The Annals of Probability
2011, Vol. 39, No. 3, 939–984
DOI: 10.1214/10-AOP574

POISSON REPRESENTATIONS OF BRANCHING MARKOV AND MEASURE-VALUED BRANCHING PROCESSES

BY THOMAS G. KURTZ[1] AND ELIANE R. RODRIGUES[2]

University of Wisconsin, Madison and UNAM

Representations of branching Markov processes and their measure-valued limits in terms of countable systems of particles are constructed for models with spatially varying birth and death rates. Each particle has a location and a "level," but unlike earlier constructions, the levels change with time. In fact, death of a particle occurs only when the level of the particle crosses a specified level r, or for the limiting models, hits infinity. For branching Markov processes, at each time t, conditioned on the state of the process, the levels are independent and uniformly distributed on $[0, r]$. For the limiting measure-valued process, at each time t, the joint distribution of locations and levels is conditionally Poisson distributed with mean measure $K(t) \times \Lambda$, where Λ denotes Lebesgue measure, and K is the desired measure-valued process.

The representation simplifies or gives alternative proofs for a variety of calculations and results including conditioning on extinction or nonextinction, Harris's convergence theorem for supercritical branching processes, and diffusion approximations for processes in random environments.

1. Introduction. Measure-valued processes arise naturally as infinite system limits of empirical measures of finite particle systems. A number of approaches have been developed which preserve distinct particles in the limit and which give a representation of the measure-valued process as a transformation of the limiting infinite particle system. Most of these representations [5–7, 35] have exploited properties of exchangeable sequences, identifying the state of the measure-valued process as a multiple of the de Finetti measure of the sequence, or, as in [35], as a transformation of the de Finetti measure of a sequence that gives both a location and a mass for the distinct particles.

The primary limitation of the representations given for measure-valued branching processes in [7] is that the branching rates must be independent of particle location. This restriction was relaxed in [31] for critical and subcritical branching, but that representation does not seem to provide much useful insight and would be

Received January 2008; revised November 2009.

[1]Supported in part by NSF Grants DMS-05-03983 and DMS-08-05793.

[2]Supported in part by DGAPA-UNAM, Mexico, Grant 968SFA.

MSC2010 subject classifications. 60J25, 60K35, 60J60, 60J80, 60K37.

Key words and phrases. Branching Markov process, Dawson–Watanabe process, superprocess, measure-valued diffusion, particle representation, Feller diffusion, exchangeability, Cox process, conditioning, random environments.

difficult to extend to supercritical processes. A second limitation of the approach in [7] is that, at least without major additional effort, it applies only to models in which the states are finite measures.

In the present paper, we give another construction which, although similar to that of [31], applies immediately to both subcritical and supercritical processes (as well as processes that are subcritical in some locations and supercritical in others). The construction also applies immediately to models with infinite mass. In addition, the new construction seems to be a much more effective tool for analyzing the measure-valued processes obtained.

We introduce the basic ideas of the construction in Section 2, giving results for population models without location or type. As in the earlier work, the justification for the representation is a consequence of a Markov mapping theorem, Theorem A.15. The Feller diffusion approximation for nearly critical branching processes is obtained as a consequence of the construction. Section 3 gives the construction for the general branching Markov process and the Dawson–Watanabe superprocess limit. Section 4 gives a variety of applications and extensions, including conditioning on nonextinction, models with heavy-tailed offspring distributions and processes in random environments. The Appendix contains background material and a number of technical lemmas.

2. Simple examples. In this section, we give particle representations for pure death and continuous time Markov branching processes and illustrate how the infinite system limit can be derived immediately. The main point is to introduce the notion of the *level* of a particle in the simplest possible settings.

2.1. *Pure death processes.* For $r > 0$, let $\xi_1(0), \dots, \xi_{n_0}(0)$ be independent random variables, uniformly distributed on $[0, r]$. For $b > 0$, let

$$\dot{\xi}_i(t) = b\xi_i(t),$$

so $\xi_i(t) = \xi_i(0)e^{bt}$, and define $N(t) = \#\{i : \xi_i(t) < r\}$ and $U(t) = (U_1(t), \dots, U_{N(t)})$, where the $U_j(t)$ are the values of the $\xi_i(t)$ that are less than r. The $U_i(t)$ will be referred to as the levels of the particles. The level of a particle being below r means that the particle is "alive," and as soon as its level reaches r the particle "dies." Note that $N(t)$ is the number of particles "alive" in the system at time t.

Let $f(u, n) = \prod_{i=1}^n g(u_i)$, where $0 \le g \le 1$, g is continuously differentiable and $g(u_i) = 1$ for $u_i > r$. [The "n" in $f(u, n)$ is, of course, redundant, but it will help clarify some of the later calculations.] Then

$$\frac{d}{dt} f(U(t), N(t)) = Af(U(t), N(t)),$$

where

$$Af(u, n) = f(u, n) \sum_{i=1}^n bu_i g'(u_i)/g(u_i).$$

Note that $Af(u,n)$ may also be written as

$$Af(u,n) = \sum_{i=1}^{n} bu_i g'(u_i) \prod_{j\neq i} g(u_j).$$

Hence, even if $g(u_i)=0$ for some i, the expression for $Af(u,n)$ still makes sense.

Let $\alpha_r(n,du)$ be the joint distribution of n i.i.d. uniform $[0,r]$ random variables. Setting $e^{-\lambda_g} = r^{-1}\int_0^r g(z)\,dz$, $\widehat{f}(n) = \int f(u,n)\alpha_r(n,du) = e^{-\lambda_g n}$ and

$$\begin{aligned}\int Af(u,n)\alpha_r(n,du) &= ne^{-\lambda_g(n-1)}br^{-1}\int_0^r zg'(z)\,dz \\ &= bne^{-\lambda_g(n-1)}r^{-1}\int_0^r \bigl(g(r)-g(z)\bigr)\,dz \\ &= bne^{-\lambda_g(n-1)}(1-e^{-\lambda_g}) \\ &= C\widehat{f}(n),\end{aligned}$$

where

$$C\widehat{f}(n) = bn[\widehat{f}(n-1)-\widehat{f}(n)],$$

that is, the generator of a linear death process. Of course, the conditional distribution of $U(t)$ given $N(t)$ is just $\alpha_r(N(t),\cdot)$.

Let $\mathcal{F}_t = \sigma(U(s): s\le t)$ and $\mathcal{F}_t^N = \sigma(N(s): s\le t)$. Then trivially,

$$f(U(0),N(0)) = f(U(t),N(t)) - \int_0^t Af(U(s),N(s))\,ds$$

is an $\{\mathcal{F}_t\}$-martingale, and Lemma A.13 implies

$$\begin{aligned}&E[f(U(t),N(t))|\mathcal{F}_t^N] - \int_0^t E[Af(U(s),N(s))|\mathcal{F}_s^N]\,ds \\ &\quad = \widehat{f}(N(t)) - \int_0^t C\widehat{f}(N(s))\,ds\end{aligned}$$

is a $\{\mathcal{F}_t^N\}$-martingale. Consequently, N is a solution of the martingale problem for C and hence is a linear death process. Of course, this observation follows immediately from the fact that τ_i^r defined by $U_i(0)e^{b\tau_i^r} = r$ is exponentially distributed with parameter b, but the martingale argument illustrates a procedure that works much more generally.

2.2. *A simple branching process.* For f and g as above, $a>0$, $r>0$ and $-\infty < b \le ra$, define the generator

$$\begin{aligned}A_r f(u,n) &= f(u,n)\sum_{i=1}^{n} 2a\int_{u_i}^{r}\bigl(g(v)-1\bigr)\,dv \\ &\quad + f(u,n)\sum_{i=1}^{n}(au_i^2 - bu_i)\frac{g'(u_i)}{g(u_i)}.\end{aligned} \tag{2.1}$$

We refer to u_i as the *level* of the ith particle, and as in the pure death example, a particle "dies" when its level reaches r. The process with generator (2.1) has the following properties. The particle levels satisfy

$$\dot{U}_i(t) = aU_i^2(t) - bU_i(t),$$

and a particle with level z gives birth at rate $2a(r-z)$ to a particle whose initial level is uniformly distributed between z and r. Uniqueness for the martingale problems for A_r follows by first checking uniqueness for the operator D given by the second term alone, and then showing that uniqueness holds up to the first time that the process includes n particles by observing that A_r truncated at n is a bounded perturbation of D. Finally, the first hitting time of n goes to infinity as $n \to \infty$ (cf. Problem 28 in Section 4.11 of [9]).

As before, $\widehat{f}(n) = \int f(u,n)\alpha_r(n,du) = e^{-\lambda_g n}$. To calculate $\int A_r(fu,n) \times \alpha_r(n,du)$, observe that

$$r^{-1}2a\int_0^r g(z)\int_z^r \big(g(v)-1\big)\,dv\,dz = are^{-2\lambda_g} - 2ar^{-1}\int_0^r g(z)(r-z)\,dz$$

and

$$\begin{aligned} r^{-1}\int_0^r (az^2-bz)g'(z)\,dz &= -r^{-1}\int_0^r (2az-b)\big(g(z)-1\big)\,dz \\ &= -2ar^{-1}\int_0^r zg(z)\,dz + ar + b(e^{-\lambda_g}-1). \end{aligned}$$

Then

$$\begin{aligned} \int Af(u,n)\alpha_r(n,du) &= ne^{-\lambda_g(n-1)}\big(are^{-2\lambda_g} - 2are^{-\lambda_g} + ar + b(e^{-\lambda_g}-1)\big) \\ &= C\widehat{f}(n), \end{aligned}$$

where

$$C\widehat{f}(n) = ran\big(\widehat{f}(n+1)-\widehat{f}(n)\big) + (ra-b)n\big(\widehat{f}(n-1)-\widehat{f}(n)\big) \tag{2.2}$$

is the generator of a branching process.

Unlike the linear death example, it is not immediately obvious that $\alpha_r(N(t),\cdot)$ is the conditional distribution of $U(t)$ given $\mathcal{F}_t^N = \sigma(N(s): s \le t)$; however, Theorem A.15 and the fact that a solution of the martingale problem for C starting from $N(0)$ exists gives the existence of a solution $(U(t), N(t))$ of the martingale problem for A_r such that for all $t \ge 0$, $\alpha_r(N(t),\cdot)$ is the conditional distribution of $U(t)$ given $\mathcal{F}_t^N$. To apply Theorem A.15, take $\psi(u,n) = n$ in (A.15) and assume $E[N(0)] < \infty$. Any solution of the martingale problem for C with $E[N(0)] < \infty$ will satisfy $E[N(t)] = E[N(0)]e^{bt}$, for all $t \ge 0$. The moment assumption can be eliminated by conditioning.

We conclude that for any distribution for $N(0)$, there is a solution (U,N) of the martingale problem for A such that N is a solution of the martingale problem for C, that is, N is a linear birth and death process with birth rate ar and death rate $ar - b$. Uniqueness holds for the martingale problem for A, so for any solution of the martingale problem for A satisfying $P\{U(0) \in \Gamma | N(0)\} = \alpha_r(N(0), \Gamma)$, we have that N is a solution of the martingale problem for C.

This representation can be used to do simple calculations. For example, let $U_*(0)$ be the minimum of $U_1(0), \dots, U_{N(0)}$. Then for all t, all levels are above

$$U_*(t) = \frac{U_*(0)e^{-bt}}{1 - (a/b)U_*(0)(1 - e^{-bt})}.$$

Let $\tau = \inf\{t : N(t) = 0\}$. Then if τ is finite, $U_*(\tau) = r$. In particular, if $N(0) = n$, then

$$\begin{aligned} P\{\tau > t\} = P\{U_*(t) < r\} &= P\left\{U_*(0) < \frac{r}{e^{-bt} - (ra/b)(e^{-bt} - 1)}\right\} \\ &= 1 - \left(e^{-bt} - \frac{ra}{b}(e^{-bt} - 1)\right)^{-n}. \end{aligned}$$

Note that the assumption that $b \le ra$ ensures that $e^{-bt} - \frac{ra}{b}(e^{-bt} - 1) \ge 1$.

In the branching process, the average lifetime of an individual is $(ar - b)^{-1}$ which will be small if r is large. Consequently, it is important to note that the levels do not represent single individuals in the branching process but whole lines of descent. For example, at least in the critical or subcritical case, the individual with level $U_*(0)$ at time zero is the individual whose line of descent lasts longer than that of any other individual alive at time zero.

2.2.1. *Conditioning on nonextinction.* If $b \le 0$, then conditioning on nonextinction, that is, conditioning on $\tau > t$ and letting $t \to \infty$, is equivalent to conditioning on $U_*(0) = 0$. Conditioned on $U_*(0) = 0$, $U_1(0), \dots, U_{N(0)}(0)$ include $N(0) - 1$ independent, uniform $[0, r]$ random variables and one that equals zero. If one of the initial levels is zero, then the solution of the martingale problems for A_r gives a solution for

$$\begin{aligned} A_r^c f(u, n) = {} & f(u, n) \sum_{i=1}^{n-1} 2a \int_{u_i}^{r} \big(g(v) - 1\big)\, dv + f(u, n) \sum_{i=1}^{n-1} (au_i^2 - bu_i) \frac{g'(u_i)}{g(u_i)} \\ & + f(u, n) 2a \int_0^r \big(g(v) - 1\big)\, dv, \end{aligned}$$

and taking $\alpha_r^c(n, du)$ to be the distribution of n independent random variables, one of which is zero and the others uniform $[0, r]$, we see that N is a solution of the martingale problem for

$$C^c \widehat{f}(n) = ra(n + 1)\big(\widehat{f}(n + 1) - \widehat{f}(n)\big) + (ra - b)(n - 1)\big(\widehat{f}(n - 1) - \widehat{f}(n)\big).$$

2.2.2. *Conditioning on extinction.* If $0 < b < ra$, then conditioning on extinction is equivalent to conditioning on $U_*(0) > \frac{b}{a}$. Conditioned on $U_*(0) > \frac{b}{a}$, $U_1(0), \dots, U_{N(0)}(0)$ are independent uniform $[\frac{b}{a}, r]$. Defining $V_i(t) = U_i(t) - \frac{b}{a}$, V is a solution of the martingale problem for

$$A_r f(v,n) = f(v,n) \sum_{i=1}^{n} 2a \int_{v_i}^{r-b/a} \big(g(z) - 1\big)\, dz + f(v,n) \sum_{i=1}^{n} (av_i^2 + bv_i) \frac{g'(v_i)}{g(v_i)},$$

so N is a solution of the martingale problem for

$$C\widehat{f}(n) = (ra - b)n\big(\widehat{f}(n+1) - \widehat{f}(n)\big) + ran\big(\widehat{f}(n-1) - \widehat{f}(n)\big),$$

which is the generator of a subcritical branching process.

2.2.3. *Convergence as $t \to \infty$.* Again, in the supercritical case, $0 < b < ra$, if $0 < U_*(0) < \frac{b}{a}$, then $N(t) \to \infty$. Observe that

$$V_*(\infty) = \lim_{t\to\infty} e^{bt} U_*(t) = \frac{U_*(0)}{1 - (a/b)U_*(0)}$$

exists, and a similar limit will hold for any level whose initial value is below $\frac{b}{a}$. Setting $\xi(t) = \sum \delta_{e^{bt} U_i(t)}$, the counting measure $\xi(t)$ converges almost surely in the sense that

$$\lim_{t\to\infty} \int_0^\infty f(u)\xi(t, du) = \lim_{t\to\infty} \sum_i f(e^{bt} U_i(t)) = \int_0^\infty f(u)\xi(\infty, du) \qquad \text{a.s.}$$

for each bounded, continuous, nonnegative f with compact support in $[0, \infty)$. Let $\{\mathcal{F}_t^N\}$ be the filtration generated by N. Then as in (A.6),

$$E\big[e^{-\int_0^{re^{bt}} f(u)\xi(t,du)} | \mathcal{F}_t^N\big] = e^{-F_f^t r^{-1} e^{-bt} N(t)}, \tag{2.3}$$

where

$$F_f^t = -re^{bt} \log\bigg(1 - r^{-1} e^{-bt} \int_0^{re^{bt}} \big(1 - e^{-f(u)}\big)\, du\bigg) \to \int_0^\infty \big(1 - e^{-f(u)}\big)\, du.$$

The left-hand side of (2.3) converges almost surely by Lemma A.14. Consequently,

$$W \equiv \lim_{t\to\infty} e^{-bt} N(t)$$

exists almost surely. Note that $W > 0$ if and only if $\lim_{t\to\infty} N(t) = \infty$.

