Statistics and Probability Letters 210 (2024) 110121

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

# Coalescence times for critical Galton–Watson processes with immigration[☆]

Rong-Li Liu[a,1,*], Yan-Xia Ren[b,2], Yingrui Wang[a]

[a] School of Mathematics and Statistics, Beijing Jiaotong University, Beijing, 100044, PR China
[b] LMAM School of Mathematical Sciences & Center for Statistical Science, Peking University, Beijing 100871, PR China

ARTICLE INFO

*MSC:*
primary 60J68
secondary 62E15
60F10
60J80

*Keywords:*
Critical Galton–Watson process
Immigration
Coalescence times

ABSTRACT

Let $X_n^I$ be the coalescence time of two particles picked at random from the $n$th generation of a critical Galton–Watson process with immigration, and let $A_n^I$ be the coalescence time of the whole population in the $n$th generation. In this paper, we study the limiting behaviors of $X_n^I$ and $A_n^I$ as $n \to \infty$.

## 1. Introduction and main results

Suppose $(Y_n)_{n\geq 0}$ is a Galton–Watson process with offspring distribution $(p_j)_{j\geq 0}$ and initial size $Y_0 = 1$. For $n \geq 1$, conditional on $\{Y_n \geq 2\}$, pick 2 distinct particles uniformly from the $n$th generation and trace their lines of descent backward in time. The common nodes in the two lines are called the common ancestors of the two particles. Let $X_n$ denote the generation of their most recent common ancestor, which is called *the pairwise coalescence time*. Next, for $n \geq 1$, conditional on $\{Y_n \geq 1\}$, we trace the lines of descent of all particles in generation $n$ backward in time. The common nodes in the $Y_n$ lines of descent are called the common ancestors of all the particles in generation $n$. Define *the total coalescence time* $A_n$ as the generation of the most recent common ancestor of all the particles in generation $n$. When $m := \sum_{n=0}^{\infty} jp_j = 1$ (critical case), $p_1 < 1$ and $\sigma^2 := \sum_{n=0}^{\infty} j^2 p_j - 1 < \infty$, Athreya (2012a) proved that for $u \in (0, 1)$,

$$\lim_{n\to\infty} P\left(\frac{X_n}{n} \geq u \middle| Y_n \geq 2\right) = E\left[\frac{\sum_{i=1}^{N_u} \eta_i^2}{(\sum_{i=1}^{N_u} \eta_i)^2}\right], \tag{1.1}$$

where $(\eta_i)_{i\geq 1}$ are independent and identically distributed exponential random variables with mean $\sigma^2/2$, and $N_u$ is independent of $(\eta_i)_{i\geq 1}$ and is a geometric random variable of parameter $1-u$ (i.e., $P(N_u = k) = (1-u)u^{k-1}, k \geq 1$). Athreya (2012a) also proved the

☆ The research of this project is supported by the National Key R&D Program of China (No. 2020YFA0712902).
* Corresponding author.
*E-mail addresses:* rlliu@bjtu.edu.cn (R.-L. Liu), yxren@math.pku.edu.cn (Y.-X. Ren), wyr2120873@gmail.com (Y. Wang).
1 The research of R. Liu is supported in part by NSFC, China (Grant No. 12271374).
2 The research of Y.-X. Ren is supported in part by NSFC, China (Grant Nos. 12071011 and 12231002) and The Fundamental Research Funds for the Central Universities, China, Peking University LMEQF, China.

https://doi.org/10.1016/j.spl.2024.110121
Received 18 December 2022; Received in revised form 24 June 2023; Accepted 22 March 2024
Available online 28 March 2024
0167-7152/

following conditional limit result:

$$\lim_{n\to\infty} P\left(\frac{A_n}{n} > u \Big| Y_n \geq 1\right) = 1 - u, \qquad \text{for } u \in (0,1).$$

The genealogy of branching processes has been widely studied. Athreya (2010, 2012b), Durrett (1978), Zubkov (1975) also investigated the distributional properties of the coalescence times for Galton–Watson processes. Kersting (2022) gave the genealogy structure of branching processes in random environment. Harris et al. (2020), Johnston (2019) and Le (2014) investigated the coalescent structure of continuous time Galton–Watson processes. Hong (2016) studied the corresponding results for multitype branching processes.

Suppose $(p_j)_{j\geq 0}$ and $(b_j)_{j\geq 0}$ are probability distributions on the set $\mathbb{N}$ of nonnegative integers. Let $(\xi_{n,i}; n \in \mathbb{N}, i \in \mathbb{N})$ be a doubly infinite family of independent random variables with common distribution $(p_j)_{j\geq 0}$, and let $(I_n)_{n\geq 0}$ be a sequence of independent random variables with common distribution $(b_j)_{j\geq 0}$ which are independent of $(\xi_{n,i}; n \in \mathbb{N}, i \in \mathbb{N})$ as well. Let $(Z_n)_{n\geq 0}$ be a Galton–Watson process with immigration (GWPI for short) defined by

$$Z_0 = I_0, \qquad Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i} + I_{n+1}, \qquad n = 0, 1, \ldots. \tag{1.2}$$

Here $Z_n$ is the population size in generation $n$, and $I_n$ is the number of immigrants in generation $n$. For each $1 \leq i \leq Z_n$, $\xi_{n,i}$ denotes the number of children of the $i$th particle in generation $n$. We assume that all the immigrants have different ancestors. Set $m = E\xi_{0,1} = \sum_{j=0}^{\infty} j p_j$. Then $(Z_n)_{n\geq 0}$ is called supercritical, critical or subcritical according to $m > 1, m = 1$ or $m < 1$, respectively. GWPI was first considered by Heathcote (1965). Recently, Wang, Li and Yao (Wang et al., 2019) found that the pairwise coalescence time $X_n$ for some supercritical GWPI converges in distribution to a $(0,\infty]$-valued random variable as $n \to \infty$.

