The Annals of Applied Probability
1995, Vol. 5, No. 1, 140–157

DIFFUSION APPROXIMATION FOR AN AGE-STRUCTURED POPULATION

BY A. BOSE[1] AND I. KAJ[2]

Carleton University and Uppsala University

We prove a diffusion limit theorem in the sense of weak convergence of measure-valued processes for a population age model first studied by Kendall. We show that in the diffusion limit scaling, the population structured in age groups behaves in the same way as the total population size, but with an exponential weight. A particular feature of the limiting process is that in general it is discontinuous at time zero.

1. Introduction. Consider a population where each element (individual) either gives birth to a single offspring or dies, with constant and equal rates λ. Let $\mathbf{N} = (\mathbf{N}_t)_{t\geq 0}$ denote the population size. Then $\mathbf{N}$ is a linear, critical birth-and-death process on the nonnegative integers with jump intensities at time t given by $\lambda \mathbf{N}_t$.

We are interested in the distribution of age among the members of the population. In order to study the age distribution we will introduce a process $\mathbf{X} = (\mathbf{X}_t)_{t\geq 0}$ on a set of measures on $\mathbf{R}^+$, which is such that

$$\mathbf{X}_t([x_1, x_2]) = \text{ number of elements with age in the interval } [x_1, x_2].$$

Thus for bounded Borel measurable functions ϕ,

$$\langle \mathbf{X}_t, \phi \rangle = \int_0^\infty \phi(y)\, \mathbf{X}_t(dy).$$

In particular, if $\mathbf{1}$ denotes the indicator of $[0, \infty[$,

$$\mathbf{N}_t = \mathbf{X}_t(\mathbf{R}^+) = \langle \mathbf{X}_t, \mathbf{1} \rangle.$$

Let $\overline{\mathbf{N}}$ denote the continuous state branching process (with parameter λ) defined by the Laplace transition function

$$\mathbf{E}[\exp(-\theta \overline{\mathbf{N}}_t) \mid \overline{\mathbf{N}}_0 = 1] = \exp(-\theta(1 + \lambda t\theta)^{-1}), \qquad \theta \geq 0.$$

This is the critical branching diffusion with zero drift, variance function $2\lambda x$ and generator $L = \lambda x d^2/dx^2$; see, for example, Athreya and Ney [(1972), Chapter VI.6]. It is a classical result that $\overline{\mathbf{N}}$ can be viewed as a diffusion approximation of $\mathbf{N}$. In terms of weak convergence of processes, one has

$$\frac{1}{n} \sum_{j=1}^{n} \mathbf{N}_{nt}^{j} \overset{\mathscr{W}}{\Longrightarrow} \overline{\mathbf{N}}_t, \qquad n \to \infty,\ \overline{\mathbf{N}}_0 = 1, \tag{1.1}$$

Received November 1990; revised June 1994.
[1]Research partially supported by NSERC of Canada.
[2]Research supported by the Swedish Natural Sciences Research Council.
AMS 1991 *subject classifications*. 60J80, 60J25.
Key words and phrases. Measure-valued processes, age distribution, diffusion limit.

where $\{\mathbf{N}^j;\ j \geq 1\}$ is a sequence of i.i.d. copies of $\mathbf{N}$ with initial condition $\mathbf{N}_0 = 1$.

Our purpose is to demonstrate an analogous result for the measure-valued age process $\mathbf{X}$. To illustrate, let $\{\mathbf{X}^j;\ j \geq 1\}$ be a sequence of i.i.d. copies of the process $\mathbf{X}$ with initial condition $\mathbf{X}_0 = \delta_{\mathbf{Y}}$, where $\mathbf{Y}$ is an exponentially distributed random variable with mean $1/\lambda$. Each copy thus describes the distribution of age among the descendants of a single ancestor assumed to be of age $\mathbf{Y}$ at time $t = 0$. Let μ_λ denote the exponential distribution such that

$$\langle \mu_\lambda, \phi \rangle = \int_0^\infty \phi(y) \lambda e^{-\lambda y}\, dy.$$

We prove that with respect to weak convergence of measure-valued processes,

$$\frac{1}{n} \sum_{j=1}^{n} \mathbf{X}^j_{nt} \overset{\mathscr{W}}{\Longrightarrow} \overline{\mathbf{Z}}_t, \qquad n \to \infty, \tag{1.2}$$

where the limiting process $\overline{\mathbf{Z}}$ is the product of the continuous state branching process and the (deterministic) exponential measure:

$$\overline{\mathbf{Z}}_t = \overline{\mathbf{N}}_t \mu_\lambda.$$

For arbitrary initial conditions, finite-dimensional distributions converge toward the same limit process $\overline{\mathbf{Z}}$, eventually modified for a discontinuity point at the origin. To obtain the weak convergence when the $\overline{\mathbf{Z}}$ process is continuous at the origin, we need to impose a condition for the rate of convergence toward μ_λ.

The interpretation of the result is that in the diffusion limit the size of age groups behaves in the same way as the total population does, but with an exponential weight. For example, the diffusion limit of elements of age exceeding some number a is given by $e^{-\lambda a}\overline{\mathbf{N}}_t$.

After some preliminaries in Section 2, we give a precise formulation of the result in Section 3. The proof consists mainly of two parts—convergence of finite-dimensional distributions and tightness—dealt with in Sections 4 and 5, respectively.

2. Preliminaries. Let $\mathscr{C}_b(\mathbf{R}^+)$ denote the continuous and bounded and $\mathscr{C}_b^+(\mathbf{R}^+)$ denote the continuous, bounded and nonnegative functions on $\mathbf{R}^+$. Similarly, $\mathscr{C}_b^1(\mathbf{R}^+)$ denotes the continuous and bounded functions on $\mathbf{R}^+$ with continuous and bounded first derivatives. Also write $\mathscr{C}_c^{+,1}(\mathbf{R}^+)$ for the subset of $\mathscr{C}_b^1(\mathbf{R}^+)$ consisting of nonnegative functions with compact supports. Let $\mathscr{M}_b^+(\mathbf{R}^+)$ denote the dual cone of positive finite measures on $\mathbf{R}^+$. Assume that the function spaces are given their supremum norm topology with norm $\|\cdot\|_\infty$ and that $\mathscr{M}_b^+(\mathbf{R}^+)$ is endowed with the topology of weak convergence generated by $\mathscr{C}_b^+(\mathbf{R}^+)$. Let $\mathscr{D}([0, \infty[, \mathscr{M}_b^+(\mathbf{R}^+))$ denote the space of càdlàg functions on the time interval $[0, \infty[$ with values in $\mathscr{M}_b^+(\mathbf{R}^+)$, equipped with the Skorokhod topology.

