Bernoulli **17**(1), 2011, 138–154
DOI: 10.3150/10-BEJ264

Supercritical age-dependent branching Markov processes and their scaling limits

KRISHNA B. ATHREYA[1], SIVA R. ATHREYA[2] and SRIKANTH K. IYER[3]

[1]*Department of Mathematics and Statistics, Iowa State University, Ames, Iowa 50011, USA.*
E-mail: kba@iastate.edu
[2]*Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560059, India.*
E-mail: athreya@isibang.ac.in
[3]*Department of Mathematics, Indian Institute of Science, Bangalore 560012, India.*
E-mail: skiyer@math.iisc.ernet.in

This paper studies the long-time behavior of the empirical distribution of age and normalized position of an age-dependent supercritical branching Markov process. The motion of each individual during its life is a random function of its age. It is shown that the empirical distribution of the age and the normalized position of all individuals alive at time t converges as $t \to \infty$ to a deterministic product measure.

Keywords: age-dependent; ancestral times; branching; empirical distribution; supercritical

1. Introduction

We consider an age-dependent branching Markov process consisting of a finite collection of particles distributed in $\mathbb{R}$ in which each particle lives for a random length of time and, upon its death, gives rise to a random number of offspring. Further, during its lifetime, each offspring migrates according to a prescribed Markov process, starting from the position where its parent died. The motion process, offspring distribution and lifetime process are all independent of each other. Further, we allow the motion process to be age-dependent and assume that the system is supercritical, that is, the mean of the offspring distribution is greater than one. We also assume that the probability of an individual producing zero offspring is zero. This implies that there is no extinction. We shall describe the model more precisely in the next section.

We study two aspects of such a system. First, at time t, we consider a randomly chosen individual from the population. We show that, asymptotically (as $t \to \infty$), the joint distribution of the position (appropriately scaled) and age (unscaled) of the randomly chosen individual decouples (see Theorem 1.1). Second, it is shown that the empirical joint distribution of the age and the normalized position of the population at time t converges in law as $t \to \infty$ to a deterministic measure (see Theorem 1.2).

Limit theorems (such as Theorems 1.1 and 1.2) for branching Markov processes where the motion depends on the age do not seem to have been considered in the literature. Our results on the age distribution (alone) in Theorems 1.1 and 1.2 are well known [2]. This has also been established for the general Crump–Mode–Jagers process by Jagers and Nerman (see [6,7]). Our focus here is on the joint distribution of the age (unscaled) and the scaled position. In this paper, we have restricted ourselves to the case where the branching part is that of a Bellman–Harris process (see [2]). If one looks at the positions of the individuals in the embedded Galton–Watson tree, it gives rise to a branching random walk. There are a number of results (see, e.g., [3,4]) on the asymptotic behavior of these random walks. Our setting is in continuous time and the question of interest is the asymptotic behavior of the joint distribution of the age and positions of the particles alive at time t as $t \to \infty$.

1.1. The model

We begin with the description of the particle system. We assume $(\Omega, \mathcal{F}, P)$ to be a canonical probability space on which all the random variables are defined. Suppose we are given the following:

(i) *lifetime distribution* $G(\cdot)$: let $G(\cdot)$ be a cumulative distribution function that is non-lattice on $[0, \infty)$ with $G(0) = 0$ and $\mu = \int_0^\infty s G(\mathrm{d}s) < \infty$;
(ii) *offspring distribution* $\mathbf{p}$: let $\mathbf{p} \equiv \{p_k\}_{k \geq 0}$ be a probability distribution such that $p_0 = 0$, $m = \sum_{k=0}^\infty k p_k > 1$ and $\sum_{k=0}^\infty k \log(k) p_k < \infty$;
(iii) *motion process* $\eta(\cdot)$: let $\eta(\cdot)$ be an $\mathbb{R}$-valued Markov process starting at 0.

Let α be the Malthusian parameter defined by $m \int_0^\infty \mathrm{e}^{-\alpha s} G(\mathrm{d}s) = 1$.

Branching Markov process $(G, \mathbf{p}, \eta)$. Suppose we start with an initial configuration $\mathcal{C}_0 = \{(a_0^i, X_0^i) : i = 1, 2, \ldots, Z_0\}$, $0 < Z_0 < \infty$, a_0^i, X_0^i denote the age and position of the ith particle at time 0, respectively. Each particle in the system lives for a random length of time L with distribution G and upon its death gives rise to a random number of offspring ξ with distribution $\mathbf{p}$. During its lifetime L, the particle will move in $\mathbb{R}$, according to the process $\{x + \eta(t) : 0 \leq t \leq L\}$, where x denotes the position of its parent at the time of its birth. More precisely, if an individual is born at time τ and at location x and has lifetime L, then it moves during $[\tau, \tau + L)$ and its movement $\{X(t) : \tau \leq t < \tau + L\}$ is distributed as $\{x + \eta(t - \tau) : \tau \leq t < \tau + L\}$, thus the movement of any individual is a random function of its age. We assume that the three objects (L, ξ, η) associated with each particle are independent and that the family of triplets (L, ξ, η) over all particles in the system are i.i.d.

Let Z_t be the number of particles alive at time t and

$$\mathcal{C}_t = \{(a_t^i, X_t^i) : i = 1, 2, \ldots, Z_t\} \tag{1.1}$$

denote the age and position configuration of all the individuals alive at time t. Since $m < \infty$ and $G(0) = 0$, there is no explosion in finite time (i.e., $P(Z_t < \infty) = 1$). Also, $P(\eta(L) \in R) = 1$ for each particle. Thus, $\mathcal{C}_t$ is well defined for each $0 \leq t < \infty$ and the process $\{\mathcal{C}_t : t \geq 0\}$ is Markov.

For a particle chosen uniformly at random from those alive at time t, let M_t, $\{L_{ti}, \{\eta_{ti}(u), 0 \leq u \leq L_{ti}\} : 1 \leq i \leq M_t\}$ denote, respectively, the number of ancestors, the lifetimes, and the motion processes of its ancestors and $\{\eta_{t(M_t+1)}(u) : 0 \leq u \leq t - \sum_{i=1}^{M_t} L_{ti}\}$ the motion of the individual concerned. If $M_t = 0$, then

$$a_t = a_0 + t \quad \text{and} \quad X_t = X_0 + \eta(a_0 + t) - \eta(a_0), \tag{1.2}$$

where a_0 is the age and X_0 is the location of the particle at time $t = 0$ and $\eta(\cdot)$ is the assumed motion process. If $M_t > 0$, then the age and position, (a_t, X_t), of the particle are given by

$$a_t = t - \sum_{i=1}^{M_t} L_{ti} \tag{1.3}$$

and

$$X_t = X_0 + \eta_{t1}(L_{t1}) - \eta_{t1}(a_0) + \sum_{i=2}^{M_t} \eta_{ti}(L_{ti}) + \eta_{t(M_t+1)}(a_t). \tag{1.4}$$

Note that, given $\{M_t, L_{ti}, 1 \le i \le M_t\}$, the collection of stochastic processes $\{\eta_{ti}(u), 0 \le u \le L_{ti}, 1 \le i \le M_t\}$ have the same distribution as η and are independent of each other.

