The Annals of Applied Probability
2001, Vol. 11, No. 2, 488–543

SUPERPROCESSES OVER A STOCHASTIC FLOW

BY GEORGIOS SKOULAKIS AND ROBERT J. ADLER[1]

University of North Carolina and Technion

We study a specific particle system in which particles undergo random branching and spatial motion. Such systems are best described, mathematically, via measure valued stochastic processes. As is now quite standard, we study the so-called superprocess limit of such a system as both the number of particles in the system and the branching rate tend to infinity. What differentiates our system from the classical superprocess case, in which the particles move independently of each other, is that the motions of our particles are affected by the presence of a global stochastic flow. We establish weak convergence to the solution of a well-posed martingale problem. Using the particle picture formulation of the flow superprocess, we study some of its properties. We give formulas for its first two moments and consider two macroscopic quantities describing its average behavior, properties that have been studied in some detail previously in the pure flow situation, where branching was absent. Explicit formulas for these quantities are given and graphs are presented for a specific example of a linear flow of Ornstein–Uhlenbeck type.

1. Introduction. The initial motivation for the mathematics to follow came from the following modelling problem in biological oceanography.

Dinoflagellates are microscopic, single cell, phytoplankton leading a rather boring life. They have some limited ability for individual, small-scale motion, but spend most of their time being moved about, on a large scale, by oceanic tides. Every few hours they relieve their boredom by reproducing, unfortunately by simple cell division, dividing obliquely to form two cells of equal size.

They would be of little interest to anyone other than pure biologists were it not for the facts that "in bloom" they reach a density of some 10^8 cells per cubic litre and that blooms can cover areas of the order of square kilometres. Their reddish-brown color then becomes visible to an observer and the corresponding phenomenon is called a "red tide." Furthermore, certain species of dinoflagellates contain potent neurotoxins. The toxins enter the food chain via ingestion of the dinoflagellates by filter-feeding shellfish and fish, eventually, unless care is taken, finding their way to human consumption. The consequences of this can include "paralytic shellfish poisoning" (PSP) in which the victim becomes paralyzed and may die. About 1,000 cases of PSP have been reported in North America, with about a 25% death rate. The phenomenon is widespread in the world's oceans and because of the dramatic effects of PSP is of major financial concern to the world's fishing industries. (See, e.g., [18] for more details, although still at an elementary level.)

Received October 1999.
[1]Supported in part by the NSF and Israel Science Foundation.
AMS 2000 *subject classifications*. Primary 60G57, 60H15; secondary 60J80, 62F05.
Key words and phrases. Superprocesses, random environment, stochastic flows.

The traditional way to model the behavior of phytoplankton blooms is via a combination of (deterministic) differential equations for the total population growth, a diffusion equation to allow for spatial spread, with everything superimposed over a fluid dynamical system of partial differential equations to describe tides and other physical oceanographical phenomena. These models describe global behavior reasonably well, but fail to reproduce a visibly observable "patchiness" in the blooms.

This paper is, in essence, about developing a model which will capture this patchiness. The central thesis centers on modelling the structure of blooms in two parts. During a bloom, which lasts up to two months and so covers trillions of phytoplankton over about 500 lifetimes, total population size is reasonably stable and so individual phytoplankton reproduction will be considered to be that of critical, or close to critical, branching. The spatial motion of the plankton will be modelled via the combination of two independent sources. The first comes from individual phytoplankton following independent Brownian motions, or other diffusion processes. The second is an an overall population drift, due to the motion of oceanic currents, modelled as a stochastic flow. From over two decades of accumulated knowledge, we know that the numbers associated with blooms are in the "domain of attraction" of a superprocess, and the one clear thing we know about superprocesses is that they generate very "patchy" pictures.

While it seems reasonable that "billions of phytoplankton" are probably enough to invoke the strong law behavior behind convergence of particle systems to superprocess limits, it is not clear that "about 500 lifetimes" are enough. Nor is it true that $500 = O(\text{"billions"})$, which, as will be clear soon, is a necessary requirement for the limit theorems to work. Nevertheless, as anyone who has ever seen a simulation of a superprocess (e.g., [1]) can testify, the fractal, "patchy" behavior of superprocesses occurs a long time before the infinite density limits are reached.

As opposed to most papers on superprocesses, we shall not treat "local patchiness" of the kind measured by Hausdorff dimension of supports, etc., but will rather concentrate on large scale motion as measured by the mean motion of mass. It is this motion which is of most interest in the modelling scenario and is all that is really quantifiable there.

This is all we shall have to say about motivation. A first attempt at reaching the biological oceanography community with these ideas is in [2]. We hope more works will follow, but now turn to a more conventional particle picture for describing the process that we will work with.

Let Y be a Feller process (e.g., Brownian motion) taking values in $\mathbb{R}^d$ and n a positive integer. Assume that, at time zero, K_n particles are placed in $\mathbb{R}^d$. Each of the K_n particles follows the path of an independent copy of the process Y until time $1/n$. At time $1/n$ each particle, independently of the others, gives birth to a number (≥ 0) of offspring according to some distribution, which is common to all particles, and then dies. Typically the expected number of offspring equals one; that is, the branching is critical. However, we shall also treat asymptotically supercritical or subcritical branching.

After branching, the individual particles in the new population follow the path of an independent copy of Y, starting at their place of birth, which is the place of death of their parent. This occurs in the time interval $[1/n, 2/n)$ and the pattern of alternating branching and spatial spreading continues as time evolves, as long as there remain particles alive. The process of interest to us is the measure-valued Markov process

$$X_t^n(B) = \frac{\text{Number of particles in } B \text{ at time } t}{n},$$

where $B \in \mathscr{B} = \sigma$-algebra of Borel sets in $\mathbb{R}^d$. Note that, for fixed t and n, X_t^n is an atomic measure. It is then well known that, under mild conditions on the process Y and the branching distribution, if $\{X_0^n\}$ converges weakly to a finite measure then the sequence $\{X^n\}$ converges, as n tends to infinity, on the Skorokhod space of cadlag (right continuous with left limits) functions from $[0, \infty)$ to the space of finite measures on $(E, \mathscr{B}(E))$ endowed with the topology of weak convergence. The limiting process is known as the Y superprocess (or super Brownian motion if Y is a Brownian motion) and it is uniquely characterized as the solution of a well-posed martingale problem. (See, for example, the extensive review of Dawson [6] for details of both this and all other facts about superprocesses that we quote without explicit reference.)

Think of the above particle picture as describing the motion of many phytoplankton in a glass jar. We now need to empty the jar into the ocean and subject them all to a common, random, motion. We shall model this motion by a stochastic flow, the details of which are in the next section.

In this case, it is possible to characterize the limit process as the measure-valued solution of the following martingale problem:

For all $f \in \mathscr{D}$,

$$Z_t(f) = X_t(f) - \nu(f) - \int_0^t X_s(Lf)\,ds - \xi \int_0^t X_s(f)\,ds$$

(1.1) is a continuous square integrable

$\{\mathscr{F}_t^X\}$-martingale such that $Z_0(f) = 0$ and

$$\langle Z(f)\rangle_t = \delta \int_0^t X_s(f^2)\,ds + \int_0^t (X_s \times X_s)(\Lambda f)\,ds,$$

where $\mathscr{D}$ is an appropriate class of functions, L is the operator describing the motion of a single particle, Λ is an operator reflecting the flow and ξ and δ are nonnegative constants.

The exact formulation of the model, in a slightly more extended form than presented here, along with the definition of all related quantities, appear in Section 2. In the same section we also give the statement the weak convergence result, the proof of which is deferred to the Appendix. In Section 3 we derive formulas for the first two moments of the flow superprocess and consider two descriptive, macroscopic, quantities related to flows in the context of our model. As mentioned above, these are the most important functions in terms of experimental quantifiability. They have also been studied in detail in a similar setting, but without the branching that is so central to our model, in two papers by Zirbel [29, 30].

The formulas of Section 3 are rather abstract, so in Section 4 we look at a specific example where, with a fair amount of work, they can be simplified and explicitly computed so as to give a good idea of what happens in general.

Before beginning in earnest, one should note that there are at least two different ways to look at the model we present. One is as a superprocess in a random environment (the flow) and one is as an extension of models of the motion of mass by flows, when the "mass" has an additional (branching) noise factor added. In the first vein, precursors can be found in the papers of Wang [26, 27], who considered a similar setup, but with a somewhat different environment, described in more detail in Section 2.3.

In the second vein, the main precursors are the thesis of Finger [12] and the works of Zirbel and Çinlar [28]–[32], who also developed a number of moment formulas for the movement of mass, akin to those that we have in Sections 3 and 4, but for motion of mass by flows alone. We tend to see our results in the latter vein and as being of more interest there.

Before turning to details, one of our referees asked for a list of notation to ease some of the pain of following technicalities. Here it is.

$\mathscr{B}(E)$: σ -algebra of Borel sets in a metric space E.
$\mathscr{P}(E)$: the space of probability measures on $(E, \mathscr{B}(E))$.
$M_F(E)$: the space of finite measures on $(E, \mathscr{B}(E))$.
$M(E)$: the space of Borel measurable functions defined on E.
$B(E)$: the space of bounded Borel measurable functions defined on E.
$C_b(E)$: the space of bounded and continuous functions defined on E.
$C_K(E)$: the space of continuous functions with compact support defined on E.
$C_0(E)$: the space of continuous functions vanishing at infinity defined on E.
$C_l(E)$: the space of continuous functions with limits at infinity defined on E.
$C_K^2(E)$: the space of continuous functions with compact support having continuous second order partial derivatives defined on E.
$C_E[0, \infty)$: the space of continuous paths taking values in the space E.
$D_E[0, \infty)$: the Skorokhod space of *cadlag* paths taking values in the space E.
$\mu(f)$: the integral $\int f(x)\mu(dx)$ of the function f with respect to measure μ.
f^i: the first order partial derivative of the function $f(x_1, \dots, x_n)$ with respect to x_i,
$f^{i,j}$: Second order partial derivative of the function $f(x_1, \dots, x_n)$ with respect to x_i and x_j,
I_A: indicator function of the set A,
$\mathrm{card}\{A\}$: cardinality of set A,
$\Rightarrow$: weak convergence,
$\mathbb{N}$: the set of natural numbers $0, 1, \dots.$
$\mathbb{N}^*$: the set of positive natural numbers $1, 2, \dots.$

2. Description of the model and weak convergence.

2.1. *Description of the model.* In this subsection we give a detailed and precise description of the branching particle system superimposed on a stochastic flow that described in the introduction. We start with some notation and, since it is all quite standard, we shall be terse. Let $\overline{\mathbb{R}_d} = \mathbb{R}^d \cup \{\Delta\}$ be the one-point Alexandroff compactification of $\mathbb{R}^d$ where Δ is the point at infinity, which we will also use as cemetery point. We use $\lambda > 0$ to denote the intensity rate of the branching mechanism.

We need a family of multiindices,

$$I \equiv \{\alpha = (\alpha_0, \alpha_1, \dots, \alpha_N)\colon N \geq 0,\ \alpha_i \in \{1, 2, \dots\},\ 0 \leq i \leq N\},$$

and set $|\alpha| = |(\alpha_0, \alpha_1, \dots, \alpha_N)| = N$, $\alpha|_i = (\alpha_0, \dots, \alpha_i)$ and $\alpha - i = (\alpha_0, \dots, \alpha_{|\alpha|-i})$. Moreover, for $t \geq 0$, write $\alpha \sim_n t$ if and only if $|\alpha|/(\lambda n) \leq t < (1 + |\alpha|)/(\lambda n)$.

Let K_n be the number of particles alive at time zero, spatially distributed in $\mathbb{R}^d$ at time zero at the points $x_1^n, x_2^n, \dots, x_{K_n}^n$ and defining the deterministic initial atomic measure

$$\nu_n = \sum_{i=1}^{K_n} \delta_{x_i^n}. \tag{2.1}$$

For each $n \geq 1$, let $\{B^{\alpha, n}\colon \alpha_0 \leq K_n, |\alpha| = 0\}$ be a collection of independent $\mathbb{R}^d$-valued Brownian motions, stopped at time $t = (\lambda n)^{-1}$, such that $B^{\alpha, n}(0) = x_{\alpha_0}^n$ and define the "tree" of processes, recursively, as follows: for each $k \geq 1$ let $\{B^{\alpha, n}\colon \alpha_0 \leq K_n, |\alpha| = k\}$ be a collection of $\mathbb{R}^d$-valued Brownian motions, stopped at time $t = (|\alpha| + 1)(\lambda n)^{-1}$, which are conditionally independent given the σ-algebra $\sigma\{B^{\alpha, n}\colon \alpha_0 \leq K_n, |\alpha| < k\}$ and for which

$$B^{\alpha, n}(t) = B^{\alpha - 1, n}(t), \qquad t \leq |\alpha|(\lambda n)^{-1}. \tag{2.2}$$

To handle the branching, for $n \geq 1$ let $\{N^{\alpha, n}\colon \alpha_0 \leq K_n\}$ be a collection of i.i.d. copies of N_n, where N_n is an $\mathbb{N}$-valued random variable such that

$$EN_n = 1 + \frac{\gamma_n}{n} = \beta_n, \qquad \gamma_n \geq 0, \quad \gamma_n \to \gamma \text{ as } n \to \infty \tag{2.3}$$

and

$$\operatorname{Var}(N_n) = {\sigma_n}^2 \to \sigma^2 \quad \text{as } n \to \infty. \tag{2.4}$$

That is, we assume the branching to be asymptotically supercritical. We further assume that

$$EN_n^p \leq M \quad \text{for all } n = 1, 2, \dots \tag{2.5}$$

for a constant $p > 2$ and a positive constant M. For each $n = 1, 2, \dots$, the collections $\{B^{\alpha, n}\colon \alpha_0 \leq K_n\}$ and $\{N^{\alpha, n}\colon \alpha_0 \leq K_n\}$ are assumed to be independent.

We now turn to the final and novel last component of our model, that of the stochastic flow. Let e and b be mappings from $\mathbb{R}^d$ to $\mathbb{R}^d$ and c a mapping from $\mathbb{R}^d$ to the space of $d\times m$ matrices, satisfying the global Lipschitz condition

$$|e(x)-e(y)|+|b(x)-b(y)|+\|c(x)-c(y)\|\le K|x-y|, \qquad x,y\in\mathbb{R}^d, \tag{2.6}$$

and the linear growth condition,

$$|e(x)|+|b(x)|+\|c(x)\|\le K(1+|x|), \qquad x\in\mathbb{R}^d, \tag{2.7}$$

for some finite positive constant K. Assume that $t\to F^n_{s,t}(x)$ is the solution of the stochastic differential equation

$$dY(t)=b(Y(t))\,dt+c(Y(t))\,dW^n(t), \qquad Y(s)=x$$

for all $t\ge s$ and $x\in\mathbb{R}^d$, where W^n is a $\mathbb{R}^m$-valued Brownian motion, independent of the collections $\{B^{\alpha,n}\}$ and $\{N^{\alpha,n}\}$. This defines a unique Brownian flow of homeomorphisms from $\mathbb{R}^d$ to $\mathbb{R}^d$. We refer to [16] for more details.

Set $a_n=1/(\lambda n)$ and $k_n=k/(\lambda n)$. Then the tree of Brownian motions over the flow is given by the collection of processes $Y^{\alpha,n}$, defined as follows: let $\alpha\sim_n k_n$ for some $k\in\mathbb{N}$. First, in the time interval $[0,k_n+a_n]$, $Y^{\alpha,n}$ is defined as the solution of the following d-dimensional stochastic differential equation:

$$\begin{aligned} &dY_i(t)=b_i(Y(t))\,dt+e_i(Y(t))\,dB_i^{\alpha,n}(t)+\sum_{l=1}^m c_{il}(Y(t))\,dW_l^n(t),\\ &Y(0)=x^n_{\alpha_0}.\end{aligned} \tag{2.8}$$

The existence and strong uniqueness of the solution is ensured by the conditions imposed on b, c and e. Now set $Y_t^{\alpha,n}=Y^{\alpha,n}_{k_n+a_n}$ for $t>k_n+a_n$. Note here that from the construction of the family $\{B^{\alpha,n}\colon\alpha_0\le K_n\}$ it follows that $Y_t^{\alpha,n}=Y_t^{\alpha-1,n}$ for $0\le t\le k_n$ in the case $k\ge 1$.

In order to "prune" this tree, we need the stopping times $\tau^{\alpha,n}$, defined as follows: for each $\alpha\in I$, let

$$\tau^{\alpha,n}=\begin{cases}0, & \text{if } \alpha_0>K_n,\\ \min\left\{\dfrac{i+1}{\lambda n}\colon 0\le i\le|\alpha|, N^{\alpha|_i}=0\right\}, & \text{if this set is } \ne\varnothing \text{ and } \alpha_0\le K_n,\\ \dfrac{1+|\alpha|}{\lambda n}, & \text{otherwise.}\end{cases} \tag{2.9}$$

The "pruned" tree of processes, with branching accounted for, is now the collection of processes $X^{\alpha,n}$ defined by

$$X_t^{\alpha,n}=\begin{cases}Y_t^{\alpha,n}, & \text{if } t<\tau^{\alpha,n},\\ \Delta, & \text{if } t\ge\tau^{\alpha,n}.\end{cases} \tag{2.10}$$

The measure-valued process for the finite system of particles is

$$X_t^n(B)=\frac{\operatorname{card}\{\alpha\sim_n t\colon X^{\alpha,n}(t)\in B\}}{n}, \tag{2.11}$$

where $B\in\mathscr{B}(\mathbb{R}^d)$.

We also need to define the corresponding filtrations,

$$\mathscr{F}_t^n \equiv \sigma(B^{\alpha,n}, N^{\alpha,n}: |\alpha| < k) \bigvee \sigma(W_s^n: s \le t) \bigvee \sigma(B_s^{\alpha,n}: s \le t,\ |\alpha| = k),$$

for $t \in [k_n, k_n + a_n)$, $k = 0, 1, \dots$.

Next we define two semigroups of operators that will be needed to describe the limit of the X^n. Let W be an $\mathbb{R}^m$-valued Brownian motion, B^1, B^2 be $\mathbb{R}^d$-valued Brownian motions and assume that all three are mutually independent. Let (Y^1, Y^2) be the solution of the stochastic differential equation,

$$\begin{aligned} dY_i^1(t) &= b_i(Y^1(t))\,dt + \sum_{l=1}^m c_{il}(Y^1(t))\,dW_l(t) + e_i(Y^1(t))\,dB_i^1(t), \\ dY_i^2(t) &= b_i(Y^2(t))\,dt + \sum_{l=1}^m c_{il}(Y^2(t))\,dW_l(t) + e_i(Y^2(t))\,dB_i^2(t), \end{aligned} \tag{2.12}$$

$i = 1, 2, \dots, d$, $t \ge 0$. For $t \ge 0$, $f \in C_l(\mathbb{R}^d)$ and $h \in C_l(\mathbb{R}^d \times \mathbb{R}^d)$, set

$$T_t f(y) = E_y f(Y^1(t)) = E(f(Y^1(t)) | Y^1(0) = y) \tag{2.13}$$

and

$$\begin{aligned} S_t h(y^1, y^2) &= E_{y^1, y^2} h(Y^1(t), Y^2(t)) \\ &= E(h(Y^1(t),\ Y^2(t)) | Y^1(0) = y^1,\ Y^2(0) = y^2). \end{aligned} \tag{2.14}$$

It follows from our global Lipschitz assumption on b, c and e and the discussion in [22], Chapter V, Section 22, that $\{T_t: t \ge 0\}$ and $\{S_t: t \ge 0\}$ are strongly continuous contraction semigroups on $C_l(\mathbb{R}^d)$ and $C_l(\mathbb{R}^d \times \mathbb{R}^d)$, respectively. In the terminology used in [22], T_t and S_t are Feller–Dynkin semigroups.

We will need three more operators. Let L be the second-order differential operator defined for $f \in C^2(\mathbb{R}^d)$ by

$$(Lf)(x) = \sum_{i=1}^d b_i(x) f^i(x) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d d_{ij}(x) f^{ij}(x), \qquad x \in \mathbb{R}^d, \tag{2.15}$$

where

$$d_{ij}(x) = \delta_{ij} e_i(x) e_j(x) + a_{ij}^{(m)}(x, x), \qquad x \in \mathbb{R}^d, \quad i, j = 1, 2, \dots, d \tag{2.16}$$

and

$$a_{ij}^{(m)}(x, y) = \sum_{l=1}^m c_{il}(x) c_{jl}(y), \qquad x, y \in \mathbb{R}^d, \quad i, j = 1, 2, \dots, d. \tag{2.17}$$

In addition, for $f \in C^2(\mathbb{R}^d)$, set

$$(\Lambda f)(x, y) = \sum_{i=1}^d \sum_{j=1}^d a_{ij}^{(m)}(x, y) f^i(x) f^j(y), \qquad x, y \in \mathbb{R}^d \tag{2.18}$$

and

$$(\Psi f)(x) = (\Lambda f)(x, x) + \sum_{i=1}^{d} (e_i(x) f^i(x))^2, \qquad x \in \mathbb{R}^d. \tag{2.19}$$

Finally, define the class of functions $\mathscr{D}$,

$$\begin{aligned} \mathscr{D} = \{ f \in C_l^2(\mathbb{R}^d)\colon Lf, \Psi f, c_{il} f^i \in C_l(\mathbb{R}^d) \\ \forall\ i = 1, 2, \ldots, d, \forall\ l = 1, 2, \ldots, m\}. \end{aligned} \tag{2.20}$$

For all $f \in \mathscr{D}$, let B_f denote a common bound for f, its first- and second-order partial derivatives, Λf, Ψf and Lf such that $B_f \geq 1$. It is clear that $\mathscr{D}_K = \{ f + c\colon f \in C_K^2(\mathbb{R}^d),\ c \in \mathbb{R}\}$ is a subset of $\mathscr{D}$.

2.2. *Weak convergence.* We need to make some further assumptions about the functions b, c and e, necessary for establishing the uniqueness of the martingale problem satisfied by the weak limit points of the $\{X^n\}$.

ASSUMPTION U. *One of the following conditions is satisfied:*

(a) *For all $i = 1, \ldots, d$, $l = 1, \ldots, m$ the functions b_i, c_{il} and e_i have bounded and continuous first and second partial derivatives. Furthermore, for each $N \geq 1$, there exists $\lambda_N > 0$ such that*

$$\sum_{i,j=1}^{d} \sum_{p,q=1}^{N} \xi_i^p d_{ij}(x_p, x_q) \xi_j^q \geq \lambda_N \sum_{p=1}^{N} \sum_{i=1}^{d} (\xi_i^p)^2$$

for all $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$ and $(\xi_1^1, \xi_2^1, \ldots, \xi_d^1, \xi_1^2, \xi_2^2, \ldots, \xi_d^2, \ldots \xi_1^N,\ \xi_2^N, \ldots, \xi_d^N) \in \mathbb{R}^{dN}$ where

$$d_{ij}(x, y) = a_{ij}^{(m)}(x, y) + \delta_{ij} e_i(x) e_j(y), \qquad x, y \in \mathbb{R}^d, \quad i, j = 1, \ldots, d$$

and $a_{ij}^{(m)}$ is given by (2.17).

