

COALESCENCE TIMES FOR THE BIENAYMÉ–GALTON–WATSON PROCESS

V. LE,* *Université de Provence*

Abstract

We investigate the distribution of the coalescence time (most recent common ancestor) for two individuals picked at random (uniformly) in the current generation of a continuous-time Bienaymé–Galton–Watson process founded t units of time ago. We also obtain limiting distributions as $t \rightarrow \infty$ in the subcritical case. We extend our results for two individuals to the joint distribution of coalescence times for any finite number of individuals sampled in the current generation.

Keywords: Bienaymé–Galton–Watson process; coalescence; discrete-state branching process; quasistationary distribution

2010 Mathematics Subject Classification: Primary 60J27

Secondary 60J80

1. Introduction

Random trees are mathematical objects that play an important role in many areas of mathematics and other sciences. One of the most celebrated random trees is the Bienaymé–Galton–Watson (BGW) tree, where the offspring of each vertex of the tree are independent and identically distributed (i.i.d.) random integers. The BGW tree plays a fundamental role in both the theory and application of stochastic processes; see, e.g. [1] and [12].

One interesting and important approach to random trees is coalescence. Lambert [7] investigated the distribution of the coalescence time for two individuals picked at random (uniformly) in the current generation of a BGW process in the discrete setting. The purpose of this note is to extend Lambert’s results to the case of the continuous-time BGW process. The basic idea is the same as that used in Lambert’s paper, but we need some other techniques. The continuous-time BGW process is started with x individuals at time 0. Its law is denoted by \mathbb{P}_x , with $\mathbb{P}_x^{(t)}$ indicating that the current time is time t . If at the current time there are at least two individuals, we uniformly choose two individuals, without replacement. We then compute the distribution of their coalescence time T (if at the current time there are less than two individuals, T is set to ∞). In the subcritical case, the law \mathbb{P}^{qs} , denoting the limit of the distributions $\mathbb{P}_x^{(t)}(\cdot \mid T < \infty)$ as $t \rightarrow \infty$, does not depend on x and is called the quasistationary distribution. In Section 3 we specify the law of T under \mathbb{P}^{qs} . In Section 4 we extend our results to multivariate coalescence when n individuals are sampled at the current time.

We do not recall Lambert’s results in this paper, but refer the reader to [7] for a comparison of the results in the discrete- and continuous-time cases. We also refer the reader to the interesting and closely related papers [5], [9], [10], [13], and [14].

Received 29 June 2012; revision received 19 March 2013.

* Postal address: Laboratoire d’Analyse Topologie et Probabilités (LATP/UMR 7353), Université d’Aix-Marseille, 39 rue F. Joliot-Curie, F-13453 Marseille cedex 13, France. Email address: levi121286@gmail.com

2. Distribution of the coalescence time

Let \mathbb{N} be the set of all natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$. We consider a continuous-time, \mathbb{N} -valued branching process $Z = \{Z_t, t \geq 0\}$, where t denotes time. Such a process is a BGW process in which to each individual is attached a random vector describing its lifetime and its number of offspring. We assume that these random vectors are i.i.d. The rate of reproduction is governed by a finite measure μ on \mathbb{N} , satisfying $\mu(1) = 0$. More precisely, each individual lives for an exponential time with parameter $\mu(\mathbb{N})$ and is replaced by a random number of children according to the probability $\mu(\mathbb{N})^{-1}\mu$. Hence, the dynamics of the continuous-time Markov process Z are entirely characterized by the measure μ . For $x \in \mathbb{N}$, denote by \mathbb{P}_x the law of Z when $Z_0 = x$.

Proposition 2.1. ([1, Chapter III, p. 106].) *The generating function of the process Z is given by*

$$\mathbb{E}_x(s^{Z_t}) = \psi_t(s)^x, \quad s \in [0, 1], x \in \mathbb{N},$$

where

$$\frac{\partial \psi_t(s)}{\partial t} = \Phi(\psi_t(s)), \quad \psi_0(s) = s,$$

and the function Φ is defined by

$$\Phi(s) = \sum_{n=0}^{\infty} (s^n - s)\mu(n), \quad s \in [0, 1].$$

The continuous-time BGW process Z is called immortal if $\mu(0) = 0$. In this paper we always assume that $\mu(0) > 0$. Let $\eta := \inf\{u > 0 : \Phi(u) = 0\}$. Since $\Phi(0) = \mu(0) > 0$, we have $\eta > 0$. Set

$$F(t) := \int_0^t \frac{du}{\Phi(u)}, \quad t < \eta.$$

Then the mapping $F: (0, \eta) \rightarrow (0, \infty)$ is bijective. Let φ be its inverse mapping. Moreover, $t \mapsto \psi_t(s)$ is the unique nonnegative solution of the integral equation

$$v(t) - \int_0^t \Phi(v(u)) du = s, \quad s \in [0, 1], t \geq 0,$$

so that

$$\int_s^{\psi_t(s)} \frac{dv}{\Phi(v)} = t, \quad s \in [0, 1], s < \eta, t \geq 0.$$

Hence,

$$\psi_t(s) = \varphi(t + F(s)), \quad s \in [0, 1], s < \eta, t \geq 0.$$

Note that the branching property implies that $\psi_{t_1+t_2} = \psi_{t_1} \circ \psi_{t_2}$.