The age distribution process $\mathbf{X}$ is a measure-valued Markov process with path space $\mathscr{D}([0,\infty[, \mathscr{M}_b^+(\mathbf{R}^+))$. Let $(\mathscr{F}_t)_{t\geq 0}$ denote the increasing sequence of σ-algebras generated by $\mathbf{X}$. The generator $\mathscr{G}$ of $\mathbf{X}$ on functionals of the form $F(\langle\cdot,\phi\rangle)$ with $F \in \mathscr{C}_b^1(\mathbf{R}^+)$, defined as

$$\lim_{h\to 0}\frac{1}{h}(\mathbf{E}[F(\langle \mathbf{X}_{t+h},\phi\rangle) \mid \mathscr{F}_t] - F(\langle \mathbf{X}_t,\phi\rangle)),$$

is given by

$$\begin{aligned}\mathscr{G}F(\langle\mu,\phi\rangle) &= \lambda\langle\mu,\mathbf{1}\rangle[F(\langle\mu,\phi\rangle+\phi(0)) - F(\langle\mu,\phi\rangle)]\\ &\quad+\lambda\int_0^\infty [F(\langle\mu,\phi\rangle-\phi(y)) - F(\langle\mu,\phi\rangle)]\mu(dy)\\ &\quad+F'(\langle\mu,\phi\rangle)\langle\mu,\phi'\rangle, \qquad \phi\in\mathscr{C}_b^1(\mathbf{R}^+),\ \mu\in\mathscr{M}_b^+(\mathbf{R}^+),\end{aligned}$$

and is such that the process

$$F(\langle \mathbf{X}_t,\phi\rangle) - F(\langle \mathbf{X}_0,\phi\rangle) - \int_0^t \mathscr{G}F(\langle \mathbf{X}_s,\phi\rangle)\,ds, \qquad t\geq 0,$$

is an $(\mathscr{F}_t)$-martingale. See Bose (1986).

The notion of the age distribution in the context of population growth models was introduced by Kendall (1949). In a subsequent paper, Kendall (1950), the following closed-form expression for the characteristic functional was given:

$$\begin{aligned}&\mathbf{E}[\exp(i\langle \mathbf{X}_t,\phi\rangle) \mid \mathbf{X}_0=\delta_x]\\ &\qquad=\frac{(e^{i\phi(x+t)}-1)e^{-\lambda t}+\int_0^t(e^{i\phi(y)}-1)\lambda e^{-\lambda y}\,dy}{1-\int_0^t(e^{i\phi(y)}-1)\lambda[1+\lambda(t-y]e^{-\lambda y}\,dy}\end{aligned} \tag{2.1}$$

We collect some of the basic properties of $\mathbf{X}$ in the next lemma. Throughout we use the notation $\mathbf{E}_x[\cdot]$ for the conditional expectation $\mathbf{E}[\cdot \mid \mathbf{X}_0=\delta_x]$.

LEMMA 2.1. *Let $\phi\in\mathscr{C}_b^+(\mathbf{R}^+)$. The age process $\mathbf{X}$ has the following properties:*

(i) *The expectation semigroup operator $T_t\phi(x) = \mathbf{E}_x[\langle \mathbf{X}_t,\phi\rangle]$ is given by*

$$T_t\phi(x) = \phi(x+t)e^{-\lambda t} + \int_0^t \phi(y)\lambda e^{-\lambda y}\,dy.$$

(ii) *The log-Laplace transition functional*

$$K_t\phi(x) = -\log \mathbf{E}_x[\exp(-\langle \mathbf{X}_t,\phi\rangle)]$$

defines a nonlinear contraction semigroup:

$$K_{t+s}\phi = K_t(K_s\phi), \qquad K_0\phi=\phi,$$

and we have

$$K_t\phi(x) = -\log\left(1 - \frac{T_t(1-e^{-\phi})(x)}{1+\lambda\int_0^t T_r(1-e^{-\phi})(0)\,dr}\right).$$

(iii) *The conditional mean and log-Laplace functions are given by*

$$\begin{aligned} \mathbf{E}[\langle \mathbf{X}_t, \phi\rangle \mid \mathscr{F}_s] &= \langle \mathbf{X}_s, T_{t-s}\phi\rangle, \\ -\log \mathbf{E}[\exp(-\langle \mathbf{X}_t, \phi\rangle) \mid \mathscr{F}_s] &= \langle \mathbf{X}_s, K_{t-s}\phi\rangle, \end{aligned} \qquad s \le t.$$

PROOF. The expression for the mean is derived in the usual way by differentiating the characteristic functional (2.1).

The semigroup property of K_t is a general feature of branching processes and is merely a reflection of the Chapman–Kolmogorov identity for the underlying transition density. The explicit form of $K_t\phi$ is obtained via a continuation of (2.1) to the Laplace functional, and then rewriting the result in terms of T_t. An immediate consequence of this representation is the contraction property $\|K_t\phi\|_\infty \le \|\phi\|_\infty$.

To compute the conditional expectations in (iii), note that particles alive at time s and of age x_s evolve according to independent copies of $\mathbf{X}$ with $\mathbf{X}_0 = \delta_{x_s}$. □

3. A diffusion approximation theorem. Let $\mathbf{X}$ be the age process discussed in the foregoing text. We consider a sequence $\{\mathbf{Y}^n;\ n \ge 1\}$ of nonnegative random variables on $\mathbf{R}^+$, which are not necessarily independent. For each $n \ge 1$ let $\{\mathbf{Y}^{j,n};\ 1 \le j \le n\}$ be i.i.d. copies of $\mathbf{Y}^n$. Let μ^n denote the mean measure defined by

$$\langle \mu^n, \phi\rangle = \mathbf{E}[\phi(\mathbf{Y}^n)].$$

Put $\mathbf{X}_0^{j,n} = \delta_{\mathbf{Y}^{j,n}}$. Given these initial conditions, let $\{\mathbf{X}^{j,n};\ 1 \le j \le n,\ n \ge 1\}$ denote a family of independent copies of $\mathbf{X}$ defined on $\mathscr{D}([0,+\infty[, \mathscr{M}_b^+(\mathbf{R}^+))$. Note that the superscript n in $\mathbf{X}^{j,n}$ only refers to a possibly varying initial state $\mathbf{Y}^{j,n}$.

We next define a sequence $\mathbf{Z}^n$, $n \ge 1$, of measure-valued processes in the state space $\mathscr{D}([0,+\infty[, \mathscr{M}_b^+(\mathbf{R}^+))$ by

$$\langle \mathbf{Z}_t^n, \phi\rangle = \frac{1}{n}\sum_{j=1}^n \langle \mathbf{X}_{nt}^{j,n}, \phi\rangle. \tag{3.1}$$

In particular,

$$\langle \mathbf{Z}_0^n, \phi\rangle = \frac{1}{n}\sum_{j=1}^n \phi(\mathbf{Y}^{j,n}).$$

Let $\overset{\mathscr{W}}{\Longrightarrow}$ denote weak convergence in the Skorokhod topology of $\mathscr{D}([0,\tau], \mathscr{M}_b^+(\mathbf{R}^+))$. Recall from the Introduction the critical branching diffusion denoted by $\overline{\mathbf{N}}$. For convenience we put $e(x) = \lambda e^{-\lambda x}$ so that

$$\langle \mu_\lambda, \phi \rangle = \int_0^\infty \phi(y) e(y)\, dy.$$

THEOREM 3.1. *Fix a time interval* $[0,\tau]$. *Suppose that the sequence* $\mathbf{Z}_0^n$ *has a limit in distribution* $\overline{\mathbf{Z}}_0$, *as* $n \to \infty$. *Then* $\mathbf{Z}^n$ *converges in finite-dimensional distributions toward a process* $\overline{\mathbf{Z}}$ *given by multiples of the continuous state branching process:*

$$\overline{\mathbf{Z}}_t = \begin{cases} \overline{\mathbf{Z}}_0, & t = 0, \\ \overline{\mathbf{N}}_t \mu_\lambda, & t > 0, \end{cases}$$

which is such that $\overline{\mathbf{N}}$ *is independent of* $\overline{\mathbf{Z}}_0$.

