The Annals of Probability
2012, Vol. 40, No. 5, 2034–2068
DOI: 10.1214/11-AOP668

SMALLER POPULATION SIZE AT THE MRCA TIME FOR STATIONARY BRANCHING PROCESSES[1]

BY YU-TING CHEN AND JEAN-FRANÇOIS DELMAS

University of British Columbia and Université Paris-Est

We consider an elementary model of random size varying population governed by a stationary continuous-state branching process. We compute the distributions of various variables related to the most recent common ancestor (MRCA): the time to the MRCA, the size of the current population and the size of the population just before the MRCA. In particular we observe a natural mild bottleneck effect as the size of the population just before the MRCA is stochastically smaller than the size of the current population. We also compute the number of individuals involved in the last coalescent event of the genealogical tree, that is, the number of individuals at the time of the MRCA who have descendants in the current population. By studying more precisely the genealogical structure of the population, we get asymptotics for the number of ancestors just before the current time. We give explicit computations in the case of the quadratic branching mechanism. In this case, the size of the population at the MRCA is, in mean, $2/3$ of the size of the current population. We also provide in this case the fluctuations for the renormalized number of ancestors.

1. Introduction. A large literature is devoted to constant size population models. It goes back to Wright [49] and Fisher [23] in discrete time, and Moran [41] in continuous time. Models for constant infinite population in continuous time with spatial motion were introduced by Fleming and Viot [24]. On the other hand, the study of the genealogical tree of constant size population was initiated by Kingman [31], and described in a more general setting by Pitman [45] and Sagitov [48]. The complete description of the genealogy of the Fleming–Viot process can be partially done using the historical super-process by Dawson and Perkins [13] and precisely by using the look-down process developed by Donnelly and Kurtz [14, 15] or the stochastic flows from Bertoin and Le Gall [9–11].

However, it is natural to consider random size varying population models. Branching population models, for which sizes of the population are random, go

Received September 2010; revised February 2011.

[1]Supported in part by National Center for Theoretical Science Mathematics Division and NSC 97-2628-M-009-014, Taiwan and by the French "Agence Nationale de la Recherche," ANR-08-BLAN-0190.

MSC2010 subject classifications. Primary 60J80, 60J85, 92D25; secondary 60G10, 60G55, 60J60.

Key words and phrases. Branching process, most recent common ancestor, bottleneck, genealogy, random size population, Feller diffusion, last coalescent event, Lévy tree.

back to Galton and Watson [25] in discrete time. Jirina [28] considered continuous-state branching process (CB) models corresponding to individuals with infinitesimal mass. The genealogy of those processes can be partially described through the historical superprocess. However the continuum Lévy tree introduced by Le Gall and Le Jan [36] and developed later by Duquesne and Le Gall [16] allows to give a complete description of the genealogy in the critical and sub-critical cases. See the approach of Abraham and Delmas [1] or Berestycki, Kyprianou and Murillo [7] for a description of the genealogy in the super-critical cases.

The two families of models—models of constant size population and models of branching population—are, in certain cases, related. The case of a quadratic branching corresponds to the fact that only two genealogical lines of the population genealogical tree can merge together. In this particular case, it is possible to establish links between the constant size population model and CB models. Thus, conditionally on having a constant population size, the Dawson–Watanabe super-process is a Fleming–Viot process; see [18]. On the other hand, using a time change (with speed proportional to the inverse of the population size), it is possible to recover a Fleming–Viot process from a Dawson–Watanabe super-process; see [43]. Birkner et al. [12] have given similar results for stable branching mechanisms. In the same spirit, Kaj and Krone [29] studied the genealogical structure of models of random size varying population models and recovered the Kingman coalescent with a random time change.

Recently, some authors studied the coalescent process (or genealogical tree) of random size varying population; in this direction see [40] and [32] for branching process, [27] for stationary random size varying population and [22] for the dynamics of the time to the most recent common ancestor in branching processes.

Our primary interest is to present an elementary model of random size varying population and exhibit some interesting properties which could not be observed in the constant size model. The most striking example is the natural mild bottleneck effect: in a stationary regime, the size of the population just before the most recent common ancestor (MRCA) is stochastically smaller than the current population size. Our second goal is to give some properties of the coalescent tree such as: time to the most recent common ancestor (TMRCA), asymptotic behavior of the number of recent ancestors, number of individuals involved in the last coalescent event (i.e., the number of individuals at the time of the MRCA who have descendants in the current population).

One of the major drawbacks of the branching population models is that either the population becomes extinct or decreases to 0, which happens with probability 1 in the (sub)critical cases, or blows up exponentially fast with positive probability in the super-critical case. In particular there is no stationary regime, and the study of the genealogy of a current population depends on the arbitrary original size and time of the initial population. To circumvent this problem, we consider a sub-critical CB, $Y = (Y_t, t \geq 0)$, with branching mechanism ψ given by (1). We get the Q-process by conditioning Y to nonextinction (which is an event of zero probability); see [38] and [33]. The Q-process can also be seen as a CB with immigration; see [47]. We take the opportunity to present a probabilistic construction of independent interest for the Q-process in Corollary 3.5 which relies on a Williams decomposition of CB described in [2]. A first study of the genealogical tree of the Q-process can be found in [32].

We consider the Q-process under its stationary distribution and defined on the real line $Z = (Z_t, t \in \mathbb{R})$. Its Laplace transform [see (3.6)] is given by

$$\mathbb{E}[\mathrm{e}^{-\lambda Z_t}] = \exp\left(-\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s))\right), \qquad \lambda \geq 0, t \in \mathbb{R},$$

where $\tilde{\psi}(\lambda) = \psi(\lambda) - \lambda\psi'(0)$. In order for Z_t to be finite, we shall assume condition (A2),

$$\int_0^1 \left(\frac{1}{v\psi'(0)} - \frac{1}{\psi(v)}\right) dv < +\infty.$$

In order for the TMRCA to be finite, we assume condition (A1),

$$\int_1^\infty \frac{dv}{\psi(v)} < +\infty.$$

Notice a very similar condition exists to characterize coalescent processes which descend from infinity; see [6].

As in the look-down representation for constant size population, we shall represent the process Z using the picture of an immortal individual which gives birth to independent sub-populations or families; see Figure 1. For fixed time $t_0 = 0$ (which we can indeed choose to be equal to 0 by stationarity), we consider A the TMRCA of the population living at time 0, $Z^A = Z_{(-A)-}$ the size of the population just before the MRCA, Z^I the size of the population at time 0 which has been generated by the immortal individual over the time interval $(-A, 0)$ and $Z^O = Z_0 - Z^I$ the size of the population at time 0 which has been generated by the immortal individual at time $-A$. In Theorem 4.1, we give the joint distribution of (A, Z^A, Z^I, Z^O). One interesting phenomenon is Corollary 4.3.

COROLLARY. *Conditionally on A, Z^A, Z^I and Z^O are independent.*

In particular, conditionally on A, Z^A and Z are independent. Conditionally on A, Z^A depends on the past before $-A$ of the process Z and has to die at time 0, Z^O corresponds to the size of the population at time 0 generated at time $-A$, and Z^I corresponds to the size of the population at time 0 generated by the immortal individual over the time interval $(-A, 0)$. Then, as the immortal individual gives birth to independent populations, the corollary is then intuitively clear.

One of the most striking results, the natural mild bottleneck effect, is stated in Proposition 4.5.

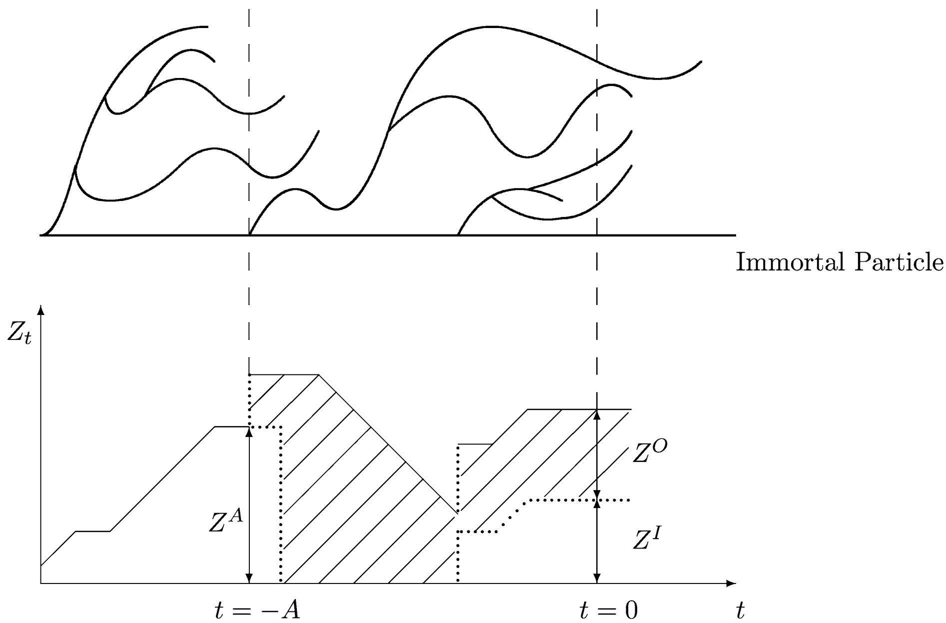

FIG. 1. *The bold lines in the first figure depict the space–time evolution of particles. The area of the shaded region in the second figure is contributed by the oldest clan alive at time $t = 0$.*

PROPOSITION. *Z^A is stochastically smaller than Z_0.*

Thus just before the MRCA, the population size is unusually small. Notice this result is not true in general if one considers the size of the population at the MRCA instead of just before; see Remark 4.6. We get nice quantitative results for the quadratic branching mechanism case; see Corollary 7.2.

COROLLARY. *Assume ψ is quadratic* [*and given by* (45)]. *We have a.s.*

$$\mathbb{P}(Z^A < Z_0|A) = \frac{11}{16} \quad \textit{and} \quad \mathbb{E}[Z^A|A] = \frac{2}{3}\mathbb{E}[Z_0|A]$$

and, *in particular*,

$$\mathbb{P}(Z^A < Z_0) = \frac{11}{16} \quad \textit{and} \quad \mathbb{E}[Z^A] = \frac{2}{3}\mathbb{E}[Z_0].$$

Notice that Z^A is stochastically smaller than Z_0; it is not a.s. smaller.

We also give in Theorem 4.7 the joint distribution of Z_0 and the TMRCA of the immortal individual and n individuals picked at random in the population at time 0. See also related results in [32].

We investigate in Proposition 5.2 the joint distribution of A, Z_0 and N^A, where $N^A + 1$ represents the number of individuals involved in the last coalescent event of the genealogical tree. Under a first moment condition on Z, we get that if the

TMRCA is large, then the last coalescent event is likely to involve only two individuals. In the stable case, this first moment condition is not satisfied, and the last coalescent event does not depend on the TMRCA; see Remark 5.6. This suggests a result similar to the one obtained in [12]: in the stable case, the topology of the genealogical tree (which does not take into account the length of the branches) may not depend on its depth given by the TMRCA.

After giving a more precise description of the genealogy of Z using continuum Lévy trees, we compute in Lemma 6.4 the joint law of Z_0 and the number of ancestors at time $-s$, M_s, of the population at time 0. Following [17], we get that a.s.

$$\lim_{s\downarrow 0} \frac{M_s}{c(s)} = Z_0,$$

where $c(s)$ is related to the extinction probability of Y and defined by $\int_{c(t)}^{\infty} \frac{dv}{\psi(v)} = t$. We can make precise the fluctuations in the asymptotic stable case [$\psi(\lambda) \sim a\lambda^{\alpha_0}$ at infinity, with $\alpha_0 \in (1, 2)$] and the quadratic case (corresponding to $\alpha_0 = 2$), as well as the fluctuation of Z_{-s} near Z_0, see Theorems 6.7 and 7.8. Notice that in the asymptotic stable case $\frac{M_s}{c(s)} - Z_0$ and $Z_{-s} - Z_0$, properly scaled, converge to the same limit, whereas this is not the case in the quadratic branching mechanism.

THEOREM. *Assume ψ is quadratic* [*and given by* (45)]. *The following convergences hold in distribution*:

$$\sqrt{c(s)\mathbb{E}[Z]}\left(\frac{M_s}{c(s)} - Z_0\right) \xrightarrow[s\downarrow 0+]{(d)} (Z_0 - Z_0')$$

and

$$\sqrt{c(s)\mathbb{E}[Z]}(Z_{-s} - Z_0) \xrightarrow[s\downarrow 0+]{(d)} \sqrt{2}(Z_0 - Z_0'),$$

where Z_0' is distributed as Z_0 and independent of Z_0.

See Theorems 7.8 for the joint distribution convergence.

The paper is organized as follows. We first recall well-known facts on CB in Section 2. We introduce in Section 3 the corresponding stationary CB, which is related to the Q-process of the CB, and give its first properties. We give the joint distribution of (A, Z^A, Z^I, Z^O) in Section 4 and prove the natural bottleneck effect, that is, Z^A is stochasitcally smaller than Z_0. We compute the number of old families (or number of individuals involved in the last coalescent event) in Section 5 and the asymptotics of the number of ancestors in Section 6. A first consequent part of the latter section is devoted to the introduction of the genealogy of CB processes using continuum random Lévy trees. We give more detailed results in the quadratic branching setting of Section 7.