Now, assume that the current generation is generation t , $t > 0$. We consider two individuals, σ_1 and σ_2 , at the present time, and ask when they coalesce, that is, how much time has elapsed since their common ancestor was present. In a more rigorous way, for $0 < u \leq t$, denote by $\tau_u(\sigma_i)$ the (unique) parent of σ_i at time $t - u$, $i = 1, 2$. The coalescence time $T(\sigma_1, \sigma_2)$ of σ_1, σ_2 is uniquely determined by

$$T(\sigma_1, \sigma_2) := \inf\{u : 0 < u \leq t, \tau_u(\sigma_1) = \tau_u(\sigma_2)\},$$

with the convention $\inf \emptyset = \infty$. We denote by T the coalescence time of two individuals picked at random (uniformly) among the individuals present in the current generation. If the current generation contains less than two individuals, T is set to ∞ .

With the notation $\mathbb{P}^{(t)}$ indicating that t is the current time, the distribution of T is given in the following statement.

Theorem 2.1. For any $0 < t_1 \leq t_2 \leq t$, $y \geq 1$, and $y \in \mathbb{N}$,

$$\mathbb{E}^{(t)}(Z_t(Z_t - 1)s^{Z_t-2}, T \leq t_1 \mid Z_{t-t_2} = y) = y\psi'_{t_2}(s)\psi_{t_2}(s)^{y-1} \frac{\psi''_{t_1}(s)}{\psi'_{t_1}(s)}, \quad s \in [0, 1).$$

The previous probability generating function can be inverted as follows. For any $p \geq 2$,

$$\begin{aligned} & \frac{\mathbb{P}^{(t)}(Z_t = p, T \in dt_1 \mid Z_{t-t_2} = y)}{dt_1} \\ &= y \sum_{n \geq 2} n\mu(n)\mathbb{E}\left(\frac{Z_{t_2}^{(1)}(1)Z_{t_1}^{(2)}(n-1)}{p(p-1)}, Z_{t_2}^{(0)}(y-1) + Z_{t_2}^{(1)}(1) + Z_{t_1}^{(2)}(n-1) = p\right), \end{aligned}$$

where $Z^{(0)}$, $Z^{(1)}$, and $Z^{(2)}$ are i.i.d. branching processes distributed as Z , and the notation $Z_{t_2}^{(0)}(y-1)$ denotes the value taken by $Z^{(0)}$ at time t_2 when started at $y-1$.

Remark 2.1. When $t_2 = t_1$, the above equation can be interpreted as follows. The population at time t (of size p) is divided into three groups. An individual is marked in generation $t - t_1$ (y possible choices) as a candidate for the common ancestor of two random individuals in generation t on $\{T \in dt_1\}$. The first group consists of the descendants at the current time of the $y - 1$ remaining individuals. On $\{T \in dt_1\}$ the marked individual must be immediately replaced by n , $n \geq 2$, offspring. Then an individual is marked among the n possible offspring of the previously marked ancestor. The descendants of this individual form the second group, and the descendants of the $n - 1$ remaining individuals form the third group. On $\{T \in dt_1\}$, one of the two individuals sampled must be in the second group, and the other in the third group.

Proof of Theorem 2.1. To obtain the first equation, we use the same argument as that used in the proof of Theorem 1 of [7]. The second equation is equivalent to

$$\begin{aligned} & \frac{\mathbb{E}^{(t)}(Z_t(Z_t - 1)s^{Z_t-2}, T \in dt_1 \mid Z_{t-t_2} = y)}{dt_1} \\ &= y \sum_{n \geq 2} n\mu(n)\mathbb{E}(Z_{t_2}^{(1)}(1)Z_{t_1}^{(2)}(n-1)s^{Z_{t_2}^{(0)}(y-1)+Z_{t_2}^{(1)}(1)+Z_{t_1}^{(2)}(n-1)-2}) \end{aligned} \tag{2.1}$$

for all $s \in (0, 1)$. Using the first result of the theorem, the left-hand side of (2.1) equals

$$y\psi'_{t_2}(s)\psi_{t_2}(s)^{y-1} \frac{\partial}{\partial t_1} \left(\frac{\psi''_{t_1}(s)}{\psi'_{t_1}(s)} \right).$$