Conditioned on W, $\xi(\infty)$ is a Poisson point process with intensity $r^{-1}W$, and $V_*(\infty)$ is exponentially distributed with parameter $r^{-1}W$, with the understanding that $V_*(\infty) = \infty$ if $W = 0$. It follows that for $\lambda > 0$,

$$\begin{aligned}
&P\{r^{-1}V_*(\infty) > \lambda | V_*(\infty) < \infty\} \\
&\quad = E[e^{-\lambda W} | W > 0] \\
&\quad = P\bigg\{r^{-1} U_*(0)\bigg(1 - \frac{a}{b} U_*(0)\bigg)^{-1} > \lambda \bigg| U_*(0) < \frac{b}{a}\bigg\}.
\end{aligned}$$

If $N(0)=1$, then $U_*(0)$ is uniformly distributed on $[0,r]$, and hence $P\{W>0\}=\frac{b}{ra}$ and

$$E[e^{-\lambda W}|W>0]=\frac{1}{1+(ra/b)\lambda},$$

that is, W is exponentially distributed with parameter $\frac{b}{ra}$. Of course, we have simply rederived a classical result of Harris [17].

2.3. *Feller diffusion approximation.* As $r\to\infty$, $A_r f$ in (2.1) converges for every continuously differentiable g such that $0\le g\le 1$ and $g(z)=1$ for $z\ge r_g$, that is, for $f(u)=\prod_i g(u_i)$, in the limit

$$Af(u)=f(u)\sum_i 2a\int_{u_i}^{r_g}\big(g(v)-1\big)\,dv+f(u)\sum_i(au_i^2-bu_i)\frac{g'(u_i)}{g(u_i)}. \tag{2.4}$$

If $n/r\to y$ as $r\to\infty$, then $\alpha_r(n,\cdot)$ converges to $\alpha(y,\cdot)$, where $\alpha(y,\cdot)$ is the distribution of a Poisson process ξ_y on $[0,\infty)$ with intensity y, in the sense that

$$\int_{[0,r]^n} f(u)\alpha_r(n,du)\to E\big[e^{\int_0^\infty \log g(z)\,d\xi_y(z)}\big]=e^{-y\int_0^\infty(1-g(z))\,dz}.$$

Note that

$$\widehat{f}(y)=\alpha f(y)=\int f(u)\alpha(y,du)=e^{-y\int_0^\infty(1-g(z))\,dz}=e^{-y\beta_g},$$

and using Lemma A.3

$$\begin{aligned}
\alpha Af(y)&=e^{-y\beta_g}\bigg(2ay\int_0^\infty g(z)\int_z^\infty\big(g(v)-1\big)\,dv\,dz\\
&\qquad\qquad+y\int_0^\infty(az^2-bz)g'(z)\,dz\bigg)\\
&=e^{-y\beta_g}\bigg(2ay\int_0^\infty g(z)\int_z^\infty\big(g(v)-1\big)\,dv\,dz\\
&\qquad\qquad-y\int_0^\infty(2az-b)\big(g(z)-1\big)\,dz\bigg)\\
&=e^{-y\beta_g}\bigg(2ay\int_0^\infty g(z)\int_z^\infty\big(g(v)-1\big)\,dv\,dz\\
&\qquad\qquad-2ay\int_0^\infty\int_z^\infty\big(g(v)-1\big)\,dv\,dz\\
&\qquad\qquad+by\int_0^\infty\big(g(z)-1\big)\,dz\bigg)\\
&=e^{-y\beta_g}(ay\beta_g^2-by\beta_g)\\
&=C\widehat{f}(y),
\end{aligned}$$

where

$$C\widehat{f}(y) = ay\widehat{f}''(y) + by\widehat{f}'(y).$$

Again, we can apply Theorem A.15 taking

$$\gamma(u) = \limsup_{z\to\infty} \frac{1}{z} \sum_i \mathbf{1}_{[0,z]}(u_i)$$

and $\psi(u) = \sum_i e^{-u_i}$, so

$$|Af(u)| \le [2ar_g + \|g'\|(ar_g^2 + |b|r_g)]e^{r_g}\psi(u)$$

and $\int \psi(u)\alpha(y, du) = y$. If $\widetilde{Y}$ is a solution of the martingale problem for C with $E[\widetilde{Y}(0)] < \infty$, then $E[\widetilde{Y}(t)] = e^{bt}E[\widetilde{Y}(0)]$ and the conditions of Theorem A.15 are satisfied. Consequently, there is a solution U of the martingale problem for A such that

$$Y(t) = \limsup_{z\to\infty} \frac{1}{z} \sum_i \mathbf{1}_{[0,z]}(U_i(t)) \tag{2.5}$$

is a solution of the martingale problem for C with the same distribution as $\widetilde{Y}$. [Note that, with probability one, the lim sup in (2.5) is actually a limit.]

If U is a solution of the martingale problem for A, then $U^r(t) = \{U_i(t) : U_i(t) < r\}$ defines a solution of the martingale problem for A_r. Uniqueness for A_r follows by the argument outlined in Section A.6, and uniqueness for A_r implies uniqueness for A. Since uniqueness holds for the martingale problem for A, by Theorem A.15(c), uniqueness holds for C also. In general, if U is a solution of the martingale problem for A and $\sum_i \delta_{U_i(0)}$ is a Poisson random measure with mean measure $y\Lambda$, where Λ denotes Lebesgue measure, then (2.5) is a solution of the martingale problem for C.

2.4. *The genealogy and the number of ancestors.* For each $T > 0$, there is a solution of

$$\dot{u}_T(t) = au_T(t)^2 - bu_T(t) \tag{2.6}$$

satisfying $u_T(t) < \infty$ for $t < T$ and $\lim_{t\to T-} u_T(t) = \infty$. Every particle alive at time T is a descendent of some particle $U_i(t)$ alive at time $t < T$ satisfying $U_i(t) < u_T(t)$. Note that the converse is also true. If $U_i(t) < u_T(t)$, then $U_i(t)$ has descendants alive at time T. In fact, a positive fraction of the particles alive at time T will be descendants of $U_i(t)$.

If $Y(T) > 0$, then there are infinitely many particles alive at time T, but since

$$\xi(t, [0, u_T(t))) < \infty,$$

they are all descendants of finitely many ancestors alive at time t. Note that $t \to \xi(t, [0, u_T(t)))$ is nondecreasing and increases by jumps of $+1$. It is not possible to recover the full genealogy just from the levels since a new individual appearing at time t with level v could be the offspring of any existing individual with level $U_i(t) < v$. In Section 3, particles will be assigned a location (or type), and if these locations evolve in such a way that two particles have the same location only if one is the offspring of the other and then only at the instant of birth, it will be possible to reconstruct the full genealogy from the levels and locations.

2.5. *Branching processes in random environments.* Assume that a and b are functions of another stochastic process ξ, say an irreducible, finite Markov chain with generator Q. Then, for functions of the form $f(l,u,n) = f_0(l)f_1(u) = f_0(l)\prod_{i=1}^n g(u_i)$, consider a scaled generator

$$A_r f(l,u,n) = r f_1(u) Q f_0(l) + f(l,u,n)\sum_{i=1}^n 2a(l)\int_{u_i}^r \big(g(v)-1\big)\,dv$$
$$+ f(l,u,n)\sum_{i=1}^n \big(a(l)u_i^2 - \sqrt{r}b(l)u_i\big)\frac{g'(u_i)}{g(u_i)},$$

which, as in (2.2), corresponds to a process with generator

$$C_r\widehat{f}(l,n) = rQ\widehat{f}(l,n) + a(l)rn\big(\widehat{f}(l,n+1) - \widehat{f}(n)\big)$$
$$+ \big(ra(l) - \sqrt{r}b(l)\big)n\big(\widehat{f}(l,n-1) - \widehat{f}(l,n)\big),$$

where $\widehat{f}(l,n) = f_0(l)e^{-\lambda_g n}$. The process corresponding to C_r is a branching process in a random environment determined by ξ. Writing the process corresponding to A_r as

$$\big(\xi(rt), U_1(t), \ldots, U_{N_r(t)}\big)$$

the process corresponding to C_r is $(\xi(rt), N_r(t))$.

Note that in this example, the levels satisfy

$$\dot{U}_i(t) = a(\xi(rt))U_i^2(t) - \sqrt{r}b(\xi(rt))U_i(t).$$

Let π be the stationary distribution for Q, and assume that $\sum_l \pi(l)b(l) = 0$. Then, by Theorem 2.1 or [3], for example,

$$Z^{(r)}(t) = \sqrt{r}\int_0^t b(\xi(rs))\,ds$$

converges to a Brownian motion Z with variance parameter

$$\sum_k\sum_l \pi(k)q_{kl}\big(h_0(l) - h_0(k)\big)^2 = -2\sum_l \pi(l)h_0(l)b(l) \equiv 2\overline{c},$$

where $h_0(l)$ is a solution of $Qh_0(l) = b(l)$. In the limit, by Theorem 5.10 of [33] (applying a truncation argument to extend the boundedness assumption), the levels will satisfy

$$dU_i(t) = \big(\overline{a}U_i(t)^2 + \overline{c}U_i(t)\big)\,dt + \sqrt{2\overline{c}}U_i(t)\,dW(t), \tag{2.7}$$

where $\overline{a} = \sum \pi(l)a(l)$.

Applying ideas from [28], we can obtain convergence for the full system by considering the asymptotic behavior of the generator. Setting

$$h_1(l,u,n) = h_0(l) f_1(u,n) \sum_{i=1}^{n} u_i \frac{g'(u_i)}{g(u_i)},$$

we have

$$\begin{aligned}
A_r&\left(f_1 + \frac{1}{\sqrt{r}} h_1\right)(l,u,n)\\
&= f_1(u,n)\sum_{i=1}^{n} 2a(l)\int_{u_i}^{r}\big(g(v)-1\big)\,dv + f_1(u,n)\sum_{i=1}^{n} a(l)u_i^2 \frac{g'(u_i)}{g(u_i)}\\
&\quad + \frac{1}{\sqrt{r}} h_0(l) f_1(u,n)\left(\sum_{i=1}^{n} u_i \frac{g'(u_i)}{g(u_i)}\right)\sum_{i=1}^{n} 2a(l)\int_{u_i}^{r}\big(g(v)-1\big)\,dv\\
&\quad + \frac{1}{\sqrt{r}}\sum_{i=1}^{n} h_0(l) f_1(u,n)\int_{u_i}^{r} v g'(v)\,dv\\
&\quad + \frac{1}{\sqrt{r}} h_0(l) f_1(u,n)\sum_{j=1}^{n}\big(a(l)u_j^2 - \sqrt{r}\,b(l)u_j\big)\\
&\qquad\qquad \times\left(\sum_{i\neq j} u_i \frac{g'(u_i)g'(u_j)}{g(u_i)g(u_j)} + \frac{g'(u_j) + u_j g''(u_j)}{g(u_j)}\right),
\end{aligned}$$

and passing to the limit as $r \to \infty$, $A_r(f_1 + \frac{1}{\sqrt{r}} h_1)$ converges to

$$\begin{aligned}
\widetilde{A}&f_1(u,l)\\
&= f_1(u)\sum_i 2a(l)\int_{u_i}^{\infty}\big(g(v)-1\big)\,dv + f_1(u)\sum_i a(l)u_i^2\frac{g'(u_i)}{g(u_i)}\\
&\quad - h_0(l)b(l) f_1(u)\sum_j\left(\sum_{i\neq j} u_j u_i \frac{g'(u_i)g'(u_j)}{g(u_i)g(u_j)} + \frac{u_j g'(u_j) + u_j^2 g''(u_j)}{g(u_j)}\right).
\end{aligned}$$

Finally, we can find an additional perturbation h_2 so that $A_r(f_1 + \frac{1}{\sqrt{r}} h_1 + \frac{1}{r} h_2)$ converges to

$$\begin{aligned}
Af_1(u) &= f_1(u)\sum_i 2\overline{a}\int_{u_i}^{\infty}\big(g(v)-1\big)\,dv + f_1(u)\sum_i \overline{a}u_i^2\frac{g'(u_i)}{g(u_i)}\\
&\qquad + \overline{c} f_1(u)\sum_j\left(\sum_{i\neq j} u_j u_i \frac{g'(u_i)g'(u_j)}{g(u_i)g(u_j)} + \frac{u_j g'(u_j) + u_j^2 g''(u_j)}{g(u_j)}\right).
\end{aligned}$$

This convergence assures convergence of the finite models to an infinite particle model. The particle birth process is the same as in Section 2.3, but the levels satisfy (2.7) where the Brownian motion W is the same for all levels.

Let α and β_g be as in Section 2.3, and note that

$$\begin{aligned}\beta_g &= \int_0^\infty \bigl(1-g(z)\bigr)\,dz = \int_0^\infty zg'(z)\,dz \\ &= -\frac{1}{2}\int_0^\infty z^2 g''(z)\,dz.\end{aligned}$$

We have from Lemma A.3

$$\begin{aligned}\alpha Af(y) = e^{-y\beta_g}\biggl(&2\overline{a}y\int_0^\infty g(z)\int_z^\infty \bigl(g(v)-1\bigr)\,dv\,dz \\ &+ y\int_0^\infty (\overline{a}z^2+\overline{c}z)g'(z)\,dz \\ &+ \overline{c}y^2\biggl(\int_0^\infty zg'(z)\,dz\biggr)^2 + \overline{c}y\int_0^\infty z^2 g''(z)\,dz\biggr) \\ = e^{-y\beta_g}&\bigl((\overline{a}y+\overline{c}y^2)\beta_g^2 - \overline{c}y\beta_g\bigr) \\ = C\widehat{f}&(y),\end{aligned}$$

where

$$C\widehat{f}(y) = (\overline{a}y+\overline{c}y^2)\widehat{f}''(y) + \overline{c}y\widehat{f}'(y),$$

which identifies the diffusion limit for $r^{-1}N_r$.

Theorem A.15 can be extended to cover models with non-Markovian environments, that is, the process ξ is specified directly rather than through a generator. The diffusion limit is then obtained by verifying convergence of the level processes and applying Theorem A.12.

For early work on diffusion approximations for branching processes in random environments, see [18, 26, 29] and also [9], Section 9.3.

3. Representations of measure-valued branching processes.

3.1. *Branching Markov processes.* We now consider particles with both a level u_i and a location x_i in a complete, separable metric space E. Since the indexing of the particles is not important, we identify a state (x,u,n) of our process with the counting measure $\mu_{(x,u)} = \sum_{i=1}^n \delta_{(x_i,u_i)}$. Let

$$f(x,u,n) = \prod_{i=1}^n g(x_i,u_i) = e^{\int \log g\,d\mu_{(x,u)}},$$

where $g:E \times [0,\infty) \to (0,1]$. We assume that as a function of x, g is in the domain $\mathcal{D}(B)$ of the generator of a Markov process in E, g is continuously differentiable in u and $g(x,u) = 1$ for $u \geq r$. We set

$$
\begin{aligned}
A_r f(x,u,n) &= f(x,u,n)\sum_{i=1}^{n}\frac{Bg(x_i,u_i)}{g(x_i,u_i)} \\
&\quad + f(x,u,n)\sum_{i=1}^{n} 2a(x_i)\int_{u_i}^{r}\big(g(x_i,v)-1\big)\,dv \qquad (3.1)\\
&\quad + f(x,u,n)\sum_{i=1}^{n}\big(a(x_i)u_i^2 - b(x_i)u_i\big)\frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}.
\end{aligned}
$$

Each particle has a location $X_i(t)$ in E and a level $U_i(t)$ in $[0,r]$. The locations evolve independently as Markov processes with generator B; the levels satisfy

$$\dot{U}_i(t) = a(X_i(t))U_i^2(t) - b(X_i(t))U_i(t); \qquad (3.2)$$

particles give birth at rates $2a(X_i(t))(r - U_i(t))$; the initial location of a new particle is the location of the parent at the time of birth; and the initial level is uniformly distributed on $[U_i(t), r]$. Particles that reach level r die. Setting $e^{-\lambda_g(x_i)} = \widehat{g}(x_i) = r^{-1}\int_0^r g(x_i,z)\,dz$ and $\widehat{f}(x,n) = \prod_{i=1}^n \widehat{g}(x_i) = e^{-\sum_{i=1}^n \lambda_g(x_i)}$, calculating as in Section 2.2, we have

$$
\begin{aligned}
C_r\widehat{f}(x,n) &= \sum_{i=1}^{n} B_{x_i}\widehat{f}(x,n) + \sum_{i=1}^{n} ra(x_i)\big(\widehat{f}\big(b(x|x_i), n+1\big) - \widehat{f}(x,n)\big) \\
&\quad + \sum_{i=1}^{n}\big(ra(x_i) - b(x_i)\big)\big(\widehat{f}\big(d(x|x_i), n-1\big) - \widehat{f}(x,n)\big),
\end{aligned}
$$

where B_{x_i} is the generator B applied to $\widehat{f}(x,n)$ as a function of x_i, $b(x|x_i)$ is the collection of $n+1$ particles in E obtained from x by adding a copy of the ith particle x_i, and $d(x|x_i)$ is the collection of $n-1$ particles obtained from x by deleting the ith particle, that is, if μ_x denotes $\sum_{i=1}^n \delta_{x_i}$, then for $z \in E$,

$$\mu_{b(x|z)} = \delta_z + \sum_{i=1}^{n}\delta_{x_i}, \qquad \mu_{d(x|x_j)} = \sum_{i=1}^{n}\delta_{x_i} - \delta_{x_j}.$$

If $ra(z) - b(z) \geq 0$ for all $z \in E$, then C is the generator of a branching Markov process with particle motion determined by B, the birth rate for a particle at $z \in E$ given by $ra(z)$ and the death rate given by $ra(z) - b(z)$.

