CRITICAL AGE DEPENDENT BRANCHING MARKOV PROCESSES AND THEIR SCALING LIMITS

KRISHNA B. ATHREYA, SIVA R. ATHREYA, AND SRIKANTH K. IYER

Abstract. This paper studies: (i) the long time behaviour of the empirical distribution of age and normalised position of an age dependent critical branching Markov process conditioned on non-extinction; and (ii) the super-process limit of a sequence of age dependent critical branching Brownian motions.

1. Introduction

Consider an age dependent branching Markov process where i) each particle lives for a random length of time and during its lifetime moves according to a Markov process and ii) upon its death it gives rise to a random number of offspring. We assume that the system is critical, i.e. the mean of the offspring distribution is one.

We study three aspects of such a system. First, at time t, conditioned on non-extinction (as such systems die out w.p. 1) 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 2.1). Second, it is shown that conditioned on non-extinction at time t, the empirical distribution of the age and the normalised position of the population converges as $t \to \infty$ in law to a random measure characterised by its moments (See Theorem 2.2). Thirdly, we establish a super-process limit of such branching Markov processes where the motion is Brownian (See Theorem 2.4).

The rest of the paper is organised as follows. In Section 2.1 we define the branching Markov process precisely and in Section 2.2 we state the three main theorems of this paper and make some remarks on various possible generalisations of our results.

2000 *Mathematics Subject Classification.* Primary: 60G57 Secondary: 60H30.

Key words and phrases. Age dependent, Branching, Ancestoral times, Measure-valued, Empirical distribution.

In Section 3 we prove four propositions on age-dependent Branching processes which are used in proving Theorem 2.1 (See Section 4). In Section 3 we also show that the joint distribution of ancestoral times for a sample of $k \geq 1$ individuals chosen at random from the population at time t converges as $t \to \infty$ (See Theorem 3.5). This result is of independent interest and is a key tool that is needed in proving Theorem 2.2 (See Section 5).

In Section 6, we prove Theorem 2.4, the key idea being to scale the age and motion parameters differently. Given this, the proof uses standard techniques for such limits. Theorem 2.1 is used in establishing the limiting log-Laplace equation. Tightness of the underlying particle system is shown in Proposition 6.4 and the result follows by the method prescribed in [7].

2. Statement of Results

2.1. **The Model.**

Each particle in our system will have two parameters, age in $\mathbb{R}_+$ and location in $\mathbb{R}$. We begin with the description of the particle system.

(i) **Lifetime Distribution $G(\cdot)$:** Let $G(\cdot)$ be a cumulative distribution function on $[0, \infty)$, with $G(0) = 0$. Let $\mu = \int_0^\infty s dG(s) < \infty$.

(ii) **Offspring Distribution $\mathbf{p}$:** Let $\mathbf{p} \equiv \{p_k\}_{k \geq 0}$ be a probability distribution such that $p_0 < 1$, $m = \sum_{k=0}^\infty k p_k = 1$ and that $\sigma^2 = \sum_{k=0}^\infty k^2 p_k - 1 < \infty$.

(iii) **Motion Process $\eta(\cdot)$:** Let $\eta(\cdot)$ be a $\mathbb{R}$ valued Markov process starting at 0.

Branching Markov Process $(G, \mathbf{p}, \eta)$: Suppose we are given a realisation of an age-dependent branching process with offspring distribution $\mathbf{p}$ and lifetime distribution G (See Chapter IV of [5] for a detailed description). We construct a branching Markov process by allowing each individual to execute an independent copy of η during its lifetime τ starting from where its parent died.

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

$$\mathcal{C}_t = \{(a_t^i, X_t^i) : i = 1, 2, \ldots, N_t\} \tag{2.1}$$

denote the age and position configuration of all the individuals alive at time t. Since $m = 1$ and $G(0) = 0$, there is no explosion in finite time (i.e. $P(N_t < \infty) = 1$) and consequently $\mathcal{C}_t$ is well defined for each $0 \leq t < \infty$ (See [5]).

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 Y_0 does so. 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

$$\langle Y_t, \phi\rangle = \int \phi\, dY_t = \sum_{i=1}^{N_t} \phi(a_t^i, X_t^i). \tag{2.2}$$

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

Note that $\mathcal{C}_t$ determines Y_t and conversely. The Laplace functional of Y_t, is given by

$$L_t\phi(a,x) := E_{a,x}[e^{-\langle\phi,Y_t\rangle}] \equiv E[e^{-\langle\phi,Y_t\rangle} \mid Y_0 = \delta_{a,x}]. \tag{2.3}$$

From the independence intrinsic in $\{Y_t : t \geq 0\}$, we have:

$$E_{\nu_1+\nu_2}[e^{-\langle\phi,Y_t\rangle}] = (E_{\nu_1}[e^{-\langle\phi,Y_t\rangle}])(E_{\nu_2}[e^{-\langle\phi,Y_t\rangle}]), \tag{2.4}$$

for any $\nu_i \in M_a(\mathbb{R}_+\times\mathbb{R})$ where $E_{\nu_i}[e^{-\langle\phi,Y_t\rangle}] := E[e^{-\langle\phi,Y_t\rangle} \mid Y_0 = \nu_i]$ for $i = 1, 2$. This is usually referred to as the branching property of Y and can be used to define the process Y as the unique measure valued Markov process with state space $M_a(\mathbb{R}_+\times\mathbb{R})$ satisfying $L_{t+s}\phi(a,x) = L_t(L_s(\phi))(a,x)$ for all $t, s \geq 0$.

2.2. **The Results.**

In this section we describe the main results of the paper. Let A_t be the event $\{N_t > 0\}$, where N_t is the number of particles alive at time t. As $p_0 < 1$, $P(A_t) > 0$ for all $0 \leq t < \infty$ provided $P(N_0 = 0) \neq 1$.

Theorem 2.1. (Limiting behaviour of a randomly chosen particle) *On the event $A_t = \{N_t > 0\}$, let (a_t, X_t) be the age and position of a randomly chosen particle from those alive at time t. Assume that $\eta(\cdot)$ is such that for all $0 \leq t < \infty$*

$$E(\eta(t)) = 0, v(t) \equiv E(\eta^2(t)) < \infty, \sup_{0\leq s\leq t} v(s) < \infty, \tag{2.5}$$

$$\textit{and } \psi \equiv \int_0^\infty v(s)G(ds) < \infty.$$

Then, conditioned on A_t, $(a_t, \frac{X_t}{\sqrt{t}})$ 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 $(1-G(\cdot))$ and V is normally distributed with mean 0 and variance $\frac{\psi}{\mu}$.

Next consider the scaled empirical measure $\tilde{Y}_t \in M_a(\mathbb{R}_+ \times \mathbb{R})$ given by $\tilde{Y}_t(A \times B) = Y_t(A \times \sqrt{t}B)$, $A \in \mathcal{B}(\mathbb{R}_+), B \in \mathcal{B}(\mathbb{R})$.

Theorem 2.2. (Empirical Measure)
Assume (2.5). Then, conditioned on $A_t = \{N_t > 0\}$, the random measures $\{\frac{\tilde{Y}_t}{N_t}\}$ converges as $t \to \infty$ in distribution to a random measure ν, characterised by its moment sequence $m_k(\phi) \equiv E[\nu(\phi)^k]$, for $\phi \in C_b(\mathbb{R}_+ \times \mathbb{R})$, $k \geq 1$.

An explicit formula for $m_k(\phi)$ is given in (5.2) below.

Our third result is on the super-process limit. We consider a sequence of branching Markov processes $(G_n, \mathbf{p}_n, \eta_n)_{\{n \geq 1\}}$ denoted by $\{Y_t^n : t \geq 0\}_{\{n \geq 1\}}$ satisfying the following:

(a) **Initial measure:** For $n \geq 1$, $Y_0^n = \pi_{n\nu}$, where $\pi_{n\nu}$ is a Poisson random measure with intensity $n\nu$, for some $\nu = \alpha \times \mu \in M(\mathbb{R}_+ \times \mathbb{R})$.
(b) **Lifetime $G^n(\cdot)$:** For all $n \geq 1$, G^n is an exponential distribution with mean $\frac{1}{\lambda}$
(c) **Branching $\mathbf{p}_n, \cdot$:** For $n \geq 1$, Let $F_n(u) = \sum_{k=0}^{\infty} p_{n,k} u^k$ be the generating function of the offspring distribution $\mathbf{p}_n \equiv \{p_{n,k}\}_{k \geq 0}$. We shall assume that F_n satisfies,

$$\lim_{n\to\infty} \sup_{0 \leq u \leq N} \parallel n^2(F_n(1-u/n) - (1-u/n)) - u^2 \parallel \to 0, \tag{2.6}$$

for all $N > 0$.
(d) **Motion Process $\eta_n(\cdot)$:** For all $n \geq 1$,

$$\eta_n(t) = \frac{1}{\sqrt{n}} \int_0^t \sigma(u) dB(u), \qquad t \geq 0, \tag{2.7}$$

where $\{B(t) : t \geq 0\}$ is a standard Brownian motion starting at 0 and $\sigma : \mathbb{R}_+ \to \mathbb{R}$ is a continuous function such that $\int_0^\infty \sigma^2(s) dG(s) < \infty$. It follows that for each $n \geq 1$, η_n satisfies (2.5).