(b) *There exists a $d \times d$ matrix $\mathbf{A}$, a $d \times m$ matrix $\mathbf{C}$, and column vectors a, ε in $\mathbb{R}^d$ such that $b(x) = \mathbf{A}x + a$, $(c_{ij}(x)) = \mathbf{C}$, and $(e_i(x)) = \varepsilon$. Furthermore the matrix $[\mathbf{E}, \mathbf{AE}, \ldots, \mathbf{A}^{d-1}\mathbf{E}]$ has rank d, where $\mathbf{E} = \mathrm{diag}(\varepsilon)$.*

The case described in (a) is referred to as the uniformly elliptic case, while the case described in (b) is referred to as the linear case. We note that under either (a) or (b) the diffusion describing the joint motion of N particles in our system has a transition density.

The weak convergence result, the proof of which we leave to the Appendix, can then be stated as follows.

THEOREM 2.2.1. *Assume that $X_0^n \Rightarrow \nu$ in $M_F(\mathbb{R}^d)$. Then, under the assumption* U, *the sequence $\{X^n\}$ converges weakly to X, where $X \in C_{M_F(\mathbb{R}^d)}[0, \infty)$ is the unique solution of the following martingale problem:*

For all $f \in \mathscr{D}$,

$$
\begin{aligned}
&Z_t(f) = X_t(f) - \nu(f) - \int_0^t X_s(Lf)\,ds - \xi \int_0^t X_s(f)\,ds \\
&\text{is a continuous square integrable} \\
&\{\mathscr{F}_t^X\}\text{-martingale such that } Z_0(f) = 0 \quad \text{and} \\
&\langle Z(f)\rangle_t = \delta \int_0^t X_s(f^2)\,ds + \int_0^t (X_s \times X_s)(\Lambda f)\,ds,
\end{aligned}
\tag{2.21}
$$

where $\xi = \lambda\gamma$ and $\delta = \lambda\sigma^2$.

2.3. *A related model.* In [26, 27], Wang independently considered a model closely related to ours. In his model, the first line of (2.8) is replaced by

$$
dY_i(t) = \int_{\mathbb{R}} g(x, Y_i(t))W(dx, dt) + dB_i^{\alpha, n}(t), \qquad i = 1, 2, \dots, d,
$$

where g is a smooth function and W is a space time white noise. This equation is somewhat simpler than ours in that the drifts b_i and diffusion coefficients e_i are missing. However, the novelty of our model, which lies primarily in the random environment provided by the stochastic flow, is also present here.

Not surprisingly, the tightness arguments for Wang's system and ours are similar. There is a significant difference when treating the uniqueness of the limit process, however, since his dual (cf. Section A.4) is simpler than ours.

What is quite different in the present analysis is the detailed study of the evolution of the moment structure of the limit process, which is made somewhat more transparent and interesting because the underlying environment is generated by a stochastic flow.

3. Moment formulas, mean and spatial covariance measures. Our main aim in this section is to derive moment formulas for certain functionals of the process we obtained in Theorem 2.2.1. The reader familiar with superprocess theory will know that, in that case, this is usually done by differentiation of the Laplace functional of the process. (cf. Theorem 1.1′ in [9] for general results.) In the current situation, however, this does not seem possible. We do not have an explicit form for the Laplace functional and, in view of the fact that, unlike in the basic superprocess case, we are not dealing with an infinitely divisible process, one would expect it to be rather complicated. (cf. the discussion in Section 5.)

Consequently, we shall adopt a technique used in Section 3 in [1], which computed moments for the limit process as a limit of moments for the particle picture. Whereas there it was primarily an expository tool, here it seems to be the only analytic tool available to us.

A referee pointed out that it may also be possible to compute moments via the dual process of Section A.4, something which we did not try, but which does not seem to be straightforward. Such an approach, however, would undoubtedly be more elegant than the one we have taken.

We shall need the following assumption about the particle motions, which, while not essential for the existence of moments, is intrinsic to the structure of the formulas that we shall obtain.

ASSUMPTION I. *For every* $a \in I$ *and* $n \geq 1$, *the* $\mathbb{R}^d$*-dimensional diffusion* $Y^{\alpha,n}$, *restricted to the interval* $[0, \tau^{\alpha,n})$, *has a transition density, denoted by* $p_t^1(x, y)$, *which is Lipschitz in* x, *uniformly in* y. *We refer to the collection* $\{Y^{\alpha,n}\}$ *as the one-point motions. Furthermore, for every* $n \geq 1$ *and every pair* (α, α') *such that* $\alpha, \alpha' \in I, \alpha \neq \alpha'$, *the* $\mathbb{R}^{2d}$*-dimensional diffusion* $(Y^{\alpha,n}, Y^{\alpha',n})$ *restricted to the interval* $[0, \min(\tau^{\alpha,n}, \tau^{\alpha',n}))$ *has a transition density, denoted by* $p_t^2(x_1, x_2; y_1, y_2)$. *We refer to the collection* $\{(Y^{a,n}, Y^{a',n})\}$ *as the two-point motions.*

This assumption is verified if the diffusions are, for example, either uniformly elliptic or linear of Ornstein–Uhlenbeck type (cf. Section 4).

3.1. *Moment formulas for the particle picture* We start with some formulas for the particle picture. These require the seemingly restrictive assumption that the branching be binary. However, since what we are really interested in is formulas for the superprocess limit, which depend only on the first two moments of the branching distribution, there is no loss of generality involved in this assumption.

ASSUMPTION II. *We assume that the branching is binary, so that the possible numbers of offspring are* 0 *and* 2, *with probabilities* $1/2 - \gamma_n/2n$ *and* $1/2 + \gamma_n/2n$, *respectively.*

Note that under this assumption we have $\sigma_n^2 = 1 - \gamma_n^2/n^2 \to 1$ as $n \to \infty$; that is, $\sigma^2 = 1$. Our arguments follow very closely those employed in Section 3 in [1]. The next proposition provides the first and second moment formulas for the particle picture.

PROPOSITION 3.1.1. *Let* $f_1, f_2 \in C_b(\mathbb{R}^d)$, $0 < t_1 \leq t_2$, $N_1 = [\lambda t_1 n]$, $N_2 = [\lambda t_2 n]$ *and* $n > [(\lambda(t_2 - t_1))^{-1}]$ *if* $t_1 < t_2$. *Then, under the set-up described in Section* 2 *and Assumptions* I *and* II *above, we have that*

$$E(X_{t_1}^n(f_1)) = (1 + \gamma_n/n)^{N_1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_{t_1}^1(x; y) f_1(y)\, dy\, X_0^n(dx) \tag{3.1}$$

and

$$
\begin{aligned}
&E(X^n_{t_1}(f_1)X^n_{t_2}(f_2))\\
&\quad= \left(1+\frac{\gamma_n}{n}\right)^{N_1+N_2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1}(x_1,x_2;y_1,y_2)\\
&\qquad\times\left(f_1(y_1)\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2;z)f_2(z)\,dz\right)dy_1\,dy_2\,X^n_0(dx_1)X^n_0(dx_2)\\
&\qquad-\frac{1}{n}\left(1+\frac{\gamma_n}{n}\right)^{N_1+N_2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1}(x,x;y_1,y_2)\\
&\qquad\times\left(f_1(y_1)\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2;z)f_2(z)\,dz\right)dy_1\,dy_2\,X^n_0(dx)\\
&\qquad+\frac{1}{n}\left(1+\frac{\gamma_n}{n}\right)^{N_2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^1_{t_1}(x;y_1)f_1(y_1)\\
&\qquad\times\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_1;y_2)f_2(y_2)\,dy_2\right)dy_1\,X^n_0(dx)\\
&\qquad+\frac{1}{n}\sum_{r=0}^{N_1-1}\left(1+\frac{\gamma_n}{n}\right)^{N_1+N_2-r-1}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^1_{r_n}(x_1;x_2)\\
&\qquad\times\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1-r_n}(x_2,x_2;y_1,y_2)f_1(y_1)\right.\\
&\qquad\times\left.\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2;z)f_2(z)\,dz\right)dy_1\,dy_2\right)dx_2\,X^n_0(dx_1)
\end{aligned}\tag{3.2}
$$

with the convention $\int_{\mathbb{R}^d} p^1_0(x;y)h(y)\,dy = h(x)$.

PROOF. We shall only give a full proof of (3.2) for the case $t_1 < t_2$. The proof of (3.2) for the case $t_1 = t_2$, as well as the proof of (3.1), follows using very similar, and actually simpler, arguments. Note that we take n large enough so that branching occurs, at least once, between the times t_1 and t_2. Let $I_{\alpha,n}(t)$ denote the indicator of the event that the particle labelled by α is alive at time t. We need to calculate

$$
\begin{aligned}
&E(X^n_t(f_1)X^n_t(f_2))\\
&\quad=\frac{1}{n^2}\sum_{\alpha\sim_n t_1}\sum_{\alpha'\sim_n t_2} E\{f_1(Y^{\alpha,n}_{t_1})f_2(Y^{\alpha',n}_{t_2})\}E\{I_{\alpha,n}(t_1)I_{\alpha',n}(t_2)\}.
\end{aligned}\tag{3.3}
$$

Let $\alpha\sim_n t_1$ and $\alpha'\sim_n t_2$. Then there are three cases one should consider:

(i) $\alpha_0 \neq \alpha'_0$. In this case the particles labelled by α and α' are the descendants of different ancestors in the initial generation.

(ii) $\alpha = \alpha' - (N_2 - N_1)$. In this case the particle labelled by α is a direct ancestor of the particle labelled by α'.

(iii) $\alpha - (N_1 - r) = \alpha' - (N_2 - r)$ for some $1 \le r \le N_1 - 1$. In this case the particles labelled by α and α' have a common ancestor, but the particle labelled by α is not a direct ancestor of the particle labelled by α'.

Note that, in what follows, all labels β we consider are such that $\beta_0 \le K_n$, that is, the particles belong to the branches starting from the initial particles that were actually born. We proceed by evaluating the contribution to (3.3) of the terms covered by each of the three cases just described.

Case (i). In this case, the particles labelled by α and α' live on separate trees. Hence,

$$E\{I_{\alpha,n}(t_1)I_{\alpha',n}(t_2)\} = EI_{\alpha,n}(t_1)EI_{\alpha',n}(t_2) = \left(\frac{1}{2} + \frac{\gamma_n}{2n}\right)^{N_1}\left(\frac{1}{2} + \frac{\gamma_n}{2n}\right)^{N_2} = \left(\frac{1}{2} + \frac{\gamma_n}{2n}\right)^{N_1+N_2}.$$

Using the Markov property of the one-point and two-point motions, we obtain

$$\begin{aligned} E\{f_1(Y^{\alpha,n}_{t_1})f_2(Y^{\alpha',n}_{t_2})\} &= E\left\{f_1(Y^{\alpha,n}_{t_1})E\left(f_2(Y^{\alpha',n}_{t_2})|\mathscr{F}^n_{t_1}\right)\right\} \\ &= E\left\{f_1(Y^{\alpha,n}_{t_1})\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(Y^{\alpha',n}_{t_1}; z)f_2(z)\,dz\right)\right\} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1}(x^n_{\alpha_0}, x^n_{\alpha'_0}; y_1, y_2)f_1(y_1) \\ &\qquad \times \left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2; z)f_2(z)\,dz\right)dy_1\,dy_2. \end{aligned}$$

Next, note that there are $2^{N_1} \times 2^{N_2}$ possible pairs (α, α') corresponding to each pair of initial ancestors. We are thus able to write the contribution to (3.3) of the terms covered by the first case as

$$\frac{1}{n^2}2^{N_1+N_2}\sum_{\substack{\alpha_0,\alpha'_0=1\\ \alpha_0\ne\alpha'_0}}\left(\frac{1}{2} + \frac{\gamma_n}{2n}\right)^{N_1+N_2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1}(x^n_{\alpha_0}, x^n_{\alpha'_0}; y_1, y_2)f_1(y_1) \times \left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2; z)f_2(z)\,dz\right)dy_1\,dy_2.$$

But this, on rearrangement, gives us the two first terms in the right-hand side of (3.2).

Case (ii). Since, in this second case, the particle labelled by α is a direct ancestor of the particle labelled by α' we have that $Y^{\alpha,n}_t = Y^{\alpha',n}_t$ for all $t \le t_1$.

Hence,

$$\begin{aligned} E\{f_1(Y_{t_1}^{\alpha,n})f_2(Y_{t_2}^{\alpha',n})\} &= E\{f_1(Y_{t_1}^{\alpha,n})E(f_2(Y_{t_2}^{\alpha',n})|\mathscr{F}_{t_1}^n)\} \\ &= E\biggl\{f_1(Y_{t_1}^{\alpha,n})\biggl(\int_{\mathbb{R}^d} p_{t_2-t_1}^1(Y_{t_1}^{\alpha,n};y_2)f_2(y_2)\,dy_2\biggr)\biggr\} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_{t_1}^1(x_{\alpha_0}^n;y_1)f_1(y_1) \\ &\qquad \times\biggl(\int_{\mathbb{R}^d} p_{t_2-t_1}^1(y_1;y_2)f_2(y_2)\,dy_2\biggr)dy_1. \end{aligned}$$

Moreover,

$$\begin{aligned} E\{I_{\alpha,n}(t_1)I_{\alpha',n}(t_2)\} &= E\{I_{\alpha,n}(t_1)E(I_{\alpha',n}(t_2)|\mathscr{F}_{t_1}^n)\} \\ &= E\biggl(I_{\alpha,n}(t_1)I_{\alpha,n}(t_1)\biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_2-N_1}\biggr) \\ &= \biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_1}\biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_2-N_1} = \biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_2}. \end{aligned}$$

The number of possible pairs (α, α') corresponding to each (common) initial ancestor is $2^{N_1} \times 2^{N_2-N_1}$. It is clear now that the contribution to (3.3) of the terms covered by the second case gives the third term in the right-hand side of (3.2). Note that this case does not appear if $t_1 = t_2$. Instead, when $t_1 = t_2$, we have the possibility that $\alpha = \alpha'$ which can be treated in a very similar fashion.

Case (iii). This is the most complicated case and also the most typical of higher moment computations. In this case the particles labelled by α and α' have a common ancestor. Let β be the label of their last common ancestor and let $r = |\beta|$. Clearly r can assume any of the values $0, 1, \dots, N_1 - 1$. Following arguments similar to those used in the previous cases we obtain

$$\begin{aligned} E\{I_{\alpha,n}(t_1)I_{\alpha',n}(t_2)\} &= \biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{r}\biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)\biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_1-r-1}\biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_2-r-1} \\ &= \biggl(\frac{1}{2}+\frac{\gamma_n}{2n}\biggr)^{N_1+N_2-r-1} \end{aligned}$$

and

$$\begin{aligned} E\{f_1(Y_{t_1}^{\alpha,n})f_2(Y_{t_2}^{\alpha' n})\} &= E\{f_1(Y_{t_1}^{\alpha,n})E(f_2(Y_{t_2}^{\alpha',n})|\mathscr{F}_{t_1}^n)\} \\ &= E\biggl\{f_1(Y_{t_1}^{\alpha,n})\biggl(\int_{\mathbb{R}^d} p_{t_2-t_1}^1(Y_{t_1}^{\alpha',n};y_2)f_2(y_2)dy_2\biggr)\biggr\} \\ &= E\biggl\{E\biggl(f_1(Y_{t_1}^{\alpha,n})\int_{\mathbb{R}^d} p_{t_2-t_1}^1(Y_{t_1}^{\alpha',n};z)f_2(z)dz|\mathscr{F}_{r_n}^n\biggr)\biggr\} \end{aligned}$$

$$= E\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1-r_n}(Y^{\beta,n}_{r_n}, Y^{\beta,n}_{r_n}; y_1, y_2) f_1(y_1)$$
$$\times\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2;z) f_2(z)\,dz\right) dy_1\,dy_2$$
$$= \int_{\mathbb{R}^d} p^1_{t_2-t_1}(x^n_{\beta_0};x_2)\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1-r_n}(x_2,x_2;y_1,y_2) f_1(y_1)\right.$$
$$\left.\times\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2;z) f_2(z)\,dz\right) dy_1\,dy_2\right) dx_2.$$

Given β_0 and r there are $2^r \times 2 \times 2^{N_1-r-1} \times 2^{N_2-r-1}$ possible corresponding pairs (α, α'). Summing over all possible values of β_0 and r immediately yields the last term in the right-hand side of (3.2) and therefore completes the proof. □

3.2. *Moment formulas for the flow superprocess.* In this subsection we let X be a weak limit point of X^n, where X^n is defined by (2.11). Under the further Assumptions I and II we derive first- and second-order moment formulas for X. We therefore assume, as in Theorem 2.2.1, that X^n_0 converges weakly to an initial measure ν. Then X is a solution to the martingale problem (2.21) with $\sigma^2 = 1$.

The next proposition is perhaps the most important result of the paper, and provides first- and second-moment formulas for X by passing to the limit in (3.1) and (3.2).

PROPOSITION 3.2.1. *Let $f \in C^2_K(\mathbb{R}^d)$ and $t > 0$. Assume $\sigma^2 = 1$ and let Assumption* I *be in effect. Then*

$$E(X_t(f)) = \exp(\lambda\gamma t)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^1_t(x, y) f(y)\,dy\,\nu(dx) \tag{3.4}$$

and

$$\begin{aligned}
&E(X_{t_1}(f_1)X_{t_2}(f_2))\\
&\quad= \exp(\lambda\gamma(t_1+t_2))\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1}(x_1, x_2; y_1, y_2)\\
&\qquad\times f_1(y_1)\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2; z) f_2(z)\,dz\right) dy_1\,dy_2\,\nu(dx_1)\nu(dx_2)\\
&\qquad+ \lambda\exp(\lambda\gamma(t_1+t_2))\int_0^{t_1}\exp(-\lambda\gamma s)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^1_s(x_1; x_2)\\
&\qquad\times\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p^2_{t_1-s}(x_2, x_2; y_1, y_2)\right)\\
&\qquad\times f_1(y_1)\left(\int_{\mathbb{R}^d} p^1_{t_2-t_1}(y_2; z) f_2(z)\,dz\right) dy_1\,dy_2\,dx_2\,\nu(dx_1)\,ds
\end{aligned} \tag{3.5}$$

with the convention $\int_{\mathbb{R}^d} p^1_0(x; y)h(y)\,dy = h(x)$.

PROOF. Since $X_0^n \Rightarrow \nu$ it follows by Proposition 4.6(b), Chapter 3 in [11] that $X_0^n \times X_0^n \Rightarrow \nu \times \nu$ as well. Hence we easily see that the first term in the right-hand side of (3.2) converges to the first term in the right-hand side of (3.5) as $n \to \infty$. Furthermore, since the multiple integrals in the second and third terms in the right-hand side of (3.2) converge to finite quantities as $n \to \infty$, it is clear that these two terms vanish in the limit due to the factor $1/n$. The last term (a sum) in the right-hand side of (3.2) tends to the last term (an integral) in the right-hand side of (3.5) as $n \to \infty$. To see why this is the case, we first let

$$g(s, x_1, x_2) = p_s^1(x_1; x_2) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_{t_1-s}^2(x_2, x_2; y_1, y_2) f_1(y_1) \times \left(\int_{\mathbb{R}^d} p_{t_2-t_1}^1(y_2; z) f_2(z)\, dz \right) dy_1\, dy_2$$

for $0 \le s \le t_1$ and $x_1, x_2 \in \mathbb{R}^d$. Note that it suffices to show that the difference between

$$(\lambda n)^{-1} \sum_{r=0}^{N_1-1} \left(1 + \frac{\gamma_n}{n}\right)^{-r} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0^n(dx_1)$$

and

$$(\lambda n)^{-1} \sum_{r=0}^{N_1-1} e^{-\lambda \gamma r_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0(dx_1)$$

converges to 0 as $n \to \infty$, where $N_1 = [\lambda t_1 n]$. But this is true since

$$\begin{aligned}
&\left| \left(1 + \frac{\gamma_n}{n}\right)^{-r} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0^n(dx_1) \right. \\
&\qquad \left. - e^{-\lambda\gamma r_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2) dx_2 X_0(dx_1) \right| \\
&\quad \le \left(1 + \frac{\gamma_n}{n}\right)^{-r} \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0^n(dx_1) \right. \\
&\qquad\qquad \left. - e^{-\lambda\gamma r_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0^n(dx_1) \right| \\
&\qquad + \left| e^{-\lambda\gamma r_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2\, X_0^n(dx_1) \right. \\
&\qquad\qquad \left. - e^{-\lambda\gamma r_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2 X_0(dx_1) \right|,
\end{aligned}$$

which converges to 0 as $n \to \infty$, uniformly in r, for $0 \le r \le [\lambda t_1 n]$. This is a consequence of the facts that $|(1+(\gamma_n/n))^{-r} - e^{-\lambda\gamma r_n}| \to 0$ as $n \to \infty$, uniformly in r, for $0 \le r \le [\lambda t_1 n]$ (easily seen by simple calculus arguments) and that

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2 X_0^n(dx_1) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2 X_0(dx_1) \right| \to 0$$

as $n \to \infty$, uniformly in r, for $0 \le r \le [\lambda t_1 n]$. (This follows from Problem 2, Chapter 3 in [11], since $\int_{\mathbb{R}^d} g(r_n, x_1, x_2)\, dx_2$ is bounded for all r and n and Lipschitz in x_1 by Assumption I.) Next we take care of the left-hand side of (3.2). Let (π_n) be the subsequence along which X^n converges weakly to X. By employing a Skorokhod representation and then using Proposition 3.2.1 and Problem 2, Chapter 3 in [11] we obtain that $(X^{\pi_n}(f_1), X^{\pi_n}(f_2)) \Rightarrow (X(f_1), X(f_2))$ in $D_{\mathbb{R}^2}[0, \infty)$. By Lemma A.3.9 in the Appendix, it follows that $X(f_1)$ and $X(f_2)$ are continuous and so by the continuous mapping theorem we obtain that $(X_{t_1}^{\pi_n}(f_1), X_{t_2}^{\pi_n}(f_2)) \Rightarrow (X_{t_1}(f_1), X_{t_2}(f_2))$ in $\mathbb{R}^2$ and so $X_{t_1}^{\pi_n}(f_1) X_{t_2}^{\pi_n}(f_2) \Rightarrow X_{t_1}(f_1) X_{t_2}(f_2)$ in $\mathbb{R}$. By Lemma A.3.1 in the Appendix, we have $\sup_{n \ge 1} E(X_{t_1}^n(1) X_{t_2}^n(1))^{p/2} < \infty$ (recall that $p > 2$) and now applying the Corollary to Theorem 25.12 in [3] provides uniform integrability and yields that $E(X_{t_1}^{\pi_n}(f_1) X_{t_2}^{\pi_n}(f_2)) \to E(X_{t_1}(f_1) X_{t_2}(f_2))$. Letting $n \to \infty$ along the subsequence (π_n) in (3.2) yields the desired formula. □

REMARK. By repeating analogous calculations on the same processes, but without the presence of branching, it is straightforward to check that (3.4) and (3.5) still hold, but now with $\lambda = 0$.

