Critical age-dependent branching Markov processes and their scaling limits

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

[1]Department of Mathematics and Statistics, Iowa State University, Ames, Iowa 50011, USA
[2]Indian Statistical Institute, 8th Mile, Mysore Road, Bangalore 560 059, India
[3]Department of Mathematics, Indian Institute of Science, Bangalore 560 012, India
E-mail: kba@iastate.edu; athreya@isibang.ac.in; skiyer@math.iisc.ernet.in

Abstract. This paper studies: (i) the long-time behaviour of the empirical distribution of age and normalized 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.

Keywords.   Age-dependent; branching; ancestoral times; measure-valued; empirical distribution.

1. Introduction

An age-dependent branching Markov process consists 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 the offspring migrates according to a prescribed Markov process starting from the position where its parent died. The motion process, offspring distribution and lifetime distribution are all independent of each other.

In our model the motion process is allowed to depend on the age of the particle. The lifetimes are not necessarily exponential. The state space $\mathbb{R}$ is chosen mainly for technical convenience and plays no role in the proofs of our results. We assume that the system is critical, i.e. the mean of the offspring distribution is one. We shall describe the model more precisely in the next section.

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 a randomly chosen individual decouples (see Theorem 1.1). Second, it is shown that conditioned on non-extinction at time t, the empirical joint distribution of the age and the normalized position of the population converges in law as $t \to \infty$ to a random measure characterized by its moments (see Theorem 1.2). Finally, we establish a super-process type limit in a special case. We assume that the motion is governed by an age dependent diffusion and the lifetimes are exponential (see Theorem 1.3).

Biological applications have motivated the study of branching Markov processes. Branching Brownian motion where the branching part evolves as a standard Markov branching process, lifetimes are exponential and the movement of each individual is Brownian motion has been studied in [13]. More recently, Athreya and Kang [2] studied the case of discrete time Galton–Watson branching process where the movement was modelled as a positive recurrent Markov chain. The above papers focussed on the super-critical case, where the population size diverges to infinity over time. There are many situations, such as in population genetics (see [12, 15]) where the population evolution is not supercritical but critical, wherein the population dies out in finite time with probability one. However, conditioned on non-extinction, the appropriately normalized process has a limiting distribution, called the Yaglom-limit (see [5] or [17]). The study of the size, age and location spread of such population is of interest.

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. These are addressed in this paper.

1.1 *The model*

We begin with the description of the particle system. Suppose we are given the following:

(i) *Lifetime distribution $G(\cdot)$.* Let $G(\cdot)$ be a cumulative distribution function on $[0, \infty)$, with $G(0) = 0$. Let $\mu = \int_0^\infty s dG(s)$, and assume $\int_0^\infty s^2 dG(s) < \infty$.

(ii) *Offspring distribution* $\mathbf{p}$. Let $\mathbf{p} \equiv \{p_k\}_{k \geq 0}$ be a probability distribution such that $p_1 < 1$, $m = \sum_{k=0}^\infty k p_k = 1$ 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 time-inhomogeneous Markov process starting at 0. We assume that for all $0 \leq t < \infty$,

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

$$\text{and } \psi = \int_0^\infty v(s) G(\mathrm{d}s) < \infty. \tag{1.1}$$

Branching Markov process $(G, \mathbf{p}, \eta)$

Suppose we start with an initial configuration $\mathcal{C}_0 = \{(a_0^i, X_0^i) : i = 1, 2, \ldots, N_0\}$, $N_0 < \infty$. a_0^i, X_0^i respectively denote the age and position of the i-th particle at time 0. The i-th particle present in the system at time 0 lives for a random length of time L_i 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 moves 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 $\{x + \eta(t - \tau) : \tau \leq t < \tau + L\}$ and 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 the family of triplets (L, ξ, η) over all particles in the system are i.i.d.

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

$$\mathcal{C}_t = \{(a_t^i, X_t^i) : i = 1, 2, \ldots, N_t\} \tag{1.2}$$

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. $P(N_t < \infty) = 1$). Also $P(\eta(L) \in \mathbb{R}) = 1$ for each particle. Thus $\mathcal{C}_t$ is well-defined for each $0 \leq t < \infty$.

Notation 1.1. For a particle chosen 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 the number of ancestors, the lifetimes, and the motion processes of its ancestors respectively and $\{\eta_{t(M_t+1)}(u): 0 \leq u \leq t - \sum_{i=1}^{M_t} L_{ti}\}$ be the motion of the individual concerned. If $M_t = 0$, then

$$a_t = a + t \quad \text{and} \quad X_t = \eta_{t1}(a + t) - \eta_{t1}(a) + X_0, \tag{1.3}$$

where a and X_0 are the age and the position of the particle at time $t = 0$. If $M_t > 0$, then the age and position (a_t, X_t), of the particle are given by

$$a_t = t - \sum_{i=1}^{M_t} L_{ti} + a, \tag{1.4}$$

and

$$X_t = X_0 - \eta_{t1}(a) + \sum_{i=1}^{M_t} \eta_{ti}(L_{ti}) + \eta_{t(M_t+1)}(a_t). \tag{1.5}$$

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

The above model of the particle movement may be contrasted with the models considered in the superprocess literature (see [7]). In those models, given the full branching tree generated by the offspring distribution $\mathbf{p}$ and the lifetime distribution G, the motion process for each line of descent is a Markov process $\tilde{\eta}$ with a given transition function. Thus, the position of a sampled individual at time t has the same distribution as $\tilde{\eta}(t)$ and this is different from (1.5).

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 ages are in A and positions 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

$$\langle \phi, Y_t \rangle \equiv \int \phi \, dY_t = \sum_{i=1}^{N_t} \phi(a_t^i, X_t^i). \tag{1.6}$$

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, is given by

$$L_t \phi(a, x) := E_{a,x}[e^{-\langle \phi, Y_t \rangle}]$$
$$\equiv E[e^{-\langle \phi, Y_t \rangle} | Y_0 = \delta_{a,x}], \quad t \geq 0, \quad \text{and} \quad \phi \in C_b^+(\mathbb{R}_+ \times \mathbb{R}). \tag{1.7}$$

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

$$E_{\nu_1+\nu_2}[\mathrm{e}^{-\langle\phi,Y_t\rangle}] = (E_{\nu_1}[\mathrm{e}^{-\langle\phi,Y_t\rangle}])(E_{\nu_2}[\mathrm{e}^{-\langle\phi,Y_t\rangle}]), \qquad (1.8)$$

where $E_\nu[\mathrm{e}^{-\langle\phi,Y_t\rangle}] := E[\mathrm{e}^{-\langle\phi,Y_t\rangle}|Y_0 = \nu]$ for $\nu \in M_a(\mathbb{R}_+ \times \mathbb{R})$. This is usually referred to as the branching property of Y. In particular, if $\nu = \sum_{i=1}^{n} \delta_{a_i,x_i}$, $n < \infty$, then (see [10])

$$E_\nu[\mathrm{e}^{-\langle\phi,Y_t\rangle}] = \prod_{i=1}^{n} L_t\phi(a_i, x_i) = \mathrm{e}^{\langle\log L_t\phi,\nu\rangle}.$$

1.2 *Main 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 $P(N_0 = 0) \neq 1$. Let ψ be as defined in (1.1).

Theorem 1.1 (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 particle chosen uniformly from those alive at time t. Assume* (1.1). *Then, conditioned on A_t, $\left(a_t, \frac{X_t}{\sqrt{t}}\right)$ 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 given by $\frac{1}{\mu}(1-G(\cdot))$ 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 $\tilde{Y}_t(A \times B) \equiv Y_t(A \times \sqrt{t}B)$, $A \in \mathcal{B}(\mathbb{R}_+)$, $B \in \mathcal{B}(\mathbb{R})$.

Theorem 1.2 (Empirical measure). *Assume* (1.1). *Then, conditioned on $A_t = \{N_t > 0\}$, the random measures $\left\{\frac{\tilde{Y}_t}{N_t}\right\}$ converge as $t \to \infty$ in distribution to a random measure ν, characterized by its moment sequence $m_k(\phi) \equiv E[\langle \nu, \phi\rangle^k]$, for $\phi \in C_b^+(\mathbb{R}_+ \times \mathbb{R})$, $k \geq 1$.*

An explicit formula for $m_k(\phi)$ is given in (4.3) below.

Remark 1.1.

(i) Theorems 1.1 and 1.2 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 $0 < \alpha \leq 2$.