Let $\mathcal{B}(\mathbb{R}_+)$ (and $\mathcal{B}(\mathbb{R})$) be the Borel σ-algebra on $\mathbb{R}_+$ (and $\mathbb{R}$). Let $M(\mathbb{R}_+ \times \mathbb{R})$ be the space of finite Borel measures on $\mathbb{R}_+ \times \mathbb{R}$ equipped with the weak topology. Let $M_a(\mathbb{R}_+ \times \mathbb{R}) := \{\nu \in M(\mathbb{R}_+ \times \mathbb{R}) : \nu = \sum_{i=1}^n \delta_{a_i, x_i}(\cdot, \cdot), n \in \mathbb{N}, a_i \in \mathbb{R}_+, x_i \in \mathbb{R}\}$. For any set $A \in \mathcal{B}(\mathbb{R}_+)$ and $B \in \mathcal{B}(\mathbb{R})$, let $Y_t(A \times B)$ be the number of particles at time t whose age is in A and position is in B. As pointed out earlier, $m < \infty$, $G(0) = 0$ implies that $Y_t \in M_a(\mathbb{R}_+ \times \mathbb{R})$ for all $t > 0$ if the same holds for Y_0. Fix a function $\phi \in C_b^+(\mathbb{R}_+ \times \mathbb{R})$ (the set of all bounded, continuous and positive functions from $\mathbb{R}_+ \times \mathbb{R}$ to $\mathbb{R}_+$) and define

$$\langle \phi, Y_t \rangle \equiv \int \phi \, \mathrm{d}Y_t = \sum_{i=1}^{Z_t} \phi(a_t^i, X_t^i). \tag{1.5}$$

Since $\eta(\cdot)$ is a Markov process, it can be seen that $\{Y_t : t \ge 0\}$ is a Markov process and we shall call $Y \equiv \{Y_t : t \ge 0\}$ the $(G, \mathbf{p}, \eta)$-branching Markov process. Note that $\mathcal{C}_t$ determines Y_t and conversely.

1.2. Main results

In this section, we describe the main results of the paper.

Theorem 1.1 (Limiting behavior of a randomly chosen particle). *Let (a_t, X_t) be the age and position of a randomly chosen particle from those alive at time t. Let $\mu_\alpha = \int_0^\infty s \mathrm{e}^{-\alpha s} \, \mathrm{d}G(s)$. Assume that $\eta(\cdot)$ is such that for all $0 \le t < \infty$,*

$$E(\eta(t)) = 0, \qquad v(t) \equiv E(\eta^2(t)) < \infty, \qquad \sup_{0 \le s \le t} v(s) < \infty \quad \textit{and} \tag{1.6}$$

$$\psi_\alpha \equiv \int_0^\infty v(s) \mathrm{e}^{-\alpha s} G(\mathrm{d}s) < \infty.$$

Then $(a_t, \frac{X_t}{\sqrt{t}})$ converges, as $t \to \infty$, to (U, V) in distribution, where U and V are independent, with U a strictly positive absolutely continuous random variable with density proportional to $\mathrm{e}^{-\alpha \cdot}(1 - G(\cdot))$ and V a normally distributed random variable with mean 0 and variance $\frac{\psi_\alpha}{\mu_\alpha}$.

Note that in the theorem, as $p_0 = 0$, $P(Z_t > 0) = 1$ for all $0 \le t < \infty$ since $Z_0 > 0$. Next, consider the scaled empirical measure $\tilde{Y}_t \in M_a(\mathbb{R}_+ \times \mathbb{R})$ given by

$$\tilde{Y}_t(A \times B) = \frac{Y_t(A \times \sqrt{t}B)}{Z_t} \tag{1.7}$$

for $A \in \mathcal{B}(\mathbb{R}_+)$, $B \in \mathcal{B}(\mathbb{R})$.

Theorem 1.2 (Empirical measure). *Assume* $p_0 = 0$ *and* (1.6). *The scaled empirical measure* $\tilde{Y}_t$, *as defined above in* (1.7), *then converges in distribution to a deterministic measure* ν, *with* $\nu(A \times B) = P(U \in A, V \in B)$, *where* U *and* V *are as in Theorem* 1.1 *for* $A \in \mathcal{B}(\mathbb{R}_+)$, $B \in \mathcal{B}(\mathbb{R})$.

1.3. Layout

The rest of the paper is organized as follows. In Section 2, we prove some preliminary results from renewal theory. In Section 3, we prove four propositions on age-dependent branching processes which are used in proving Theorem 1.1 in Section 4. This is required because sampling an individual from those alive at time t introduces dependencies in the lifetimes of the ancestors of the chosen individual. In Section 3, we also show that the joint distribution of coalescent times for a sample of two individuals chosen at random from the population at time t converges as $t \to \infty$ (see Theorem 3.1). This result is of independent interest and is a key tool that is needed in proving Theorem 1.2 in Section 5. Some of the results presented in Sections 2 and 3 are known in the literature. They are included here to make the paper more accessible and also in a form in which they are needed for the proofs of the main results.

2. Results from renewal theory

Let $\{X_i : i \ge 1\}$ be an i.i.d. sequence of positive random variables with cumulative distribution function G that is non-lattice and satisfies $G(0) = 0$. Let $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, $n \ge 1$. For $t \ge 0$, let $N(t) = k$ if $S_k \le t < S_{k+1}$, $k \ge 0$. Further, let $A_t = t - S_{N(t)}$ and $R_t = S_{N(t)+1} - t$ be, respectively, the age and residual lifetime at time t. Let $\mu \equiv \int_0^\infty x \, \mathrm{d}G(x) \le \infty$.

Lemma 2.1. *Let* A_t, $N(t)$, R_t *and* μ *be as above. Then*:

(i) $\frac{N(t)}{t} \xrightarrow{\text{a.e.}} \frac{1}{\mu}$;

(ii) *if* $\theta \in \mathbb{R}$ *is such that* $g(\theta) = \int_0^\infty \mathrm{e}^{\theta t}(1 - G(t)) \, \mathrm{d}t < \infty$ *and* $\mu < \infty$, *then*

$$\lim_{t\to\infty} E(\mathrm{e}^{\theta A_t}) = \lim_{t\to\infty} E(\mathrm{e}^{\theta R_t}) = \frac{g(\theta)}{\mu} \tag{2.1}$$

and, for any $0 < l < \infty$,

$$\lim_{t\to\infty} E(\mathrm{e}^{\theta A_t} : A_t > l) = \lim_{t\to\infty} E(\mathrm{e}^{\theta R_t} : R_t > l) = \frac{1}{\mu} \int_l^\infty \mathrm{e}^{\theta u}\big(1 - G(u)\big) \, \mathrm{d}u, \tag{2.2}$$

hence

$$\lim_{l\to\infty}\lim_{t\to\infty} E(\mathrm{e}^{\theta R_t} : R_t > l) = 0. \tag{2.3}$$

Proof. Part (i) is a well-known result.