In this paper, we consider the coalescence times for critical GWPI $(Z_n)_{n\geq 0}$. Unlike the case of a Galton–Watson process starting with one particle, two randomly picked distinct particles (all particles) from generation $n$ of a GWPI may not have a common ancestor. Conditional on $\{Z_n > 1\}$, we pick two distinct particles, say $\mathrm{v}_1$ and $\mathrm{v}_2$, uniformly from the $n$th generation and trace their lines of descent backward in time. Define the *pairwise coalescence time* for GWPI

$$X_n^I = \begin{cases} |\mathrm{v}|, & \text{if the most recent common ancestor of } \mathrm{v}_1 \text{ and } \mathrm{v}_2 \text{ is } \mathrm{v}, \\ \infty, & \text{otherwise,} \end{cases} \tag{1.3}$$

where $|\mathrm{v}|$ is the generation of $\mathrm{v}$. Note that even if $\mathrm{v}_1$ and $\mathrm{v}_2$ are descendants of two distinct particles immigrated to the system at the same time, we do not say they have a common ancestor. Similarly, conditional on $\{Z_n > 0\}$, define *the total coalescence time* for GWPI

$$A_n^I = \begin{cases} |\mathrm{v}|, & \text{if the most recent common ancestor of all particles alive at } n \text{ is } \mathrm{v}, \\ \infty, & \text{otherwise.} \end{cases} \tag{1.4}$$

We will study the asymptotic behaviors of the distribution of $X_n^I$ conditioned on $\{Z_n > 1\}$ and the distribution of $A_n^I$ conditioned on $\{Z_n > 0\}$. We will explore the effect of the immigrations on the coalescence times. Throughout this paper we suppose the following assumption holds.

**Assumption 1.** $0 < p_0 + p_1 < 1$, $m = 1$, $\sigma^2 = \sum_j (j^2 - 1) p_j < \infty$. $b_0 < 1$ and $\beta = \sum_j j b_j < \infty$.

We use $\langle g, \mu \rangle$ to denote the integral of a function $g$ with respect to a Radon measure $\mu$ whenever this integral makes sense.

**Theorem 1.1.** *Suppose Assumption 1 holds. Let $\gamma = 2\beta/\sigma^2$. Define*

$$\phi(j,\mu) = E\Big[\frac{\sum_{i=1}^{j} \omega_i^2 + \langle f^2, \mu\rangle}{(\sum_{i=1}^{j} \omega_i + \langle f, \mu\rangle)^2}\Big], \tag{1.5}$$

*where $f(r) = r, r > 0$, and $(\omega_i)_{i\geq 1}$ are independent exponential random variables with parameter $\frac{2}{\sigma^2}$.*

*(1) For $0 < u < 1$,*

$$\lim_{n\to\infty, k/n\to u} P\left(k \leq X_n^I < n \Big| Z_n > 1\right) = E\phi(N_u^I, W),$$

*where $N_u^I$ is a negative binomial random variable with law*

$$P(N_u^I = k) = \frac{(-\gamma)(-\gamma-1)\cdots(-\gamma-k+1)}{k!}(1-u)^{\gamma}(-u)^k, \qquad k = 0, 1, 2, \ldots, \tag{1.6}$$

*with the convention $\frac{(-\gamma)(-\gamma-1)\cdots(-\gamma-k+1)}{k!} = 1$ when $k = 0$, $W$ is a Poisson random measure on $(0,\infty)$ with intensity $\frac{\gamma}{r} e^{-\frac{2}{\sigma^2} r} dr$, and $N_u^I$ and $W$ are independent.*

*(2)*

$$\lim_{n\to\infty} P\left(X_n^I < \infty \Big| Z_n > 1\right) = E\left[\frac{\langle f^2, W\rangle}{\langle f, W\rangle^2}\right].$$

Note that $N_u$ in (1.1) for a critical Galton–Watson process only takes positive integer values, while $N_u^I$ in Theorem 1.1 can take value 0 with positive probability. In the special case $\gamma = 1$, the random number $N_u^I + 1$ and $N_u$ have the same distribution.

We conclude from Pakes (1971b, Theorem 3) (see Lemma 2.2) that $Z_n$ diverges to infinity in probability as $n \to \infty$. Our second result says that as $n \to \infty$, the probability that all the particles of generation $n$ have a common ancestor goes to 0.

**Theorem 1.2.** *Suppose Assumption 1 holds. Then*

$$\lim_{n\to\infty} P(A_n^I < \infty | Z_n > 0) = 0.$$

## 2. Some preliminary results

Recall that $(Y_n)_{n\geq 0}$ is a critical Galton–Watson process with offspring distribution $(p_j)_{j\geq 0}$ starting with $Y_0 = 1$. The following result was proved in Athreya and Ney (1972).

**Lemma 2.1.** *When $m = 1$, $p_1 < 1$, $\sigma^2 = \sum_j (j^2 - j)p_j < \infty$,*

$$\lim_{n\to\infty} nP(Y_n > 0) = \frac{2}{\sigma^2}, \tag{2.1}$$

*and for any $t > 0$,*

$$\lim_{n\to\infty} P\left(\frac{Y_n}{n} > t \middle| Y_n > 0\right) = e^{-\frac{2t}{\sigma^2}}. \tag{2.2}$$

The following result for critical GWPI is from Pakes (1971b, Theorem 3).

**Lemma 2.2.** *Suppose Assumption 1 holds. Put $\gamma = \frac{2\beta}{\sigma^2}$. Then, as $n \to \infty$, $\frac{Z_n}{n}$ converges in distribution to a Gamma random variable with parameters $(2/\sigma^2, \gamma)$, whose density function is*

$$h(t) = \frac{2}{\sigma^2 \Gamma(\gamma)} \left(\frac{2t}{\sigma^2}\right)^{\gamma-1} e^{-\frac{2t}{\sigma^2}}, \qquad t > 0. \tag{2.3}$$

The above lemma implies that $\lim_{n\to\infty} P(Z_n > 0) = 1$. The rate that $1 - P(Z_n > 0)$ converges to 0 was investigated in Pakes (1971a).