With Theorem A.15 in mind, we make the following assumptions on B, a, b and r. $C(E)$ is the space of continuous functions on E, $\overline{C}(E)$ the space of bounded continuous functions on E and $M(E)$ the space of Borel measurable functions on E.

CONDITION 3.1.

(i) $B \subset \overline{C}(E) \times C(E)$, $\mathcal{D}(B)$ is closed under multiplication and is separating.
(ii) $f \in \mathcal{D}(A)$ satisfies $f(x,u,n) = \prod_{i=1}^n g(x_i,u_i)$, where $g(z,v) = \prod_{l=1}^m (1 - g_1^l(z) g_2^l(v))$ for $g_1^l \in \mathcal{D}(B)$, g_2^l differentiable with support in $[0,r]$ and $0 \le g_1^l(z) g_2^l(v) \le \rho_g < 1$ for all l, z and v.
(iii) There exists $\psi_B \in C(E)$, $\psi_B \ge 0$ and constants $c_g \ge 0$, for each g in (ii), such that

$$\sup_u |Bg(x,u)| \le c_g \psi_B(x), \qquad x \in E.$$

(iv) Defining $B_0 = \{(g, (\psi_B \vee 1)^{-1} Bg) : g \in \mathcal{D}(B)\}$, B_0 is graph separable (see Section A.5).
(v) $a, b \in M(E)$, $a \ge 0$, $r > 0$ and $ra - b \ge 0$.

We have the following generalization of the results of Section 2.2.

THEOREM 3.2. *Assume Condition 3.1. For $x \in E^n$ and $u \in [0,\infty)^n$, let*

$$\psi(x,u) = 1 + \sum_{i=1}^n \bigl(\psi_B(x_i) + a(x_i) + |b(x_i)|\bigr) e^{-u_i}$$

and

$$\widetilde{\psi}(x) = 1 + \sum_{i=1}^n \bigl(\psi_B(x_i) + a(x_i) + |b(x_i)|\bigr)(1 - e^{-r}).$$

If X is a solution of the martingale problem for C satisfying

$$E\left[\int_0^t \widetilde{\psi}(X(s))\,ds\right] < \infty \qquad \textit{for all } t \ge 0, \tag{3.3}$$

then there is a solution $(\widetilde{X}, \widetilde{U})$ of the martingale problem for A_r such that X and $\widetilde{X}$ have the same distribution.

REMARK 3.3. For many models, ψ_B, a and b will be uniformly bounded, and the moment conditions (3.3) will hold as long as $E[X(0)] < \infty$.

PROOF OF THEOREM 3.2. Note that

$$|A_r f(x,u,n)| \le \bigl(2r + (1 + r^2 + r) d_g\bigr) e^r \psi(x,u),$$

where d_g depends on the g_1^l, g_2^l, $\partial_v g_2^l$ and $B \prod g_1^{l_k}$ for all choices of $\{l_1, \dots, l_j\} \subset \{1, \dots, m\}$. The result then follows by application of Theorem A.15. $\square$

Theorem 3.2 applies to finite branching Markov processes. Similar results also hold for locally finite processes.

THEOREM 3.4. *In addition to Condition 3.1, assume that*

$$\int_0^\infty |g(x,u)-1|\,du + \sup_u (u+u^2)\,\partial_u g(x,u) \le c_g \psi_B(x).$$

For $x \in E^\infty$ and $u \in [0,\infty)^\infty$, let

$$\psi(x,u) = 1 + \sum_{i=1}^\infty \psi_B(x_i)\big(1 + a(x_i) + |b(x_i)|\big)e^{-u_i}$$

and

$$\widetilde{\psi}(x) = 1 + \sum_{i=1}^\infty \psi_B(x_i)\big(1 + a(x_i) + |b(x_i)|\big)(1 - e^{-r}).$$

If X is a solution of the martingale problem for C satisfying

$$E\left[\int_0^t \widetilde{\psi}(X(s))\,ds\right] < \infty \qquad \textit{for all } t \ge 0, \tag{3.4}$$

then there is a solution $(\widetilde{X}, \widetilde{U})$ of the martingale problem for A_r such that X and $\widetilde{X}$ have the same distribution.

PROOF. Note that

$$A_r f(x,u,n) \le 2c_g e^r \psi(x,u).$$

The result then follows by application of Theorem A.15. □

EXAMPLE 3.5. Suppose that a and b are bounded, and B is a diffusion operator with bounded drift and diffusion coefficients. Then we can take $\mathcal{D}(B)$ to be the collection of nonnegative C^2-functions with compact support and $\psi_B(z) = \frac{1}{(1+|z|^2)^\beta}$, for $\beta > 0$. If $E[\sum_i \psi_B(X_i(0))] < \infty$, then there exists a solution of the martingale problems for C satisfying $\sup_{s \le t} E[\sum_i \psi_B(X_i(s))] < \infty$ and hence $E[\int_0^t \widetilde{\psi}(X(s))\,ds] < \infty$.

3.2. *Basic limit theorem.* As in Section 2.3, if $r \to \infty$, Af given by (3.1) becomes

$$\begin{aligned} Af(x,u) = {} & f(x,u) \sum_i \frac{Bg(x_i,u_i)}{g(x_i,u_i)} \\ & + f(x,u) \sum_i 2a(x_i) \int_{u_i}^{r_g} \big(g(x_i,v) - 1\big)\,dv \\ & + f(x,u) \sum_i \big(a(x_i)u_i^2 - b(x_i)u_i\big) \frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}, \end{aligned} \tag{3.5}$$

where $g : E \times [0, \infty) \to (0, 1]$ has the property that there exists r_g such that $g(z, v) = 1$ for $v > r_g$, and $f(x, u) = \prod_i g(x_i, u_i)$. We can identify the state space of the corresponding process with a subset of $(E \times [0, \infty))^\infty$ or, since order is not important, with a subset of $\mathcal{N}(E \times [0, \infty))$, the counting measures on $E \times [0, \infty)$. Define

$$\mathcal{N}_f\big(E \times [0, \infty)\big) = \big\{\mu \in \mathcal{N}\big(E \times [0, \infty)\big) : \mu\big(E \times [0, u]\big) < \infty \ \forall 0 < u < \infty\big\},$$

where the topology for $\mathcal{N}_f(E \times \infty)$ is given by the requirement that $\mu_n \to \mu$ if and only if $\int f \, d\mu_n \to \int f \, d\mu$ for all $f \in \overline{C}(E \times [0, \infty)$ for which there exists $u_f > 0$ such that $f(x, u) = 0$ for $u \geq u_f$. (See Section A.3 for a discussion of the appropriate topology to use in the infinite measure setting.)

As $r \to \infty$, the particle process converges to a process in which particle locations evolve as independent Markov processes with generator B, levels satisfy (3.2), a particle with level $U_i(t)$ gives birth to new particles at its location $X_i(t)$ and level in the interval $[U_i(t) + c, U_i(t) + d]$ at rate $2a(X_i(t))(d - c)$. A particle dies when its level hits ∞. The level of a particle born at time t_0 (or in the initial population, if $t_0 = 0$) with initial level $U_i(t_0)$ satisfies

$$U_i(t) = \frac{U_i(t_0) e^{-\int_{t_0}^t b(X_i(s))\, ds}}{1 - U_i(t_0) \int_{t_0}^t e^{-\int_{t_0}^v b(X_i(s))\, ds} a(X_i(v))\, dv},$$

until it hits infinity. If $b \leq 0$ and a is bounded away from zero, then U_i will hit infinity in finite time.

If we extend the path X_i back along its ancestral path to time zero, we would have

$$U_i(t) \geq \frac{u_0 e^{-\int_0^t b(X_i(s))\, ds}}{1 - u_0 \int_0^t e^{-\int_0^v b(X_i(s))\, ds} a(X_i(v))\, dv},$$

where u_0 is the level of the particle's ancestor at time zero, and $X_i(s)$ is the position of the ancestor at time $s \leq t$. Since we are assuming that the initial position of an offspring is that of the parent, X_i is a solution of the martingale problem for B.

PROPOSITION 3.6. *Let (X, U) be a solution of the martingale problem for A given by* (3.5). *Let (X^r, U^r) consist of the subset of particles for which $U_i < r$, that is*,

$$\sum \delta_{(X_j^r(t), U_j^r(t))} = \sum \delta_{(X_i(t), U_i(t))} \mathbf{1}_{[0,r)}(U_i(t)).$$

If $a(z)r - b(z) \geq 0$ for all $z \in E$, then (X^r, U^r) is a solution of the martingale problem for A_r given by (3.1).

REMARK 3.7. The condition $a(z)r - b(z) \geq 0$ for all $z \in E$ ensures that any particle that is above level r at time t will stay above level r at all future times.

PROOF OF PROPOSITION 3.6. The proposition follows by the observation that A_r can be obtained from A by restricting the domain to $f(x,u) = \prod_i g(x_i, u_i)$ for which $g(z,v) = 1$ for $v \geq r$. $\square$

3.3. *The genealogy.* Assume for the moment that a and b do not depend on x. If the location process has the property that at the time of a birth only the offspring has the same location as the parent (e.g., if the location process is Brownian motion), then the full genealogy can be recovered from knowledge of the levels and locations. The collection of ancestors at time $t < T$ of the particles alive at time T is $\{(X_i(t), U_i(t)) : U_i(t) < u_T(t)\}$, where $u_T(t)$ is given by (2.6). The number of particles in this collection is nondecreasing in t and increases only by jumps of $+1$. The parent of the new particle is identifiable by the fact that only the parent and the offspring will be at the same location.

If a and b depend on x, the full genealogy is still determined by the locations and levels of the particles, but recovering the genealogy is more complicated since it may not be possible to tell whether or not a particle $(X_i(t), U_i(t))$ has descendants alive at time $T > t$ just from information available at time t. However, some easy observations can be made. For example, if $\inf_x a(x) > 0$ and $\sup_x b(x) < \infty$, then for $t < T$, all particles alive at time T are descendants of finitely many particles alive at time t.

3.4. *The measure-valued limit.* The generator for a Dawson–Watanabe superprocess is typically of the form

$$C\widehat{f}(\mu) = \exp\{-\langle h, \mu\rangle\} \int_E \bigl(-Bh(y) - F(h(y), y)\bigr)\mu(dy),$$

for $\widehat{f}(\mu) = \exp\{-\langle h, \mu\rangle\}$, where $\langle h, \mu\rangle = \int_E h(y)\mu(dy)$ and h is an appropriate function in $\mathcal{D}(B)$ (see, e.g., Theorem 9.4.3 of [9]). For superprocesses arising from branching models with offspring distributions having finite variances, F should be of the form $F(h(y), y) = -a(y)h(y)^2 + b(y)h(y)$.

For $\mu \in \mathcal{M}_f(E)$, let $\alpha(\mu, dx \times du)$ be the distribution of a Poisson random measure on $E \times [0, \infty)$ with mean measure $\mu \times \Lambda$. Then setting $h(y) = \int_0^\infty (1 - g(y, v))\,dv$,

$$\begin{aligned}
\widehat{f}(\mu) = \alpha f(\mu) &= \int f(x,u)\alpha(\mu, dx \times du) \\
&= \exp\biggl\{\int_E \int_0^\infty \bigl(g(y,v) - 1\bigr)\,dv\,\mu(dy)\biggr\} \\
&= \exp\{-\langle h, \mu\rangle\}.
\end{aligned}$$

Using Lemma A.3, we have

$$\begin{aligned}
\alpha A f(\mu) &= \int_E \int_0^\infty Bg(y,v)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty 2a(y)g(y,z)\int_z^\infty \bigl(g(y,v)-1\bigr)\,dv\,dz\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty \bigl(a(y)v^2-b(y)v\bigr)\,\partial_v g(y,v)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&= \int_E \int_0^\infty Bg(y,v)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty 2a(y)g(y,z)\int_z^\infty \bigl(g(y,v)-1\bigr)\,dv\,dz\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad - \int_E \int_0^\infty \bigl(2a(y)v-b(y)\bigr)\bigl(g(y,v)-1\bigr)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&= \int_E \int_0^\infty Bg(y,v)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty 2a(y)g(y,z)\int_z^\infty \bigl(g(y,v)-1\bigr)\,dv\,dz\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad - \int_E \int_0^\infty 2a(y)\int_z^\infty \bigl(g(y,v)-1\bigr)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty b(y)\bigl(g(y,v)-1\bigr)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&= \int_E \int_0^\infty Bg(y,v)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E a(y)\biggl(\int_0^\infty \bigl(g(y,v)-1\bigr)\,dv\biggr)^2 \mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&\quad + \int_E \int_0^\infty b(y)\bigl(g(y,v)-1\bigr)\,dv\,\mu(dy)\exp\{-\langle h,\mu\rangle\} \\
&= \int_E \bigl(-Bh(y)+a(y)h(y)^2-b(y)h(y)\bigr)\mu(dy)\exp\{-\langle h,\mu\rangle\} = C\widehat{f}(\mu).
\end{aligned}$$

But C is the generator for a superprocess, so for each $\mu \in \mathcal{M}_f(E)$, there exists a solution Z of the martingale problem for C with $Z(0)=\mu$ and hence a solution (X,U) of the martingale problem for A with initial distribution $\alpha(\mu,\cdot)$.

The mapping $\gamma : \mathcal{N}_f(E\times[0,\infty)) \to \mathcal{M}_f(E)$ used in the application of Theorem A.15 is given by

$$\gamma\Bigl(\sum \delta_{(x_i,u_i)}\Bigr) = \begin{cases} \displaystyle\lim_{r\to\infty} \frac{1}{r}\sum_{u_i\le r} \delta_{x_i}, & \text{if the measures converge in the weak topology,} \\ \mu_0, & \text{otherwise,} \end{cases}$$

where μ_0 is a fixed element of $\mathcal{M}_f(E)$. The solution $\sum \delta_{(X_i(t),U_i(t))}$ of the martingale problem for A is a conditionally Poisson random measure (see Section A.2) with Cox measure $Z(t)$. Consequently, the particles determine Z by

$$Z(t) = \lim_{r\to\infty} \frac{1}{r} \sum_{U_i(t)\le r} \delta_{X_i(t)}. \tag{3.6}$$

Since by Proposition 3.6 and Theorem 3.2, $Z^r(t) = \frac{1}{r}\sum_{U_i(t)\le r} \delta_{X_i(t)}$ is the normalized empirical measure for a branching Markov process, (3.6) gives the convergence of the normalized branching Markov process to the corresponding Dawson–Watanabe superprocess (cf. [4, 49]).