2. Continuous-state branching process (CB). We recall some well-known facts on continuous-state branching process (CB); see, for example, [37] and references therein. We consider a *sub-critical* branching mechanism ψ: for $\lambda \geq 0$,

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,+\infty)} \pi(d\ell)[\mathrm{e}^{-\lambda\ell} - 1 + \lambda\ell], \tag{1}$$

where $\alpha = \psi'(0) > 0$, $\beta \geq 0$ and π is a Radon measure on $(0, +\infty)$ such that $\int_{(0,+\infty)}(\ell \wedge \ell^2)\pi(d\ell) < +\infty$. We consider the nontrivial case, that is either $\beta > 0$ or $\pi((0, 1)) = +\infty$. Notice that ψ is convex, of class $\mathcal{C}^1$ on $[0, +\infty)$ and of class $\mathcal{C}^\infty$ on $(0, +\infty)$ and $\psi''(0+) \in (0, +\infty]$.

Let P_x be the law of a CB $Y = (Y_t, t \geq 0)$ started at mass $x \geq 0$ and with branching mechanism ψ, and let E_x be the corresponding expectation. The process Y is a càdlàg $\mathbb{R}_+$-valued Feller process, and 0 is a cemetery point. The process Y has no fixed discontinuities. For every $\lambda > 0$ and for every $t \geq 0$, we have

$$\mathrm{E}_x[\mathrm{e}^{-\lambda Y_t}] = \mathrm{e}^{-xu(\lambda,t)}, \tag{2}$$

where the function u is the unique nonnegative solution of

$$u(\lambda, t) + \int_0^t \psi(u(\lambda, s))\, ds = \lambda, \qquad \lambda \geq 0, t \geq 0. \tag{3}$$

Note that the function u is equivalently characterized as the unique nonnegative solution of

$$\int_{u(\lambda,t)}^{\lambda} \frac{dr}{\psi(r)} = t, \qquad \lambda \geq 0, t \geq 0, \tag{4}$$

or as the unique nonnegative solution of, for $\lambda \geq 0$,

$$\begin{cases} \partial_t u + \psi(u) = 0, & t > 0, \\ u(\lambda, 0), = \lambda. \end{cases} \tag{5}$$

The Markov property of Y implies that for all $\lambda, s, t \geq 0$,

$$u(u(\lambda, t), s) = u(\lambda, t + s). \tag{6}$$

Let $\mathbb{N}$ be the canonical measure (we shall also call it excursion measure) associated to Y. It is a σ-finite measure which intuitively describes the distribution of Y started at an infinitesimal mass. We recall that if

$$\sum_{i\in I} \delta_{x_i, Y^i}(dx, dY)$$

is a Poisson point measure with intensity $\mathbf{1}_{[0,+\infty)}(x)\, dx\, \mathbb{N}[dY]$, then

$$\sum_{i\in I} \mathbf{1}_{\{x_i \leq x\}} Y^i \tag{7}$$

is distributed as Y under P_x. In particular, we have, for $\lambda \geq 0$,

$$\mathbb{N}[1 - \mathrm{e}^{-\lambda Y_t}] = \lim_{x \downarrow 0} \frac{1}{x} \mathrm{E}_x[1 - \mathrm{e}^{-\lambda Y_t}] = u(\lambda, t).$$

For convenience, we put $Y_t = 0$ for $t < 0$.

Let $\zeta = \inf\{t; Y_t = 0\}$ be the extinction time of Y. We consider the function

$$c(t) = \mathbb{N}[\zeta > t] = \mathbb{N}[Y_t > 0] = \lim_{\lambda \to \infty} \uparrow u(\lambda, t). \tag{8}$$

We shall assume throughout this paper, but for Sections 3.1 and 3.3, that the following strong extinction property holds:

$$\int_1^\infty \frac{dv}{\psi(v)} < +\infty. \tag{A1}$$

It follows from (4) and (8) that c is the unique nonnegative solution of

$$\int_{c(t)}^\infty \frac{dv}{\psi(v)} = t, \qquad t > 0. \tag{9}$$

Thanks to (A1), we get that $c(t)$ is finite for all $t > 0$ and $\mathbb{N}[\zeta = +\infty] = 0$. We also get that c is continuous decreasing and thus one-to-one from $(0, +\infty)$ to $(0, +\infty)$. Letting λ go to infinity in (6) yields that for $s, t \geq 0$

$$u(c(t), s) = c(t + s). \tag{10}$$

3. Stationary CB. In contrast to the Wright–Fisher population model, CB models do not exhibit stationary distributions. However, by conditioning subcritical CB to nonextinction (see [21, 33] and [47] for details), one gets the so-called Q-process, which we denote by Y''. This process is a CB process with immigration in the sense of [30] and may have a stationary distribution. As pointed out in [3] (see also [19] or [20]), this process has a heuristic interpretation by introducing a fixed infinite ancestral lineage. Namely, it is an independent sum of the process Y and the size of families thrown off by an "immortal individual" where the law of each family coincides with that of a generic family of Y.

We introduce the process Y'' in Section 3.1 as well as its stationary version Z. Then we check in Section 3.2, that under (A1) the process Y'' is indeed the Q-process associated to Y. This gives then a natural interpretation of Z. We give preliminary results on the process Z in Sections 3.3 and 3.4.

3.1. *Poisson point measure of CB.* We consider the following Poisson point measures:

- Let $\mathcal{N}_0(dr, dt) = \sum_{i \in I} \delta_{(r_i, t_i)}(dr, dt)$ be a Poisson point measure on $(0, +\infty) \times \mathbb{R}$ with intensity

$$r\pi(dr)\, dt.$$

- Conditionally on $\mathcal{N}_0$, let $(\mathcal{N}_{1,i}, i \in I)$, where $\mathcal{N}_{1,i}(dt, dY) = \sum_{j \in J_{1,i}} \delta_{t_j, Y^j}(dt, dY)$, be independent Poisson point measures with respective intensity

$$r_i \delta_{t_i}(dt)\mathbb{N}[dY].$$

Notice that for all $j \in J_{1,i}$, we have $t_j = t_i$. We set $J_1 = \bigcup_{i \in I} J_{1,i}$ and $\mathcal{N}_1(dt, dY) = \sum_{j \in J_1} \delta_{t_j, Y^j}(dt, dY)$.

- Let $\mathcal{N}_2(dt, dY) = \sum_{j \in J_2} \delta_{t_j, Y^j}(dt, dY)$ be a Poisson point measure independent of $(\mathcal{N}_0, \mathcal{N}_1)$ and with intensity

$$2\beta\, dt\, \mathbb{N}[dY].$$

We set $\mathcal{J} = J_1 \cup J_2$. We shall call Y^j, with $j \in \mathcal{J}$ a *family* and t_j its birth time.

We will consider the two following processes $Y'' = (Y''_t, t \geq 0)$ and their stationary version $Z = (Z_t, t \in \mathbb{R})$:

$$Y''_t = \sum_{j \in \mathcal{J}, t_j > 0} Y^j_{t - t_j}, \tag{11}$$

$$Z_t = \sum_{j \in \mathcal{J}} Y^j_{t - t_j}. \tag{12}$$

We will denote by $\mathbb{P}$ the probability measure under which Y'' and Z are defined and $\mathbb{E}$ the corresponding expectation.

At this stage, let us emphasize there is another natural decomposition of Y'' and Z. For $i \in I$, set $Y^i = \sum_{j \in J_{1,i}} Y^j$ and $\mathcal{I} = I \cup J_2$. The random measure

$$\mathcal{N}_3(dt, dY) = \sum_{i \in \mathcal{I}} \delta_{t_i, Y^i}(dt, dY) \tag{13}$$

is a Poisson point measure with intensity $dt\, \mu(dY)$ and

$$\mu(dY) = 2\beta \mathbb{N}[dY] + \int_{(0,+\infty)} \ell \pi(d\ell) \mathrm{P}_\ell(dY). \tag{14}$$

We have

$$Y''_t = \sum_{i \in \mathcal{I}; t_i > 0} Y^i_{t - t_i}, \tag{15}$$

$$Z_t = \sum_{i \in \mathcal{I}} Y^i_{t - t_i}. \tag{16}$$

We shall call Y^i, with $i \in \mathcal{I}$, a *clan* and t_i its birth time. For $j \in J_2$, Y^j is a clan and a family. Notice that a.s. two clans have different birth time, but families in the same clan have the same birth time.

The presentation with clans is simpler than the representation with families, and most of the results can be obtained by using the following representation by Poisson random measures. We will use the family representation in Sections 5 and 6.

We define $\tilde{\psi}$ by

$$\tilde{\psi}(\lambda) = \psi(\lambda) - \lambda\psi'(0) = \psi(\lambda) - \alpha\lambda. \tag{17}$$

The next lemma is the exponential formula for Poisson point measure; see Section XII.1 of [46].

LEMMA 3.1. *Let F be a nonnegative measurable function. We have*

$$\mathbb{E}\big[\mathrm{e}^{-\sum_{j\in\mathcal{J}} F(t_j,Y^j)}\big] = \exp\biggl(-\int_{\mathbb{R}} dt\, \tilde{\psi}'\bigl(\mathbb{N}\big[1-\mathrm{e}^{-F(t,Y)}\big]\bigr)\biggr). \tag{18}$$

PROOF. Using basic properties of Poisson point measures, we get

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{e}^{-\sum_{j\in\mathcal{J}} F(t_j,Y^j)}\big]\\
&\quad = \mathbb{E}\big[\mathrm{e}^{-\sum_{j\in J_1} F(t_j,Y^j)}\big]\mathbb{E}\big[\mathrm{e}^{-\sum_{j\in J_2} F(t_j,Y^j)}\big]\\
&\quad = \mathbb{E}\big[\mathrm{e}^{-\sum_{i\in I} r_i \mathbb{N}[1-\mathrm{e}^{-F(t_i,Y)}]}\big]\mathrm{e}^{-2\beta\int dt \mathbb{N}[1-\mathrm{e}^{-F(t,Y)}]}\\
&\quad = \mathrm{e}^{-\int dt \int_{(0,+\infty)} \ell\pi(d\ell)(1-\exp(-\ell\mathbb{N}[1-\mathrm{e}^{-F(t,Y)}]))}\mathrm{e}^{-2\beta\int dt\, \mathbb{N}[1-\mathrm{e}^{-F(t,Y)}]}\\
&\quad = \mathrm{e}^{-\int dt\, \tilde{\psi}'(\mathbb{N}[1-\mathrm{e}^{-F(t,Y)}])}.
\end{aligned}$$

This gives the desired formula. □

PROPOSITION 3.2. *The process Y'' is a CB with branching mechanism ψ and immigration function $\tilde{\psi}'$*

$$\tilde{\psi}'(\lambda) = 2\beta\lambda + \int_{(0,+\infty)} \ell\pi(d\ell)(1-\mathrm{e}^{-\lambda\ell})$$

started at $Y''_0 = 0$.

PROOF. This is a direct consequence of Lemma 3.1 and results from [30]. □

In particular, Y'' is a strong Markov process started at 0, and its transition kernel is characterized by the following: for $\lambda \geq 0$, $t \geq 0$, $r \geq 0$

$$\mathbb{E}[\mathrm{e}^{-\lambda Y''_t} | Y''_0 = r] = \exp\biggl(-ru(\lambda,t) - \int_0^t \tilde{\psi}'(u(\lambda,s))\, ds\biggr).$$

The next result is then straightforward.

COROLLARY 3.3. *For each $t \in \mathbb{R}$, $(Z_s; s \geq t)$ has the same law as a CB with branching mechanism ψ and immigration function $\tilde{\psi}'$ started at the invariant distribution $\mathbb{P}(Z_t \in \cdot)$.*

3.2. *Q-process.* We check the process Y'' is indeed the Q-process for CB using Williams's decomposition.

Let $m>0$ and $\nu_m(dt)=\sum_{i\in I_m} r_i\delta_{t_i}(dt)$, where $\sum_{i\in I_m}\delta_{(r_i,t_i)}(dr,dt)$ is a Poisson point measure with intensity

$$\mathbf{1}_{[0,m]}(t)\mathrm{e}^{-rc(m-t)}r\pi(dr)\,dt.$$

Conditionally on ν_m, let $\mathcal{N}^{(m)}(dt,dY)=\sum_{j\in\mathcal{J}^m}\delta_{t_j,Y^j}(dt,dY)$ be a Poisson point measure with intensity

$$\big(\nu_m(dt)+2\beta\mathbf{1}_{[0,m]}(t)\,dt\big)\mathbb{N}[dY,\zeta<m-t].$$

The next proposition is a consequence of Theorem 3.3 in [2].

PROPOSITION 3.4. *Assume* (A1) *holds. Under* $\mathbb{N}$, *conditionally on* $\{\zeta=m\}$, Y *is distributed as* $(Y'_t,t\ge 0)$ *where*

$$Y'_t=\sum_{j\in\mathcal{J}^m}Y^j_{t-t_j}.$$

It is then easy to deduce the following corollary using representation (15) of Y''.

COROLLARY 3.5. *Assume* (A1) *holds. The limit distribution of* Y *under* $\mathbb{N}$, *conditionally on* $\{\zeta=m\}$, *as* m *goes to infinity, is the distribution of* Y'' *from Proposition* 3.2.