From the construction (1.2) of the GWPI $(Z_n)_{n\geq 0}$, for any $0 \leq k < n$, $Z_n$ can be rewritten as

$$Z_n = \sum_{i=1}^{Z_k} Y_{n,k,i} + \sum_{j=k+1}^{n} \sum_{l=1}^{I_j} Y_{n-j,l}^{(j)}, \tag{2.4}$$

where $Y_{n,k,i}, i = 1, 2, \ldots$, are independent and have the same distribution as $Y_{n-k}$, and for $0 \leq j \leq n$, $Y_{n-j,l}^{(j)}, l = 1, 2 \ldots$, are independent and have the same distribution as $Y_{n-j}$. Note that $Y_{n,k,i}$ represents the number of descendants in generation $n$ of the $i$th particle in generation $k$, and $Y_{n-j,l}^{(j)}$ represents the number of descendants in generation $n$ of the $l$th particle in the $I_j$ immigrants in generation $j$. For any non-negative integer $m$, set $(m)_2 = m(m-1)$. Notice that $(m)_2 = 0$ when $m = 0$ or 1. Starting from the representation (2.4), the distribution of the pairwise coalescence time $X_n^I$, given $\{Z_n > 1\}$, has the following expression.

**Lemma 2.3.** *For any $0 \leq k < n$,*

$$P\left(k \leq X_n^I < n \middle| Z_n > 1\right) = E\left[\frac{\sum_{i=1}^{Z_k} (Y_{n,k,i})_2 + \sum_{j=1+k}^{n-1} \sum_{l=1}^{I_j} (Y_{n-j,l}^{(j)})_2}{(Z_n)_2} \middle| Z_n > 1\right],$$

*with the convention that the second term in the numerator equals 0 when $k > n-2$. In particular,*

$$P(X_n^I < \infty | Z_n > 1) = E\left[\frac{\sum_{j=0}^{n-1} \sum_{l=1}^{I_j} (Y_{n-j,l}^{(j)})_2}{(Z_n)_2} \middle| Z_n > 1\right].$$

**Proof.** For $0 \leq k < n$, the event $\{k \leq X_n^I < n\}$ occurs if and only if either the two randomly picked particles from generation $n$ are both descendants of a particle in the $k$th generation, or they are both descendants of a particle immigrated into the system between generation $k+1$ and generation $n-1$. The number of choices of the two particles from the descendants of the $i$th particle in generation $k$ is $(Y_{n,k,i})_2$, and therefore the total number is $\sum_{i=1}^{Z_k} (Y_{n,k,i})_2$ with the convention that the sum is 0 if $Z_k = 0$. The number of choices of the two particles from the descendants of the $l$th particle immigrated into the system in generation $j$ for $k+1 \leq j < n$ is $(Y_{n-j,l}^{(j)})_2$, and the total number is $\sum_{j=k+1}^{n-1} \sum_{l=1}^{I_j} (Y_{n-j,l}^{(j)})_2$. Also, the total number of choices of the two particles from the $n$th generation is $(Z_n)_2$. Thus for any $n \geq 1$ and $0 \leq k < n$, conditional on $\{Z_n > 1\}$, the probability of $\{k \leq X_n^I < n\}$ is given by

$$P(k \leq X_n^I < n | Z_n > 1) = E\left[\frac{\sum_{i=1}^{Z_k} (Y_{n,k,i})_2 + \sum_{j=k+1}^{n-1} \sum_{l=1}^{I_j} (Y_{n-j,l}^{(j)})_2}{(Z_n)_2} \middle| Z_n > 1\right].$$

Since $Z_0 = I_0$, we have $Y_{n,0,i} = Y^{(0)}_{n,i}$, $i = 1, \dots, I_0$. Taking $k = 0$ in the above identity, we obtain

$$P(X_n^I < \infty | Z_n > 1) = P(X_n^I < n | Z_n > 1) = E\Big[\frac{\sum_{j=0}^{n-1}\sum_{l=1}^{I_j}(Y^{(j)}_{n-j,l})_2}{(Z_n)_2}\Big| Z_n > 1\Big]. \quad \square$$

Let $\mathcal{M}$ be the space of finite measures on $[0, \infty)$ equipped with the topology of weak convergence. Let $C_b[0,\infty)(C_b^+[0,\infty))$ be the space of bounded continuous (nonnegative bounded continuous) functions on $[0,\infty)$. Then for any $g \in C_b[0,\infty)$, the map $\pi_g : \mu \to \langle g, \mu\rangle$ on $\mathcal{M}$ is continuous. For random measures $\eta_n, \eta \in \mathcal{M}$, $n = 1, 2, \dots$, $\eta_n$ converges to $\eta$ in distribution as $n \to \infty$ is equivalent to $\langle g, \eta_n\rangle \xrightarrow{d} \langle g, \eta\rangle$ for all $g \in C_b^+[0,\infty)$. We refer the readers to Kallenberg (2017, p.109) for more details. Let $\mathcal{F}_k$ be the $\sigma$-algebra generated by $\xi_{i,j}, i < k, j = 1, 2, \dots$, and $I_j, j = 0, 1, \dots, k$. Then $\mathcal{F}_k$ contains all information up to generation $k$. For $k \geq 0$, given $\mathcal{F}_k$, $(Y_{n,k,i})_{n\geq k}$, $i = 1, 2, \dots$, are independent critical Galton–Watson processes with initial value 1 at generation $k$.

**Lemma 2.4.** *Suppose Assumption 1 holds. If $\frac{k}{n} \to u$ as $n \to \infty$ for some $u \in (0,1)$, then as $n \to \infty$, the random measure*

$$V_{n,k}(\cdot) = \sum_{i=1}^{Z_k} \mathrm{I}_{\{Y_{n,k,i}>0\}}\delta_{\frac{Y_{n,k,i}}{n-k}}(\cdot) \in \mathcal{M}$$

*converges in distribution to the random measure $V_u := \sum_{i=1}^{N_u^I}\delta_{\omega_i}(\cdot) \in \mathcal{M}$ with the convention that $V_u = 0$ when $N_u^I = 0$, where $(\omega_i)_{i\geq 1}$ are independent exponential random variables with parameter $\frac{2}{\sigma^2}$, and $N_u^I \in \mathbb{N}$ is independent of $(\omega_i)_{i\geq 1}$ with the law given by (1.6).*

