beta coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):214–238, 2008.

[8] N. Berestycki. Recent progress in coalescent theory. *Ensaios Matematicos*, 16(1):1–193, 2009.

[9] J. Bertoin and J.-F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probability Theory and Related Fields*, 117(2):249–266, Jun 2000.

[10] H. Bi and J.-F. Delmas. Total length of the genealogical tree for quadratic stationary continuous-state branching processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1321–1350, 2016.

[11] N. H. Bingham. Continuous branching processes and spectral positivity. *Stochastic Processes Appl.*, 4(3):217–242, 1976.

[12] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Communications in Mathematical Physics*, 197(2):247–276, Oct 1998.

[13] L. Breuer. *From Markov jump processes to spatial queues*. Springer Netherlands, 1 edition, 2003.

[14] Y.-T. Chen and J.-F. Delmas. Smaller population size at the MRCA time for stationary branching processes. *Ann. Probab.*, 40(5):2034–2068, 2012.

[15] A. Dress, V. Moulton, and W. Terhalle. T-theory: an overview. *European J. Combin.*, 17(2-3):161–175, 1996.

[16] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[17] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.

[18] S. N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[19] S. Feng. *The Poisson-Dirichlet distribution and related topics*. Probability and its Applications (New York). Springer, Heidelberg, 2010.

[20] C. Foucart, C. Ma, and B. Mallein. Coalescences in continuous-state branching processes. *Electron. J. Probab.*, 24:52 pp., 2019.

[21] C. Goldschmidt and J. B. Martin. Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab*, 10(21):718–745, 2005.

[22] D. R. Grey. Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probability*, 11:669–677, 1974.

[23] R. C. Griffiths and S. Tavare. Sampling theory for neutral alleles in a varying environment. *Phil. Trans. R. Soc. Lond. B*, 344, 1994.

[24] S. C. Harris, S. G. Johnston, and M. I. Roberts. The coalescent structure of continuous-time Galton-Watson trees. *To appear in Ann. Appl. Probab.*, 2019.

[25] S. Ho and B. Shapiro. Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, 11:423–434, 2011.

[26] S. D. Jacka and G. O. Roberts. On strong forms of weak convergence. *Stochastic Process. Appl.*, 67(1):41–53, 1997.

[27] L. F. James. Lamperti-type laws. *Ann. Appl. Probab.*, 20(4):1303–1340, 2010.

[28] S. G. G. Johnston. The genealogy of Galton-Watson trees. *Electronic Journal of Probability*, 24(4):1–35, 2019.

[29] S. G. G. Johnston and A. Lambert. The coalescent structure of branching processes: a unifying Poissonization approach. arXiv:1912.00198, 2019.

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

[31] J. F. Kingman. Random discrete distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(1):1–15, 1975.

[32] A. Lambert. Coalescence times for the branching process. *Adv. in Appl. Probab.*, 35(4):1071–1089, 2003.

[33] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, 12:no. 14, 420–446, 2007.

[34] A. Lambert. The contour of splitting trees is a Lévy process. *Ann. Probab.*, 38(1):348–395, 2010.

[35] Z. Li. *Measure-valued branching Markov processes*. Probability and its Applications (New York). Springer, Heidelberg, 2011.
[36] Z. Li. Continuous-state branching processes with immigration. In *From Probability to Finance*, pages 1–69. Springer, 2020.
[37] S. Nee, R. M. May, and P. H. Harvey. The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B*, 344, 1994.
[38] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields*, 92(1):21–39, 1992.
[39] J. Pitman. Coalescents with multiple collisions. *The Annals of Probability*, 27(4):1870–1902, Oct 1999.
[40] J. Pitman. Poisson-Kingman partitions. In *Statistics and science: a Festschrift for Terry Speed*, volume 40 of *IMS Lecture Notes Monogr. Ser.*, pages 1–34. Inst. Math. Statist., Beachwood, OH, 2003.
[41] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25(2):855–900, 1997.
[42] S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4):1116–1125, 1999.
[43] J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Processes and their Applications*, 106(1):107–139, 2003.
[44] H.-M. Wang, L. Li, and H. Yao. Coalescence for supercritical Galton-Watson processes with immigration. arXiv:1906.03945, 2019.

Romain Abraham, Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS, France
Email address: romain.abraham@univ-orleans.fr

Jean-François Delmas, CERMICS, Ecole des Ponts, France
Email address: delmas@cermics.enpc.fr

Hui He, School of Mathematical Sciences, Beijing Normal University, P.R. China
Email address: hehui@bnu.edu.cn