From the Proposition 1 we have

$$\begin{aligned} \frac{\partial \psi_{t_1}(s)}{\partial t_1} &= \Phi(\psi_{t_1}(s)), \\ \frac{\partial \psi'_{t_1}(s)}{\partial t_1} &= \Phi'(\psi_{t_1}(s))\psi'_{t_1}(s), \\ \frac{\partial \psi''_{t_1}(s)}{\partial t_1} &= \Phi''(\psi_{t_1}(s))\psi'_{t_1}(s)^2 + \Phi'(\psi_{t_1}(s))\psi''_{t_1}(s), \end{aligned}$$

so that

$$\frac{\partial}{\partial t_1} \left(\frac{\psi''_{t_1}(s)}{\psi'_{t_1}(s)} \right) = \frac{\psi'_{t_1}(s) \partial \psi''_{t_1}(s) / \partial t_1 - \psi''_{t_1}(s) \partial \psi'_{t_1}(s) / \partial t_1}{\psi'_{t_1}(s)^2} = \Phi''(\psi_{t_1}(s)) \psi'_{t_1}(s).$$

Then

$$\begin{aligned} & \frac{\mathbb{E}^{(t)}(Z_t(Z_t - 1)s^{Z_t-2}, T \in dt_1 \mid Z_{t-t_2} = y)}{dt_1} \\ &= y \psi'_{t_2}(s) \psi_{t_2}(s)^{y-1} \Phi''(\psi_{t_1}(s)) \psi'_{t_1}(s) \\ &= y \psi'_{t_2}(s) \psi_{t_2}(s)^{y-1} \psi'_{t_1}(s) \sum_{n \geq 2} n(n-1) \mu(n) \psi_{t_1}(s)^{n-2}. \end{aligned}$$

Finally, the right-hand side of (2.1) equals

$$\begin{aligned} & y \sum_{n \geq 2} n \mu(n) \mathbb{E}(s^{Z_{t_2}^{(0)}(y-1)}) \mathbb{E}(Z_{t_2}^{(1)}(1)s^{Z_{t_2}^{(1)}(1)-1}) \mathbb{E}(Z_{t_1}^{(2)}(n-1)s^{Z_{t_1}^{(2)}(n-1)-1}) \\ &= y \sum_{n \geq 2} n \mu(n) \mathbb{E}_{y-1}(s^{Z_{t_2}}) \mathbb{E}_1(Z_{t_2} s^{Z_{t_2}-1}) \mathbb{E}_{n-1}(Z_{t_1} s^{Z_{t_1}-1}) \\ &= y \sum_{n \geq 2} n \mu(n) \psi_{t_2}(s)^{y-1} \psi'_{t_2}(s) (n-1) \psi_{t_1}(s)^{n-2} \psi'_{t_1}(s), \end{aligned}$$

completing the proof.

Corollary 2.1. *For any $0 < t_1 \leq t$,*

$$\mathbb{P}_x^{(t)}(T \leq t_1) = x \int_0^1 ds (1-s) \frac{\psi''_{t_1}(s)}{\psi'_{t_1}(s)} \psi'_{t_1}(s) \psi_t(s)^{x-1}.$$

In particular,

$$\begin{aligned} & \mathbb{P}_x^{(t)}(\text{at least two extant individuals, a random pair has no common ancestor}) \\ &= x(x-1) \int_0^1 ds (1-s) \psi'_t(s)^2 \psi_t(s)^{x-2}. \end{aligned}$$

Proof. See the proof of Corollary 1 of [7].

3. Quasistationary distribution

In this section we consider the limiting distribution of the coalescence time when the process is conditioned on $\{Z_t \geq 2\}$ and $t \rightarrow \infty$. Informally, this limit represents the situation where the genealogy was founded a long time ago and is still not extinct, with at least two descendants at the present time. We will need some results on quasistationary distributions for the continuous-time BGW process, which can be found in [1], [4], and [16]. The reader is referred to [15] and [3] for more general results on quasistationary distributions respectively obtained for continuous-time Markov chains and semi-Markov processes. We also refer the reader to [2], [8], and [11] for results on quasistationary distributions for population processes.