**Proof.** Suppose $g \in C_b^+[0,\infty)$. For any $0 \leq k < n$, let

$$L_{n,k}(g) = \exp\{-\langle g, V_{n,k}\rangle\} = \exp\Big\{-\sum_{i=1}^{Z_k} g\Big(\frac{Y_{n,k,i}}{n-k}\Big)\mathrm{I}_{\{Y_{n,k,i}>0\}}\Big\},$$

and set $S_{n,k}g = E\Big(\exp\{-g(\frac{Y_{n-k}}{n-k})\mathrm{I}_{\{Y_{n-k}>0\}}\}\Big)$. Then we have

$$E[L_{n,k}(g)|\mathcal{F}_k] = E[L_{n,k}(g)|Z_k] = \Big[E\Big(\exp\Big\{-g\Big(\frac{Y_{n-k}}{n-k}\Big)\mathrm{I}_{\{Y_{n-k}>0\}}\Big\}\Big)\Big]^{Z_k} = (S_{n,k}g)^{Z_k}.$$

Let $q_n = P(Y_n > 0)$ be the survival probability of the process $(Y_k)_{k\geq 0}$ in generation $n$. Then we have

$$\begin{aligned} S_{n,k}g &= E\Big[\exp\Big\{-g\Big(\frac{Y_{n-k}}{n-k}\Big)\Big\}\Big| Y_{n-k} > 0\Big]q_{n-k} + (1 - q_{n-k}) \\ &= 1 - q_{n-k}\Big[1 - E\Big(\exp\Big\{-g\Big(\frac{Y_{n-k}}{n-k}\Big)\Big\}\Big| Y_{n-k} > 0\Big)\Big]. \end{aligned}$$

It follows from (2.2) that for any $g \in C_b^+[0,\infty)$ and $u \in (0,1)$,

$$\lim_{n\to\infty, k/n\to u} E\Big[\exp\Big\{-g\Big(\frac{Y_{n-k}}{n-k}\Big)\Big\}\Big| Y_{n-k} > 0\Big] = \frac{2}{\sigma^2}\int_0^\infty e^{-g(r)}e^{-\frac{2r}{\sigma^2}}dr =: L(g).$$

From these we derive that

$$\lim_{n\to\infty, k/n\to u} \frac{\ln S_{n,k}g}{q_{n-k}} = -(1 - L(g)). \tag{2.5}$$

Meanwhile, using (2.1) and Lemma 2.2, we obtain that $Z_k q_{n-k}$ converges to $\xi_u$ weakly as $n \to \infty, k/n \to u$ for $u \in (0,1)$, where $\xi_u$ is a random variable having Gamma distribution with parameters $\Big(\frac{1-u}{u}, \gamma\Big)$. As a result, $-Z_k \ln(S_{n,k}g)$ converges weakly to $\xi_u(1 - L(g))$. Since $e^{-x}$ is a bounded continuous function for $x \in [0,\infty)$, we have

$$\begin{aligned} &\lim_{n\to\infty, k/n\to u} E\big[\exp\{-\langle g, V_{n,k}\rangle\}\big] = \lim_{n\to\infty, k/n\to u} E\big[E\left(L_{n,k}(g)|\mathcal{F}_k\right)\big] \\ &= \lim_{n\to\infty, k/n\to u} E\big[(S_{n,k}g)^{Z_k}\big] = \lim_{n\to\infty, k/n\to u} E\big[\exp\{-(-Z_k\ln(S_{n,k}g))\}\big] \\ &= E[\exp\{-\xi_u(1 - L(g))\}]. \end{aligned} \tag{2.6}$$

The Laplace transform of $\xi_u$ is given by (c.f. Sato (2014, Example 2.15))

$$L_{\xi_u}(\lambda) = Ee^{-\lambda\xi_u} = \Big(1 + \frac{u\lambda}{1-u}\Big)^{-\gamma}, \qquad \lambda > 0.$$

Therefore,

$$\begin{aligned} &E\big[\exp\{-\xi_u(1 - L(g))\}\big] = \Big[1 + \frac{u}{1-u}(1 - L(g))\Big]^{-\gamma} = (1-u)^\gamma[1 - uL(g)]^{-\gamma} \\ &= \sum_{k=0}^\infty \frac{(-\gamma)(-\gamma-1)\cdots(-\gamma-k+1)}{k!}(1-u)^\gamma(-u)^k L(g)^k \\ &= Ee^{-\sum_{j=1}^{N_u^I} g(w_j)} = E\big[e^{-\langle g, V_u\rangle}\big]. \end{aligned}$$

In conclusion, $V_{n,k}$ converges to $V_u$ in distribution as $n \to \infty, k/n \to u$. $\square$

For $r > 0$, put

$$f(r) = r, \quad g_1(r) = r \wedge r^{-1}, \quad g_2(r) = 1 \wedge r^2. \tag{2.7}$$

**Remark 2.5.** Using the same argument as in the proof of Lemma 2.4 for the random measure

$$\widetilde{V}_{n,k}(\cdot) := \sum_{i=1}^{Z_k} \mathrm{I}_{\{Y_{n,k,i}>0\}} \Big(1 \vee \Big(\frac{Y_{n,k,i}}{n-k}\Big)^2\Big) \delta_{\frac{Y_{n,k,i}}{n-k}}(\cdot),$$

and using the fact that $h(r) := (1 \vee r^2)$ is a continuous function on $[0,\infty)$, we obtain that

$$\widetilde{V}_{n,k}(dr) \xrightarrow{d} (1 \vee r^2) V_u(dr) =: \widetilde{V}_u(dr) \quad \text{in } \mathcal{M}.$$

Since $g_1, g_2 \in C_b^+[0,\infty)$, $\langle g_1, \widetilde{V}_{n,k}\rangle = \langle f, V_{n,k}\rangle$ and $\langle g_2, \widetilde{V}_{n,k}\rangle = \langle f^2, V_{n,k}\rangle$, we have

$$\begin{aligned} &(\langle f, V_{n,k}\rangle,\ \langle f^2, V_{n,k}\rangle) = (\langle g_1, \widetilde{V}_{n,k}\rangle,\ \langle g_2, \widetilde{V}_{n,k}\rangle) \\ &\xrightarrow{d} (\langle g_1, \widetilde{V}_u\rangle,\ \langle g_2, \widetilde{V}_u\rangle) = (\langle f, V_u\rangle,\ \langle f^2, V_u\rangle) = \Big(\sum_{k=1}^{N_u^I} \omega_k,\ \sum_{k=1}^{N_u^I} \omega_k^2\Big), \end{aligned} \tag{2.8}$$

as $n \to \infty, k/n \to u$ with $u \in (0,1)$.