3.3. *The mean and the spatial covariance measures and their densities.* In this section we consider two descriptive, macroscopic, quantities related to flows and investigate them in the context of our model. Let X be a weak limit point of the sequence X^n. Following [32], we define two deterministic measures as follows:

1. The *mean measure* m_t defined on $\mathscr{B}(\mathbb{R}^d)$ by

$$m_t(A) = E(X_t(A)). \tag{3.6}$$

2. The *spatial covariance measure* c_t defined on $\mathscr{B}(\mathbb{R}^d \times \mathbb{R}^d)$ by

$$c_t(A \times B) = E(X_t(A) X_t(B)). \tag{3.7}$$

Note that Assumption I, made in the introduction of this section, is still in effect and $\sigma^2 = 1$. The main result of this subsection is contained in the following proposition.

PROPOSITION 3.3.1. *For all $t > 0$ the measures m_t and c_t, defined by* (3.6) *and* (3.7)*, have densities with respect to Lebesgue measure, which we denote by η_t and ζ_t, respectively. The densities are*

$$\eta_t(y) = \exp(\lambda \gamma t) \int_{\mathbb{R}^d} p_t^1(x; y) \nu(dx) \tag{3.8}$$

for all $y \in \mathbb{R}^d$ *and*

$$\begin{aligned}\zeta_t(y_1, y_2) = {}& \exp(2\lambda\gamma t)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_t^2(x_1, x_2; y_1, y_2)\nu(dx_1)\nu(dx_2)\\ &+ \lambda\exp(2\lambda\gamma t)\int_0^t \exp(-\lambda\gamma s)\\ &\times \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_s^1(u; x)p_{t-s}^2(x, x; y_1, y_2)\,dx\,\nu(du)\,ds\end{aligned} \tag{3.9}$$

for all $y_1, y_2 \in \mathbb{R}^d$.

PROOF. We first show that $EX_t(f) = \int_{\mathbb{R}^d} f(x)m_t(dx)$ for all $f \in C_b(\mathbb{R}^d)$. But since this is true for all indicator functions $f = I_A$, $A \in \mathscr{B}(\mathbb{R}^d)$ by definition, a standard monotone convergence argument gives us that it is also true for nonnegative $f \in C_b(\mathbb{R}^d)$ and so for all $f \in C_b(\mathbb{R}^d)$. Clearly $C_K^2(\mathbb{R}^d)$ separates points and therefore from Proposition 4.5(a), Chapter 3 in [11] we see that $C_K^2(\mathbb{R}^d)$ is separating. The existence and the form of the density η_t of m_t now follow immediately from (3.4). Arguing in a similar fashion we obtain that $EX_t(f_1)X_t(f_2) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f_1(y_1)f_2(y_2)c_t(dy_1, dy_2)$ for all $f_1, f_2 \in C_b(\mathbb{R}^d)$. Moreover Proposition 4.5(a) and Proposition 4.6(a), Chapter 3 in [11] imply that the collection $\{h\colon h(x_1, x_2) = f_1(x_1)f_2(x_2), x_1, x_2 \in \mathbb{R}^d, \text{where} f_1, f_2 \in C_K^2(\mathbb{R}^d)\}$ is separating. Clearly the existence and the form of the density ζ_t of c_t now follow from (3.5) and the proof is complete. □

REMARK. Using (3.8) and (3.9) one can calculate the covariance between $X_t(A)$ and $X_t(B)$ where A, B are Borel subsets of $\mathbb{R}^d$ since

$$\begin{aligned}\operatorname{Cov}(X_t(A), X_t(B)) &= c_t(A \times B) - m_t(A)m_t(B)\\ &= \int_A\int_B \theta_t(y_1, y_2)\,dy_1\,dy_2,\end{aligned} \tag{3.10}$$

where

$$\theta_t(y_1, y_2) = \zeta_t(y_1, y_2) - \eta_t(y_1)\eta_t(y_2). \tag{3.11}$$

We now turn to a special case, to get a better feel for what the formulas of this section are telling us.

4. Application to a flow of Ornstein–Uhlenbeck type. In this section we look at a very specific example, derive explicit formulas and show a few graphs, so as to give a feel for what the general results of the previous section are saying. Even though we have cut out a lot of details, the computations are still long. Nevertheless, they are tractable, as they should be for most cases. For more complicated examples, computer algebra and/or numerical computation might be necessary.

The example that we shall treat is that in which the stochastic differential equations satisfied by the diffusions that describe the one-point and two-point motions [see (2.12)] are linear, of Ornstein–Uhlenbeck type (so that Assumption I holds) and driven by an $(m+d)$-dimensional Brownian motion. More specifically, we start with a flow of Ornstein–Uhlenbeck type. That is, in the notation of Section 2, we assume $b(x) = \mathbf{A}x + a$, $(c_{ij}(x)) = \mathbf{C}$, and $(e_i(x)) = \varepsilon$, where $\mathbf{A}$ is a $d \times d$ matrix, $\mathbf{C}$ is a $d \times m$ matrix and a, ε are column vectors in $\mathbb{R}^d$. Set $\mathbf{E} = \operatorname{diag}(\varepsilon)$ and $\mathbf{\Sigma} = [\mathbf{C}, \mathbf{E}]$ [a $d \times (m+d)$ matrix] and assume that the matrix $[\mathbf{E}, \mathbf{AE}, \dots, \mathbf{A}^{d-1}\mathbf{E}]$ has rank d.

Furthermore, assume that the initial measure ν has the Gaussian density

$$g(x) = (2\pi)^{-d/2}|\mathbf{V}_0|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu_0)^T \mathbf{V}_0^{-1}(x-\mu_0)\right\}, \quad x \in \mathbb{R}^d, \tag{4.1}$$

where μ_0 is a column vector in $\mathbb{R}^d$ and $\mathbf{V}_0$ is a positive definite $d \times d$ matrix.

This is all we need to be able to compute.

4.1. *Calculation of the density η_t.* From the discussion in the introduction of Section 5.6 and Problem 6.1 in [15] it follows that, for $t > 0$, we have

$$p_t^1(x; y) = (2\pi)^{-d/2}|\mathbf{V}_t|^{-1/2} \exp\left\{-\frac{1}{2}(y - \mathbf{\Psi}_t x - \mathbf{K}_t)^T \mathbf{V}_t^{-1}(y - \mathbf{\Psi}_t x - \mathbf{K}_t)\right\},$$

where

$$\mathbf{\Psi}_t = \exp(t\mathbf{A}) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n, \qquad \mathbf{K}_t = \mathbf{\Psi}_t\left(\int_0^t \mathbf{\Psi}_s^{-1}\, ds\right)a,$$

$$\mathbf{V}_t = \mathbf{\Psi}_t \mathbf{\Theta}_t \mathbf{\Psi}_t^T, \quad \text{and} \quad \mathbf{\Theta}_t = \int_0^t \mathbf{\Psi}_u^{-1}\mathbf{\Sigma}(\mathbf{\Psi}_u^{-1}\Sigma)^T\, du.$$

From the assumption that the matrix $[\mathbf{E}, \mathbf{AE}, \dots, \mathbf{A}^{d-1}\mathbf{E}]$ has rank d it follows that the matrix $[\mathbf{\Sigma}, \mathbf{A\Sigma}, \dots, \mathbf{A}^{d-1}\mathbf{\Sigma}]$ has rank d as well. Then, Propositions 6.4 and 6.5, Section 5.6 in [15] imply that the matrix $\mathbf{V}_t$ is positive definite and so nonsingular. We now start computing the density η_t of the mean measure. By (3.8) we obtain, for $t > 0$,

$$\begin{aligned}\eta_t(y) = {} & \exp(\lambda\gamma t)(2\pi)^{-d}|\mathbf{V}_0|^{-1/2}|\mathbf{V}_t|^{-1/2} \\ & \times \int_{\mathbb{R}^d} \exp\left\{-\tfrac{1}{2}\Big[(x-\mu_0)^T \mathbf{V}_0^{-1}(x-\mu_0)\right. \\ & \left.\qquad + (y - \mathbf{\Psi}_t x - \mathbf{K}_t)^T\, \mathbf{V}_t^{-1}(y - \mathbf{\Psi}_t x - \mathbf{K}_t)\Big]\right\} dx.\end{aligned}$$

In order to be able to integrate with respect to x we simplify the expression in the exponent. For this we set

$$\begin{aligned}l(t, x, y) &= (x-\mu_0)^T \mathbf{V}_0^{-1}(x-\mu_0) + (y - \mathbf{\Psi}_t x - \mathbf{K}_t)^T \mathbf{V}_t^{-1}(y - \mathbf{\Psi}_t x - \mathbf{K}_t) \\ &= x^T(\mathbf{\Psi}_t^T \mathbf{V}_t^{-1}\mathbf{\Psi}_t + \mathbf{V}_0^{-1})x - 2(y - \mathbf{K}_t)^T \mathbf{V}_t^{-1}\mathbf{\Psi}_t x \\ &\quad + (y - \mathbf{K}_t)^T \mathbf{V}_t^{-1}(y - \mathbf{K}_t) - 2\mu_0^T \mathbf{V}_0^{-1}x + \mu_0^T \mathbf{V}_0^{-1}\mu_0.\end{aligned}$$

Under our assumptions, $\mathbf{\Theta}_t$ and $\mathbf{V}_0$ are positive definite. Hence $\mathbf{\Theta}_t^{-1} + \mathbf{V}_0^{-1}$ is positive definite and so nonsingular. If we then let $\mathbf{U}_t = (\mathbf{\Theta}_t^{-1} + \mathbf{V}_0^{-1})^{-1}$ we trivially have that $\mathbf{U}_t^{-1} = \mathbf{\Psi}_t^T \mathbf{V}_t^{-1} \mathbf{\Psi}_t + \mathbf{V}_0^{-1}$. Therefore,

$$l(t, x, y) = \Big(x - \mu_t(y)\Big)^T \mathbf{U}_t^{-1} \Big(x - \mu_t(y)\Big) + r(t, y),$$

where

$$\mu_t(y) = \mathbf{U}_t \Big({\mathbf{\Psi}_t}^T \mathbf{V}_t^{-1} (y - \mathbf{K}_t) + \mathbf{V}_0^{-1} \mu_0 \Big)$$

and

$$r(t, y) = -\mu_t(y)^T \mathbf{U}_t^{-1} \mu_t(y) + (y - \mathbf{K}_t)^T \mathbf{V}_t^{-1} (y - \mathbf{K}_t) + {\mu_0}^T \mathbf{V}_0^{-1} \mu_0.$$

Furthermore,

$$\begin{aligned} r(t, y) = &-\Big(\mu_0^T \mathbf{V}_0^{-1} + (y - \mathbf{K}_t)^T (\mathbf{\Psi}_t^{-1})^T \mathbf{\Theta}_t^{-1}\Big) \mathbf{U}_t \Big(\mathbf{\Theta}_t^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{K}_t) + \mathbf{V}_0^{-1} \mu_0\Big) \\ &+ (y - \mathbf{K}_t)^T (\mathbf{\Psi}_t^{-1})^T \mathbf{\Theta}_t^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{K}_t) + \mu_0^T \mathbf{V}_0^{-1} \mu_0 \\ = &-(y - \mathbf{K}_t)^T (\mathbf{\Psi}_t^{-1})^T \mathbf{\Theta}_t^{-1} \mathbf{U}_t \mathbf{\Theta}_t^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{K}_t) - \mu_0^T \mathbf{V}_0^{-1} \mathbf{U}_t \mathbf{V}_0^{-1} \mu_0 \\ &- 2\mu_0^T \mathbf{V}_0^{-1} \mathbf{U}_t \mathbf{\Theta}_t^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{K}_t) + (y - \mathbf{K}_t)^T (\mathbf{\Psi}_t^{-1})^T \mathbf{\Theta}_t^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{K}_t) \\ &+ \mu_0^T \mathbf{V}_0^{-1} \mu_0. \end{aligned}$$

A few lines of calculations give us that $\Theta_t^{-1} \mathbf{U}_t \mathbf{\Theta}_t^{-1} = \mathbf{\Theta}_t^{-1} - (\Theta_t + \mathbf{V}_0)^{-1}$, $\mathbf{V}_0^{-1} \mathbf{U}_t \mathbf{V}_0^{-1} = \mathbf{V}_0^{-1} - (\mathbf{\Theta}_t + \mathbf{V}_0)^{-1}$ and $\mathbf{V}_0^{-1} \mathbf{U}_t \mathbf{\Theta}_t^{-1} = (\mathbf{\Theta}_t + \mathbf{V}_0)^{-1}$ and by substitution we obtain

$$r(t, y) = (y - \mathbf{\Psi}_t \mu_0 - \mathbf{K}_t)^T (\mathbf{\Psi}_t^{-1})^T (\mathbf{\Theta}_t + \mathbf{V}_0)^{-1} \mathbf{\Psi}_t^{-1} (y - \mathbf{\Psi}_t \mu_0 - \mathbf{K}_t).$$

Thus integrating out x yields

$$\eta_t(y) = \exp(\lambda \gamma t) (2\pi)^{-d/2} |\mathbf{V}_0|^{-1/2} |\mathbf{V}_t|^{-1/2} |\mathbf{U}_t|^{1/2} \exp\left\{-\tfrac{1}{2} r(t, y)\right\}$$

and since $|\mathbf{V}_0||\mathbf{V}_t||\mathbf{U}_t^{-1}| = |\mathbf{\Psi}_t (\mathbf{\Theta}_t + \mathbf{V}_0) \mathbf{\Psi}_t^T|$ we finally conclude

$$\begin{aligned} \eta_t(y) = &(2\pi)^{-d/2} \exp(\lambda \gamma t) |\mathbf{V}_t^\eta|^{-1/2} \\ &\times \exp\left\{-\tfrac{1}{2} (y - \mu_t^\eta)^T (\mathbf{V}_t^\eta)^{-1} (y - \mu_t^\eta)\right\}, \end{aligned} \tag{4.2}$$

where

$$\mu_t^\eta = \mathbf{\Psi}_t \mu_0 + \mathbf{K}_t \tag{4.3}$$

and

$$\mathbf{V}_t^\eta = \mathbf{V}_t + \mathbf{\Psi}_t \mathbf{V}_0 \mathbf{\Psi}_t^T. \tag{4.4}$$

REMARK 1. The density η_t is a multiple of a Gaussian density with mean μ_t^η and covariance matrix $\mathbf{V}_t^\eta$. The coefficient $e^{\lambda\gamma t}$ appearing in (4.2) represents the rate of the mean mass creation in the case of supercritical branching ($\lambda\gamma > 0$). It should be noted here that one could, in a straightforward fashion, extend this result to the case in which ν is a pure atomic measure or a mixture of pure atomic measures and Gaussian measures.

REMARK 2. If we make the simplifying assumption that the matrix $\mathbf{A}$ has the special diagonal form $\mathbf{A} = \operatorname{diag}(r_1, r_2, \dots, r_d)$, we can obtain, after some lines of routine calculations, the following formulas:

$$\mu_t^\eta = \left(\mu_1^0 e^{r_1 t} + a_1 \frac{e^{r_1 t} - 1}{r_1}, \dots, \mu_d^0 e^{r_d t} + a_d \frac{e^{r_d t} - 1}{r_d}\right)^T \tag{4.5}$$

and

$$\mathbf{V}_t^\eta = \left(e^{(r_i + r_j)t} V_{ij}^0 + \frac{e^{(r_i+r_j)t} - 1}{r_i + r_j}\left(\sum_{k=1}^m c_{ik} c_{jk} + \delta_{ij}\varepsilon_i \varepsilon_j\right)\right)_{i,\, j=1,\dots,d}, \tag{4.6}$$

where $\mu_0 = (\mu_1^0, \dots, \mu_d^0)$ and $\mathbf{V}_0 = (V_{ij}^0)_{i,\, j=1,\dots,d}$. If $r_i = 0$ for some i we replace $(e^{r_i t} - 1)/r_i$ by t in (4.5). Similarly, if $r_i + r_j = 0$ for some i, j we replace $(e^{(r_i+r_j)t} - 1)/(r_i + r_j)$ by t in (4.6). Note that in this case our assumption that the matrix $[\mathbf{E}, \mathbf{AE}, \dots, \mathbf{A}^{d-1}\mathbf{E}]$ has rank d reduces to the assumption that $\varepsilon_i \neq 0$ for all $i = 1, 2, \dots, d$.

In the case $r_i < 0$ for all $i = 1, 2, \dots, d$, which can be described as the mean reverting flow case, we observe that the median μ_t^η of η_t tends, at a rate depending on r_i's, to $(-a_1/r_1, \dots, -a_d/r_d)$ as t increases. In other words, on average, the mass moves to a region around $(-a_1/r_1, \dots, -a_d/r_d)$ as times evolves. In general, if we denote

$$\lim_{t\to\infty} \mu_t^\eta = \mu_\infty^\eta = \left(\mu_\infty^{\eta,1}, \mu_\infty^{\eta,2}, \dots, \mu_\infty^{\eta,d}\right),$$

then we have

$$\mu_\infty^{\eta,\, i} = \begin{cases} \infty, & \text{if } r_i = 0,\ a_i > 0 \quad \text{or} \quad r_i > 0,\ \mu_i^0 + a_i/r_i > 0, \\ -\infty, & \text{if } r_i = 0,\ a_i < 0 \quad \text{or} \quad r_i > 0,\ \mu_i^0 + a_i/r_i < 0, \\ -a_i/r_i, & \text{if } \mu_i^0 + a_i/r_i = 0 \quad \text{or} \quad r_i < 0, \\ \mu_i^0, & \text{if } r_i = a_i = 0. \end{cases} \tag{4.7}$$

The densities μ_t^η and μ_∞^η describe the (mean) distribution of mass. Of independent interest are the corresponding centers of mass, given by

$$m_t^{\eta,\, i} = \int_{\mathbb{R}^d} x_i \eta_t(x)\, dx \quad \text{and} \quad m_\infty^{\eta,\, i} = \lim_{t\to\infty} m_t^{\eta,\, i}.$$

Then if we set $m_t^\eta = (m_t^{\eta,\, 1}, \dots, m_t^{\eta,\, d})$ we have $m_t^\eta = e^{\lambda\gamma t}\mu_t^\eta$. To describe the limiting behavior of m_t^η we need to consider the set

$$\mathbf{\Pi} = \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \text{ and its partition } \mathbf{\Pi} = \mathbf{\Pi}_1 \cup \mathbf{\Pi}_2 \cup \mathbf{\Pi}_3 \cup \mathbf{\Pi}_4 \cup \mathbf{\Pi}_5,$$

where

$$\begin{aligned}
\mathbf{\Pi}_1 &= \{(\theta,\mu,r,a)\in\mathbf{\Pi}\colon \mu=a=0 \ or\ \theta>0, r<0, a=0, \theta+r<0\},\\
\mathbf{\Pi}_2 &= \{(\theta,\mu,r,a)\in\mathbf{\Pi}\colon \theta+r=0,\ r\le 0,\ a=0\},\\
\mathbf{\Pi}_3 &= \{(\theta,\mu,r,a)\in\mathbf{\Pi}\colon \theta=0, r<0 \ or\ \theta=0,\ r>0, \mu+a/r=0\},\\
\mathbf{\Pi}_4 &= \{(\theta,\mu,r,a)\in\mathbf{\Pi}\colon r=0, a>0 \ or\ \theta>0, \mu>0, r=a=0\\
&\qquad or\quad r>0, \mu+a/r>0 \ or\ \theta>0,\ \mu>0, r>0, \mu+a/r=0\\
&\qquad or\quad \theta>0, r>0, a>0 \ or\ \theta>0, \mu>0, r<0,\ a=0, \theta+r>0\},\\
\mathbf{\Pi}_5 &= \{(\theta,\mu,r,a)\in\mathbf{\Pi}\colon r=0,\ a<0 \ or\ \theta>0, \mu<0, r=a=0\\
&\qquad or\quad r>0, \mu+a/r<0 \ or\ \theta>0, \mu<0, r>0, \mu+a/r=0\\
&\qquad or\quad \theta>0, r>0, a<0 \ or\ \theta>0, \mu<0, r<0, a=0, \theta+r>0\}.
\end{aligned}$$

Then we have that

$$m_\infty^{\eta,i} = \begin{cases} 0, & \text{if } (\lambda\gamma,\mu_i^0,r_i,a_i)\in\mathbf{\Pi}_1,\\ \mu_i^0, & \text{if } (\lambda\gamma,\mu_i^0,r_i,a_i)\in\mathbf{\Pi}_2,\\ -a_i/r_i, & \text{if } (\lambda\gamma,\mu_i^0,r_i,a_i)\in\mathbf{\Pi}_3,\\ \infty, & \text{if } (\lambda\gamma,\mu_i^0,r_i,a_i)\in\mathbf{\Pi}_4,\\ -\infty, & \text{if } (\lambda\gamma,\mu_i^0,r_i,a_i)\in\mathbf{\Pi}_5. \end{cases}$$

4.2. *Calculation of the density* ζ_t. We first need to introduce some notation. Let

$$\tilde a = \begin{bmatrix} a\\ a\end{bmatrix},\qquad \widetilde{\mathbf{A}} = \begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}\end{bmatrix},\qquad \widetilde{\mathbf{\Sigma}} = \begin{bmatrix}\mathbf{C} & \mathbf{E} & \mathbf{0}\\ \mathbf{C} & \mathbf{0} & \mathbf{E}\end{bmatrix}$$

and

$$\tilde\mu = \begin{bmatrix}\mu_0\\ \mu_0\end{bmatrix},\quad \widetilde{\mathbf{V}}_0 = \begin{bmatrix}\mathbf{V}_0 & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_0\end{bmatrix}.$$