Definition 2.3. *Let $\mathcal{E}$ be an independent exponential random variable with mean $\frac{1}{\lambda}$, $0 < \lambda < \infty$. For $f \in C_l^+(\mathbb{R}_+ \times \mathbb{R})$ let $U_t f(x) = E(f(\mathcal{E}, x + \sqrt{\lambda\psi} B_t))$ where ψ is defined in (2.5). For $t \geq 0$, let $u_t(f)$ be the unique solution of* *the non linear integral equation*

$$u_t f(x) = U_t f(x) - \lambda \int_0^t U_{t-s}(u_s(f)^2)(x)ds. \tag{2.8}$$

Let $\{\mathcal{Y}_t : t \geq 0\}$ *be a* $M(\mathbb{R}_+ \times \mathbb{R})$ *valued Markov process whose Laplace functional is given by*

$$E_{\mathcal{E}\times\mu}[e^{-\langle f,\mathcal{Y}_t\rangle}] = e^{-\langle V_t f,\mu\rangle}, \tag{2.9}$$

where $f \in C_l^+(\mathbb{R}_+ \times \mathbb{R}^d)$ *(the set of all continuous functions from* $\mathbb{R}_+ \times \mathbb{R}$ *to* $\mathbb{R}$ *with finite limits as* $(a, x) \to \infty$*) and* $V_t(f)(x) \equiv u_t(f(x))$ *for* $x \in \mathbb{R}$ *(See* [7] *for existence of* $\mathcal{Y}$ *satisfying (2.9).*

Note that in the process $\{\mathcal{Y}_t : t \geq 0\}$ defined above, the distribution of the age (i.e. the first coordinate) is deterministic. The spatial evolution behaves like that of a super-process where the motion of particles is like that of a Brownian motion with variance equal to the average variance of the age-dependent particle displacement over its lifetime. Also, $u_s(f)$ in second term of (2.8) is interpreted in the natural way as a function on $\mathbb{R}_+ \times \mathbb{R}$ with $u_s(f)(a, x) = u_s(f)(x)$ for all $a > 0, x \in \mathbb{R}$.

Theorem 2.4. (Age Structured Super-process)
Let $\epsilon > 0$*. Let* $\{Y_t^n : t \geq 0\}$ *be the sequence of branching Markov processes defined above(i.e.in (a), (b), (c), (d)). Then as* $n \to \infty$*,* $\{\mathcal{Y}_t^n \equiv \frac{1}{n}Y_{nt}^n, t \geq \epsilon\}$ *converges weakly on the Skorokhod space* $D([\epsilon, \infty), M(\mathbb{R}_+ \times \mathbb{R}))$ *to* $\{\mathcal{Y}_t : t \geq \epsilon\}$*.*

2.3. **Remarks.**

(a) If $\eta(\cdot)$ is not Markov then $\tilde{C}_t = \{a_t^i, X_t^i, \tilde{\eta}_{t,i} \equiv \{\eta_{t,i}(u) : 0 \leq u \leq a_t^i\} : i = 1, 2 \ldots, N_t\}$ is a Markov process where $\{\tilde{\eta}_{t,i}(u) : 0 \leq u \leq a_t^i\}$ is the history of $\eta(\cdot)$ of the individual i during its lifetime. Theorem 2.1 and Theorem 2.2 extends to this case.

(b) Most of the above results also carry over to the case when the motion process is $\mathbb{R}^d$ valued ($d \geq 1$) or is Polish space valued and where the offspring distribution is age-dependent.

(c) Theorem 2.1 and Theorem2.2 can also be extended to the case when $\eta(L_1)$, with $L_1 \stackrel{d}{=} G$, is in the domain of attraction of a stable law of index $0 < \alpha \leq 2$.

(d) In Theorem 2.4 the convergence should hold on $D([0,\infty), M(\mathbb{R}_+ \times \mathbb{R}))$ if we take α in the sequence of branching Markov processes to be $\mathcal{E}$ (i.e. Exponential with mean $\frac{1}{\lambda}$).

(e) The super-process limit obtained in Theorem 2.4 has been considered in two special cases in the literature. One is in [6] where an age-dependent Branching process is rescaled (i.e. the particles do not perform any motion). The other is in [8] where a general non-local super-process limit is obtained when the offspring distribution is given by $p_1 = 1$. In our results, to obtain a super-process limit the age-parameter is scaled differently when compared to the motion parameter giving us an age-structured super-process.

(f) Limit theorems for critical branching Markov processes where the motion depends on the age does not seem to have been considered in the literature before.

3. Results on Branching Processes

Let $\{N_t : t \geq 0\}$ be an age-dependent branching process with offspring distribution $\{p_k\}_{k\geq 0}$ and lifetime distribution G (see [5] for detailed discussion). 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, $k \geq 0$. Let A_t be the event $\{N_t > 0\}$. On this event, choose an individual uniformly from those alive at time t. Let M_t be the generation number and a_t be the age of this individual.

Proposition 3.1. *Let A_t, a_t, M_t and N_t be as above. Let μ and σ be as in Section 2.1. Then*

$$(a) \qquad \lim_{t\to\infty} tP(A_t) = \frac{2\mu}{\sigma^2}$$

$$(b) \qquad \text{For all } x > 0,\ \lim_{t\to\infty} P(\frac{N_t}{t} > x | A_t) = e^{-\frac{2\mu x}{\sigma^2}},$$

$$(c) \qquad \text{For all } \epsilon > 0,\ \lim_{t\to\infty} P(|\frac{M_t}{t} - \frac{1}{\mu}| > \epsilon | A_t) = 0$$

$$(d) \qquad \text{For all } x > 0,\ \lim_{t\to\infty} P(a_t \leq x | A_t) = \frac{1}{\mu}\int_0^x (1 - G(s))ds.$$

Proof : For (a) and (b) see chapter 4 in [5]. For (c) see [9] and for (d) see [3]. □

Proposition 3.2. *(Law of large numbers) Let $\epsilon > 0$ be given. For the randomly chosen individual at time t, let $\{L_{ti} : 1 \leq i \leq M_t\}$, be the lifetimes* *of its ancestors. Let* $h : [0, \infty) \to \mathbb{R}$ *be Borel measurable and* $E(\mid h(L_1) \mid) < \infty$ *with* $L_1 \stackrel{d}{=} G$*. Then, as* $t \to \infty$

$$P(|\frac{1}{M_t}\sum_{i=1}^{M_t} h(L_{ti}) - E(h(L_1))| > \epsilon | A_t) \to 0.$$

Proof : Let ϵ and $\epsilon_1 > 0$ be given and let $k_1(t) = t(\frac{1}{\mu} - \epsilon)$ and $k_2(t) = t(\frac{1}{\mu} + \epsilon)$. By Proposition 3.1 there exists $\delta > 0$, $\eta > 0$ and $t_0 > 0$ such that for all $t \geq t_0$,

$$tP(N_t > 0) > \delta \text{ and } P(N_t \leq t\eta | A_t) < \epsilon_1; \tag{3.1}$$

$$P(M_t \in [k_1(t), k_2(t)]^c | A_t) < \epsilon_1. \tag{3.2}$$

Also by the law of large numbers for any $\{L_i\}_{i\geq 1}$ i.i.d. G with $E|h(L_1)| < \infty$

$$\lim_{k\to\infty} P(\sup_{j\geq k} \frac{1}{j} |\sum_{i=1}^{j} h(L_i) - E(h(L_1))| > \epsilon) = 0. \tag{3.3}$$

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 k-th generation alive at time t (i.e. the successive life times $\{L_i\}_{i\geq 1}$ 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 $\{\zeta_{ktj} : 1 \leq j \leq \zeta_{kt}\}$. Call ζ_{ktj} *bad* if

$$|\frac{1}{k}\sum_{i=1}^{k} h(L_{ktji}) - E(h(L_1)))| > \epsilon, \tag{3.4}$$

where $\{L_{ktji}\}_{i\geq 1}$ 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 $\{\zeta_{ktj} : 1 \leq$