Now suppose $\overline{\mathbf{Z}}_0 = \mu_\lambda$ *so that* $\overline{\mathbf{Z}}_t$, $0 \le t \le \tau$, *is continuous. Suppose also that* μ^n *has a density* $g_n(x)$ *with respect to the Lebesgue measure on* $\mathbf{R}^+$ *and put* $a_n(x) = g_n(x)/e(x)$. *If* (a) $\|g_n - e\|_\infty \le C/n$ *and* (b) $\|a_n(\cdot) - a_n(\cdot + ns)\|_\infty \le Cs$, *then*

$$\mathbf{Z}^n \overset{\mathscr{W}}{\Longrightarrow} \overline{\mathbf{Z}}. \tag{3.2}$$

REMARKS.

1. Suppose the initial age variables $\mathbf{Y}^{j,n}$ are i.i.d. exponentially distributed with parameter λ (independent of n). Then by the law of large numbers $\mathbf{Z}_0^n \to \mu_\lambda$, almost surely in the topology of weak convergence of finite measures. Moreover, in this case the assumptions (a) and (b) are trivially fulfilled and hence (3.2) holds. This is the situation alluded to in the Introduction. Also,

$$\mathbf{E}[\langle \mathbf{X}_t, \phi \rangle] = \mathbf{E}[T_t \phi(\mathbf{Y})] = \langle \mu_\lambda, \phi \rangle, \qquad 0 \le t < \infty, \tag{3.3}$$

that is, the exponential distribution is invariant for the expectation operator T_t.

2. For large n and $t > 0$, $\mathbf{Z}_t^n$ is "close" to the limit $\mu_\lambda \overline{\mathbf{N}}_t$ with $\overline{\mathbf{N}}_t \to 1$ as $t \to 0$. The last part of the theorem reflects this behaviour. Indeed, we are not only forced to assume that $\overline{\mathbf{Z}}_0 = \mu_\lambda$, but also to impose a suitable rate of convergence toward the invariant measure in order to obtain weak convergence (3.2). The assumptions (a) and (b) can be weakened, but the relatively simple form above mirrors well the type of condition needed in the proof. As an alternative, we could let the $\mathbf{X}_0^{j,n}$, $1 \le j \le n$, be Poisson point measures on $\mathbf{R}^+$ with intensity measures μ^n converging to μ_λ. The same restrictions on the rate of convergence apply.

3. Other generalized versions of relation (1.1) are known. For example, under the corresponding scaling as in (3.1), the limiting behaviour of $\mathbf{N}$ is known for general branching mechanisms such as state-dependent rates and also for multivalued situations; see Joffe and Metiviér (1986). Concerning the age distribution, also in a very general setting, limit results are known for the case of *unscaled time* t. Then, typically, a law of large numbers holds with convergence toward a measure determined by the initial measure. As an example, consider the case we are interested in here with constant and equal jump rates. Then the measure-valued process $\mathbf{W}^n$, defined by

$$\langle \mathbf{W}_t^n, \phi\rangle = \frac{1}{n}\sum_{j=1}^{n}\langle \mathbf{X}_t^{j,n}, \phi\rangle,$$

converges weakly to a measure-valued process $\overline{\mathbf{W}}$, given by $\langle \overline{\mathbf{W}}_t, \phi\rangle = \langle \overline{\mathbf{Z}}_0, T_t\phi\rangle$, $t \geq 0$. For such results, see Bose (1986), Oelschläger (1990) and Borde-Boussion (1990).

As usual in a situation like this, a proof of weak convergence has two parts. We show next that all finite-dimensional distributions converge to the correct limits. In the last section we state and establish the required tightness property.

4. Finite-dimensional distributions. To prove convergence of finite-dimensional distributions for the population size process in (1.1) we may proceed from the observation that the Laplace function of $\mathbf{N}_t$ is given by

$$\psi(\theta, t) = \mathbf{E}[e^{-\theta \mathbf{N}_t} \mid \mathbf{N}_0 = 1] = 1 - \frac{1 - e^{-\theta}}{1 + \lambda t(1 - e^{-\theta})},$$

and then note that the limit

$$\lim_{n\to\infty} \psi^n(\theta/n, nt) = \exp(-\theta(1 + \lambda t\theta)^{-1})$$

is the Laplace function of $\overline{\mathbf{N}}_t$. To identify the limiting distribution of $\langle \mathbf{Z}_t^n, \phi\rangle$, we shall follow the same route.

For the proofs it is convenient to introduce the auxiliary notations

$$\overline{\phi} = \langle \mu_\lambda, \phi\rangle, \qquad \ell_t(\theta) = \frac{\theta}{1 + \lambda t\theta}$$

and

$$S_t\phi(x) = \frac{T_t(1 - e^{-\phi})(x)}{1 + \lambda \int_0^t T_r(1 - e^{-\phi})(0)\, dr}.$$

By the representation of K_t in Lemma 2.1(ii), then

$$K_t\phi = -\log(1 - S_t\phi).$$

LEMMA 4.1. *Fix $0 \le t_1 < \cdots < t_m \le \tau$ and positive constants $\theta_1, \ldots, \theta_m$. Then*

(i)
$$-\log \mathbf{E}\left[\exp\left(-\sum_{k=1}^{m} \theta_k \overline{\mathbf{N}}_{t_k}\right) \middle| \overline{\mathbf{N}}_0 = 1\right]$$
$$= \ell_{t_1}(\theta_1 + \ell_{t_2-t_1}(\cdots(\theta_{m-1} + \ell_{t_m-t_{m-1}}(\theta_m))\cdots)).$$

Furthermore, for ϕ_k, $k = 1, \ldots, m$, in $\mathscr{C}_b^+(\mathbf{R}^+)$,

(ii)
$$-\log \mathbf{E}_x\left[\exp\left(-\sum_{k=1}^{m} \langle \mathbf{X}_{t_k}, \phi_k \rangle\right)\right]$$
$$= K_{t_1}(\phi_1 + K_{t_2-t_1}(\cdots(\phi_{m-1} + K_{t_m-t_{m-1}}\phi_m)\cdots))(x),$$

and, for any $n \ge 1$,

(iii)
$$-\log \mathbf{E}_x\left[\exp\left(-\sum_{k=1}^{m} \langle \mathbf{Z}_{t_k}^n, \phi_k \rangle\right)\right]$$
$$= n\, K_{nt_1}\left(\frac{\phi_1}{n} + K_{n(t_2-t_1)}\left(\cdots\left(\frac{\phi_{m-1}}{n} + K_{n(t_m-t_{m-1})}\left(\frac{\phi_m}{n}\right)\right)\cdots\right)\right)(x).$$

PROOF. The first statement is a consequence of the conditioning formula

$$-\log \mathbf{E}[\exp(-\theta \overline{\mathbf{N}}_t) \mid \mathscr{F}_s] = \overline{\mathbf{N}}_s \ell_{t-s}(\theta), \qquad s \le t,$$

and a simple induction on m.

The corresponding relation (ii) for $\mathbf{X}$ follows in the same way by using Lemma 2.1(iii).