(ii) For $t \ge 0$, $\theta \in \mathbb{R}$, let $f(\theta, t) = E(\mathrm{e}^{\theta A_t})$ and $g(\theta, t) = E(\mathrm{e}^{\theta R_t})$. It is easy to see that

$$f(\theta, t) = \mathrm{e}^{\theta t}\bigl(1 - G(t)\bigr) + \int_0^t f(\theta, t-u)\,\mathrm{d}G(u),$$

$$g(\theta, t) = h(\theta, t)\bigl(1 - G(t)\bigr) + \int_0^t g(\theta, t-u)\,\mathrm{d}G(u),$$

where $h(\theta, t) = E(\mathrm{e}^{\theta(X-t)}|X > t)$ with $X \stackrel{d}{=} G$. By the key renewal theorem, as $t \to \infty$,

$$f(\theta, t) \to \frac{\int_0^\infty \mathrm{e}^{\theta u}(1 - G(u))\,\mathrm{d}u}{\mu}$$

and

$$\begin{aligned} g(\theta, t) &\to \frac{\int_0^\infty h(\theta, u)(1 - G(u))\,\mathrm{d}u}{\mu} \\ &= \frac{\int_0^\infty E(\mathrm{e}^{\theta(X-u)}; X > u)\,\mathrm{d}u}{\mu} \\ &= \frac{E(\int_0^X \mathrm{e}^{\theta u}\,\mathrm{d}u)}{\mu} = \frac{\int_0^\infty \mathrm{e}^{\theta u} P(X > u)\,\mathrm{d}u}{\mu} \\ &= \frac{\int_0^\infty \mathrm{e}^{\theta u}(1 - G(u))\,\mathrm{d}u}{\mu}. \end{aligned}$$

This proves (2.1). This being true for all $\theta \le 0$, it follows that as $t \to \infty$, both $A_t \xrightarrow{d} A_\infty$ and $R_t \xrightarrow{d} R_\infty$, where A_∞ and R_∞ have the same distribution, namely, an absolutely continuous one on $(0, \infty)$ with probability density function $\frac{(1-G(\cdot))}{\mu}$. Since $P(R_\infty = l) = 0$,

$$E(\mathrm{e}^{\theta R_t}; R_t > l) \to E(\mathrm{e}^{\theta R_\infty}; R_\infty > l) = \frac{1}{\mu}\int_l^\infty \mathrm{e}^{\theta t}\bigl(1 - G(t)\bigr)\,\mathrm{d}t,$$

which proves (2.2). Under the hypothesis $\int_0^\infty \mathrm{e}^{\theta t}(1 - G(t))\,\mathrm{d}t < \infty$, (2.3) follows easily from (2.2) by the dominated convergence theorem. □

The first part of the following lemma is known in various forms in the literature (see [3–5]). A proof is given here for our setup in the precise form in which we will need it.

Lemma 2.2. *Let $\{X_i\}_{i\geq 1}$ be i.i.d. positive random variables with distribution function G and $G(0)=0$. Let $m>1$ and $0<\alpha<\infty$ be the Malthusian parameter given by $m\int_0^\infty \mathrm{e}^{-\alpha x}\,\mathrm{d}G(x)=1$. Let $\{\tilde{X}_i\}_{i\geq 1}$ be i.i.d. positive random variables with distribution function $\tilde{G}(x)=m\times\int_0^x \mathrm{e}^{-\alpha y}\,\mathrm{d}G(y)$, $x\geq 0$. Let $N(t), A_t, R_t$ be as defined above for $\{X_i\}_{i\geq 1}$ and $\tilde{N}(t), \tilde{A}_t, \tilde{R}_t$ be the corresponding objects for $\{\tilde{X}_i\}_{i\geq 1}$. Then*:

(i) *for any $k\geq 1$ and bounded Borel measurable function $\phi:\mathbb{R}^k\to\mathbb{R}$,*

$$E(\phi(\tilde{X}_1,\tilde{X}_2,\dots,\tilde{X}_k))=E(\mathrm{e}^{-\alpha S_k}m^k\phi(X_1,X_2,\dots,X_k)),\tag{2.4}$$

where $S_k=\sum_{i=1}^k X_i$;

(ii) *for any $t\geq 0$, $k\geq 0$, $c\in\mathbb{R}$ and Borel measurable function h, and $\varepsilon>0$,*

$$\begin{aligned}&E\left(\mathrm{e}^{\alpha R_t}\mathrm{e}^{-\alpha S_{k+1}}m^{k+1}I\left(N(t)=k,\left|\frac{1}{k}\sum_{i=1}^k h(X_i)-c\right|>\varepsilon\right)\right)\\&\quad=E\left(\mathrm{e}^{\alpha\tilde{R}_t}I\left(\tilde{N}(t)=k,\left|\frac{1}{k}\sum_{i=1}^k h(\tilde{X}_i)-c\right|>\varepsilon\right)\right);\end{aligned}\tag{2.5}$$

(iii)

$$\lim_{l\to\infty}\lim_{t\to\infty}E(\mathrm{e}^{\alpha\tilde{R}_t}:\tilde{R}_t>l)=0.\tag{2.6}$$

Proof. From the definition of $\tilde{G}$, for any Borel sets B_i, $i=1,2,\dots,k$, we have

$$P(\tilde{X}_i\in B_i, i=1,2,\dots,k)=\prod_{i=1}^k m\int_{B_i}\mathrm{e}^{-\alpha x}\,\mathrm{d}G(x)=E\left(\mathrm{e}^{-\alpha S_k}m^k\prod_{i=1}^k I(X_i\in B_i)\right).$$

Therefore, (2.4) follows and (2.5) follows from it by setting, for given $t\geq 0$, $k\geq 0$, $(x_1,x_2,\dots,x_{k+1})\in\mathbb{R}^{k+1}$,

$$\phi_t(x_1,x_2,\dots,x_{k+1})=\mathrm{e}^{\alpha r_t}I\left(s_k\leq t<s_{k+1},\left|\frac{1}{k}\sum_{i=1}^k h(x_i)-c\right|>\varepsilon\right),$$

where $r_t=\sum_{i=1}^{k+1}x_i-t$, $s_j=\sum_{i=1}^j x_i$, $1\leq j\leq k+1$. Now,

$$\begin{aligned}&\int_0^\infty \mathrm{e}^{\alpha x}\big(1-\tilde{G}(x)\big)\,\mathrm{d}x\\&\quad=m\int_0^\infty \mathrm{e}^{\alpha x}\left(\int_x^\infty \mathrm{e}^{-\alpha y}\,\mathrm{d}G(y)\right)\mathrm{d}x=m\int_0^\infty\left(\int_0^y \mathrm{e}^{\alpha x}\,\mathrm{d}x\right)\mathrm{e}^{-\alpha y}\,\mathrm{d}G(y)\\&\quad=m\int_0^\infty\frac{\mathrm{e}^{\alpha y}-1}{\alpha}\mathrm{e}^{-\alpha y}\,\mathrm{d}G(y)=\frac{m}{\alpha}\left(1-\frac{1}{m}\right)<\infty.\end{aligned}\tag{2.7}$$