$j \leq \zeta_{kt}$ and ζ_{ktj} is bad$\}$. Now,

$$
\begin{aligned}
& P(|\frac{1}{M_t}\sum_{i=1}^{M_t} h(L_{ti}) - E(h(L_1))| > \epsilon | A_t) \\
&= P(\text{ The chosen line of descent at time } t \text{ is } \textit{bad } | A_t) \\
&\leq P(\text{ The chosen line of descent at time } t \text{ is } \textit{bad}, M_t \in [k_1(t), k_2(t)]) | A_t) \\
&\quad + P(M_t \in [k_1(t)), k_2(t)]^c | A_t) \\
&= \frac{1}{P(N_t > 0)} E(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; A_t) + P(M_t \in [k_1(t)), k_2(t)]^c | A_t) \\
&= \frac{1}{P(N_t > 0)} E(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; N_t > t\eta) + \\
&\quad + \frac{1}{P(N_t > 0)} E(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; N_t \leq t\eta) + P(M_t \in [k_1(t)), k_2(t)]^c | A_t) \\
&\leq \frac{1}{P(N_t > 0)} E(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{t\eta}; N_t > t\eta) + \\
&\qquad + \frac{P(N_t \leq t\eta)}{P(N_t > 0)} + P(M_t \in [k_1(t)), k_2(t)]^c | A_t) \\
&= \frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b}) + \\
&\quad + P(N_t \leq t\eta | N_t > 0) + P(M_t \in [k_1(t)), k_2(t)]^c | A_t)
\end{aligned}
\tag{3.5}
$$

For $t \geq t_0$ by (3.2) and (3.3), the last two terms in (3.5) are less than ϵ_1. The first term is equal to

$$
\frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b}) \quad = \quad \frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\sum_{i=1}^{\zeta_j} 1_{\{\zeta_{jti} \text{ is bad.}\}})
$$

$$= \frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_j) \times$$
$$\times P\left(\sum_{i=1}^{j} L_i \le t < \sum_{i=1}^{j+1} L_i, \frac{1}{j}|\sum_{i=1}^{j} h(L_i) - E(h(L_1))| > \epsilon\right),$$

where the $\{L_i\}_{i\ge1}$ are i.i.d. G.

Using (3.1) and (since $m=1$) $E(\zeta_j) = E(\zeta_0)$ we can conclude that

$$\frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b})$$
$$\le E(\zeta_0)\frac{P(\sup_{j\ge k_1(t)} \frac{1}{j}|\sum_{i=1}^{j} h(L_i) - E(h(L_1))| > \epsilon)}{t\eta P(N_t > 0)}$$
$$\le E(\zeta_0)\frac{P(\sup_{j\ge k_1(t)} \frac{1}{j}|\sum_{i=1}^{j} h(L_i) - E(h(L_1))| > \epsilon)}{\eta\delta}, \tag{3.6}$$

which by (3.3) goes to zero. So we have shown that for $t \ge t_0$,

$$P(|\frac{1}{M_t}\sum_{i=1}^{M_t} h(L_{ti}) - E(h(L_1))| > \epsilon | A_t) < 3\epsilon_1.$$

Since $\epsilon_1 > 0$ is arbitrary, the proof is complete. □

Proposition 3.3. *Assume (2.5) holds. Let $\{L_i\}_{i\ge1}$ be i.i.d G and $\{\eta_i\}_{i\ge1}$ be i.i.d copies of η and independent of the $\{L_i\}_{i\ge1}$. For $\theta \in \mathbb{R}, t \ge 0$ define $\phi(\theta, t) = Ee^{i\theta\eta(t)}$. Then there exists an event D, with $P(D) = 1$ and on D for all $\theta \in \mathbb{R}$,*

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

where ψ is as in (2.5).

Proof: Recall from (2.5) that $v(t) = E(\eta^2(t))$ for $t \ge 0$. Consider

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

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 $\epsilon > 0$ be given. Let

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

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

$$\begin{aligned}
\limsup_{n\to\infty} L_n(\epsilon) &\leq \limsup_{n\to\infty} \frac{\psi}{2n}\sum_{i=1}^{n} E(|\eta_i(L_i)|^2) :| \eta_i(L_i)|^2 > \frac{\epsilon n\psi}{2}|\mathcal{F}) \\
&\leq \limsup_{k\to\infty} \frac{\psi}{2} E(|\eta_1(L_1)|^2 :| \eta_1(L_1) |^2 > k) \\
&= 0.
\end{aligned}$$

Thus the Linderberg-Feller Central Limit Theorem (see [4]) implies, that on D, for all $\theta \in \mathbb{R}$

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

Combining this with (3.7) yields the result. □

Proposition 3.4. *For the randomly chosen individual at time t, let $\{L_{ti}, \{\eta_{ti}(u) : 0 \leq u \leq L_{ti}\} : 1 \leq i \leq M_t\}$, 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 $\mathcal{L}_t = \sigma\{M_t, L_{ti} : 1 \leq i \leq M_t\}$. Then*

$$E\left(|E(e^{i\theta Z_{t1}}|\mathcal{L}_t) - e^{-\frac{\theta^2\psi}{2}}| \,|A_t\right) \to 0 \tag{3.8}$$

Proof: Fix $\theta \in \mathbb{R}, \epsilon_1 > 0$ and $\epsilon > 0$. Replace the definition of "bad" in (3.4) by

$$|\prod_{i=1}^{k} \phi(\frac{\theta}{\sqrt{k}}, L_{ktji}) - e^{-\frac{\theta^2\psi}{2}}| > \epsilon \tag{3.9}$$

By Proposition 3.3 we have,

$$\lim_{k\to\infty} P(\sup_{j\geq k} |\prod_{i=1}^{j} \phi(\frac{\theta}{\sqrt{j}}, L_i) - e^{-\frac{\theta^2\psi}{2}}| > \epsilon) = 0. \tag{3.10}$$

Using this in place of (3.3) and imitating the proof of Proposition 3.2, (since the details mirror that proof we avoid repeating them here), we obtain that for t sufficiently large

$$P(|\prod_{i=1}^{M_t}\phi(\frac{\theta}{\sqrt{M_t}},L_{ti})-e^{-\frac{\theta^2\psi}{2}}|>\epsilon_1|A_t)<\epsilon. \tag{3.11}$$

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

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

So,

$$\begin{aligned}
&\limsup_{t\to\infty}E(|E(e^{i\theta\frac{1}{\sqrt{M_t}}\sum_{i=1}^{M_t}\eta_i(L_{ti})}|\mathcal{L}_t)-e^{-\frac{\theta^2\psi}{2}}||A_t)\\
&= \limsup_{t\to\infty}E(|\prod_{i=1}^{M_t}\phi(\frac{\theta}{\sqrt{M_t}},L_{ti})-e^{-\frac{\theta^2\psi}{2}}||A_t)\\
&< \epsilon_1+2\limsup_{t\to\infty}P(|\prod_{i=1}^{M_t}\phi(\frac{\theta}{\sqrt{M_t}},L_{ti})-e^{-\frac{\theta^2\psi}{2}}|>\epsilon_1|A_t)\\
&= \epsilon_1+2\epsilon.
\end{aligned}$$

Since $\epsilon>0,\epsilon_1>0$ are arbitrary we have the result. □

The above four Propositions will be used in the proof of Theorem 2.1. For the proof of Theorem 2.2 we will need a result on coalescing times of the lines of descent.

Fix $k\geq 2$. On the event $A_t=\{N_t>0\}$, pick k individuals $C_1,C_2,\ldots,C_k$ from those alive at time t by simple random sampling without replacement. For any two particles C_i,C_j, let $\tau_{C_j,C_i,t}$ be the birth time of their most recent common ancestor. Let $\tau_{k-1,t}=\sup\{\tau_{C_j,C_i,t}: i\neq j, 1\leq i,j\leq k\}$. Thus $\tau_{k-1,t}$ is the first time there are $k-1$ ancestors of the k individuals $C_1,C_2,\ldots,C_k$. More generally, for $1\leq j\leq k-1$ let $\tau_{j,t}$ as the first time there are j ancestors of the k individuals $C_1,C_2,\ldots C_k$.

Theorem 3.5.

(i) *For any i,j, $\lim_{t\to\infty}P(\frac{\tau_{C_i,C_j,t}}{t}\leq x|A_t)\equiv H(x)$ exists for all $x\geq 0$ and $H(\cdot)$ is an absolutely continuous distribution function on $[0,\infty]$*

(ii) *Conditioned on A_t the vector $\tilde{\tau}_t=\frac{1}{t}(\tau_{j,t}:1\leq j\leq k-1)$ as $t\to\infty$ converges in distribution to a random vector $\tilde{T}=(T_1,\ldots,T_{k-1})$*

with $0 < T_1 < T_2 < \ldots < T_{k-1} < 1$ *and having an absolutely continuous distribution on* $[0,1]^{k-1}$.