Given the initial configuration, the summands in $\mathbf{Z}^n$ are independent. Thus relation (ii) immediately implies (iii). □

LEMMA 4.2. *Let $\phi \in \mathscr{C}_b^+(\mathbf{R}^+)$. For $t > 0$,*

$$\|nK_{nt}(\phi/n) - \ell_t(\overline{\phi})\|_\infty \to 0, \qquad n \to \infty.$$

PROOF. We first note that if $(h_n)_{n\ge1}$ is a sequence in $\mathscr{C}_b^+(\mathbf{R}^+)$ with $\|h_n\|_\infty \le C$ for all $n \ge 1$, then

$$\| -n\log(1 - h_n/n) - h_n\|_\infty \le C^2/n, \qquad n > 2C. \tag{4.1}$$

For $h_n = nS_{nt}(\phi/n)$ and $C = \|\phi\|_\infty$, we see that to prove the lemma it is enough to show

$$\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\|_\infty \to 0, \qquad n \to \infty.$$

However, we have

$$\begin{aligned}\|nS_{nt}(\phi/n) - \ell_t(\overline{\phi})\|_\infty &\le \|nT_{nt}(1 - e^{-\phi/n}) - \overline{\phi}\|_\infty \\ &\quad + |\overline{\phi}A_n(\phi)^{-1} - \ell_t(\overline{\phi})|,\end{aligned} \tag{4.2}$$

where $A_n(\phi) = 1 + \lambda \int_0^t nT_{nr}(1 - e^{-\phi/n})(0)\,dr$. Moreover, for $n > 2C$,

$$\|nT_{nt}(1 - e^{-\phi/n}) - \overline{\phi}\|_\infty \le \|T_{nt}\phi - \overline{\phi}\|_\infty + C^2/n \le 2e^{-\lambda nt}\|\phi\|_\infty + C^2/n.$$

Hence $A_n(\phi)$ converges to $1+\lambda t\overline{\phi}$ and both terms in (4.2) vanish as $n \to \infty$. □

LEMMA 4.3. *For a sequence $f_n \in \mathscr{C}_b^+(\mathbf{R}^+)$, $n \ge 0$, with $C = \sup_n \|f_n\|_\infty$ and $t > 0$, one has*

$$\begin{aligned}&\|nK_{nt}(f_n/n) - nK_{nt}(f_0/n)\|_\infty\\ &\quad\le \|T_{nt}(f_n - f_0)\|_\infty + \lambda C \int_0^t |T_{ns}(f_n - f_0)(0)|\,ds\\ &\qquad + 2C^2(2 + \lambda tC)/n, \qquad n > 2C.\end{aligned}$$

PROOF. Take $n > 2C$. By (4.1),

$$\|nK_{nt}(f_n/n) - nK_{nt}(f_0/n)\|_\infty \le \|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty + 2C^2/n.$$

A straightforward estimate gives

$$\begin{aligned}&\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty\\ &\quad\le \|T_{nt}(n(1 - \exp(-f_n/n)) - n(1 - \exp(-f_0/n)))\|_\infty\\ &\qquad + \|nT_{nt}(1 - \exp(-f_0/n))\|_\infty |A_n(f_n)^{-1} - A_n(f_0)^{-1}|,\end{aligned}$$

with A_n as in the previous proof. Use (4.1) twice with $h_n = n(1 - e^{-f_n/n})$ and $h_n' = n(1 - e^{-f_0/n})$ to get

$$\|T_{nt}(n(1 - \exp(-f_n/n)) - n(1 - \exp(-f_0/n)))\|_\infty \le \|T_{nt}(f_n - f_0)\|_\infty + 2C^2/n.$$

Similarly,

$$|A_n(f_n)^{-1} - A_n(f_0)^{-1}| \le \lambda \int_0^t |T_{ns}(f_n - f_0)(0)|\,ds + 2\lambda tC^2/n.$$

Hence

$$\begin{aligned}&\|nS_{nt}(f_n/n) - nS_{nt}(f_0/n)\|_\infty\\ &\quad\le \|T_{nt}(f_n - f_0)\|_\infty + \lambda C \int_0^t |T_{ns}(f_n - f_0)(0)|\,ds + 2C^2(1 + \lambda tC)/n.\end{aligned}$$

Add the obtained remainder terms to finish the proof. □

LEMMA 4.4. *For all $0 = t_0 < t_1 < \cdots < t_m \le \tau$ and ϕ_k in $\mathscr{C}_b^+(\mathbf{R}^+)$, $k = 1, \dots, m$, we have*

$$\begin{aligned}\sup_x \Bigg|&\log \mathbf{E}_x\left[\exp\left(-\sum_{k=1}^m \langle \mathbf{Z}_{t_k}^n, \phi_k\rangle\right)\right]\\ &- \log \mathbf{E}\left[\exp\left(-\sum_{k=1}^m \theta_k \overline{\mathbf{N}}_{t_k}\right)\Bigg|\,\overline{\mathbf{N}}_0 = 1\right]\Bigg| \to 0, \qquad n \to \infty.\end{aligned}$$

PROOF. The proof is by induction on m and we use the shorthand notations

$$L^m(t_0,t_1,\dots,t_m) = \ell_{t_1-t_0}(\overline{\phi}_1 + \ell_{t_2-t_1}(\cdots(\overline{\phi}_{m-1} + \ell_{t_m-t_{m-1}}(\overline{\phi}_m))\cdots)),$$

$$H_n^m(t_0,t_1,\dots,t_m) = n\,K_{n(t_1-t_0)}\left(\frac{\phi_1}{n} + K_{n(t_2-t_1)}\left(\cdots\left(\frac{\phi_{m-1}}{n} + K_{n(t_m-t_{m-1})}\left(\frac{\phi_m}{n}\right)\right)\cdots\right)\right)(x).$$

According to Lemma 4.1 we need to prove that $\|H_n^m(t_0,t_1,\dots,t_m) - L^m(t_0, t_1,\dots,t_m)\|_\infty$ vanishes as $n\to\infty$. Lemma 4.2 provides a first induction step:

$$\|H_n^1(t_0,t_1) - L^1(t_0,t_1)\|_\infty \to 0, \qquad n\to\infty.$$

Now

$$\begin{aligned}\|H_n^m(t_0,t_1,\dots,t_m) - L^m(t_0,t_1,\dots,t_m)\|_\infty &\le \|nK_{nt_1}((\phi_1 + H_n^{m-1}(t_1,\dots,t_m))/n)\\ &\quad - nK_{nt_1}((\phi_1 + L^{m-1}(t_1,\dots,t_m))/n)\|_\infty\\ &\quad + \|nK_{nt_1}((\phi_1 + L^{m-1}(t_1,\dots,t_m))/n) - L^m(t_0,t_1,\dots,t_m)\|_\infty.\end{aligned}$$

Take $\phi = \phi_1 + L^{m-1}(t_1,\dots,t_m)$ in Lemma 4.2 to see that the last term on the right side converges to zero as $n\to\infty$. Moreover, if we take $n > 2C$ with $C = \sum_{j=1}^m \|\phi_j\|_\infty$, then by Lemma 4.3 and the contraction property of T_t,

$$\begin{aligned}&\|nK_{nt_1}((\phi_1 + H_n^{m-1}(t_1,\dots,t_m))/n) - nK_{nt_1}((\phi_1 + L^{m-1}(t_1,\dots,t_m))/n)\|_\infty\\ &\quad\le (1+\lambda tC)\big(\|H_n^{m-1}(t_1,\dots,t_m) - L^{m-1}(t_1,\dots,t_m)\|_\infty + 2C^2/n\big),\end{aligned}$$

which vanishes as $n\to\infty$, by the induction hypothesis. □

PROPOSITION 4.5. *For all $0 = t_0 < t_1 < \cdots < t_m \le \tau$ we have the joint convergence in finite-dimensional distributions:*

$$(\mathbf{Z}_0^n, \mathbf{Z}_{t_1}^n, \dots, \mathbf{Z}_{t_m}^n) \Rightarrow (\overline{\mathbf{Z}}_0, \overline{\mathbf{N}}_{t_1}\mu_\lambda, \dots, \overline{\mathbf{N}}_{t_m}\mu_\lambda), \qquad n\to\infty,$$

where the limiting continuous state branching process $\overline{\mathbf{N}}$ with $\overline{\mathbf{N}}_0 = 1$ is independent of $\overline{\mathbf{Z}}_0$.