PROOF. The proof relies on the monotonic convergence of the intensities of Poisson point measures. Let $\nu(dt,dr,du)=\sum_{i\in I}r_i\delta_{r_i}(dr)\delta_{t_i}(dt)\delta_{u_i}(du)$, where $\sum_{i\in I}\delta_{(r_i,t_i,u_i)}(dr,dt,du)$ is a Poisson point measure with intensity

$$\mathbf{1}_{[0,+\infty)}(t)\mathbf{1}_{[0,1]}(u)r\pi(dr)\,dt\,du.$$

Conditionally on ν, let $\sum_{j\in\mathcal{J}}\delta_{(t_j,Y^j,r_j,u_j)}(dt,dY,dr,du)$ be a Poisson point measure with intensity

$$\big(\nu(dt,dr,du)+2\beta\mathbf{1}_{[0,+\infty)}(t)\,dt\,\delta_0(dr)\delta_0(du)\big)\mathbb{N}[dY].$$

We denote by ζ^j the extinction time of Y^j. For $m>0$, we set

$$\mathcal{M}^{(m)}(dt,dY)=\sum_{j\in\mathcal{J}}\mathbf{1}_{\{\zeta^j<m-t_j\}}\mathbf{1}_{\{t_j<m\}}\mathbf{1}_{\{u_j<\exp(-rc(m-t_j))\}}\delta_{t_j,Y^j}(dt,dY).$$

Notice that $\mathcal{M}^{(m)}$ is distributed as $\mathcal{N}^{(m)}$ and that $(\mathcal{M}^{(m)},m>0)$ is an increasing sequence with limit

$$\mathcal{M}^{(\infty)}(dt,dY)=\sum_{j\in\mathcal{J}}\delta_{t_j,Y^j}(dt,dY).$$

Notice that $\mathcal{M}^{(\infty)}(dt, dY)$ is distributed as $\mathbf{1}_{\{t\geq 0\}}(\mathcal{N}_1 + \mathcal{N}_2)(dt, dY)$. Let us consider the processes $Y^{(m)} = (Y_t^{(m)}, t \geq 0)$ and $Y^{(\infty)} = (Y_t^{(\infty)}, t \geq 0)$ defined by

$$Y_t^{(m)} = \sum_{j\in\mathcal{J}} \mathbf{1}_{\{\zeta^j < m-t_j\}} \mathbf{1}_{\{t_j<m\}} \mathbf{1}_{\{u_j<\exp(-rc(m-t_j))\}} Y^j_{t-t_j} \quad \text{and} \quad Y_t^{(\infty)} = \sum_{j\in\mathcal{J}} Y^j_{t-t_j}.$$

Then we deduce from Proposition 3.4 that $Y^{(m)}$ is defined as Y under $\mathbb{N}$ conditionally on $\{\zeta = m\}$. Furthermore the process $Y^{(\infty)}$ is defined as Y''. By construction, we get that a.s. the sequence $(Y^{(m)}, m \geq 0)$ increases to $Y^{(\infty)}$. This gives the result. $\square$

Corollary 3.5 readily implies that the Q-process associated to Y, that is, the limit distribution of Y under $\mathbb{N}$, conditionally on $\{\zeta \geq m\}$, as m goes to infinity, is the distribution of Y'' from Proposition 3.2.

3.3. *Stationary CB*. We first give an interpretation of Z in terms of the underlying populations. At time t, Z_t correspond to the size of a population generated by an immortal individual (with zero mass) which gives birth at rate 2β to clans (or families) which sizes evolve independently as Y under $\mathbb{N}$ and at rate $r\pi(dr)$ to clans with initial size r which evolve independently as Y under P_r.

By construction the process Z is stationary. The next lemma which gives the Laplace transform of Z is a direct consequence of the construction of Z.

LEMMA 3.6. *For all $t \in \mathbb{R}$ and $\lambda \geq 0$, the Laplace transform of Z_t is given by*

$$\mathbb{E}[\mathrm{e}^{-\lambda Z_t}] = \exp\biggl(-\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s))\biggr). \tag{19}$$

PROOF. Using Lemma 3.1, we have

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{-\lambda Z_t}] &= \exp\biggl(-\int_{\mathbb{R}} ds\, \tilde{\psi}'(\mathbb{N}[1 - \mathrm{e}^{-\lambda Y_{t-s}}])\biggr)\\ &= \exp\biggl(-\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s))\biggr).\end{aligned}$$ $\square$

We shall consider the following assumption:

$$\int_1^{+\infty} \ell \log(\ell) \pi(d\ell) < +\infty. \tag{A2}$$

The next lemma is well known [notice condition (A1) is not assumed].

LEMMA 3.7. *In the sub-critical case, the following conditions are equivalent*:

(i) (A2) *holds*;

(ii) $\int_0^1 (\frac{1}{\alpha v} - \frac{1}{\psi(v)})\, dv < +\infty$;
(iii) $\mathbb{E}_r[Y_t \log(Y_t)] < +\infty$ *for some* $t > 0$ *and* $r > 0$;
(iv) $\mathbb{E}_r[Y_t \log(Y_t)] < +\infty$ *for all* $t > 0$ *and* $r > 0$.

PROOF. For (i) $\Leftrightarrow$ (ii), see [26], proof of Theorem 4a, and for (ii) $\Leftrightarrow$ (iii) [or (iv)] use Lemma 1, page 25, of [5]. $\square$

The next proposition gives a condition for finiteness of Z; see also [44] in a more general framework.

PROPOSITION 3.8. *We have* $\mathbb{P}(Z_0 < +\infty) = 1$ *if and only if* (A2) *holds.*

PROOF. Thanks to (19), we get $\mathbb{P}(Z_0 < +\infty) = 1$ if and only if

$$\lim_{\lambda \to 0} \int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s)) = 0.$$

As $\lambda \mapsto u(\lambda, s)$ decreases to 0 as λ goes down to 0 for all $s \ge 0$, we deduce by dominated convergence that $\mathbb{P}(Z_0 < +\infty) = 1$ if and only if $\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s)) < +\infty$ for at least one $\lambda > 0$.

Notice that $\partial_t u + \psi(u) = 0$ implies $\psi'(u) = -\partial_t^2 u / \partial_t u$, and hence for every $0 \le t < T < +\infty$ we have

$$\int_t^T \tilde{\psi}'(u(\lambda, s))\, ds = \log\left(\frac{\psi(u(\lambda, t))\mathrm{e}^{\alpha t}}{\psi(u(\lambda, T))\mathrm{e}^{\alpha T}}\right). \tag{20}$$

We deduce that $T \mapsto \psi(u(\lambda, T))\mathrm{e}^{\alpha T}$ is decreasing. Thus, we get that

$$\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s)) < +\infty$$

if and only if $\lim_{T \to +\infty} \psi(u(\lambda, T))\mathrm{e}^{\alpha T} > 0$. Thanks to (4) we have $\lim_{T \to +\infty} u(\lambda, T) = 0$. Since $\lim_{\lambda \downarrow 0} \psi(\lambda)/\lambda = \alpha > 0$, we get that $\lim_{T \to +\infty} \psi(u(\lambda, T))\mathrm{e}^{\alpha T} > 0$ if and only if $\lim_{T \to +\infty} u(\lambda, T)\mathrm{e}^{\alpha T} > 0$.

We deduce from (4) that

$$u(\lambda, T)\mathrm{e}^{\alpha T} = \lambda \exp\left(\alpha \int_{u(\lambda, T)}^{\lambda} dr \left(\frac{1}{\psi(r)} - \frac{1}{\alpha r}\right)\right). \tag{21}$$

Thus we deduce from Lemma 3.7 that $\mathbb{P}(Z_0 < +\infty) = 1$ if and only if (A2) holds. $\square$

COROLLARY 3.9. *Assume* (A2) *holds. We have for* $\lambda > 0$, $t \in \mathbb{R}$,

$$\mathbb{E}[Z_t \mathrm{e}^{-\lambda Z_t}] = \frac{\tilde{\psi}'(\lambda)}{\psi(\lambda)} \mathbb{E}[\mathrm{e}^{-\lambda Z_t}]. \tag{22}$$

In particular, we have

$$\mathbb{E}[Z_t] = \frac{\psi''(0+)}{\psi'(0)} \in (0, +\infty]. \tag{23}$$

PROOF. We deduce from (19) that

$$\mathbb{E}[Z_t e^{-\lambda Z_t}] = \mathbb{E}[e^{-\lambda Z_t}]\,\partial_\lambda \int_0^\infty \tilde{\psi}'(u(\lambda, s))\,ds.$$

We deduce from (4) that $\lambda \mapsto u(\lambda, s)$ is increasing and of class $\mathcal{C}^\infty$ on $(0, +\infty)$ and that

$$\partial_\lambda u(\lambda, s) = \frac{\psi(u(\lambda, s))}{\psi(\lambda)} = \frac{-\partial_s u(\lambda, s)}{\psi(\lambda)}. \tag{24}$$

Thus, we get

$$\begin{aligned}\partial_\lambda \int_0^\infty \tilde{\psi}'(u(\lambda, s))\,ds &= \int_0^\infty \psi''(u(\lambda, s))\partial_\lambda u(\lambda, s)\,ds \\ &= -\frac{1}{\psi(\lambda)} \int_0^\infty \psi''(u(\lambda, s))\partial_s u(\lambda, s)\,ds \\ &= \frac{\tilde{\psi}'(\lambda)}{\psi(\lambda)}.\end{aligned}$$

The last part of the corollary is immediate. □

REMARK 3.10. Assumption (A1) is not needed to define the process Y'' or the stationary process Z. However, the study of MRCA for Z is not relevant if (A1) does not hold.

Notice, we will introduce a complete genealogical structure for Z in Section 6 by using a genealogical structure of the families $(Y^j, j \in \mathcal{J})$.

From now on, we shall assume that (A1) *and* (A2) *are in force.*

3.4. *Further property for stationary CB.* By construction, we deduce that for all $t \in \mathbb{R}$, the process $(Z_{s+t}, s \geq 0)$ is a CB with branching mechanism ψ and immigration function $\tilde{\psi}'$ started as the stationary distribution whose Laplace transform is given by (19). Then Proposition 1.1 in [30] implies that Z is a Hunt process, and, in particular, it is càdlàg and strongly Markov taking values in $[0, +\infty]$. By stationarity and since $+\infty$ is a cemetery point for Z, we deduce that a.s. for all $t \in \mathbb{R}$, Z_t is finite.

Next, we recall some asymptotic properties of the functions u and c given in Lemma 3.1 of [32].

LEMMA 3.11. *For every* $\lambda \in (0, \infty)$, *we have*

$$\lim_{t\to\infty} \frac{u(\lambda, t)}{c(t)} = e^{-\alpha c^{-1}(\lambda)}, \tag{25}$$

and there exists $\kappa_* \in (0, \infty)$ *such that*

$$\lim_{t\to\infty} c(t)e^{\alpha t} = \kappa_*. \tag{26}$$

We compute some integrals of $\tilde{\psi}'$.

PROPOSITION 3.12. *The followings hold for every* $0 \le t < \infty$:

$$\int_t^\infty \tilde{\psi}'(u(\lambda, s))\, ds = \log\left(\frac{\psi(u(\lambda, t))e^{\alpha t + \alpha c^{-1}(\lambda)}}{\kappa_* \alpha}\right), \qquad \lambda > 0, \tag{27}$$

$$\int_t^\infty \tilde{\psi}'(c(s))\, ds = \log\left(\frac{\psi(c(t))e^{\alpha t}}{\kappa_* \alpha}\right), \tag{28}$$

where the constant κ_* *is defined in Lemma* 3.11.

PROOF. We deduce from (20), (25) and (26) that

$$\lim_{T\to\infty} \psi(u(\lambda, T))e^{\alpha T} = \lim_{T\to\infty} \frac{\psi(u(\lambda, T))}{u(\lambda, T)} \frac{u(\lambda, T)}{c(T)} c(T)e^{\alpha T} = \alpha e^{-\alpha c^{-1}(\lambda)} \kappa_*,$$

and (27) follows by letting $T \to \infty$ for both sides of (20). Then, let λ go to infinity in (27) to get (28), and use the monotone convergence theorem. □

As a consequence of (27) with $t = 0$ and Lemma 3.6, we get the following corollary.

COROLLARY 3.13. *For all* $t \in \mathbb{R}$ *and* $\lambda \ge 0$, *the Laplace transform of* Z_t *is given by*

$$\mathbb{E}[e^{-\lambda Z_t}] = \exp\left(-\int_0^\infty ds\, \tilde{\psi}'(u(\lambda, s))\right) = \frac{e^{-\alpha c^{-1}(\lambda)} \kappa_* \alpha}{\psi(\lambda)}. \tag{29}$$

Eventually, we check that Z is nonzero. Recall our notations in Section 3.1. Let $\zeta_i = \inf\{t > 0; Y_t^i = 0\}$ be the duration of the family or clan Y^i and $t_i + \zeta_i$ its extinction time, with i in I, J_1 or J_2.

PROPOSITION 3.14. *We have*

$$\mathbb{P}\left(\sum_{i\in\mathcal{I}} \mathbf{1}_{(t_i, t_i+\zeta_i)}(t) > 0, \forall t \in \mathbb{R}\right) = 1.$$

In particular, we have $\mathbb{P}(\exists t \in \mathbb{R}; Z_t = 0) = 0$.

For $-\infty < a < b < +\infty$, we will consider in the forthcoming proof

$$N_{a,b} = \sum_{i\in\mathcal{I}} \mathbf{1}_{\{t_i < a; b < t_i + \zeta_i\}}, \tag{30}$$

the number of clans born before a and still alive at time b. Notice $N_{a,b}$ is a Poisson random variable with parameter

$$\begin{aligned}\Lambda(b-a) &:= \int dr\, \mu(dY)\mathbf{1}_{(-\infty,a)}(r)\mathbf{1}_{\{\zeta+r>b\}}\\ &= \int_{b-a}^{\infty} dr\, \tilde{\psi}'(c(r))\\ &= \log\left(\frac{\psi(c(b-a))\mathrm{e}^{\alpha(b-a)}}{\kappa_*\alpha}\right),\end{aligned} \tag{31}$$

where we have used (14) the definition of μ for the first equality and (28) for the last equality.