Therefore, (2.6) follows from (2.3). □

Corollary 2.1. *Assume the setup of Lemma 2.2. Let* $\mu_\alpha \equiv m\int_0^\infty x\mathrm{e}^{-\alpha x}\,\mathrm{d}G(x)$. *Then, for all* $\varepsilon > 0$,

$$\lim_{t\to\infty} \mathrm{e}^{-\alpha t} E\left(m^{N(t)} : \left|\frac{N(t)}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right) = 0. \tag{2.8}$$

Proof. By the definition of R_t,

$$\begin{aligned} &\mathrm{e}^{-\alpha t} E\left(m^{N(t)}; \left|\frac{N(t)}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right) \\ &\quad = E\left(\mathrm{e}^{-\alpha(t-S_{N(t)+1})}\mathrm{e}^{-\alpha S_{N(t)+1}} m^{N(t)} I\left(\left|\frac{N(t)}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right)\right) \\ &\quad = \frac{1}{m}\sum_{k=0}^{\infty} E\left(\mathrm{e}^{\alpha R_t}\mathrm{e}^{-\alpha S_{k+1}} m^{k+1} I\left(N(t) = k, \left|\frac{k}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right)\right) \\ &\quad = \frac{1}{m}\sum_{k=0}^{\infty} E\left(\mathrm{e}^{\alpha \tilde{R}_t} I\left(\tilde{N}(t) = k, \left|\frac{k}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right)\right) \\ &\quad = \frac{1}{m} E\left(\mathrm{e}^{\alpha \tilde{R}_t}; \left|\frac{\tilde{N}(t)}{t} - \frac{1}{\mu_\alpha}\right| > \varepsilon\right), \end{aligned} \tag{2.9}$$

where the second-to-last equality follows from Lemma 2.2(ii). Since $\tilde{X}_1$ satisfies the hypothesis of Lemma 2.1(ii), the family of random variables $\{\mathrm{e}^{\alpha\tilde{R}_t} : t \geq 0\}$ is uniformly integrable. Also, by Lemma 2.1(i), $\frac{\tilde{N}(t)}{t} \to \frac{1}{\mu_\alpha}$ with probability 1. Therefore, $E(\mathrm{e}^{\alpha\tilde{R}_t}; |\frac{\tilde{N}(t)}{t} - \frac{1}{\mu_\alpha}| > \varepsilon) \to 0$ as $t \to \infty$. □

3. Results on branching processes

Let $\{Z_t : t \geq 0\}$ be an age-dependent branching process with offspring distribution $\{p_k\}_{k\geq 0}$ and lifetime distribution G (see [2] for details). Let $\{\zeta_k\}_{k\geq 0}$ be the embedded discrete-time Galton–Watson branching process with ζ_k being the size of the kth generation, $k \geq 0$. Since $p_0 = 0$, $P(Z_t > 0) = 1$ for all $t > 0$. On this event, choose an individual uniformly from those alive at time t. Let M_t be the generation number and a_t the age of this individual.

Proposition 3.1. *Let* $\alpha > 0$ *be the Malthusian parameter defined by* $m\int_0^\infty \mathrm{e}^{-\alpha u}\,\mathrm{d}G(u) = 1$.

(i) *For* $x \geq 0$,

$$\lim_{t\to\infty} P(a_t \leq x) = \frac{\int_0^x \mathrm{e}^{-\alpha u}(1 - G(u))\,\mathrm{d}u}{\int_0^\infty \mathrm{e}^{-\alpha u}(1 - G(u))\,\mathrm{d}u} \equiv \tilde{A}(x). \tag{3.1}$$

(ii) *Let* $\mu_\alpha = m\int_0^\infty x\mathrm{e}^{-\alpha x}\,\mathrm{d}G(x)$. *Then*

$$\frac{M_t}{t} \xrightarrow{\text{a.e.}} \frac{1}{\mu_\alpha} \tag{3.2}$$

as $t \to \infty$.

Proof. (i) Fix $x \geq 0$. Then $P(a_t \leq x) = E(A(t,x))$, where $A(t,x)$ is the proportion of individuals alive at time t with age less than or equal to x. From [1], it is known that $A(t,x) \xrightarrow{p} \tilde{A}(x)$. By the bounded convergence theorem, the result follows.

(ii) Let $\{\zeta_k\}_{k\geq 0}$ be the embedded Galton–Watson process. For each $t > 0$ and $k \geq 1$, let ζ_{kt} denote the number of lines of descent in the kth generation alive at time t (i.e., the successive lifetimes $\{L_i\}_{i\geq 1}$ of the individuals in that line of descent satisfying $\sum_{i=1}^{k} L_i \leq t \leq \sum_{i=1}^{k+1} L_i$). Denote the lines of descent of these individuals by $\{\zeta_{ktj} : 1 \leq j \leq \zeta_{kt}\}$. Now, for any $k > 0$, $\eta > 0$, let $N(t)$ be as in Lemma 2.1. Then

$$P(M_t > kt) = E\left(\frac{1}{Z_t}\sum_{j>kt}\zeta_{jt}\right) \leq P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{1}{\eta\mathrm{e}^{\alpha t}}\sum_{j>kt} E\zeta_{jt}.$$

However, $E\zeta_{jt} = m^j P(N(t) = j)$. Hence,

$$\begin{aligned} P(M_t > kt) &= P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{1}{\eta}\mathrm{e}^{-\alpha t}\sum_{j>kt} m^j P\big(N(t) = j\big) \\ &= P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{1}{\eta}\mathrm{e}^{-\alpha t} E\big(m^{N(t)}; N(t) > kt\big) \\ &= P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{1}{\eta m} E\big(\mathrm{e}^{\alpha\tilde{R}_t}; \tilde{N}(t) > kt\big) \\ &\leq P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{\mathrm{e}^{\alpha l}}{\eta m} P\big(\tilde{N}(t) > kt\big) + \frac{1}{\eta m} E(\mathrm{e}^{\alpha\tilde{R}_t}; \tilde{R}_t > l), \end{aligned} \tag{3.3}$$

where the second-to-last step follows from an argument similar to that used in the proof of Corollary 2.1. Now, by Lemma 2.1(i), $\limsup_{t\to\infty} P(\tilde{N}(t) > kt) = 0$ if $k > \frac{1}{\mu_\alpha}$, and by Lemma 2.1(ii), $\limsup_{l\to\infty}\limsup_{t\to\infty} E(\mathrm{e}^{\alpha R_t} : R_t > l) = 0$. Also, as $\sum_{k=1}^{\infty} k\log(k)p_k < \infty$ (see [2], Chapter 4), there exists a random variable W such that $Z_t\mathrm{e}^{-\alpha t} \xrightarrow{\text{a.e.}} W$ as $t \to \infty$ and $P(W < \eta) \to 0$ as $\eta \downarrow 0$. Consequently, from (3.3), for $k > \frac{1}{\mu_\alpha}$, we obtain

$$\begin{aligned} \limsup_{t\to\infty} P(M_t > kt) \leq{} & \limsup_{\eta\downarrow 0}\limsup_{t\to\infty} P(Z_t < \mathrm{e}^{\alpha t}\eta) + \frac{\mathrm{e}^{\alpha l}}{\eta m}\limsup_{t\to\infty} P\big(\tilde{N}(t) > kt\big) \\ & + \frac{1}{\eta m}\limsup_{l\to\infty}\limsup_{t\to\infty} E(\mathrm{e}^{\alpha\tilde{R}_t} : \tilde{R}_t > l) = 0. \end{aligned}$$