We consider the case $\psi'_1(1) = \mathbb{E}_1(Z_1) < 1$ (the subcritical case) when $\mathbb{E}_1(Z_1 \log(Z_1)) < \infty$. According to Theorem 6 of [16], there is a nonnegative sequence $(\alpha_k, k \geq 1)$ summing to 1 such that

$$\lim_{t \rightarrow \infty} \mathbb{P}_x(Z_t = j \mid Z_t > 0) = \alpha_j \quad \text{for all } x \in \mathbb{N}, j \geq 1. \tag{3.1}$$

The sequence $(\alpha_k, k \geq 1)$ is called the Yaglom limit of the process Z . Define

$$g(s) = \sum_{k \geq 1} \alpha_k s^k, \quad s \in [0, 1].$$

Then (3.1) yields

$$g(s) = \lim_{t \rightarrow \infty} \mathbb{E}_x(s^{Z_t} \mid Z_t > 0) = \lim_{t \rightarrow \infty} \frac{\psi_t(s) - \psi_t(0)}{1 - \psi_t(0)}, \quad s \in [0, 1].$$

We have the following result.

Proposition 3.1. *In the subcritical case when $\mathbb{E}_1(Z_1 \log(Z_1)) < \infty$, we have, for any $s \in [0, 1]$,*

$$\lim_{t \rightarrow \infty} \mathbb{E}_x(Z_t s^{Z_t - 1} \mid Z_t > 0) = g'(s) \leq g'(1) < \infty. \tag{3.2}$$

For a proof of Proposition 3.1, see [1, Chapter IV, p. 170]. Under the more restrictive hypothesis $\mathbb{E}_1(Z_1^2) < \infty$, we use the following two lemmas to give a very elementary and interesting proof of (3.2).

Lemma 3.1. *For $t \geq 0$, let $\varepsilon_t(s)$ be the function defined by*

$$\frac{1 - \psi_t(s)}{1 - s} = \psi'_t(1) - \varepsilon_t(s), \quad s \in [0, 1). \tag{3.3}$$

Then $\varepsilon_t(s)$ is monotone decreasing and tends to 0 as s tends to 1.

Proof. The proof follows from the fact that, for each t , $\psi_t(s)$ is increasing, convex, and $\psi_t(1) = 1$.

Equality (3.3) is equivalent to

$$\frac{1 - \psi_t(s)}{(1 - s)\psi'_t(1)} = 1 - \frac{\varepsilon_t(s)}{\psi'_t(1)}. \tag{3.4}$$

Replacing s by $\psi_h(s)$ in (3.4) yields

$$\frac{1 - \psi_t(\psi_h(s))}{(1 - \psi_h(s))\psi'_t(1)} = 1 - \frac{\varepsilon_t(\psi_h(s))}{\psi'_t(1)} \leq 1, \quad t, h > 0.$$

Note that $\psi_{t+h}(s) = \psi_t(\psi_h(s))$ and $\psi'_{t+h}(1) = \psi'_t(1)\psi'_h(1)$. Therefore,

$$\frac{1 - \psi_{t+h}(s)}{(1 - s)\psi'_{t+h}(1)} = \frac{1 - \psi_t(\psi_h(s))}{(1 - \psi_h(s))\psi'_t(1)} \frac{1 - \psi_h(s)}{(1 - s)\psi'_h(1)} \leq \frac{1 - \psi_h(s)}{(1 - s)\psi'_h(1)}, \quad t, h > 0.$$

This implies that the sequence $(1 - \psi_t(s))/(1 - s)\psi'_t(1)$ is monotone decreasing in t and, thus, converges to a function $\chi(s)$. Letting $s = 0$ we have

$$\chi(0) = \lim_{t \rightarrow \infty} \frac{\mathbb{P}_1(Z_t > 0)}{\psi'_t(1)} \geq 0.$$

Lemma 3.2. *We have $\chi(0)$ is positive and for all $x \in \mathbb{N}$,*

$$\lim_{t \rightarrow \infty} \mathbb{E}_x(Z_t \mid Z_t > 0) = g'(1) = \frac{1}{\chi(0)}.$$

Proof. We will follow the same arguments as used in the proof of Lemma 2 of [6]. Note that

$$\chi(0) = \lim_{t \rightarrow \infty} \frac{1 - \psi_t(0)}{\psi_t'(1)} = \lim_{n \rightarrow \infty, n \in \mathbb{N}} \frac{1 - \psi_n(0)}{\psi_n'(1)} = \lim_{n \rightarrow \infty} \prod_{k=0}^{n-1} \left[1 - \frac{\varepsilon_1(\psi_k(0))}{\psi_1'(1)} \right].$$

Hence, it follows that $\chi(0) > 0$ if and only if the series $\sum_{k=0}^{\infty} \varepsilon_1(\psi_k(0))$ converges. Since $\varepsilon_t(s) \geq 0$, we obtain

$$\frac{1 - \psi_t(s)}{1 - s} \leq \psi_t'(1), \quad t \geq 0, s \in [0, 1).$$

Letting $s = 0$ we obtain

$$\begin{aligned} \psi_t(0) &\geq 1 - \psi_t'(1), \quad t \geq 0, \\ \varepsilon_1(\psi_k(0)) &\leq \varepsilon_1(1 - \psi_k'(1)), \quad k \geq 0. \end{aligned} \tag{3.5}$$