Define the birth time $\tau_n$ of the oldest clan in generation $n$ by

$$\tau_n = \inf\Big\{0 \le j \le n; \sum_{l=1}^{I_j} Y_{n-j,l}^{(j)} > 0\Big\}$$

with the convention $\inf \emptyset = +\infty$. The birth time of the oldest clan for stationary continuous state branching processes is studied in Chen and Delmas (2012, Corollary 4.2). Using Lemma 2.4, it is easy to get the limit distribution of $\tau_n$. Recall that $\gamma = 2\beta/\sigma^2$.

**Corollary 2.6.** *Suppose Assumption 1 holds. We have*

$$\lim_{n\to\infty, k/n\to u} P(\tau_n > k) = P(N_u^I = 0) = (1-u)^\gamma, \qquad 0 < u < 1.$$

**Proof.** The event $\{\tau_n > k\}$ can be written as $\{V_{n,k}(1) = 0\}$. Thus

$$\lim_{n\to\infty, k/n\to u} P(\tau_n > k) = \lim_{n\to\infty, k/n\to u} P(V_{n,k}(1) = 0) = P(N_u^I = 0) = (1-u)^\gamma. \quad \square$$

Define a function $w$ by

$$w(r) = r \vee r^2, \quad r \in (0,\infty). \tag{2.9}$$

We next consider the following random measures related to immigrations after generation $k$,

$$W_{n,k}(\cdot) := \sum_{j=k+1}^{n} \sum_{l=1}^{I_j} \mathrm{I}_{\big\{Y_{n-j,l}^{(j)}>0\big\}} w\Big(\frac{Y_{n-j,l}^{(j)}}{n-k}\Big) \delta_{\frac{Y_{n-j,l}^{(j)}}{n-k}}(\cdot), \qquad n > k.$$

For each $(n,k)$ with $k < n$, thanks to (2.4), we see that $W_{n,k}(\cdot)$ has the same distribution as the random measure

$$\widetilde{W}_{n-k}(\cdot) := \sum_{j=0}^{n-k-1} \sum_{l=1}^{I_j} \mathrm{I}_{\big\{Y_{j,l}>0\big\}} w\Big(\frac{Y_{j,l}}{n-k}\Big) \delta_{\frac{Y_{j,l}}{n-k}}(\cdot), \tag{2.10}$$

where $Y_{j,l}, j \in \mathbb{N}, l = 1,2,\ldots$, are independent and for each $j$, $Y_{j,l}, l = 1,2,\ldots$, are identically distributed as $Y_j$, and where $(Y_{j,l})_{j\ge0,l\ge1}$ are independent of the immigration process $(I_j)_{j\ge0}$. By an argument very similar to that used in the proof of Lemma 2.4, we get the following convergence in distribution result for the random measures $(W_{n,k})_{n\ge k}$.

**Lemma 2.7.** *Suppose Assumption 1 holds. Let $\zeta$ be the random measure defined by*

$$\zeta(dr) = w(r) W(dr),$$

*where $W$ is a Poisson random measure with intensity $\frac{\gamma}{r} e^{-\frac{2r}{\sigma^2}} dr$ on $(0,\infty)$ and $w$ is the function defined in (2.9). Then $W_{n,k} \xrightarrow{d} \zeta$ in $\mathcal{M}$ as $n-k \to \infty$.*

**Proof.** Since $W_{n,k} \overset{d}{=} \widetilde{W}_{n-k}$, we have for $g \in C_b^+[0,\infty)$,

$$E\big[\exp\{-\langle g, W_{n,k}\rangle\}\big] = E\Big[\exp\{-\langle g, \widetilde{W}_{n-k}\rangle\}\Big], \tag{2.11}$$

which means that we only need to consider the limit of the Laplace functional of $\widetilde{W}_n$ as $n \to \infty$. For any $g \in C_b^+[0,\infty)$, put

$$T_{n,j}(g) = E\Big[\exp\Big\{-w\left(\frac{Y_j}{n}\right) g\left(\frac{Y_j}{n}\right) \mathrm{I}_{\{Y_j>0\}}\Big\}\Big], \quad j = 0, 1, \ldots, n-1.$$

Then $0 < T_{n,j}(g) < 1$. By the definition (2.10) of $\widetilde{W}_n$,

$$\exp\big\{-\langle g, \widetilde{W}_n\rangle\big\} = \exp\Big\{-\sum_{j=0}^{n-1}\sum_{l=1}^{I_j} w\left(\frac{Y_{j,l}}{n}\right) g\left(\frac{Y_{j,l}}{n}\right) \mathrm{I}_{\{Y_{j,l}>0\}}\Big\}.$$

The Laplace transform of $\widetilde{W}_n$ can be written as

$$E\big[\exp\{-\langle g, \widetilde{W}_n\rangle\}\big] = \prod_{j=0}^{n-1} E\big[T_{n,j}(g)^{I_j}\big] = \prod_{j=0}^{n-1} B\big(T_{n,j}(g)\big) = \exp\Big\{\sum_{j=0}^{n-1} \ln B\big(T_{n,j}(g)\big)\Big\}, \tag{2.12}$$

where $B(s) = \sum_j b_j s^j, |s| < 1$, is the probability generating function of $I_k, k \ge 0$. We claim that

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} \ln B\big(T_{n,j}(g)\big) = \gamma \int_0^\infty \big(e^{-w(r)g(r)} - 1\big)\frac{1}{r} e^{-\frac{2r}{\sigma^2}}\, dr. \tag{2.13}$$

Suppose for the moment the claim is true. Then by (2.12), for any $g \in C_b^+[0,\infty)$,

$$\lim_{n\to\infty} E\big[\exp\{-\langle g, \widetilde{W}_n\rangle\}\big] = \exp\Big\{\gamma \int_0^\infty \big(e^{-w(r)g(r)} - 1\big)\frac{1}{r} e^{-\frac{2r}{\sigma^2}}\, dr\Big\}.$$