(ii) Equations (1.4) and (1.5) suggest that Theorem 1.1 should follow from the central limit theorem. But conditioning on non-extinction introduces dependencies in the lifetimes of the ancestors of the chosen individual. To take care of this we prove in §2 four propositions on age-dependent branching processes which are used in proving the result (see §3).

We now study the super-process scaling limit for the age-dependent branching process. For a review of the super-process literature we refer the reader to [7], [10], [11] and [16]. One may view a super-process as a renormalized limit of the empirical measure of a sequence of branching Markov processes. When motion depends on age or in the branching Markov process the position of the parent at the time of its death and the positions of its offspring (at the time of their birth) are not the same, the system is said to be non-local. In such systems the standard super-process scaling limit procedure does not work.

References [6] and [8] have considered two special cases where such a scaling limit has been obtained. In [6] an age-dependent branching process is rescaled (i.e. the particles do not perform any motion) and in [8] a general non-local super-process limit is obtained when the offspring distribution is given by $p_1 = 1$.

In order to obtain a super-process type limit, we scale the age and the motion differently. In the limit we obtain an age-'averaged' super-process. To illustrate this idea we consider a specific sequence of branching Markov processes $(G_n, \mathbf{p}_n, \eta_n)_{\{n \geq 1\}}$ denoted by $\{Y_t^n : t \geq 0\}_{\{n \geq 1\}}$ given by:

Notation 1.2.

(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 finite product measure $\nu = \alpha \times \mu \in M(\mathbb{R}_+ \times \mathbb{R})$.

(b) *Lifetime $G^n(\cdot)$.* For all $n \geq 1$, $G^n \equiv G$ is an exponential distribution with mean $\frac{1}{\lambda}$, $0 < \lambda < \infty$.

(c) *Offspring distribution $\mathbf{p}_n$.* For each $n \geq 1$, 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,

$$\lim_{n \to \infty} \sup_{0 \leq u \leq N} |n^2(F_n(1 - u/n) - (1 - u/n)) - cu^2| \to 0, \qquad (1.9)$$

for some $c > 0$ and all $N > 0$.

(d) *Motion process $\eta_n(\cdot)$.* For all $n \geq 1$,

$$\left\{ \eta_n(t) \equiv \frac{1}{\sqrt{n}} \int_0^t \sigma(u) \mathrm{d}B_n(u), \quad t \geq 0, \right\} \qquad (1.10)$$

where $\{B_n(t) : t \geq 0\}$ are independent standard Brownian motions starting at 0 and $\sigma : \mathbb{R}_+ \to \mathbb{R}$ is a continuous function such that for each $n \geq 1$, η_n satisfies (1.1).

We now define the limiting age-averaged super-process.

DEFINITION 1.1

Let $f \in C_l^+(\mathbb{R}_+ \times \mathbb{R})$ (the set of all nonnegative, continuous functions from $\mathbb{R}_+ \times \mathbb{R}$ to $\mathbb{R}$ with finite limits at infinity). Let $\{B_t : t \geq 0\}$ be a standard Brownian motion and $\mathcal{E}$ an exponential random variable with mean $\frac{1}{\lambda}$, $0 < \lambda < \infty$, independent of $\{B_t : t \geq 0\}$ and let ψ be as defined in (1.1). Let

$$U_t f(x) \equiv E(f(\mathcal{E}, x + \sqrt{\lambda \psi} B_t)), \quad x \in \mathbb{R}, t \geq 0.$$

Let $u_t(f)$ be the unique solution of the integral equation

$$u_t f(x) = U_t f(x) - c\lambda \int_0^t U_{t-s}(u_s(f)^2)(x) \, \mathrm{d}s, \qquad x \in \mathbb{R}, t \geq 0. \quad (1.11)$$

Let $\mathcal{Y} \equiv \{\mathcal{Y}_t : t \geq 0\}$ be a $M(\mathbb{R}_+ \times \mathbb{R})$-valued Markov process such that $\mathcal{Y}_0 \equiv \mathcal{E} \times \mu$, i.e. $\mathcal{Y}_0 \in M(\mathbb{R}_+ \times \mathbb{R})$ and for any $f \in C_b(\mathbb{R}_+ \times \mathbb{R})$,

$$\langle f, \mathcal{Y}_0 \rangle = \lambda \int_{\mathbb{R}_+ \times \mathbb{R}} f(a, x) \mathrm{e}^{-\lambda a} \, \mathrm{d}a \, \mu(\mathrm{d}x),$$

and whose Laplace functional is given by

$$E_{\mathcal{E} \times \mu}[e^{-\langle f, \mathcal{Y}_t \rangle}] \equiv E[e^{-\langle f, \mathcal{Y}_t \rangle}|\mathcal{Y}_0 = \mathcal{E} \times \mu] = e^{-\langle u_t(f), \mu \rangle}. \tag{1.12}$$

The function $u_s(f)$ in the second term of (1.11) 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 $a > 0, x \in \mathbb{R}$. It can be shown that there is a unique solution of the nonlinear integral equation (1.11) and there exists a $M(\mathbb{R}_+ \times \mathbb{R})$-valued Markov process $\mathcal{Y}$ satisfying (1.12) (see [7]).

Note that in the process $\mathcal{Y} \equiv \{\mathcal{Y}_t : t \geq 0\}$ defined above, the distribution of the age (i.e. the first coordinate) is exponential and does not change with time. 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 averaged (over lifetime) variance of the age-dependent motion.

Theorem 1.3 (Age structured super-process). *Let $\epsilon > 0$. Let $\{Y_t^n : t \geq 0\}_{n \geq 1}$ be the sequence of branching Markov processes defined above (i.e. in (a), (b), (c) and (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 $\{\mathcal{Y}_t : t \geq \epsilon\}$.*

The rest of the paper is organized as follows. In §2, we prove some preliminary results concerning age-dependent branching processes. Using a central limit type theorem (Proposition 2.3) proved in this section we prove Theorem 1.1 in §3. In §2, we also show that the joint distribution of coalescent times for a sample of $k \geq 1$ individuals chosen at random from the population at time t scaled by t converges, as $t \to \infty$ (see Theorem 2.1). This result is of independent interest and is a key tool in the proof of Theorem 1.2 in §4. Finally in §5 we prove Theorem 1.3.

2. Results on branching processes

Let $\{N_t : t \geq 0\}$ be the age-dependent branching process with offspring distribution $\{p_k\}_{k \geq 0}$ and lifetime distribution G as defined earlier. Let $\{\zeta_k\}_{k \geq 0}$ be the embedded discrete time Galton–Watson branching process with ζ_k the size of the k-th generation, $k \geq 0$. Recall that $A_t = \{N_t > 0\}$. On this event, choose an individual at random with uniform distribution from those alive at time t. Let M_t be the generation number and a_t be the age of this individual, respectively.

PROPOSITION 2.1

Let A_t, a_t, M_t and N_t be as above with $N_0 = 1$. Let μ and σ be as in §1.1. Then

(a) $\lim\limits_{t \to \infty} t P(A_t) = \frac{2\mu}{\sigma^2}$;

(b) *for all $x > 0$,* $\lim\limits_{t \to \infty} P\left(\frac{N_t}{t} > x | A_t\right) = e^{-\frac{2\mu x}{\sigma^2}}$;

(c) *for all $\epsilon > 0$,* $\lim\limits_{t \to \infty} P\left(\left|\frac{M_t}{t} - \frac{1}{\mu}\right| > \epsilon | A_t\right) = 0$;

(d) *for all $x > 0$,* $\lim\limits_{t \to \infty} P(a_t \leq x | A_t) = \frac{1}{\mu} \int_0^x (1 - G(s)) ds$.