Similarly, one can show that $\limsup_{t\to\infty} P(M_t < kt) \to 0$ for $k < \frac{1}{\mu_\alpha}$, thereby obtaining the result. This result has also been proven in [9], using a different method. $\square$

Proposition 3.2 (Law of large numbers). *Let $\varepsilon > 0$ be given. For the randomly chosen individual at time t, let $\{L_{ti} : 1 \le i \le M_t\}$ be the lifetimes of its ancestors. Let $h : [0, \infty) \to \mathbb{R}$ be Borel measurable such that $m \int_0^\infty |h(x)| \mathrm{e}^{-\alpha x} \, \mathrm{d}G(x) < \infty$. Then, as $t \to \infty$,*

$$P\left(\left| \frac{1}{M_t} \sum_{i=1}^{M_t} h(L_{ti}) - c \right| > \varepsilon \right) \to 0, \tag{3.4}$$

where $c = m \int_0^\infty h(x) \mathrm{e}^{-\alpha x} \, \mathrm{d}G(x)$.

Proof. Let ζ_k, ζ_{kjt} be as in proof of Proposition 3.1(ii). We call ζ_{ktj} *bad* if

$$\left| \frac{1}{k} \sum_{i=1}^{k} h(L_{ktji}) - c \right| > \varepsilon, \tag{3.5}$$

where $\{L_{ktji}\}_{i \ge 1}$ are the successive lifetimes in the line of descent ζ_{ktj} starting from the ancestor. Let $\zeta_{kt,b}$ denote the cardinality of the set $\{\zeta_{ktj} : 1 \le j \le \zeta_{kt}$ and ζ_{ktj} is bad$\}$. Then

$$\begin{aligned} &P\left(\left| \frac{1}{M_t} \sum_{i=1}^{M_t} h(L_{ti}) - c \right| > \varepsilon \right) \\ &\quad = E\left(\frac{\sum_{j=0}^\infty \zeta_{jt,b}}{Z_t} \right) = E\left(\frac{\sum_{j=0}^\infty \zeta_{jt,b}}{Z_t}; Z_t < \mathrm{e}^{\alpha t} \eta \right) + E\left(\frac{\sum_{j=0}^\infty \zeta_{jt,b}}{Z_t}; Z_t \ge \mathrm{e}^{\alpha t} \eta \right) \\ &\quad \le P(Z_t < \mathrm{e}^{\alpha t} \eta) + \frac{1}{\eta \mathrm{e}^{\alpha t}} E\left(\sum_{j=0}^\infty \zeta_{jt,b} \right), \end{aligned} \tag{3.6}$$

where $\eta > 0$. Now, note that for $j \ge 0$ and $t \ge 0$,

$$\begin{aligned} E(\zeta_{jtb}) &= m^j P\left(\sum_{i=1}^{j} L_i \le t < \sum_{i=1}^{j+1} L_i, \left| \frac{\sum_{i=1}^j h(L_i)}{j} - c \right| > \varepsilon \right) \\ &= E\big(m^j I(S_j \le t < S_{j+1}, |\bar{Y}_j| > \varepsilon) \big), \end{aligned}$$

where $\{L_i\}$ are i.i.d. G, $S_j = \sum_{i=1}^j L_i$ and $\bar{Y}_j = \frac{\sum_{i=1}^j h(L_i)}{j} - c$. Thus,

$$\begin{aligned} &\sum_{j=0}^\infty \frac{1}{\eta} \mathrm{e}^{-\alpha t} E(\zeta_{jt,b}) \\ &\quad = \frac{1}{\eta} \sum_{j=0}^\infty E\big(\mathrm{e}^{-\alpha t} m^j I(S_j \le t < S_{j+1}, |\bar{Y}_j| > \varepsilon) \big) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\eta m} \sum_{j=0}^{\infty} E\big(\mathrm{e}^{\alpha R_t} \mathrm{e}^{-\alpha S_{j+1}} m^{j+1} I(S_j \le t < S_{j+1}, |\bar{Y}_j| > \varepsilon)\big) \\
&= \frac{1}{\eta m} \sum_{j=0}^{\infty} E\big(\mathrm{e}^{\alpha \tilde{R}_t} I\big(\tilde{N}(t) = j, |\bar{\tilde{Y}}_j| > \varepsilon\big)\big) \qquad \text{(by Lemma 2.2(ii))} \\
&= \frac{1}{\eta m} E\big(\mathrm{e}^{\alpha \tilde{R}_t} I\big(\big|\bar{\tilde{Y}}_{\tilde{N}(t)}\big| > \varepsilon\big)\big) \\
&\le \frac{\mathrm{e}^{\alpha l}}{\eta m} P\big(\big|\bar{\tilde{Y}}_{\tilde{N}(t)}\big| > \varepsilon\big) + \frac{1}{\eta m} E(\mathrm{e}^{\alpha \tilde{R}_t}; \tilde{R}_t > l).
\end{aligned}
$$

By the strong law of large numbers, $|\bar{\tilde{Y}}_{\tilde{N}(t)}| \xrightarrow{\text{a.e.}} 0$ and, consequently, $\limsup_{t\to\infty} P(|\bar{\tilde{Y}}_{\tilde{N}(t)}| > \varepsilon) = 0$. Applying (2.3) with $\theta = \alpha$ shows that the second term on the right-hand side of (3.6) converges to zero as $t \to \infty$. Further, since $\limsup_{\eta\to 0} \limsup_{t\to\infty} P(Z_t < \mathrm{e}^{\alpha t}\eta) = 0$, (3.4) follows. □

Proposition 3.3. *Assume that (1.6) holds. Let $\{L_i\}_{i\ge 1}$ be i.i.d. G and $\{\eta_i\}_{i\ge 1}$ be i.i.d. copies of η and independent of the $\{L_i\}_{i\ge 1}$. For $\theta \in \mathbb{R}$, $t \ge 0$, define $\phi(\theta, t) = E\mathrm{e}^{\mathrm{i}\theta\eta(t)}$. There then exists an event D with $P(D) = 1$, and on D for all $\theta \in \mathbb{R}$,*

$$
\prod_{j=1}^{n} \phi\left(\frac{\theta}{\sqrt{n}}, L_j\right) \to \mathrm{e}^{-\theta^2\psi/2} \qquad \textit{as } n \to \infty,
$$

where $\Psi = \int_0^\infty v(s) G(\mathrm{d}s)$.