Proof : The proof of (i) and (ii) for cases $k = 2, 3$ is in [9]. The following is an outline of a proof of (ii) for the case $k > 3$ (for a detailed proof see [3]).

Below, for $1 \leq j \leq k-1$, $\tau_{j,t}$ will be denoted by τ_j. It can be shown that it suffices to show that for any $1 \leq i_1 < i_2 \ldots < i_p < k$ and $0 < r_1 < r_2 < \ldots < r_p < 1$,

$$\lim_{t\to\infty} P(\frac{\tau_{i_1}}{t} < r_1 < \frac{\tau_{i_2}}{t} < r_2 < \ldots < \frac{\tau_{i_p}}{t} < r_p < \frac{\tau_{k-1}}{t} < r_{k-1} < 1|A_t)$$

exists. We shall now condition on the population size at time tr_1. Suppose that at time tr_1 there are n_{11} particles of which k_{11} have descendants that survive till time tr_2. For each $1 \leq j \leq k_{11}$, suppose there are n_{2j} descendants alive at time tr_2 and for each such j, let k_{2j} out of the n_{2j} have descendants that survive till time tr_3. Let $k_2 = (k_{21}, \ldots, k_{2|k_1|})$ and $|k_2| = \sum_{j=1}^{|k_1|} k_{2j}$. Inductively at time tr_i, there are n_{ij} descendants for the j-th particle, $1 \leq j \leq |k_{i-1}|$. For each such j, let k_{ij} out of n_{ij} have descendants that survive up till time tr_{i+1} (See Figure 3 for an illustration).

It will be useful to use the following notation: Let

$$n_{11}, k_{11} \in \mathbb{N}, k_{11} \leq n_{11}, |k_1| = k_{11}, n_1 = (n_{11}).$$

For $i = 2, \ldots i_p$ let $(n_i, k_i) \in \mathbb{N}_i$, where $N_i \equiv \mathbb{N}^{|k_{i-1}|} \times \mathbb{N}^{|k_{i-1}|}$

$$k_{ij} \leq n_{ij}, |\, k_i \,| \equiv \sum_{j=1}^{|k_{i-1}|} k_{ij}, \binom{n_i}{k_i} \equiv \prod_{j=1}^{|k_{i-1}|} \binom{n_{ij}}{k_{ij}}.$$

Let $f_s = P(N_s > 0)$. Now,

$$\begin{aligned}
&P(\frac{\tau_{i_1}}{t} < r_1 < \frac{\tau_{i_2}}{t} < r_2 < \ldots < \frac{\tau_{i_p}}{t} < r_p < \frac{\tau_{k-1}}{t} < r_{k-1} < t|A_t) = \\
&= \frac{f_{tr_1}}{f_t} \sum_{(n_i,k_i)\in\mathbb{N}_i} \left(\binom{n_{11}}{k_{11}} (f_{tr_1})^{k_{11}} (1-f_{tr_1})^{n_{11}-k_{11}} \right) \frac{P(N_{tr_1} = n_1)}{f_{tr_1}} \times \\
&\times \prod_{i=1}^{p+1} \prod_{j=1}^{|k_{i-1}|} \binom{n_{ij}}{k_{ij}} (f_{tu_i})^{k_{ij}} (1-f_{u_i})^{n_{ij}-k_{ij}} P(N^j_{tu_i} = n_{i,j} | N^j_{tu_i} > 0) \times \\
&\times g(\mathbf{k}) E \frac{\prod_{j=1}^k X^j}{S^k},
\end{aligned}$$

with $u_i = r_{i+1} - r_i, i = 1, 2, \ldots, p-1$, $u_p = 1 - r_p$, $N^j_{tu_i}$ is number of particles alive at time tu_i of the age-dependent branching process starting with one particle namely j, $g(\mathbf{k}) = g(k_1, \dots, k_p)$ is the proportion of configurations that have the desired number of ancestors corresponding to the given event, $X^j \stackrel{d}{=} N^j_{tu_p} | N^j_{tu_p} > 0$ and $S = \sum_{j=1}^{|k_{p+1}|} X^j$.

Let $q_i = \frac{u_i}{u_{i+1}}$ for $1 \le i \le p-1$. Then following [9] and using Proposition 3.1 (i), (ii) repeatedly we can show that $P(\frac{\tau_{i_1}}{t} < r_1 < \frac{\tau_{i_2}}{t} < r_2 < \ldots < \frac{\tau_{i_p}}{t} < r_p < \frac{\tau_{k-1}}{t} < r_{k-1} < t | A_t)$ converges to

$$
\begin{aligned}
&\frac{1}{q_1} \sum_{k_i \in \mathbb{N}^{|k_{i-1}|}} \int dx e^{-x} (q_1 x)^{k_{11}} \frac{1}{k_{11}!} e^{-x q_1} \times \\
&\quad \times \prod_{i=2}^{p+1} \prod_{j=1}^{|k_{i-1}|} \int dx e^{-x} \frac{(q_i x)^{k_i j}}{k_{ij}!} e^{-x q_i}) g(\mathbf{k}) \\
&\quad \times \int \prod_{i=1}^{k+1} dx_i \left(\frac{\prod_{i=1}^{k} x_i}{(\sum_{i=1}^{k+1} x_i)^k} \right) e^{-\sum_{i=1}^{k+1} x_i} \frac{(x_{k+1})^{|k_{p+1}|-k}}{(|k_{p+1}| - k)!}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \frac{1}{q_1} \sum_{k_i \in \mathbb{N}^{|k_{i-1}|}} \prod_{i=1}^{p+1} \frac{(q_i)^{|k_i|}}{(1+q_i)^{|k_i| - |k_{i-1}|}} g(\mathbf{k}) \times \\
&\qquad \times \int \prod_{i=1}^{k+1} dx_i \left(\frac{\prod_{i=1}^{k} x_i}{(\sum_{i=1}^{k+1} x_i)^k} \right) e^{-\sum_{i=1}^{k+1} x_i} \frac{(x_{k+1})^{|k_{p+1}|-k}}{(|k_{p+1}| - k)!}.
\end{aligned}
$$

Consequently, we have shown that the random vector $\tilde{\tau}_t$ converges in distribution to a random vector $\tilde{T}$. From the above limiting quantity, one can show that the $\tilde{T}$ has an absolutely continuous distribution on $[0,1]^{k-1}$. See [3] for a detailed proof.

□

4. Proof of Theorem 2.1

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 the motion processes of the ancestors of this individual and $\{\eta_{t(M_t+1)}(u) : 0 \le u \le t - \sum_{i=1}^{M_t} L_{ti}\}$ be the motion 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:

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

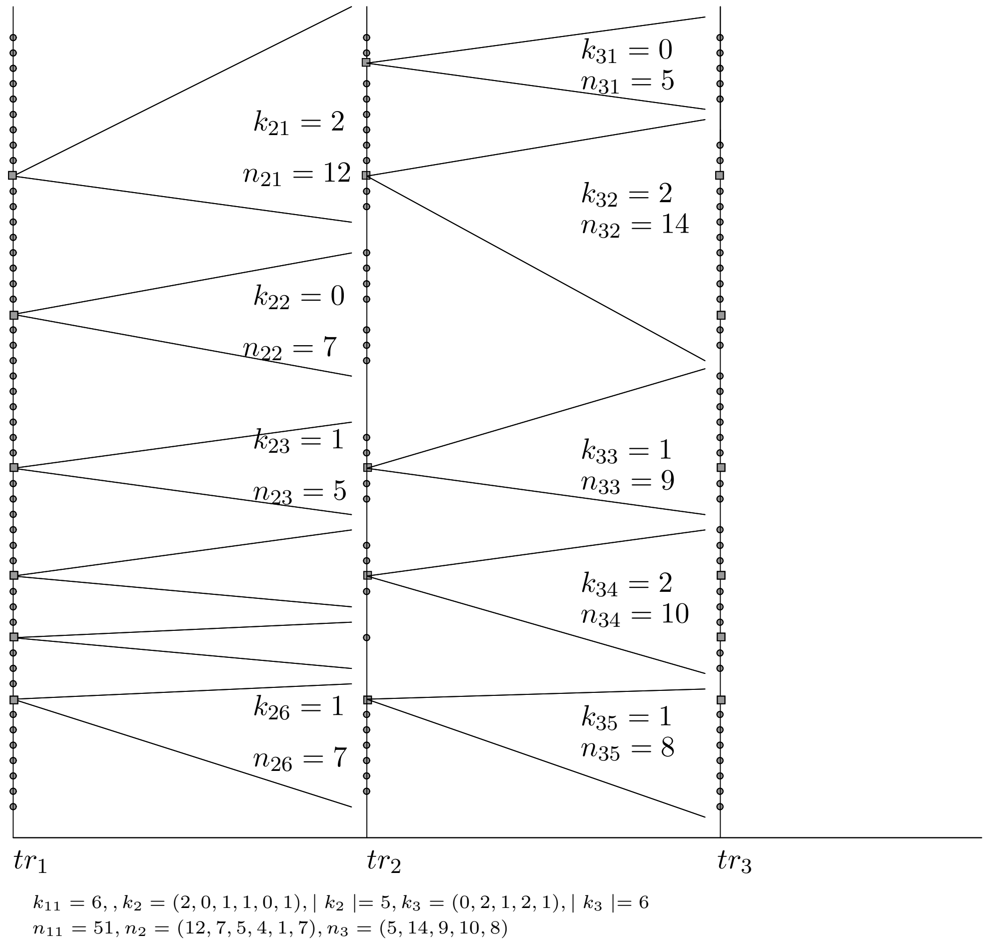

FIGURE 1. Tracking particles surviving at various times

whenever $M_t > 0$ and is equal to $a + t$ otherwise; and that

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

Rearranging the terms, we obtain

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

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 $\epsilon > 0$ be given.

$$\begin{aligned}
P(|Z_{t2}| > \epsilon | A_t) &\leq P(|Z_{t2}| > \epsilon, a_t \leq k | A_t) + P(|Z_{t2}| > \epsilon, a_t > k | A_t) \\
&\leq P(|Z_{t2}| > \epsilon, a_t \leq k | A_t) + P(a_t > k | A_t) \\
&\leq \frac{E(|Z_{t2}|^2 I_{a_t \leq k} | A_t)}{\epsilon^2} + P(a_t > k | A_t)
\end{aligned}$$

By Proposition 3.1 and the ensuing tightness, for any $\eta > 0$ there is a k_η

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

for all $k \geq k_\eta, t \geq 0$. Next,

$$\begin{aligned}
E(|Z_{t2}|^2 I_{a_t \leq k_\eta} | A_t) &= E(I_{a_t \leq k_\eta} E(|Z_{t2}|^2 | \mathcal{L}_t) | A_t) \\
&= E(I_{a_t \leq k_\eta} \frac{v(a_t)}{t} | A_t) \\
&\leq \frac{\sup_{u \leq k_\eta} v(u)}{t}.
\end{aligned}$$

Hence,

$$P(|Z_{t2}| > \epsilon | A_t) \leq \frac{\sup_{u \leq k_\eta} v(u)}{t\epsilon^2} + \frac{\eta}{2}$$

Since $\epsilon > 0$ and $\eta > 0$ are arbitrary this shows that as $t \to \infty$

$$Z_{t2} | A_t \xrightarrow{d} 0, \tag{4.1}$$

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

$$\begin{aligned}
E(e^{-\lambda a_t} e^{-i\frac{\theta}{\sqrt{\mu}} Z_{t1}} | A_t) &= E(e^{-\lambda a_t} (E(e^{-i\theta Z_{t1}} | \mathcal{L}_t) - e^{-\frac{\theta^2 \psi}{2\mu}}) | A_t) + \\
&\quad + {}^{-\frac{\theta^2 \psi}{2\mu}} E(e^{-\lambda a_t} | A_t)
\end{aligned}$$

Proposition 3.3 shows that the first term above converges to zero and using Proposition 3.1 we can conclude that as $t \to \infty$

$$(a_t, \frac{1}{\sqrt{\mu}} Z_{t1}) | A_t \xrightarrow{d} (U, V) \tag{4.2}$$

As X_0 is a constant, by Proposition 3.1 (c), (4.2), (4.1) and Slutsky's Theorem, the proof is complete. □

5. Proof of Theorem 2.2

Let $\phi \in C_b(\mathbb{R} \times \mathbb{R}_+)$. We shall show, for each $k \geq 1$, that the moment-functions of $E(\frac{<\tilde{Y}_t,\phi>^k}{N_t^k}|A_t)$ converges as $t \to \infty$. Then by Theorem 16.16 in [11] the result follows.

The case $k = 1$ follows from Theorem 2.1 and the bounded convergence theorem. We shall next consider the case $k = 2$. 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 Proposition 3.5 (i) conditioned on A_t, $\frac{\tau_t}{t}$ converges in distribution to an absolutely continuous random variable T (say) in $[0, 1]$. Also by Theorem 2.1 conditioned on $\mathcal{G}_t$ and A_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 2.1. Also $\frac{\tilde{X}_{\tau_t}}{\sqrt{\tau_t}}$ conditioned on A_{τ_t} converges in distribution to a random variable S distributed as V.

Combining these one can conclude that $\{(a_t^i, \frac{X_t^i}{\sqrt{t}}), i = 1, 2\}$ conditioned on A_t converges in distribution to $\{(U_i, \sqrt{T}S + \sqrt{(1-T)}V_i), i = 1, 2\}$ where U_1, U_2, S, V_1, V_2 are all independent. Thus for any $\phi \in C_b(\mathbb{R}_+ \times \mathbb{R})$ we have, by the bounded convergence theorem,

(5.1)
$$\lim_{t\to\infty} E(\prod_{i=1}^{2} \phi(a_t^i, \frac{X_t^i}{\sqrt{t}})|A_t) = E\prod_{i=1}^{2} \phi(U_i, \sqrt{T}S + \sqrt{(1-T)}V_i) \equiv m_2(\phi) \text{ (say)}$$

Now,

$$\begin{aligned} E(\left(\frac{\tilde{Y}_t(\phi)}{N_t}\right)^2|A_t) &= E(\frac{(\phi(a_t, \frac{X_t}{\sqrt{t}}))^2}{N_t}|A_t) \\ &\quad + E(\prod_{i=1}^{2} \phi(a_t^i, \frac{X_t^i}{\sqrt{t}})\frac{N_t(N_t-1)}{N_t^2}|A_t) \end{aligned}$$

Using Proposition 3.1 (b) and the fact that ϕ is bounded we have $\lim_{t\to\infty} E((\frac{\tilde{Y}_t(\phi)}{N_t})^2|A_t)$ exists in $(0, \infty)$ and equals $m_2(\phi)$. The case $k > 2$

can be proved in a similar manner but we use Theorem 3.5 (ii) as outlined below. First we observe that as ϕ is bounded,

$$E\left(\frac{<\tilde{Y}_t,\phi>^k}{N_t^k}|A_t\right)+=\sum_{\mathbf{i}} Eh(N_t,k)\left(\prod_{j=1}^k \phi(a_t^{i_j},\frac{X_t^{i_j}}{\sqrt{t}})|A_t\right)+g(\phi,\mathcal{C}_t,N_t),$$

where $h(N_t,k)\to 1$ and $g(\phi,\mathcal{C}_t,N_t)\to 0$ as $t\to\infty$; and $\mathbf{i}=\{i_1,i_2,\ldots,i_k\}$ is the index of $k$particles sampled without replacement from $\mathcal{C}_t$ (see (2.1)). Consider one such sample, and re-trace the genealogical tree $\mathcal{T}_{\mathbf{i}}\in\mathcal{T}(k)$,($\mathcal{T}(k)$ is the collection of all possible trees with k leaves given by $\mathbf{i}$), until their most common ancestor. For any leaf i_j in $\mathcal{T}_{\mathbf{i}}$, let $1=n(i_j,1)<n(i_j,2)<\cdots<n(i_j,N_{i_j})$ be the labels of the internal nodes on the path from leaf i_j to the root. We list the ancestoral times on this by $\{\tau_1,\tau_{n(i_j,1)},\ldots,\tau_{n(i_j,N_{i_j})}$. Finally we denote the net displacement of the ancestors in the time intervals

$$[0,\tau_1],[\tau_1,\tau_{n(i_j,2)}],\ldots,[\tau_{n(i_j,N_{i_j}-1)},\tau_{n(i_j,N_{i_j})}],[\tau_{n(i_j,N_{i_j})},t]$$

by

$$\tilde{\eta}^1_{i_j}(\tau_1),\tilde{\eta}^2_{i_j}(\tau_{n(i_j,2)},\tau_1),\ldots,\tilde{\eta}^{N_{i_j}}_{i_j}(\tau_{n(i_j,N_{i_j})},\tau_{n(i_j,N_{i_j}-1)}),\tilde{\eta}'_{i_j}(t,\tau_{n(i_j,N_{i_j})}).$$