Proof. For (a) and (b), see chapter 4 in [5]. For (c), see [9] and for (d) the proof from [3] for the super-critical case can be adapted. $\square$

PROPOSITION 2.2 (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 $E(|h(L_1)|) < \infty$ with $L_1 \overset{d}{=} G$. Then, as $t \to \infty$,

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

Proof. Let ϵ and $\epsilon_1 > 0$ be given and let $k_1(t) = t\left(\frac{1}{\mu} - \epsilon\right)$ and $k_2(t) = t\left(\frac{1}{\mu} + \epsilon\right)$. By Proposition 2.1 there exists $\delta > 0$, $\eta > 0$ and $t_0 > 0$ such that for all $t \geq t_0$,

$$t P(N_t > 0) > \delta \quad \text{and} \quad P(N_t \leq t\eta | A_t) < \epsilon_1; \tag{2.1}$$

$$P(M_t \in [k_1(t), k_2(t)]^c | A_t) < \epsilon_1. \tag{2.2}$$

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

$$\lim_{k \to \infty} P\left(\sup_{j \geq k}\left|\frac{1}{j}\sum_{i=1}^{j}h(L_i) - E(h(L_1))\right| > \epsilon\right) = 0. \tag{2.3}$$

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 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 $\{\zeta_{ktj}: 1 \leq j \leq \zeta_{kt}\}$. Call ζ_{ktj} *bad* if

$$\left|\frac{1}{k}\sum_{i=1}^{k}h(L_{ktji}) - E(h(L_1)))\right| > \epsilon, \tag{2.4}$$

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*$\}$. So,

$$P \text{ (the chosen line of descent at time } t \text{ is } bad, M_t \in [k_1(t), k_2(t)]) | A_t)$$

$$= \frac{1}{P(A_t)}E\left(\frac{\sum_{j=k_1(t)}^{k_2(t)}\zeta_{jt,b}}{N_t}; A_t\right).$$

Consequently,

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

$$= P \text{ (the chosen line of descent at time } t \text{ is } bad|A_t)$$

$$\leq P \text{ (the chosen line of descent at time } t \text{ is } bad, M_t \in [k_1(t), k_2(t)]) | A_t)$$

$$+ P(M_t \in [k_1(t)), k_2(t)]^c | A_t)$$

$$= \frac{1}{P(N_t > 0)} E\left(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; A_t\right) + P(M_t \in [k_1(t)), k_2(t)]^c | A_t)$$

$$= \frac{1}{P(N_t > 0)} E\left(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; N_t > t\eta\right)$$

$$+ \frac{1}{P(N_t > 0)} E\left(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{N_t}; 0 < N_t \leq t\eta\right)$$

$$+ P(M_t \in [k_1(t)), k_2(t)]^c | A_t)$$

$$\leq \frac{1}{P(N_t > 0)} E\left(\frac{\sum_{j=k_1(t)}^{k_2(t)} \zeta_{jt,b}}{t\eta}; N_t > t\eta\right) +$$

$$+ \frac{P(0 < N_t \leq t\eta)}{P(N_t > 0)} + P(M_t \in [k_1(t)), k_2(t)]^c | A_t)$$

$$= \frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b})$$

$$+ P(0 < N_t \leq t\eta | N_t > 0) + P(M_t \in [k_1(t)), k_2(t)]^c | A_t). \qquad (2.5)$$

For $t \geq t_0$ by (2.1) and (2.2), the last two terms in (2.5) are less than ϵ_1. The first term is equal to

$$\frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b})$$

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

$$= \frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_j)$$

$$\times P\left(\sum_{i=1}^{j} L_i \leq t < \sum_{i=1}^{j+1} L_i, \left|\frac{1}{j}\sum_{i=1}^{j} h(L_i) - E(h(L_1))\right| > \epsilon\right),$$

where the $\{L_i\}_{i \geq 1}$ are i.i.d. G. Using (2.1) and (since $m = 1$) $E(\zeta_j) = E(\zeta_0)$ we can conclude that

$$\frac{1}{t\eta P(N_t > 0)} \sum_{j=k_1(t)}^{k_2(t)} E(\zeta_{jt,b})$$

$$\leq E(\zeta_0) \frac{P\left(\sup_{j\geq k_1(t)} \left|\frac{1}{j}\sum_{i=1}^{j} h(L_i) - E(h(L_1))\right| > \epsilon\right)}{t\eta P(N_t > 0)}$$

$$\leq E(\zeta_0) \frac{P\left(\sup_{j\geq k_1(t)} \left|\frac{1}{j}\sum_{i=1}^{j} h(L_i) - E(h(L_1))\right| > \epsilon\right)}{\eta\delta}, \tag{2.6}$$

which by (2.3) goes to zero as $t \to \infty$. So we have shown that

$$\limsup_{t\to\infty} P\left(\left|\frac{1}{M_t}\sum_{i=1}^{M_t} h(L_{ti}) - E(h(L_1))\right| > \epsilon|A_t\right) < 2\epsilon_1.$$

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

In the following, for any collection $\mathcal{A}$ of random variables, $\sigma(\mathcal{A})$ will denote the σ-algebra generated by $\mathcal{A}$.

PROPOSITION 2.3

Assume (1.1) *holds. Let* $\{L_i\}_{i\geq 1}$ *be i.i.d. G and* $\{\eta_i\}_{i\geq 1}$ *be i.i.d. copies of* η *and independent of* $\{L_i\}_{i\geq 1}$. *For* $\theta \in \mathbb{R}$, $t \geq 0$ *define* $\phi(\theta, t) = Ee^{i\theta\bar{\eta}(t)}$. *Then there exists an event D, with* $P(D) = 1$ *and on D, for all* $\theta \in \mathbb{R}$,

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

where ψ *is as in* (1.1).

Proof. Recall from (1.1) that $v(t) = E(\eta^2(t)) < \infty$ for $t \geq 0$. By the strong law of large numbers and (1.1),

$$\frac{\sum_{j=1}^{n} v(L_j)}{n} \to \psi \quad \text{w.p. 1.} \tag{2.7}$$

Note that,

$$\prod_{i=1}^{n} \phi\left(\frac{\theta}{\sqrt{\sum_{j=1}^{n} v(L_j)}}, L_j\right) = E(e^{i\theta\sum_{j=1}^{n} X_{nj}}|\mathcal{F}), \tag{2.8}$$

where

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

and $\mathcal{F} \equiv \sigma(L_i: i \geq 1)$. We need to prove a central limit theorem for the triangular array $\{X_{nj}: 1 \leq j \leq n\}$. Given $\mathcal{F}$, $\{X_{ni}: 1 \leq i \leq n\}$ is a triangular array of independent random variables such that for $1 \leq i \leq n$, $E(X_{ni}|\mathcal{F}) = 0$, $\sum_{i=1}^{n} E(X_{ni}^2|\mathcal{F}) = 1$.

Let $\epsilon > 0$ be given. Let

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

Let D be the event on which (2.7) holds. Using the dominated convergence theorem, we have

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

on D. Thus, the Lindeberg–Feller central limit theorem (see [4]) implies, that on D, for all $\theta \in \mathbb{R}$,

$$
\lim_{n\to\infty} E(e^{i\theta \sum_{j=1}^{n} X_{nj}} | \mathcal{F}) \to e^{\frac{-\theta^2}{2}}.
$$

Combining this with (2.7) and (2.8) we have the result. $\qquad\square$

PROPOSITION 2.4

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 $H_t = \frac{1}{\sqrt{M_t}} \sum_{i=1}^{M_t} \eta_{ti}(L_{ti})$, and $\mathcal{L}_t \equiv \sigma(M_t, L_{ti} : 1 \leq i \leq M_t)$. Then

$$
E(|E(e^{i\theta H_t} | \mathcal{L}_t) - e^{-\frac{\theta^2 \psi}{2}}| | A_t) \to 0. \tag{2.9}
$$

Proof. Fix $\theta \in \mathbb{R}$, $\epsilon_1 > 0$ and $\epsilon > 0$. Replace the definition of 'bad' in (2.4) by

$$
\left| \prod_{i=1}^{k} \phi\left(\frac{\theta}{\sqrt{k}}, L_{ktji}\right) - e^{-\frac{\theta^2 \psi}{2}} \right| > \epsilon. \tag{2.10}
$$

By Proposition 2.3 we have

$$
\lim_{k\to\infty} P\left(\sup_{j\geq k} \left| \prod_{i=1}^{j} \phi\left(\frac{\theta}{\sqrt{j}}, L_i\right) - e^{-\frac{\theta^2 \psi}{2}} \right| > \epsilon \right) = 0. \tag{2.11}
$$

Using this in place of (2.3) and imitating the proof of Proposition 2.2 (since the details mirror that proof we avoid repeating them here), we obtain that for t sufficiently large

$$
P\left(\left| \prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - e^{-\frac{\theta^2 \psi}{2}} \right| > \epsilon_1 | A_t \right) < \epsilon. \tag{2.12}
$$

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

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

So,

$$\limsup_{t \to \infty} E(|E(e^{i\theta \frac{1}{\sqrt{M_t}} \sum_{i=1}^{M_t} \eta_i(L_{ti})}|\mathcal{L}_t) - e^{-\frac{\theta^2 \psi}{2}}||A_t)$$

$$= \limsup_{t \to \infty} E\left(\left|\prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - e^{-\frac{\theta^2 \psi}{2}}\right||A_t\right)$$

$$< \epsilon_1 + 2\limsup_{t \to \infty} P\left(\left|\prod_{i=1}^{M_t} \phi\left(\frac{\theta}{\sqrt{M_t}}, L_{ti}\right) - e^{-\frac{\theta^2 \psi}{2}}\right| > \epsilon_1|A_t\right)$$

$$= \epsilon_1 + 2\epsilon.$$

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

The above four 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.