4. Examples and extensions.

4.1. *A model with immigration.* The simplest immigration process assumes that the space–time point process giving the arrival times and locations of the immigrants is a Poisson process. Assuming temporal homogeneity, immigration is introduced by adding the generator of a space–time-level Poisson random measure. Let ν be the intensity of immigration, that is, $\nu(A)\Delta t$ is approximately the probability that an individual immigrates into $A \subset E$ in a time interval of length Δt. The generator becomes

$$\begin{aligned} Af(x,u) &= f(x,u)\sum_i \frac{Bg(x_i,u_i)}{g(x_i,u_i)} \\ &\quad + f(x,u)\sum_i 2a(x_i)\int_{u_i}^{r_g} \bigl(g(x_i,v)-1\bigr)\,dv \\ &\quad + f(x,u)\int_0^{r_g}\int_E \bigl(g(z,v)-1\bigr)\nu(dz)\,dv \\ &\quad + f(x,u)\sum_i \bigl(a(x_i)u_i^2 - b(x_i)u_i\bigr)\frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}. \end{aligned} \tag{4.1}$$

Noting that the generator for finite r is obtained from A by restricting the domain to the collection of g with $r_g \le r$, if $ra(x) - b(x) \ge 0$, for all x, the generator of the corresponding branching Markov process with immigration is

$$\begin{aligned} C_r\widehat{f}(x,n) &= \sum_{i=1}^n B_{x_i}\widehat{f}(x,n) + \sum_{i=1}^n ra(x_i)\bigl(\widehat{f}\bigl(b(x|x_i),n+1\bigr) - \widehat{f}(x,n)\bigr) \\ &\quad + \sum_{i=1}^n \bigl(ra(x_i)-b(x_i)\bigr)\bigl(\widehat{f}\bigl(d(x|x_i),n-1\bigr) - \widehat{f}(x,n)\bigr) \\ &\quad + \int_E \bigl(\widehat{f}\bigl(b(x|z),n+1\bigr) - \widehat{f}(x,n)\bigr)\nu(dz). \end{aligned}$$

For $r=\infty$, setting $h(x)=\int_0^\infty(1-g(x,v))\,dv$ as before, the generator for the measure-valued process is

$$C\widehat{f}(\mu)=\alpha Af(\mu)=(\langle -Bh+ah^2-bh,\mu\rangle-\langle h,\nu\rangle)\exp\{-\langle h,\mu\rangle\},$$

for $\widehat{f}(\mu)=\exp\{-\langle h,\mu\rangle\}$.

Early results on branching Markov processes with immigration include [19, 27]. Work on limiting measure-valued processes with immigration includes [8, 37–39, 48].

4.2. *Conditioning on nonextinction.* In the limiting model considered in Section 3.2, let a and b be constant and $b<0$. Let τ be the time of extinction and let $U_*(0)$ be the minimum of the initial levels. Then τ is the solution of $1-U_*(0)\frac{a}{b}(1-e^{-b\tau})=0$, so

$$\tau=-\frac{1}{b}\log\frac{U_*(0)a-b}{U_*(0)a}.$$

If $Z(0)=\mu_0$, then $U_*(0)$ is exponentially distributed with parameter $\mu_0(E)$ and

$$P\{\tau>T\}=P\left\{U_*(0)<\frac{b}{a(1-e^{-bT})}\right\}=1-\exp\left\{-\frac{b\mu_0(E)}{a(1-e^{-bT})}\right\}.$$

The case $b=0$ is obtained by passing to the limit so that $P\{\tau>T\}=1-\exp\{-\mu_0(E)/(aT)\}$.

As in Section 2.2.1, conditioning on $\{\tau>T\}$ and letting $T\to\infty$ is equivalent to conditioning on the initial Poisson process having a level at zero. The resulting generator becomes

$$\begin{aligned}Af(x,u)&=f(x,u)\sum_i\frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\&\quad+f(x,u)\sum_{u_i>0}2a\int_{u_i}^{r_g}\big(g(x_i,v)-1\big)\,dv\\&\quad+f(x,u)2a\int_0^{r_g}\big(g(x_0,v)-1\big)\,dv\\&\quad+f(x,u)\sum_{u_i>0}(au_i^2-bu_i)\frac{\partial_{u_i}g(x_i,u_i)}{g(x_i,u_i)},\end{aligned}\tag{4.2}$$

where the u_i are the nonzero levels, and the generator for the conditioned measure-valued process is given by setting

$$\begin{aligned}\alpha_0 f(\mu)&=\int f(x,u)\alpha_0(\mu,dx\times du)\\&=\frac{1}{|\mu|}\int_E g(z,0)\mu(dz)\exp\{-\langle h,\mu\rangle\}\end{aligned}$$

and

$$\begin{aligned}\alpha_0 Af(\mu) = \langle -Bh(y) + ah(y)^2 - b(y)h(y), \mu\rangle \\ \times \frac{1}{|\mu|}\int_E g(z,0)\mu(dz)\exp\{-\langle h,\mu\rangle\} \\ + \frac{1}{|\mu|}\int_E \bigl(Bg(z,0) - 2ag(z,0)h(z)\bigr)\mu(dz)\exp\{-\langle h,\mu\rangle\}.\end{aligned}$$

Note that α_0 is the distribution of $\xi = \delta_{0,Z_0} + \sum_{i=1}^{\infty}\delta_{(U_i,Z_i)}$, where $\{U_i, i \geq 1\}$ is a Poisson process with intensity $\mu(E)$, and $Z_0, Z_1, \ldots$ are i.i.d. with distribution $\mu(\cdot)/\mu(E)$.

For earlier work, see [10, 12, 20, 36, 40]. In particular, the particle at level zero in the construction above is the "immortal particle" of Evans [10].

4.3. *Conditioning on extinction*. As in Section 2.2.2, if a and b are constant and $b > 0$, then conditioning on extinction is equivalent to conditioning on $U_*(0) > \frac{b}{a}$. Defining $V_i(t) = U_i(t) - \frac{b}{a}$, the generator for the conditioned process is

$$\begin{aligned}Af(x,v) = f(x,v)\sum_i \frac{Bg(x_i,v_i)}{g(x_i,v_i)} \\ + f(x,v)\sum_{i>0} 2a\int_{v_i}^{r_g}\bigl(g(x_i,v) - 1\bigr)\,dv \\ + f(x,v)\sum_{i>0}(av_i^2 - bv_i)\frac{\partial_{v_i} g(x_i,v_i)}{g(x_i,v_i)},\end{aligned}\tag{4.3}$$

and the generator of the measure-valued process is

$$C\widehat{f}(\mu) = \int_E \bigl(-Bh(y) + ah(y)^2 + bh(y)\bigr)\mu(dy)\exp\{-\langle h,\mu\rangle\},$$

for $\widehat{f}(\mu) = \exp\{-\langle h,\mu\rangle\}$. In other words, conditioning a supercritical process on extinction replaces the supercritical process by a subcritical one. This result is originally due to Evans and O'Connell [11].

4.4. *Models with multiple simultaneous births*. We now consider continuous-time, branching Markov processes with general offspring distributions. The general theory of branching Markov processes was developed by Ikeda, Nagasawa and Watanabe in a long series of papers [21–24] following earlier work by several authors. The particle representation is substantially more complicated and passage to the infinite population limit more delicate.

As above, the particles move independently in E according to a generator B. A particle at position $x \in E$ with level u gives birth to k offspring at rate $(k+1)a_k^{(r)}(x)(r-u)^k r^{-(k-1)}$. New particles have the location of the parent, but their levels are uniformly distributed on $[u, r)$. Then for $f(x,u,n) = \prod_{i=1}^n g(x_i, u_i)$,

$$
\begin{aligned}
&A_r f(x,u,n)\\
&\quad = f(x,u,n)\sum_{i=1}^n \frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\
&\qquad + f(x,u,n)\sum_{i=1}^n\sum_{k=1}^\infty \frac{(k+1)a_k^{(r)}(x_i)}{r^{k-1}}\\
&\qquad\qquad \times \int_{[u_i,r)^k}\Biggl[\Biggl(\prod_{l=1}^k g(x_i,v_l)\Biggr)-1\Biggr]\,dv_1\cdots dv_k\\
&\qquad + f(x,u,n)\sum_{i=1}^n\Biggl(\sum_{k=1}^\infty r^2 a_k^{(r)}(x_i)\Biggl[\biggl(1-\frac{u_i}{r}\biggr)^{k+1}\\
&\qquad\qquad\qquad -1+(k+1)\frac{u_i}{r}\Biggr]\\
&\qquad\qquad\qquad\qquad - b(x_i)u_i\Biggr)\frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}.
\end{aligned}
\tag{4.4}
$$

Now, the levels satisfy the equation

$$
\begin{aligned}
\dot U_i(t) = {}&\sum_{k=1}^\infty r^2 a_k^{(r)}(X_i(t))\Biggl[\biggl(1-\frac{U_i(t)}{r}\biggr)^{k+1} - 1 + (k+1)\frac{U_i(t)}{r}\Biggr]\\
&- b(X_i(t))U_i(t).
\end{aligned}
\tag{4.5}
$$

Defining $g(x) = \frac{1}{r}\int_0^r g(x,v)\,dv$ and integrating (4.4) with respect to $\alpha(n, du)$, the uniform measure on $[0, r]^n$, we have that

$$
\begin{aligned}
&\int A_r f(x,u,n)\alpha(n,du)\\
&\quad = C_r f(x,n)\\
&\quad = f(x,n)\sum_{i=1}^n \frac{Bg(x_i)}{g(x_i)}\\
&\qquad + f(x,n)\sum_{i=1}^n\sum_{k=1}^\infty r a_k^{(r)}(x_i)[g(x_i)^k - 1]\\
&\qquad + f(x,n)\sum_{i=1}^n\Biggl(r\sum_{k=1}^\infty k a_k^{(r)}(x_i) - b(x_i)\Biggr)\Biggl[\frac{1}{g(x_i)} - 1\Biggr],
\end{aligned}
\tag{4.6}
$$

which is the generator of a branching process with multiple births with birth rates $ra_k^{(r)}(\cdot)$, death rate $r\sum_{k=1}^\infty ka_k^{(r)}(\cdot)-b(\cdot)$ (provided this expression is nonnegative) and particles moving according to the generator B. The analog of Theorem 3.2 holds with $a(x_i)$ replaced by $\sum_k(k+1)a_k^{(r)}(x_i)$ in the definition of ψ and $\widetilde{\psi}$.

We define

$$\Lambda^{(r)}(x,u)=\sum_{k=1}^\infty r(k+1)a_k^{(r)}(x)\left[1-\left(1-\frac{u}{r}\right)^k\right]$$

and assume that

$$\begin{aligned}\lim_{r\to\infty}&\Lambda^{(r)}(x,u)\\ &=\lim_{r\to\infty}\sum_{k=1}^\infty r(k+1)a_k^{(r)}(x)\left[\sum_{l=1}^k\binom{k}{l}(-1)^{l+1}\left(\frac{u}{r}\right)^l\right]\\ &\equiv\Lambda(x,u)\end{aligned}\tag{4.7}$$

exists uniformly for $x\in E$ and u in bounded intervals. This condition is essentially (9.4.36) of [9].

Observe that

$$\int_0^u\Lambda^{(r)}(x,v)\,dv=\sum_{k=1}^\infty r^2a_k^{(r)}(x)\left[\left(1-\frac{u}{r}\right)^{k+1}-1+(k+1)\frac{u}{r}\right],$$

so that (4.5) becomes

$$\dot{U}_i(t)=\int_0^{U_i(t)}\Lambda^{(r)}(X_i(s),v)\,dv-b(X_i(t))U_i(t)\tag{4.8}$$

and

$$\int_0^r\Lambda^{(r)}(x,v)\,dv=\sum_{k=1}^\infty r^2ka_k^{(r)}(x),$$

so that the death rate for the branching process can be written as $r^{-1}\int_0^r\Lambda^{(r)}(x,v)\,dv-b(x)$.

As in [31],

$$\begin{aligned}\partial^m\Lambda^{(r)}(x,u)&\equiv\frac{\partial^m}{\partial u^m}\Lambda^{(r)}(x,u)\\ &=(-1)^{m+1}\sum_{k=m}^\infty a_k^{(r)}(x)\frac{(k+1)k\cdots(k-m+1)}{r^{m-1}}\\ &\qquad\times\left(1-\frac{u}{r}\right)^{k-m}.\end{aligned}\tag{4.9}$$

Each of the derivatives is a monotone function of u, and since

$$\begin{aligned}\partial^{m-1}\Lambda^{(r)}(x,b) - \partial^{m-1}\Lambda^{(r)}(x,a)\\ = \int_a^b \partial^m\Lambda^{(r)}(x,u)\,du,\end{aligned}$$

it follows by the convergence of $\Lambda^{(r)}$ and induction on m that each $\partial^m\Lambda^{(r)}(x,u)$ converges uniformly in u on bounded intervals that are bounded away from zero. Consequently, Λ is infinitely differentiable in u, and for $0 < u_1 < u_2 < \infty$,

$$\lim_{r\to\infty}\sup_{u_1\le u\le u_2}\left|\partial^m\Lambda^{(r)}(x,u) - \partial^m\Lambda(x,u)\right| = 0.$$

The fact that the derivatives alternate in sign implies that $\partial^1\Lambda(x,\cdot)$ is completely monotone and hence can be represented as

$$\partial^1\Lambda(x,u) = \int_0^\infty e^{-uz}\widehat{\nu}(x,dz)$$

for some σ-finite measure $\widehat{\nu}(x,\cdot)$. Writing $\widehat{\nu}(x,\cdot) = 2a_0(x)\delta_0 + \nu(x,\cdot)$ with $\nu(x,\{0\}) = 0$,

$$\Lambda(x,u) = 2a_0(x)u + \int_0^\infty z^{-1}(1-e^{-uz})\nu(x,dz).$$

Let g satisfy $g(x,v) = 1$, for $v \ge u_g$, and define

$$h(x,u) = \int_u^{u_g}\bigl(1-g(x,v)\bigr)\,dv.$$

If $r > u_g$ and there are n particles below level r, then (4.4) may be written as

$$\begin{aligned}&A_r f(x,u,n)\\ &\quad= f(x,u,n)\sum_{i=1}^n \frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\ &\qquad+ f(x,u,n)\\ &\qquad\quad\times\sum_{i=1}^n\sum_{k=1}^\infty r(k+1)\\ &\qquad\qquad\times a_k^{(r)}(x_i)\left\{\left(1-\frac{u_i+h(x_i,u_i)}{r}\right)^k - \left(1-\frac{u_i}{r}\right)^k\right\}\\ &\qquad+ f(x,u,n)\sum_{i=1}^n\left(\int_0^{u_i}\Lambda^{(r)}(x_i,v)\,dv - b(x_i)u_i\right)\frac{\partial_{u_i}g(x_i,u_i)}{g(x_i,u_i)}.\end{aligned}$$

Then, by (4.7) and the definition of h, we have

$$
\begin{aligned}
&Af(x,u)\\
&\quad\equiv \lim_{r\to\infty} A_r f(x,u)\\
&\quad= f(x,u)\sum_i \frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\
&\qquad + f(x,u)\sum_i \bigl(\Lambda(x_i,u_i) - \Lambda\bigl(x_i,u_i+h(x_i,u_i)\bigr)\bigr)\\
&\qquad + f(x,u)\sum_i \biggl(\int_0^{u_i}\Lambda(x_i,v)\,dv - b(x_i)u_i\biggr)\frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}\\
&\quad= f(x,u)\sum_i \frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\
&\qquad + f(x,u)\sum_i 2a_0(x_i)\int_{u_i}^\infty \bigl(g(x_i,v)-1\bigr)\,dv\\
&\qquad + f(x,u)\sum_i \int_0^\infty \bigl(e^{z\int_{u_i}^\infty (g(x_i,v)-1)\,dv} - 1\bigr)z^{-1}e^{-zu_i}\nu(x_i,dz)\\
&\qquad + f(x,u)\sum_i \biggl(a_0(x_i)u_i^2 - b(x_i)u_i\\
&\qquad\qquad\qquad + \int_0^\infty z^{-1}\bigl(u_i - z^{-1}(1-e^{-u_i z})\bigr)\nu(x_i,dz)\biggr)\frac{\partial_{u_i} g(x_i,u_i)}{g(x_i,u_i)}.
\end{aligned}
$$

Note that the second term on the right-hand side has the same interpretation as the second term on the right-hand side of (3.5). To understand the third term, recall that if $\xi = \sum_i \delta_{\tau_i}$ is a Poisson process on $[0,\infty)$ with parameter λ, then

$$E\Bigl[\prod g(\tau_i)\Bigr] = e^{\lambda\int_0^\infty (g(v)-1)\,dv}.$$

Consequently, the third term determines bursts of simultaneous offspring at the location x_i of the parent and with levels forming a Poisson process with intensity z on $[u_i,\infty)$.

Setting $h(x) = h(x,0) = \int_0^\infty (1-g(x,v))\,dv$ and $\widehat{f}(\mu) = \exp\{-\langle h,\mu\rangle\}$,

$$
\begin{aligned}
C\widehat{f}(\mu) &= \alpha Af(\mu)\\
&= \int_E \biggl(-Bh(y) + \int_0^{h(y)} \Lambda(y,z)\,dz - b(y)h(y)\biggr)\mu(dy)\exp\{-\langle h,\mu\rangle\}.
\end{aligned}
$$

Based on the above calculations, we have the following theorem.

THEOREM 4.1. *Let $B \subset \overline{C}(E) \times \overline{C}(E)$ satisfy Condition 3.1, and let the martingale problem for B be well posed. Assume that for $r \geq r_0$,*

$$\inf_x \left(r \sum_{k=1}^{\infty} k a_k^{(r)}(x) - b(x) \right) \geq 0$$

and that the convergence in (4.7) is uniform in x. Let $K(0)$ be a finite random measure on E, and let ξ^r be a solution of the martingale problem for A_r such that $\xi^r(0)$ is conditionally Poisson on $E \times [0, r]$ with mean measure $K(0) \times \Lambda$. Then $\xi^r \Rightarrow \xi$ where ξ is a solution of the martingale problem for A.