On the other hand, since $\mathbb{E}_1(Z_1^2) < \infty$ implies that $\psi_1''(1) < \infty$, there exists a constant $C > 0$ such that

$$\varepsilon_1(s) < C(1 - s), \quad s \in [0, 1). \tag{3.6}$$

From (3.5) and (3.6), we deduce that the series $\sum_{k=0}^{\infty} \varepsilon_1(\psi_k(0))$ converges, so $\chi(0) > 0$. This implies that $\psi_t(0) \rightarrow 1$ as $t \rightarrow \infty$. Therefore,

$$\begin{aligned} g(\psi_t(0)) &= \lim_{h \rightarrow \infty} \frac{\psi_{t+h}(0) - \psi_h(0)}{1 - \psi_h(0)} \\ &= \lim_{h \rightarrow \infty} \frac{-(1 - \psi_{t+h}(0)) + (1 - \psi_h(0))}{1 - \psi_h(0)} \\ &= \frac{-\psi'_{t+h}(1) + \psi'_h(1)}{\psi'_h(1)} \\ &= -\psi'_t(1) + 1. \end{aligned}$$

Thus,

$$g'(1) = \lim_{t \rightarrow \infty} \frac{g(\psi_t(0)) - 1}{\psi_t(0) - 1} = \lim_{t \rightarrow \infty} \frac{-\psi'_t(1)}{\psi_t(0) - 1} = \frac{1}{\chi(0)}.$$

We have the following theorem, where \tilde{Z} denotes the limiting value of Z_t conditioned on $\{Z_t \geq 2\}$ as $t \rightarrow \infty$.

Theorem 3.1. *In the subcritical case when $\mathbb{E}_1(Z_1 \log(Z_1)) < \infty$, the quasistationary distribution \mathbb{P}^{qs} of T and \tilde{Z} is defined by*

$$\mathbb{P}^{qs}(\tilde{Z} = p, T \in dh) = \lim_{t \rightarrow \infty} \mathbb{P}_x^{(t)}(Z_t = p, T \in dh \mid Z_t \geq 2), \quad p \geq 2, h > 0.$$

Then \mathbb{P}^{qs} defines a probability distribution which does not depend on x and satisfies

$$\mathbb{E}^{qs}(\tilde{Z}(\tilde{Z} - 1)s^{\tilde{Z}-2}, T \leq h) = \frac{g'(s)}{1 - g'(0)} \frac{\psi_h''(s)}{\psi_h'(s)}.$$

In particular,

$$\mathbb{P}^{qs}(T \leq h) = \frac{1}{1 - g'(0)} \int_0^1 ds(1 - s) \frac{\psi_h''(s)}{\psi_h'(s)} g'(s).$$

Proof. See the proof of Theorem 2 of [7].

4. Multivariate coalescence

Assume that the current generation contains at least $n + 1$, $n \geq 1$, individuals. We now present the distribution of the coalescence times when $n + 1$ individuals are sampled uniformly and independently at the current time t . For $k = 1, 2, \dots, n$, we denote by T_k the coalescence time of the first individual and the $(k + 1)$ th individual, and by T_k^* the k th coalescence time.

Theorem 4.1. *For any $0 < t_1 < t_2 < \dots < t_n \leq t$, the joint distribution of coalescence times T_k is given by*

$$\frac{\mathbb{E}_x^{(t)}(Z_t(Z_t - 1) \cdots (Z_t - n)s^{Z_t - n - 1}, T_1 \in dt_1, \dots, T_n \in dt_n)}{dt_1 \cdots dt_n} = x\psi'_t(s)\psi_t(s)^{x-1} \prod_{i=1}^n \psi'_{t_i}(s) \left[\sum_{k \geq 2} k(k-1)\mu(k)\psi_{t_i}(s)^{k-2} \right], \quad s \in [0, 1).$$

Proof. We proceed by induction since the formula holds when $n = 1$ by Theorem 2.1. We first condition on $\{Z_{t-t_n} = y\}$. Applying the second formula of Theorem 2.1 to the last coalescence time T_n yields

$$\frac{\mathbb{P}^{(t)}(Z_t = p, T_1 \in dt_1, \dots, T_n \in dt_n \mid Z_{t-t_n} = y)}{dt_n} = y \sum_{k \geq 2} k\mu(k)\mathbb{E} \left(\frac{Z_{t_n}^{(1)}(1)Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + 1)}{p(p-1) \cdots (p-n)}, \right. \\ \left. Z_{t_n}^{(0)}(y-1) + Z_{t_n}^{(1)}(1) + Z_{t_n}^{(2)}(k-1) = p, T_i \in dt_i, i \leq n-1 \right),$$