PROOF OF PROPOSITION 3.14. Observe that no clan surviving at time $t \in (a,b)$ implies that there are no clan surviving on any nondegenerate interval containing t. Hence, for any $n \geq 1$, we have

$$\left\{\exists t \in (a,b), \sum_{i\in\mathcal{I}} \mathbf{1}_{(t_i,t_i+\zeta_i)}(t) = 0\right\} \subset \bigcup_{j=1}^{n}\{N_{u_{j-1},u_j} = 0\} \cup \bigcup_{j=1}^{n+1}\{N_{v_{j-1},v_j} = 0\},$$

where $u_j = a + j(b-a)/n$ and $v_j = a + (2j-1)(b-a)/2n$. Notice that N_{u_{j-1},u_j} and N_{v_{j-1},v_j} are Poisson random variables with parameter $\theta_n = \Lambda((b-a)/n)$. We deduce that

$$\mathbb{P}\left(\exists t \in (a,b), \sum_{i\in\mathcal{I}} \mathbf{1}_{(t_i,t_i+\zeta_i)}(t) = 0\right) \leq (2n+1)\mathrm{e}^{-\theta_n}. \tag{32}$$

Therefore the first part of the proposition will be proved as soon as $\lim_{n\to+\infty} n \times \exp(-\theta_n) = 0$ which, thanks to formula (31), will be implied by $\lim_{t\to 0} t\psi(c(t)) = +\infty$ and thus by

$$\lim_{\lambda\to+\infty} \int_{\lambda}^{+\infty} \frac{dr}{\psi(r)}\, \psi(\lambda) = +\infty. \tag{33}$$

Hypothesis on β and π imply there exists a constant $c_0 > 0$ such that

$$\alpha\lambda \leq \psi(\lambda) \leq c_0\lambda^2 \quad \text{and} \quad \lim_{\lambda\to+\infty} \psi(\lambda)/\lambda = +\infty.$$

Therefore (33) is in force.

The second part of the proposition is clear by definition of ζ_i and representation (16). $\square$

4. TMRCA and populations sizes. We consider the coalescence of the genealogy at a fixed time t_0. Thanks to stationarity, we may assume that $t_0 = 0$, and we write Z instead of Z_0. There are infinitely many clans contributing to the population at time 0. The Poisson random variable introduced in (30), with $b = 0$, gives the number of clans born before a and still alive at time 0. Notice its parameter is finite; see (31). Therefore, there are only finitely many clans born before a and alive at time 0. In particular, this implies that there is one unique oldest clan alive at time 0. We denote by $-A$ the birth time of this unique oldest clan at time 0,

$$A = -\inf\{t_i \leq 0; Y^i_{-t_i} > 0, i \in \mathcal{I}\}.$$

We set Z^O the population size of this clan at time 0.

$$Z^O := Y^i_{-t_i} \qquad \text{if } A = -t_i.$$

The time A is also the time to the most recent common ancestor (TMRCA) of the population at time 0. The size of all the clans alive at time 0 with birth time in $(-A, 0)$ is given by

$$Z^I := Z - Z^O.$$

We are also interested in the size of the population just before the most recent common ancestor (MRCA).

$$Z^A := Z_{(-A)-} = \sum_{i \in \mathcal{I}} Y^i_{(-A-t_i)} \mathbf{1}_{\{t_i < -A\}}.$$

THEOREM 4.1. *The joint distribution of* (A, Z^A, Z^I, Z^O) *is characterized by the following*: *for* $\lambda, \gamma, \eta \geq 0$ *and* $t \geq 0$,

$$\begin{aligned} &\mathbb{E}[\mathrm{e}^{-\lambda Z^A - \gamma Z^I - \eta Z^O}; A \in dt] \\ &\quad = dt\big(\tilde{\psi}'(c(t)) - \tilde{\psi}'(u(\eta, t))\big) \\ &\qquad \times \exp\biggl(-\int_0^t ds\, \tilde{\psi}'(u(\gamma, s)) - \int_0^\infty ds\, \tilde{\psi}'\big(u\big(\lambda + c(t), s\big)\big)\biggr). \end{aligned} \tag{34}$$

PROOF. Given f, a nonnegative Borel measurable function defined on $\mathbb{R}$, we have

$$\begin{aligned} &\mathbb{E}[\mathrm{e}^{-\lambda Z^A - \gamma Z^I - \eta Z^O} f(A)] \\ &\quad = \mathbb{E}\biggl[\sum_{j \in \mathcal{I}} \exp\biggl(-\lambda \sum_{i \in \mathcal{I}, t_i < t_j} Y^i_{(t_j - t_i)} - \gamma \sum_{i \in \mathcal{I}, t_i > t_j} Y^i_{-t_i} - \eta Y^j_{-t_j}\biggr) \\ &\qquad\qquad \times f(-t_j) \mathbf{1}_{\{Y^j_{-t_j} > 0, \sum_{i \in \mathcal{I}, t_i < t_j} \mathbf{1}_{\{Y^i_{-t_i} > 0\}} = 0\}}\biggr] \\ &\quad = \int_0^\infty dt\, \mu(\mathrm{e}^{-\eta Y_t}; Y_t > 0) f(t) \mathbb{E}\biggl[\exp\biggl(-\gamma \sum_{i \in \mathcal{I}, t_i > -t} Y^i_{-t_i}\biggr)\biggr] \\ &\qquad \times \lim_{K \to \infty} \mathbb{E}\biggl[\exp\biggl(-\lambda \sum_{t_i < -t} \big(Y^i_{(-t-t_i)} + K \mathbf{1}_{\{Y^i_{-t_i} > 0\}}\big)\biggr)\biggr], \end{aligned}$$

where we used that Poisson point measures over disjoint sets are independent. We have

$$\begin{aligned}\mu(\mathrm{e}^{-\eta Y_t}; Y_t > 0) &= \mu\big(\mathbf{1}_{\{Y_t>0\}} - (1 - \mathrm{e}^{-\eta Y_t})\big) \\ &= \tilde{\psi}'(c(t)) - \tilde{\psi}'(u(\eta, t)).\end{aligned}$$

Using Lemma 3.1, we get

$$\mathbb{E}\Big[\exp\Big(-\gamma \sum_{i\in\mathcal{I}, t_i>-t} Y^i_{-t_i}\Big)\Big] = \exp\Big(-\int_0^t ds\, \tilde{\psi}'(u(\gamma, s))\Big).$$

We also have

$$\begin{aligned}&\lim_{K\to\infty} \mathbb{E}\Big[\exp\Big(-\lambda \sum_{i\in\mathcal{I}, t_i<-t} \big(Y^i_{(-t-t_i)} + K\mathbf{1}_{\{Y^i_{-t_i}>0\}}\big)\Big)\Big] \\ &\quad = \exp\Big(-\int ds\, \mathbf{1}_{\{s>0\}}\mu\big(1 - \mathrm{e}^{-\lambda Y_s}\mathbf{1}_{\{Y_{s+t}=0\}}\big)\Big) \\ &\quad = \exp\Big(-\int ds\, \mathbf{1}_{\{s>0\}}\mu\big(1 - \mathrm{e}^{-\lambda Y_s}\mathrm{P}_{Y_s}(Y_t = 0)\big)\Big) \\ &\quad = \exp\Big(-\int ds\, \mathbf{1}_{\{s>0\}}\mu\big(1 - \mathrm{e}^{-(\lambda+c(t))Y_s}\big)\Big) \\ &\quad = \exp\Big(-\int_0^\infty ds\, \tilde{\psi}'\big(u(\lambda + c(t), s)\big)\Big),\end{aligned}$$

where we used exponential formulas for Poisson point measure in the first equality and the Markov property of Y for the second equality. Putting things together, we then get (34). □

It is then easy to derive the distribution of the TMRCA A; see also [22].

COROLLARY 4.2. *The distribution function of A is given by*

$$\mathbb{P}(A \le t) = \mathbb{E}\big[\mathrm{e}^{-c(t)Z}\big] = \exp\Big(-\int_t^\infty ds\, \tilde{\psi}'(c(s))\Big),$$

and A has density f_A, with respect to the Lebesgue measure given by

$$f_A(t) = \tilde{\psi}'(c(t)) \exp\Big(-\int_t^\infty ds\, \tilde{\psi}'(c(s))\Big)\mathbf{1}_{\{t>0\}} = \frac{\tilde{\psi}'(c(t))}{\psi(c(t))}\mathrm{e}^{-\alpha t}\kappa_*\alpha\mathbf{1}_{\{t>0\}}. \tag{35}$$

PROOF. This is a direct consequence of Theorem 4.1 and (10). Use Lemma 3.6 to get (35). □

The next result is a direct consequence of Theorem 4.1.

COROLLARY 4.3. *Conditionally on A, the three random variables Z^I, Z^A and Z^O are independent.*

We can also give the mean of the population size just before the most recent common ancestor (MRCA) [to be compared to the mean size of the current population given by (23)].

COROLLARY 4.4. *Let $t>0$. We have*

$$\mathbb{E}[\mathrm{e}^{-\lambda Z^A}|A=t]=\frac{\mathbb{E}[\mathrm{e}^{-(\lambda+c(t))Z}]}{\mathbb{E}[\mathrm{e}^{-c(t)Z}]} \quad \text{and} \quad \mathbb{E}[Z^A|A=t]=\frac{\tilde{\psi}'(c(t))}{\psi(c(t))}. \tag{36}$$

PROOF. We deduce from Theorem 4.1 that

$$\mathbb{E}[\mathrm{e}^{-\lambda Z^A}; A\in dt]=dt\,\tilde{\psi}'(c(t))\exp\biggl(-\int_0^\infty ds\,\tilde{\psi}'\bigl(u\bigl(\lambda+c(t),s\bigr)\bigr)\biggr).$$

Thanks to (19), this implies that

$$\mathbb{E}[\mathrm{e}^{-\lambda Z^A}|A=t]=\frac{\mathrm{e}^{-\int_0^\infty ds\tilde{\psi}'(u(\lambda+c(t),s))}}{\mathrm{e}^{-\int_0^\infty ds\tilde{\psi}'(u(\lambda+c(t),s))}}=\frac{\mathbb{E}[\mathrm{e}^{-(\lambda+c(t))Z}]}{\mathbb{E}[\mathrm{e}^{-c(t)Z}]}.$$

The second part of the corollary is then a consequence of (22). $\square$

We deduce from (36) that the distribution of Z^A conditionally on $\{A=t\}$ converges, as t goes to infinity, to the distribution of Z.

As another application of Theorem 4.1, we get that the population just before the MRCA, Z^A, is stochastically smaller than the current population, Z. Note that strong inequality, namely inequality in the almost-surely sense, does not hold in general (see Section 7).

PROPOSITION 4.5. *We have $\mathbb{P}(Z^A\le z|A=t)\ge\mathbb{P}(Z\le z)$ for all $z\ge 0$ and $t\ge 0$. Hence, the population size Z^A is stochastically smaller than Z: $\mathbb{P}(Z^A\le z)\ge\mathbb{P}(Z\le z)$ for all $z\ge 0$. In particular, we have*

$$\mathbb{E}[Z^A|A]\le\mathbb{E}[Z] \qquad \text{a.s.}$$

PROOF. The first equality of (36) implies that for any nonnegative measurable function F defined on $\mathbb{R}$,

$$\mathbb{E}[F(Z^A)|A=t]=\frac{\mathbb{E}[F(Z)\mathrm{e}^{-c(t)Z}]}{\mathbb{E}[\mathrm{e}^{-c(t)Z}]}.$$

Note that $\mathrm{e}^{-c(t)Z}-\mathbb{E}[\mathrm{e}^{-c(t)Z}]$ is nonnegative for Z less than $\frac{1}{-c(t)}\log(\mathbb{E}[\mathrm{e}^{-c(t)Z}])$ and nonpositive otherwise, and that $\lim_{z\to\infty}\mathbb{E}[\mathrm{e}^{-c(t)Z}; Z\le z]-\mathbb{E}[\mathrm{e}^{-c(t)Z}]\mathbb{P}(Z\le z)=0$. We deduce that

$$\mathbb{P}(Z^A\le z|A=t)=\frac{\mathbb{E}[\mathrm{e}^{-c(t)Z}; Z\le z]}{\mathbb{E}[\mathrm{e}^{-c(t)Z}]}\ge\mathbb{P}(Z\le z).$$

For the last assertion, recall that for any nonnegative random variable, we have $\mathbb{E}[X]=\int_0^\infty \mathbb{P}(X>x)\,dx$. $\square$

REMARK 4.6. Instead of considering Z^A, the size of the population just before the MRCA, we could consider the size of the population at the MRCA, Z_+^A, which is formally given by

$$Z_+^A = Z^A + \sum_{i\in I} Y_0^i \mathbf{1}_{\{t_i=-A\}}.$$

Notice we do not take into account the contribution of $i \in J_2$ as for these indices we have $Y_0^i=0$. (In particular if $\pi=0$, then Z is continuous and $Z^A=Z_+^A$.) Similar computations as those in the proof of Theorem 4.1 yield the following: for $\lambda, t>0$

$$\mathbb{E}[\mathrm{e}^{-\lambda Z_+^A}|A=t]=\mathbb{E}[\mathrm{e}^{-\lambda Z^A}|A=t]\frac{\psi'(\lambda+c(t))-\psi'(\lambda)}{\psi'(c(t))-\psi'(0)}.$$

If $\psi''(0)=+\infty$, then we get that $\lim_{t\to+\infty}\mathbb{E}[\mathrm{e}^{-\lambda Z_+^A}|A=t]=0$. Thus, conditionally on $\{A=t\}$ with t large, we have that Z_+^A is likely to be very large. (Intuitively, a clan is born at time $-t$ which has survived up to time 0, and if t is large, it is very likely to have a large initial size.) Therefore, Z_+^A is not stochastically smaller than Z in the general case.