PROOF. For $r > q$, let $\xi^{(q),r}$ denote the restriction of ξ^r to $E \times [0, q]$ and similarly for $\xi^{(q)}$. It is enough to prove that $\xi^{(q),r} \Rightarrow \xi^{(q)}$ for each $q > r_0$. The generator for $\xi^{(q),r}$ is the restriction of A_r to functions $f \in \mathcal{D}(A_r)$ such that the corresponding g satisfies $g(x, u) = 1$ for $u \geq q$. For f of this form, by (4.9), $A_{r,q} f = A_r f$ satisfies

$$\begin{aligned}
&A_{r,q} f(x, u, n) \\
&\quad = f(x, u, n) \sum_{i=1}^{n} \frac{Bg(x_i, u_i)}{g(x_i, u_i)} \\
&\qquad + f(x, u, n) \sum_{i=1}^{n} \sum_{m=1}^{\infty} \frac{1}{m!} (-1)^{m+1} \partial^m \Lambda^{(r)}(x_i, q) \\
&\qquad\qquad \times \int_{[u_i, q)^m} \left[\left(\prod_{l=1}^{m} g(x_i, v_l) \right) - 1 \right] dv_1 \cdots dv_m \\
&\qquad + f(x, u, n) \sum_{i=1}^{n} \left(\int_0^{u_i} \Lambda^{(r)}(x, v)\, dv - b(x_i) u_i \right) \frac{\partial_{u_i} g(x_i, u_i)}{g(x_i, u_i)},
\end{aligned}$$

and the corresponding branching Markov process has generator

$$\begin{aligned}
&C_{r,q} f(x, n) \\
&\quad = f(x, n) \sum_{i=1}^{n} \frac{Bg(x_i)}{g(x_i)} \\
&\qquad + f(x, n) \sum_{i=1}^{n} \sum_{m=1}^{\infty} \frac{1}{(m+1)!} (-1)^{m+1} \partial^m \Lambda^{(r)}(x_i, q) q^m [g(x_i)^m - 1] \\
&\qquad + f(x, n) \sum_{i=1}^{n} \left(\frac{1}{q} \int_0^{q} \Lambda^{(r)}(x_i, v)\, dv - b(x_i) \right) \left[\frac{1}{g(x_i)} - 1 \right].
\end{aligned}$$

The convergence of $\xi^{(q),r}$ follows by the convergence assumptions on $\Lambda^{(r)}$. $\square$

The measure $\nu(x,\cdot)$ is nonzero only if the offspring distribution has a "heavy tail." If $a_k^{(r)}(x)=a_k(x)$ and

$$\sum_{k=1}^{\infty}(k+1)ka_k(x)<\infty,$$

then

$$\Lambda(x,u)=\lim_{r\to\infty}\sum_{k=1}^{\infty}r(k+1)a_k(x)\left[1-\left(1-\frac{u}{r}\right)^k\right]=\sum_{k=1}^{\infty}(k+1)ka_k(x)u$$

and

$$\begin{aligned}Af(x,u)&=f(x,u)\sum_i\frac{Bg(x_i,u_i)}{g(x_i,u_i)}\\&\qquad+f(x,u)\sum_i\sum_{k=1}^{\infty}(k+1)ka_k(x_i)\int_{u_i}^{u_g}[g(x_i,v)-1]\,dv\\&\qquad+f(x,u)\sum_i\left(\sum_{k=1}^{\infty}\frac{(k+1)ka_k(x_i)}{2}u_i^2-b(x_i)u_i\right)\frac{\partial_{u_i}g(x_i,u_i)}{g(x_i,u_i)},\end{aligned}$$

which is essentially (3.5).

For scalar branching processes with general offspring distributions, convergence to possibly discontinuous continuous state branching processes was proved by Grimvall [16] (see [9], Section 9.1). Convergence in the measure-valued setting is given in [49] and [4] for offspring distributions with finite second moment and more generally in [9], Theorem 9.4.3. Fitzsimmons [13] gives a very general construction of these processes.

If ν is not zero, then the genealogy of the process is much more complicated than that described in Sections 2.4 and 3.3. Assume that Λ and b do not depend on x, and define

$$\widehat{\Lambda}(u)\equiv\int_0^u\Lambda(v)\,dv-bu=2a_0u^2-bu+\int_0^{\infty}z^{-2}(zu-1+e^{-uz})\nu(dz).$$

If u_T satisfies

$$\dot{u}_T(t)=\widehat{\Lambda}(u_T(t)),$$

$u_T(t)<\infty$ for $t<T$ and $\lim_{t\to T-}u_T(t)=\infty$, then it is still the case that the collection of ancestors at time $t<T$ of the population alive at time T is $\{(X_i(t),U_i(t)):U_i(t)<u_T(t)\}$, but u_T may not exist. In fact, since if $\dot{u}=\widehat{\Lambda}(u)$

$$\int_{u(t_1)}^{u(t_2)}\frac{1}{\widehat{\Lambda}(v)}\,dv=t_2-t_1,$$

u_T exists if and only if

$$\int_u^\infty \frac{1}{\widehat{\Lambda}(v)}\,dv < \infty$$

for u sufficiently large, which always holds if $a_0 > 0$. In the critical and subcritical cases, this condition is equivalent to extinction with probability one as was noted by Bertoin and Le Gall ([2], page 167).

This *finite ancestry* property or *coming down from infinity* of the genealogy has been studied for a variety of population models. See [47] and [1] for results in the Fleming–Viot setting. The equivalence of the conditions for Fleming–Viot and Dawson–Watanabe processes is given in ([2], page 171).

The argument in Section 2.2.3 can undoubtedly be extended to the present setting. This development will be carried out elsewhere.

4.5. *Model with exponentially distributed levels.* The discrete models that we have considered have been formulated with levels that are uniformly distributed on an interval. That is not necessary, and other distributions may be convenient in other contexts. We illustrate this flexibility by formulating a model for a simple branching process with levels that are exponentially distributed. The dynamics of the levels change, and the correct dynamics are determined by essentially working backwards from the answer.

As before, let $f(u,n) = \prod_{i=1}^n g(u_i)$ where $0 \le g \le 1$ and $g(u_i) = 1$ for $u_i \ge u_g$. Let

$$A_r f(u,n) = f(u,n)\sum_{i=1}^n 2a\int_{u_i}^\infty e^{-v/r}\big(g(v)-1\big)\,dv + f(u,n)\sum_{i=1}^n G_r(u_i)\frac{g'(u_i)}{g(u_i)},$$

where G_r will be determined below. Note that a particle at level u_i is giving birth at rate $2are^{-u_i/r}$, and the levels satisfy

$$\dot{U}_i(t) = G_r(U_i(t)).$$

Let $\alpha_r(n,du)$ be the distribution of n independent exponential random variables with mean r, and define $e^{-\lambda_g} = r^{-1}\int_0^\infty g(v)e^{-v/r}\,dv$ so

$$\widehat{f}(n) = \int f(u,n)\alpha_r(n,du) = e^{-\lambda_g n}.$$

To calculate $\int A_r f(u,n)\alpha_r(n,du)$, observe that

$$\begin{aligned} r^{-1}2a&\int_0^\infty e^{-z/r}g(z)\int_z^\infty e^{-v/r}\big(g(v)-1\big)\,dv\,dz \\ &= are^{-2\lambda_g} - 2a\int_0^\infty e^{-2z/r}g(z)\,dz, \end{aligned}$$

and assuming $G_r(0)=0$,

$$\begin{aligned}
&r^{-1}\int_0^\infty e^{-z/r}G_r(z)g'(z)\,dz\\
&\qquad = -r^{-1}\int_0^\infty e^{-z/r}\bigl(G_r'(z)-r^{-1}G_r(z)\bigr)\bigl(g(z)-1\bigr)\,dz\\
&\qquad = -r^{-1}\int_0^\infty e^{-z/r}\bigl(G_r'(z)-r^{-1}G_r(z)\bigr)g(z)\,dz\\
&\qquad\qquad + r^{-1}\int_0^\infty e^{-z/r}\bigl(G_r'(z)-r^{-1}G_r(z)\bigr)\,dz.
\end{aligned}$$

Then for

$$G_r'(z)-r^{-1}G_r(z)=e^{z/r}\frac{d}{dz}(e^{-z/r}G_r(z))=2ar(1-e^{-z/r})-b,$$

we have

$$e^{-z/r}G_r(z)=2ar\biggl(r(1-e^{-z/r})-\frac{r}{2}(1-e^{-2z/r})\biggr)-br(1-e^{-z/r})$$

and

$$\begin{aligned}
&\int A_r f(u,n)\alpha_r(n,du)\\
&\quad = ne^{-\lambda_g(n-1)}\biggl(are^{-2\lambda_g}-r^{-1}\int_0^\infty e^{-z/r}\bigl(G_r'(z)-r^{-1}G_r(z)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad + 2are^{-z/r}\bigr)g(z)\,dz\\
&\qquad\qquad\qquad\qquad + r^{-1}\int_0^\infty e^{-z/r}\bigl(G_r'(z)-r^{-1}G_r(z)\bigr)\,dz\biggr)\\
&\quad = C_r\widehat{f}(n),
\end{aligned}$$

where

$$C_r\widehat{f}(n)=ran\bigl(\widehat{f}(n+1)-\widehat{f}(n)\bigr)+(ra-b)n\bigl(\widehat{f}(n-1)-\widehat{f}(n)\bigr) \tag{4.10}$$

is the generator of a branching process.

Note that as $r\to\infty$, $G_r(z)$ converges to az^2-bz, and hence, A_r converges to A given by (2.4).

4.6. *Multitype branching processes*. We now consider a branching particle system with m possible types, $S=\{1,2,\dots,m\}$. We assume that a particle of type $\zeta_1\in S$ gives birth to a particle of type $\zeta_2\in S$ at rate $ra^{(r)}(\zeta_1,\zeta_2)$ and dies at rate $ra^{(r)}(\zeta_1)-b^{(r)}(\zeta_1)$, where $a^{(r)}(\zeta_1)=\sum_{j\in S}a^{(r)}(\zeta_1,j)$.

The fact that the ordered representations constructed for the previous examples give the correct measure-valued processes depends on the fact that observing a birth event in the measure-valued process gives no information about the levels of the particles after the birth event. That, in turn, depends on the offspring being indistinguishable from the parent. Since in the current model, the type of an offspring may differ from the type of the parent, we need to find a way to "preserve ignorance" about the levels when the type of the offspring is different. We accomplish this goal by randomizing the assignment of the parent and offspring to the original level of the parent and a new level. Let $f(\zeta, u, n)$ be of the form

$$f(\zeta, u, n) = \prod_{i=1}^{n} g(\zeta_i, u_i).$$

Then the generator of the ordered representation of the branching process described above is given by

$$\begin{aligned} &A_r f(\zeta, u, n) \\ &\quad = f(\zeta, u, n) \sum_{i=1}^{n} \sum_{j \in S} 2a^{(r)}(\zeta_i, j) \\ &\qquad\qquad \times \int_{u_i}^{r} \left[\frac{1}{2} \left(\frac{g(\zeta_i, u_i) g(j, v) + g(\zeta_i, v) g(j, u_i)}{g(\zeta_i, u_i)} \right) - 1 \right] dv \\ &\qquad + f(\zeta, u, n) \sum_{i=1}^{n} \left[a^{(r)}(\zeta_i) u_i^2 - b^{(r)}(\zeta_i) u_i \right] \frac{\partial_{u_i} g(\zeta_i, u_i)}{g(\zeta_i, u_i)}, \end{aligned}$$

where as before, each level satisfies

$$\frac{d}{dt} U_i^{(r)}(t) = a^{(r)}(X_i(t)) U_i^2(t) - b^{(r)}(X_i(t)) U_i(t).$$

Let $Q^{(r)} h(\zeta) = \sum_{j \in S} a^{(r)}(\zeta, j)[h(j) - h(\zeta)]$. Because of the randomization of the level assignments at each birth event, it follows that

$$h(X_i(t)) - \int_{\tau_i}^{t} \big(r - U_i(s)\big) Q^{(r)} h(X_i(s))\, ds$$

is a martingale.

Taking $\alpha_r(n, du)$ as before, we have that

$$\begin{aligned} \int A_r f(\zeta, u, n) \alpha(n, du) &= C_r \overline{f}(\zeta, n) \\ &= \overline{f}(\zeta, n) \sum_{i=1}^{n} \sum_{j \in S} r a^{(r)}(\zeta_i, j)[\overline{g}(j) - 1] \\ &\quad + \overline{f}(\zeta, n) \sum_{i=1}^{n} \left[r a^{(r)}(\zeta_i) - b^{(r)}(\zeta_i) \right] \left[\frac{1}{\overline{g}(\zeta_i)} - 1 \right], \end{aligned}$$

where $\overline{f}(\zeta, n) = \prod_{i=1}^{n} \overline{g}(\zeta_i)$ and $\overline{g}(\zeta_i) = \frac{1}{r}\int_0^r g(\zeta_i, v)\,dv$. Hence, $C_r \overline{f}(\zeta, n)$ is the generator of a multitype branching process.

Assume that

$$\begin{aligned} a(\zeta, j) &= \lim_{r\to\infty} a^{(r)}(\zeta, j),\\ a(\zeta) &= \lim_{r\to\infty} a^{(r)}(\zeta) = \sum_{j\in S} a(\zeta, j),\\ b(\zeta) &= \lim_{r\to\infty} b^{(r)}(\zeta) \end{aligned}$$

and that

$$Qh(\zeta) = \sum_{j\in S} a(\zeta, j)[h(j) - h(\zeta)]$$

is the generator of an irreducible, finite state Markov chain. Let π denote the unique stationary distribution for Q. It is clear from the ergodicity of the Markov chain that in the limit, the levels must satisfy

$$\frac{d}{dt} U_i(t) = \overline{a} U_i^2(t) - \overline{b} U_i(t),$$

where $\overline{a} = \sum_j \pi(j)a(j)$ and $\overline{b} = \sum_j \pi(j)b(j)$.

We can make this observation precise by analyzing the asymptotic behavior of the generator. Taking $g(\zeta, u) = \exp(-h_0(u) + \frac{1}{r}h(\zeta, u))$, where $h(\zeta, u)$ and $h_0(u)$ are equal to zero if $u \geq u_g$, and letting $r \to \infty$, we have that

$$\begin{aligned} \lim_{r\to\infty} f(\zeta, u) &= \lim_{r\to\infty} \exp\biggl(-\sum_i h_0(u_i) + \frac{1}{r}\sum_i h(\zeta_i, u_i)\biggr)\\ &= \exp\biggl(-\sum_i h_0(u_i)\biggr) \equiv \overline{f}(u) \end{aligned}$$

and since

$$\begin{aligned} &\frac{g(\zeta_i, u_i)g(j, v) + g(\zeta_i, v)g(j, u_i)}{g(\zeta_i, u_i)}\\ &\qquad = e^{-h_0(v)}\bigl(e^{r^{-1}h(j,v)} + e^{r^{-1}(h(\zeta_i,v)+h(j,u_i)-h(\zeta_i,u_i))}\bigr),\\ &\lim_{r\to\infty} A_r f(\zeta, u)\\ &\qquad = \overline{f}(u)\sum_i \biggl\{2a(\zeta_i)\int_{u_i}^{u_g}[e^{-h_0(v)} - 1]\,dv\\ &\qquad\qquad + \sum_{j\in S} a(\zeta_i, j)[h(j, u_i) - h(\zeta_i, u_i)]\\ &\qquad\qquad - [a(\zeta_i)u_i^2 - b(\zeta_i)u_i]\,\partial_{u_i} h_0(u_i)\biggr\}. \end{aligned} \tag{4.11}$$

If $\sum \pi(j)G(j,u) \equiv 0$ for all u, then there exists h such that

$$\sum_{j\in S} a(\zeta,j)[h(j,u)-h(\zeta,u)] = G(\zeta,u).$$

Consequently, there exists h such that the right-hand side of (4.11) becomes

$$A\overline{f}(u) = \overline{f}(u)\sum_i \left\{2\overline{a}\int_{u_i}^{u_g}[e^{-h_0(v)}-1]\,dv - [\overline{a}u_i^2 - \overline{b}u_i]\,\partial_{u_i}h_0(u_i)\right\},$$

which is just a rewriting of (2.4), and hence we have convergence of the normalized total population to the Feller diffusion.

For earlier work, see [14, 25, 29] and Section 9.2 of [9].