PROOF. Fix ϕ_k in $\mathscr{C}_b^+(\mathbf{R}^+)$, $k = 0,\dots,m$. We must show that

$$\mathbf{E}\left[\exp\left(-\sum_{k=0}^m \langle \mathbf{Z}_{t_k}^n, \phi_k\rangle\right)\right]$$

converges as $n\to\infty$, toward

$$\mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0\rangle)]\,\mathbf{E}\left[\exp\left(-\sum_{k=1}^m \overline{\phi}_k\,\overline{\mathbf{N}}_{t_k}\right)\right].$$

Note that the case $m = 0$ corresponds to our assumption that $\mathbf{Z}_0^n$ has a limit $\overline{\mathbf{Z}}_0$ in distribution as $n \to \infty$. So we let $m = 1$. By Lemma 2.1(iii) and Lemma 4.1(iii) for $m = 1$, we have for $t > 0$,

$$\begin{aligned}\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0\rangle - \langle \mathbf{Z}_t^n, \phi_1\rangle)] &= \mathbf{E}\left[\exp\left(-\langle \mathbf{Z}_0^n, \phi_0\rangle - \frac{1}{n}\sum_{j=1}^n nK_{nt}\left(\frac{\phi_1}{n}\right)(\mathbf{Y}^{j,n})\right)\right] \\ &= \mathbf{E}\left[\exp\left(-\left\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}\left(\frac{\phi_1}{n}\right)\right\rangle\right)\right].\end{aligned}$$

Also, without making any assumptions on the dependence structure,

$$\begin{aligned}&\mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0\rangle)]\,\mathbf{E}[\exp(-\overline{\mathbf{N}}_t\,\overline{\phi}_1)] \\ &\quad= \mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0\rangle)]\exp(-\ell_t(\overline{\phi}_1)) \\ &\quad= \mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0 + \ell_t(\overline{\phi}_1)\rangle)],\end{aligned}$$

because $\ell_t(\overline{\phi}_1)$ is independent of x and $\langle \overline{\mathbf{Z}}_0, \mathbf{1}\rangle = 1$.

Hence,

$$\begin{aligned}&\left|\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0\rangle - \langle \mathbf{Z}_t^n, \phi_1\rangle)] - \mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0\rangle)]\,\mathbf{E}[\exp(-\overline{\mathbf{N}}_t\,\overline{\phi}_1)]\right| \\ &\quad= \left|\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}(\phi_1/n)\rangle)] - \mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0 + \ell_t(\overline{\phi}_1)\rangle)]\right| \\ &\quad\le \left|\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}(\phi_1/n)\rangle)] - \mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0 + \ell_t(\overline{\phi}_1)\rangle)]\right| \\ &\qquad+ \left|\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0\rangle)] - \mathbf{E}[\exp(-\langle \overline{\mathbf{Z}}_0, \phi_0\rangle)]\right|\exp(-\ell_t(\overline{\phi}_1)).\end{aligned}$$

Clearly, the last term goes to zero as $n \to \infty$ by our assumption on $\mathbf{Z}_0^n$. For the remaining term, we have

$$\begin{aligned}&\left|\mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0 + nK_{nt}(\phi_1/n)\rangle)] - \mathbf{E}[\exp(-\langle \mathbf{Z}_0^n, \phi_0 + \ell_t(\overline{\phi}_1)\rangle)]\right| \\ &\quad\le \mathbf{E}\left|\langle \mathbf{Z}_0^n, nK_{nt}(\phi_1/n) - \ell_t(\overline{\phi}_1)\rangle\right| \\ &\quad\le \|nK_{nt}(\phi_1/n) - \ell_t(\overline{\phi}_1)\|_\infty,\end{aligned}$$

which goes to zero by Lemma 4.2.

For the general case $m \ge 2$, the functions $nK_{nt}(\phi_1/n)$ and $\ell_t(\overline{\phi}_1)$ are replaced by the corresponding iterated functions in Lemma 4.1. In the final estimate above the norm $\|nK_{nt}(\phi_1/n) - \ell_t(\overline{\phi}_1)\|_\infty$ is then replaced by $\|H_n^m(t_0, t_1, \dots, t_m) - L^m(t_0, t_1, \dots, t_m)\|_\infty$. By Lemma 4.4, we are done. □

5. Tightness. To establish the tightness properties of the sequence $\{\mathbf{Z}^n\}$ required for Theorem 3.1, we will use the following general criterion. See Gorostiza and Lopez-Mimbela (1990), Theorem 2.1 ($p = 0$).

LEMMA 5.1. *The sequence* $\{\mathbf{Z}^n\}$ *is tight in* $\mathscr{D}([0, \tau], \mathscr{M}_b^+(\mathbf{R}^+))$ *if and only if for each* $\phi \in \mathscr{C}_b^+(\mathbf{R}^+)$ *the associated sequence of real-valued processes* $\{\langle \mathbf{Z}_t^n, \phi\rangle;\ 0 \le t \le \tau\}$ *is tight on* $\mathbf{R}^+$.

So we turn to the proof of (3.2), which relies on Lemma 5.1 and a fourth-moment calculation of the increments of $\langle \mathbf{Z}_t^n, \phi \rangle$.

For this purpose we introduce for $\nu \geq 1$ the functions

$$d_{t,s}^{(\nu)} = \mathbf{E}[(\langle \mathbf{X}_{t+s}, \phi \rangle - \langle \mathbf{X}_t, \phi \rangle)^\nu],$$

the moment functions

$$u_t^{(\nu)}(x) = \mathbf{E}_x[\langle \mathbf{X}_t, \phi \rangle^\nu] = (-1)^{\nu-1} \frac{\partial^\nu}{\partial \theta^\nu} S_t(\theta\phi)(x)\big|_{\theta=0},$$

the cumulants

$$k_t^{(\nu)} = (-1)^{\nu-1} \frac{\partial^\nu}{\partial \theta^\nu} K_t(\theta\phi)\big|_{\theta=0}$$

and the conditional moments

$$m_{t,s}^{(\nu)}(\mathbf{X}_t) = \mathbf{E}[\langle \mathbf{X}_{t+s}, \phi \rangle^\nu \mid \mathscr{F}_t].$$

We have suppressed the upper indices on $\mathbf{X}_t$, but it is understood that $\mathbf{X}_0$ still depends on the scaling parameter n. We write $\mu^n = \mathbf{E}[\mathbf{X}_0]$ so that, for example,

$$d_{t,s}^{(1)} = \langle \mu^n, T_{t+s}\phi - T_t\phi \rangle.$$

In particular, if $\mu^n = \mu_\lambda$, then $d_{t,s}^{(1)} = 0$, by (3.3). With this convention we can write

$$\begin{aligned}
&\mathbf{E}[|\langle \mathbf{Z}_{t+s}^n, \phi \rangle - \langle \mathbf{Z}_t^n, \phi \rangle|^4] \\
&\quad = \frac{1}{n^3} d_{nt,ns}^{(4)} + \frac{3(n-1)}{n^3} (d_{nt,ns}^{(2)})^2 + \frac{4(n-1)}{n^3} d_{nt,ns}^{(3)} d_{nt,ns}^{(1)} \\
&\quad\quad + \frac{6(n-1)(n-2)}{n^3} d_{nt,ns}^{(2)} (d_{nt,ns}^{(1)})^2 \\
&\quad\quad + \frac{(n-1)(n-2)(n-3)}{n^3} (d_{nt,ns}^{(1)})^4.
\end{aligned} \tag{5.1}$$