From the discussion in the introduction of Section 5.6 in [15] and Problem 6.1 there it follows that, for $t>0$, we have

$$\begin{aligned} &p_t^2(x_1,x_2;y_1,y_2)\\ &\quad = (2\pi)^{-d}|\widetilde{\mathbf{V}}_t|^{-1/2}\exp\{-\tfrac12(\tilde y-\widetilde{\mathbf{\Psi}}_t\tilde x-\widetilde{\mathbf{K}}_t)^T\widetilde{\mathbf{V}}_t^{-1}(\tilde y-\widetilde{\mathbf{\Psi}}_t\tilde x-\widetilde{\mathbf{K}}_t)\}, \end{aligned}$$

where

$$\tilde x = \begin{bmatrix}x_1\\ x_2\end{bmatrix},\qquad \tilde y = \begin{bmatrix}y_1\\ y_2\end{bmatrix},\qquad \widetilde{\mathbf{\Psi}}_t = \exp(t\widetilde{\mathbf{A}}) = \sum_{n=0}^\infty \frac{t^n}{n!}\widetilde{\mathbf{A}}^n,$$

$$\widetilde{\mathbf{K}}_t = \widetilde{\mathbf{\Psi}}_t\left(\int_0^t \widetilde{\mathbf{\Psi}}_s^{-1}\,ds\right)\tilde a,\quad \widetilde{\mathbf{V}}_t = \widetilde{\mathbf{\Psi}}_t\widetilde{\Theta}_t\widetilde{\mathbf{\Psi}}_t^T \quad\text{and}\quad \mathbf{\Theta}_t = \int_0^t \widetilde{\mathbf{\Psi}}_u^{-1}\widetilde{\mathbf{\Sigma}}\left(\widetilde{\mathbf{\Psi}}_u^{-1}\widetilde{\mathbf{\Sigma}}\right)^T du.$$

The assumption that the matrix $[\mathbf{E}, \mathbf{A}\mathbf{E}, \dots, \mathbf{A}^{d-1}\mathbf{E}]$ has rank d implies that the matrix $[\widetilde{\boldsymbol{\Sigma}}, \widetilde{\mathbf{A}}\widetilde{\boldsymbol{\Sigma}}, \dots, \widetilde{\mathbf{A}}^{d-1}\widetilde{\Sigma}]$ has rank $2d$ and therefore by Propositions 6.4 and 6.5, Section 5.6 in [15] it follows that the matrix $\widetilde{\mathbf{V}}_t$ is positive definite and so nonsingular. Thus, the first term in the right-hand side of (3.9) equals

$$\begin{aligned}
&\exp(2\lambda\gamma t)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_t^2(x_1, x_2; y_1, y_2)(2\pi)^{-d/2}|\mathbf{V}_0|^{-1/2}\\
&\qquad\qquad \times \exp\bigl\{-\tfrac{1}{2}(x_1-\mu_0)^T\mathbf{V}_0^{-1}(x_1-\mu_0)\bigr\}(2\pi)^{-d/2}|\mathbf{V}_0|^{-1/2}\\
&\qquad\qquad \times \exp\bigl\{-\tfrac{1}{2}(x_2-\mu_0)^T\mathbf{V}_0^{-1}(x_2-\mu_0)\bigr\}\,dx_1dx_2\\
&\quad = \exp(2\lambda\gamma t)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(2\pi)^{-d}|\widetilde{\mathbf{V}}_t|^{-1/2}\\
&\qquad \times \exp\bigl\{-\tfrac{1}{2}(\tilde{y}-\widetilde{\boldsymbol{\Psi}}_t\tilde{x}-\widetilde{\mathbf{K}}_t)^T\widetilde{\mathbf{V}}_t^{-1}(\tilde{y}-\widetilde{\boldsymbol{\Psi}}_t\tilde{x}-\widetilde{\mathbf{K}}_t)\bigr\}\\
&\qquad \times (2\pi)^{-d}|\widetilde{\mathbf{V}}_0|^{-1/2}\exp\bigl\{-\tfrac{1}{2}(\tilde{x}-\tilde{\mu}_0)^T\widetilde{\mathbf{V}}_0^{-1}(\tilde{x}-\tilde{\mu}_0)\bigr\}\,dx_1dx_2\\
&\quad = \exp(2\lambda\gamma t)(2\pi)^{-d}|\mathbf{V}_t^{\zeta}|^{-1/2}\exp\bigl\{-\tfrac{1}{2}(\tilde{y}-\mu_t^{\zeta})^T(\mathbf{V}_t^{\zeta})^{-1}(\tilde{y}-\mu_t^{\zeta})\bigr\},
\end{aligned}$$

where $\mu_t^{\zeta} = \widetilde{\boldsymbol{\Psi}}_t\tilde{\mu}_0 + \widetilde{\mathbf{K}}_t$ and $\mathbf{V}_t^{\zeta} = \widetilde{\mathbf{V}}_t + \widetilde{\boldsymbol{\Psi}}_t\widetilde{\mathbf{V}}_0\widetilde{\boldsymbol{\Psi}}_t^T$. The last equality follows from a calculation almost identical to the calculation done in Section 4.1. After some elementary matrix computations, we obtain

$$\widetilde{\boldsymbol{\Psi}}_t = \begin{bmatrix}\boldsymbol{\Psi}_t & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Psi}_t\end{bmatrix}, \qquad \widetilde{\mathbf{K}}_t = \begin{bmatrix}\mathbf{K}_t\\ \mathbf{K}_t\end{bmatrix}, \qquad \widetilde{\boldsymbol{\Sigma}}_t\widetilde{\boldsymbol{\Sigma}}_t^T = \begin{bmatrix}\Sigma\boldsymbol{\Sigma}^T & \mathbf{C}\mathbf{C}^T\\ \mathbf{C}\mathbf{C}^T & \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\end{bmatrix},$$

$$\widetilde{\boldsymbol{\Psi}}_t^{-1}\widetilde{\Sigma}_t\widetilde{\Sigma}_t^T(\widetilde{\boldsymbol{\Psi}}_t^{-1})^T = \begin{bmatrix}\boldsymbol{\Psi}_t^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T(\boldsymbol{\Psi}_t^{-1})^T & \boldsymbol{\Psi}_t^{-1}\mathbf{C}\mathbf{C}^T(\boldsymbol{\Psi}_t^{-1})^T\\ \boldsymbol{\Psi}_t^{-1}\mathbf{C}\mathbf{C}^T(\boldsymbol{\Psi}_t^{-1})^T & \boldsymbol{\Psi}_t^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T(\boldsymbol{\Psi}_t^{-1})^T\end{bmatrix}$$

and therefore

$$\widetilde{\mathbf{V}}_t = \begin{bmatrix}\mathbf{V}_t & \mathbf{N}_t\\ \mathbf{N}_t & \mathbf{V}_t\end{bmatrix} \tag{4.8}$$

where

$$\mathbf{N}_t = \boldsymbol{\Psi}_t\biggl(\int_0^t \boldsymbol{\Psi}_u^{-1}\mathbf{C}\mathbf{C}^T(\boldsymbol{\Psi}_u^{-1})^T\,du\biggr)\boldsymbol{\Psi}_t^T. \tag{4.9}$$

Consequently,

$$\mu_t^{\zeta} = \begin{bmatrix}\boldsymbol{\Psi}_{t\mu_0} + \mathbf{K}_t\\ \boldsymbol{\Psi}_{t\mu_0} + \mathbf{K}_t\end{bmatrix} = \begin{bmatrix}\mu_t^{\eta}\\ \mu_t^{\eta}\end{bmatrix} \tag{4.10}$$

and

$$\mathbf{V}_t^{\zeta} = \begin{bmatrix}\mathbf{V}_t + \boldsymbol{\Psi}_t\mathbf{V}_0\boldsymbol{\Psi}_t^T & \mathbf{N}_t\\ \mathbf{N}_t & \mathbf{V}_t + \boldsymbol{\Psi}_t\mathbf{V}_0\boldsymbol{\Psi}_t^T\end{bmatrix} = \begin{bmatrix}\mathbf{V}_t^{\eta} & \mathbf{N}_t\\ \mathbf{N}_t & \mathbf{V}_t^{\eta}\end{bmatrix}. \tag{4.11}$$

Next we take care of the second term in the right-hand side of (3.9). To calculate the integral appearing in this term we first integrate with respect to u [exactly as we did in Section 4.1 where r is defined] to get

$$\int_{\mathbb{R}^d} p_s^1(u;x)g(u)\,du = (2\pi)^{-d/2}|\mathbf{V}_s^\eta|^{-1/2}\exp\left\{-\tfrac{1}{2}r(s,x)\right\}.$$

In order to continue an expression for the inverse of the matrix $\widetilde{\mathbf{V}}_t$ is needed. But, if we let

$$\mathbf{M}_t = \mathbf{\Psi}_t\Big(\int_0^t \mathbf{\Psi}_u^{-1}\mathbf{E}\mathbf{E}^T(\mathbf{\Psi}_u^{-1})^T\,du\Big)\mathbf{\Psi}_t^T \tag{4.12}$$

and

$$\mathbf{X}_t = -(\mathbf{V}_t+\mathbf{N}_t)^{-1}\mathbf{N}_t\mathbf{M}_t^{-1} \tag{4.13}$$

then it is easily verified that

$$\widetilde{\mathbf{V}}_t^{-1} = \begin{bmatrix} \mathbf{X}_t+\mathbf{M}_t^{-1} & \mathbf{X}_t \\ \mathbf{X}_t & \mathbf{X}_t+\mathbf{M}_t^{-1}\end{bmatrix}. \tag{4.14}$$

The following two observations are necessary. First, the assumption that the matrix has rank d implies that $\mathbf{M}_t$ is positive definite. Moreover, we observe that $\mathbf{V}_t+\mathbf{N}_t = \mathbf{\Psi}_t\int_0^t\mathbf{\Psi}_u^{-1}\ \mathbf{\Sigma}_*\mathbf{\Sigma}_*^T(\mathbf{\Psi}_u^{-1})^T\,du\ \ \mathbf{\Psi}_t^T$, where $\mathbf{\Sigma}_* = [\sqrt{2}\mathbf{C},\mathbf{E}]$, and so conclude that $\mathbf{V}_t+\mathbf{N}_t$ is positive definite as well. This is true since rank $([\mathbf{\Sigma}_*,\mathbf{A}\mathbf{\Sigma}_*,\dots,\mathbf{A}^{d-1}\ \mathbf{\Sigma}_*]) = d$ which follows from the assumption rank$([\mathbf{E},\mathbf{A}\mathbf{E},\dots,\mathbf{A}^{d-1}\mathbf{E}]) = d$. Therefore,

$$\begin{aligned}
&p_{t-s}^2(x,x;y_1,y_2)\int_{\mathbb{R}^d} p_s^1(u;x)g(u)\,du\\
&\quad= (2\pi)^{-3d/2}\big|\widetilde{\mathbf{V}}_{t-s}\big|^{-1/2}|\mathbf{V}_s^\eta|^{-1/2}\exp\Big\{-\frac{1}{2}\Big(q(t-s,x,y_1,y_2)+r(s,x)\Big)\Big\},
\end{aligned}$$

where

$$\begin{aligned}
q(u,x,y_1,y_2) &= \begin{bmatrix} y_1-\mathbf{\Psi}_u x-\mathbf{K}_u \\ y_2-\mathbf{\Psi}_u x-\mathbf{K}_u\end{bmatrix}^T\begin{bmatrix} \mathbf{X}_u+\mathbf{M}_u^{-1} & \mathbf{X}_u \\ \mathbf{X}_u & \mathbf{X}_u+\mathbf{M}_u^{-1}\end{bmatrix}\\
&\qquad\times\begin{bmatrix} y_1-\mathbf{\Psi}_u x-\mathbf{K}_u \\ y_2-\mathbf{\Psi}_u x-\mathbf{K}_u\end{bmatrix}.
\end{aligned}$$

Hence

$$\begin{aligned}
&q(t-s,x,y_1,y_2)+r(s,x)\\
&\quad= (y_1-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_1-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})\\
&\qquad+(y_1-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}(y_2-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})\\
&\qquad+(y_2-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}(y_1-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})\\
&\qquad+(y_2-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_2-\mathbf{\Psi}_{t-s}x-\mathbf{K}_{t-s})+r(s,x)
\end{aligned}$$

$$\begin{aligned}
&= 2x^T\boldsymbol{\Psi}_{t-s}^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})\boldsymbol{\Psi}_{t-s}x + 2x^T\boldsymbol{\Psi}_{t-s}^T\mathbf{X}_{t-s}\boldsymbol{\Psi}_{t-s}x\\
&\quad - 2(y_1-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})\boldsymbol{\Psi}_{t-s}x - 2(y_2-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}\boldsymbol{\Psi}_{t-s}x\\
&\quad - 2(y_1-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}\boldsymbol{\Psi}_{t-s}x - 2(y_2-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})\boldsymbol{\Psi}_{t-s}x\\
&\quad + (y_1-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_1-\mathbf{K}_{t-s}) + 2(y_1-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}(y_2-\mathbf{K}_{t-s})\\
&\quad + (y_2-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_2-\mathbf{K}_{t-s})\\
&\quad + x^T(\mathbf{V}_s^\eta)^{-1}x - 2(\mu_s^\eta)^T(\mathbf{V}_s^\eta)^{-1}x + (\mu_s^\eta)^T(\mathbf{V}_s^\eta)^{-1}\mu_s^\eta\\
&= x^T\mathbf{R}_{s,t}x - 2d(s,t,y_1,y_2)^Tx + z(s,t,y_1,y_2),
\end{aligned}$$

where

$$\mathbf{R}_{s,t} = 2\boldsymbol{\Psi}_{t-s}^T(\mathbf{V}_{t-s}+\mathbf{N}_{t-s})^{-1}\boldsymbol{\Psi}_{t-s} + (\mathbf{V}_s^\eta)^{-1}, \tag{4.15}$$

$$d(s,t,y_1,y_2) = \boldsymbol{\Psi}_{t-s}^T(\mathbf{V}_{t-s}+\mathbf{N}_{t-s})^{-1}(y_1+y_2-2\mathbf{K}_{t-s}) + (\mathbf{V}_s^\eta)^{-1}\mu_s^\eta \tag{4.16}$$

and

$$\begin{aligned}
z(s,t,y_1,y_2) &= (y_1-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_1-\mathbf{K}_{t-s})\\
&\quad + 2(y_1-\mathbf{K}_{t-s})^T\mathbf{X}_{t-s}(y_2-\mathbf{K}_{t-s})\\
&\quad + (y_2-\mathbf{K}_{t-s})^T(\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1})(y_2-\mathbf{K}_{t-s})\\
&\quad + (\mu_s^\eta)^T(\mathbf{V}_s^\eta)^{-1}\mu_s^\eta.
\end{aligned} \tag{4.17}$$

In the last step we have used the fact $2\mathbf{X}_{t-s}+\mathbf{M}_{t-s}^{-1} = (\mathbf{V}_{t-s}+\mathbf{N}_{t-s})^{-1}$. Thus by integrating out x we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} p_{t-s}^2(x,x;y_1,y_2)\int_{\mathbb{R}^d} p_s^1(u;x)g(u)\,du\,dx\\
&\qquad = (2\pi)^{-d}|\mathbf{V}_s^\eta|^{-1/2}|\widetilde{\mathbf{V}}_{t-s}|^{-1/2}|\mathbf{R}_{s,t}|^{-1/2}\exp\left\{-\tfrac{1}{2}w(s,t,y_1,y_2)\right\}
\end{aligned}$$

where

$$w(s,t,y_1,y_2) = z(s,t,y_1,y_2) - d(s,t,y_1,y_2)^T\mathbf{R}_{s,t}^{-1}d(s,t,y_1,y_2). \tag{4.18}$$

Finally, by collecting all the terms in (3.9) together, we obtain

$$\begin{aligned}
&\zeta_t(y_1,y_2)\\
&\quad = (2\pi)^{-d}\exp(2\lambda\gamma t)|\mathbf{V}_t^\zeta|^{-1/2}\exp\{-\tfrac{1}{2}(\tilde{y}-\mu_t^\zeta)^T(\mathbf{V}_t^\zeta)^{-1}(\tilde{y}-\mu_t^\zeta)\}\\
&\qquad + \lambda(2\pi)^{-d}\exp(2\lambda\gamma t)\int_0^t \exp(-\lambda\gamma s)|\mathbf{V}_s^\eta|^{-1/2}|\widetilde{\mathbf{V}}_{t-s}|^{-1/2}|\mathbf{R}_{s,t}|^{-1/2}\\
&\qquad\qquad \times \exp\{-\tfrac{1}{2}w(s,t,y_1,y_2)\}\,ds.
\end{aligned} \tag{4.19}$$

REMARK 3. If, as in Remark 2, we assume that the matrix $\mathbf{A}$ has the special diagonal form $\mathbf{A} = \operatorname{diag}(r_1, r_2, \ldots, r_d)$, we obtain the following simplified formula:

$$\mathbf{\Psi}_t = \operatorname{diag}(e^{r_1 t}, \ldots, e^{r_d t}), \tag{4.20}$$

$$\mathbf{K}_t = \left(a_1 \frac{e^{r_1 t} - 1}{r_1}, \ldots, a_d \frac{e^{r_d t} - 1}{r_d} \right)^T, \tag{4.21}$$

$$\mathbf{V}_t = \left(\frac{e^{(r_i + r_j)t} - 1}{r_i + r_j} \left(\sum_{k=1}^m c_{ik} c_{jk} + \delta_{ij} \varepsilon_i \varepsilon_j \right) \right)_{i,\, j=1,\ldots,d}, \tag{4.22}$$

$$\mathbf{N}_t = \left(\frac{e^{(r_i + r_j)t} - 1}{r_i + r_j} \left(\sum_{k=1}^m c_{ik} c_{jk} \right) \right)_{i,\, j=1,\ldots,d}, \tag{4.23}$$

$$\mathbf{M}_t = \operatorname{diag}\left(\varepsilon_1^2 \frac{e^{2r_1 t} - 1}{2r_1}, \ldots, \varepsilon_d^2 \frac{e^{2r_d t} - 1}{2r_d} \right). \tag{4.24}$$

If $r_i = 0$ for some i we replace $(e^{r_i t} - 1)/r_i$ by t in (4.21) and $(e^{2r_i t} - 1)/(2r_i)$ by t in (4.24). Similarly, if $r_i + r_j = 0$ for some i, j we replace $(e^{(r_i + r_j)t} - 1)/(r_i + r_j)$ by t in (4.22) and (4.23). These formulas, along with (4.5), (4.6), (4.8), (4.2.3), (4.11), (4.13), (4.15), (4.16), (4.17) and (4.18) enable us to fairly easily compute the covariance density ζ_t for certain choices of the model parameters. The final calculation involves a numerical integration for the time integral in (4.19).

4.3. *Graphs of the mean and the covariance densities.* The time has come to see what these formulas all mean and so here are some some graphs of the mean density η_t and the function θ_t [see (3.8) and (3.11)] for $d = 1$, and the mean density η_t for $d = 2$. We consider several choices for t and the model parameters. The matrix $\mathbf{A}$ is taken to be diagonal, as in the simplifying assumptions of Remarks 2 and 3.

In order to understand the graphs, one should first look at (4.3) and (4.4). Letting t increase in these two formulas we can see which area the mass moves to and how spread out it becomes in average as time evolves. In addition we could derive the rate at which this happens. Note that, as it follows from (4.7), two types of behavior are possible: the mass might move toward a fixed, finite region (i.e., mean reverting case) or move toward $\pm\infty$ and this might happen in any dimension independently. As one would expect, the graphs of the mean density differ between the three cases of no branching ($\lambda = 0$), critical branching ($\gamma = 0$) and supercritical branching ($\lambda\gamma > 0$), in the last instance due to nontrivial mass creation. The same is true for the function θ_t.

We do not have any graphs for the non-flow scenario. These will look much like those for the flow case, but without the motion towards the asymptotic mean evident in all the figures.

In Figure 1 we give six graphs of the mean density for $d = 1$. For all graphs, $\lambda = 1$ and $m = 3$. Otherwise, the parameters are as follows:

Graph 1: $\gamma = 0,\ \mu_0 = -4,\ \mathbf{V}_0 = 9,\ a = 5,\ r = -0.6,$
$\mathbf{C} = [0.8 \quad -1.5 \quad 0.5],\ \varepsilon = 2.$

Graph 2: $\gamma = 0,\ \mu_0 = 25,\ \mathbf{V}_0 = 12,\ a = -5,\ r = -0.25,$
$\mathbf{C} = [-0.4 \quad 1.5 \quad 0.5],\ \varepsilon = 5.$

Graph 3: $\gamma = 0.1,\ \mu_0 = -10,\ \mathbf{V}_0 = 3,\ a = 5,\ r = -1,$
$\mathbf{C} = [1 \quad -0.5 \quad 0],\ \varepsilon = 1.$

Graph 4: $\gamma = 0.01,\ \mu_0 = -5,\ \mathbf{V}_0 = 9,\ a = -6,\ r = -0.25,$
$\mathbf{C} = [-0.4 \quad 1.5 \quad 2.5],\ \varepsilon = 4.$

Graph 5: $\gamma = 0.02,\ \mu_0 = 5,\ \mathbf{V}_0 = 16,\ a = -5,\ r = -1,$
$\mathbf{C} = [-1.5 \quad 2.5 \quad 3.5],\ \varepsilon = 1.$

Graph 6: $\gamma = 0.05,\ \mu_0 = 0,\ \mathbf{V}_0 = 25,\ a = 9,\ r = -0.2,$
$\mathbf{C} = [1 \quad -2 \quad 4],\ \varepsilon = 1.$

In all graphs in Figure 1 the flows are mean reverting; that is, the mass moves, in expectation, toward the asymptotic mean as time evolves. In the first two graphs the branching is either absent or critical. In the last four the branching is supercritical. Observe that the system exhibits several patterns of behavior depending on the parameters. In particular, the mass might move from the left to the right ($a > 0$) or vice versa ($a < 0$) and the variance might increase or decrease. Note also the crucial effect of the supercritical branching.

In Figure 2 we give eight graphs of the mean density for $d = 2$ at times $t = 0$, $t = 1$, $t = 4$ and $t = 10$ in the case of supercritical branching. We use the following values for the parameters: $\lambda = 1$, $\gamma = 0.15$, $m = 3$,

$$\mu_0 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \qquad \mathbf{V}_0 = \begin{bmatrix} 1 & 0.81 \\ 0.81 & 4 \end{bmatrix}, \qquad a = \begin{bmatrix} -0.5 \\ 1 \end{bmatrix}, \qquad r = \begin{bmatrix} -0.25 \\ -0.125 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1.1 & 0.5 & -1.4 \\ -1.3 & 0.1 & 0.2 \end{bmatrix} \quad \text{and} \quad \varepsilon = \begin{bmatrix} 0.3 \\ 0.2 \end{bmatrix}.$$

The third column in Figure 2 contains the corresponding contour plots while the first column contains the graphs as viewed from the lower left corner of the contour plot and the second column contains the graphs as viewed from the lower right corner of the contour plot. In this graph observe that the mass moves, on average, towards to an area around $(-2, 8)$ as time evolves. The mean total mass increases and is spread out as a result of supercritical branching. If the branching were critical the mass would shrink to smaller and smaller areas around $(-2, 8)$.