4.7. *Models with catastrophic death.* Now consider

$$\begin{aligned}
A_r f(x,u,n) &= f(x,u,n)\sum_{i=1}^n \frac{Bg(x_i,u_i)}{g(x_i,u_i)} \\
&\quad + f(x,u,n)\sum_{i=1}^n 2a(x_i)\int_{u_i}^r \big(g(x_i,v)-1\big)\,dv \\
&\quad + f(x,u,n)\sum_{i=1}^n \big(a(x_i)u_i^2 - b(x_i)u_i\big)\frac{\partial_{u_i}g(x_i,u_i)}{g(x_i,u_i)} \\
&\quad + \int_V \big(f(x,c(u,x,v),n) - f(x,u,n)\big)\gamma(dv),
\end{aligned}\tag{4.12}$$

where γ is a σ-finite measure on a measurable space $(V,\mathcal{V})$,

$$c(u,x,v) = (u_1\rho(x_1,v), u_2\rho(x_2,v),\dots)$$

and $\rho(x_i,v)\ge 1$. Then as in Section 3.1

$$\begin{aligned}
C_r\widehat{f}(x,n) &= \sum_{i=1}^n B_{x_i}\widehat{f}(x,n) + \sum_{i=1}^n ra(x_i)\big(\widehat{f}\big(b(x|x_i),n+1\big) - \widehat{f}(x,n)\big) \\
&\quad + \sum_{i=1}^n \big(ra(x_i)-b(x_i)\big)\big(\widehat{f}\big(d(x|x_i),n-1\big) - \widehat{f}(x,n)\big) \\
&\quad + \int_V \left(\prod_{i=1}^n \big(\rho^{-1}(x_i,v)\widehat{g}(x_i) + \big(1-\rho^{-1}(x_i,v)\big)\big) - \widehat{f}(x,n)\right)\gamma(dv).
\end{aligned}$$

For simplicity, assume that $\gamma(V)<\infty$. Then at rate $\gamma(V)$ an event occurs in which an element v is selected from V, and given v, particles are independently killed, with the probability that a particle at x_i survives being $\rho^{-1}(x_i,v)$.

Letting $r \to \infty$ to obtain A and integrating,

$$\begin{aligned}\alpha A f &= \int_E \big(-Bh(y) + a(y)h(y)^2 - b(y)h(y)\big)\mu(dy) \exp\{-\langle h, \mu\rangle\} \\ &\qquad + \int_V \big(\exp\{-\langle \rho^{-1}(\cdot, v)h, \mu\rangle\} - \exp\{-\langle h, \mu\rangle\}\big)\gamma(dv) \\ &= C\widehat{f}(\mu).\end{aligned}$$

Branching processes with catastrophes have been considered in a series of papers by Pakes [41–46] and by Grey [15].

APPENDIX

A.1. Poisson random measures. Let $(S, \mathcal{S})$ be a measurable space, and let ν be a σ-finite measure on $\mathcal{S}$. ξ is a *Poisson random measure* with mean measure ν if:

(a) ξ is a random counting measure on S;
(b) for each $A \in \mathcal{S}$ with $\nu(A) < \infty$, $\xi(A)$ is Poisson distributed with parameter $\nu(A)$;
(c) for $A_1, A_2, \ldots \in \mathcal{S}$ disjoint, $\xi(A_1), \xi(A_2), \ldots$ are independent.

LEMMA A.1. *If $H : S \to S_0$ is Borel measurable and $\widehat{\xi}(A) = \xi(H^{-1}(A))$, then $\widehat{\xi}$ is a Poisson random measure on S_0 with mean measure $\widehat{\nu}$ given by $\widehat{\nu}(A) = \nu(H^{-1}(A))$.*

REMARK A.2. $\widehat{\nu}$ need not be σ-finite even if ν is, but the meaning of the lemma should still be clear. σ-finite or not, $\widehat{\nu}(A) = \infty$ if and only if $\widehat{\xi}(A) = \infty$ a.s.

PROOF OF LEMMA A.1. The lemma follows from the fact that $A_1, A_2, \ldots$ disjoint implies $H^{-1}(A_1), H^{-1}(A_2), \ldots$ are disjoint. $\square$

LEMMA A.3. *If ξ is a Poisson random measure with mean measure ν and $f \in L^1(\nu)$, then*

$$E\big[e^{\int f(z)\xi(dz)}\big] = e^{\int (e^f - 1)\, d\nu}, \tag{A.1}$$

$$E\left[\int f(z)\xi(dz)\right] = \int f\, d\nu, \qquad \operatorname{Var}\left(\int f(z)\xi(dz)\right) = \int f^2\, d\nu, \tag{A.2}$$

allowing $\infty = \infty$.

Letting $\xi = \sum_i \delta_{Z_i}$, for $g \geq 0$ with $\log g \in L^1(\nu)$,

$$E\left[\prod_i g(Z_i)\right] = e^{\int (g-1)\, d\nu}.$$

Similarly, if $hg, g-1 \in L^1(\nu)$, *then*

$$E\Big[\sum_j h(Z_j)\prod_i g(Z_i)\Big] = \int hg\,d\nu\, e^{\int (g-1)\,d\nu}$$

and

$$E\Big[\sum_{i\neq j} h(Z_i)h(Z_j)\prod_k g(Z_k)\Big] = \Big(\int hg\,d\nu\Big)^2 e^{\int (g-1)\,d\nu}.$$

PROOF. The independence properties of ξ imply (A.1) for simple functions. The general case follows by approximation. The other identities follow in a similar manner. Note that the integrability of the random variables in the expectations above can be verified by replacing g by $(|g| \wedge a)\mathbf{1}_A + \mathbf{1}_{A^c}$ and h by $(|h| \wedge a)\mathbf{1}_A$ for $0 < a < \infty$ and $\nu(A) < \infty$ and passing to the limit as $a \to \infty$ and $A \nearrow E$. $\square$

LEMMA A.4. *If* $\xi_0 = \sum_i \delta_{U_i}$ *is a Poisson random measure on* $[0,\infty)$ *with mean measure* $\lambda\Lambda$, Λ *Lebesgue measure, and* $\{X_i\}$ *are i.i.d. positive random variables, independent of* ξ_0, *then*

$$\xi = \sum_i \delta_{(X_i,U_i)}$$

is a Poisson random measure on $[0,\infty)^2$ *with mean measure* $\lambda\mu_X \times \Lambda$, *were* μ_X *is the law of* X_1.

If $\kappa = E[\frac{1}{X_i}] < \infty$, *then*

$$\widehat{\xi} = \sum \delta_{X_i U_i}$$

is a Poisson random measure on $[0,\infty)$ *with mean measure* $\lambda\kappa\Lambda$.

PROOF. By Lemma A.1, $\widehat{\xi}$ is a Poisson random measure with mean measure given by

$$\begin{aligned}\widehat{\nu}[0,c] &= \lambda\mu_X \times \Lambda\{(x,u) : xu \le c\} = \lambda\int_0^\infty P\{X^{-1} \ge uc^{-1}\}\,du \\ &= \lambda c E[X^{-1}].\end{aligned}$$

$\square$

A.2. Conditionally Poisson systems. We begin by considering general conditionally Poisson systems or Cox processes. Consider (S,d) a metric space, and let ξ be a random counting measure on S and Ξ be a locally finite random measure on S. [A measure ν on S is locally finite if for each $x \in S$, there exists an $\epsilon > 0$ such that $\nu(B_\varepsilon(x)) < \infty$.] We say that ξ is conditionally Poisson with Cox measure Ξ if, conditioned on Ξ, ξ is a Poisson random measure with mean measure Ξ. This requirement is equivalent to

$$E[e^{-\int_S f\,d\xi}] = E\big[e^{-\int_S (1-e^{-f})\,d\Xi}\big],$$

for all nonnegative $f \in M(S)$, where $M(S)$ is the set of all Borel measurable functions on S. Since the collection of functions $F_f(\mu) = e^{-\int_S f\,d\mu}$ is closed under multiplication and separates points in the space of locally finite measures, the distribution of Ξ determines the distribution of ξ.

We are actually interested in the conditionally Poisson system on $S \times [0, \infty)$ with Cox measure $\Xi \times \Lambda$, where Λ is Lebesgue measure. Then for nonnegative $f \in M(S)$, we have

$$E[e^{-\int_{S\times[0,K]} f\,d\xi}] = E\big[e^{-K\int_S (1-e^{-f})\,d\Xi}\big],$$

and the distribution of ξ determines the distribution of Ξ, where we consider $f \in M(S)$ to be a function on $S \times [0, K]$ satisfying $f(x, u) = f(x)$. In particular,

$$\Xi(f) = \lim_{K\to\infty} \frac{1}{K} \int_{S\times[0,K]} f\,d\xi \qquad \text{a.s.}$$

LEMMA A.5. *Suppose ξ is a conditionally Poisson random measure on $S \times [0, \infty)$ with Cox measure $\Xi \times \Lambda$, and let $f \in M(S)$, $0 \le f \le 1$. Then for $C, D > 0$,*

$$P\left\{\int_{S\times[0,K]} f\,d\xi \ge C\right\} \le \frac{KD}{C} + P\left\{\int_S f\,d\Xi \ge D\right\} \tag{A.3}$$

and

$$P\left\{\int_S f\,d\Xi \ge C\right\} \le \frac{E[1 - e^{-C^{-1}\int_{S\times[0,K]} f\,d\xi}]}{1 - e^{-Ke^{-C^{-1}}}}. \tag{A.4}$$

Let $\{\xi_\alpha, \alpha \in \mathcal{A}\}$ be a collection of conditionally Poisson random measures on $S \times [0, \infty)$ with Cox measures $\Xi_\alpha \times \Lambda$, and let $f \in M(S)$, $0 \le f \le 1$. Then $\{\int_{S\times[0,K]} f\,d\xi_\alpha, \alpha \in \mathcal{A}\}$ is stochastically bounded if and only if $\{\int_S f\,d\Xi_\alpha, \alpha \in \mathcal{A}\}$ is stochastically bounded.

PROOF. Since $E[\int_{S\times[0,K]} f\,d\xi \,|\, \Xi] = K\int_S f\,d\Xi$,

$$P\left\{\int_{S\times[0,K]} f\,d\xi \ge C \,\middle|\, \Xi\right\} \le \frac{K\int_S f\,d\Xi}{C} \wedge 1 \le \frac{KD}{C} + \mathbf{1}_{\{\int_S f\,d\Xi \ge D\}},$$

and taking expectations gives (A.3).

By (A.1)

$$\begin{aligned} E[1 - e^{-\int_{S\times[0,K]} \varepsilon f\,d\xi}] &= E\big[1 - e^{-K\int_S (1-e^{-\varepsilon f})\,d\Xi}\big] \\ &\ge E[1 - e^{-\varepsilon K e^{-\varepsilon}\int_S f\,d\Xi}] \\ &\ge (1 - e^{-\varepsilon K e^{-\varepsilon} C}) P\left\{\int_S f\,d\Xi \ge C\right\}, \end{aligned}$$

and taking $\varepsilon = C^{-1}$ gives (A.4).

The final statement follows from the two inequalities. $\square$

Let $\widehat{\xi} = \sum \delta_{X_i}$ be a point process on S, and let $\{U_i\}$ be independent random variables, uniformly distributed on $[0, r]$ and independent of $\widehat{\xi}$. Define

$$\xi = \sum \delta_{(X_i, U_i)}, \qquad \Xi_r = r^{-1}\widehat{\xi}. \tag{A.5}$$

Then for $f \geq 0$ on $S \times [0, r]$,

$$E[e^{-\int_{S\times[0,r]} f\, d\xi} | \Xi_r] = \prod_i \left(r^{-1} \int_0^r e^{-f(X_i,u)}\, du \right) = e^{-\int_S F_f^r(x)\Xi_r(dx)}, \tag{A.6}$$

where

$$F_f^r(x) = -r \log \frac{1}{r} \int_0^r e^{-f(x,u)}\, du = -r \log\left(1 - \frac{1}{r} \int_0^r (1 - e^{-f(x,u)})\, du\right).$$

We have the following analog of Lemma A.5.

LEMMA A.6. *Suppose ξ and Ξ_r are given by* (A.5), *and let $f \in M(S)$, $0 \leq f \leq 1$. Then for $C, D > 0$ and $K \leq r$,*

$$P\left\{\int_{S\times[0,K]} f\, d\xi \geq C\right\} \leq \frac{KD}{C} + P\left\{\int_S f\, d\Xi_r \geq D\right\} \tag{A.7}$$

and

$$P\left\{\int_S f\, d\Xi_r \geq C\right\} \leq \frac{E[1 - e^{-C^{-1}\int_{S\times[0,K]} f\, d\xi}]}{1 - e^{-Ke^{-C^{-1}}}}. \tag{A.8}$$

PROOF. Since $E[\int_{S\times[0,K]} f\, d\xi | \Xi_r] = K \int_S f\, d\Xi_r$,

$$P\left\{\int_{S\times[0,K]} f\, d\xi \geq C \middle| \Xi_r\right\} \leq \frac{K \int_S f\, d\Xi_r}{C} \wedge 1 \leq \frac{KD}{C} + \mathbf{1}_{\{\int_S f\, d\Xi_r \geq D\}},$$

and taking expectations gives (A.7).

Defining

$$G_{K,\varepsilon,f}^r(x) = -r \log\left(1 - \frac{K}{r}(1 - e^{-\varepsilon f(x)})\right) \geq \varepsilon K e^{-\varepsilon} f(x),$$

by (A.6)

$$\begin{aligned} E[1 - e^{-\int_{S\times[0,K]} \varepsilon f\, d\xi}] &= E[1 - e^{-\int_S G_{K,\varepsilon,f}^r\, d\Xi_r}] \\ &\geq E[1 - e^{-\varepsilon K e^{-\varepsilon}\int_S f\, d\Xi_r}] \\ &\geq (1 - e^{-\varepsilon K e^{-\varepsilon} C}) P\left\{\int_S f\, d\Xi_r \geq C\right\}, \end{aligned}$$

and taking $\varepsilon = C^{-1}$ gives (A.8). $\square$

LEMMA A.7. *Suppose ξ is a conditionally Poisson random measure on $S \times [0, \infty)$ with Cox measure $\Xi \times \Lambda$. If $\Xi(S) < \infty$ a.s., then we can write $\xi = \sum_{i=1}^{\infty} \delta_{(X_i, U_i)}$ with $U_1 < U_2 < \cdots$ a.s. and $\{X_i\}$ exchangeable.*

PROOF. Let $\{\widetilde{X}_i\}$ be exchangeable with de Finetti measure $\frac{\Xi}{|\Xi|}$, and let Y be a unit Poisson process with jump times $\{S_i\}$ independent of of $\{\widetilde{X}_i\}$ and Ξ. Define $\widetilde{\xi} = \sum_{i=1}^{\infty} \delta_{(\widetilde{X}_i, |\Xi|^{-1} S_i)}$, and note that

$$
\begin{aligned}
E[e^{-\int f d\widetilde{\xi}}] &= E\left[\prod_i e^{-f(\widetilde{X}_i, |\Xi|^{-1} S_i)}\right] \\
&= E\left[\prod_i \int e^{-f(z, |\Xi|^{-1} S_i)} |\Xi|^{-1} \Xi(dz)\right] \\
&= E\left[\exp\left\{-\int_0^{\infty} \left(1 - \int e^{-f(z, |\Xi|^{-1} s)} |\Xi|^{-1} \Xi(dz)\right) ds\right\}\right] \\
&= E\left[\exp\left\{-\int_0^{\infty} \int (1 - e^{-f(z,u)}) \Xi(dz)\, du\right\}\right].
\end{aligned}
$$

Consequently, $\widetilde{\xi}$ is conditionally Poisson with Cox measure $\Xi \times \Lambda$, and ξ and $\widetilde{\xi}$ have the same distribution. □

As in Lemma A.4, we have the following.