We may also consider the TMRCA of the immortal individual and individuals taken independently and uniformly among the current population living at time t. Let $J_t^n \subset \mathcal{I}$ be the indices of the clans of the randomly chosen n individuals alive at time t. (One individual chosen at random in the population at time t belongs to the clan, i with probability $Y_{t-t_i}^i/Z_t$.) Notice that $\mathrm{Card}(J_t^n)\leq n$. The TMRCA for the n individuals alive at time t and the immortal individual is given by

$$A_t^n := -\inf\{t_i; i\in J_t^n, i\in\mathcal{I}\}.$$

Because of stationarity, we shall focus on $t=0$ and write A^n for A_t^n. The joint distribution law of Z and A^n can be characterized by the following result.

THEOREM 4.7. *For any $n\geq 1$ and any $\lambda, T\geq 0$, we have*

$$\mathbb{E}\big[Z^n\mathrm{e}^{-\lambda Z}\mathbf{1}_{\{A^n\leq T\}}\big]=\frac{\mathrm{e}^{-\alpha c^{-1}(\lambda)}\kappa_*\alpha}{\psi(u(\lambda,T))}(-1)^n\frac{\partial^n}{\partial^n\eta}\left(\frac{\psi(u(\lambda+\eta,T))}{\psi(\lambda+\eta)}\right)\bigg|_{\eta=0}. \tag{37}$$

By integrating (37) n times in λ over $[\lambda,+\infty)$ we get an expression of $\mathbb{E}[\mathrm{e}^{-\lambda Z}\mathbf{1}_{\{A^n\leq T\}}]$ for all $\lambda\geq 0$ and $T\geq 0$, which characterizes the joint distribution of (Z, A^n). Thus, Theorem 4.7 indeed characterizes the joint distribution of Z and A^n.

PROOF OF THEOREM 4.7. By definition, we have

$$
\begin{aligned}
&\mathbb{E}\big[Z^n \mathrm{e}^{-\lambda Z}\mathbf{1}_{\{A^n\le T\}}\big]\\
&\quad=\mathbb{E}\Bigg[Z^n\sum_{i_1,\dots,i_n}\frac{Y^{i_1}_{-t_{i_1}}}{Z}\cdots\frac{Y^{i_n}_{-t_{i_n}}}{Z}\prod_{k=1}^n\mathbf{1}_{\{-t_{i_k}\le T\}}\mathrm{e}^{-\lambda Z}\Bigg]\\
&\quad=\mathbb{E}\bigg[\bigg(\int\mathcal{N}_3(ds,dY)Y_{-s}\mathbf{1}_{\{-s\le T\}}\bigg)^n\exp\bigg(-\lambda\int\mathcal{N}_3(ds,dY)Y_{-s}\bigg)\bigg]\\
&\quad=(-1)^n\frac{\partial^n}{\partial^n\eta}\mathbb{E}\bigg[\exp\bigg(-\int\mathcal{N}_3(ds,dY)\big(\eta Y_{-s}\mathbf{1}_{\{-s\le T\}}+\lambda Y_{-s}\big)\bigg)\bigg]\bigg|_{\eta=0}\\
&\quad=(-1)^n\frac{\partial^n}{\partial^n\eta}\exp\bigg(-\int_T^\infty ds\,\tilde{\psi}'(u(\lambda,s))-\int_0^T ds\,\tilde{\psi}'\big(u(\lambda+\eta,s)\big)\bigg)\bigg|_{\eta=0},
\end{aligned}
$$

where $\mathcal{N}_3$ in the second equality is defined by (13). The result then follows from (20) and (27). $\square$

REMARK 4.8. Following almost the same lines as the proof of Theorem 4.7, one can characterize explicitly the joint distribution of $\{(Z_{r_j}, A^{n_j}_{r_j}); 1\le j\le m\}$ for any $m,n_1,\dots,n_m\in\mathbb{N}^*$ and $-\infty<r_1<r_2<\cdots<r_m<\infty$.

5. Number of old families. We now consider the number families in the oldest clan alive at time 0. This corresponds to the number of individuals involved in the last coalescent event of the genealogical tree. To this end, we take representation (12) for Z.

DEFINITION 5.1. The number of oldest families alive at time 0 (excluding the immortal particle) is defined by

$$
N^A=\sum_{j\in\mathcal{J}}\mathbf{1}_{\{A=-t_j,Y^j_{-t_j}>0\}}=\sum_{j\in\mathcal{J}}\mathbf{1}_{\{A=-t_j,\zeta_j>-t_j\}}. \tag{38}
$$

We have $N^A\ge 1$. In the particular case $\pi=0$ and $\beta>0$, we have $\mathcal{J}=J_2$ and $N^A=1$.

The following proposition gives the joint law of A, N^A and Z.

PROPOSITION 5.2. *We have for $a\in[0,1]$, $\lambda\ge 0$, $t\ge 0$,*

$$
\mathbb{E}[a^{N^A}\mathrm{e}^{-\lambda Z}|A=t]=\frac{\psi'(c(t))-\psi'((1-a)c(t)+au(\lambda,t))}{\tilde{\psi}'(c(t))}\mathrm{e}^{-\int_0^t\tilde{\psi}'(u(\lambda,r))\,dr}
$$

and

$$
\mathbb{E}[a^{N^A}|A=t]=\frac{\psi'(c(t))-\psi'((1-a)c(t))}{\tilde{\psi}'(c(t))}=1-\frac{\tilde{\psi}'((1-a)c(t))}{\tilde{\psi}'(c(t))}.
$$

PROOF. Recall notations from Section 3.1. For $i \in \mathcal{I}$, we set $J_i^* = J_{1,i}$ if $i \in I$ and $J_i^* = \{i\}$ if $i \in J_2$. Given any nonnegative function f, we have, using (12) and (16),

$$\begin{aligned}
&\mathbb{E}[a^{N^A} \mathrm{e}^{-\lambda Z} f(A)] \\
&\quad = \mathbb{E}\Big[\mathrm{e}^{-\lambda \sum_{k\in\mathcal{I}} Y^k_{-t_k}} \sum_{i\in\mathcal{I}} a^{\sum_{j\in J_i^*} \mathbf{1}_{\{\zeta_j > -t_i\}}} f(-t_i) \mathbf{1}_{\{Y^i_{-t_i} \neq 0\}} \mathbf{1}_{\{\sum_{k'\in\mathcal{I}, t_{k'} < t_i} \mathbf{1}_{\{Y^{k'}_{-t_i} > 0\}} = 0\}}\Big] \\
&\quad = \int_0^\infty ds\, f(s) \mathbb{E}\big[\mathrm{e}^{-\lambda \sum_{k\in\mathcal{I}} Y^k_{-t_k} \mathbf{1}_{\{t_k > -s\}}}\big] \mathbb{P}\Big(\sum_{k\in\mathcal{I}} \mathbf{1}_{\{t_k < -s, Y^k_s > 0\}} = 0\Big) \\
&\qquad \times \Big(2\beta \mathbb{N}[a \mathrm{e}^{-\lambda Y_s} \mathbf{1}_{\{Y_s > 0\}}] \\
&\qquad\qquad + \int_{(0,+\infty)} \ell\pi(d\ell) \mathbb{E}_\ell\big[a^{\sum_{j\in J_3} \mathbf{1}_{\{Y^j_s > 0\}}} \mathrm{e}^{-\lambda \sum_{j\in J_3} Y^j_s} \mathbf{1}_{\{\sum_{j\in J_3} Y^j_s > 0\}}\big]\Big),
\end{aligned}$$

where $\sum_{j\in J_3} \delta_{Y^j}(dY)$ is under $\mathbb{E}_\ell$ a Poisson point measure with intensity $\ell\mathbb{N}[dY]$. We have

$$\mathbb{E}\big[\mathrm{e}^{-\lambda \sum_{k\in\mathcal{I}} Y^k_{-t_k} \mathbf{1}_{\{t_k > -s\}}}\big] \mathbb{P}\Big(\sum_{k\in\mathcal{I}} \mathbf{1}_{\{t_k < -s, Y^k_s > 0\}} = 0\Big) = \mathrm{e}^{-\int_0^s dr\, \tilde{\psi}'(u(\lambda,r)) - \int_s^\infty dr\, \tilde{\psi}'(c(r))}.$$

We also have

$$\mathbb{N}\big[\mathrm{e}^{-\lambda Y_s} \mathbf{1}_{\{Y_s > 0\}}\big] = \mathbb{N}[Y_s > 0] - \mathbb{N}[1 - \mathrm{e}^{-\lambda Y_s}] = c(s) - u(\lambda, s)$$

and

$$\begin{aligned}
&\mathbb{E}_\ell\big[a^{\sum_{j\in J_3} \mathbf{1}_{\{Y^j_s > 0\}}} \mathrm{e}^{-\lambda \sum_{j\in J_3} Y^j_s} \mathbf{1}_{\{\sum_{j\in J_3} Y^j_s > 0\}}\big] \\
&\quad = \mathbb{E}_\ell\big[a^{\sum_{j\in J_3} \mathbf{1}_{\{Y^j_s > 0\}}} \mathrm{e}^{-\lambda \sum_{j\in J_3} Y^j_s}\big] - \mathbb{P}_\ell\Big(\sum_{j\in J_3} Y^j_s = 0\Big) \\
&\quad = \exp\big(-\ell \mathbb{N}\big[(1 - a\mathrm{e}^{-\lambda Y_s}) \mathbf{1}_{\{Y_s > 0\}}\big]\big) - \exp(-\ell \mathbb{N}[Y_s > 0]) \\
&\quad = \exp\big(-\ell \mathbb{N}[Y_s > 0] + \ell a \mathbb{N}[\mathrm{e}^{-\lambda Y_s}] \mathbf{1}_{\{Y_s > 0\}}\big) - \exp(-\ell \mathbb{N}[Y_s > 0]) \\
&\quad = \exp\big(-\ell\big((1-a)c(s) - a u(\lambda, s)\big)\big) - \exp(-\ell c(s)).
\end{aligned}$$

Thus, we get

$$\begin{aligned}
&2\beta \mathbb{N}\big[a \mathrm{e}^{-\lambda Y_s} \mathbf{1}_{\{Y_s > 0\}}\big] \\
&\quad + \int_{(0,+\infty)} \ell\pi(d\ell) \mathbb{E}_\ell\big[a^{\sum_{j\in J_3} \mathbf{1}_{\{Y^j_s > 0\}}} \mathrm{e}^{-\lambda \sum_{j\in J_3} Y^j_s} \mathbf{1}_{\{\sum_{j\in J_3} Y^j_s > 0\}}\big] \\
&\quad = \psi'(c(s)) - \psi'\big((1-a)c(s) + a u(\lambda, s)\big).
\end{aligned}$$

Putting things together, we obtain

$$\begin{aligned}\mathbb{E}[a^{N^A} \mathrm{e}^{-\lambda Z} f(A)] \\ &= \int_0^\infty ds\, f(s) \mathrm{e}^{-\int_0^s dr\, \tilde{\psi}'(u(\lambda,r)) - \int_s^\infty dr\, \tilde{\psi}'(c(r))} \\ &\qquad \times \bigl[\psi'(c(s)) - \psi'\bigl((1-a)c(s) + au(\lambda, s)\bigr)\bigr].\end{aligned}$$

Then, use (35) for the density of A to get the result. $\square$

COROLLARY 5.3. *We have*

$$\mathbb{P}(N^A = n | A = t) = (-1)^{n+1} \frac{c(t)^n \psi^{(n+1)}(c(t))}{n! \tilde{\psi}'(c(t))}, \qquad n \in \mathbb{N}^*. \tag{39}$$

Suppose that $\psi''(0+) < \infty$ (*i.e.*, $\mathbb{E}[Z] < +\infty$). *Then*, *we have*

$$\mathbb{E}[N^A | A = t] = \psi''(0) \frac{c(t)}{\tilde{\psi}'(c(t))}.$$

Furthermore the function $t \mapsto \mathbb{E}[N^A | A = t]$ *is nonincreasing*.

PROOF. The first two assertions are straightforward consequences of Proposition 5.2. To get the monotonicity of $t \mapsto \mathbb{E}[N^A | A = t]$, we simply notice that both $t \mapsto c(t)$ and

$$x \mapsto \frac{\tilde{\psi}'(x)}{x} = 2\beta + \int_0^\infty \pi(d\ell) \ell \frac{1 - \mathrm{e}^{-x\ell}}{x}$$

are nonincreasing. $\square$

REMARK 5.4. Suppose that $\psi''(0+) < \infty$. We deduce from (39) that

$$\lim_{t \to +\infty} \mathbb{P}(N^A = 1 | A = t) = 1.$$

Thus, the distribution of N^A conditionally on $\{A = t\}$ converges as t goes from infinity to 1. So roughly speaking N^A is likely to be equal to 1 if the TMRCA (or age of the oldest clan alive) is large. Notice that if $\psi''(0+) = +\infty$, this result may be false (see the next remark).

REMARK 5.5. While the foregoing corollary shows that the conditional expectation of N^A given $A = t$ is monotonic, it is not true, in general, that the conditional distribution of N^A given $A = t$ is stochastically monotonic. For example, this is not the case if $\psi''(0) < \infty$, $\beta > 0$, $\pi \neq 0$ and $\tilde{\psi}'(\lambda) \sim 2\beta\lambda$ as λ goes to infinity. Indeed, using the Laplace transform of N^A, one gets that, conditionally on $\{A = t\}$, N^A converges in distribution to 1 as t goes to 0 or infinity, whereas N^A is not equal to 1 a.s. as $\pi \neq 0$.

REMARK 5.6. Let us consider the stable cases, $\psi(\lambda) = \alpha\lambda + c_0\lambda^{1+\alpha_0}$, with $c_0 > 0$ and $\alpha_0 \in (0, 1]$. We deduce from Corollary 5.3 that

$$\mathbb{E}[a^{N^A}|A = t] = 1 - (1 - a)^{\alpha_0}.$$

In particular, N^A is independent of A. The case $\alpha_0 = 1$ corresponds to the quadratic branching mechanism, and we get that a.s. $N^A = 1$. For $\alpha_0 \in (0, 1)$, we deduce from (39) that for $n \in \mathbb{N}^*$

$$\mathbb{P}(N^A = n|A = t) = \frac{1}{n!}\alpha_0 \prod_{k=1}^{n-1}(k - \alpha_0).$$

For $\alpha_0 \in (0, 1)$, we have $\psi''(0+) = +\infty$, and the result of Remark 5.4 does not hold.