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 death 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 last time and before t 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 $\tau_{j,t}$ be the last time before t when there are j ancestors of the k individuals $C_1, C_2, \ldots, C_k$.

Theorem 2.1.

(i) *For any i, j, $\lim_{t \to \infty} P\left(\frac{\tau_{C_i,C_j,t}}{t} \leq x|A_t\right) \equiv H(x)$ exists for all $x \geq 0$ and $H(\cdot)$ is an absolutely continuous distribution function on $[0, 1]$.*

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

Proof. The proof of (i) and (ii) (Case $k = 3$) can be found in [9]. It suffices to prove (ii).

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 $1 \leq i_1 < i_2 \cdots < i_p < k$ and $0 < r_1 < r_2 < \cdots < r_p < r_{p+1} < 1 = r_{p+2}$,

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

(2.13)

exists and is absolutely continuous.

Suppose that at time tr_1 there are $n_1 \in \mathbb{N}$ particles of which k_1 have descendants that survive till time tr_2. For each $1 \leq j \leq k_1$, suppose there are $n_{2j} \in \mathbb{N}$ 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 $\mathbf{k_2} = (k_{21}, \ldots, k_{2k_1})$ and $|\mathbf{k_2}| = \sum_{j=1}^{k_1} k_{2j}$. Inductively, for $i = 3, \ldots, p + 1$, at time tr_i, there are n_{ij} descendants for the j-th particle, $1 \leq j \leq |\mathbf{k_{i-1}}|$. For each such j, let k_{ij} out of n_{ij} have descendants that survive up till time tr_{i+1} (see figure 1 for an illustration).
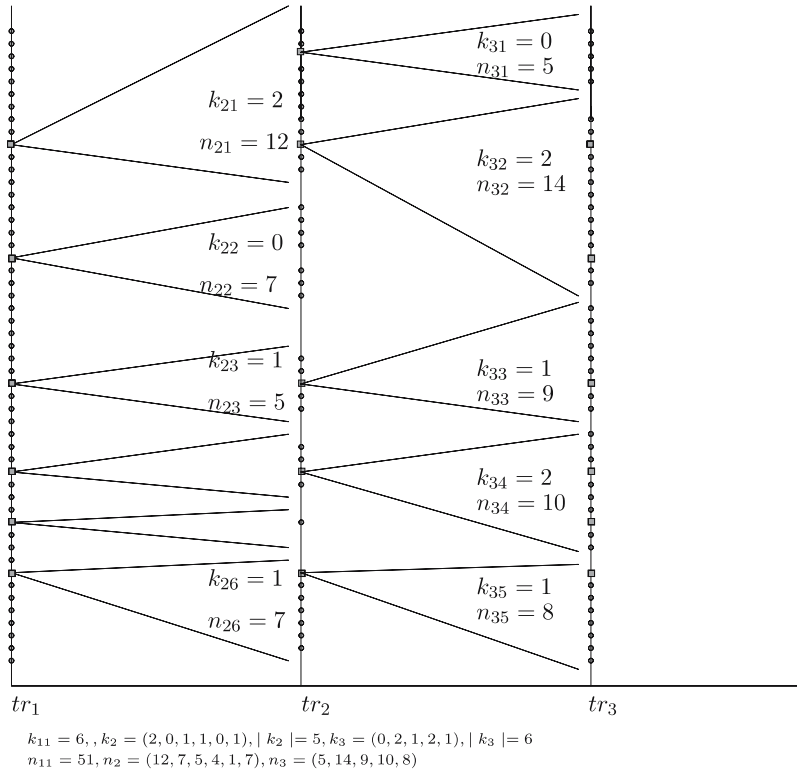
$k_{31} = 0$
$n_{31} = 5$

$k_{21} = 2$
$n_{21} = 12$

$k_{32} = 2$
$n_{32} = 14$

$k_{22} = 0$
$n_{22} = 7$

$k_{23} = 1$
$n_{23} = 5$

$k_{33} = 1$
$n_{33} = 9$

$k_{34} = 2$
$n_{34} = 10$

$k_{26} = 1$
$n_{26} = 7$

$k_{35} = 1$
$n_{35} = 8$

tr_1 tr_2 tr_3

$k_{11} = 6, , k_2 = (2, 0, 1, 1, 0, 1), | k_2 | = 5, k_3 = (0, 2, 1, 2, 1), | k_3 | = 6$
$n_{11} = 51, n_2 = (12, 7, 5, 4, 1, 7), n_3 = (5, 14, 9, 10, 8)$

Figure 1. Tracking particles surviving at various times.

Let $\mathbb{N}_1 = \{(n_1, k_1): n_1, k_1 \in \mathbb{N}, k_1 \leq n_1\}$. For $i = 2, \ldots p + 1$ let $\mathbf{n_i} = (n_{i1}, n_{i2}, \ldots, n_{i|\mathbf{k_{i-1}}|})$, $\mathbf{k_i} = (k_{i1}, k_{i2}, \ldots, k_{i|\mathbf{k_{i-1}}|})$, such that $n_{ij}, k_{ij} \in \mathbb{N}$, and $|\mathbf{k_i}| \equiv \sum_{j=1}^{|\mathbf{k_{i-1}}|} k_{ij}$.
Let $\mathbb{N}_i = \{(\mathbf{n_i}, \mathbf{k_i}): k_{ij} \leq n_{ij} \text{ for } 2 \leq j \leq |\mathbf{k_{i-1}}|\}$ and $\hat{\mathbb{N}} = \prod_{i=1}^{p+1} \mathbb{N}_i$.
Let $f_s \equiv P(N_s > 0 | N_0 = 1)$. Now,

$$P\left(\frac{\tau_{i_1}}{t} < r_1 < \frac{\tau_{i_2}}{t} < r_2 < \cdots < \frac{\tau_{i_p}}{t} < r_p < \frac{\tau_{k-1}}{t} < r_{p+1} < 1 | A_t\right)$$

$$= \frac{f_{tr_1}}{f_t} \sum_{\hat{n} \in \hat{\mathbb{N}}} \left(\binom{n_1}{k_1} (f_{tu_1})^{k_1} (1 - f_{tu_1})^{n_1 - k_1}\right) P(N_{tr_1} = n_1 | N_{tr_1} > 0)$$

$$\times \prod_{i=2}^{p+1} \prod_{j=1}^{|\mathbf{k_{i-1}}|} \left(\binom{n_{ij}}{k_{ij}} (f_{tu_i})^{k_{ij}} (1 - f_{tu_i})^{n_{ij} - k_{ij}}\right)$$

$$\times P(N_{tu_i}^{(j)} = n_{ij} | N_{tu_i}^{(j)} > 0) g(\mathbf{k}) E\left(\frac{\prod_{j=1}^{k} X^{(j)}}{S^k}\right),$$

with $u_i = r_{i+1} - r_i$, $i = 1, 2, \ldots, p + 1$, $N_{tu_i}^{(j)}$ is the number of particles alive at time tu_i of the age-dependent branching process starting with one particle namely j, $g(\mathbf{k}) = g(k_1, \ldots, k_p)$ is the proportion of configurations that have the desired number of

ancestors corresponding to the event of interest in (2.13), $X^{(j)} \stackrel{d}{=} N_{tu_{p+1}}^{(j)} | (N_{tu_{p+1}}^{(j)} > 0)$ and $S = \sum_{j=1}^{|\mathbf{k_{p+1}}|} X^{(j)}$.