LEMMA A.8. *Suppose ξ is a conditionally Poisson random measure on $S \times [0, \infty)^2$ with Cox measure $\Xi \times \Lambda$, where Ξ is a random measure on $S \times [0, \infty)$. Suppose*

$$
\widehat{\Xi}(A) = \int_{S \times [0,\infty)} \frac{1}{y} \mathbf{1}_A(x) \Xi(dx \times dy)
$$

defines a locally finite random measure on S. Then writing $\xi = \sum_i \delta_{(X_i, Y_i, U_i)}$, $\widehat{\xi} = \sum_i \delta_{(X_i, Y_i U_i)}$ is a conditionally Poisson random measure on $S \times [0, \infty)$ with Cox measure $\widehat{\Xi} \times \Lambda$, and hence

$$
\int_{S \times [0,\infty)} y^{-1} f(x) \Xi(dx \times dy) = \widehat{\Xi}(f) = \lim_{K \to \infty} \frac{1}{K} \int_{S \times [0,K]} f\, d\widehat{\xi} \qquad a.s.
$$

A.3. Convergence results. Let $\{h_k, k = 1, 2, \ldots\} \subset \overline{C}(S)$ satisfy $0 \le h_k \le 1$ and $\bigcup_k \{x : h_k(x) > 0\} = S$, where $\overline{C}(S)$ denotes the space of bounded continuous functions on S, and let $\mathcal{M}_{\{h_k\}}(S)$ be the collection of Borel measures on S satisfying $\int_S h_k\, d\nu < \infty$, for all k, topologized by the requirement that $\nu_n \to \nu$ if and only if $\int_S f h_k\, d\nu_n \to \int_S f h_k\, d\nu$ for all $f \in \overline{C}(S)$ and k; that is, the measures $d\nu_n^k = h_k\, d\nu_n$ converge weakly for each k. Similarly, let $\mathcal{M}_{\{h_k\}}(S \times [0, \infty))$ be the space of Borel measures on $S \times [0, \infty)$ satisfying $\int_{S\times[0,K]} h_k \, d\mu < \infty$ for all $k = 1, 2, \ldots$ and $K > 0$, topologized by the requirement that $\mu_n \to \mu$ if and only if

$$\int_{S\times[0,\infty)} f h_k \, d\mu_n \to \int_{S\times[0,\infty)} f h_k \, d\mu,$$

for all k and $f \in \overline{C}(S \times [0, \infty))$ such that the support of f is contained in $S \times [0, K]$ for some $K > 0$. Note that in both cases, $\mathcal{M}_{\{h_k\}}$ is metrizable. To simplify notation, let

$$\mathcal{C}_{\{h_k\}}(S) = \{f \in \overline{C}(S) : |f| \leq c h_k \text{ for some } c > 0 \text{ and } h_k\}.$$

Then convergence in $\mathcal{M}_{\{h_k\}}(S)$ is equivalent to convergence of $\int_S f \, d\nu_n$ for all $f \in \mathcal{C}_{\{h_k\}}(S)$.

THEOREM A.9. *Let $\{\xi^n\}$ be a sequence of conditionally Poisson random measures on $S \times [0, \infty)$ with Cox measures $\{\Xi^n \times \Lambda\}$. Then $\xi^n \Rightarrow \xi$ in $\mathcal{M}_{\{h_k\}}(S \times [0, \infty))$ if and only if $\Xi^n \Rightarrow \Xi$ in $\mathcal{M}_{\{h_k\}}(S)$. If the limit holds, then ξ is conditionally Poisson with Cox measure $\Xi \times \Lambda$.*

PROOF. Suppose $\xi^n \Rightarrow \xi$ in $\mathcal{M}_{\{h_k\}}(S \times [0, \infty))$. Then for each $f \in \overline{C}(S)$, $f \geq 0$, each k, and all but countably many K

$$\begin{aligned} E[e^{-\int_{S\times[0,K]} f h_k \, d\xi}] &= \lim_{n\to\infty} E[e^{-\int_{S\times[0,K]} f h_k \, d\xi^n}] \\ &= \lim_{n\to\infty} E\bigl[e^{-K \int_S h_k^{-1}(1-e^{-f h_k}) h_k \, d\Xi^n}\bigr]. \end{aligned}$$

For $g \geq 0$ and K satisfying $\sup_x K^{-1} g(x) h_k(x) < 1$, let

$$f(x) = \begin{cases} -h_k^{-1}(x) \log\bigl(1 - K^{-1} g(x) h_k(x)\bigr), & h_k(x) > 0, \\ K^{-1} g(x), & h_k(x) = 0, \end{cases}$$

and we see that

$$\lim_{n\to\infty} E[e^{-\int_S g h_k \, d\Xi^n}] = E[e^{-\int_{S\times[0,K]} f h_k \, d\xi}]$$

exists. Since $\xi^n \Rightarrow \xi$ in $\mathcal{M}_{\{h_k\}}(S \times [0, \infty))$, $\{\int_{S\times[0,K]} h_k \, d\xi^n\}$ is stochastically bounded and by Lemma A.5, $\{\int h_k \, d\Xi^n\}$ must be stochastically bounded. Tightness follows similarly. Consequently, $\{\Xi^n\}$ is relatively compact in $\mathcal{M}_{\{h_k\}}(S)$ in distribution, and the unique limit Ξ is determined by the fact that

$$E[e^{-\int_S g h_k \, d\Xi}] = E[e^{-\int_{S\times[0,K]} f h_k \, d\xi}],$$

for g and f related as above. The proof of the converse is similar. $\square$

THEOREM A.10. *For each $n = 1, 2, \dots,$ let $r_n > 0$ and ξ^n be a point process on $S \times [0, r_n]$, and define*

$$\Xi^n(dx) = \frac{1}{r_n}\xi^n(dx \times [0, r_n]). \tag{A.9}$$

Suppose for $f \geq 0$, $E[e^{-\int f(x,u)\xi^n(dx\times du)}] = E[e^{-\int F_f^n(x)\Xi^n(dx)}]$, where

$$F_f^n(x) = -r_n \log \frac{1}{r_n}\int_0^{r_n} e^{-f(x,u)}\,du = -r_n \log\biggl(1 - \frac{1}{r_n}\int_0^{r_n}(1 - e^{-f(x,u)})\,du\biggr),$$

that is, the $[0, r_n]$ components are independent, uniformly distributed, and independent of Ξ^n. Then assuming $r_n \to \infty$, $\xi^n \Rightarrow \xi$ in $\mathcal{M}_{\{h_k\}}(S \times [0, \infty))$ if and only if $\Xi^n \Rightarrow \Xi$ in $\mathcal{M}_{\{h_k\}}(S)$. If the limit holds, then ξ is conditionally Poisson with Cox measure $\Xi \times \Lambda$.

PROOF. For $g_0, f_0 \geq 0$, $g_0 \in C_c([0, \infty))$, $f_0 \in \overline{C}(S)$ and $f(x, u) = h_k(x) \times f_0(x)g_0(u)$, $F_f^n(x) \to \int_0^\infty (1 - e^{-f(x,u)})\,du$, and the remainder of the proof is similar to that of Theorem A.9. $\square$

These convergence theorems apply only to the one-dimensional distributions of the models considered in this paper. To address convergence as processes, note that for finite r and $\Xi_r(t, dx) = r^{-1}\xi(t, dx \times [0, r])$, the models satisfy

$$E\bigl[e^{-\int_{S\times[0,r]} f(x,u)\xi(t,dx\times du)}|\mathcal{F}_t^{\Xi_r}\bigr] = e^{-\int F_f^r(x)\Xi_r(t,dx)}, \tag{A.10}$$

where

$$F_f^r(x) = -r \log \frac{1}{r}\int_0^r e^{-f(x,u)}\,du = -r \log\biggl(1 - \frac{1}{r}\int_0^r (1 - e^{-f(x,u)})\,du\biggr),$$

and the $r = \infty$ models satisfy

$$E\bigl[e^{-\int_{S\times[0,K]} f\,d\xi(t)}|\mathcal{F}_t^{\Xi}\bigr] = e^{-K\int_S (1-e^{-f})\,d\Xi(t)}, \tag{A.11}$$

for $f \in \mathcal{C}_{\{h_k\}}(S)$. The following estimates imply that convergence of the finite-dimensional distributions for ξ^n imply convergence of the finite-dimensional distributions for Ξ^n (or $\Xi_{r_n}^n$, assuming $r_n \to \infty$); however, convergence of the finite-dimensional distributions of Ξ^n may not imply convergence of the finite-dimensional distributions of of ξ^n.

LEMMA A.11. *Suppose ξ is a conditionally Poisson random measure on $S \times [0, \infty)$ with Cox measure $\Xi \times \Lambda$, Ξ with values in $\mathcal{M}_{\{h_k\}}(S)$. Then for each $f \in$*

$\mathcal{C}_{\{h_k\}}(S)$ *and* $\delta, K, K' > 0$,

$$
\begin{aligned}
&P\left\{\left|K^{-1}\int_{S\times[0,K]} f\,d\xi - \int_S f\,d\Xi\right| \geq \delta\right\} \\
&\quad \leq \frac{C}{K\delta^2} + P\left\{\int f^2\,d\Xi > C\right\} \qquad \text{(A.12)}\\
&\quad \leq \frac{C}{K\delta^2} + \frac{E[1-e^{-C^{-1}\int_{S\times[0,K']} f^2\,d\xi}]}{1-e^{-K'e^{-C^{-1}}}}.
\end{aligned}
$$

Suppose ξ *satisfies* (A.10). *Then for each* $f \in \mathcal{C}_{\{h_k\}}(S)$, $\delta > 0$ *and* $0 < K$, $K' < r$

$$
\begin{aligned}
&P\left\{\left|K^{-1}\int_{S\times[0,K]} f\,d\xi - \int_S f\,d\Xi_r\right| \geq \delta\right\} \\
&\quad \leq \frac{(r-K)C}{rK\delta^2} + P\left\{\int f^2\,d\Xi_r > C\right\} \qquad \text{(A.13)}\\
&\quad \leq \frac{(r-K)C}{rK\delta^2} + \frac{E[1-e^{-C^{-1}\int_{S\times[0,K']} f^2\,d\xi}]}{1-e^{-K'e^{-C^{-1}}}}.
\end{aligned}
$$

PROOF. By (A.2) and the Chebyshev inequality,

$$
P\left\{\left|K^{-1}\int_{S\times[0,K]} f\,d\xi - \int_S f\,d\Xi\right| \geq \delta \middle| \Xi\right\} \leq \frac{\int f^2\,d\Xi}{K\delta^2} \wedge 1 \leq \frac{C}{K\delta^2} + \mathbf{1}_{\{\int f^2\,d\Xi > C\}},
$$

and taking expectations gives the first inequality in (A.12). The second inequality follows by (A.4).

Similarly, for the second part,

$$
\begin{aligned}
P\left\{\left|K^{-1}\int_{S\times[0,K]} f\,d\xi - \int_S f\,d\Xi_r\right| \geq \delta \middle| \Xi_r\right\} &\leq \frac{\int (1-K/r) f^2\,d\Xi_r}{K\delta^2} \wedge 1 \\
&\leq \frac{(r-K)C}{rK\delta^2} + \mathbf{1}_{\{\int f^2\,d\Xi_r > C\}},
\end{aligned}
$$

and taking expectations gives the first inequality in (A.13). The second inequality follows by (A.8). □

The estimates in Lemma A.11 allow verifying convergence of measure-valued processes satisfying (A.11) or (A.10) by verifying convergence of the corresponding particle representations.

THEOREM A.12. *Let* $\{\xi^n\}$ *be a sequence of cadlag* $\mathcal{M}_{\{h_k\}}(S\times[0,\infty))$*-valued processes satisfying* (A.11) *for cadlag* $\mathcal{M}_{\{h_k\}}(S)$*-valued processes* $\{\Xi^n\}$. *If the* *finite-dimensional distributions of ξ^n converge to the finite-dimensional distributions of ξ, then the finite-dimensional distributions of Ξ^n converge to the finite-dimensional distributions of Ξ satisfying*

$$E\big[e^{-\int_{S\times[0,K]} f\,d\xi(t)}|\mathcal{F}_t^{\Xi}\big]=e^{-K\int_S(1-e^{-f})\,d\Xi(t)}.$$

For $n=1,2,\ldots,$ let ξ^n be a cadlag $\mathcal{M}_{\{h_k\}}(S\times[0,r_n])$-valued process satisfying (A.10) for cadlag $\mathcal{M}_{\{h_k\}}(S)$-valued processes $\{\Xi_{r_n}^n\}$. If $r_n\to\infty$ and the finite-dimensional distributions of ξ^n converge to the finite-dimensional distributions of ξ, then the finite-dimensional distributions of $\Xi_{r_n}^n$ converge to the finite-dimensional distributions of Ξ satisfying

$$E\big[e^{-\int_{S\times[0,K]} f\,d\xi(t)}|\mathcal{F}_t^{\Xi}\big]=e^{-K\int_S(1-e^{-f})\,d\Xi(t)}.$$

PROOF. Convergence of the finite-dimensional distributions follows easily from the estimates in Lemma A.11. □

A.4. Martingale lemmas.

LEMMA A.13. *Let $\{\mathcal{F}_t\}$ and $\{\mathcal{G}_t\}$ be filtrations with $\mathcal{G}_t\subset\mathcal{F}_t$. Suppose that $E[|X(t)|+\int_0^t|Y(s)|\,ds]<\infty$ for each t, and*

$$M(t)=X(t)-\int_0^t Y(s)\,ds$$

is an $\{\mathcal{F}_t\}$-martingale. Then

$$\widehat{M}(t)=E[X(t)|\mathcal{G}_t]-\int_0^t E[Y(s)|\mathcal{G}_s]\,ds$$

is a $\{\mathcal{G}_t\}$-martingale.

PROOF. Let $D\in\mathcal{G}_t\subset\mathcal{F}_t$. Then

$$\begin{aligned}
&E\big[\big(\widehat{M}(t+r)-\widehat{M}(t)\big)\mathbf{1}_D\big]\\
&\quad=E\bigg[\bigg(E[X(t+r)|\mathcal{G}_{t+r}]-E[X(t)|\mathcal{G}_t]-\int_t^{t+r}E[Y(s)|\mathcal{G}_s]\,ds\bigg)\mathbf{1}_D\bigg]\\
&\quad=E\bigg[\bigg(X(t+r)-X(t)-\int_t^{t+r}Y(s)\,ds\bigg)\mathbf{1}_D\bigg]\\
&\quad=0,
\end{aligned}$$

giving the martingale property. □

LEMMA A.14. *Let $\{\mathcal{F}_n\}$ be an increasing sequence of σ-algebras and $\{X_n\}$ a sequence of random variables satisfying $E[\sup_n|X_n|]<\infty$ and $\lim_{n\to\infty}X_n=X$ a.s. Then*

$$\lim_{n\to\infty}E[X_n|\mathcal{F}_n]=E\Big[X\Big|\bigvee_n\mathcal{F}_n\Big].$$

PROOF. By the martingale convergence theorem, we have

$$E\Big[\inf_{k\geq m} X_k \Big| \bigvee_n \mathcal{F}_n\Big] \leq \liminf_{n\to\infty} E[X_n|\mathcal{F}_n] \leq \limsup_{n\to\infty} E[X_n|\mathcal{F}_n] \leq E\Big[\sup_{k\geq m} X_k \Big| \bigvee_n \mathcal{F}_n\Big],$$

and the result follows by letting $m \to \infty$. □

A.5. Markov mapping theorem. The following theorem (extending Corollary 3.5 from [30]) plays an essential role in justifying the particle representations and can also be used to prove uniqueness for the corresponding measure-valued processes. Let (S, d) and (S_0, d_0) be complete, separable metric spaces, $B(S) \subset M(S)$ be the Banach space of bounded measurable functions on S, with $\|f\| = \sup_{x\in S}|f(x)|$ and $\overline{C}(S) \subset B(S)$ be the subspace of bounded continuous functions. An operator $A \subset B(S) \times B(S)$ is *dissipative* if $\|f_1 - f_2 - \varepsilon(g_1 - g_2)\| \geq \|f_1 - f_2\|$ for all $(f_1, g_1), (f_2, g_2) \in A$ and $\varepsilon > 0$; A is a *pre-generator* if A is dissipative and there are sequences of functions $\mu_n : S \to \mathcal{P}(S)$ and $\lambda_n : S \to [0,\infty)$ such that for each $(f, g) \in A$

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_S \big(f(y) - f(x)\big)\mu_n(x, dy) \tag{A.14}$$

for each $x \in S$. A is *graph separable* if there exists a countable subset $\{g_k\} \subset \mathcal{D}(A) \cap \overline{C}(S)$ such that the graph of A is contained in the bounded, pointwise closure of the linear span of $\{(g_k, Ag_k)\}$. [More precisely, we should say that there exists $\{(g_k, h_k)\} \subset A \cap \overline{C}(S) \times B(S)$ such that A is contained in the bounded pointwise closure of $\{(g_k, h_k)\}$, but typically A is single-valued, so we use the more intuitive notation Ag_k.] These two conditions are satisfied by essentially all operators A that might reasonably be thought to be generators of Markov processes. Note that A is graph separable if $A \subset L \times L$, where $L \subset B(S)$ is separable in the sup norm topology, for example, if S is locally compact, and L is the space of continuous functions vanishing at infinity.