In Figures 3 and 4 we present six graphs of the function θ_t for $d = 1$ at times $t = 0.5$ and $t = 2.5$. The graphs include the cases of no branching

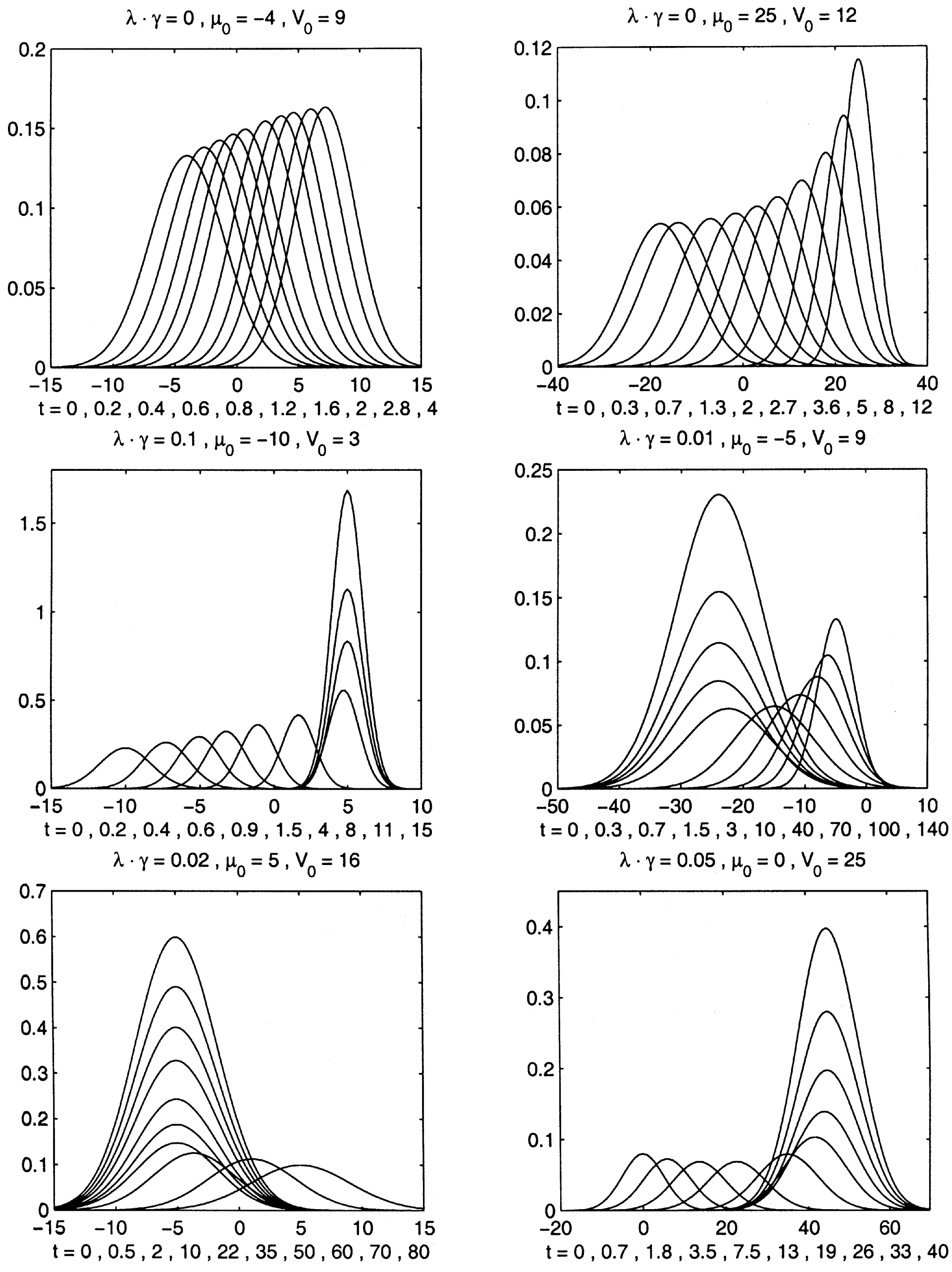

FIG. 1. *Graphs of the mean density η_t for $d = 1$.*

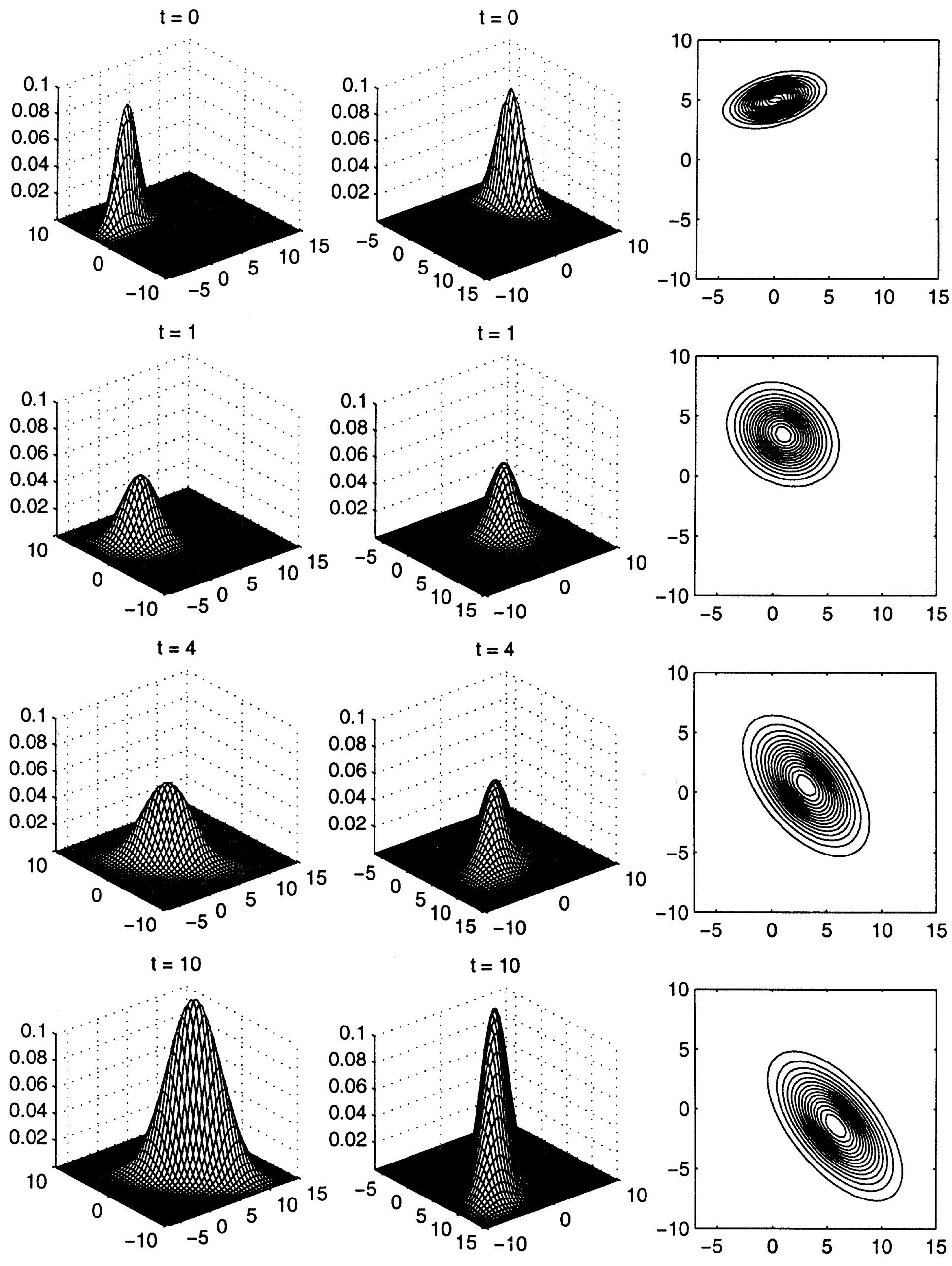

FIG. 2. *Graphs of the mean density η_t for $d = 2$ and supercritical branching.*

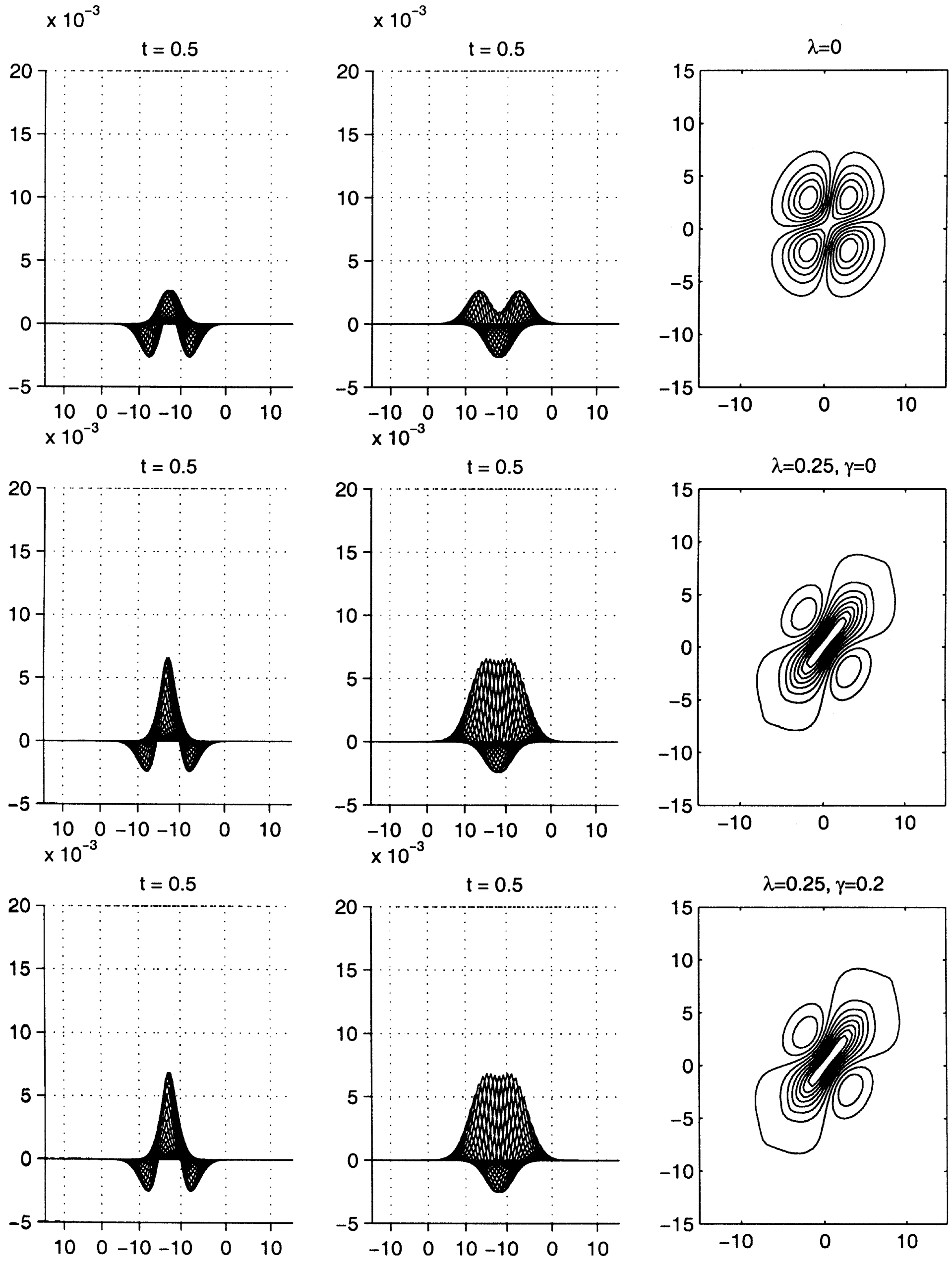

FIG. 3. *Graphs of the function θ_t for $d = 1$.*

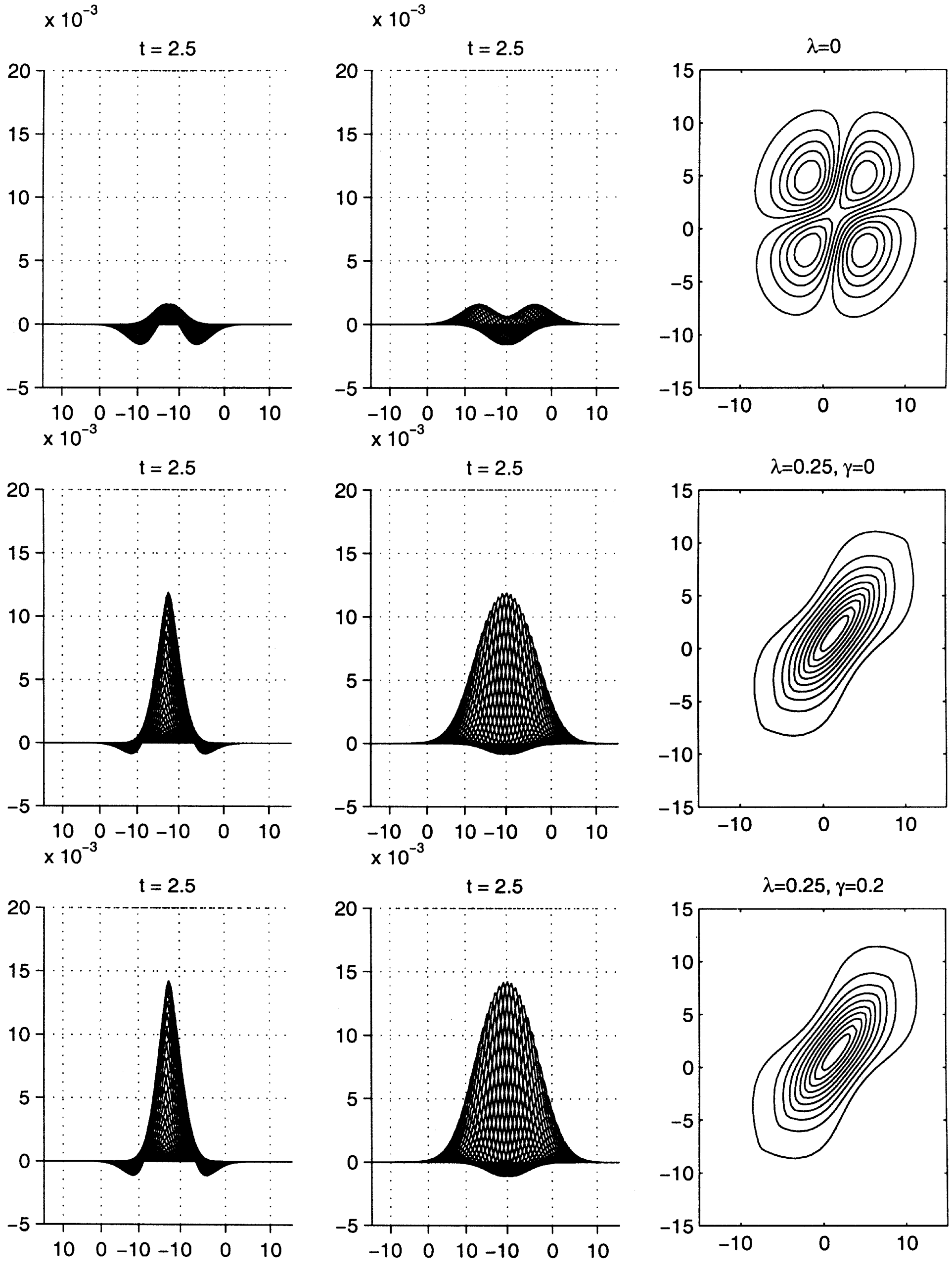

FIG. 4. *Graphs of the function* θ_t *for* $d = 1$.

($\lambda = 0$), of critical branching in which we take $\lambda = 0.25$ and $\gamma = 0$ and of supercritical branching for which we use $\lambda = 0.25$ and $\gamma = 0.2$. In addition, in all three cases, we use the following values for the rest of the parameters:

$$\mu_0 = 0, \quad \mathbf{V}_0 = 1.5, \quad a = 1, \quad r = -0.5, \quad \mathbf{C} = 2, \quad \varepsilon = 3.$$

The first row in each of Figures 3 and 4 contains the graphs of the function θ_t that correspond to the case of no branching, the second row contains the graphs of the function θ_t that correspond to the case of critical branching, and the third row contains the graphs of the function θ_t that correspond to the case of supercritical branching. The third column in each of these figures, which is the easiest column to understand, contains the corresponding contour plots. What is not evident from these plots is that there are regions over which θ takes negative values. To see this, imagine walking along the axes in this column, starting at $(-15, 15)$, walking down to $(-15, -15)$ and then along to $(-15, 15)$, all the time at the zero level of the function θ and looking towards the function. What one sees appears in the first column. Doing the same, but walking from $(-15, 15) \to (15, 15) \to (15, -15)$ gives the second column.

Note that there is limited variability in the case of no branching compared to the cases of branching, critical or supercritical. As expected, supercritical branching yields higher variability than critical branching. Furthermore, disjoint intervals close to the center of mass have negative correlation in the case of no branching. This is also observed in the case of branching for small t. However, the negative correlation decreases as time evolves due to the effect of branching.

5. On random environments. A natural way to think of the process that we have been working with is to condition on the stochastic flow in the equation (cf. 2.8) defining the one-point motions,

$$\begin{aligned} dY_i(t) = b_i(Y(t))\,dt + \sum_{l=1}^{m} c_{il}(Y(t))\,dW_l(t) \\ + e_i(Y(t))\,dB_i^{\alpha}(t), \qquad i = 1, 2, \ldots, d, \end{aligned}$$

that is, to condition on the Brownian motions W_l and then to think of the resulting limit process as a superprocess built over the diffusion with drift $b_i(Y(t))\,dt + \sum_{l=1}^{m} c_{il}(Y(t))\,dW_l(t)$. In fact, when simulating the particle picture, so that time and space become discrete, this is exactly what one does.

Such a direct approach, is not, however, plausible, since the "drift term" coming from the flow is not smooth enough (in t) to be acceptable as the drift of a diffusion process.

One way around this would be to approximate the global Brownian motion W by a smooth (differentiable in t) process W^{ε} (which converges a.s. to W as $\varepsilon \to 0$) with time derivative denoted by V^{ε}. Then, in the spirit of the Wong–Zakai approximation, we could approximate the one-point motions by the solution of the SDE,

$$dY_i(t) = \left(\bar{b}_i(Y(t)) + \sum_{l=1}^{m} V_l^{\varepsilon}(t) c_{il}(Y(t))\right) dt$$
$$+ e_i(Y(t))\, dB_i^{\alpha}(t), \qquad i = 1, 2, \dots, d,$$

construct the corresponding (time-inhomogeneous) superprocess and then lift the smoothness on the superprocess by sending $\varepsilon \to 0$, taking care of the Stratonovich–Itô correction term en passant. Results that would come out of this approach would presumably hold for almost every flow W.

Since we are interested only in distributional results, there is also an averaging over W to be done at some stage, which can be done either before we send $\varepsilon \to 0$, or after. We had little success either way, but it may be instructive for the reader to sketch the argument and point out where the difficulties lie in the more interesting of these approaches, that of sending $\varepsilon \to 0$ and then averaging over the flow. Details can be found in the thesis in [24].

For fixed $\varepsilon > 0$ one can apply the theory of nonstationary superprocesses (cf. [10]) to construct a measure-valued diffusion whose function-valued dual solves the SDE,

$$y_r(x) = f(x) + \int_0^r \left[\sum_{i=1}^{d} \bar{b}_i(x) y_v^i(x) + \frac{1}{2}\sum_{i=1}^{d} e_i^2(x) y_v^{ii}(x) - y_v^2(x)\right] dv$$
$$+ \sum_{l=1}^{m} \int_0^r N_l(y_v)(x) V_l^{\varepsilon}(s+t-v)\, dv,$$

for all $x \in \mathbb{R}^d$, where $y_0 \equiv f$, $\bar{b}_i = b_i - \frac{1}{2}\sum_{j=1}^{d}\sum_{l=1}^{m} c_{il}^j c_{jl}$ and $N_l(f) = \sum_{i=1}^{d} c_{il} f^i$, $l = 1, \dots, m$.

This is straightforward to do in a rigorous fashion. Now send $\varepsilon \to 0$. Then, using the usual Wong–Zakai correction term (see Section 3 in [25]) we expect to find (after some work, some algebra and some convenient cancellations) a corresponding limiting dual that satisfies

$$y_r(x) = f(x) + \int_0^r \big((Ly_v)(x) - y_v^2(x)\big)\, dv$$
$$+ \sum_{l=1}^{m} \int_0^r (N_l y_v)(x)\, d\widetilde{W}_v^l, \quad x \in \mathbb{R}^d, \tag{5.1}$$

where $\widetilde{W}$ is an m-dimensional Brownian motion, and L is defined by (2.15). As usual, this equation is linked to a martingale problem and finding a solution of (5.1) and solving a martingale problem amount to more or less the same thing. Either way, we would have established our "superprocess in a random environment" for a fixed environment.

There are, however, two problems with the procedure we just described. The first and probably more serious one is to rigorously establish the passage to the Wong–Zakai limit above as $\varepsilon \to 0$. The second problem lies in solving (5.1). While it is reminiscent of similar equations in [5] and [23], it fits into none of these set-ups and we could go no further. The main reason for both problems seems to be the quadratic term in the drift of (5.1). Had we been able to proceed in the described fashion, then averaging over W would yield the kind of results that we have in this paper.

Had this route been successful, we would also have an approach for establishing results for superprocesses over flows that would hold almost every (fixed) flow, something which is beyond the tools of our current approach.

While this lack of progress on our part certainly does not imply that this path is doomed (rather, it opens up an interesting challenge) it does indicate that this seemingly more direct approach to superprocesses over flows is not likely to be more "direct" than the one taken in this paper for the kind of results we have proven.

APPENDIX

A. Proof of the weak convergence. The main result stated in Theorem 2.2.1 follows by combining Theorem A.1.1, which is about tightness and is proved in the first subsection, with Theorem A.4.1 which is about uniqueness and is the subject of the second subsection.