where the interpretation is as for $n = 1$ (see Remark 2.1): y corresponds to the choice of the common ancestor of all individuals in generation $t - t_n$, k is the number of offspring this ancestor has instantaneously at time $t - T_n$ and corresponds to the choice of the ancestor of the last individual within this offspring. The n remaining individuals have to be found in the descendants of the $k - 1$ remaining offspring. Then

$$\frac{\mathbb{E}^{(t)}(Z_t(Z_t - 1) \cdots (Z_t - n)s^{Z_t - n - 1}, T_1 \in dt_1, \dots, T_n \in dt_n \mid Z_{t-t_n} = y)}{dt_n} = y \sum_{k \geq 2} k\mu(k)\mathbb{E}(Z_{t_n}^{(1)}(1)Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + 1) \\ \times s^{Z_{t_n}^{(0)}(y-1) + Z_{t_n}^{(1)}(1) + Z_{t_n}^{(2)}(k-1) - n - 1}, T_i \in dt_i, i \leq n-1) \\ = y \sum_{k \geq 2} k\mu(k)\mathbb{E}(s^{Z_{t_n}^{(0)}(y-1)})\mathbb{E}(Z_{t_n}^{(1)}(1)s^{Z_{t_n}^{(1)}(1)-1}) \\ \times \mathbb{E}(Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + 1)s^{Z_{t_n}^{(2)}(k-1) - n}, T_i \in dt_i, i \leq n-1) \\ = y\psi_{t_n}(s)^{y-1}\psi'_{t_n}(s) \sum_{k \geq 2} k\mu(k)\mathbb{E}(Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + 1) \\ \times s^{Z_{t_n}^{(2)}(k-1) - n}, T_i \in dt_i, i \leq n-1).$$

By the induction hypothesis, the last expression equals

$$\begin{aligned}
 & y\psi_{t_n}(s)^{y-1}\psi'_{t_n}(s)\sum_{k\geq 2}k\mu(k)(k-1)\psi'_{t_n}(s)\psi_{t_n}(s)^{k-2} \\
 & \times \prod_{i=1}^{n-1}\psi'_{t_i}(s)\left[\sum_{j\geq 2}j(j-1)\mu(j)\psi_{t_i}(s)^{j-2}\right]dt_1\cdots dt_{n-1} \\
 & = y\psi_{t_n}(s)^{y-1}\psi'_{t_n}(s)\prod_{i=1}^n\psi'_{t_i}(s)\left[\sum_{k\geq 2}k(k-1)\mu(k)\psi_{t_i}(s)^{k-2}\right]dt_1\cdots dt_{n-1}.
 \end{aligned}$$

Hence, the result follows by integrating with respect to the distribution of Z_{t-t_n} conditional on $\{Z_0 = x\}$.

Theorem 4.2. For any $0 < t_1 < t_2 < \cdots < t_n \leq t$, the joint distribution of coalescence times T_k^* is given by

$$\begin{aligned}
 & \frac{\mathbb{E}_x^{(t)}(Z_t(Z_t-1)\cdots(Z_t-n)s^{Z_t-n-1}, T_1^* \in dt_1, \dots, T_n^* \in dt_n)}{dt_1\cdots dt_n} \\
 & = \frac{n!(n+1)!}{2^n}x\psi'_t(s)\psi_t(s)^{x-1}\prod_{i=1}^n\psi'_{t_i}(s)\left[\sum_{k\geq 2}k(k-1)\mu(k)\psi_{t_i}(s)^{k-2}\right], \quad s \in [0, 1).
 \end{aligned}$$

Proof. The proof is similar to that of Theorem 4.1. We reason by induction since the formula holds when $n = 1$ by Theorem 2.1. We first condition on $\{Z_{t-t_n} = y\}$ and apply the second formula of Theorem 2.1 to the last coalescence time T_n^* , giving

$$\begin{aligned}
 & \frac{\mathbb{P}^{(t)}(Z_t = p, T_1^* \in dt_1, \dots, T_n^* \in dt_n \mid Z_{t-t_n} = y)}{dt_n} \\
 & = \frac{1}{2}y\sum_{k\geq 2}k\mu(k)\sum_{i=1}^n\binom{n+1}{i} \\
 & \times \sum_{1\leq j_1 < \cdots < j_{i-1}\leq n-1}\mathbb{E}\left(\frac{Z_{t_n}^{(1)}(1)\cdots Z_{t_n}^{(1)}(1-i+1)Z_{t_n}^{(2)}(k-1)\cdots Z_{t_n}^{(2)}(k-1-n+i)}{p(p-1)\cdots(p-n)}, \right. \\
 & \quad Z_{t_n}^{(0)}(y-1) + Z_{t_n}^{(1)}(1) + Z_{t_n}^{(2)}(k-1) = p, \\
 & \quad T_h^*(i) \in dt_h \text{ for } h \in \{j_1, \dots, j_{i-1}\} \text{ and} \\
 & \quad \left. T_h^*(n+1-i) \in dt_h \text{ for } h \notin \{j_1, \dots, j_{i-1}\}, h \leq n-1\right),
 \end{aligned}$$