In case $\mathbf{E}[\mathbf{X}_0] = \mu_\lambda$, this formula simplifies to

$$\mathbf{E}[|\langle \mathbf{Z}_{t+s}^n, \phi \rangle - \langle \mathbf{Z}_t^n, \phi \rangle|^4] = \frac{1}{n^3} d_{nt,ns}^{(4)} + \frac{3(n-1)}{n^3} (d_{nt,ns}^{(2)})^2.$$

LEMMA 5.2. *For $\nu \geq 2$,*

(i) $$u_t^{(\nu)} = T_t(\phi^\nu) + \lambda \sum_{k=1}^{\nu-1} \binom{\nu}{k} \int_0^t T_{t-s}(u_s^{(k)}) u_s^{(\nu-k)}(0)\, ds, \qquad u_t^{(1)} = T_t\phi,$$

(ii) $$\begin{aligned} d_{t,s}^{(\nu)} &= \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi \rangle^\nu] \\ &\quad + \sum_{k=2}^{\nu} (-1)^{\nu-k} \binom{\nu}{k} \mathbf{E}[\langle \mathbf{X}_t, \phi \rangle^{\nu-k} (m_{t,s}^{(k)}(\mathbf{X}_t) - \langle \mathbf{X}_t, T_s\phi \rangle^k)]. \end{aligned}$$

For $\nu \geq 1$ *(and with* $m_{t,s}^{(0)} = 1$*),*

(iii)
$$m_{t,s}^{(\nu)}(\mathbf{X}_t) = \sum_{k=0}^{\nu-1} \binom{\nu-1}{k} \langle \mathbf{X}_t, k_s^{(\nu-k)} \rangle m_{t,s}^{(k)}(\mathbf{X}_t), \quad a.e.$$

PROOF. It is easy to verify that $S_t\phi$ satisfies the integral equation

$$S_t\phi = T_t(1-e^{-\phi}) - \lambda \int_0^t T_{t-s}(S_s\phi S_s\phi(0))\, ds.$$

Apply S_t to $\theta\phi$ and differentiate ν times with respect to θ to get

$$u_t^{(\nu)} = T_t(\phi^\nu) + \lambda \int_0^t T_{t-s}\left(\sum_{k=1}^{\nu-1} \binom{\nu}{k} u_s^{(k)} u_s^{(\nu-k)}(0) \right) ds$$

which is (i).

Next, by the binomial expansion,

(5.2)
$$d_{t,s}^{(\nu)} = \sum_{k=0}^{\nu} (-1)^{\nu-k} \binom{\nu}{k} \mathbf{E}[\langle \mathbf{X}_t, \phi \rangle^{\nu-k} m_{t,s}^{(k)}(\mathbf{X}_t)].$$

Obviously, the moment functions u_t^ν defined previously for $\phi \in \mathscr{C}_b^+(\mathbf{R}^+)$ can be extended to $\phi \in \mathscr{C}_b(\mathbf{R}^+)$. For example,

$$\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi \rangle^\nu] = \sum_{k=0}^{\nu} (-1)^{\nu-k} \binom{\nu}{k} \mathbf{E}[\langle \mathbf{X}_t, \phi \rangle^{\nu-k} \langle \mathbf{X}_t, T_s\phi \rangle^k].$$

Subtract this from (5.2) to obtain (ii).

To see (iii), use Lemma 2.1(iii) and note that

$$\begin{aligned} &\frac{\partial^\nu}{\partial\theta^\nu} \mathbf{E}[\exp(-\langle \mathbf{X}_{t+s}, \theta\phi \rangle) \mid \mathscr{F}_t] \\ &\quad = \frac{\partial^\nu}{\partial\theta^\nu} \exp(-\langle \mathbf{X}_t, K_s(\theta\phi) \rangle) \\ &\quad = \frac{\partial^{\nu-1}}{\partial\theta^{\nu-1}} \left(\exp(-\langle \mathbf{X}_t, K_s(\theta\phi) \rangle) \left\langle \mathbf{X}_t, -\frac{\partial}{\partial\theta} K_s(\theta\phi) \right\rangle \right). \end{aligned}$$

Then differentiate the product $\nu - 1$ times and put $\theta = 0$. □

Now we look more closely at the first four moments. By Lemma 5.2(iii),

$$\begin{aligned} m_{t,s}^{(1)}(\mathbf{X}_t) &= \langle \mathbf{X}_t, T_s\phi \rangle, \\ m_{t,s}^{(2)}(\mathbf{X}_t) &= \langle \mathbf{X}_t, k_s^{(2)} \rangle + \langle \mathbf{X}_t, k_s^{(1)} \rangle^2, \\ m_{t,s}^{(3)}(\mathbf{X}_t) &= \langle \mathbf{X}_t, k_s^{(3)} \rangle + 3\langle \mathbf{X}_t, k_s^{(1)} \rangle \langle \mathbf{X}_t, k_s^{(2)} \rangle + \langle \mathbf{X}_t, k_s^{(1)} \rangle^3, \\ m_{t,s}^{(4)}(\mathbf{X}_t) &= \langle \mathbf{X}_t, k_s^{(4)} \rangle + 4\langle \mathbf{X}_t, k_s^{(1)} \rangle \langle \mathbf{X}_t, k_s^{(3)} \rangle + 6\langle \mathbf{X}_t, k_s^{(1)} \rangle^2 \langle \mathbf{X}_t, k_s^{(2)} \rangle \\ &\quad + \langle \mathbf{X}_t, k_s^{(1)} \rangle^4 + 3\langle \mathbf{X}_t, k_s^{(2)} \rangle^2. \end{aligned}$$

Then by Lemma 5.2(ii),

$$
\begin{aligned}
d_{t,s}^{(1)} &= \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle],\\
d_{t,s}^{(2)} &= \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2] + \mathbf{E}[\langle \mathbf{X}_t, k_s^{(2)}\rangle],\\
d_{t,s}^{(3)} &= \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^3] + 3\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle\langle \mathbf{X}_t, k_s^{(2)}\rangle]\\
&\quad + \mathbf{E}[\langle \mathbf{X}_t, k_s^{(3)}\rangle],\\
d_{t,s}^{(4)} &= \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^4] + 6\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2\langle \mathbf{X}_t, k_s^{(2)}\rangle]\\
&\quad + 4\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle\langle \mathbf{X}_t, k_s^{(3)}\rangle] + 3\mathbf{E}[\langle \mathbf{X}_t, k_s^{(2)}\rangle^2]\\
&\quad + \mathbf{E}[\langle \mathbf{X}_t, k_s^{(4)}\rangle].
\end{aligned}
\tag{5.3}
$$