Given the above notation we have:

$$E\left(\prod_{j=1}^k \phi(a_t^{i_j},\frac{X_t^{i_j}}{\sqrt{t}})|A_t\right)=E\left(\sum_{T\in\mathcal{T}_{\mathbf{i}}}\prod_{j=1}^k f(\phi,j,t)|A_t\right),$$

where

$$f(\phi,j,t)=\phi(a_t^{i_j},\frac{1}{\sqrt{t}}(\tilde{\eta}^1_{i_j}(\tau_1)+\sum_{m=2}^{N_{i_j}}\tilde{\eta}^m_{i_j}(\tau_{n(i_j,m)},\tau_{n(i_j,m-1)})+\tilde{\eta}'_{i_j}(t,\tau_{n(i_j,N_{i_j})}).$$

Now by Theorem 3.5,

$$\frac{(\tau_1,\tau_{n(i_j,2)},\ldots,\tau_{n(i_j,N_{i_j})})}{\sqrt{t}}|A_t\xrightarrow{d}(T_1,T_{n(i_j,2)},\ldots,T_{n(i_j,N_{i_j})}).$$

So by Theorem 2.1

$$\lim_{t\to\infty}E(\left(\frac{\tilde{Y}_t(\phi)}{N_t}\right)^2|A_t)=E\left(\sum_{\mathbf{i}}\sum_{T\in\mathcal{T}_{\mathbf{i}}}\prod_{j=1}^k g(\phi,j,t)|A_t\right)\equiv m_k(\phi)$$

(5.2)

where

$$
\begin{aligned}
g(\phi, j, t) = & \\
= \quad & \phi\left(U, S\sqrt{T_1} + \sum_{m=2}^{N_{i_j}} Z_{i_j}^m \sqrt{T_{n(i_j,m)} - T_{n(i_j,m-1)}} + Z'_{i_j}\sqrt{1 - T_{n(i_j,N_{i_j})}}\right)
\end{aligned}
$$

with $S, Z'_{i_j}, Z_{i_j}^m$, $m = 2, \ldots, N_{i_j}$, are i.i.d.V, U is an independent random variable given in Theorem 2.1 and T_i's are as in Theorem 3.5 (ii). Since ϕ is bounded, the sequence $\{m_k(\phi) \equiv \lim_{t\to\infty} E(\frac{<\tilde{Y}_t,\phi>^k}{N_t^k})\}$ is necessarily a moment sequence of a probability distribution on $\mathbb{R}$. This being true for each ϕ, by Theorem 16.16 in [11] we are done. □

6. Proof of Theorem 2.4

Let Z be the Branching Markov process Y described earlier, with lifetime G exponential with mean λ, $p_1 = 1$ and $\eta \overset{d}{=} \eta_1$ (see (2.7)). Then it is easy to see that for any bounded continuous function, $S_t\phi(a,x) = E_{(a,x)} < Z_t, \phi >= E_{(a,x)}\phi(a_t, X_t)$ satisfies the following equation:

$$
S_t\phi(a,x) = e^{-\lambda t}W_t\phi(a,x) + \int_0^t ds\lambda e^{-\lambda s}W_s(S_{t-s}(\phi)(0,\cdot))(a,x), \tag{6.1}
$$

where W_t is the semi-group associated to η_1. Let $\mathcal{L}$ be the generator of η_1. Making a change of variable $s \to t-s$ in the second term of the above and then differentiating it with respect to t, we have

$$
\begin{aligned}
\frac{d}{dt}S_t(\phi)(a,x) = & -\lambda e^{-\lambda t}W_t\phi(a,x) + e^{-\lambda t}\mathcal{L}W_t\phi(a,x) + \lambda S_t(\phi)(0,x) \\
& + \int_0^t ds\lambda(-\lambda e^{-\lambda(t-s)})W_{t-s}(S_s(\phi)(0,\cdot))(a,x) \\
& + \int_0^t ds\lambda e^{-\lambda(t-s)}\mathcal{L}W_{t-s}(S_s(\phi)(0,\cdot))(a,x) \\
= \quad & \lambda S_t(\phi)(0,x) \\
& +(\mathcal{L}-\lambda)\left[e^{-\lambda t}W_t\phi(a,x) + \int_0^t ds\lambda e^{-\lambda(t-s)}W_{t-s}(S_s(\phi)(0,\cdot))(a,x)\right] \\
= \quad & \lambda S_t(\phi)(0,x) + (\mathcal{L}-\lambda)S_t(\phi)(a,x) \\
= \quad & \frac{\partial S_t\phi}{\partial a}(a,x) + \frac{\sigma^2(a)}{2}\Delta S_t\phi(a,x) + \lambda(S_t(\phi)(0,x) - S_t(\phi)(a,x)),
\end{aligned}
$$

For each $n \geq 1$ define (another semigroup) $R_t^n \phi(a,x) = E_{a,0}(\phi(a_t, x + \frac{X_t}{\sqrt{n}})$. Now note that,

$$\begin{aligned} R_t^n \phi(a,x) &= E_{a,0}(\phi(a_t, x + \frac{X_t}{\sqrt{n}}) \\ &= E_{a,\sqrt{n}x}(\phi(a_t, \frac{X_t}{\sqrt{n}}) \\ &= S_t \phi_n(a, \sqrt{n}x), \end{aligned}$$

where $\phi_n(a,x) = \phi(a, \frac{x}{\sqrt{n}})$. Differentiating w.r.t. t, we have that the generator of R_t^n is

$$\mathcal{R}^n \phi(a,x) = \frac{\partial \phi}{\partial a}(a,x) + \frac{\sigma^2(a)}{2n} \Delta \phi(a,x) + \lambda(\phi(0,x) - \phi(a,x)). \tag{6.2}$$

Proposition 6.1. *Let $\epsilon > 0$ and $t \geq \epsilon$. Let $\phi \in C_l^+(\mathbb{R}_+ \times \mathbb{R}^d)$. Then,*

$$\sup_{(a,x) \in \mathbb{R}_+ \times \mathbb{R}} \mid R_{nt}^n(\phi)(a,x) - U_t(\phi)(x) \mid \to 0. \tag{6.3}$$

Proof: Let $t \geq \epsilon$. Applying Theorem 2.1 to the process Z, we have $(a_{nt}, \frac{X_{nt}}{\sqrt{n}}) \xrightarrow{d} (U, V)$. The proposition is then immediate from the bounded convergence theorem and the fact that $\phi \in C_l^+(\mathbb{R}_+ \times \mathbb{R})$ □

Proposition 6.2. *Let $\pi_{n\nu}$ be a Poisson random measure with intensity $n\nu$ and $t \geq 0$. The log-Laplace functional of $\mathcal{Y}_t^n$,*

$$E_{\pi_{n\nu}}[e^{-\langle \phi, \mathcal{Y}_t^n \rangle}] = e^{-\langle u_t^n \phi, \nu \rangle}, \tag{6.4}$$

where

$$u_t^n \phi(a,x) = R_{nt}^n n(1 - e^{-\frac{\phi}{n}})(a,x) - \lambda \int_0^t ds R_{n(t-s)}^n (n^2 \Psi_n(\frac{u_s^n \phi}{n}))(a,x), \tag{6.5}$$

where

$$\Psi_n(\phi)(a,x) := \left[F_n(1 - \phi(0,x)) - (1 - \phi(0,x))\right].$$

Proof: For any $n \in \mathbb{N}$, let Y_t^n be the sequence of branching Markov processes defined in Section 2.2. It can be shown that its log-Laplace functional L_t^n satisfies,

$$L_{nt}^n \phi(a,x) = e^{-\lambda nt} W_{nt}^n[e^{-\phi}](a,x) + \int_0^{nt} ds \lambda e^{-\lambda s} W_s^n \left[F_n(L_{nt-s}^n \phi(0,\cdot))\right](a,x) ds, \tag{6.6}$$

where $t \geq 0$ and W_t^n is the semigroup associated with η_n. Using the fact that $e^{-\lambda u} = 1 - \int_0^u ds \lambda e^{-\lambda s}$ for all $u \geq 0$ and a routine simplification, as done in [10], will imply that

(6.7)
$$L^n_{nt}\phi(a,x) = W^n_{nt}[e^{-\phi}](a,x) + \lambda \int_0^{nt} W^n_{nt-s}(F_n(L^n_s\phi(0,\cdot)) - L^n_s\phi)(a,x)ds$$