Let $q_1 = \frac{r_1}{u_1}$ and $q_i = \frac{u_i}{u_{i+1}}$ for $2 \leq i \leq p+1$. Then following [9] and using Proposition 2.1(a) and (b) repeatedly we can show that $P\left(\frac{\tau_{i_1}}{t} < r_1 < \frac{\tau_{i_2}}{t} < r_2 < \cdots < \frac{\tau_{i_p}}{t} < r_p < \frac{\tau_{k-1}}{t} < r_{k-1} < 1 | A_t \right)$ converges to

$$\frac{1}{r_1} \sum_{k_i \in \mathbb{N}^{|\mathbf{k_{i-1}}|}} \int dx e^{-x} (q_1 x)^{k_1} \frac{1}{k_1!} e^{-xq_1}$$

$$\times \prod_{i=2}^{p+1} \prod_{j=1}^{|\mathbf{k_{i-1}}|} \int dx e^{-x} \frac{(q_i x)^{k_{ij}}}{k_{ij}!} e^{-xq_i} \times g(\mathbf{k})$$

$$\times \int \prod_{i=1}^{k+1} dx_i \left(\frac{\prod_{i=1}^{k} x_i}{(\sum_{i=1}^{k+1} x_i)^k} \right) e^{-\sum_{i=1}^{k+1} x_i} \frac{(x_{k+1})^{|\mathbf{k_{p+1}}|-(k+1)}}{(|\mathbf{k_{p+1}}| - (k+1))!}$$

$$= \frac{1}{r_1} \sum_{\mathbf{k_i} \in \mathbb{N}^{|\mathbf{k_{i-1}}|}} \prod_{i=1}^{p+1} \frac{(q_i)^{|\mathbf{k_i}|}}{(1+q_i)^{|\mathbf{k_i}|-|k_{i-1}|}} g(\mathbf{k})$$

$$\times \int \prod_{i=1}^{k+1} dx_i \left(\frac{\prod_{i=1}^{k} x_i}{(\sum_{i=1}^{k+1} x_i)^k} \right) e^{-\sum_{i=1}^{k+1} x_i} \frac{(x_{k+1})^{|\mathbf{k_{p+1}}|-(k+1)}}{(|\mathbf{k_{p+1}}| - (k+1))!}. \tag{2.14}$$

Consequently, we have shown that the random vector $\tilde{\tau}_t$ converges in distribution to a random vector $\tilde{T}$. Since the right-hand side of (2.14) is a differentiable function of $(r_1, r_2, \ldots, r_{k-1})$, it follows that $\tilde{T}$ has an absolutely continuous distribution on $[0, 1]^{k-1}$. □

3. Proof of Theorem 1.1

Using Notation 1.1 we have

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

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

$$\left(a_t, \frac{X_t}{\sqrt{t}} \right)$$

$$= \left(a_t, \sqrt{\frac{1}{\mu}} H_t \right) + \left(0, \left(\sqrt{\frac{M_t}{t}} - \sqrt{\frac{1}{\mu}} \right) H_t \right) + (0, J_t) + \left(0, \frac{X_0}{\sqrt{t}} \right),$$

where $H_t = \frac{\sum_{i=1}^{M_t} \eta_{ti}(L_{t_i})}{\sqrt{M_t}}$ and $J_t = \frac{1}{\sqrt{t}} \eta_{t(M_t+1)}(a_t)$. We shall show that as $t \to \infty$,

$$J_t | A_t \xrightarrow{p} 0, \tag{3.1}$$

and

$$\left(a_t, \frac{1}{\sqrt{\mu}} H_t \right) | A_t \xrightarrow{d} (U, V). \tag{3.2}$$

As $\frac{X_0}{\sqrt{t}} \xrightarrow{p} 0$, an application of Proposition 2.1(c) along with Slutsky's theorem will yield the result. So we now prove (3.1) and (3.2).

Proof of (3.1). Let $\epsilon > 0$ be given.

$$P(|J_t| > \epsilon | A_t) \le P(|J_t| > \epsilon, a_t \le k | A_t) + P(|J_t| > \epsilon, a_t > k | A_t)$$

$$\le P(|J_t| > \epsilon, a_t \le k | A_t) + P(a_t > k | A_t)$$

$$\le \frac{E(|J_t|^2 I_{a_t \le k} | A_t)}{\epsilon^2} + P(a_t > k | A_t).$$

By Proposition 2.1 and the ensuing tightness, for any $\eta > 0$ there is a k_η such that

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

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

$$E(|J_t|^2 I_{a_t \le k_\eta} | A_t) = E(I_{a_t \le k_\eta} E(|J_t|^2 | \mathcal{L}_t) | A_t)$$

$$= E\left(I_{a_t \le k_\eta} \frac{v(a_t)}{t} | A_t \right)$$

$$\le \frac{\sup_{u \le k_\eta} v(u)}{t}.$$

Hence,

$$P(|J_t| > \epsilon | A_t) \le \frac{\sup_{u \le k_\eta} v(u)}{t \epsilon^2} + \frac{\eta}{2}.$$

Since $\epsilon > 0$ and $\eta > 0$ are arbitrary this proves (3.1).

Proof of (3.2). Now we consider the second term. For $\lambda > 0$, $\theta \in \mathbb{R}$, as a_t is $\mathcal{L}_t$ measurable we have

$$E(e^{-\lambda a_t} e^{-i \frac{\theta}{\sqrt{\mu}} H_t} | A_t) = E(e^{-\lambda a_t}(E(e^{-i\theta H_t} | \mathcal{L}_t) - e^{-\frac{\theta^2 \psi}{2\mu}}) | A_t)$$

$$+ e^{-\frac{\theta^2 \psi}{2\mu}} E(e^{-\lambda a_t} | A_t).$$

Proposition 2.4 shows that the first term above converges to zero and using Proposition 2.1(d), (3.2) is immediate. $\qquad \square$

4. Proof of Theorem 1.2

Let $\phi \in C_b^+(\mathbb{R} \times \mathbb{R}_+)$. We shall show, for each $k \geq 1$, that the moment-functions $E\left(\frac{\langle \phi, \tilde{Y}_t \rangle^k}{N_t^k} | A_t\right)$ converge as $t \to \infty$. Then, by Theorem 16.16 in [14] the result follows.

The case $k = 1$ follows from Theorem 1.1 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), $i = 1, 2$.

$$
E\left(\left(\frac{\langle \phi, \tilde{Y}_t \rangle}{N_t}\right)^2 | A_t\right)
$$

$$
= E\left(\frac{\sum_{i=1}^{N_t} \phi^2\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right) + \sum_{\substack{i,j=1 \\ i \neq j}}^{N_t} \phi\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right) \phi\left(a_t^j, \frac{X_t^j}{\sqrt{t}}\right)}{N_t^2} | A_t\right)
$$

$$
= E\left(\frac{1}{N_t^2} \sum_{i=1}^{N_t} \phi^2\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right) | A_t\right)
$$

$$
+ E\left(\frac{1}{N_t^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_t} \phi\left(a_t^i, \frac{X_t^i}{\sqrt{t}}\right) \phi\left(a_t^j, \frac{X_t^j}{\sqrt{t}}\right) | A_t\right)
$$

$$
= E\left(\frac{1}{N_t} E\left(\phi^2\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) | N_t\right) | A_t\right)
$$

$$
+ E\left(\frac{N_t(N_t - 1)}{N_t^2} E\left(\phi\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) \phi\left(a_t^2, \frac{X_t^2}{\sqrt{t}}\right) | N_t\right) | A_t\right)
$$

$$
= E\left(\frac{1}{N_t} E\left(\phi^2(a_t^1, X_t^1) - \phi\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) \phi\left(a_t^2, \frac{X_t^2}{\sqrt{t}}\right) | N_t\right) | A_t\right)
$$

$$
+ E\left(E\left(\phi\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) \phi\left(a_t^2, \frac{X_t^2}{\sqrt{t}}\right) | N_t\right) | A_t\right)
$$

$$
= E\left(\frac{1}{N_t} E\left(\phi^2(a_t^1, X_t^1) - \phi\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) \phi\left(a_t^2, \frac{X_t^2}{\sqrt{t}}\right) | N_t\right) | A_t\right)
$$

$$
+ E\left(\phi\left(a_t^1, \frac{X_t^1}{\sqrt{t}}\right) \phi\left(a_t^2, \frac{X_t^2}{\sqrt{t}}\right) | A_t\right). \tag{4.1}
$$

As ϕ is bounded, the first term on the right hand side in (4.1) is bounded above by a constant times $E\left(\frac{1}{N_t} | A_t\right)$, which goes to 0 as $t \to \infty$ by Proposition 2.1(b). We will now analyse the second term in (4.1). Let $\tau_t = \tau_{C_1, C_2, t}$ be the death time of the common ancestor, say D, of C_1, C_2. Let the position of D at time τ_t be given 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 $X_t^i = \tilde{X}_{\tau_t} + X_{t-\tau_t}^i$, $i = 1, 2$.

Conditioned on the history up to the death of D, say $\mathcal{G}_t$, the random variables $(a_t^i, X_{t-\tau_t}^i), i = 1, 2$ are independent. By Theorem 2.1(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 1.1 conditioned on $\mathcal{G}_t$ and A_t, $\left\{ \left(a_t^i, \frac{X_{t-\tau_t}^i}{\sqrt{t-\tau_t}} \right), i = 1, 2 \right\}$ 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. Further, using Theorem 1.1 again, $\frac{\tilde{X}_{\tau_t}}{\sqrt{\tau_t}}$ conditioned on A_{τ_t} converges in distribution to a random variable S distributed as V and independent of $\{(U_i, V_i), i = 1, 2\}$.