A.1. *Tightness.* In this subsection we will prove the following theorem, in which X^n is as defined by (2.11) and L, Λ are defined by (2.15) and (2.18) at the end of Section 2.1.

THEOREM A.1.1. *Assume that* $X_0^n = \frac{1}{n}\nu_n \Rightarrow \nu$ *in* $M_F(\mathbb{R}^d)$. *Then the sequence* $\{X^n\}$ *is tight in* $D_{M_F(\mathbb{R}^d)}[0,\infty)$, *each weak limit point* X *is in* $C_{M_F(\mathbb{R}^d)}[0,\infty)$ *and satisfies the following martingale problem:*

For all $f \in \mathscr{D}$,

$$Z_t(f) = X_t(f) - \nu(f) - \int_0^t X_s(Lf)\,ds - \xi \int_0^t X_s(f)\,ds \tag{A.1}$$

is a continuous square integrable $\{\mathscr{F}_t^X\}$*-martingale such that*

$$Z_0(f) = 0 \text{ and } \langle Z(f)\rangle_t = \delta \int_0^t X_s(f^2)\,ds + \int_0^t (X_s \times X_s)(\Lambda f)\,ds,$$

where $\xi = \lambda\gamma$ *and* $\delta = \lambda\sigma^2$.

A.2. *General tools.* First we state two theorems that are the basic tools used in the derivation of both Theorem A.1.1 and the analogous result to Theorem A.1.1 in the regular superprocess case (see [20]).

The first theorem gives the construction of a continuous time martingale from a sequence of discrete time martingales by passing to a limit and provides information regarding tightness and quadratic variation of the limiting process.

THEOREM A.2.1 ([20], Lemma II.4.5). *Let* $\{(M_k^{(n)}, \mathscr{F}_k^{(n)})\colon k \in \mathbb{N}\}, n = 1, 2, \dots$ *be a sequence of discrete time square integrable martingales and* $\{\lambda_n\}$ *be a sequence of positive numbers such that* $\lambda_n \uparrow \infty$ *as* $n \to \infty$. *Define*

$$X_t^{(n)} = M_{[\lambda_n t]}^{(n)}, \qquad t \geq 0$$

and let

$$\langle X^{(n)} \rangle_t = \sum_{j=1}^{[\lambda_n t]} E\Big((M_j^{(n)} - M_{j-1}^{(n)})^2 | \mathscr{F}_{j-1}^{(n)}\Big) + E\Big(M_0^{(n)}\Big)^2, \qquad t \geq 0.$$

(a) *If*

$$\big\{\langle X^{(n)} \rangle\colon n = 1, 2, \dots\big\} \textit{ is a C-tight sequence in } D_{\mathbb{R}}[0, \infty)$$

and

$$\sup_{1 \leq k \leq \lambda_n J} |M_k^{(n)} - M_{k-1}^{(n)}| \longrightarrow 0 \textit{ in probability} \quad \textit{ for all } J \in \mathbb{N}^*,$$

then

$$\big\{X^{(n)}\colon n = 1, 2, \dots\big\} \textit{ is a C-tight sequence in } D_{\mathbb{R}}[0, \infty).$$

(b) *If, in addition to the assumptions in part* (a) *the family*

$$\left\{\sup_{1 \leq k \leq \lambda_n J} M_k^{(n)}\colon n = 1, 2, \dots\right\} \textit{ is uniformly integrable, for all } J \in \mathbb{N}^*$$

and if

$$X^{(n_k)} \Rightarrow X \ \big[\textit{i.e., } X \textit{ is a limit point of } (X^{(n)})\big],$$

then X *is a continuous square integrable martingale and*

$$\big\langle X^{(n_k)} \big\rangle \Rightarrow \big\langle X \big\rangle + E({X_0}^2) \quad \textit{as } k \to \infty.$$

The second theorem describes how the task of proving tightness for a sequence of measure-valued processes can be reduced to the easier task of proving tightness for sequences of real-valued processes. Let K be a compact Polish space and $\{X^n\}$ be a sequence of $M_F(K)$-valued processes. Then, if we can prove that the sequence of real-valued processes $\{X^n(f)\}$ is tight for all f in an appropriate class of functions, the tightness of the sequence $\{X^n\}$ generally follows. This idea is contained in the following theorem. We note that versions of this theorem have been used by several authors in similar settings. See, for example, Theorem 3.7.1 in [6] or Theorem 2.1 in [21].

THEOREM A.2.2. *Assume* K *is a compact Polish space. Let* S *be a countable set dense in* $C(K)$ *such that the constant function* $1 \in S$ *and let* $\{X^n\}$ *be a sequence of* $M_F(K)$*-valued processes. Then* $\{X^n\}$ *is tight in* $D_{M_F(K)}[0, \infty)$ *and all limit points are in* $C_{M_F(K)}[0, \infty)$ (*i.e.,* $\{X^n\}$ *is C-tight*) *if and only if, for each* $f \in S$, *the sequence of real-valued process* $\{X^n(f)\}$ *is tight in* $D_{\mathbb{R}}[0, \infty)$ *and all limit points are in* $C_{\mathbb{R}}[0, \infty)$ (*i.e.,* $\{X^n(f)\}$ *is C-tight*).

A.3. *Proof of Theorem* A.1.1. The proof consists of several steps. First we consider a sufficiently large class of functions and then we fix a member f of this class. The objective, then, is to prove C-tightness of the sequence $\{X^n(f)\}$ so we can use Theorem A.2.2 to obtain C-tightness of the sequence $\{X^n\}$. For this, we provide a decomposition of $\{X^n(f)\}$ and prove C-tightness by proving C-tightness for the several terms in the decomposition and using Theorem A.2.1 and standard results about tightness. Then we pass to the limit, along converging subsequences of $\{X^n\}$, describe the limit points of $\{X^n(f)\}$ as semimartingales and give the quadratic variation of the martingale part. Finally, an application of Theorem A.2.2 yields the result of C-tightness for $\{X^n\}$. In order to proceed we need some notation. We denote by $I_{\mathbb{R}^d}$ the indicator function of $\mathbb{R}^d$, by $\hat{g}$ the extension of g to $\overline{\mathbb{R}}$ such that $\hat{g}(\Delta) = 0$ and by $\hat{h}$ the extension of $h \in C_b(\mathbb{R}^d \times \mathbb{R}^d)$ to $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ such that $\hat{h}(x, y) = 0$ if $x = \Delta$ or $y = \Delta$. Let k be a nonnegative integer and $f \in \mathscr{D}$. Then, for $\alpha \sim_n k_n$, by applying Itô's formula to (2.8) we obtain, for $t \in [k_n, k_n + a_n]$, that

$$
\begin{aligned}
f(Y_t^{\alpha,n}) = f(Y_{k_n}^{\alpha,n}) &+ \sum_{i=1}^d \int_{k_n}^t f^i(Y_u^{\alpha,n}) b_i(Y_u^{\alpha,n})\,du \\
&+ \sum_{i=1}^d \int_{k_n}^t f^i(Y_u^{\alpha,n}) e_i(Y_u^{\alpha,n})\,dB_i^{\alpha,n}(u) \\
&+ \sum_{i=1}^d \sum_{l=1}^m \int_{k_n}^t f^i(Y_u^{\alpha,n}) c_{il}(Y_u^{\alpha,n})\,dW_l^n(u) \\
&+ \tfrac{1}{2} \sum_{i=1}^d \int_{k_n}^t f^{ii}(Y_u^{\alpha,n}) e_i^2(Y_u^{\alpha,n})\,du \\
&+ \tfrac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \sum_{l=1}^m \int_{k_n}^t f^{ij}(Y_u^{\alpha,n}) c_{il}(Y_u^{\alpha,n}) c_{jl}(Y_u^{\alpha,n})\,du.
\end{aligned}
$$

Thus $f(Y_t^{\alpha,n}) - f(Y_{k_n}^{\alpha,n}) - \int_{k_n}^t (Lf)(Y_u^{\alpha,n})\,du$ is a martingale with respect to the filtration $\{\mathscr{F}_t^n\}$ in the interval $[k_n, k_n + a_n]$ where L is the operator defined by (2.15). For $t \in [k_n, k_n + a_n]$ and $\alpha \sim_n k_n$ we define

$$
M_t^{\alpha, k_n}(f) = \begin{cases} f(Y_t^{\alpha,n}) - f(Y_{k_n}^{\alpha,n}) - \displaystyle\int_{k_n}^t (Lf)(Y_u^{\alpha,n})\,du, & \text{if } X_{k_n}^{\alpha,n} \neq \Delta, \\ 0, & \text{if } X_{k_n}^{\alpha,n} = \Delta. \end{cases}
$$

Then, clearly, we have that $\{(M_t^{\alpha,k_n}(f), \mathscr{F}_t^n)\colon t \in [k_n, k_n + a_n]\}$ is a martingale for each $k \in \mathbb{N}$ and $\alpha \sim_n k_n$. Let $r \in \mathbb{N}$. Then

$$
\begin{aligned}
&X_{r_n + a_n}^n(f) - X_{r_n}^n(f) \\
&\quad = n^{-1} \sum_{\alpha \sim_n r_n} I_{\mathbb{R}^d}(X_{r_n}^{\alpha,n}) \Big(f(Y_{r_n + a_n}^{\alpha,n}) N^{\alpha,n} - f(Y_{r_n}^{\alpha,n}) \Big)
\end{aligned}
$$

$$= n^{-1} \sum_{\alpha \sim_n r_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n}) \Big(f(Y^{\alpha,n}_{r_n+a_n}) - f(Y^{\alpha,n}_{r_n}) \Big) N^{\alpha,n}$$

$$+ n^{-1} \sum_{\alpha \sim_n r_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n}) f(Y^{\alpha,n}_{r_n})(N^{\alpha,n} - 1)$$

$$= n^{-1} \sum_{\alpha \sim_n r_n} M^{\alpha, r_n}_{r_n+a_n}(f) N^{\alpha,n} + n^{-1} \sum_{\alpha \sim_n r_n} \Big[\int_{r_n}^{r_n+a_n} (\widehat{Lf})(X^{\alpha,n}_u)\, du \Big] N^{\alpha,n}$$

$$+ n^{-1} \sum_{\alpha \sim_n r_n} \hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n} - \beta_n) + (\beta_n - 1) n^{-1} \sum_{\alpha \sim_n r_n} \hat{f}(X^{\alpha,n}_{r_n})$$

$$= n^{-1} \sum_{\alpha \sim_n r_n} M^{\alpha, r_n}_{r_n+a_n}(f)(N^{\alpha,n} - \beta_n)$$

$$+ n^{-1} \sum_{\alpha \sim_n r_n} \Big[\int_{r_n}^{r_n+a_n} (\widehat{Lf})(X^{\alpha,n}_u)\, du \Big] (N^{\alpha,n} - \beta_n)$$

$$+ n^{-1} \sum_{\alpha \sim_n r_n} \Big[\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n} - \beta_n) + \beta_n M^{\alpha, r_n}_{r_n+a_n}(f) \Big]$$

$$+ \beta_n n^{-1} \sum_{\alpha \sim_n r_n} \int_{r_n}^{r_n+a_n} (\widehat{Lf})(X^{\alpha,n}_u)\, du + (\beta_n - 1) n^{-1} \sum_{\alpha \sim_n r_n} \hat{f}(X^{\alpha,n}_{r_n}).$$

Moreover, if $r_n \le t < r_n + a_n$, then

$$X^n_t(f) - X^n_{r_n}(f) = n^{-1} \sum_{\alpha \sim_n r_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n}) \Big(f(Y^{\alpha,n}_t) - f(Y^{\alpha,n}_{r_n}) \Big)$$

$$= n^{-1} \sum_{\alpha \sim_n r_n} M^{\alpha, r_n}_t(f) + n^{-1} \sum_{\alpha \sim_n r_n} \int_{r_n}^t (\widehat{Lf})(X^{\alpha,n}_u)\, du.$$

Hence if, for $t \in [k_n, k_n + a_n),\ k = 0, 1, 2 \ldots$, we define

$$M^{(n)}_t(f) = n^{-1} \sum_{r<k} \sum_{\alpha \sim_n r_n} M^{\alpha, r_n}_{r_n+a_n}(f)(N^{\alpha,n} - \beta_n),$$

$$J^{(n)}_t(f) = n^{-1} \sum_{\alpha \sim_n k_n} M^{\alpha, k_n}_t(f) + (1 - \beta_n) n^{-1} \sum_{\alpha \sim_n k_n} \int_{k_n}^t (\widehat{Lf})(X^{\alpha,n}_u)\, du,$$

$$N^{(n)}_t(f) = n^{-1} \sum_{r<k} \sum_{\alpha \sim_n r_n} \Big[\int_{r_n}^{r_n+a_n} (\widehat{Lf})(X^{\alpha,n}_u)\, du \Big] (N^{\alpha,n} - \beta_n),$$

$$Z^{(n)}_t(f) = n^{-1} \sum_{r<k} \sum_{\alpha \sim_n r_n} \Big[\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n} - \beta_n) + \beta_n M^{\alpha, r_n}_{r_n+a_n}(f) \Big],$$

$$C^{(n)}_t(f) = \int_0^t X^n_s(Lf)\, ds,$$

$$H^{(n)}_t(f) = n^{-2} \sum_{r<k} \sum_{\alpha \sim_n r_n} \hat{f}(X^{\alpha,n}_{r_n}) = \lambda \int_0^{k_n} X^n_{[\lambda n s]_n}(f)\, ds,$$

then we have

$$X_t^n(f) = X_0^n(f) + M_t^{(n)}(f) + J_t^{(n)}(f) + N_t^{(n)}(f) + Z_t^{(n)}(f) + \beta_n C_t^{(n)}(f) + \gamma_n H_t^{(n)}(f). \tag{A.3}$$

We will show that $M_{k_n}^{(n)}(f)$, $N_{k_n}^{(n)}(f)$ and $Z_{k_n}^{(n)}(f)$ are $\{\mathscr{F}_{k_n}^n\}$-martingales. $M^{(n)}(f)$ and $N^{(n)}(f)$ are related to the motion of the particles, and $Z^{(n)}(f)$ is related to the branching and the stochastic component of the flow. Note that $M^{(n)}(f)$, $N^{(n)}(f)$, and $Z^{(n)}(f)$ are merely cadlag extensions of discrete time martingales considered at times $k(\lambda n)^{-1}$, $k \in \mathbb{N}$.

The next lemma, which is related to the total mass process $X^n(1)$, will be used extensively in the sequel in proving C-tightness for a number of processes. Its proof follows by using, in a straightforward fashion, standard facts from the theory of branching processes (see [13], Chapter 1, Sections 5 and 8) and Theorem 21.1 in [4].

LEMMA A.3.1. *For each $T > 0$,*

$$C'_T = \sup_{n\geq 1} E \sup_{0\leq t\leq T} (X_t^n(1))^p < \infty$$

and therefore

$$C_T = \sup_{n\geq 1} E \sup_{0\leq t\leq T} (X_t^n(1))^2 < \infty,$$

where p is the number satisfying (2.5).

Using the previous lemma and the fact that the motion of the particles between time points k_n and $k_n + a_n$ is a diffusion, we can prove the following lemma.

LEMMA A.3.2. *For all $f \in \mathscr{D}$, $\{(M_{k_n}^{(n)}(f), \mathscr{F}_{k_n}^n)\colon k = 0, 1, \dots\}$ and $\{(N_{k_n}^{(n)}(f), \mathscr{F}_{k_n}^n)\colon k = 0, 1, \dots\}$ are martingales, and for all $T > 0$,*

$$\lim_{n\to\infty} E \sup_{0\leq t\leq T} (M_t^{(n)}(f))^2 = 0 \quad \textit{and} \quad \lim_{n\to\infty} E \sup_{0\leq t\leq T} (N_t^{(n)}(f))^2 = 0.$$

In addition, for all $f \in \mathscr{D}$ and $T > 0$,

$$\lim_{n\to\infty} E \sup_{0\leq t\leq T} (J_t^{(n)}(f))^2 = 0. \tag{A.4}$$

PROOF. We only provide the proof of the last equality. First we note that, for $k_n \leq t < k_n + a_n$, $\quad k = 0, 1, \dots$, we have

$$\left(J_t^{(n)}(f)\right)^2 \leq 2\left(n^{-1} \sum_{\alpha\sim_n k_n} M_t^{\alpha, k_n}(f)\right)^2 + 2\left(\gamma_n n^{-1}(\lambda n)^{-1} B_f X_t^n(1)\right)^2$$

and

$$n^{-2}E\sup_{0\le t\le T}\big(X_t^n(1)\big)^2\le n^{-2}C_T\to 0\quad \text{as } n\to\infty.$$

Hence, if we let $r(n)=[\lambda nT]$ and

$$Q_t^{n,k}(f)=n^{-1}\sum_{\alpha\sim_n k_n}M_t^{\alpha,k_n}(f) \tag{A.5}$$

for $k_n\le t\le k_n+a_n$, $\quad k=0,1,\dots$, it suffices to show

$$\lim_{n\to\infty}E\max_{0\le k\le r(n)}\sup_{k_n\le t\le k_n+a_n}(Q_t^{n,k}(f))^p=0,$$

where p is the number satisfying (2.5). It follows that $\{(Q_t^{n,k}(f),\mathscr{F}_t^n):t\in[k_n,k_n+a_n]\}$ is a martingale, and a simple computation yields that

$$\begin{aligned}\langle Q^{n,k}(f)\rangle_t&=n^{-1}\int_{k_n}^t X_u^n(\Psi f)\,du\\&\qquad+\int_{k_n}^t(X_u^n\times X_u^n)(\Lambda f)\,du-n^{-1}\int_{k_n}^t\int_{\mathbb{R}^d}(\Lambda f)(x,x)X_u^n(dx)\,du\end{aligned}$$

for $k_n\le t\le k_n+a_n$ and thus

$$\langle Q^{n,k}(f)\rangle_{k_n+a_n}\le 2B_f(\lambda n)^{-1}\Big[n^{-1}\sup_{0\le t\le T+1}X_t^n(1)+\sup_{0\le t\le T+1}(X_t^n(1))^2\Big]$$

for all $k=0,1,\dots,r(n)$. Hence, by Theorem 42.1 in [22] and Minkowski's inequality, we have

$$\begin{aligned}&E\max_{0\le k\le r(n)}\sup_{k_n\le t\le k_n+a_n}(Q_t^{n,k}(f))^p\\&\quad\le\sum_{k=0}^{r(n)}E\sup_{k_n\le t\le k_n+a_n}(Q_t^{n,k}(f))^p\\&\quad\le\sum_{k=0}^{r(n)}c_pE\Big(\langle Q^{n,k}(f)\rangle_{k_n+a_n}\Big)^{p/2}\\&\quad\le(r(n)+1)c_p(2B_f(\lambda n)^{-1})^{p/2}\\&\qquad\times\bigg[n^{-1}\Big(E\sup_{0\le t\le T+1}(X_t^n(1))^{p/2}\Big)^{2/p}+\Big(E\sup_{0\le t\le T+1}(X_t^n(1))^p\Big)^{2/p}\bigg]^{p/2}\\&\quad\le\frac{r(n)+1}{n^{p/2}}c_p(2B_f\lambda^{-1})^{p/2}\Big[n^{-1}(C'_{T+1})^{1/p}+(C'_{T+1})^{2/p}\Big]^{p/2},\end{aligned}$$

which converges to 0 as $n\to\infty$, since $p>2$. Here c_p is a positive constant depending on p. The proof of the last equality is then complete. The martingale property of the two sequences follows in a straightforward fashion from the structure of the model. Moreover, using the Markov property of the motion of the particles, the strong continuity of the semigroup T_t (see 2.13) and Lemma A.3.1 we can derive the first and the second equality. □

From Lemma A.3.2 we can immediately obtain, using standard facts about tightness and convergence, the following.

LEMMA A.3.3. *For all $f \in \mathscr{D}$, the sequences $\{M^{(n)}(f): n \in \mathbb{N}^*\}$, $\{N^{(n)}(f): n \in \mathbb{N}^*\}$, and $\{J^{(n)}(f): n \in \mathbb{N}^*\}$ are tight in $D_{\mathbb{R}}[0,\infty)$ and $M^{(n)}(f) \Rightarrow \mathbf{0}$, $N^{(n)}(f) \Rightarrow \mathbf{0}$, and $J^{(n)}(f) \Rightarrow \mathbf{0}$, where $\mathbf{0}$ is the zero process in $D_{\mathbb{R}}[0,\infty)$.*

LEMMA A.3.4. *For all $f \in \mathscr{D}$, the sequences $\{C^{(n)}(f): n = 1, 2, \ldots\}$ and $\{H^{(n)}(f): n = 1, 2, \ldots\}$ are C-tight in $D_{\mathbb{R}}[0,\infty)$.*

PROOF. Only the proof for the second sequence is given since the proof for the first one follows in a similar, and actually easier, fashion. We will use Proposition 3.26 of Chapter VI in [14]. Let $N \in \mathbb{N}^*$, $\varepsilon > 0$, and $\eta > 0$. First note that

$$\sup_{0\le t\le N} \left|H_t^{(n)}(f)\right| \le \lambda N B_f \sup_{0\le t\le N} X_t^n(1)$$

and so

$$P\left(\sup_{0\le t\le N} \left|H_t^{(n)}(f)\right| > K\right) \le P\left(\lambda^2 N^2 B_f^2 \sup_{0\le t\le N} (X_t^n(1))^2 > K^2\right) \le \varepsilon$$

for all $n \ge 1$ and $K \ge \lambda N B_f \sqrt{C_N/\varepsilon}$ by Chebyshev's inequality and Lemma A.3.1. Now let $0 \le u \le s \le t \le u+\theta \le N$ where θ is to be determined and define $k = [\lambda n s]$, $l = [\lambda n t]$. Then

$$\left|H_t^{(n)}(f) - H_s^{(n)}(f)\right| \le n^{-1} B_f \sum_{k\le r<l} X_{r_n}^n(1) \le n^{-1} B_f (l-k) \sup_{0\le u\le t} X_u^n(1).$$

Next we observe that $n^{-1}(l-k) = \lambda(l_n - k_n) \le \lambda(\theta + a_n) = \lambda\theta + n^{-1}$ and so for any choice of θ we can find $n_0(\theta) \in \mathbb{N}^*$ such that $n^{-1}(l-k) \le 2\lambda\theta$ for all $n \ge n_0(\theta)$. Then, using the notation in Proposition 3.26 of Chapter VI in [14], we have

$$w_N(H^{(n)}(f), \theta) \le 2\lambda\theta B_f \sup_{0\le t\le N} X_t^n(1)$$

for all $n \ge n_0(\theta)$, from we which we conclude

$$P\big(w_N(H^{(n)}(f), \theta) > \eta\big) \le \varepsilon$$

for all $\theta > 0$ such that $\theta B_f \le (2\lambda)^{-1}\eta\sqrt{\varepsilon/C_N}$ and $n \ge n_0(\theta)$ again by Chebyshev's inequality and Lemma A.3.1. Applying Proposition 3.26 of Chapter VI in [14] yields the conclusion. □

The next lemma is the most important in this proof since it explains where the second term in the quadratic variation in Theorem A.1.1 comes from and also describes the difference between our model and the regular superprocess. For this reason we give a detailed proof.