Proof. Recall from (1.6) that $v(t) = E(\eta^2(t))$ for $t \ge 0$. Consider

$$
X_{ni} = \frac{\eta_i(L_i)}{\sqrt{\sum_{j=1}^{n} v(L_j)}} \qquad \text{for } 1 \le i \le n
$$

and $\mathcal{F} = \sigma(L_i : i \ge 1)$. Given $\mathcal{F}$, $\{X_{ni} : 1 \le i \le n\}$ is a triangular array of independent random variables such that for $1 \le i \le n$, $E(X_{ni}|\mathcal{F}) = 0$, $\sum_{i=1}^{n} E(X_{ni}^2|\mathcal{F}) = 1$.

Let $\varepsilon > 0$ be given. Let

$$
L_n(\varepsilon) = \sum_{i=1}^{n} E(X_{ni}^2; X_{ni}^2 > \varepsilon|\mathcal{F}).
$$

By the strong law of large numbers,

$$
\frac{\sum_{j=1}^{n} v(L_j)}{n} \to \psi \qquad \text{w.p. 1.} \tag{3.7}
$$

Let D be the event on which (3.7) holds. Then, on D,

$$\begin{aligned}
\limsup_{n\to\infty} L_n(\varepsilon) &\le \limsup_{n\to\infty} \frac{2}{\psi n}\sum_{i=1}^{n} E\left(\left|\eta_i(L_i)|^2 : |\,\eta_i(L_i)|^2 > \frac{\varepsilon n\psi}{2}\right|\mathcal{F}\right)\\
&\le \limsup_{k\to\infty}\limsup_{n\to\infty}\frac{2}{\psi n}\sum_{i=1}^{n} E\big(|\eta_i(L_i)|^2 : |\,\eta_i(L_i)|^2 > k|\mathcal{F}\big)\\
&= \frac{2}{\psi}\limsup_{k\to\infty} E\big(|\eta_1(L_1)|^2 : |\,\eta_1(L_1)\,|^2 > k\big) \qquad \text{(by SLLN)}\\
&= 0.
\end{aligned}$$

Thus, the Lindeberg–Feller central limit theorem implies that on D, for all $\theta \in \mathbb{R}$,

$$\prod_{i=1}^{n} \phi\left(\frac{\theta}{\sqrt{\sum_{j=1}^{n} v(L_j)}}, L_j\right) = E(\mathrm{e}^{\mathrm{i}\theta\sum_{j=1}^{n} X_{nj}}|\mathcal{F}) \to \mathrm{e}^{-\theta^2/2}.$$

Combining this with (3.7) yields the result. □

Proposition 3.4. *For the randomly chosen individual at time t, let $\{L_{ti}, \{\eta_{ti}(u) : 0 \le u \le L_{ti}\} : 1 \le i \le M_t\}$ be the lifetimes and motion processes of its ancestors. Let $Z_{t1} = \frac{1}{\sqrt{M_t}}\sum_{i=1}^{M_t}\eta_{ti}(L_{ti})$ and $\mathcal{L}_t = \sigma\{M_t, L_{ti} : 1 \le i \le M_t\}$. Then*

$$E\big(|E(\mathrm{e}^{\mathrm{i}\theta Z_{t1}}|\mathcal{L}_t) - \mathrm{e}^{-\theta^2\psi_\alpha/2}|\big) \to 0 \qquad \textit{as } t \to \infty. \tag{3.8}$$

Proof. Fix $\theta \in \mathbb{R}$, $\varepsilon_1 > 0$ and $\varepsilon > 0$. Replace the definition of "bad" in (3.5) by

$$\left|\prod_{i=1}^{k}\phi\left(\frac{\theta}{\sqrt{k}}, L_{ktji}\right) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\right| > \varepsilon. \tag{3.9}$$

By Proposition 3.3, we have, if the $\{L_i\}_{i\ge1}$ are i.i.d. $\tilde{G}$ (as in Lemma 2.2),

$$\lim_{k\to\infty} P\left(\sup_{j\ge k}\left|\prod_{i=1}^{j}\phi\left(\frac{\theta}{\sqrt{j}}, L_i\right) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\right| > \varepsilon\right) = 0. \tag{3.10}$$

Using this and following along the lines of the proof of Proposition 3.2 (since the details mirror that proof we avoid repeating them here), we obtain that for t sufficiently large,

$$P\left(\left|\prod_{i=1}^{M_t}\phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\right| > \varepsilon_1\right) < \varepsilon. \tag{3.11}$$

Now, for all $\theta \in \mathbb{R}$,

$$E(\mathrm{e}^{\mathrm{i}\theta Z_{t1}} | \mathcal{L}_t) = \prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right).$$

Therefore,

$$\begin{aligned}
&\limsup_{t\to\infty} E\big(\big|E\big(\mathrm{e}^{\mathrm{i}\theta(1/\sqrt{M_t})\sum_{i=1}^{M_t}\eta_i(L_{ti})}\big|\mathcal{L}_t\big) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\big|\big) \\
&= \limsup_{t\to\infty} E\left(\left|\prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\right|\right) \\
&< \varepsilon_1 + 2\limsup_{t\to\infty} P\left(\left|\prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - \mathrm{e}^{-\theta^2\psi_\alpha/2}\right| > \varepsilon_1\right) \\
&= \varepsilon_1 + 2\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$, $\varepsilon_1 > 0$ are arbitrary, we have the result. □

The above propositions will be used in the proof of Theorem 1.1. For the proof of Theorem 1.2, we will need a result on coalescing times of the lines of descent.

Theorem 3.1. *For $t \geq 0$, choose two individuals uniformly at random from those alive at time t and trace their lines of descents backward in time to find the time of death τ_t of their last common ancestor, also known as the coalescent time. Then, for $0 < s < \infty$,*

$$\lim_{t\to\infty} P(\tau_t < s) = H(s) \quad \textit{exists and} \quad \lim_{s\to\infty} H(s) = 1. \tag{3.12}$$

Proof. For $s \geq 0$ and $t \geq s$, let $\{Z_{t-s,i} : t \geq s\}$, $i = 1, 2, \ldots, Z_s$, denote the branching processes initiating from the Z_s individuals at time s. Then

$$P(\tau_t < s) = E\left(\frac{\sum_{i\neq j=1}^{Z_s} Z_{t-s,i} Z_{t-s,j}}{Z_t(Z_t - 1)}\right). \tag{3.13}$$