where the interpretation is as follows: y corresponds to the choice of the common ancestor of all individuals in generation $t - t_n$, k is the number of offspring this ancestor has instantaneously at time $t - T_n^*$ and corresponds to the choice of the ancestor of the last i individuals within this offspring (there are $\binom{n+1}{i}$ possible choices for the last i individuals). The $n + 1 - i$ remaining individuals have to be found from the descendants of the $k - 1$ remaining offspring. For $m = 1, \dots, i - 1$, $T_{j_m}(i)$ is the m th coalescence time of the last i individuals, and, for $h \notin \{j_1, \dots, j_{i-1}\}$, $h \leq n - 1$, $T_h^*(n + 1 - i)$ is a coalescence time of the $n + 1 - i$ remaining

individuals. Note that we have to divide the expression by 2 because each sample has been counted twice. We then have

$$\begin{aligned}
 & \frac{\mathbb{E}^{(t)}(Z_t(Z_t - 1) \cdots (Z_t - n) s^{Z_t - n - 1}, T_1^* \in dt_1, \dots, T_n^* \in dt_n \mid Z_{t-t_n} = y)}{dt_n} \\
 &= \frac{1}{2} y \sum_{k \geq 2} k \mu(k) \sum_{i=1}^n \binom{n+1}{i} \\
 & \quad \times \sum_{1 \leq j_1 < \dots < j_{i-1} \leq n-1} \mathbb{E}(Z_{t_n}^{(1)}(1) \cdots (Z_{t_n}^{(1)}(1) - i + 1) Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + i) \\
 & \quad \quad \quad \times s^{Z_{t_n}^{(0)}(y-1) + Z_{t_n}^{(1)}(1) + Z_{t_n}^{(2)}(k-1) - n - 1}, \\
 & \quad \quad \quad T_h^*(i) \in dt_h \text{ for } h \in \{j_1, \dots, j_{i-1}\} \text{ and} \\
 & \quad \quad \quad T_h^*(n+1-i) \in dt_h \text{ for } h \notin \{j_1, \dots, j_{i-1}\}, h \leq n-1) \\
 &= \frac{1}{2} y \sum_{k \geq 2} k \mu(k) \sum_{i=1}^n \binom{n+1}{i} \\
 & \quad \times \sum_{1 \leq j_1 < \dots < j_{i-1} \leq n-1} \mathbb{E}(s^{Z_{t_n}^{(0)}(y-1)}) \mathbb{E}(Z_{t_n}^{(1)}(1) \cdots (Z_{t_n}^{(1)}(1) - i + 1) s^{Z_{t_n}^{(1)}(1) - i}, \\
 & \quad \quad \quad T_h^*(i) \in dt_h \text{ for } h \in \{j_1, \dots, j_{i-1}\}) \\
 & \quad \quad \times \mathbb{E}(Z_{t_n}^{(2)}(k-1) \cdots (Z_{t_n}^{(2)}(k-1) - n + i) s^{Z_{t_n}^{(2)}(k-1) - n + i - 1}, \\
 & \quad \quad \quad T_h^*(n+1-i) \in dt_h \text{ for } h \notin \{j_1, \dots, j_{i-1}\}, h \leq n-1).
 \end{aligned}$$

By the induction hypothesis, the last expression equals

$$\begin{aligned}
 & \frac{y}{2} \sum_{k \geq 2} k \mu(k) \sum_{i=1}^n \binom{n+1}{i} \\
 & \quad \times \sum_{1 \leq j_1 < \dots < j_{i-1} \leq n-1} \psi_{t_n}(s)^{y-1} \frac{(i-1)! i!}{2^{i-1}} \psi'_{t_n}(s) \\
 & \quad \quad \times \prod_{h \in \{j_1, \dots, j_{i-1}\}} \psi'_{t_h}(s) \left[\sum_{j \geq 2} j(j-1) \mu(j) \psi_{t_h}(s)^{j-2} \right] \\
 & \quad \quad \times \frac{(n-i)! (n-i+1)!}{2^{n-i}} (k-1) \psi'_{t_n}(s) \psi_{t_n}(s)^{k-2} \\
 & \quad \quad \times \prod_{1 \leq h \leq n-1, h \notin \{j_1, \dots, j_{i-1}\}} \psi'_{t_h}(s) \left[\sum_{j \geq 2} j(j-1) \mu(j) \psi_{t_h}(s)^{j-2} \right] dt_1 dt_2 \cdots dt_{n-1} \\
 &= \frac{y}{2} \sum_{k \geq 2} k \mu(k) \sum_{i=1}^n \binom{n+1}{i} \\
 & \quad \times \sum_{1 \leq j_1 < \dots < j_{i-1} \leq n-1} \frac{(i-1)! i! (n-i)! (n-i+1)!}{2^{n-1}} \psi'_{t_n}(s) \psi_{t_n}(s)^{y-1} \\
 & \quad \times (k-1) \psi'_{t_n}(s) \psi_{t_n}(s)^{k-2} \\
 & \quad \times \prod_{h=1}^{n-1} \psi'_{t_h}(s) \left[\sum_{j \geq 2} j(j-1) \mu(j) \psi_{t_h}(s)^{j-2} \right] dt_1 dt_2 \cdots dt_{n-1}
 \end{aligned}$$