6. Asymptotics for the number of ancestors. The number $N_{-s,0}$ defined by (30) of clans born before time $-s$ and alive at time 0 is nondecreasing and is distributed as a Poisson random variable with parameter $\Lambda(s)$ given by (31). As $\Lambda(s)$ goes to infinity as s goes down to 0, we deduce that $N_{-s,0}$ tends to infinity almost surely as $s \downarrow 0+$. A natural question is then how fast the numbers $N_{-s,0}$ tend to infinity. It follows from the definition of the Poisson random measure $\mathcal{N}_3$ in (13) that $\{N_{-\Lambda^{-1}(s),0}; s \geq 0\}$ is Poisson process with parameter 1, and by the strong law of large numbers for Lévy processes (see [8]), we deduce that

$$\lim_{s\downarrow 0+} \frac{N_{-s,0}}{\Lambda(s)} = 1 \qquad \text{almost surely.}$$

One can also ask how fast the number M_s of ancestors at time $-s$ of the current population living at time 0 tends to infinity. To answer this question, we need to introduce the genealogy of the families. Notice the genealogy of a CB contains more information than the CB itself.

6.1. *Genealogy of CB.* We recall here the definition of the Lévy continuum random tree (CRT) introduced in [35, 36] and developed later in [16] for critical or sub-critical branching mechanism. See also [17, 34] for a real trees setting.

We first recall the coding of a compact real tree by a continuous function $g : [0, +\infty) \to [0, +\infty)$ with compact support and such that $g(0) = 0$. We also assume that g is not identically 0. For every $0 \leq s \leq t$, we set

$$m_g(s, t) = \inf_{u\in[s,t]} g(u) \quad \text{and} \quad d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

We then introduce the equivalence relation $s \sim t$ if and only if $d_g(s, t) = 0$. Let $\mathcal{T}_g$ be the quotient space $[0, +\infty)/\sim$. It is easy to check that d_g induces a distance on $\mathcal{T}_g$. Moreover, $(\mathcal{T}_g, d_g)$ is a compact real tree (see [17], Theorem 2.1). We say that g is the height process of the tree $\mathcal{T}_g$. For instance, when g is a normalized Brownian excursion, the associated real tree is Aldous's CRT [4].

We get from Section 3.2 in [17] and Theorem 1.4.3 in [16] [for the continuity of H under (A1)] the following result.

THEOREM 6.1. *Under* (A1), *there exists a continuous process* $H = (H_s, s \geq 0)$, *called height process*, *and a càdlàg process* $L(H) = (L^a, a > 0)$, *called local time of* H, *defined under a* σ*-finite measure* $\mathbf{N}$, *called the excursion measure of* H, *such that* $\mathbf{N}$*-a.e.*:

- $H_0 = 0$, $H_s = 0$ *for* $s \geq \sigma$ *where* σ *is finite and defined by*

$$\sigma = \inf\{s > 0; H_s = 0\};$$

- *for all* $a > 0$,

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^\sigma \mathbf{1}_{\{a-\varepsilon < H_r \leq a\}} \, dr = L^a \qquad \text{in } L^1(\mathbf{N}).$$

- *The process* $L(H)$ *is distributed under the excursion measure* $\mathbf{N}$ *as* Y *under its excursion measure* $\mathbb{N}$.

In order to simplify notations, we shall identify Y with $L(H)$ as well as $\mathbb{N}$ with $\mathbf{N}$.

The tree $(\mathcal{T}_H, d_H)$ corresponding to H is called a Lévy tree. Informally, L^a measures the number of vertices (in fact leaves) of $\mathcal{T}_H$ at level a under $\mathbf{N}$.

Let $a > 0$ be fixed. We consider the excursions of the height process H above a under the excursion measure $\mathbb{N}$. Precisely, let (u_k, v_k), $k \in \mathcal{K}$ be the excursions of H above a over the time interval $[0, \sigma]$. We set $H^k = (H_{(u_k+s)\wedge v_k} - a, s \geq 0)$.

The next result is a consequence of Proposition 4.2.3 in [16].

PROPOSITION 6.2. *Conditionally on* $(L^r, r \leq a)$, *the measure*

$$\sum_{k \in \mathcal{K}} \delta_{H^k}(dH)$$

is a Poisson point measure with intensity $L^a \mathbb{N}[dH]$.

We give a definition for the number of ancestors, which will be used in the next section.

DEFINITION 6.3. The number of ancestors at time a of the population (coded by H) alive at time b is the number of excursions of H above level a which reach level $b > a$.

$$R_{a,b}(H) = \sum_{k \in \mathcal{K}} \mathbf{1}_{\{\zeta_k \geq b-a\}},$$

where $\zeta_k = \max\{H_s^k, s \geq 0\}$.

6.2. *Genealogy of Z.* Recall notations from Section 6.1.

We use formulation (12) to construct the genealogy of Z. Recall notation $\mathcal{N}_0$ from Section 3.1:

- Conditionally on $\mathcal{N}_0$, let $\tilde{\mathcal{N}_1}(dt, dH) = \sum_{j\in J_1} \delta_{t_j, H^j}(dt, dH)$ be a Poisson point measure with intensity $\nu(dt)\mathbb{N}[dH]$ with $\nu(dt) = \sum_{i\in I} r_i \delta_{t_i}(dt)$.
- Let $\tilde{\mathcal{N}_2}(dt, dH) = \sum_{j\in J_2} \delta_{t_j, H^j}(dt, dH)$ be a Poisson point measure independent of $(\mathcal{N}_0, \tilde{\mathcal{N}_1})$ and with intensity $2\beta\, dt\, \mathbb{N}[dH]$.

We will write Y^j for $L(H^j)$ for $j \in \mathcal{J} = J_1 \cup J_2$. Thus notation (12) is still consistent with the previous sections, thanks to Proposition 6.2. And the process $\sum_{j\in\mathcal{J}} \delta_{t_j, H^j}$ allows to code for the genealogy of the families of Z.

Let $s > 0$. Following Definition 6.3, we consider M_s the number of ancestors at time $-s$ of the current population living at time 0, excluding the immortal individual.

$$M_s = \sum_{j\in\mathcal{J}} \mathbf{1}_{\{t_j<-s\}} R_{-s-t_j,-t_j}(H^j).$$

6.3. *Asymptotics for the number of ancestors.* We first give a technical lemma, whose proof is postponed to the end of this section.

LEMMA 6.4. *The joint distribution of M_s and Z_0, conditionally on Z_{-s}, is characterized by the following equation: for $\eta, \lambda \geq 0$ $s > 0$,*

$$\mathbb{E}[\mathrm{e}^{-\eta M_s - \lambda Z_0} | Z_{-s}] = \mathrm{e}^{-\int_0^s dr\, \tilde{\psi}'(u(\lambda,r))} \mathrm{e}^{-Z_{-s}[(1-\mathrm{e}^{-\eta})c(s)+\mathrm{e}^{-\eta}u(\lambda,s)]}. \tag{40}$$

In particular, M_s is, conditionally on Z_{-s}, distributed as a Poisson random variable with parameter $c(s)Z_{-s}$.

In a sense, M_s counts the number of excursions of the height process at time $-s$ above level s. It is well known, see the second equality in (21) of [17], that for CSBP processes, the number of excursions at level $t-\varepsilon$ which reach level t divided by $c(\varepsilon)$ (i.e., the excursion measure of all the excursions with height larger than ε) converge a.s. to the local time at level t. Mimicking the proof of the second equality in (21) of [17], which relies on the fact that M_s is increasing and distributed as a Poisson random variable with (random) parameters which converge, we get the following result.

COROLLARY 6.5. *The following convergence holds:*

$$\lim_{s\to 0} \frac{M_s}{c(s)} = Z_0 \qquad \textit{almost surely.}$$

REMARK 6.6. Suppose in addition that $\int_0^\infty \ell^2\pi(d\ell) < \infty$. Set $\tilde{\pi}(d\ell) = \ell^2\pi(d\ell)$. Then the $\tilde{\pi}$-coalescent N^μ defined in [6], where N_t^μ is the number of ancestors at time t for the coalescent process, comes down from infinity by the assumption (A2) (see [6] and the references therein). It was shown in [6] (see also [39]) that the speed of coming down from infinity satisfies

$$\lim_{t\downarrow 0+} \frac{N_t^\mu}{c(t)} = 1 \qquad \text{almost surely.} \tag{41}$$

From the heuristic duality between coalescence and branching processes, our result in Corollary 6.5 can be seen as a duality to (41).

The next theorem gives the speed of convergence of Z_{-s} and $M_s/c(s)$ to Z_0 when ψ behaves like a power at infinity; for the quadratic case, see Theorem 7.8. Notice the behavior is different in the asymptotic stable case and in the quadratic case.

THEOREM 6.7. *Assume there exists $a > 0$ and $\alpha_0 \in (1, 2)$ such that $\lim_{\lambda\to+\infty} \lambda^{-\alpha_0}\psi(\lambda) = a$. Set $h(s) = s^{-1/\alpha_0}$. Then the following convergence holds in distribution*:

$$\left(Z_{-s}, h(s)\left(Z_0 - \frac{M_s}{c(s)}\right), h(s)(Z_0 - Z_{-s})\right) \xrightarrow[s\downarrow 0+]{(d)} (Z_0, V_{Z_0}, V_{Z_0}),$$

where V is a stable process independent of Z_0 with Laplace transform $\mathbb{E}[\mathrm{e}^{-\lambda V_t}] = \mathrm{e}^{at\lambda^{\alpha_0}}$.

The Lévy measure of V is $\mathbf{1}_{\{x>0\}} x^{-\alpha_0-1}\,dx$ up to a multiplicative constant.

PROOF OF THEOREM 6.7. Let $\lambda \geq 0$ and $\eta \geq 0$. We get from Lemma 6.4 that

$$\mathbb{E}\big[\mathrm{e}^{-\eta h(s)(Z_0 - c(s)^{-1}M_s) - \lambda h(s)(Z_0 - Z_{-s})} | Z_{-s}\big] = \mathrm{e}^{-\int_0^s dr \tilde{\psi}'(u((\lambda+\eta)h(s),r))}\mathrm{e}^{-\Delta_s Z_{-s}},$$

where we set

$$\Delta_s = \big(1 - \mathrm{e}^{\eta h(s)/c(s)}\big)c(s) + \mathrm{e}^{\eta h(s)/c(s)} u\big((\lambda+\eta)h(s), s\big) - \lambda h(s). \tag{42}$$

Let $q > 0$ be fixed. Notice that $s\psi(rh(s))$ is bounded near 0. Then we deduce from (4) that $u(qh(s), s) \leq qh(s)$ and

$$u(qh(s), s) = qh(s) - aq^{\alpha_0} + o(1),$$

where $o(1)$ denotes any function of s which converges to 0 as s goes down to 0. Using (20), we get

$$\int_0^s dr\, \tilde{\psi}'(u(qh(s), r)) = \log\left(\frac{\psi(qh(s))}{\psi(u(qh(s), s))\mathrm{e}^{\alpha s}}\right) = o(1). \tag{43}$$

We deduce from (9) that $c(s) \sim (as(\alpha_0 - 1))^{-1/(\alpha_0-1)}$ at infinity. Thus, we get $h^2(s)/c(s)$ is of order $s^{-2/\alpha_0+1/(\alpha_0-1)} = s^{(2-\alpha_0)/(\alpha_0-1)\alpha_0}$ and thus $h^2(s)/c(s) = o(1)$. We compute

$$\begin{aligned}\Delta_s &= \left(-\eta\frac{h(s)}{c(s)} + \frac{1}{c(s)}o(1)\right)c(s)\\ &\quad + \left(1 + \frac{1}{\sqrt{c(s)}}o(1)\right)\bigl((\lambda+\eta)h(s) - a(\lambda+\eta)^{\alpha_0} + o(1)\bigr) - \lambda h(s)\\ &= -a(\lambda+\eta)^{\alpha_0} + o(1).\end{aligned}$$

We deduce that for any $\eta \geq 0$, $\lambda \geq 0$ and $b > a(\lambda+\eta)^{\alpha_0}$, we have, for s small enough,

$$\begin{aligned}&\mathbb{E}\bigl[\mathrm{e}^{-bZ_{-s}}\mathrm{e}^{-\eta h(s)(Z_0 - c(s)^{-1}M_s) - \lambda h(s)(Z_0 - Z_{-s})}\bigr]\\ &\quad = \mathbb{E}\bigl[\mathrm{e}^{-bZ_{-s}}\mathrm{e}^{(a(\lambda+\eta)^{\alpha_0}+o(1))Z_{-s}}\bigr] + o(1)\\ &\quad = \mathbb{E}\bigl[\mathrm{e}^{-bZ_0}\mathrm{e}^{(a(\lambda+\eta)^{\alpha_0}+o(1))Z_0}\bigr] + o(1)\\ &\quad \underset{s\downarrow 0+}{\longrightarrow} \mathbb{E}\bigl[\mathrm{e}^{-bZ_0}\mathrm{e}^{a(\lambda+\eta)^{\alpha_0}Z_0}\bigr]\\ &\quad = \mathbb{E}\bigl[\mathrm{e}^{-bZ_0}\mathrm{e}^{-(\lambda+\eta)V_{Z_0}}\bigr].\end{aligned}$$

An easy adaptation of [42] to multidimensional Laplace transform yields the result. □

PROOF OF LEMMA 6.4. For any $b, \eta, \lambda \geq 0$, we have

$$\begin{aligned}&\mathbb{E}[\mathrm{e}^{-bZ_{-s}-\eta M_s - \lambda Z_0}]\\ &\quad = \mathbb{E}\bigl[\mathrm{e}^{-\lambda\sum_{j\in\mathcal{J}}\mathbf{1}_{\{-s\leq t_j\leq 0\}}Y^j_{-t_j}}\bigr]\\ &\qquad \times \mathbb{E}\Bigl[\exp\Bigl(-bZ_{-s} - \eta M_s - \lambda\sum_{j\in\mathcal{J}}\mathbf{1}_{\{t_j<-s\}}Y^j_{-t_j}\Bigr)\Bigr]\\ &\quad = \exp\Bigl(-\int_0^s dr\, \tilde{\psi}'(u(\lambda,r))\Bigr) \qquad (44)\\ &\qquad \times \mathbb{E}\Bigl[\exp\Bigl(-\sum_{j\in\mathcal{J}}\mathbf{1}_{\{t_j<-s\}}\bigl(bY^j_{-s-t_j} + \eta R_{-s-t_j,-t_j}(H^j) + \lambda Y^j_{-t_j}\bigr)\Bigr)\Bigr]\\ &\quad = \exp\Bigl(-\int_0^s dr\, \tilde{\psi}'(u(\lambda,r))\Bigr)\\ &\qquad \times \exp\Bigl(-\int_0^\infty da\, \tilde{\psi}'(\mathbb{N}[1 - \exp(-bY_a - \eta R_{a,a+s}(H) - \lambda Y_{a+s})])\Bigr),\end{aligned}$$

where we used that Poisson random measures over disjoint sets are independent in the first equality, Lemma 3.1 in the second equality and a immediate generalization of Lemma 3.1 to genealogies in the third equality.