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

$$\lim_{t \to \infty} E \left(\prod_{i=1}^2 \phi \left(a_t^i, \frac{X_t^i}{\sqrt{t}} \right) | A_t \right)$$

$$= E \prod_{i=1}^2 \phi(U_i, \sqrt{T}S + \sqrt{(1-T)}V_i) \equiv m_2(\phi) \quad \text{(say)}. \tag{4.2}$$

Hence we have shown that the second term of (4.1) converges to $m_2(\phi)$ and we are done.

Next, we consider the case $k > 2$. For any $\phi \in C_b^+(\mathbb{R} \times \mathbb{R}_+)$ one has

$$E \left(\frac{\langle \phi, \tilde{Y}_t \rangle^k}{N_t^k} | A_t \right) = E \left(\frac{N_t(N_t-1)(N_t-2) \cdots (N_t-k)}{N_t^k} \sum_{\mathbf{i} \in C_k} \prod_{j=1}^k \phi \left(a_t^{i_j}, \frac{X_t^{i_j}}{\sqrt{t}} \right) | A_t \right)$$

$$+ E \left(O \left(\frac{1}{N_t} \right) | A_t \right),$$

where C_k is the collection of indices of the k particles sampled without replacement from the particles alive at time t. We shall deem $C_k = $ on the event $\{N_t < k\}$. Consider one such sample, $\mathbf{i}$, and trace the genealogical tree $T \in \mathcal{T}_{\mathbf{i}}(k)$, ($\mathcal{T}_{\mathbf{i}}(k)$ is the collection of all possible labelled trees with k leaves given by $\mathbf{i}$), until their most recent common ancestor. For any leaf i_j in T, 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 ancestral death 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

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

Given the above notation we have

$$E \left(\prod_{j=1}^k \phi \left(a_t^{i_j}, \frac{X_t^{i_j}}{\sqrt{t}} \right) | A_t \right) = E \left(\sum_{T \in \mathcal{T}_{\mathbf{i}}} \prod_{j=1}^k \phi \left(a_t^{i_j}, \frac{1}{\sqrt{t}} \left(\tilde{\eta}_{i_j}^1(\tau_1) \right. \right. \right.$$

$$\left. \left. \left. + \sum_{m=2}^{N_{i_j}} \tilde{\eta}_{i_j}^m(\tau_{n(i_j,m)}, \tau_{n(i_j,m-1)}) + \tilde{\eta}_{i_j}'(t, \tau_{n(i_j, N_{i_j})}) | A_t \right) \right. \right).$$

Now by Theorem 2.1,

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

Further using Theorem 1.1 repeatedly, with appropriate conditioning as in the case $k = 2$ we obtain

$$\lim_{t \to \infty} E\left(\frac{\langle \phi, \tilde{Y}_t \rangle^k}{N_t^k} | A_t\right) = E\left(\sum_{\mathbf{i} \in C_k} \sum_{T \in \mathcal{T}_{\mathbf{i}}} \prod_{j=1}^{k} \phi\left(U_{i_j}, S_{i_j}\sqrt{T_1}\right.\right.$$

$$\left.\left. + \sum_{m=2}^{N_{i_j}} Z_{i_j}^m \sqrt{T_{n(i_j,m)} - T_{n(i_j,m-1)}} + Z'_{i_j}\sqrt{1 - T_{n(i_j,N_{i_j})}}\right) | A_t\right)$$

$$\equiv m_k(\phi), \tag{4.3}$$

where $S_{i_j}, Z'_{i_j}, Z_{i_j}^m$, $m = 2, \ldots, N_{i_j}$ are i.i.d. random variables distributed as V, and U_{i_j} are independent random variables distributed as U. For $\phi \in C_b(\mathbb{R})$, the sequence $\{m_k(\phi) : k \geq 1\}$ is necessarily a moment sequence of a unique probability distribution on $\mathbb{R}$. This being true for each ϕ, by Theorem 16.16 in [14] we are done. $\square$

5. Proof of Theorem 1.3

As stated in the theorem we scale the age and motion parameters differently, to obtain a super-process limit. Theorem 1.1 is used in establishing the limiting log-Laplace equation (Propositions 5.2 and 5.3). Tightness of the underlying particle system is shown in Proposition 5.4 and the result then follows by the standard techniques (see [7]).

Let $C_b(\mathbb{R}_+ \times \mathbb{R})$ be the set of bounded continuous function on $\mathbb{R}_+ \times \mathbb{R}$ and $C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ be the set of bounded functions which are differentiable in the first variable and are twice differentiable in the second variable. Let Z be the branching Markov process Y described earlier, with $Y_0 = \delta_{(a,x)}$, lifetime G exponential with mean λ, $p_1 = 1$ and $\eta \stackrel{d}{=} \eta_1$ (see (1.10)). Note that Z_t has the representation given by (1.3)–(1.5) with $X_0 = x$. For any $\phi \in C_b(\mathbb{R}_+ \times \mathbb{R})$, define the semigroup

$$S_t\phi(a, x) \equiv E_{(a,x)}[\langle Z_t, \phi \rangle] = E_{(a,x)}[\phi(a_t, X_t)].$$

Conditioning on the first birth we obtain the following evolution equation for S_t:

$$S_t\phi(a, x) = e^{-\lambda t} T_t\phi(a, x) + \int_0^t ds \, \lambda \, e^{-\lambda s} T_s(S_{t-s}(\phi)(0, \cdot))(a, x), \tag{5.1}$$

where T_t is the semigroup associated to $\mathcal{L}$ given by

$$\mathcal{L}f(a, x) = \frac{\partial f}{\partial a}(a, x) + \frac{\sigma^2(a)}{2}\Delta f(a, x), \tag{5.2}$$

with $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ and Δ being the one-dimensional Laplacian. Making a change of variable $s \to t - s$ in the second term of (5.1) and then differentiating it with respect to t, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t(\phi)(a, x) = -\lambda \mathrm{e}^{-\lambda t} T_t \phi(a, x) + \mathrm{e}^{-\lambda t} \mathcal{L} T_t \phi(a, x) + \lambda S_t(\phi)(0, x)$$

$$+ \int_0^t \mathrm{d}s \lambda (-\lambda \mathrm{e}^{-\lambda(t-s)}) T_{t-s}(S_s(\phi)(0, \cdot))(a, x)$$

$$+ \int_0^t \mathrm{d}s \lambda \mathrm{e}^{-\lambda(t-s)} \mathcal{L} T_{t-s}(S_s(\phi)(0, \cdot))(a, x)$$

$$= \lambda S_t(\phi)(0, x) + (\mathcal{L} - \lambda) \left[\mathrm{e}^{-\lambda t} T_t \phi(a, x) \right.$$

$$\left. + \int_0^t \mathrm{d}s \lambda \mathrm{e}^{-\lambda(t-s)} T_{t-s}(S_s(\phi)(0, \cdot))(a, x) \right]$$

$$= \lambda S_t(\phi)(0, x) + (\mathcal{L} - \lambda) S_t(\phi)(a, x)$$

$$= \frac{\partial S_t \phi}{\partial a}(a, x) + \frac{\sigma^2(a)}{2} \Delta S_t \phi(a, x) + \lambda (S_t(\phi)(0, x) - S_t(\phi)(a, x)).$$

Hence the generator of S_t is given by

$$\mathcal{R}_t f(a, x) = \mathcal{L} f(a, x) + \lambda (f(0, x) - f(a, x)), \quad f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}). \tag{5.3}$$

For each $n \geq 1$, define (another semigroup)

$$R_t^n \phi(a, x) \equiv E_{a,0} \left(\phi \left(a_t, x + \frac{X_t}{\sqrt{n}} \right) \right), \quad \phi \in C_b(\mathbb{R}_+ \times \mathbb{R}).$$

Now using the representation (1.3)–(1.4) and (1.10), it is easily seen that the semigroup R_t^n is the semigroup S_t with σ replaced by $\frac{\sigma}{\sqrt{n}}$. Hence, from (5.2), (5.3) it follows that the generator of the semigroup R_t^n will be given by

$$\mathcal{R}^n f(a, x) = \mathcal{L}^n f(a, x) + \lambda (f(0, x) - f(a, x)), \tag{5.4}$$

where

$$\mathcal{L}^n f(a, x) = \frac{\partial f}{\partial a}(a, x) + \frac{\sigma^2(a)}{2n} \Delta f(a, x), \quad f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}). \tag{5.5}$$