LEMMA A.3.5. *For all $f \in \mathscr{D}$ and for all $n = 1, 2, \dots$, $\{(Z^{(n)}_{k_n}(f), \mathscr{F}^n_{k_n}): k = 0, 1, 2, \dots,\}$ is a square integrable discrete time martingale with quadratic variation*

$$\begin{aligned}\langle Z^{(n)}(f)\rangle_{k_n} &= \lambda\sigma_n^2 \int_0^{k_n} X^n_{[\lambda ns]_n}(f^2)\,ds \\ &\quad + \beta_n^2 \int_0^{k_n} \int_{\mathbb{R}^d\times\mathbb{R}^d} \left(a_n^{-1}\int_0^{a_n} (S_u(\Lambda f))(x, y)\,du\right) \\ &\quad \times X^n_{[\lambda ns]_n}(dx) X^n_{[\lambda ns]_n}(dy)\,ds \\ &\quad + n^{-1}\beta_n^2 \int_0^{k_n}\int_{\mathbb{R}^d}\left(a_n^{-1}\int_0^{a_n}\Big[(T_u(\Psi f))(x) \right. \\ &\quad \left. - (S_u(\Lambda f))(x, x)\Big]\,du\right) X^n_{[\lambda ns]_n}(dx)\,ds,\end{aligned} \tag{A.6}$$

where T_t, S_t are the semigroup operators defined by (2.13) *and* (2.14)*, and Λ and Ψ are the operators defined by* (2.18) *and* (2.19).

PROOF. Let $k \in \mathbb{N}$. Then

$$\begin{aligned}&E\Big(Z^{(n)}_{k_n+a_n}(f) - Z^{(n)}_{k_n}(f)\,\big|\,\mathscr{F}^n_{k_n}\Big) \\ &\quad = n^{-1}\sum_{\alpha\sim_n k_n} E\Big(\Big[\hat{f}(X^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n) + \beta_n M^{\alpha,k_n}_{k_n+a_n}(f)\Big]\,\Big|\,\mathscr{F}^n_{k_n}\Big) \\ &\quad = n^{-1}\sum_{\alpha\sim_n k_n}\Big[\hat{f}(X^{\alpha,n}_{k_n})E(N^{\alpha,n} - \beta_n|\mathscr{F}^n_{k_n}) + \beta_n E(M^{\alpha,k_n}_{k_n+a_n}(f)|\mathscr{F}^n_{k_n})\Big],\end{aligned}$$

which equals 0, since $E(N^{\alpha,n} - \beta_n|\mathscr{F}^n_{k_n}) = E(N^{\alpha,n} - \beta_n) = 0$ and $\{(M^{\alpha,k_n}_t(f), \mathscr{F}^n_t): t \in [k_n, k_n + a_n]\}$ is a martingale for each $k = 0, 1, 2, \dots$ and $\alpha \sim_n k_n$. Thus $\{(Z^{(n)}_{k_n}(f), \mathscr{F}^n_{k_n}): k = 0, 1, 2, \dots\}$ is a martingale. Next we calculate its quadratic variation,

$$\langle Z^{(n)}(f)\rangle_{k_n} = \sum_{r<k} E\Big(\Big(Z^{(n)}_{r_n+a_n}(f) - Z^{(n)}_{r_n}(f)\Big)^2\Big|\mathscr{F}^n_{r_n}\Big), \qquad k = 0, 1, 2, \dots.$$

First we note that for $r = 0, 1, 2, \dots,$

$$
\begin{aligned}
&\Big(Z^{(n)}_{r_n+a_n}(f) - Z^{(n)}_{r_n}(f)\Big)^2\\
&= \bigg[n^{-1}\sum_{\alpha\sim_n r_n}\Big[\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n) + \beta_n M^{\alpha,r_n}_{r_n+a_n}(f)\Big]\bigg]^2\\
&= n^{-2}\sum_{\alpha\sim_n r_n}\Big(\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)\Big)^2 + n^{-2}\beta_n^2\sum_{\alpha\sim_n r_n}\Big(M^{\alpha,r_n}_{r_n+a_n}(f)\Big)^2\\
&\quad + n^{-2}\beta_n\sum_{\alpha,\beta\sim_n r_n}\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)M^{\beta,r_n}_{r_n+a_n}(f)\\
&\quad + n^{-2}\beta_n^2\sum_{\alpha\neq\beta\sim_n r_n}M^{\alpha,r_n}_{r_n+a_n}(f)M^{\beta,r_n}_{r_n+a_n}(f)\\
&\quad + n^{-2}\sum_{\alpha\neq\beta\sim_n r_n}\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)\hat{f}(X^{\beta,n}_{r_n})(N^{\beta,n}-\beta_n)\\
&:= \sum_1(r) + \sum_2(r) + \sum_3(r) + \sum_4(r) + \sum_5(r).
\end{aligned}
$$

For $\alpha \sim_n r_n$ and $\beta \sim_n r_n$ we have that

$$
\begin{aligned}
&E\Big(\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)M^{\beta,r_n}_{r_n+a_n}(f)|\mathscr{F}^n_{r_n}\Big)\\
&\qquad = \hat{f}(X^{\alpha,n}_{r_n})E(N^{\alpha,n}-\beta_n|\mathscr{F}^n_{r_n})E(M^{\beta,r_n}_{r_n+a_n}(f)|\mathscr{F}^n_{r_n}) = 0
\end{aligned}
$$

and so, for $r = 0, 1, 2, \dots$ we have $E(\sum_3(r)|\mathscr{F}^n_{r_n}) = 0$. Similarly for $\alpha \sim_n r_n, \beta \sim_n r_n, \alpha \neq \beta$ we have

$$
E\Big(\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)\hat{f}(X^{\beta,n}_{r_n})(N^{\beta,n}-\beta_n)\,\Big|\,\mathscr{F}^n_{r_n}\Big) = 0,
$$

and so, for $r = 0, 1, 2, \dots$ we have $E(\sum_5(r)|\mathscr{F}^n_{r_n}) = 0$. On the other hand,

$$
E\Big(\Big(\hat{f}(X^{\alpha,n}_{r_n})(N^{\alpha,n}-\beta_n)\Big)^2\,\Big|\,\mathscr{F}^n_{r_n}\Big) = \big(\hat{f}(X^{\alpha,n}_{r_n})\big)^2\sigma_n^2,
$$

and so, for $r = 0, 1, 2, \dots$ we have

$$
E\bigg(\sum_1(r)|\mathscr{F}^n_{r_n})\bigg) = n^{-2}\sigma_n^2\sum_{\alpha\sim_n r_n}\big(\hat{f}(X^{\alpha,n}_{r_n})\big)^2 = n^{-1}\sigma_n^2 X^{\alpha,n}_{r_n}(f^2).
$$

Moreover, from the martingale structure of $M^{\alpha,r_n}(f)$, we have that for $\alpha \sim_n r_n$ and $r_n \leq t \leq r_n + a_n$,

$$
\begin{aligned}
\big\langle M^{\alpha,r_n}(f)\big\rangle_t &= \sum_{l=1}^m\int_{r_n}^t\bigg[\sum_{i=1}^d\hat{c_{il}}(X^{\alpha,n}_u)\hat{f}^i(X^{\alpha,n}_u)\bigg]^2du\\
&\quad + \sum_{i=1}^d\int_{r_n}^t\big[\hat{e}_i(X^{\alpha,n}_u)\hat{f}^i(X^{\alpha,n}_u)\big]^2du
\end{aligned}
$$

and so

$$\begin{aligned}E\big((M^{\alpha,r_n}_{r_n+a_n}(f))^2|\mathscr{F}^n_{r_n}\big) &= E\Big(\langle M^{\alpha,r_n}(f)\rangle_{r_n+a_n}\,\Big|\,\mathscr{F}^n_{r_n}\Big)\\&= \int_{r_n}^{r_n+a_n} E\Big((\widehat{\Psi f})(X^{\alpha,n}_u)\,\Big|\,\mathscr{F}^n_{r_n}\Big)\,du\\&= \int_{r_n}^{r_n+a_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})E\Big((\Psi f)(Y^{\alpha,n}_u)\,\Big|\,\mathscr{F}^n_{r_n}\Big)\,du\\&= I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})\int_{r_n}^{r_n+a_n} T_{u-r_n}(\Psi f)(Y^{\alpha,n}_{r_n})\,du\\&= I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})\int_0^{a_n} T_u(\Psi f)(Y^{\alpha,n}_{r_n})\,du\end{aligned}$$

which, in turn, implies that

$$E\Big((M^{\alpha,r_n}_{r_n+a_n}(f))^2\,\Big|\,\mathscr{F}^n_{r_n}\Big) = \int_0^{a_n} \widehat{T_u(\Psi} f)(X^{\alpha,n}_{r_n})\,du. \tag{A.7}$$

Hence, for $r = 0, 1, 2, \ldots$ we have

$$\begin{aligned}E\Big(\sum_2(r)\,\Big|\,\mathscr{F}^n_{r_n}\Big) &= n^{-2}\beta_n^2\sum_{\alpha\sim_n r_n}\int_0^{a_n}\widehat{T_u(\Psi} f)(X^{\alpha,n}_{r_n})\,du\\&= n^{-1}\beta_n^2\int_0^{a_n} X^n_{r_n}(T_u(\Psi f))\,du.\end{aligned}$$

Finally, for $\alpha\sim_n r_n$, $\beta\sim_n r_n$, $\alpha\neq\beta$ and $r_n\le t\le r_n+a_n$,

$$\langle M^{\alpha,r_n}(f), M^{\beta,r_n}(f)\rangle_t = \sum_{i=1}^d\sum_{j=1}^d\sum_{l=1}^m\int_{r_n}^t \hat f^i(X^{\alpha,n}_u)\hat c_{il}(X^{\alpha,n}_u)\hat f^j(X^{\beta,n}_u)\hat c_{jl}(X^{\beta,n}_u)du$$

and so

$$\begin{aligned}&E\Big(M^{\alpha,r_n}_{r_n+a_n}(f)M^{\beta,r_n}_{r_n+a_n}(f)\,|\,\mathscr{F}^n_{r_n}\Big)\\&\quad= E\Big(\langle M^{\alpha,r_n}(f), M^{\beta,r_n}(f)\rangle_{r_n+a_n}\,|\,\mathscr{F}^n_{r_n}\Big)\\&\quad= \int_{r_n}^{r_n+a_n} E\Big((\widehat{\Lambda f})(X^{\alpha,n}_u, X^{\beta,n}_u)\,|\,\mathscr{F}^n_{r_n}\Big)\,du\\&\quad= \int_{r_n}^{r_n+a_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})I_{\mathbb{R}^d}(X^{\beta,n}_{r_n})E\Big((\Lambda f)(Y^{\alpha,n}_u, Y^{\beta,n}_u)\,|\,\mathscr{F}^n_{r_n}\Big)\,du\\&\quad= I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})I_{\mathbb{R}^d}(X^{\beta,n}_{r_n})\int_{r_n}^{r_n+a_n}\big(S_{u-r_n}(\Lambda f)\big)(Y^{\alpha,n}_{r_n}, Y^{\beta,n}_{r_n})\,du\\&\quad= I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})I_{\mathbb{R}^d}(X^{\beta,n}_{r_n})\int_0^{a_n}\big(S_u(\Lambda f)\big)(Y^{\alpha,n}_{r_n}, Y^{\beta,n}_{r_n})\,du,\end{aligned}$$

which, in turn, implies

$$E\Big(M^{\alpha,r_n}_{r_n+a_n}(f)M^{\beta,r_n}_{r_n+a_n}(f)\,|\,\mathscr{F}^n_{r_n}\Big) = \int_0^{a_n}\widehat{S_u(\Lambda} f)(X^{\alpha,n}_{r_n}, X^{\beta,n}_{r_n})\,du. \tag{A.8}$$

Hence, for $r = 0, 1, 2, \dots,$ we have

$$E\biggl(\sum_4(r)\Big|\mathscr{F}^n_{r_n}\biggr) = n^{-2}\beta_n^2 \sum_{\alpha\neq\beta\sim_n r_n} \int_0^{a_n} S_u\widehat{(\Lambda f)}(X^{\alpha,n}_{r_n}, X^{\beta,n}_{r_n})\,du$$

$$= n^{-2}\beta_n^2 \sum_{\alpha,\,\beta\sim_n r_n} \int_0^{a_n} S_u\widehat{(\Lambda f)}(X^{\alpha,n}_{r_n}, X^{\beta,n}_{r_n})\,du$$

$$- n^{-2}\beta_n^2 \sum_{\alpha\sim_n r_n} \int_0^{a_n} S_u\widehat{(\Lambda f)}(X^{\alpha,n}_{r_n}, X^{\alpha,n}_{r_n})\,du$$

$$= \beta_n^2 \int_0^{a_n} \biggl(\int_{\mathbb{R}^d\times\mathbb{R}^d} (S_u(\Lambda f))(x,y) X^n_{r_n}(dx) X^n_{r_n}(dy)\biggr)\,du$$

$$- n^{-1}\beta_n^2 \int_0^{a_n} \biggl(\int_{\mathbb{R}^d} (S_u(\Lambda f))(x,x) X^n_{r_n}(dx)\biggr)\,du.$$

Therefore, by collecting all the terms together, we obtain

$$\langle Z^{(n)}(f)\rangle_{k_n}$$

$$= \sum_{r<k}\biggl(E\biggl(\sum_1(r)|\mathscr{F}^n_{r_n}\biggr) + E\biggl(\sum_2(r)|\mathscr{F}^n_{r_n}\biggr) + E\biggl(\sum_4(r)|\mathscr{F}^n_{r_n}\biggr)\biggr)$$

$$= \sigma_n^2 n^{-1} \sum_{r<k} X^n_{r_n}(f^2) + \beta_n^2 n^{-1} \sum_{r<k} \int_{\mathbb{R}^d} \biggl[\int_0^{a_n} (T_u(\Psi f))(x)\,du\biggr] X^n_{r_n}(dx)$$

$$+ \beta_n^2 n^{-2} \sum_{r<k} \biggl\{ \sum_{\alpha,\,\beta\sim_n r_n} \int_0^{a_n} S_u\widehat{(\Lambda f)}(X^{\alpha,n}_{r_n}, X^{\beta,n}_{r_n})\,du$$

$$- \sum_{\alpha\sim_n r_n} \int_0^{a_n} S_u\widehat{(\Lambda f)}(X^{\alpha,n}_{r_n}, X^{\alpha,n}_{r_n})\,du \biggr\}$$

$$= \sigma_n^2 \lambda \int_0^{k_n} X^n_{[\lambda n s]_n}(f^2)\,ds$$

$$+ \beta_n^2 \int_0^{k_n} \int_{\mathbb{R}^d\times\mathbb{R}^d} \biggl(a_n^{-1}\int_0^{a_n} (S_u(\Lambda f))(x,y)\,du\biggr) X^n_{[\lambda n s]_n}(dx) X^n_{[\lambda n s]_n}(dy)\,ds$$

$$+ n^{-1}\beta_n^2 \int_0^{k_n} \int_{\mathbb{R}^d} \biggl(a_n^{-1}\int_0^{a_n} \Bigl[(T_u(\Psi f))(x) - (S_u(\Lambda f))(x,x)\Bigr]\,du\biggr)$$

$$\times X^n_{[\lambda n s]_n}(dx)\,ds$$

which completes the proof. □

Next we consider the cadlag extensions of the discrete time processes $\langle Z^{(n)}(f)\rangle_{k_n}$, $k = 0, 1, \dots,$ for $f \in \mathscr{D}$ and $n = 1, 2, \dots,$ by defining $\langle Z^{(n)}(f)\rangle_t = \langle Z^{(n)}(f)\rangle_{k_n}$ for $k_n \le t < k_n + a_n$ and $k = 0, 1, 2, \dots.$ Then, proceeding as in the proof of Lemma A.3.4 and using Chebyshev's inequality, Lemma A.3.1 and Proposition 3.26 of Chapter VI in [14] we can prove the following.

LEMMA A.3.6. *For all* $f \in \mathscr{D}$, $\{\langle Z^{(n)}(f)\rangle: n \geq 1\}$ *is a C-tight sequence of processes in* $D_{\mathbb{R}}[0,\infty)$.

The following lemma will be used in proving C-tightness of the sequence $\{Z^{(n)}(f): n \geq 1\}$.

LEMMA A.3.7. *For all* $f \in \mathscr{D}$ *and* $J \in \mathbb{N}^*$,

$$\lim_{n\to\infty} E \sup_{0\le k\le \lambda n J} \left(Z^{(n)}_{(k+1)_n}(f) - Z^{(n)}_{k_n}(f)\right)^2 = 0.$$

PROOF. First we note that, for $k = 0, 1, \dots,$ we have

$$\left(Z^{(n)}_{(k+1)_n}(f) - Z^{(n)}_{k_n}(f)\right)^2 \le 2\left(n^{-1}\sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{k_n}) f(Y^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n)\right)^2 + 2\beta_n^2\left(n^{-1}\sum_{\alpha\sim_n k_n} M^{\alpha,k_n}_{k_n+a_n}(f)\right)^2.$$

Let $r(n) = [\lambda n T]$. Then, as in the treatment of (A.4), the last term can be seen to tend to 0 as $n\to\infty$. Thus, it suffices to prove that

$$\lim_{n\to\infty} E \max_{0\le k\le r(n)} \left| n^{-1}\sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{k_n}) f(Y^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n)\right|^p = 0,$$

where p is the number satisfying (2.5). But

$$E \max_{0\le k\le r(n)} \left| n^{-1}\sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{k_n}) f(Y^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n)\right|^p \le \sum_{k=0}^{r(n)} EE\left(\left| n^{-1}\sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{k_n}) f(Y^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n)\right|^p \middle| \mathscr{F}^n_{k_n}\right).$$

Now let $n \geq 1$ and $k \geq 0$ be fixed. Then

$$E\left(\left| n^{-1}\sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X^{\alpha,n}_{k_n}) f(Y^{\alpha,n}_{k_n})(N^{\alpha,n} - \beta_n)\right|^p \middle| \mathscr{F}^n_{k_n}\right) = E\left| n^{-1}\sum_{i=1}^{N} f(Y^{\alpha_i,n}_{k_n})(N^{\alpha_i,n} - \beta_n)\right|^p,$$

where $\alpha_1, \dots, \alpha_N$ are the labels of the particles alive at time k_n. Clearly $N = nX^n_{k_n}(1)$. Next let, for $i = 1, 2, \dots, \alpha_i \sim_n k_n$ such that $\alpha_i \neq \alpha_j$ when $i \neq j$ and define

$$M_m(f) = n^{-1}\sum_{i=1}^{m} f(Y^{\alpha_i,n}_{k_n})(N^{\alpha_i,n} - \beta_n),$$

$$\mathscr{G}_m = \sigma(N^{\alpha_i,n}, Y^{\alpha_i,n}_{k_n}: i = 1, \dots, m)$$

for $m = 1, 2, \ldots$. Then, it follows that $\{(M_m(f), \mathscr{G}_m): m = 1, 2, \ldots\}$ is a martingale and a simple computation yields $\langle M(f)\rangle_m \le m B_f^2 \sigma_n^2/n^2$. Hence, by Theorem 21.1 in [4] we obtain

$$\begin{aligned} E|M_m(f)|^p &\le c_1 E(\langle M(f)\rangle_m)^{p/2} + c_1 E \max_{1\le i\le m} \left| n^{-1} f(Y_{k_n}^{\alpha_i, n})(N^{\alpha_i, n} - \beta_n)\right|^p \\ &\le c_1 \left(\frac{\sqrt{m}}{n} B_f \sigma_n\right)^p + c_2 \frac{m}{n^p} B_f^p \\ &\le c_3 \left(\frac{\sqrt{m}}{n}\right)^p, \end{aligned}$$

where c_1, c_2, and c_3 are positive constants that depend only on p, M and f. Therefore,

$$\begin{aligned} &E \max_{0\le k\le r(n)} \left| n^{-1} \sum_{\alpha\sim_n k_n} I_{\mathbb{R}^d}(X_{k_n}^{\alpha, n}) f(Y_{k_n}^{\alpha, n})(N^{\alpha, n} - \beta_n)\right|^p \\ &\quad \le (r(n)+1) c_3 \left(n^{-1} \sqrt{n^{-1} \sup_{0\le t\le J} X_t^n(1)}\right)^p \\ &\quad \le c_3 \sqrt{C'_J} \frac{r(n)+1}{n^{p/2}} \to 0 \quad \text{as } n \to \infty \end{aligned}$$

since $p > 2$ and $r(n) \sim n$. □

Next we consider the cadlag extensions of the discrete time processes $Z_{k_n}^{(n)}(f)$, $k \ge 0$, for $f \in \mathscr{D}$ and $n \ge 1$, by defining $Z_t^{(n)}(f) = Z_{k_n}^{(n)}(f)$ for $k_n \le t < k_n + a_n$ and $k \ge 0$. The following lemma, which is an immediate consequence of Lemmas A.3.6, A.3.7 and Theorem A.2.1(a), will be the final step in proving C-tightness of the sequence $\{X^{(n)}(f): n = 1, 2, \ldots\}$.

LEMMA A.3.8. *For all $f \in \mathscr{D}$, $\{Z^{(n)}(f): n \ge 1\}$ is a C-tight sequence of processes in $D_{\mathbb{R}}[0, \infty)$.*

From the decomposition in (A.3), using Lemmas A.3.3, A.3.4 and A.3.8 and Corollary 3.33 of Chapter VI in [14], we have

LEMMA A.3.9. *For all $f \in \mathscr{D}$, $\{X^n(f): n \ge 1\}$ is a C-tight sequence of processes in $D_{\mathbb{R}}[0, \infty)$.*

In what follows, we denote by $\bar{h}$ the extension of $h \in C_l(\mathbb{R}^d)$ to $\overline{\mathbb{R}}$ such that $\bar{h}(\Delta) = \lim_{|x|\to\infty} h(x)$. Then clearly $\bar{h} \in C(\overline{\mathbb{R}})$. Now we can establish C-tightness of the sequence $\{X^n\}$.