Using notations from Section 6.1 on the Poissonian representation of the height process above level a from Proposition 6.2, we get

$$\begin{aligned}\mathbb{N}\big[1-\mathrm{e}^{-bY_a-\eta R_{a,a+s}(H)-\lambda Y_{a+s}}\big] &= \mathbb{N}\big[1-\mathrm{e}^{-bY_a-\sum_{k\in\mathcal{K}}\eta\mathbf{1}_{\{\zeta_k\geq s\}}-\lambda Y(H^k)_s}\big]\\ &= \mathbb{N}\big[1-\mathrm{e}^{-Y_a(b+\mathbb{N}[1-\exp(-\eta\mathbf{1}_{\{\zeta\geq s\}}-\lambda Y_s)])}\big].\end{aligned}$$

Recall that on $\{\zeta < s\}$ we have $Y_s = 0$. As $1-\exp(-\eta\mathbf{1}_{\{\zeta\geq s\}}-\lambda Y_s) = (1-\mathrm{e}^{-\eta})\mathbf{1}_{\{\zeta\geq s\}}+\mathrm{e}^{-\eta}(1-\mathrm{e}^{-\lambda Y_s})$, we deduce that

$$\mathbb{N}\big[1-\mathrm{e}^{-bY_a-\eta R_{a,a+s}(H)-\lambda Y_{a+s}}\big] = \mathbb{N}[1-\mathrm{e}^{-\lambda' Y_a}] = u(\lambda', a)$$

with $\lambda' = b+(1-\mathrm{e}^{-\eta})c(s)+\mathrm{e}^{-\eta}u(\lambda,s)$. Then we use Lemma 3.6 to write

$$\begin{aligned}&\exp\Big(-\int_0^\infty da\,\tilde{\psi}'(\mathbb{N}[1-\exp(-bY_a-\eta R_{a,a+s}(H)-\lambda Y_{a+s})])\Big)\\ &\quad=\exp\Big(-\int_0^\infty da\,\tilde{\psi}'(u(\lambda',a))\Big)\\ &\quad=\mathbb{E}[\mathrm{e}^{-\lambda' Z_{-s}}].\end{aligned}$$

Plugging this in (44), we deduce (40). $\square$

7. The quadratic branching mechanism.

Let $(\mathbf{e}_k; k\in\mathbb{N})$ be independent exponential random variables with mean 1.

7.1. *Preliminaries.* In this section we give some explicit distributions and more precise results for the case of quadratic branching mechanism.

$$\psi(\lambda) = \beta\lambda^2 + 2\beta\theta\lambda, \tag{45}$$

where $\beta > 0$ and $\theta > 0$. We have

$$u(\lambda,t) = \frac{2\theta\lambda}{(2\theta+\lambda)\mathrm{e}^{2\theta\beta t}-\lambda},\qquad c(t) = \frac{2\theta}{\mathrm{e}^{2\theta\beta t}-1},\qquad \kappa_* = 2\theta.$$

For every $t\in\mathbb{R}$, it follows from Corollary 3.3 that the process $\{Z_{s+t}; s\geq 0\}$ has the same distribution as the strong solution of the following stochastic differential equation:

$$dX_s = \sqrt{2\beta X_s}\,dW_s + 2\beta(1-\theta X_s)\,ds$$

with initial law $\mathbb{P}(Z_0\in\cdot)$, where W is a standard Brownian motion; see [46], Section XI.3, for the existence of strong solution.

7.2. *Joint law of the TMRCA and populations sizes.* We have the following representations.

THEOREM 7.1. *Assume ψ is given by* (45).

(i) *We have, for $\lambda \geq 0$,*

$$\mathbb{E}[\mathrm{e}^{-\lambda Z}] = \left(\frac{2\theta}{2\theta+\lambda}\right)^2 \quad \textit{and} \quad Z \overset{(d)}{=} \frac{1}{2\theta}(\mathbf{e}_1 + \mathbf{e}_2). \tag{46}$$

(ii) *We have, for $t \geq 0$,*

$$\mathbb{P}(A \leq t) = (1 - \mathrm{e}^{-2\theta\beta t})^2 \quad \textit{and} \quad A \overset{(d)}{=} \frac{1}{2\theta\beta} \max(\mathbf{e}_1, \mathbf{e}_2). \tag{47}$$

(iii) *Conditionally on $\{A = t\}$, we have the following distribution representation*:

$$(Z^A, Z^I, Z^O) \overset{(d)}{=} \left(\frac{\mathbf{e}_1 + \mathbf{e}_2}{2\theta + c(t)}, \frac{\mathbf{e}_3 + \mathbf{e}_4}{2\theta + c(t)}, \frac{\mathbf{e}_5}{2\theta + c(t)}\right). \tag{48}$$

PROOF. By Lemma 19, we have

$$\mathbb{E}[\mathrm{e}^{-\lambda Z}] = \left(\frac{2\theta}{2\theta+\lambda}\right)^2.$$

This gives (i). Using Theorem 4.1, we obtain

$$\begin{aligned}
&\mathbb{E}[\mathrm{e}^{-\lambda Z^A - \gamma Z^I - \eta Z^O}; A \in dt] \\
&\quad = \frac{2\beta(2\theta)^6 \mathrm{e}^{6\theta\beta t}(\mathrm{e}^{2\theta\beta t} - 1)}{[(2\theta+\eta)\mathrm{e}^{2\theta\beta t} - \eta][(2\theta+\gamma)\mathrm{e}^{2\theta\beta t} - \gamma]^2[(2\theta+\lambda)\mathrm{e}^{2\theta\beta t} - \lambda]^2}\, dt.
\end{aligned}$$

We then deduce (ii) and (iii). □

We then are able to compare more precisely the size of the current population $Z = Z^I + Z^O$ with the size of the population Z^A just before the birth time of the MRCA. As $(Z_t, t \in \mathbb{R})$ is continuous, notice that that Z^A is also the size of the population at the birth time of the MRCA. Recall that Z^A is stochastically smaller than Z. The next corollary indicates that Z^A is, however, not a.s. smaller than Z.

COROLLARY 7.2. *Assume ψ is given by* (45). *We have a.s.*

$$\mathbb{P}(Z^A < Z|A) = \tfrac{11}{16} \quad \textit{and} \quad \mathbb{E}[Z^A|A] = \tfrac{2}{3}\mathbb{E}[Z|A]$$

as well as

$$\mathbb{P}(Z^A < Z) = \tfrac{11}{16} \quad \textit{and} \quad \mathbb{E}[Z^A] = \tfrac{2}{3}\mathbb{E}[Z].$$

PROOF. We have

$$\mathbb{P}(Z^A < Z|A) = \mathbb{P}(\mathbf{e}_1 + \mathbf{e}_2 < \mathbf{e}_3 + \mathbf{e}_4 + \mathbf{e}_5) = \tfrac{11}{16}.$$

The other equalities are obvious. □

There is also an interesting result (which is not valid for general branching mechanism) which can be interpreted by time reversal. Recall ζ is the extinction time of Y.

PROPOSITION 7.3. *Assume ψ is given by* (45). *Conditionally on Z, A is distributed as ζ under P_Z: for all $t \geq 0$*

$$\mathbb{P}(A \leq t|Z) = \mathrm{e}^{-c(t)Z} = \mathrm{P}_Z(\zeta \leq t). \tag{49}$$

PROOF. We deduce from (46) and (47) that the densities of Z and A are

$$f_A(t) = 4\theta\beta \mathrm{e}^{-2\theta\beta t}(1 - \mathrm{e}^{-2\theta\beta t})\mathbf{1}_{\{t>0\}} \quad \text{and} \quad f_Z(z) = (2\theta)^2 z \mathrm{e}^{-2\theta z}\mathbf{1}_{\{z>0\}}. \tag{50}$$

We also deduce from (48) the density of Z, conditionally on $A = t$.

$$f_{Z|A=t}(z) = \big(2\theta + c(t)\big)^3 z^2 \mathrm{e}^{-(2\theta+c(t))z}\mathbf{1}_{\{z>0\}}.$$

Using Bayes's rule, we get the density of A conditionally on $Z = z$: for $z, t > 0$

$$\begin{aligned} f_{A|Z=z}(t) = f_{Z|A=t}(z)\frac{f_A(t)}{f_Z(z)} &= \frac{z(2\theta)^2\beta}{(\mathrm{e}^{2\theta\beta t} - 1)^2}\mathrm{e}^{2\theta\beta t}\exp\biggl(-\frac{2\theta z}{\mathrm{e}^{2\theta\beta t} - 1}\biggr) \\ &= -c'(t) z \mathrm{e}^{-c(t)z}. \end{aligned}$$

We obtain $\mathbb{P}(A \leq t|Z) = \mathrm{e}^{-c(t)Z}$. Then, we conclude as

$$\mathrm{P}_r(\zeta \leq t) = \mathrm{e}^{-r\mathbb{N}[\zeta \geq t]} = \mathrm{e}^{-rc(t)},$$

where we used the Poissonian representation of Y given by (7). □

Notice that (49) implies that

$$\mathbb{P}\big(c(A)Z \geq c(t)Z|Z\big) = \mathbb{P}(A \leq t|Z) = \mathrm{e}^{-c(t)Z}.$$

We obtain that $c(A)Z$ is independent of Z and $c(A)Z \overset{(\mathrm{d})}{=} \mathbf{e}_1$. We thus deduce the following corollary.

COROLLARY 7.4. *Assume ψ is given by* (45). *We have the following representation*:

$$(Z, c(A), Z^A) \overset{(d)}{=} \biggl(\frac{\mathbf{e}_1 + \mathbf{e}_2}{2\theta}, 2\theta\frac{\mathbf{e}_3}{\mathbf{e}_1 + \mathbf{e}_2}, \frac{1}{2\theta}\frac{\mathbf{e}_1 + \mathbf{e}_2}{\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3}(\mathbf{e}_4 + \mathbf{e}_5)\biggr).$$

REMARK 7.5. It is also easy to check that conditionally on $\{Z = z\}$, A is distributed as $\frac{1}{2\beta\theta}\log(1 + \frac{2\theta z}{\mathbf{e}_3})$. In particular, we deduce that A is distributed as $\frac{1}{2\beta\theta}\log(1 + \frac{\mathbf{e}_1+\mathbf{e}_2}{\mathbf{e}_3})$.

7.3. *TMRCA for n individuals.* Next, we consider the joint distribution of Z and A^n the TMRCA of the immortal individual and n individuals chosen at random among the current population. The next result is a direct application of Theorem 4.7.

PROPOSITION 7.6. *Assume ψ is given by (45). We set $s = 1 - \mathrm{e}^{-2\beta\theta t}$. We have, for $n \in \mathbb{N}^*$,*

$$\mathbb{E}\big[Z^n \mathrm{e}^{-\lambda Z} \mathbf{1}_{\{A^n \in [0,t]\}}\big] = \frac{(n+1)! s^n}{(2\theta + \lambda s)^n} \left(\frac{2\theta}{2\theta + \lambda}\right)^2,$$

and the size-biased distribution of A^n is the maximum of n independent exponential random variables with mean 1,

$$\mathbb{E}\big[Z^n \mathbf{1}_{\{A^n \in [0,t]\}}\big] = \mathbb{E}[Z^n](1 - \mathrm{e}^{-2\beta\theta t})^n.$$

We can compute explicitly the distribution of A^1. See also [32], Section 3, for similar computations in a slightly different setting.

PROPOSITION 7.7. *Assume ψ is given by (45). We set $s = 1 - \mathrm{e}^{-2\beta\theta t}$. We have*

$$\begin{aligned} &\mathbb{P}(A^1 \le t) = 2\frac{s}{1-s}\left(1 + \frac{s}{1-s}\log(s)\right) \quad \textit{and} \\ &\mathbb{P}\big(c(A^1)Z \ge x | Z\big) = \frac{2}{x} - \frac{2}{x^2}(1 - \mathrm{e}^{-x}). \end{aligned} \tag{51}$$

In particular, $c(A^1)Z$ is independent of Z.