And then using (2.11), we have

$$\lim_{n-k\to\infty} E\big[\exp\{-\langle g, W_{n,k}\rangle\}\big] = \exp\Big\{\gamma \int_0^\infty \big(e^{-w(r)g(r)} - 1\big)\frac{1}{r} e^{-\frac{2r}{\sigma^2}}\, dr\Big\}.$$

Since $\int_0^\infty (w(r)\wedge 1)\frac{1}{r} e^{-\frac{2r}{\sigma^2}}\, dr < \infty$, it follows from Kallenberg (2017, Theorem 3.20) that there is an infinitely divisible random measure $\zeta \in \mathcal{M}$ represented as $\zeta(dr) = w(r)W(dr), r > 0$, where $W$ is a Poisson random measure with intensity $\mathrm{I}_{\{r>0\}}\frac{\gamma}{r} e^{-\frac{2r}{\sigma^2}}\, dr$. The Laplace functional of $\zeta$ is given by

$$E\big[\exp\{-\langle g, \zeta\rangle\}\big] = \exp\Big\{\gamma \int_0^\infty \big(e^{-w(r)g(r)} - 1\big)\frac{1}{r} e^{-\frac{2r}{\sigma^2}}\, dr\Big\}, \quad \forall g \in C_b^+[0,\infty).$$

Thus $W_{n,k} \xrightarrow{d} \zeta$ as $n - k \to \infty$.

Now we prove the claim (2.13). By the mean value theorem, there exists $\xi_{n,j} \in (T_{n,j}(g), 1)$ such that

$$\begin{aligned} B\big(T_{n,j}(g)\big) - 1 &= B'(\xi_{n,j})\big(T_{n,j}(g) - 1\big) \\ &= \beta\big(T_{n,j}(g) - 1\big) + (B'(\xi_{n,j}) - \beta)\big(T_{n,j}(g) - 1\big). \end{aligned} \tag{2.14}$$

Thanks to the inequality $0 < 1 - e^{-x} \le x$ for $x > 0$ and the fact that $\mathrm{Var}(Y_j) = j\sigma^2$ (see Athreya and Ney (1972, Section $I$.2)), we have that for $0 \le j \le n-1$,

$$0 \le 1 - T_{n,j}(g) \le \|g\|_\infty E\Big[w\big(\frac{Y_j}{n}\big)\Big] \le \|g\|_\infty E\Big[\frac{Y_j}{n} + \Big(\frac{Y_j}{n}\Big)^2\Big] \le \frac{a\|g\|_\infty}{n}, \tag{2.15}$$

for some constant $a > 0$. Thus $n(1 - T_{n,j}(g))$ is bounded for $n > 0$ and $j \le n$. Moreover from (2.1) and (2.2), it follows that for any $0 < t < 1$,

$$\begin{aligned} \lim_{n\to\infty} n[1 - T_{n,[nt]}(g)] &= \lim_{n\to\infty} nP(Y_{[nt]} > 0)E\Big[1 - \exp\Big\{-w\big(\frac{Y_{[nt]}}{n}\big)g\big(\frac{Y_{[nt]}}{n}\big)\Big\}\Big|Y_{[nt]} > 0\Big] \\ &= \frac{4}{(\sigma^2)^2 t}\int_0^\infty \big(1 - e^{-w(rt)g(rt)}\big)\, e^{-\frac{2r}{\sigma^2}}\, dr. \end{aligned}$$

Then by the dominated convergence theorem,

$$\begin{aligned} \lim_{n\to\infty} \sum_{j=0}^{n-1}(T_{n,j}(g) - 1) &= \lim_{n\to\infty} \int_0^1 n(T_{n,[nt]}(g) - 1)dt \\ &= \int_0^1 \frac{4}{(\sigma^2)^2 t}dt \int_0^\infty \big(e^{-w(rt)g(rt)} - 1\big)e^{-\frac{2r}{\sigma^2}}\, dr \\ &= \frac{2}{\sigma^2}\int_0^\infty \big(e^{-w(r)g(r)} - 1\big)\frac{1}{r}e^{-\frac{2r}{\sigma^2}}\, dr. \end{aligned} \tag{2.16}$$

Using (2.15) and the continuity of $B'(s)$ on $[0,1]$, we get that $B'(\xi_{n,j}) - \beta$ converges to 0 uniformly for $0 \le j \le n$, as $n \to \infty$. It has been shown in (2.16) that $\sum_{j=0}^{n-1}\big|T_{n,j}(g) - 1\big|$ converges. Therefore, $\sum_{j=0}^{n-1}(B'(\xi_{n,j}) - \beta)\big(T_{n,j}(g) - 1\big)$ converges to 0. Thus, by (2.14),

$\sum_{j=0}^{n-1}\big(B\big(T_{n,j}(g)\big)-1\big)$ and $\beta\sum_{j=0}^{n-1}\big(T_{n,j}(g)-1\big)$ have the same limit. More precisely, from (2.16), it follows that

$$\lim_{n\to\infty}\sum_{j=0}^{n-1}(B\big(T_{n,j}(g)\big)-1)=\lim_{n\to\infty}\beta\sum_{j=0}^{n-1}\big(T_{n,j}(g)-1\big)$$
$$=\gamma\int_0^\infty\big(e^{-w(r)g(r)}-1\big)\frac{1}{r}e^{-\frac{2r}{\sigma^2}}\,dr.$$

Meanwhile, since $-x\geq\ln(1-x)\geq -x-\dfrac{x^2}{1-x}$ for $0<x<1$, if

$$\lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{\big[B\big(T_{n,j}(g)\big)-1\big]^2}{B\big(T_{n,j}(g)\big)}=0,\tag{2.17}$$

then $\sum_{j=0}^{n-1}\ln B\big(T_{n,j}(g)\big)$ and $\beta\sum_{j=0}^{n-1}\big(T_{n,j}(g)-1\big)$ have the same limit as $n\to\infty$, and thus the claim is true. Now we prove (2.17). By (2.15), for any $1/2<\delta<1$, there is $N>0$, such that for any $n>N, 0<j\leq n$, $T_{n,j}(g)>\delta$. Since $B(s)$ is an increasing continuous function on $[0,1]$ and $B(1)=1$, for any $\varepsilon>0$, we can choose $\delta$ above such that when $1>s>\delta$, $B(s)>1-\varepsilon$. Therefore when $n>N$,

$$0\leq\sum_{j=0}^{n-1}\frac{\big[B\big(T_{n,j}(g)\big)-1\big]^2}{B\big(T_{n,j}(g)\big)}\leq\frac{\varepsilon}{B(\frac{1}{2})}\sum_{j=0}^{n-1}\big[1-B\big(T_{n,j}(g)\big)\big].$$