Therefore $v^n_{nt}(\phi)(a,x) = 1 - L^n_t\phi(a,x)$, satisfies,

(6.8)
$$v^n_{nt}\phi(a,x) = W^n_{nt}(1-e^{-\phi})(a,x)+\int_0^{nt} dsW^n_{nt-s}((1-v^n_s\phi)-F_n(1-v^n_s)\phi(0,\cdot)))(a,x)\lambda ds.$$

Let $\mathcal{L}^n$ be the generator of η_n. Then for $0 \leq s < t$

$$\begin{aligned}
&\frac{d}{ds}R^n_{n(t-s)}(v^n_{ns}(\phi))(a,x) = \\
&= -(n\mathcal{R}^n)R^n_{n(t-s)}\left(v^n_{ns}(\phi)\right)(a,x) + R^n_{n(t-s)}\left(\frac{\partial}{\partial s}v^n_{ns}(\phi)\right)(a,x) \\
&= -(n\mathcal{R}^n)R^n_{n(t-s)}\left(v^n_{ns}(\phi)\right)(a,x) \\
&\quad +R^n_{n(t-s)}\left(n\mathcal{L}^nW^n_{ns}(1-e^{-\phi}) + n\lambda((1-v^n_{ns}\phi) - F_n(1-v^n_{ns})\phi(0,\cdot)))(a,x)\right) \\
&\quad +R^n_{n(t-s)}\left(\int_0^{ns} drn\mathcal{L}^n(W^n_{ns-r}((1-v^n_r(\phi)) - F_n(1-v^n_r)\phi(0,\cdot))\right)(a,x) \\
&= R^n_{n(t-s)}n\left(-\lambda(v^n_{ns}(\phi)(0,\cdot) - v^n_{ns}(\phi)) + \lambda((1-v^n_{ns}\phi) - F_n(1-v^n_{ns})\phi(0,\cdot)))(a,x)\right) \\
&= -R^n_{n(t-s)}(n\Psi_n(v^n_{ns}\phi))(a,x),
\end{aligned}$$

Integrating both sides with respect to s from 0 to t, we obtain that

$$v^n_{nt}(\phi)(a,x) = R^n_{nt}(1-e^{-\phi})(a,x) - \int_0^t dsR^n_{n(t-s)}(n\Psi_n(v^n_{ns}\phi))(a,x). \tag{6.9}$$

If $\pi_{n\nu}$ is a Poisson random measure with intensity $n\nu$, then

$$E_{\pi_{n\nu}}[e^{-\langle\phi,\mathcal{Y}^n_t\rangle}] = E_{\pi_{n\nu}}[e^{-\langle\frac{\phi}{n},Y^n_{nt}\rangle}] = e^{\langle L^n_t(\frac{\phi}{n})-1,n\nu\rangle} = e^{-\langle nv^n_t(\frac{\phi}{n}),\nu\rangle}.$$

Therefore if we set $u^n_t(\phi) \equiv nv^n_{nt}(\frac{\phi}{n})$. From (6.9), it is easy to see that $u^n_t(\phi)$ satisfies (6.4). □

For any $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, , we let $\| f \|_\infty = \sup_{(a,x)\in\mathbb{R}_+\times\mathbb{R}} | f(a,x) |$. With a little abuse of notation we shall let $\| f \|_\infty = \sup_{x\in\mathbb{R}} | f(x) |$ when $f : \mathbb{R} \to \mathbb{R}$ as well.

Proposition 6.3. *Let $\epsilon > 0$. $\phi \in C^+_l(\mathbb{R}_+ \times \mathbb{R}^d)$ and $u^n_t(\phi)$ be as in Proposition 6.2 and $u_t(\phi)$ be as in Theorem 2.4. Then for $t \geq \epsilon$,*

$$\sup_{(a,x)\in\mathbb{R}_+\times\mathbb{R}} | u^n_t(\phi)(a,x) - u_t(\phi)(x) | \to 0 \tag{6.10}$$

Proof: For any real $u \in \mathbb{R}$, define, $\varepsilon_n(u) = \lambda n^2(F_n(1-\frac{u}{n}) - (1-\frac{u}{n})) - u^2$. So,

$$\begin{aligned} u_t^n(\phi)(a,x) &= R_{nt}^n n(1-e^{-\frac{\phi}{n}})(a,x) - \lambda \int_0^t ds R_{n(t-s)}^n (n^2\Psi_n(\frac{u_s^n\phi}{n}))(a,x) \\ &= R_{nt}^n n(1-e^{-\frac{\phi}{n}})(a,x) - \int_0^t ds R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0\cdot))))(a,x) \\ &\quad -\lambda \int_0^t ds R_{n(t-s)}^n (u_s^n\phi(0,\cdot)^2)(a,x) \end{aligned}$$

Now

$$\begin{aligned} & u_t^n(\phi)(a,x) - u_t(\phi)(x) = \\ &= R_{nt}^n(n(1-e^{-\frac{\phi}{n}}))(a,x) - U_t(\phi)(x) \\ &- \int_0^t ds R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0\cdot))))(a,x) \\ &+ \lambda \int_0^t ds \left(U_{(t-s)}((u_s\phi)^2)(a,x) - R_{n(t-s)}^n(u_s^n\phi(0,\cdot)^2)(a,x)\right) \end{aligned}$$

$$\begin{aligned} = \quad & R_{nt}^n(n(1-e^{-\frac{\phi}{n}}))(a,x) - U_t(\phi)(x) - \int_0^t ds R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0,\cdot))))(a,x) \\ & +\lambda \int_0^t ds R_{n(t-s)}^n((u_s\phi)^2 - u_s^n\phi(0,\cdot)^2)(a,x) \\ & \qquad +\lambda \int_0^t ds \left(U_{t-s}(u_s\phi)^2)(x) - R_{n(t-s)}^n(u_s\phi)^2)(a,x)\right) \end{aligned}$$

Observe that, $R_\cdot^n$ is a contraction, $\| u_\cdot^n(\phi) \|_\infty \leq \| \phi \|_\infty$ and $\| u_\cdot(\phi) \|_\infty \leq \| \phi \|_\infty$ for $\phi \in C_l(\mathbb{R}_+ \times \mathbb{R})$. Therefore, we have

$$\begin{aligned} \| u_t^n(\phi) - u_t(\phi) \|_\infty \quad \leq \quad & \| R_{nt}^n(n(1-e^{-\frac{\phi}{n}})) - U_t(\phi) \|_\infty + t \| \epsilon_n(u_s^n(\phi(0,\cdot)) \|_\infty \\ & +2\lambda \| \phi \|_\infty \int_0^t ds \| u_s^n(\phi) - u_s(\phi) \|_\infty \\ & +\lambda \int_0^t ds \| (U_{t-s} - R_{n(t-s)}^n)(u_s\phi)^2 \|_\infty . \end{aligned}$$

For $\phi \in C_l(\mathbb{R}_+ \times \mathbb{R}^d)$, note that, U_t, is a strongly continuous semi-group implies that $u_s(\phi)$ is a uniformly continuous function. So using Proposition 6.3 the first term and the last term go to zero. By our assumption on F, $\| \epsilon_n(u_s^n(\phi(0,\cdot)) \|_\infty$ will go to zero as $n \to \infty$. Now using the standard Gronwall argument we have the result. $\square$

Proposition 6.4. *Let $\epsilon > 0$. The processes $\mathcal{Y}^n_\cdot$ are tight in $D([\epsilon,\infty), M(\mathbb{R}_+\times\mathbb{R}))$.*

Proof By Theorem 3.7.1 and Theorem 3.6.5 (Aldous Criterion) in [7] , it is enough to show

$$\langle \mathcal{Y}^n_{\tau_n+\delta_n}, \phi\rangle - \langle \mathcal{Y}^n_{\tau_n}, \phi\rangle \xrightarrow{d} 0, \tag{6.11}$$

where $\phi \in C_l^+(\mathbb{R}_+\times\mathbb{R})$, δ_n is a sequence of positive numbers that converge to 0 and τ_n is any stop time of the process $\mathcal{Y}^n$ with respect to the canonical filtration, satisfying $0 < \epsilon \leq \tau_n \leq T$ for some $T < \infty$.