Notice that $\mathbb{P}(A \le t) = s^2$ so that we recover from (51) the trivial inequality $\mathbb{P}(A^1 \le t) \ge \mathbb{P}(A \le t)$ as $A \ge A^1$.

PROOF OF PROPOSITION 7.6. Applying Theorem 4.7, we get

$$\begin{aligned} &\mathbb{E}[\mathrm{e}^{-\lambda Z} \mathbf{1}_{\{A^1 \le t\}}] \\ &\quad = \int_\lambda^\infty d\eta\, \mathbb{E}\big[Z \mathrm{e}^{-\eta Z} \mathbf{1}_{\{A^1 \le t\}}\big] \\ &\quad = 2(\mathrm{e}^{2\theta\beta t} - 1)^2 \left(\frac{1}{(\mathrm{e}^{2\theta\beta t} - 1)} \frac{2\theta}{2\theta + \lambda} - \log\left(1 + \frac{1}{(\mathrm{e}^{2\theta\beta t} - 1)} \frac{2\theta}{2\theta + \lambda}\right)\right). \end{aligned} \tag{52}$$

In particular, the distribution of A^1 is given by

$$\mathbb{P}(A^1 \le t) = 2(\mathrm{e}^{2\theta\beta t} - 1)^2 \left(\frac{1}{(\mathrm{e}^{2\theta\beta t} - 1)} - \log\left(1 + \frac{1}{(\mathrm{e}^{2\theta\beta t} - 1)}\right)\right).$$

Applying inverse Laplace transforms to (52) and using the density of Z given in (50), we get that the conditional law of A^1 given Z,

$$\mathbb{P}(A^1 \le t|Z) = \frac{2(\mathrm{e}^{2\theta\beta t}-1)^2}{(2\theta)^2 Z}\left(\frac{2\theta}{\mathrm{e}^{2\theta\beta t}-1} + \frac{\mathrm{e}^{-2\theta Z/(\mathrm{e}^{2\theta\beta t}-1)}-1}{Z}\right),$$

which implies that

$$\mathbb{P}\big(2\theta Z/(\mathrm{e}^{2\theta\beta A^1}-1) > x\big) = \frac{2}{x} - \frac{2}{x^2}(1-\mathrm{e}^{-x}). \qquad \square$$

7.4. *Fluctuations for the renormalized number of ancestors*. Finally, we complete corollary 6.5 by giving the fluctuations for the renormalized number of ancestors (to be compared with Theorem 6.7 in the stable case, recall that $Z_0 = Z$).

THEOREM 7.8. *Assume ψ is given by* (45). *Then the following convergence holds in distribution*:

$$\left(Z_{-s}, \sqrt{c(s)}\left(Z - \frac{M_s}{c(s)}\right), \sqrt{c(s)}(Z - Z_{-s})\right) \xrightarrow[s\downarrow 0+]{(d)} \left(Z, \frac{B_Z}{\sqrt{2}}, \frac{B_Z + W_Z}{\sqrt{2}}\right),$$

where $(B_t, t \ge 0)$ and $(W_t, t \ge 0)$ are two independent standard Brownian motions indepedent of Z. In particular, the following convergences hold in distribution:

$$\sqrt{c(s)\mathbb{E}[Z]}\left(\frac{M_s}{c(s)} - Z\right) \xrightarrow[s\downarrow 0+]{(d)} (Z - Z')$$

and

$$\sqrt{c(s)\mathbb{E}[Z]}(Z_{-s} - Z) \xrightarrow[s\downarrow 0+]{(d)} \sqrt{2}(Z - Z'),$$

where Z' is distributed as Z and independent of Z.

PROOF. We follow the proof of Theorem 6.7, with $h(s) = \sqrt{c(s)}$, up to formula (43). Then notice that $h^2(s)/c(s) = 1$ [instead of $o(1)$ in the proof of Theorem 6.7]. We have, for $r > 0$,

$$u\big(r\sqrt{c(s)}, s\big) = r\sqrt{c(s)}\left(1 - \frac{r}{\sqrt{c(s)}} + o\big(1/\sqrt{c(s)}\big)\right).$$

So we have for Δ_s defined by (42) the following approximation:

$$\begin{aligned}\Delta_s &= -\left(\frac{\eta}{\sqrt{c(s)}} + \frac{\eta^2}{2c(s)} + o\big(1/c(s)\big)\right)c(s) - \lambda\sqrt{c(s)} \\ &\quad + \left(1 + \frac{\eta}{\sqrt{c(s)}} + o\big(1/\sqrt{(c(s))}\big)\right)(\lambda+\eta)\sqrt{c(s)}\left(1 - \frac{\lambda+\eta}{\sqrt{c(s)}} + o\big(1/\sqrt{c(s)}\big)\right) \\ &= -\left(\frac{\eta^2}{2} + \lambda^2 + \lambda\eta\right) + o(1).\end{aligned}$$

We deduce that for any $\eta \geq 0$, $\lambda \geq 0$ and $b > (\frac{\eta^2}{2} + \lambda^2 + \lambda\eta)$, we have, for s small enough,

$$\begin{aligned}
&\mathbb{E}\big[\mathrm{e}^{-bZ_{-s}}\mathrm{e}^{-\eta\sqrt{c(s)}(Z_0-c(s)^{-1}M_s)-\lambda\sqrt{c(s)}(Z_0-Z_{-s})}\big]\\
&\quad=\mathbb{E}\big[\mathrm{e}^{-bZ_{-s}}\mathrm{e}^{(\eta^2/2+\lambda^2+\lambda\eta+o(1))Z_{-s}}\big]+o(1)\\
&\quad\underset{s\downarrow 0+}{\longrightarrow}\mathbb{E}\big[\mathrm{e}^{-bZ_0}\mathrm{e}^{(\eta^2/2+\lambda^2+\lambda\eta)Z_0}\big]\\
&\quad=\mathbb{E}\big[\mathrm{e}^{-bZ_0}\mathrm{e}^{-(\eta/\sqrt{2})B_{Z_0}-(\lambda/\sqrt{2})(B_{Z_0}+W_{Z_0})}\big].
\end{aligned}$$

An easy adaptation of [42] to multidimensional Laplace transform yields the first part of the theorem. Then notice that B_{Z_0} is distributed as $\sqrt{2\theta}(Z_0 - Z_0')$ to conclude. $\square$

Acknowledgment. The first author would like to thank the CMMSC of NCTU in Taiwan for its support.

REFERENCES

[1] Abraham, R. and Delmas, J. F. (2008). A continuum-tree-valued Markov process. *Ann. Probab.* **40** 1167–1211.

[2] Abraham, R. and Delmas, J.-F. (2009). Williams' decomposition of the Lévy continuum random tree and simultaneous extinction probability for populations with neutral mutations. *Stochastic Process. Appl.* **119** 1124–1143. MR2508567

[3] Aldous, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (*Durham*, 1990). *London Mathematical Society Lecture Note Series* **167** 23–70. Cambridge Univ. Press, Cambridge. MR1166406

[4] Aldous, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[5] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes. Die Grundlehren der Mathematischen Wissenschaften* **196** Springer, New York. MR0373040

[6] Berestycki, J., Berestycki, N. and Limic, V. (2010). The Λ-coalescent speed of coming down from infinity. *Ann. Probab.* **38** 207–233. MR2599198

[7] Berestycki, J., Kyprianou, A. E. and Murillo, A. (2009). The prolific backbone for supercritical superdiffusions. Available at ArXiv:0912.4736.

[8] Bertoin, J. (1996). *Lévy Processes. Cambridge Tracts in Mathematics* **121**. Cambridge Univ. Press, Cambridge. MR1406564

[9] Bertoin, J. and Le Gall, J.-F. (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126** 261–288. MR1990057

[10] Bertoin, J. and Le Gall, J.-F. (2005). Stochastic flows associated to coalescent processes. II. Stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 307–333. MR2139022

[11] Bertoin, J. and Le Gall, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* **50** 147–181 (electronic). MR2247827

[12] Birkner, M., Blath, J., Capaldo, M., Etheridge, A., Möhle, M., Schweinsberg, J. and Wakolbinger, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10** 303–325 (electronic). MR2120246

[13] Dawson, D. A. and Perkins, E. A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93** iv+179. MR1079034

[14] DONNELLY, P. and KURTZ, T. G. (1996). A countable representation of the Fleming–Viot measure-valued diffusion. *Ann. Probab.* **24** 698–742. MR1404525

[15] DONNELLY, P. and KURTZ, T. G. (1999). Particle representations for measure-valued population models. *Ann. Probab.* **27** 166–205. MR1681126

[16] DUQUESNE, T. and LE GALL, J.-F. (2002). *Random Trees, Lévy Processes and Spatial Branching Processes*. *Astérisque* **281**. MR1954248

[17] DUQUESNE, T. and LE GALL, J.-F. (2005). Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** 553–603. MR2147221

[18] ETHERIDGE, A. and MARCH, P. (1991). A note on superprocesses. *Probab. Theory Related Fields* **89** 141–147. MR1110534

[19] ETHERIDGE, A. M. and WILLIAMS, D. R. E. (2003). A decomposition of the $(1+\beta)$-superprocess conditioned on survival. *Proc. Roy. Soc. Edinburgh Sect. A* **133** 829–847. MR2006204

[20] EVANS, S. N. (1993). Two representations of a conditioned superprocess. *Proc. Roy. Soc. Edinburgh Sect. A* **123** 959–971. MR1249698

[21] EVANS, S. N. and PERKINS, E. (1990). Measure-valued Markov branching processes conditioned on nonextinction. *Israel J. Math.* **71** 329–337. MR1088825

[22] EVANS, S. N. and RALPH, P. L. (2010). Dynamics of the time to the most recent common ancestor in a large branching population. *Ann. Appl. Probab.* **20** 1–25. MR2582640

[23] FISHER, R. A. (1930). *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

[24] FLEMING, W. H. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843. MR0542340

[25] GALTON, F. and WATSON, H. W. (1874). On the probability of the extinction of families. *J. Roy. Anthropol. Inst.* **4** 138–144.

[26] GREY, D. R. (1974). Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Probab.* **11** 669–677. MR0408016

[27] JAGERS, P. and SAGITOV, S. (2004). Convergence to the coalescent in populations of substantially varying size. *J. Appl. Probab.* **41** 368–378. MR2052578

[28] JIŘINA, M. (1958). Stochastic branching processes with continuous state space. *Czechoslovak Math. J.* **83** 292–313. MR0101554

[29] KAJ, I. and KRONE, S. M. (2003). The coalescent process in a population with stochastically varying size. *J. Appl. Probab.* **40** 33–48. MR1953766

[30] KAWAZU, K. and WATANABE, S. (1971). Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen.* **16** 34–51. MR0290475

[31] KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR0671034

[32] LAMBERT, A. (2003). Coalescence times for the branching process. *Adv. in Appl. Probab.* **35** 1071–1089. MR2014270

[33] LAMBERT, A. (2007). Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.* **12** 420–446. MR2299923

[34] LE GALL, J.-F. (2006). Random real trees. *Ann. Fac. Sci. Toulouse Math.* (6) **15** 35–62. MR2225746

[35] LE GALL, J.-F. and LE JAN, Y. (1998). Branching processes in Lévy processes: Laplace functionals of snakes and superprocesses. *Ann. Probab.* **26** 1407–1432. MR1675019

[36] LE GALL, J.-F. and LE JAN, Y. (1998). Branching processes in Lévy processes: The exploration process. *Ann. Probab.* **26** 213–252. MR1617047

[37] LI, Z. (2011). *Measure-Valued Branching Markov Processes*. Springer, Heidelberg. MR2760602

[38] LI, Z.-H. (2000). Asymptotic behaviour of continuous time and state branching processes. *Austral. Math. Soc. Lect. Ser.* **68** 68–84. MR1727226

[39] LIMIC, V. (2010). On the speed of coming down from infinity for Ξ-coalescent processes. *Electron. J. Probab.* **15** 217–240. MR2594877
[40] MÖHLE, M. (2002). The coalescent in population models with time-inhomogeneous environment. *Stochastic Process. Appl.* **97** 199–227. MR1875333
[41] MORAN, P. A. P. (1958). Random processes in genetics. *Math. Proc. Cambridge Philos. Soc.* **54** 60–71. MR0127989
[42] MUKHERJEA, A., RAO, M. and SUEN, S. (2006). A note on moment generating functions. *Statist. Probab. Lett.* **76** 1185–1189. MR2270543
[43] PERKINS, E. A. (1992). Conditional Dawson–Watanabe processes and Fleming–Viot processes. In *Seminar on Stochastic Processes*, 1991 (*Los Angeles*, *CA*, 1991). *Progress in Probability* **29** 143–156. Birkhäuser, Boston, MA. MR1172149
[44] PINSKY, M. A. (1972). Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.* (*N.S.*) **78** 242–244. MR0295450
[45] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892
[46] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357
[47] ROELLY-COPPOLETTA, S. and ROUAULT, A. (1989). Processus de Dawson–Watanabe conditionné par le futur lointain. *C. R. Acad. Sci. Paris Sér. I Math.* **309** 867–872. MR1055211
[48] SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** 1116–1125. MR1742154
[49] WRIGHT, S. (1931). Evolution in Mendelian populations. *Genetics* **16** 97–159.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
1984 MATHEMATICS ROAD
VANCOUVER, B.C.
CANADA V6T 1Z2
E-MAIL: ytchen@math.ubc.ca

UNIVERSITÉ PARIS-EST
CERMICS, 6-8 AV. BLAISE PASCAL
CHAMPS-SUR-MARNE
77455 MARNE LA VALLÉE
FRANCE
E-MAIL: delmas@cermics.enpc.fr
URL: http://cermics.enpc.fr/~delmas/home.html