Then (2.17) follows from the convergence of $\sum_{j=0}^{n-1}\big[1-B\left(T_{n,j}(g)\right)\big]$ and the arbitrariness of $\varepsilon$. $\square$

**Remark 2.8.** (1) Let $\tilde g_1(r)=1\wedge r^{-1}, r>0$ and $\tilde g_2(r)=1\wedge r, r>0$. Then $\tilde g_1,\tilde g_2\in C_b^+[0,\infty)$. Thanks to Lemma 2.7 and the facts $\tilde g_1(r)w(r)=r=f(r)$ and $\tilde g_2(r)w(r)=r^2=f^2(r)$ for $r>0$, we get that

$$\big(\langle\tilde g_1,W_{n,k}\rangle,\langle\tilde g_2,W_{n,k}\rangle\big)\xrightarrow{d}\big(\langle\tilde g_1,\zeta\rangle,\langle\tilde g_2,\zeta\rangle\big)=\big(\langle f,W\rangle,\langle f^2,W\rangle\big),\quad\text{as } n-k\to\infty.$$

(2) We observe that

$$\frac{n-k}{n}\big[\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle\big]=\frac{Z_n}{n},$$

where $f$ and $\tilde g_1$ are defined as above. Since $V_{n,k}$ and $W_{n,k}$ are independent, from Lemmas 2.4 and 2.7, it follows that for any $\lambda>0$.

$$\lim_{n\to\infty}E\Big[\exp\Big\{-\lambda\frac{Z_n}{n}\Big\}\Big]=\lim_{n\to\infty,k/n\to u}E\exp\Big\{-\lambda\frac{n-k}{n}(\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle)\Big\}$$
$$=\lim_{n\to\infty,k/n\to u}E\Big[\exp\{-\lambda\frac{n-k}{n}\langle f,V_{n,k}\rangle\}\Big]\lim_{n\to\infty,k/n\to u}E\Big[\exp\{-\lambda\frac{n-k}{n}\langle\tilde g_1,W_{n,k}\rangle\}\Big]$$
$$=E[\exp\{-\lambda(1-u)\langle f,V_u\rangle\}]E[\exp\{-\lambda(1-u)\langle f,W\rangle\}]$$
$$=\Big(\frac{\lambda+\frac{2}{\sigma^2}}{\lambda(1-u)+\frac{2}{\sigma^2}}\Big)^{-\gamma}\Big(\frac{\lambda(1-u)+\frac{2}{\sigma^2}}{\frac{2}{\sigma^2}}\Big)^{-\gamma}=\Big(1+\frac{\lambda}{\frac{2}{\sigma^2}}\Big)^{-\gamma},$$

where the last term is the Laplace transform of the Gamma distribution with parameters $(\frac{2}{\sigma^2},\gamma)$. This is consistent with Lemma 2.2.

## 3. Proofs of the main results

**Proof of Theorem 1.1.** Let $f$ be the function defined in (2.7), and let $\tilde g_1,\tilde g_2$ be the functions defined in Remark 2.8(1). The random variable in Lemma 2.3 can be expressed in terms of the random measures defined in Section 2, and then we have

$$\frac{\sum_{i=1}^{Z_k}\big(Y_{n,k,i}\big)_2+\sum_{j=1+k}^{n-1}\sum_{l=1}^{I_j}\big(Y_{n-j,l}^{(j)}\big)_2}{(Z_n)_2}=\frac{\langle f^2,V_{n,k}\rangle-\frac{1}{n-k}\langle f,V_{n,k}\rangle+\langle\tilde g_2,W_{n,k}\rangle-\frac{1}{n-k}\langle\tilde g_1,W_{n,k}\rangle}{\big[\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle\big]^2-\frac{1}{n-k}\big[\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle\big]}.$$

Since $(V_{n,k})_{n>k}$ and $(W_{n,k})_{n>k}$ are independent and $0<\dfrac{\sum_{i=1}^{Z_k}\big(Y_{n,k,i}\big)_2+\sum_{j=1+k}^{n}\sum_{l=1}^{I_j}\big(Y_{n-j,l}^{(I_j)}\big)_2}{(Z_n)_2}\leq 1$ is a bounded continuous function of $(\langle f,V_{n,k}\rangle,\langle f^2,V_{n,k}\rangle,\langle\tilde g_1,W_{n,k}\rangle,\langle\tilde g_2,W_{n,k}\rangle)$, according to Remark 2.5 and Remark 2.8, for $u\in(0,1)$,

$$\lim_{n\to\infty,k/n\to u}\frac{\langle f^2,V_{n,k}\rangle-\frac{1}{n-k}\langle f,V_{n,k}\rangle+\langle\tilde g_2,W_{n,k}\rangle-\frac{1}{n-k}\langle\tilde g_1,W_{n,k}\rangle}{\big[\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle\big]^2-\frac{1}{n-k}\big[\langle f,V_{n,k}\rangle+\langle\tilde g_1,W_{n,k}\rangle\big]}=\frac{\langle f^2,V_u\rangle+\langle f^2,W\rangle}{\big[\langle f,V_u\rangle+\langle f,W\rangle\big]^2}$$

in distribution. It follows from Lemma 2.2 that $\lim_{n\to\infty}P(Z_n>1)=1$. The results of this theorem follow from Lemma 2.3.