Furthermore, we note that the functions $u_t^{(\nu)}$ for $\nu = 1, 2, 3$ and 4 as they unfold from the recursion formula in Lemma 5.2 are given by

$$
\begin{aligned}
u_t^{(1)} &= T_t\phi,\\
u_t^{(2)} &= T_t(\phi^2) + 2\lambda T_t\phi \int_0^t T_s\phi(0)\,ds,\\
u_t^{(3)} &= T_t(\phi^3) + 3\lambda T_t\phi \int_0^t T_s(\phi^2)(0)\,ds + 3\lambda T_t(\phi^2)\int_0^t T_s\phi(0)\,ds\\
&\qquad + 6\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)^2,\\
u_t^{(4)} &= T_t(\phi^4) + 4\lambda T_t\phi \int_0^t T_s(\phi^3)(0)\,ds + 6\lambda T_t(\phi^2)\int_0^t T_s(\phi^2)(0)\,ds\\
&\qquad + 4\lambda T_t(\phi^3)\int_0^t T_s\phi(0)\,ds + 24\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)\\
&\qquad \times\left(\int_0^t T_s(\phi^2)(0)\,ds\right)\\
&\qquad + 12\lambda^2 T_t(\phi^2)\left(\int_0^t T_s\phi(0)\,ds\right)^2 + 24\lambda^3 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)^3,
\end{aligned}
$$

and since the first four elements in the sequences of functions $u^{(\nu)}$ and $k^{(\nu)}$ are related by

$$
\begin{aligned}
k_t^{(1)} &= u_t^{(1)},\\
k_t^{(2)} &= u_t^{(2)} - (u_t^{(1)})^2,\\
k_t^{(3)} &= u_t^{(3)} - 3u_t^{(2)}u_t^{(1)} + 2(u_t^{(1)})^3,\\
k_t^{(4)} &= u_t^{(4)} - 4u_t^{(3)}u_t^{(1)} - 3(u_t^{(2)})^2 + 12u_t^{(2)}(u_t^{(1)})^2 - 6(u_t^{(1)})^4,
\end{aligned}
$$

we have

$$k_t^{(1)} = T_t\phi,$$

$$k_t^{(2)} = T_t(\phi^2) - (T_t\phi)^2 + 2\lambda T_t\phi \int_0^t T_s\phi(0)\,ds,$$

$$\begin{aligned} k_t^{(3)} = {} & T_t(\phi^3) - 3T_t\phi T_t(\phi^2) + 2(T_t\phi)^3 + 3\lambda T_t\phi \int_0^t T_s(\phi^2)(0)\,ds \\ & + 3\lambda T_t(\phi^2)\int_0^t T_s\phi(0)\,ds - 6\lambda(T_t\phi)^2\int_0^t T_s\phi(0)\,ds, \\ & + 6\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)^2, \end{aligned}$$

$$\begin{aligned} k_t^{(4)} = {} & T_t(\phi^4) - 4T_t\phi T_t(\phi^3) - 3(T_t(\phi^2))^2 + 12T_t(\phi^2)(T_t\phi)^2 - 6(T_t\phi)^4 \\ & + 4\lambda T_t\phi \int_0^t T_s(\phi^3)(0)\,ds + 6\lambda T_t(\phi^2)\int_0^t T_s(\phi^2)(0)\,ds \\ & + 4\lambda T_t(\phi^3)\int_0^t T_s\phi(0)\,ds - 24\lambda T_t\phi T_t(\phi^2)\int_0^t T_s\phi(0)\,ds \\ & - 12\lambda(T_t\phi)^2\int_0^t T_s(\phi^2)(0)\,ds + 24\lambda(T_t\phi)^3\int_0^t T_s\phi(0)\,ds \\ & + 24\lambda^2 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)\left(\int_0^t T_s(\phi^2)(0)\,ds\right) \\ & + 12\lambda^2 T_t(\phi^2)\left(\int_0^t T_s\phi(0)\,ds\right)^2 - 36\lambda^2(T_t\phi)^2\left(\int_0^t T_s\phi(0)\,ds\right)^2 \\ & + 24\lambda^3 T_t\phi\left(\int_0^t T_s\phi(0)\,ds\right)^3. \end{aligned}$$

From now on, $\phi \in \mathscr{C}_c^{+,1}(\mathbf{R}^+)$ and C denotes a positive constant whose numerical value may change from one occurrence to another and that value may depend on ϕ and ϕ'. From the explicit formulas for $u_t^{(\nu)}$, $\nu = 1, \ldots, 4$, we first conclude

$$u_t^{(\nu)} \le \begin{cases} C(\|\phi\|_\infty + \int_0^t T_r\phi(0)\,dr), & \nu = 2, \\ C(\|\phi\|_\infty^2 + t(\|\phi\|_\infty + \int_0^t T_r\phi(0)\,dr)), & \nu = 3, \\ C(\|\phi\|_\infty^3 + t(\|\phi\|_\infty + \int_0^t T_r\phi(0)\,dr)^2), & \nu = 4. \end{cases}$$

Then we note the estimates

$$\|T_s\phi - \phi\|_\infty \le Cs, \qquad \left|\int_0^t (T_{r+s}\phi - T_r\phi)(0)\,dr\right| \le Cs.$$

In fact, the integral expression equals

$$\int_0^t (\exp(-\lambda(r+s))\phi(r+s) - \exp(-\lambda r)\phi(r))\,dr + \int_0^t \int_r^{r+s} \lambda \exp(-\lambda v)\phi(v)\,dv\,dr.$$

Clearly the second term is bounded by $\|\phi\|_\infty s$. Add and subtract $e^{-\lambda(r+s)}\phi(r)$ to get the upper bound $(\|\phi'\|_\infty + \|\phi\|_\infty)\, s$ for the first term. The argument for $T_s\phi - \phi$ is similar and simpler.

Writing $u_t(x) = u_t[\phi](x)$, in order to emphasize the particular function ϕ that is used to form the moment functions, and also remembering that $\mathbf{X}_0 = \delta_{\mathbf{Y}}$, we conclude

$$\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^\nu] = \langle \mu^n, u_t^{(\nu)}[T_s\phi - \phi]\rangle.$$

Hence we also obtain

$$\big|\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^\nu]\big| \le \begin{cases} Cs, & \nu = 2, \\ C(s^2 + st), & \nu = 3, \\ C(s^3 + s^2 t), & \nu = 4. \end{cases} \tag{5.4}$$

Moreover, from the preceding explicit formulas for the functions $k_t^{(\nu)}$, $\nu = 1, \dots, 4$, in terms of T_t, using again $\|T_t\phi - \phi\|_\infty \le Ct$, obtain

$$|k_t^{(\nu)}| \le \begin{cases} Ct, & \nu = 2, \\ C(t + t^2), & \nu = 3, \\ C(t + t^2 + t^3), & \nu = 4. \end{cases}$$

Consequently,

$$\big|\mathbf{E}[\langle \mathbf{X}_t, k_s^{(\nu)}\rangle]\big| \le \begin{cases} Cs, & \nu = 2, \\ C(s + s^2), & \nu = 3, \\ C(s + s^2 + s^3), & \nu = 4. \end{cases} \tag{5.5}$$

Furthermore, because $\mathbf{E}[\mathbf{N}_t^2] = (1 + 2\lambda t)$,

$$\begin{aligned} &\big|\mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle\langle \mathbf{X}_t, k_s^{(\nu)}\rangle]\big| \\ &\quad\le \begin{cases} C\mathbf{E}[\mathbf{N}_t^2]s \le C(s + st), & \nu = 2, \\ C\mathbf{E}[\mathbf{N}_t^2](s + s^2) \le C(s + s^2 + st + s^2 t), & \nu = 3, \end{cases} \end{aligned} \tag{5.6}$$

and

$$\mathbf{E}[\langle \mathbf{X}_t, k_s^{(2)}\rangle^2] \le C\mathbf{E}[\mathbf{N}_t^2]s^2 \le C(s^2 + s^2 t). \tag{5.7}$$

Finally, by the Schwarz inequality, (5.4) and (5.7),

$$\begin{aligned} \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^2\langle \mathbf{X}_t, k_s^{(2)}\rangle] &\le \mathbf{E}[\langle \mathbf{X}_t, T_s\phi - \phi\rangle^4]^{1/2}\mathbf{E}[\langle \mathbf{X}_t, k_s^{(2)}\rangle^2]^{1/2} \\ &\le C(s^3 + s^2 t)^{1/2}(s^2 + s^2 t)^{1/2} \\ &= Cs^2(t + s + st + t^2)^{1/2}. \end{aligned} \tag{5.8}$$