As $\sum_{k=1}^{\infty} k\log(k)p_k < \infty$ (see [2], Chapter 4), conditional on Z_s, $\mathrm{e}^{-\alpha(t-s)}Z_{t-s,i} \xrightarrow{\text{a.e.}} W_i\mathrm{e}^{\alpha a_{s,i}}$ as $t \to \infty$ for all $i = 1, \ldots, Z_s$, where $\{W_i\}_{i\geq 1}$ is a sequence of i.i.d. random variables with $W_i < \infty$ a.e. and $E(W_i) = 1$. $\{a_{s,i}\}_{i=1,\ldots,Z_s}$ are the ages of the individuals alive at time s. Hence,

$$\frac{\sum_{i\neq j=1}^{Z_s} Z_{t-s,i} Z_{t-s,j}}{Z_t(Z_t - 1)} \xrightarrow{\text{a.e.}} \frac{\sum_{i\neq j=1}^{Z_s} W_i W_j \mathrm{e}^{\alpha a_{s,i}} \mathrm{e}^{\alpha a_{s,j}}}{(\sum_{i=1}^{Z_s} W_i \mathrm{e}^{\alpha a_{s,i}})^2}. \tag{3.14}$$

Define the function $g : \mathbb{N} \to \mathbb{N}$ by

$$g(k) = E\left(\frac{\sum_{i\neq j=1}^{k} W_i W_j}{(\sum_{i=1}^{k} W_i)^2}\right) = 1 - E\left(\frac{\sum_{i=1}^{k} W_i^2}{(\sum_{i=1}^{k} W_i)^2}\right).$$

Since $\sum_{i=1}^k W_i^2 \le (\sum_{i=1}^k W_i)^2$, we have $0 \le g(k) \le 1$. So, by the bounded convergence theorem,

$$E\left(\frac{\sum_{i\neq j=1}^{Z_s} Z_{t-s,i} Z_{t-s,j}}{Z_t(Z_t-1)}\middle| Z_s\right) \xrightarrow{\text{a.e.}} g(Z_s).$$

It follows from (3.13) and the bounded convergence theorem that

$$\lim_{t\to\infty} P(\tau_t < s) = E(g(Z_s)) \equiv H(s).$$

We now claim that $g(k) \to 1$ as $k \to \infty$. Again, by the bounded convergence theorem, it suffices to show that

$$\frac{\sum_{i=1}^k W_i^2}{(\sum_{i=1}^k W_i)^2} \xrightarrow{p} 0.$$

Since $\sum_{j=1}^\infty j \log j p_j < \infty$, $E(W_i) \equiv 1$, and hence, by the strong law,

$$\frac{1}{k}\sum_{i=1}^k W_i \to 1 \qquad \text{a.e.} \tag{3.15}$$

Consequently, it suffices to show that

$$\frac{1}{k^2}\sum_{i=1}^k W_i^2 \xrightarrow{p} 0. \tag{3.16}$$

Since $E(W_i) = 1$ and the $\{W_i\}_{i\ge 1}$ are i.i.d., for every $\varepsilon > 0$, $\sum_{i=1}^\infty P(W_i \ge i\varepsilon) < \infty$. So, by the Borel–Cantelli lemma, with probability 1, $W_i \le i\varepsilon$ for all large i. This implies that if $W_k^* \equiv \max_{1\le i\le k} W_i$, then

$$\limsup_{k\to\infty} \frac{W_k^*}{k} \le \varepsilon \qquad \text{w.p. 1.}$$

This, together with (3.15), yields

$$\limsup_{k\to\infty} \frac{1}{k^2}\sum_{i=1}^k W_i^2 \le \varepsilon \qquad \text{w.p. 1.}$$

(3.16) follows from the above and the fact that the empirical age distribution converges. Since $Z_s \xrightarrow{\text{a.e.}} \infty$ as $s \to \infty$, it follows by the bounded convergence theorem that $\lim_{s\to\infty} H(s) = 1$. $\square$

A similar argument to the above leads to the following corollary. This, however, is not required for the proof of Theorem 1.2.

Corollary 3.1. *Suppose r individuals are chosen at time t by simple random sampling without replacement. Let $\tau_{r,t}$ be the last time they have a common ancestor. Then*

$$\lim_{t\to\infty} P(\tau_{r,t} < s) = H_r(s) \quad \textit{exists and} \quad \lim_{s\to\infty} H_r(s) = 1. \tag{3.17}$$

4. Proof of Theorem 1.1

For the individual chosen, let (a_t, X_t) be the age and position at time t. As in Proposition 3.4, let $\{L_{ti}, \{\eta_{ti}(u), 0 \le u \le L_{ti}\} : 1 \le i \le M_t\}$ be the lifetimes and motion processes of the ancestors of this individual and $\{\eta_{t(M_t+1)}(u) : 0 \le u \le t - \sum_{i=1}^{M_t} L_{ti}\}$ the motion of this individual. Let $\mathcal{L}_t = \sigma(M_t, L_{ti}, 1 \le i \le M_t)$. It is immediate from the construction of the process that

$$a_t = t - \sum_{i=1}^{M_t} L_{ti}$$

whenever $M_t > 0$, that it is equal to $a_0 + t$ otherwise and that

$$X_t = X_0 - \eta_{t1}(a_0) + \sum_{i=1}^{M_t} \eta_{ti}(L_{ti}) + \eta_{t(M_t+1)}(a_t).$$

Rearranging the terms, we obtain

$$\left(a_t, \frac{X_t}{\sqrt{t}}\right) = \left(a_t, \sqrt{\frac{1}{\mu_\alpha}} Z_{t1}\right) + \left(0, \left(\sqrt{\frac{M_t}{t}} - \sqrt{\frac{1}{\mu_\alpha}}\right) Z_{t1}\right) + \left(0, \frac{X_0 + \eta_{t1}(a_0)}{\sqrt{t}} + Z_{t2}\right),$$

where $Z_{t1} = \frac{\sum_{i=1}^{M_t} \eta_{ti}(L_{t_i})}{\sqrt{M_t}}$ and $Z_{t2} = \frac{1}{\sqrt{t}} \eta_{t(M_t+1)}(a_t)$. Let $\varepsilon > 0$ be given,

$$\begin{aligned} P(|Z_{t2}| > \varepsilon) &\le P(|Z_{t2}| > \varepsilon, a_t \le k) + P(|Z_{t2}| > \varepsilon, a_t > k) \\ &\le P(|Z_{t2}| > \varepsilon, a_t \le k) + P(a_t > k) \\ &\le \frac{E(|Z_{t2}|^2 I(a_t \le k))}{\varepsilon^2} + P(a_t > k). \end{aligned}$$

By Proposition 3.1(i), a_t converges in distribution to U and hence, for any $\eta > 0$, there is a k_η such that for all $k \ge k_\eta$, $t \ge 0$,

$$P(a_t > k) < \frac{\eta}{2}.$$

Next,

$$\begin{aligned} E\big(|Z_{t2}|^2 I(a_t \le k_\eta)\big) = E\big(I(a_t \le k_\eta) E(|Z_{t2}|^2 | \mathcal{L}_t)\big) &= E\left(I(a_t \le k_\eta) \frac{v(a_t)}{t}\right) \\ &\le \frac{\sup_{u \le k_\eta} v(u)}{t}. \end{aligned}$$