PROPOSITION 5.1

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

$$\sup_{(a,x) \in \mathbb{R}_+ \times \mathbb{R}} |R_{nt}^n(\phi)(a, x) - U_t(\phi)(x)| \to 0. \tag{5.6}$$

Proof. Let $t \geq \epsilon$. Applying Theorem 1.1 to the process Z, we have $\left(a_{nt}, \frac{X_{nt}}{\sqrt{n}} \right) \xrightarrow{d} (U, tV)$. The proposition is then immediate from the bounded convergence theorem and the fact that $\phi \in C_l^+(\mathbb{R}_+ \times \mathbb{R})$. $\qquad \square$

PROPOSITION 5.2

Let π_{nv} be as in Notation 1.2. *The log-Laplace functional of $\mathcal{Y}_t^n$ is given by*

$$E_{\pi_{nv}}[e^{-\langle \phi, \mathcal{Y}_t^n \rangle}] = e^{-\langle u_t^n \phi, v \rangle}, \qquad (5.7)$$

where

$$u_t^n \phi(a, x) = R_{nt}^n n(1 - e^{-\frac{\phi}{n}})(a, x) - \lambda \int_0^t ds\, R_{n(t-s)}^n \left(n^2 \Psi_n \left(\frac{u_s^n \phi}{n} \right) \right)(a, x), \qquad (5.8)$$

where

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

Proof. For any $n \in \mathbb{N}$, let Y_t^n be the sequence of branching Markov processes defined in §1.2. It can be shown that its log-Laplace functional L_t^n satisfies

$$L_{nt}^n \phi(a, x) = e^{-\lambda nt} T_{nt}^n[e^{-\phi}](a, x) + \int_0^{nt} ds\, \lambda e^{-\lambda s} T_s^n[F_n(L_{nt-s}^n \phi(0, \cdot))](a, x), \qquad (5.9)$$

where $t \geq 0$ and T_t^n is the semigroup associated with $\mathcal{L}^n$. 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

$$L_{nt}^n \phi(a, x) = T_{nt}^n[e^{-\phi}](a, x) + \lambda \int_0^{nt} T_{nt-s}^n (F_n(L_s^n \phi(0, \cdot)) - L_s^n \phi)(a, x) ds \qquad (5.10)$$

Therefore $v_t^n(\phi)(a, x) = 1 - L_t^n \phi(a, x)$, satisfies

$$v_{nt}^n \phi(a, x)$$

$$= T_{nt}^n(1 - e^{-\phi})(a, x) + \int_0^{nt} ds\, T_{nt-s}^n((1 - v_s^n \phi) - F_n(1 - v_s^n)\phi(0, \cdot)))(a, x)\lambda. \qquad (5.11)$$

Then for $0 \leq s < t$,

$$\frac{d}{ds} R_{n(t-s)}^n (v_{ns}^n(\phi))(a, x)$$

$$= -(n\mathcal{R}^n) R_{n(t-s)}^n (v_{ns}^n(\phi))(a, x) + R_{n(t-s)}^n \left(\frac{\partial}{\partial s} v_{ns}^n(\phi) \right)(a, x)$$

$$= -(n\mathcal{R}^n) R_{n(t-s)}^n (v_{ns}^n(\phi))(a, x)$$

$$\quad + R_{n(t-s)}^n (n\mathcal{L}^n T_{ns}^n(1 - e^{-\phi}) + n\lambda((1 - v_{ns}^n \phi) - F_n(1 - v_{ns}^n)\phi(0, \cdot)))(a, x))$$

$$\quad + R_{n(t-s)}^n \left(\int_0^{ns} dr\, n\mathcal{L}^n(T_{ns-r}^n((1 - v_r^n(\phi)) - F_n(1 - v_r^n)\phi(0, \cdot)) \right)(a, x)$$

$$= R_{n(t-s)}^n n(-\lambda(v_{ns}^n(\phi)(0, \cdot) - v_{ns}^n(\phi)) + \lambda((1 - v_{ns}^n \phi) - F_n(1 - v_{ns}^n)\phi(0, \cdot)))(a, x))$$

$$= -R_{n(t-s)}^n (n\Psi_n(v_{ns}^n \phi))(a, x).$$

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

$$v_{nt}^n(\phi)(a, x) = R_{nt}^n(1 - e^{-\phi})(a, x) - \int_0^t ds\, R_{n(t-s)}^n(n\Psi_n(v_{ns}^n\phi))(a, x).$$

(5.12)

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

$$E_{\pi_{n\nu}}[e^{-\langle\phi,\mathcal{Y}_t^n\rangle}] = E_{\pi_{n\nu}}[e^{-\langle\frac{\phi}{n},Y_{nt}^n\rangle}] = e^{\langle L_t^n(\frac{\phi}{n})-1,n\nu\rangle} = e^{-\langle nv_t^n(\frac{\phi}{n}),\nu\rangle}.$$

Set $u_t^n(\phi) \equiv nv_{nt}^n\left(\frac{\phi}{n}\right)$. From (5.12), it is easy to see that $u_t^n(\phi)$ satisfies (5.7). □

For any $f\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, 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 $f\colon \mathbb{R} \to \mathbb{R}$ as well.

PROPOSITION 5.3

Let $\epsilon > 0$. Let $\phi \in C_l^+(\mathbb{R}_+ \times \mathbb{R})$ (from Definition 1.1) and $u_t^n(\phi)$ be as in Proposition 5.2 and $u_t(\phi)$ be as in Theorem 1.3. Then for $t \geq \epsilon$,

$$\sup_{(a,x)\in\mathbb{R}_+\times\mathbb{R}} |u_t^n(\phi)(a, x) - u_t(\phi)(x)| \to 0.$$

(5.13)

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

$$u_t^n(\phi)(a, x)$$

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

$$= R_{nt}^n n(1 - e^{-\frac{\phi}{n}})(a, x)$$

$$\quad - \int_0^t ds\, R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0\cdot))))(a, x)$$

$$\quad - \lambda c \int_0^t ds\, R_{n(t-s)}^n(u_s^n\phi(0, \cdot)^2)(a, x).$$

Now

$$u^n t(\phi)(a, x) - u_t(\phi)(x)$$

$$= R_{nt}^n(n(1 - e^{-\frac{\phi}{n}}))(a, x) - U_t(\phi)(x)$$

$$\quad - \int_0^t ds\, R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0, \cdot))))(a, x)$$

$$\quad + \lambda c \int_0^t ds\, (U_{t-s}((u_s\phi)^2)(a, x) - R_{n(t-s)}^n(u_s^n\phi(0, \cdot)^2)(a, x))$$

$$= R_{nt}^n(n(1 - e^{-\frac{\phi}{n}}))(a, x) - U_t(\phi)(x)$$

$$\quad - \int_0^t ds\, R_{n(t-s)}^n(\varepsilon_n(u_s^n(\phi(0, \cdot))))(a, x)$$

$$+ \lambda c \int_0^t \mathrm{d}s\, R^n_{n(t-s)}((u_s\phi)^2 - u^n_s \phi(0, \cdot)^2)(a, x)$$

$$+ \lambda c \int_0^t \mathrm{d}s\, (U_{t-s}((u_s\phi)^2)(x) - R^n_{n(t-s)}(u_s\phi)^2)(a, x)).$$

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

$$\|u^n_t(\phi) - u_t(\phi)\|_\infty \le \|R^n_{nt}(n(1 - \mathrm{e}^{-\frac{\phi}{n}})) - U_t(\phi)\|_\infty + t\|\epsilon_n(u^n_s(\phi(0, \cdot))\|_\infty$$

$$+ 2\lambda c\|\phi\|_\infty \int_0^t \mathrm{d}s \|u^n_s(\phi) - u_s(\phi)\|_\infty$$

$$+ \lambda c \int_0^t \mathrm{d}s \|(U_{t-s} - R^n_{n(t-s)})(u_s\phi)^2\|_\infty.$$

For $\phi \in C_l(\mathbb{R}_+ \times \mathbb{R})$, note that U_t is a strongly continuous semigroup which implies that $u_s(\phi)$ is a uniformly continuous function. So using Proposition 5.1 the first term and the last term go to zero. By our assumption on F, $\|\epsilon_n(u^n_s(\phi(0, \cdot))\|_\infty$ will go to zero as $n \to \infty$. Now using the standard Gronwall argument we have the result. $\qquad\square$