**Proof of Theorem 1.2.** If all the particles in generation $n$ have the same ancestor, then they must be descendants of one immigrant before generation $n$. Thus

$$\big\{A_n^I<\infty, Z_n>0\big\}\subset\big\{Y_{n-j,l}^{(j)}=0\,\text{for all but one pair }(j,l), 0\leq j\leq n, 1\leq l\leq I_j\big\}.$$

Then we only need to prove that the probability of the event on the right hand side converges to 0. Recall that $q_n = P(Y_n > 0)$. Set $a_n = 1 - q_n = P(Y_n = 0)$. Then

$$\begin{aligned}
&P\Big(Y^{(j)}_{n-j,l} = 0 \text{ for all but one pair}(j,l), 0 \le j \le n, 1 \le l \le I_j\Big) \\
&= E\Big[\sum_{j=0}^{n}\prod_{k\neq j} P(Y_{n-k} = 0)^{I_k} I_j P(Y_{n-j} = 0)^{I_j - 1} P(Y_{n-j} > 0)\Big] \\
&= \Big[\prod_{k=0}^{n} B(a_k)\Big]\Big[\sum_{j=0}^{n} \frac{B'(a_j)}{B(a_j)} q_j\Big],
\end{aligned} \tag{3.1}$$

where $B(a_0) = B(0) = b_0$ and $B'(a_0) = B'(0) = b_1$. From (2.1), we know $q_k = 1 - a_k \sim \frac{2}{\sigma^2 k}$ as $k \to \infty$. In addition, since $B(s) = 1 + \beta(s-1) + o(1-s)$ as $s \to 1-$,

$$\lim_{j\to\infty} j(1 - B(a_j)) = \lim_{j\to\infty} \beta j(1-a_j) + o(j(1-a_j)) = \gamma > 0. \tag{3.2}$$

Therefore, there exists some $N \in \mathbb{N}$, such that when $k \ge N$, $k(1 - B(a_k)) > \gamma/2$, which implies that $B(a_k) < 1 - \frac{\gamma}{2k}$ for $k \ge N$. Noticing that $B(a_k) \le 1$, the first factor on the right-hand side of (3.1) can be estimated as follows:

$$\prod_{j=0}^{n} B(a_j) \le \prod_{j=N}^{n} B(a_j) \le \prod_{j=N}^{n}\big(1 - \frac{\gamma}{2j}\big) = \exp\Big\{\sum_{j=N}^{n} \ln(1 - \frac{\gamma}{2j})\Big\}, \quad n > N.$$

Since $\ln(1-x) < -x$ for $0 < x < 1$, we have

$$\sum_{k=N}^{n} \ln(1 - \frac{\gamma}{2k}) \le -\sum_{k=N}^{n} \frac{\gamma}{2k} \le -L(\ln n - \ln N),$$

for some constant $L > 0$. As a result, there exists $C_1 > 0$, such that

$$\prod_{k=0}^{n} B(a_k) \le C_1 \cdot n^{-L}. \tag{3.3}$$

Since $a_k$ is nondecreasing in $k$ and converges to 1 as $k \to \infty$, and $B'(s)$ is a continuous function on $[0,1]$,

$$\lim_{j\to\infty} \frac{B'(a_j)}{B(a_j)} = B'(1) = \beta.$$

The second factor on the right-hand side of (3.1) has the following upper bound:

$$\sum_{j=1}^{n} \frac{B'(a_j)}{B(a_j)} q_j \le C_2 \sum_{j=1}^{n} q_j \le C_3 \sum_{j=1}^{n} \frac{1}{j} \le C_3(1 + \ln n), \tag{3.4}$$

for some positive constants $C_2$ and $C_3$. Combining (3.3) and (3.4), we obtain

$$\lim_{n\to\infty} P\Big(Y^{(j)}_{n-j,l} = 0 \text{ for all but one pair } (j,l),\ 0 \le j \le n, 1 \le l \le I_j\Big) = 0.$$

We finish the proof. □

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## Acknowledgments

We thank the referees for very helpful comments and suggestions.

## References

Athreya, K.B., 2010. Ancestor problem for branching trees. In: Mathematics Newsletter: Special Issue Commemorating ICM 2010 in India, vol. 19, (1), pp. 1–10.
Athreya, K.B., 2012a. Coalescence in critical and subcritical Galton–Watson branching processes. J. Appl. Probab. 49 (03), 627–638.
Athreya, K.B., 2012b. Coalescence in the recent past in rapidly growing populations. Stoch. Process. Appl. 122, 3757–3766.
Athreya, K.B., Ney, P.E., 1972. Branching Processes. Springer Verlag, New York-Heidelberg.
Chen, Y.T., Delmas, J.F., 2012. Smaller population size at the mrca time for stationary branching processes. Ann. Probab. 40, 2034–2068.
Durrett, R., 1978. The genealogy of critical branching processes. Stoch. Process. Appl. 8, 101–116.

Harris, S.C., Johnston, S.G.G., Roberts, M.I., 2020. The coalescent structure of continuous-time Galton–Watson trees. Ann. Appl. Probab. 30, 1368–1414.
Heathcote, C.R., 1965. A branching process allowing immigration. J. R. Stat. Soc. Ser. B Stat. Methodol. 27, 138–143.
Hong, J.-I., 2016. Coalescence on critical and subcritical multitype branching processes. J. Appl. Probab. 53, 802–817.
Johnston, S.G.G., 2019. The genealogy of Galton–Watson trees. Electron. J. Probab. 24, 1–35.
Kallenberg, O., 2017. Random Measures, Theory and Applications. In: Probability Theory and Stochastic Modelling, vol. 77, Springer, Cham.
Kersting, G., 2022. On the genealogical structure of critical branching processes in a varying environment. Proc. Steklov Inst. Math 316, 209–219.
Le, V., 2014. Coalescence times for the Bienaymé-Galton–Watson process. J. Appl. Probab. 51 (1), 209–218.
Pakes, A.G., 1971a. Branching processes with immigration. J. Appl. Probab. 8, 32–42.
Pakes, A.G., 1971b. On the critical Galton–Watson process with immigration. J. Austral. Math. Soc. 12, 476–482.
Sato, K., 2014. Lévy Processes and Infinitely Divisible Distributions. In: Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University.
Wang, H.M., Li, L., Yao, H., 2019. Coalescence for supercritical Galton–Watson processes with immigration. arXiv:1906.03945.
Zubkov, A.M., 1975. Limit distributions of the distance to the nearest common ancestor. Theory Probab. Appl. 20, 602–612.