First we note that, as $\langle \mathcal{Y}^n_t, 1\rangle$ is a martingale, for $\gamma > 0$ by Chebyschev's inequality and Doob's maximal inequality we have

$$P(\langle \mathcal{Y}^n_{\tau_n}, \phi\rangle > \gamma) \leq \frac{1}{\gamma}c_1 \parallel \phi \parallel_\infty E(\sup_{\epsilon\leq t\leq T}\langle \mathcal{Y}^n_t, 1\rangle) \leq \frac{1}{\gamma}c_2 \parallel \phi \parallel_\infty \tag{6.12}$$

By the strong Markov Property applied to the process $\mathcal{Y}^n$ we obtain that for $\alpha,\beta \geq 0$, we have

$$\begin{aligned} L_n(\delta_n;\alpha,\beta) &= E(\exp(-\alpha\langle \mathcal{Y}^n_{\tau_n+\delta_n}, \phi\rangle - \beta\langle \mathcal{Y}^n_{\tau_n}, \phi\rangle)) \\ &= E(\exp(-\langle \mathcal{Y}^n_{\tau_n}, u^n_{\delta_n}(\alpha\phi) + \beta\phi\rangle)) \\ &= E(\exp(-\langle \mathcal{Y}^n_{\tau_n-\epsilon}, u^n_\epsilon(u^n_{\delta_n}(\alpha\phi) + \beta\phi)\rangle)) \end{aligned}$$

Therefore

$$\begin{aligned} &\mid L_n(0;\alpha,\beta) - L_n(\delta_n;\alpha,\beta) \mid \leq \\ &\quad\leq \parallel u^n_\epsilon(u^n_{\delta_n}(\alpha\phi)+\beta\phi) - u^n_\epsilon((\alpha+\beta)\phi) \parallel_\infty E(\sup_{t\leq T}\langle \mathcal{Y}^n_t, 1\rangle) \\ &\quad\leq c_1 \parallel u^n_\epsilon(u^n_{\delta_n}(\alpha\phi)+\beta\phi) - u^n_\epsilon((\alpha+\beta)\phi) \parallel_\infty \end{aligned}$$

where is the last inequality is by Doob's maximal inequality. Now,

$$\begin{aligned} &\parallel u^n_\epsilon(u^n_{\delta_n}(\alpha\phi)+\beta\phi) - u^n_\epsilon((\alpha+\beta)\phi) \parallel_\infty \leq \parallel R^n_{n\epsilon}(u^n_{\delta_n}(\alpha\phi) - \alpha\phi) \parallel_\infty + \\ &\quad + c_2 \parallel \phi \parallel_\infty \int_0^\epsilon da \parallel u^n_a(u^n_{\delta_n}(\alpha\phi)+\beta\phi) - u^n_a((\alpha+\beta)\phi) \parallel_\infty + d_n(\phi), \end{aligned}$$

where $d_n(\phi) = \lambda \int_0^\epsilon da \parallel \epsilon_n(u_a^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi) + \epsilon_n(u_a^n((\alpha+\beta)\phi)) \parallel_\infty$. Observe that

$$
\begin{aligned}
\parallel R^n_{n\epsilon}(u^n_{\delta_n}(\alpha\phi) - \alpha\phi) \parallel_\infty &\leq \parallel R^n_{n\epsilon}(u^n_{\delta_n}(\alpha\phi) - R^n_{n\delta_n}(\alpha\phi)) \parallel_\infty \\
&\quad + \parallel R^n_{n\epsilon}(R^n_{n\delta_n}(\alpha\phi) - \alpha\phi) \parallel_\infty \\
\leq \quad & \parallel u^n_{\delta_n}(\alpha\phi) - R^n_{n\delta_n}(\theta_2\phi) \parallel_\infty + \parallel R^n_{n(\epsilon+\delta_n)}(\alpha\phi) - R^n_{n\epsilon}(\alpha\phi) \parallel_\infty \\
\leq \quad & \parallel R^n_{n\delta_n}(n(1-e^{\frac{s\phi}{n}}) - \alpha\phi) \parallel_\infty + \int_0^{\delta_n} da \parallel R^n_{n(\delta_n - a)}(n^2\Psi(\frac{u^n_a\phi}{n})) \parallel_\infty \\
& \quad + \parallel R^n_{n(\epsilon+\delta_n)}(\alpha\phi) - R^n_{n\epsilon}(\alpha\phi) \parallel_\infty \\
\leq \quad & \parallel n(1-e^{\frac{s\phi}{n}}) - \alpha\phi \parallel_\infty + \delta_n c_2(\parallel \phi \parallel_\infty^2 + 1) + \parallel R^n_{n(\epsilon+\delta_n)}(\alpha\phi) - R^n_{n\epsilon}(\alpha\phi) \parallel_\infty, \\
& \equiv e_n(\phi)
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&\parallel u^n_\epsilon(u^n_{\delta_n}(\alpha\phi) + \beta\phi) - u^n_\epsilon((r+s)\phi) \parallel_\infty \leq e_n(\phi) + d_n(\phi) \\
&\qquad + c_2 \parallel \phi \parallel_\infty \int_0^\epsilon da \parallel u^n_a(u^n_{\delta_n}(\alpha\phi) + \beta\phi) - u^n_a((r+s)\phi) \parallel_\infty .
\end{aligned}
$$

By Proposition 6.1, $e_n(\phi) \to 0$ and $d_n(\phi) \to 0$ by our assumption F_n. Hence by a standard Gronwall argument we have that,

$$
| L_n(0; s, r) - L_n(\delta_n; s, r) | \to 0 \tag{6.13}
$$

By (6.12), $\{\langle \mathcal{Y}^n_{\tau_n}, \phi\rangle; n = 1, 2, \ldots\}$ is tight in $\mathbb{R}_+$. Take an arbitrary subsequence. Then there is a further subsequence of it indexed by $\{n_k; k = 1, 2, \ldots\}$ such that $\langle \mathcal{Y}^{n_k}_{\tau_{n_k}}, \phi\rangle$ converges in distribution to some random limit b. Thus we get

$$
(\mathcal{Y}^{n_k}_{\tau_{n_k}}(\phi), \mathcal{Y}^{n_k}_{\tau_{n_k}}(\phi)) \xrightarrow{d} (b, b) \text{ as } k \to \infty.
$$

But (6.13) implies that

$$
(\mathcal{Y}^{n_k}_{\tau_{n_k}}(\phi), \mathcal{Y}^{n_k}_{\tau_{n_k}+\delta_{n_k}}(\phi)) \xrightarrow{d} (b, b) \text{ as } k \to \infty.
$$

This implies that $\langle \mathcal{Y}^{n_k}_{\tau_{n_k}+\delta_{n_k}}, \phi\rangle - \langle \mathcal{Y}^{n_k}_{\tau_{n_k}}, \phi\rangle \xrightarrow{d} 0$ as $k \to \infty$. So (6.11) holds and the proof is complete. □

Proof of Theorem 2.4 Proposition 6.3 shows that the log-Laplace functionals of the process $\mathcal{Y}^n_t$ converge to $\mathcal{Y}_t$ for every $t \geq \epsilon$. Proposition 6.4 implies tightness for the processes. As the solution to (2.8) is unique, we are done. □

References

[1] D. Aldous, Stopping times and tightness, Ann. Probab. 6 (1978) 335-340.

[2] K.B. Athreya, On the Supercritical Age-dependent Branching Process *Ann. Math. Statist.* **40** 743-763.

[3] K.B. Athreya, Common ancestor problem in branching processes, *in preparation*, 2007

[4] K.B. Athreya and S. Lahiri, Probability Theory, **41**, TRIM Series, *Hindustan Book Agency* 2006.

[5] K.B. Athreya and P. Ney, Branching Processes, *Dover, New York* 2000.

[6] A. Bose and I. Kaj, Age structured super-processes, *Technical Report*, 161, 1991.

[7] D.A. Dawson, Measure-valued Markov processes. École d´Éte de Probabilités de Saint-Flour XXI—1991, 1–260, Lecture Notes in Math., 1541, —em Springer, Berlin, 1993.

[8] D.A. Dawson, L.G.Gorostiza, and Z. Li, Nonlocal branching super-processes and some related models. *Acta Appl. Math.* 74 (2002), no. 1, 93–112.

[9] R. Durrett, The Genealogy of Critical Branching Processes *Stoc. Proc. App* 8 (1978) 101-116.

[10] Dynkin, Eugene B. An introduction to branching measure-valued processes. CRM Monograph Series, 6. American Mathematical Society, Providence, RI, 1994. x+134 pp.

[11] O. Kallenberg, Foundation of Modern Probability Theory, Springer, 2002.

Krishna B. Athreya, Department of Mathematics and Statistics, Iowa State University, Ames, Iowa, 50011, U.S.A.

E-mail address: `kba@iastate.edu`

Siva R. Athreya, 8th Mile Mysore Road, Indian Statistical Institute, Bangalore 560059, India.

E-mail address: `athreya@isibang.ac.in`

Srikanth K. Iyer, Department of Mathematics, Indian Institute of Science, Bangalore 560012, India.

E-mail address: `skiyer@math.iisc.ernet.in`