Hence,

$$P(|Z_{t2}| > \varepsilon) \le \frac{\sup_{u \le k_\eta} v(u)}{t\varepsilon^2} + \frac{\eta}{2}.$$

Using (1.6), since $\varepsilon > 0$ and $\eta > 0$ are arbitrary, this shows that as $t \to \infty$,

$$Z_{t2} \xrightarrow{p} 0. \tag{4.1}$$

Now, for $\lambda > 0$, $\theta \in \mathbb{R}$, as a_t is $\mathcal{L}_t$ measurable, we have

$$\begin{aligned} &E\big(\mathrm{e}^{-\lambda a_t}\mathrm{e}^{-\mathrm{i}(\theta/\sqrt{\mu_\alpha})Z_{t1}}\big) \\ &\quad = E\big(\mathrm{e}^{-\lambda a_t}\big(E\big(\mathrm{e}^{-\mathrm{i}(\theta/\sqrt{\mu_\alpha})Z_{t1}}\big|\mathcal{L}_t\big) - \mathrm{e}^{-\theta^2\psi_\alpha/(2\mu_\alpha)}\big)\big) + \mathrm{e}^{-\theta^2\psi_\alpha/(2\mu_\alpha)}E(\mathrm{e}^{-\lambda a_t}). \end{aligned}$$

Proposition 3.4 shows that the first term above converges to zero and, using Proposition 3.1(i), we can conclude that as $t \to \infty$,

$$\left(a_t, \frac{1}{\sqrt{\mu_\alpha}}Z_{t1}\right) \xrightarrow{d} (U, V). \tag{4.2}$$

As X_0, a_0 are constants, by Proposition 3.1, (4.2), (4.1) and Slutsky's theorem, the proof is complete.

5. Proof of Theorem 1.2

Let $\phi \in C_b(\mathbb{R} \times \mathbb{R}_+)$. By Theorem 1.1 and the bounded convergence theorem,

$$E(\langle \tilde{Y}_t, \phi\rangle) \to E(\phi(U, V)). \tag{5.1}$$

We shall show that

$$\mathrm{Var}(\langle \tilde{Y}_t, \phi\rangle) \equiv E(\langle \tilde{Y}_t, \phi\rangle^2) - (E(\langle \tilde{Y}_t, \phi\rangle))^2 \to 0 \tag{5.2}$$

as $t \to \infty$. This will yield that

$$\langle \tilde{Y}_t, \phi\rangle \xrightarrow{d} \phi(U, V)$$

for all $\phi \in C_b(\mathbb{R} \times \mathbb{R}_+)$. The result then follows from [8], Theorem 16.16. We now proceed to establish (5.2).

Pick two individuals C_1, C_2 at random (i.e., by simple random sampling without replacement) from those alive at time t. Let the age and position of the two individuals be denoted by (a_t^i, X_t^i), $i = 1, 2$. Let $\tau_t = \tau_{C_1,C_2,t}$ be the birth time of their common ancestor, say D, whose position we denote by $\tilde{X}_{\tau_t}$. Let the net displacement of C_1 and C_2 from D be denoted by $X_{t-\tau_t}^i$, $i = 1, 2$, respectively. Then $X_t^i = \tilde{X}_{\tau_t} + X_{t-\tau_t}^i$, $i = 1, 2$.

Next, conditioned on this history up to the birth of $D(\equiv \mathcal{G}_t)$, the random variables $(a_t^i, X_{t-\tau_t}^i)$, $i = 1, 2$, are independent. By Theorem 3.1, $\frac{\tau_t}{t} \xrightarrow{d} 0$. Also, by Theorem 1.1, conditioned on $\mathcal{G}_t$, $\{(a_t^i, \frac{X_{t-\tau_t}^i}{\sqrt{t-\tau_t}}), i = 1, 2\}$ converges in distribution to $\{(U_i, V_i), i = 1, 2\}$, which are i.i.d. with distribution (U, V), as in Theorem 1.1. Next, by Theorem 3.1, $\{\tau_t : t \geq 0\}$ is tight and hence $\frac{\tilde{X}_{\tau_t}}{\sqrt{t}} \xrightarrow{p} 0$ as $t \to \infty$.

Combining these, we conclude that $\{(a_t^i, \frac{X_t^i}{\sqrt{t}}), i = 1, 2\}$ converges in distribution to $\{(U_i, V_i), i = 1, 2\}$. Thus, for any $\phi \in C_b(\mathbb{R}_+ \times \mathbb{R})$, we have, by the bounded convergence theorem,

$$\lim_{t\to\infty} E\left(\prod_{i=1}^{2} \phi\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right)\right) = E\prod_{i=1}^{2} \phi(U_i, V_i) = (E\phi(U, V))^2.$$

Now,

$$E(\tilde{Y}_t(\phi))^2 = E\left(\frac{(\phi(a_t, X_t/\sqrt{t}))^2}{Z_t}\right) + E\left(\prod_{i=1}^{2} \phi\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right)\frac{Z_t(Z_t - 1)}{Z_t^2}\right).$$

Using the fact that ϕ is bounded and $Z_t \xrightarrow{\text{a.e.}} \infty$, we have

$$\lim_{t\to\infty} E(\tilde{Y}_t(\phi))^2 \to (E\phi(U, V))^2.$$

This, along with (5.1), implies (5.2) and we have thus proven the theorem.

Acknowledgement

Siva Athreya's research was supported in part by CSIR Grant in Aid Scheme and Homi Bhabha Fellowship. Srikanth Iyer's research was supported in part by UGC-SAP IV.

References

[1] Athreya, K.B. and Kang, H.J. (1998). Some limit theorems for positive recurrent Markov Chains I and II. *Adv. in Appl. Probab.* **30** 693–722. MR1663545

[2] Athreya, K.B. and Ney, P. (2004). *Branching Processes*. New York: Dover. MR2047480

[3] Biggins, J.D. and Kyprianou, A.E. (1997). Seneta–Heyde norming in the branching random walk. *Ann. Probab.* **25** 337–360. MR1428512

[4] Bingham, N.H. and Doney, R.A. (1975). Asymptotic properties of supercritical branching processes, II. Crump–Mode and Jirina processes. *Adv. in Appl. Probab.* **7** 66–82. MR0378125

[5] Durrett, R. and Liggett, T.M. (1983). Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete* **64** 275–301. MR0716487

[6] Jagers, P. and Nerman, O. (1984). The growth and composition of branching populations. *Adv. in Appl. Probab.* **16** 221–259. MR0742953

[7] Nerman, O. and Jagers, P. (1984). The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete* **65** 445–460. MR0731231
[8] Kallenberg, O. (2002). *Foundation of Modern Probability Theory*. New York: Springer. MR1876169
[9] Samuels, M. (1971). Distribution of the branching process population among generations. *J. Appl. Probab.* **8** 655–667. MR0292187

Received November 2008 and revised December 2009