$$\begin{aligned}
&= \frac{y}{2} \sum_{k \geq 2} k \mu(k) \sum_{i=1}^n \binom{n+1}{i} \binom{n-1}{i-1} \frac{(i-1)! i! (n-i)! (n-i+1)!}{2^{n-1}} \psi'_{t_n}(s) \psi_{t_n}(s)^{y-1} \\
&\quad \times (k-1) \psi'_{t_n}(s) \psi_{t_n}(s)^{k-2} \\
&\quad \times \prod_{h=1}^{n-1} \psi'_{t_h}(s) \left[\sum_{j \geq 2} j(j-1) \mu(j) \psi_{t_h}(s)^{j-2} \right] dt_1 dt_2 \cdots dt_{n-1} \\
&= \frac{n!(n+1)!}{2^n} y \psi'_{t_n}(s) \psi_{t_n}(s)^{y-1} \prod_{h=1}^n \psi'_{t_h}(s) \left[\sum_{j \geq 2} j(j-1) \mu(j) \psi_{t_h}(s)^{j-2} \right] dt_1 dt_2 \cdots dt_{n-1}.
\end{aligned}$$

Hence the result follows by integrating with respect to the distribution of Z_{t-t_n} conditional on $\{Z_0 = x\}$.

Acknowledgements

I am deeply grateful to my thesis advisor Professor Etienne Pardoux for useful suggestions and constant encouragement during the preparation of this work. I would also like to thank the anonymous referee for a careful reading of the paper, as well as many valuable comments.

References

- [1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [2] CATTIAUX, P. *et al.* (2009). Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Prob.* **37**, 1926–1969.
- [3] CHEONG, C. K. (1970). Quasi-stationary distributions in semi-Markov processes. *J. Appl. Prob.* **7**, 388–399, 788.
- [4] CHEONG, C. K. (1972). Quasi-stationary distributions for the continuous-time Galton–Watson process. *Bull. Soc. Math. Belg.* **24**, 343–350.
- [5] JAGERS, P., KLEBANER, F. AND SAGITOV, S. (2007). Markovian paths to extinction. *Adv. Appl. Prob.* **39**, 569–587.
- [6] JOFFE, A. (1967). On the Galton–Watson branching processes with mean less than one. *Ann. Math. Statist.* **38**, 264–266.
- [7] LAMBERT, A. (2003). Coalescence times for the branching process. *Adv. Appl. Prob.* **35**, 1071–1089.
- [8] LAMBERT, A. (2007). Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Prob.* **12**, 420–446.
- [9] LAMBERT, A. (2009). The allelic partition for coalescent point processes. *Markov Process. Relat. Fields* **15**, 359–386.
- [10] LAMBERT, A. AND POPOVIC, L. (2013). The coalescent point process of branching trees. *Ann. Appl. Prob.* **23**, 99–144.
- [11] MÉLÉARD, S. AND VILLEMONAIS, D. (2012). Quasi-stationary distributions for population processes. *Prob. Surveys* **9**, 340–410.
- [12] PARDOUX, E. (2014). *Probabilistic Models of Population Genetics*. In preparation.
- [13] PFAFFELHUBER, P. AND WAKOLBINGER, A. (2006). The process of most recent common ancestors in an evolving coalescent. *Stoch. Process. Appl.* **116**, 1836–1859.
- [14] PFAFFELHUBER, P., WAKOLBINGER, A. AND WEISSHAUPT, H. (2011). The tree length of an evolving coalescent. *Prob. Theory Relat. Fields* **151**, 529–557.
- [15] VERE-JONES, D. (1969). Some limit theorems for evanescent processes. *Austral. J. Statist.* **11**, 67–78.
- [16] ZOLOTAREV, V. M. (1957). More exact statements of several theorems in the theory of branching processes. *Theory Prob. Appl.* **2**, 245–253.