PROPOSITION 5.4

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

Proof. By Theorems 3.7.1 and 3.6.5 (Aldous criterion) in [7], it is enough to show

$$\langle \mathcal{Y}^n_{\tau_n+\delta_n}, \phi \rangle - \langle \mathcal{Y}^n_{\tau_n}, \phi \rangle \xrightarrow{d} 0, \tag{5.14}$$

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 stopping time of the process $\mathcal{Y}^n$ with respect to the canonical filtration, satisfying $0 < \epsilon \le \tau_n \le T$ 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

$$P(\langle \mathcal{Y}^n_{\tau_n}, \phi \rangle > \gamma) \le \frac{1}{\gamma} c_1 \|\phi\|_\infty E(\sup_{\epsilon \le t \le T} \langle \mathcal{Y}^n_t, 1 \rangle) \le \frac{1}{\gamma} c_2 \|\phi\|_\infty. \tag{5.15}$$

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

$$L_n(\delta_n; \alpha, \beta) := E(\exp(-\alpha\langle \mathcal{Y}^n_{\tau_n+\delta_n}, \phi \rangle - \beta\langle \mathcal{Y}^n_{\tau_n}, \phi \rangle))$$

$$= E(\exp(-\langle \mathcal{Y}^n_{\tau_n}, u^n_{\delta_n}(\alpha\phi) + \beta\phi \rangle))$$

$$= E(\exp(-\langle \mathcal{Y}^n_{\tau_n-\epsilon}, u^n_\epsilon(u^n_{\delta_n}(\alpha\phi) + \beta\phi) \rangle)).$$

Therefore

$$|L_n(0; \alpha, \beta) - L_n(\delta_n; \alpha, \beta)|$$

$$\le \|u^n_\epsilon(u^n_{\delta_n}(\alpha\phi) + \beta\phi) - u^n_\epsilon((\alpha + \beta)\phi)\|_\infty E(\sup_{t \le T}\langle \mathcal{Y}^n_t, 1 \rangle)$$

$$\le c_1 \|u^n_\epsilon(u^n_{\delta_n}(\alpha\phi) + \beta\phi) - u^n_\epsilon((\alpha + \beta)\phi)\|_\infty,$$

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

$$\|u_\epsilon^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi) - u_\epsilon^n((\alpha + \beta)\phi)\|_\infty$$

$$\leq \|R_{n\epsilon}^n(u_{\delta_n}^n(\alpha\phi) - \alpha\phi)\|_\infty$$

$$+ c_2\|\phi\|_\infty \int_0^\epsilon \mathrm{d}a\|u_a^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi) - u_a^n((\alpha + \beta)\phi)\|_\infty + d_n(\phi),$$

where $d_n(\phi) = \lambda \int_0^\epsilon \mathrm{d}a\|\epsilon_n(u_a^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi)) + \epsilon_n(u_a^n((\alpha + \beta)\phi))\|_\infty$. Observe that

$$\|R_{n\epsilon}^n(u_{\delta_n}^n(\alpha\phi) - \alpha\phi)\|_\infty$$

$$\leq \|R_{n\epsilon}^n(u_{\delta_n}^n(\alpha\phi) - R_{n\delta_n}^n(\alpha\phi))\|_\infty + \|R_{n\epsilon}^n(R_{n\delta_n}^n(\alpha\phi) - \alpha\phi)\|_\infty$$

$$\leq \|u_{\delta_n}^n(\alpha\phi) - R_{n\delta_n}^n(\alpha\phi)\|_\infty + \|R_{n(\epsilon+\delta_n)}^n(\alpha\phi) - R_{n\epsilon}^n(\alpha\phi)\|_\infty$$

$$\leq \|R_{n\delta_n}^n(n(1 - \mathrm{e}^{\frac{\alpha\phi}{n}}) - \alpha\phi)\|_\infty$$

$$+ \int_0^{\delta_n} \mathrm{d}a \left\| R_{n(\delta_n-a)}^n \left(n^2\Psi_n\left(\frac{u_a^n(\alpha\phi)}{n}\right) \right) \right\|_\infty$$

$$+ \|R_{n(\epsilon+\delta_n)}^n(\alpha\phi) - R_{n\epsilon}^n(\alpha\phi)\|_\infty$$

$$\leq \|n(1 - \mathrm{e}^{\frac{\alpha\phi}{n}}) - \alpha\phi\|_\infty + \delta_n c_2(\|\phi\|_\infty^2 + 1)$$

$$+ \|R_{n(\epsilon+\delta_n)}^n(\alpha\phi) - R_{n\epsilon}^n(\alpha\phi)\|_\infty$$

$$\equiv e_n(\phi).$$

Consequently,

$$\|u_\epsilon^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi) - u_\epsilon^n((\alpha + \beta)\phi)\|_\infty$$

$$\leq e_n(\phi) + d_n(\phi) + c_2\|\phi\|_\infty \int_0^\epsilon \mathrm{d}a\|u_a^n(u_{\delta_n}^n(\alpha\phi) + \beta\phi) - u_a^n((\alpha + \beta)\phi)\|_\infty.$$

By Proposition 5.1, $e_n(\phi) \to 0$ and $d_n(\phi) \to 0$ by our assumption F_n. Hence by a standard Gronwall argument we have that

$$|L_n(0; s, r) - L_n(\delta_n; s, r)| \to 0. \tag{5.16}$$

By (5.15), $\{\langle \mathcal{Y}_{\tau_n}^n, \phi \rangle; n = 1, 2, \dots\}$ is tight in $\mathbb{R}_+$. Take an arbitrary subsequence. Then there is a further subsequence of it indexed by $\{n_k; k = 1, 2, \dots\}$ such that $\langle \mathcal{Y}_{\tau_{n_k}}^{n_k}, \phi \rangle$ converges in distribution to some random limit b. Thus we get

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

But (5.16) implies that

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

This implies that $\langle \mathcal{Y}_{\tau_{n_k}+\delta_{n_k}}^{n_k}, \phi \rangle - \langle \mathcal{Y}_{\tau_{n_k}}^{n_k}, \phi \rangle \xrightarrow{d} 0$ as $k \to \infty$. So (5.14) holds and the proof is complete. $\qquad\square$

Proof of Theorem 1.3. Proposition 5.3 shows that the log-Laplace functionals of the process $\mathcal{Y}_t^n$ converge to $\mathcal{Y}_t$ for every $t \geq \epsilon$. Proposition 5.4 implies tightness for the processes. As the solution to (1.11) is unique, we are done. □

Acknowledgements

References

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

[2] Athreya K B and Kang M, Some limit theorems for positive recurrent Markov chains, I and II, *Adv. Appl. Probab.* **30** (1998) 693–710, 711–722

[3] Athreya K B and Kaplan N, Convergence of the age distribution in the one-dimensional supercritical age-dependent branching process, *Ann. Probab.* **4(1)** (1976) 38–50

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

[5] Athreya K B and Ney P, Branching Processes (New York: Dover) (2000)

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

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

[8] Dawson D A, Gorostiza L G and Li Z, Nonlocal branching super-processes and some related models, *Acta Appl. Math.* **74(1)** (2002) 93–112

[9] Durrett R, The genealogy of critical branching processes, *Stoc. Proc. App.* **8** (1978)

[10] Dynkin E B, An introduction to branching measure-valued processes, CRM Monograph Series 6 (American Mathematical Society, Providence, RI) (1994)

[11] Etheridge A, An Introduction to Super-Processes, University Lecture Series 20 (American Mathematical Society, Providence, RI) (2000).

[12] Feller W, An introduction to probability theory and its applications, Vol. II (New York-London-Sydney: John Wiley & Sons, Inc.) (1966) xviii+636 pp.

[13] Ikeda N, Nagasawa M and Watanabe S, Branching Markov processes, I, *J. Math. Kyoto Univ.* **8** (1968) 365–410

[14] Kallenberg O, Foundation of Modern Probability Theory (Springer) (2002)

[15] Karlin S and McGregor J, Direct product branching processes and related induced Markov chains I, Calculations of rates of approach to homozygosity, Proc. Internat. Res. Sem. Stat. Lab., Univ. of Calif. Berkeley (California: Springer) (1963) pp. 111–145

[16] Perkins E A, Dawson-Watanabe superprocesses and measure-valued diffusions, Lectures on probability theory and statistics (Saint-Flour) (1999) pp. 125–324, Lecture Notes in Math. 1781 (Berlin: Springer) (2002)

[17] Seneta E and Vere-Jones D, On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states, *J. Appl. Probab.* **3** (1966) 403–434