Putting the estimates (5.4) to (5.8) into (5.3), we obtain

$$(5.9)\quad |d_{t,s}^{(\nu)}| \le \begin{cases} Cs, & \nu = 2, \\ C(s + s^2 + st), & \nu = 3, \\ C(s^3 + ts^2 + s^2(t + s + st + t^2)^{1/2} + s + s^2 + st, & \nu = 4. \end{cases}$$

LEMMA 5.3. *Fix* $\phi \in \mathscr{C}_c^{+,1}(\mathbf{R}^+)$ *and suppose that* $\overline{\mathbf{Z}}_0 = \mu_\lambda$. *Under the assumptions* (a) *and* (b) *of Theorem* 3.1 *we have, for* $0 \le t \le \tau$ *and* $0 \le s \le 1$,

$$(5.10)\qquad \mathbf{E}[|\langle \mathbf{Z}_{t+s}^n, \phi\rangle - \langle \mathbf{Z}_t^n, \phi\rangle|^4] \le C\left(s^2 + \frac{s}{n}\right).$$

PROOF. By (5.9),

$$\frac{d_{nt,ns}^{(2)}}{n} \le Cs, \qquad \left|\frac{d_{nt,ns}^{(3)}}{n^2}\right| \le C\left(s^2 + st + \frac{s}{n}\right)$$

and

$$\begin{aligned}\frac{d_{nt,ns}^{(4)}}{n^3} &\le C\left(s^3 + ts^2 + s^2\sqrt{st + t^2 + \frac{s+t}{n}} + \frac{s^2 + st}{n} + \frac{s}{n^2}\right) \\ &\le C\left(s^2 + \frac{s}{n}\right).\end{aligned}$$

Now insert these estimates into (5.1) to obtain

$$(5.11)\qquad \begin{aligned}&\mathbf{E}[|\langle \mathbf{Z}_{t+s}^n, \phi\rangle - \langle \mathbf{Z}_t^n, \phi\rangle|^4] \\ &\quad\le C\left(s^2 + \frac{s}{n} + s\,|d_{nt,ns}^{(1)}| + s(d_{nt,ns}^{(1)})^2 + (d_{nt,ns}^{(1)})^4\right).\end{aligned}$$

It remains to estimate $d_{nt,ns}^{(1)}$. However, we have

$$\begin{aligned}d_{nt,ns}^{(1)} &= \int_0^\infty (T_{n(t+s)}\phi - T_{nt}\phi)(x)\, g_n(x)\, dx \\ &= \exp(-\lambda nt)\left(\int_0^\infty \phi(x + nt + ns)(\exp(-\lambda ns) g_n(x) - g_n(x + ns))\, dx\right. \\ &\qquad\qquad \left. + \int_0^{ns} \phi(x + nt)(e(x) - g_n(x))\, dx\right).\end{aligned}$$

Hence,

$$|d_{nt,ns}^{(1)}| \le \|\phi\|_\infty\left(\int_0^\infty |a_n(x) - a_n(x + ns)| e(x)\, dx + \int_0^{ns} |g_n(x) - e(x)|\, dx\right).$$

Thus, by assumptions (a) and (b),

$$|d_{nt,ns}^{(1)}| \le \|\phi\|_\infty(\|a_n(\cdot) - a_n(\cdot + ns)\|_\infty + ns\|g_n - e\|_\infty) \le Cs.$$

This, together with (5.11), allows us to conclude the bound stated in the lemma. □

PROPOSITION 5.4. *Suppose* $\overline{\mathbf{Z}}_0$ *equals the exponential measure* μ_λ. *Then under the assumptions* (a) *and* (b) *of Theorem* 3.1, *the sequence* $\{\mathbf{Z}^n\}$ *is tight in* $\mathscr{D}([0,\tau],\mathscr{M}_b^+(\mathbf{R}^+))$.

PROOF. We apply Lemma 5.1 and note that to test for weak convergence in $\mathscr{M}_b^+(\mathbf{R}^+)$, it suffices to check convergence only for $\phi \in \mathscr{C}_c^{+,1}(\mathbf{R}^+) \cup \mathbf{1}$. The case $\phi = \mathbf{1}$ corresponds to (1.1). By Lemma 5.3, the upper bound (5.10) on the fourth moment of the increments of the $\langle \mathbf{Z}^n, \phi\rangle$ process holds for $\phi \in \mathscr{C}_c^{+,1}(\mathbf{R}^+)$. For $0 \le v \le s$ and $0 \le t \le \tau$,

$$\mathbf{E}[(\langle \mathbf{Z}^n_{t+s}, \phi\rangle - \langle \mathbf{Z}^n_{t+v}, \phi\rangle)^2 (\langle \mathbf{Z}^n_{t+v}, \phi\rangle - \langle \mathbf{Z}^n_{t}, \phi\rangle)^2] \le C\left(s^2 + \frac{s}{n}\right).$$

To prove the proposition, it then suffices, by Theorem 15.4 of Billingsley (1968), to show that for each positive ε and η there exist a positive δ and an integer n_0 such that

$$\mathbf{P}(w''(\langle \mathbf{Z}^n, \phi\rangle, \delta) \ge \varepsilon) \le \eta, \qquad n \ge n_0, \tag{5.12}$$

where w'' denotes the second modulus of continuity [Billingsley (1968), page 118] over a partition of $[0,\tau]$ of mesh size at most 2δ. However, we can mimic the proof of Theorem 15.6, in particular inequality 15.30, in Billingsley to obtain that

$$\mathbf{P}(w''(\langle \mathbf{Z}^n, \phi\rangle, \delta) \ge \varepsilon) \le \frac{2C\tau}{\varepsilon^4}\left(2\delta + \frac{1}{n}\right).$$

The right member of this inequality goes to zero as $n \to \infty$ and $\delta \to 0$, which proves (5.12).

We can now complete the proof of Theorem 3.1. In Proposition 4.5 we have established that all finite-dimensional distributions converge and in Proposition 5.4, under the additional assumptions, that the sequence of approximating laws is tight so that $\mathbf{Z}^n$ converges weakly in $\mathscr{D}([0,\tau],\mathscr{M}_b^+(\mathbf{R}^+))$, which is (3.2). Usually one establishes the inequality (5.12) from the second moment considerations. This was tried without success.

Acknowledgment. The detailed comments from an Associate Editor substantially improved on several earlier versions of this manuscript.

REFERENCES

ATHREYA, K. and NEY, P. (1972). *Branching Processes*. Springer, New York.
BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
BORDE-BOUSSION, A.-M. (1990). Stochastic demographic models: Age of a population. *Stochastic Process. Appl.* **35** 279–291.

BOSE, A. (1986). A law of large numbers for the scaled age distribution of linear birth-and-death processes. *Canad. J. Statist.* **14** 233–244.

GOROSTIZA, L. G. and LOPEZ-MIMBELA J. A. (1990). The multitype measure branching process. *Adv. in Appl. Probab.* **22** 49–67.

JOFFE, A. and MÉTIVIER, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching process. *Adv. in Appl. Probab.* **18** 20–65.

KENDALL, D. G. (1949). Stochastic processes and population growth. *J. Roy. Statist. Soc. Ser. B* **11** 267–283.

KENDALL, D. G. (1950). Random fluctuations in the age distribution of a population whose development is controlled by simple "birth-and-death" process. *J. Roy. Statist. Soc. Ser. B* **12** 278–285.

OELSCHLÄGER, K. (1990). Limit theorems for age-structured populations. *Ann. Probab.* **18** 290–318.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
CARLETON UNIVERSITY
OTTAWA
CANADA K1S 5B6

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
BOX 480
S-751 06 UPPSALA
SWEDEN