A collection of functions $D \subset \overline{C}(S)$ is *separating* if $\nu, \mu \in \mathcal{P}(S)$ and $\int_S f\,d\nu = \int_S f\,d\mu$ for all $f \in D$ imply $\mu = \nu$.

For an S_0-valued, measurable process Y, $\widehat{\mathcal{F}}_t^Y$ will denote the completion of the σ-algebra $\sigma(Y(0), \int_0^r h(Y(s))\,ds, r \leq t, h \in B(S_0))$. For almost every t, $Y(t)$ will be $\widehat{\mathcal{F}}_t^Y$-measurable, but in general, $\widehat{\mathcal{F}}_t^Y$ does not contain $\mathcal{F}_t^Y = \sigma(Y(s) : s \leq t)$. Let $\mathbf{T}^Y = \{t : Y(t) \text{ is } \widehat{\mathcal{F}}_t^Y \text{ measurable}\}$. If Y is cadlag and has no fixed points of discontinuity [i.e., for every t, $Y(t) = Y(t-)$ a.s.], then $\mathbf{T}^Y = [0,\infty)$. $D_S[0,\infty)$ denotes the space of cadlag, S-valued functions with the Skorohod topology, and $M_S[0,\infty)$ denotes the space of Borel measurable functions, $x : [0,\infty) \to S$, topologized by convergence in Lebesgue measure.

THEOREM A.15. *Let (S, d) and (S_0, d_0) be complete, separable metric spaces. Let $A \subset \overline{C}(S) \times C(S)$ and $\psi \in C(S)$, $\psi \geq 1$. Suppose that for each $f \in \mathcal{D}(A)$ there exists $c_f > 0$ such that*

$$|Af(x)| \leq c_f \psi(x), \qquad x \in A, \tag{A.15}$$

and define $A_0 f(x) = Af(x)/\psi(x)$.

Suppose that A_0 *is a graph-separable pre-generator, and suppose that* $\mathcal{D}(A) = \mathcal{D}(A_0)$ *is closed under multiplication and is separating. Let* $\gamma : S \to S_0$ *be Borel measurable, and let* α *be a transition function from* S_0 *into* S [$y \in S_0 \to \alpha(y, \cdot) \in \mathcal{P}(S)$ *is Borel measurable*] *satisfying* $\int h \circ \gamma(z)\alpha(y, dz) = h(y)$, $y \in S_0$, $h \in B(S_0)$, *that is,* $\alpha(y, \gamma^{-1}(y)) = 1$. *Assume that* $\widetilde{\psi}(y) \equiv \int_S \psi(z)\alpha(y, dz) < \infty$ *for each* $y \in S_0$, *and define*

$$C = \left\{\left(\int_S f(z)\alpha(\cdot, dz), \int_S Af(z)\alpha(\cdot, dz)\right) : f \in \mathcal{D}(A)\right\}.$$

Let $\mu_0 \in \mathcal{P}(S_0)$, *and define* $\nu_0 = \int \alpha(y, \cdot)\mu_0(dy)$.

(a) *If* $\widetilde{Y}$ *satisfies* $\int_0^t E[\widetilde{\psi}(\widetilde{Y}(s))]\,ds < \infty$ *for all* $t \geq 0$, *and* $\widetilde{Y}$ *is a solution of the martingale problem for* (C, μ_0), *then there exists a solution* X *of the martingale problem for* (A, ν_0) *such that* $\widetilde{Y}$ *has the same distribution on* $M_{S_0}[0, \infty)$ *as* $Y = \gamma \circ X$. *If* Y *and* $\widetilde{Y}$ *are cadlag, then* Y *and* $\widetilde{Y}$ *have the same distribution on* $D_{S_0}[0, \infty)$.

(b) *For* $t \in \mathbf{T}^Y$,

$$P\{X(t) \in \Gamma | \widehat{\mathcal{F}}_t^Y\} = \alpha(Y(t), \Gamma), \qquad \Gamma \in \mathcal{B}(S). \tag{A.16}$$

(c) *If, in addition, uniqueness holds for the martingale problem for* (A, ν_0), *then uniqueness holds for the* $M_{S_0}[0, \infty)$*-martingale problem for* (C, μ_0). *If* $\widetilde{Y}$ *has sample paths in* $D_{S_0}[0, \infty)$, *then uniqueness holds for the* $D_{S_0}[0, \infty)$*-martingale problem for* (C, μ_0).

(d) *If uniqueness holds for the martingale problem for* (A, ν_0), *then* Y *restricted to* $\mathbf{T}^Y$ *is a Markov process.*

REMARK A.16. Theorem A.15 can be extended to cover a large class of generators whose range contains discontinuous functions. (See [30], Corollary 3.5 and Theorem 2.7.) In particular, suppose $A_1, \ldots, A_m$ satisfy the conditions of Theorem A.15 for a common domain $\mathcal{D} = \mathcal{D}(A_1) = \cdots = \mathcal{D}(A_m)$, and $\beta_1, \ldots, \beta_m$ are nonnegative functions in $M(S)$. Then the conclusions of Theorem A.15 hold for

$$Af = \beta_1 A_1 f + \cdots + \beta_m A_m f.$$

PROOF OF THEOREM A.15. Theorem 3.2 of [30] can be extended to operators satisfying (A.15) by applying Corollary 1.12 of [34] (with the operator B in that corollary set equal zero) in place of Theorem 2.6 of [30]. Alternatively, see Corollary 3.3 of [32]. $\square$

A.6. Uniqueness for martingale problems. Assume that $B \subset \overline{C}(E) \times \overline{C}(E)$, that $\mathcal{D}(B)$ is closed under multiplication and is separating, and that existence and uniqueness hold for the $D_E[0, \infty)$ martingale problem for (B, ν) for each initial distribution $\nu \in \mathcal{P}(E)$. Without loss of generality, we can assume $g \in \mathcal{D}(B)$ satisfies $0 \le g \le 1$.

By Theorem 4.10.1 of [9], existence and uniqueness then follows for the n-particle motion martingale problem with generator

$$B^n = \left\{ \left(f(x), f(x) \sum_{i=1}^n \frac{Bg(x_i)}{g(x_i)} \right) : f(x) = \prod_{i=1}^n g(x_i) \right\}. \tag{A.17}$$

Actually, the cited theorem implies uniqueness for the ordered n-particle motion with generator

$$\widetilde{B}^n = \left\{ \left(f(x), f(x) \sum_{i=1}^n \frac{Bg_i(x_i)}{g_i(x_i)} \right) : f(x) = \prod_{i=1}^n g_i(x_i), g_i \in \mathcal{D}(B) \right\},$$

but Theorem A.15 can be applied to obtain uniqueness for B^n from uniqueness for $\widetilde{B}^n$. Define $\gamma(x) = \sum_{i=1}^n \delta_{x_i}$ and let $\alpha(y, \cdot)$ average over all permutations of the x_i in $y = \sum_{i=1}^n \delta_{x_i}$.

Now consider a generator for a process with state space

$$S = \bigcup_n \left\{ \sum_{i=1}^n \delta_{x_i} : x_i \in E \right\},$$

where we allow $n = 0$, that is, no particles exist.

For $f(x, n) = \prod_{i=1}^n g(x_i)$, let

$$\begin{aligned} Af(x,n) &= B^n f(x,n) \\ &\quad + f(x,n) \sum_k \lambda_k(x) \int_{E^k} \left(\prod_{i=1}^k g(z_i) - 1 \right) \eta_k(x, dz_1, \dots, dz_k) \\ &\quad + f(x,n) \sum_{(z_1,\dots,z_l) \subset \{x_1,\dots,x_n\}} \mu(x, z_1, \dots, z_l) \left(\frac{1}{\prod_{i=1}^l g(z_i)} - 1 \right), \end{aligned}$$

where $\lambda_k, \mu \ge 0$ and η_k is a transition function from S to E^k. The generator has the following simple interpretation: in between birth and death events the particles move independently with motion determined by B. At rate $\lambda_k(x)$, k new particles are created with locations in E determined by η_k. At rate $\mu(x, z_1, \dots, z_l)$, the particles at $z_1, \dots, z_l$ are removed.

Let

$$\beta(x) = \sum_k \lambda_k(x) + \sum_{(z_1,\dots,z_l) \subset \{x_1,\dots,x_n\}} \mu(x, z_1, \dots, z_l).$$

Then for each initial distribution ν_0 and each $m > 0$, a localization argument and Theorem 4.10.3 of [9] imply existence and uniqueness of the martingale problem for (A, ν_0) up to the first time the solution leaves $\{x : \beta(x) < m\}$. Consequently, existence and uniqueness hold provided that there is a solution X satisfying $\sup_{s \le t} \beta(X(s)) < \infty$ a.s. for each $t > 0$.

Essentially the same argument gives existence and uniqueness for generators of the form (3.1) provided $\inf_x(a(x)r - b(x)) > 0$ and there exists a solution satisfying $\sup_{s\le t}\sum a(X_i(s)) < \infty$ a.s. for each $t > 0$.

REFERENCES

[1] BERESTYCKI, J., BERESTYCKI, N. and LIMIC, V. The λ-coalescent speed of coming down from infinity. Preprint.

[2] BERTOIN, J. and LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math*. **50** 147–181 (electronic). MR2247827

[3] BHATTACHARYA, R. N. (1982). On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete* **60** 185–201. MR0663900

[4] DAWSON, D. A. (1975). Stochastic evolution equations and related measure processes. *J. Multivariate Anal*. **5** 1–52. MR0388539

[5] DONNELLY, P. and KURTZ, T. G. (1996). A countable representation of the Fleming–Viot measure-valued diffusion. *Ann. Probab.* **24** 698–742. MR1404525

[6] DONNELLY, P. and KURTZ, T. G. (1999). Genealogical processes for Fleming–Viot models with selection and recombination. *Ann. Appl. Probab.* **9** 1091–1148. MR1728556

[7] DONNELLY, P. and KURTZ, T. G. (1999). Particle representations for measure-valued population models. *Ann. Probab.* **27** 166–205. MR1681126

[8] ETHIER, S. N. and GRIFFITHS, R. C. (1993). The transition function of a measure-valued branching diffusion with immigration. In *Stochastic Processes* 71–79. Springer, New York. MR1427302

[9] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[10] EVANS, S. N. (1993). Two representations of a conditioned superprocess. *Proc. Roy. Soc. Edinburgh Sect. A* **123** 959–971. MR1249698

[11] EVANS, S. N. and O'CONNELL, N. (1994). Weighted occupation time for branching particle systems and a representation for the supercritical superprocess. *Canad. Math. Bull.* **37** 187–196. MR1275703

[12] EVANS, S. N. and PERKINS, E. (1990). Measure-valued Markov branching processes conditioned on nonextinction. *Israel J. Math.* **71** 329–337. MR1088825

[13] FITZSIMMONS, P. J. (1988). Construction and regularity of measure-valued Markov branching processes. *Israel J. Math.* **64** 337–361 (1989). MR0995575

[14] GOROSTIZA, L. G. and LÓPEZ-MIMBELA, J. A. (1990). The multitype measure branching process. *Adv. in Appl. Probab.* **22** 49–67. MR1039376

[15] GREY, D. R. (1988). Supercritical branching processes with density independent catastrophes. *Math. Proc. Cambridge Philos. Soc.* **104** 413–416. MR0948925

[16] GRIMVALL, A. (1974). On the convergence of sequences of branching processes. *Ann. Probab.* **2** 1027–1045. MR0362529

[17] HARRIS, T. E. (1951). Some mathematical models for branching processes. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950 305–328. Univ. of California Press, Berkeley and Los Angeles. MR0045331

[18] HELLAND, I. S. (1981). Minimal conditions for weak convergence to a diffusion process on the line. *Ann. Probab.* **9** 429–452. MR0614628

[19] HERING, H. (1978). The non-degenerate limit for supercritical branching diffusions. *Duke Math. J.* **45** 561–600. MR0507459

[20] HERING, H. and HOPPE, F. M. (1981). Critical branching diffusions: Proper normalization and conditioned limit. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **17** 251–274. MR0631242

[21] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1965). On branching Markov processes. *Proc. Japan Acad.* **41** 816–821. MR0202195

[22] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1968). Branching Markov processes. I. *J. Math. Kyoto Univ.* **8** 233–278. MR0232439

[23] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1968). Branching Markov processes. II. *J. Math. Kyoto Univ.* **8** 365–410. MR0238401

[24] IKEDA, N., NAGASAWA, M. and WATANABE, S. (1969). Branching Markov processes. III. *J. Math. Kyoto Univ.* **9** 95–160. MR0246376

[25] JOFFE, A. and MÉTIVIER, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.* **18** 20–65. MR0827331

[26] KEIDING, N. (1975). Extinction and exponential growth in random environments. *Theoret. Population Biol.* **8** 49–63.

[27] KULPERGER, R. (1979). Brillinger type mixing conditions for a simple branching diffusion process. *Stochastic Process. Appl.* **9** 55–66. MR0544715

[28] KURTZ, T. G. (1973). A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Funct. Anal.* **12** 55–67. MR0365224

[29] KURTZ, T. G. (1978). Diffusion approximations for branching processes. In *Branching Processes* (*Conf.*, *Saint Hippolyte*, *Que.*, 1976). *Adv. Probab. Related Topics* **5** 269–292. Dekker, New York. MR0517538

[30] KURTZ, T. G. (1998). Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.* **3** 29 pp. (electronic). MR1637085

[31] KURTZ, T. G. (2000). Particle representations for measure-valued population processes with spatially varying birth rates. In *Stochastic Models* (*Ottawa*, *ON*, 1998). *CMS Conf. Proc.* **26** 299–317. Amer. Math. Soc., Providence, RI. MR1765017

[32] KURTZ, T. G. and NAPPO, G. (2010). The filtered martingale problem. In *Handbook on Nonlinear Filtering* (D. Crisan and B. Rozovsky, eds.). Oxford Univ. Press. To appear.

[33] KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** 1035–1070. MR1112406

[34] KURTZ, T. G. and STOCKBRIDGE, R. H. (2001). Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.* **6** 52 pp. (electronic). MR1873294

[35] KURTZ, T. G. and XIONG, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* **83** 103–126. MR1705602

[36] LAMPERTI, J. and NEY, P. (1968). Conditioned branching processes and their limiting diffusions. *Teor. Verojatnost. i Primenen.* **13** 126–137. MR0228073

[37] LI, Z. H. (1992). Measure-valued branching processes with immigration. *Stochastic Process. Appl.* **43** 249–264. MR1191150

[38] LI, Z. H., LI, Z. B. and WANG, Z. K. (1993). Asymptotic behavior of the measure-valued branching process with immigration. *Sci. China Ser. A* **36** 769–777. MR1247000

[39] LI, Z. and WANG, Z. (1999). Measure-valued branching processes and immigration processes. *Adv. Math.* (*China*) **28** 105–134. MR1723025

[40] MELLEIN, B. (1982). Diffusion limits of conditioned critical Galton–Watson processes. *Rev. Colombiana Mat.* **16** 125–140. MR0685248

[41] PAKES, A. G. (1986). The Markov branching-catastrophe process. *Stochastic Process. Appl.* **23** 1–33. MR0866285

[42] PAKES, A. G. (1987). Limit theorems for the population size of a birth and death process allowing catastrophes. *J. Math. Biol.* **25** 307–325. MR0900324

[43] PAKES, A. G. (1988). The Markov branching process with density-independent catastrophes. I. Behaviour of extinction probabilities. *Math. Proc. Cambridge Philos. Soc.* **103** 351–366. MR0923688

[44] PAKES, A. G. (1989). Asymptotic results for the extinction time of Markov branching processes allowing emigration. I. Random walk decrements. *Adv. in Appl. Probab.* **21** 243–269. MR0997723

[45] PAKES, A. G. (1989). The Markov branching process with density-independent catastrophes. II. The subcritical and critical cases. *Math. Proc. Cambridge Philos. Soc.* **106** 369–383. MR1002548

[46] PAKES, A. G. (1990). The Markov branching process with density-independent catastrophes. III. The supercritical case. *Math. Proc. Cambridge Philos. Soc.* **107** 177–192. MR1021881

[47] SCHWEINSBERG, J. (2000). A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.* **5** 1–11 (electronic). MR1736720

[48] STANNAT, W. (2003). On transition semigroups of (A, Ψ)-superprocesses with immigration. *Ann. Probab.* **31** 1377–1412. MR1989437

[49] WATANABE, S. (1968). A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.* **8** 141–167. MR0237008

DEPARTMENTS OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF WISCONSIN, MADISON
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706-1388
USA
E-MAIL: kurtz@math.wisc.edu

INSTITUTO DE MATEMÁTICAS
UNAM
MÉXICO, DF 04510
MEXICO
E-MAIL: eliane@math.unam.mx
eliane@math.cinvestav.mx