PROPOSITION A.3.10. *$\{X^n: n \ge 1\}$ is a C-tight sequence of processes in $D_{M_F(\overline{\mathbb{R}})}[0, \infty)$.*

PROOF. Let $h \in C_l(\mathbb{R}^d)$. Since X^n does not charge Δ we see that $X^n(\bar{h}) = X^n(h)$. An application of the Stone-Weierstrass theorem gives the existence of a countable subset $\mathscr{R}$ of $\mathscr{D}$ which is dense in $C_0(\mathbb{R}^d)$. We can assume that $0 \in \mathscr{R}$. Thus if we let $S = \{\bar{f}: f = g + c,\ g \in \mathscr{R},\ c \text{ rational}\}$, then it is easily seen that S is countable and dense in $C(\overline{\mathbb{R}})$ and $1 \in S$. Furthermore, for $g \in \mathscr{D}$, c rational and $f = g + c$ we have that $X^n(\bar{f}) = X^n(f) = X^n(g) + cX^n(1)$ is a C-tight sequence of processes in $D_{\mathbb{R}}[0, \infty)$, by Lemma A.3.9 and Corollary VI.3.33 in [14]. Finally the conclusion follows from Theorem A.2.2. □

The following lemma will be used in identifying the quadratic variation of martingales that are weak limit points of $\{Z^{(n)}(f)\}$.

LEMMA A.3.11. *For all $f \in \mathscr{D}$ and $J \in \mathbb{N}^*$, the family* $\{\sup_{1 \le k \le \lambda n J} Z^{(n)}_{k_n}(f)$: $n \ge 1\}$ *is uniformly integrable.*

PROOF. By the corollary to Proposition 6.3.3 in [17] we see that it suffices to show that $\sup_n E(\sup_{1\le k_n \le J} Z^{(n)}_{k_n}(f))^2 < \infty$. First, by Lemma A.3.5 and Doob's inequality, we obtain

$$E\Bigl(\sup_{1\le k_n\le J} Z^{(n)}_{k_n}(f)\Bigr)^2 \le 4E\bigl(Z^{(n)}_J(f)\bigr)^2.$$

Let $q \in \mathbb{N}^*$. For $l \le r < q$ and $\alpha \sim_n l_n, \beta \sim_n r_n$ such that $\alpha \ne \beta$, it is then easily deduced that

$$E\Bigl(f(X^{\alpha,n}_{l_n})(N^{\alpha,n} - \beta_n)f(X^{\beta,n}_{r_n})(N^{\beta,n} - \beta_n)\Bigr) = 0,$$

$$E\Bigl(M^{\alpha,n}_{l_n+a_n}(f)f(X^{\beta,n}_{r_n})(N^{\beta,n} - \beta_n)\Bigr) = 0,$$

$$E\Bigl(f(X^{\alpha,n}_{l_n})(N^{\alpha,n} - \beta_n)M^{\beta,n}_{r_n+a_n}(f)\Bigr) = 0.$$

Similarly, for $l < r < q$ and $\alpha \sim_n l_n, \beta \sim_n r_n$ we have

$$E\Bigl(M^{\alpha,n}_{l_n+a_n}(f)M^{\beta,n}_{r_n+a_n}(f)\Bigr) = 0,$$

and for $r < q$ and $\alpha \sim_n r_n$

$$E\Bigl(f(X^{\alpha,n}_{r_n})(N^{\alpha,n} - \beta_n)M^{\alpha,n}_{r_n+a_n}(f)\Bigr) = 0.$$

Hence,

$$E\Bigl(Z^{(n)}_{q_n}(f)\Bigr)^2 = n^{-2}\sum_{r<q}\sum_{\alpha\sim_n r_n} E\Bigl((f(X^{\alpha,n}_{r_n}))^2(N^{\alpha,n} - \beta_n)^2\Bigr) + n^{-2}\beta_n^2\sum_{r<q}\sum_{\alpha,\beta\sim_n r_n} E\Bigl(M^{\alpha,n}_{r_n+a_n}(f)M^{\beta,n}_{r_n+a_n}(f)\Bigr).$$

Now note that

$$E\Big((f(X^{\alpha,n}_{r_n}))^2(N^{\alpha,n}-\beta_n)^2\Big) = E\Big(E\Big((f(X^{\alpha,n}_{r_n}))^2(N^{\alpha,n}-\beta_n)^2|\mathscr{F}^n_{r_n}\Big)\Big)$$
$$\le \sigma_n^2 {B_f}^2 E(I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n})).$$

Furthermore by using expressions (A.7) and (A.8) in the proof of Lemma A.3.5 we can obtain that for $r < q$ and $\alpha \sim_n r_n$, $\beta \sim_n r_n$

$$E\Big(M^{\alpha,n}_{r_n+a_n}(f)M^{\beta,n}_{r_n+a_n}(f)|\mathscr{F}^n_{r_n}\Big) \le a_n B_f I_{\mathbb{R}^d}(X^{\alpha,n}_{r_n}) I_{\mathbb{R}^d}(X^{\beta,n}_{r_n})$$

which now implies

$$E\Big(Z^{(n)}_{q_n}(f)\Big)^2 \le n^{-1}\sigma_n^2 {B_f}^2 \sum_{r<q} E(X^n_{r_n}(1)) + a_n\beta_n^2 B_f \sum_{r<q} E(X^n_{r_n}(1))^2.$$

Taking $q = [\lambda n J]$ and using Lemma A.3.1 yields

$$E\big(Z^{(n)}_J(f)\big)^2 \le n^{-1}\sigma_n^2 {B_f}^2[\lambda n J]\sqrt{C_J} + a_n\beta_n^2 B_f[\lambda n J]C_J$$
$$\to \sigma^2 {B_f}^2 \lambda J\sqrt{C_J} + JB_f C_J \quad \text{as } n\to\infty$$

completing the proof. □

The following proposition and lemma provide the last steps in the proof of Theorem A.1.1.

PROPOSITION A.3.12. *Let $X \in C_{M_F(\mathbb{R})}[0,\infty)$ be a weak limit point of $\{X^n\}$. Then, for all $f \in \mathscr{D}$,*

$$Z_t(f) = X_t(\bar f) - \nu(f) - \int_0^t X_s(\overline{Lf})\,ds - \xi\int_0^t X_s(\bar f)\,ds$$

is a continuous square integrable $\{\mathscr{F}^X_t\}$-martingale such that $Z_0(f)=0$ and

$$\langle Z(f)\rangle_t = \delta\int_0^t X_s(\bar f^2)\,ds + \sum_{l=1}^m \int_0^t \left(\sum_{i=1}^d X_s\Big(\overline{c_{il}f^i}\Big)\right)^2 ds,$$

where $\xi = \lambda\gamma$ and $\delta = \lambda\sigma^2$.

PROOF. Let $f \in \mathscr{D}$ be fixed and $\{X^{\pi_n}\}$ be a subsequence of $\{X^n\}$ such that $X^{\pi_n} \Rightarrow X$. Then (A.3) can be rewritten as

(A.9)
$$X^n_t(\bar f) = X^n_0(f) + M^{(n)}_t(f) + J^{(n)}_t(f) + N^{(n)}_t(f)$$
$$+ Z^{(n)}_t(f) + \beta_n C^{(n)}_t(f) + \gamma_n H^{(n)}_t(f),$$

where

$$C^{(n)}_t(f) = \int_0^t X^n_s(\overline{Lf})\,ds \quad \text{and} \quad H^{(n)}_t(f) = \lambda\int_0^{k_n} X^n_{[\lambda ns]_n}(\bar f)\,ds.$$

By Lemmas A.3.3, A.3.4, A.3.8 all sequences of processes appearing in (A.9) are C-tight, and by Lemma A.3.6 $\{\langle Z^{(n)}(f)\rangle\}$ is C-tight as well. Therefore it can be assumed that the subsequences indexed by (π_n) of all processes appearing in (A.9) along with $\{\langle Z^{(\pi_n)}(f)\rangle\}$ and $\{X^{\pi_n}\}$ converge weakly jointly on the appropriate Skorokhod product space. This can always be done by taking further subsequences if necessary. Let $Z(f)$ denote the weak limit of $\{Z^{(\pi_n)}(f)\}$. Then by Theorem A.2.1(b) and Lemma A.3.11, we obtain that $Z(f)$ is a continuous square integrable $\{\mathscr{F}_t^X\}$-martingale and that $\langle Z^{(\pi_n)}(f)\rangle \Rightarrow \langle Z(f)\rangle$. The joint convergence just mentioned can be assumed to be almost sure Skorokhod convergence on the Skorokhod product space by using a Skorokhod representation. Now taking limits as $n \to \infty$ in A.9, along subsequences indexed by (π_n), we obtain, by using Lemma A.3.3 again to see which terms disappear in the limit, that

$$X_t(\bar{f}) = \nu(f) + Z_t(f) + \int_0^t X_s(\overline{Lf})\,ds + \xi \int_0^t X_s(\overline{f})\,ds.$$

Finally, the conclusion follows by taking limits as $n \to \infty$ in (A.6) of Lemma A.3.5, using the fact $s\text{-}\lim_{t\downarrow 0} t^{-1}\int_0^t S_u h\,du = h$ for $h \in C_l(\mathbb{R}^d \times \mathbb{R}^d)$ and noticing that the last term in (A.6) vanishes in the limit. □

LEMMA A.3.13. *Let $X \in C_{M_F(\overline{\mathbb{R}})}[0,\infty)$ be a weak limit point of $\{X^n\}$. Then, with probability 1, $X_t(\{\Delta\}) = 0$ for all $t \geq 0$.*

PROOF. First we define a sequence of functions on $\mathbb{R}^d$ as follows. For $n = 1, 2, \ldots$, let

$$g_n(x) = \begin{cases} \exp\left\{-\dfrac{1}{|x|^2 - n^2}\right\}, & \text{if } |x| > n,\\ 0, & \text{if } |x| \leq n. \end{cases}$$

Then one can easily check, using the linear growth assumption on b, c and e, that $g_n \in \mathscr{D}$, $\lim_{|x|\to\infty} g_n(x) = 1$, $\lim_{|x|\to\infty} Lg_n(x) = 0$, $\lim_{|x|\to\infty} c_{il}(x)g_n^i(x) = 0$, $\overline{g_n} \to I_{\{\Delta\}}$ b.p., as $n \to \infty$, $\overline{Lg_n} \to 0$ b.p., as $n \to \infty$ and $\overline{c_{il}g_n^i} \to 0$ b.p., as $n \to \infty$, where "b.p." stands for bounded pointwise. Moreover, by Proposition A.3.12 we have

$$X_t(\overline{g_n}) = \nu(g_n) + Z_t(g_n) + \int_0^t X_s(\overline{Lg_n})\,ds + \xi \int_0^t X_s(\overline{g_n})\,ds. \tag{A.10}$$

Let $T > 0$ and $g_{n,r} = g_n - g_r$ for $n, r = 1, 2, \ldots$. Then, using Doob's inequality and Proposition A.3.12, we obtain

$$\begin{aligned} E \sup_{0\leq t\leq T} (Z_t(g_n) - Z_t(g_r))^2 &= E \sup_{0\leq t\leq T} (Z_t(g_{n,r}))^2 \leq 4E(Z_T(g_{n,r}))^2 \\ &= 4E\langle Z(g_{n,r})\rangle_T = 4\delta E \int_0^T X_s\left(\overline{g_{n,r}^{\,2}}\right) ds \\ &\quad + 4\sum_{l=1}^m E\int_0^T \left(X_s\left(\sum_{i=1}^d \overline{c_{il}g_{n,r}^i}\right)\right)^2 ds. \end{aligned}$$

Therefore, by bounded pointwise convergence, $E\sup_{0\le t\le T}(Z_t(g_n)-Z_t(g_r))^2\to\infty$ as $n, r\to\infty$ and so $Z_t(g_n)$ converges in mean square as $n\to\infty$ to a limit that we denote by Z_t^Δ and with probability 1 along an appropriate subsequence. Here Z_t^Δ is a continuous square integrable martingale with $Z_0^\Delta=0$. Then from (A.10) we deduce that

$$X_t(\{\Delta\}) = Z_t^\Delta + \xi\int_0^t X_s(\{\Delta\})\,ds$$

and thus $X_t(\{\Delta\})$ is a continuous process. Taking expectations in the last equality and using Gronwall's inequality yields $EX_t(\{\Delta\})=0$, for all $t\ge 0$. Hence $X_t(\{\Delta\})=0$ with probability one, for all $t\ge 0$ and now the conclusion follows from the continuity of $X_t(\{\Delta\})$. □

Finally, Theorem A.1.1 follows from Proposition A.3.10, Proposition A.3.12 and Lemma A.3.13. □

A.4. *Uniqueness.* In this subsection we will prove that uniqueness holds for solutions of the martingale problem $(A.1)$ using the duality technique developed by Dawson and Kurtz in [7]. We will first show that every solution of the martingale problem $(A.1)$ solves the $(\mathscr{L},\delta_\nu)$-martingale problem where $\mathscr{L}$ is the second-order differential operator defined by

$$\begin{aligned}\mathscr{L}F(\mu) &= \int_{\overline{\mathbb{R}}}(L+\xi)\left(\frac{\delta F(\mu)}{\delta\mu(x)}\right)\mu(dx) + \frac{1}{2}\delta\int_{\overline{\mathbb{R}}}\frac{\delta^2 F(\mu)}{\delta\mu(x)^2}\mu(dx)\\ &\quad + \frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d\int_{\overline{\mathbb{R}}}\int_{\overline{\mathbb{R}}} a_{ij}^{(m)}(x,y)\frac{\partial^2}{\partial x_i\partial y_j}\left(\frac{\delta^2 F(\mu)}{\delta\mu(x)\delta\mu(y)}\right)\mu(dx)\mu(dy)\end{aligned}$$

for F in some appropriate domain $\mathscr{D}(\mathscr{L})\subset C(M_F(\overline{\mathbb{R}}))$, where the so-called variational derivatives are defined by

$$\frac{\delta F(\mu)}{\delta\mu(x)} := \lim_{h\downarrow 0}\frac{F(\mu+h\delta_x)-F(\mu)}{h} = \frac{\partial}{\partial h}F(\mu+h\delta_x)|_{h=0},\qquad x\in\overline{\mathbb{R}}$$

and

$$\frac{\delta^2 F(\mu)}{\delta\mu(x)\delta\mu(y)} := \frac{\partial^2}{\partial h_1\partial h_2}F(\mu+h_1\delta_x+h_2\delta_y)|_{h_1=h_2=0},\qquad x, y\in\overline{\mathbb{R}}.$$

The notation δ_x stands for the measure with unit mass at x. Then, employing the duality method, we will prove the well-posedness for the martingale problem for $(\mathscr{L},\delta_\nu)$, which in turn will imply the uniqueness for solutions of the martingale problem $(A.1)$.

Let us start with some preliminary notation and definitions. By $\overline{\mathbb{R}}^N$ we denote the N-fold Cartesian product of $\overline{\mathbb{R}}$ (the one-point compactification of E) and define

$$\mathbf{C} := \bigcup_{N=0}^{\infty} C(\overline{\mathbb{R}}^N)\text{ (disjoint union)}\quad\text{where } C(\overline{\mathbb{R}}^0) := \mathbb{R}.$$

For $f \in \mathbf{C}$ we define $N(f) := N$ if $f \in C(\overline{\mathbb{R}}^N)$. Similarly we define

$$\mathbf{D} := \bigcup_{N=0}^{\infty} D(\overline{\mathbb{R}}^N),$$

where for each N, $D(\overline{\mathbb{R}}^N)$ is assumed to be a dense subspace of $C(\overline{\mathbb{R}}^N)$. We will specify the spaces $D(\overline{\mathbb{R}}^N)$ later. For every function $f \in \mathbf{C}$ we define the function F_f on $M_F(\overline{\mathbb{R}})$ by

$$F_f(\mu) = \int_{\overline{\mathbb{R}}} \cdots \int_{\overline{\mathbb{R}}} f(x_1, \ldots, x_{N(f)}) \mu(dx_1) \cdots \mu(dx_{N(f)}),$$

which is said to be a monomial on $M_F(\overline{\mathbb{R}})$. Then it is easy to see that the variational derivatives of a monomial F_f are given by

$$\frac{\delta F_f(\mu)}{\delta \mu(x)} = \sum_{p=1}^{N(f)} \int_{\overline{\mathbb{R}}} \cdots \int_{\overline{\mathbb{R}}} f(x_1, \ldots, x_{p-1}, x, x_{p+1}, \ldots, x_{N(f)}) \mu^{(N(f)/p)}(dx)$$

and

$$\frac{\delta^2 F_f(\mu)}{\delta \mu(x) \delta \mu(y)} = \sum_{\substack{p,q=1 \\ p \neq q}}^{N(f)} \int_{\overline{\mathbb{R}}} \cdots \int_{\overline{\mathbb{R}}} f\big(x_1, \ldots, x_{p-1}, x, x_{p+1}, \ldots, x_{q-1}, y, x_{q+1} x_{N(f)}\big) \times \mu^{(N(f)/pq)}(dx),$$

where

$$\mu^{(N/p)}(dx) = \prod_{\substack{l=1 \\ l \neq p}}^{N} \mu(dx_l) \quad \text{and} \quad \mu^{(N/pq)}(dx) = \prod_{\substack{l=1 \\ l \neq p,\, q}}^{N} \mu(dx_l).$$

Now we restrict ourselves to a specific choice of f. Let $N \geq 1$ and $f_i \in \mathscr{D}$ for $i = 1, 2, \ldots, N$ and define the function f on $\overline{\mathbb{R}}^N$ by $f(x_1, x_2, \ldots, x_N) = \prod_{i=1}^N f_i(x_i)$. Then, for this specific choice of f, we have $F_f(\mu) = \mu(f_1) \cdots \mu(f_N)$,

$$\frac{\delta F_f(\mu)}{\delta \mu(x)} = \sum_{p=1}^{N} \left(\prod_{\substack{l=1 \\ l \neq p}}^{N} \mu(f_l) \right) f_p(x), \qquad x \in \overline{\mathbb{R}}$$

and

$$\frac{\delta^2 F_f(\mu)}{\delta \mu(x) \delta \mu(y)} = \sum_{\substack{p,q=1 \\ p \neq q}}^{N} \left(\prod_{\substack{l=1 \\ l \neq p,\, q}}^{N} \mu(f_l) \right) f_p(x) f_q(y), \qquad x, y \in \overline{\mathbb{R}}.$$

Let X be a solution to the martingale problem $(A.1)$. Then by applying Itô's formula we obtain

$$F_f(X_t) = X_t(f_1)\cdots X_t(f_N) = F_f(X_0) + \sum_{p=1}^{N}\int_0^t \left(\prod_{\substack{l=1\\ l\neq p}}^{N} X_s(f_l)\right) dZ_s(f_p)$$

$$+ \sum_{p=1}^{N}\int_0^t \left(\prod_{\substack{l=1\\ l\neq p}}^{N} X_s(f_l)\right) X_s((L+\xi)f_p)\,ds$$

$$+\tfrac{1}{2}\sum_{\substack{p,\,q=1\\ p\neq q}}^{N}\int_0^t \left(\prod_{\substack{l=1\\ l\neq p,\,q}}^{N} X_s(f_l)\right) d\langle Z(f_p), Z(f_q)\rangle_s.$$

Also from $(A.1)$ it follows that

$$\langle Z(f_p), Z(f_q)\rangle_t = \delta\int_0^t X_s(f_pf_q)\,ds + \int_0^t (X_s\times X_s)(\Lambda_{f_p,f_q})\,ds,$$

where

$$\Lambda_{g,h}(x,y) = \sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}^{(m)}(x,y)g^i(x)h^j(y), \qquad x, y\in E$$

for $g, h\in C^1(E)$. Hence we see that, for all $t\geq 0$,

$$\begin{aligned} F_f(X_t) \ &-\int_0^t\int_{\overline{\mathbb{R}}}(L+\xi)\left(\frac{\delta F_f(X_s)}{\delta X_s(x)}\right)X_s(dx)\,ds\\ &-\frac{1}{2}\delta\int_0^t\int_{\overline{\mathbb{R}}}\frac{\delta^2F_f(X_s)}{\delta X_s(x)^2}X_s(dx)\,ds\\ &-\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\int_0^t\int_{\overline{\mathbb{R}}}\int_{\overline{\mathbb{R}}} a_{ij}^{(m)}(x,y)\frac{\partial^2}{\partial x_i\partial y_j}\left(\frac{\delta^2 F_f(X_s)}{\delta X_s(x)\delta X_s(y)}\right)\\ &\times X_s(dx)X_s(dy)\,ds \end{aligned}$$

is an $\{\mathscr{F}_t^X\}$-martingale, or, more compactly,

(A.20) $$F_f(X_t) - \int_0^t(\mathscr{L}F_f)(X_s)\,ds \text{ is an } \{\mathscr{F}_t^X\}\text{-martingale.}$$

At this point we have to introduce the following spaces. For $N\geq 1$ and $\alpha\geq 0$ we define

$$C_l^{2,\alpha}(E^N) := \{f + c\colon c\in\mathbb{R},\ f\in C^2(E^N),\ \phi_\alpha D^k f\in C_0(E^N),\ 0\leq |k|\leq 2\}$$

and

$$D(\overline{\mathbb{R}}^N) := \{\bar{g}\colon g\in C_l^{2,\alpha}(E^N)\},$$

where $\bar{g}$ stands for the continuous extension to $\overline{\mathbb{R}}^N$ of $g \in C_l(E^N)$ and $\phi_\alpha(x) = (1+|x|)^\alpha$. In the uniformly elliptic case we take $\alpha = 0$, when in the linear case we take $\alpha = 1$. Furthermore, we define $\Gamma(E^N) := \{f\colon \exists\ n \geq 1,\ f_{ij} \in C_l^{2,\alpha}(E),\ i = 1, \dots, n,\ j = 1, \dots, N$ such that $f(x_1, x_2, \dots, x_N) = \sum_{i=1}^n f_{i1} \times (x_1) f_{i2}(x_2) \cdots f_{iN}(x_N)\}$. Since F_f is linear in $f \in C(\overline{\mathbb{R}}^N)$ and $\mathscr{L}$ is a linear operator, we see that (3.8) holds for all $\bar{f}$ with $f \in \Gamma(E^N)(C_l^{2,\alpha}(E) \subset \mathscr{D})$. By the Stone–Weierstrass theorem, it follows that $\Gamma(E^N)$ is dense in $C_l(E^N)$ [and as a consequence $D(\overline{\mathbb{R}}^N)$ is dense in $C(\overline{\mathbb{R}}^N)$]. Hence, by an approximating procedure, we can obtain that (3.8) holds for all $f \in D(\overline{\mathbb{R}}^N)$ which leads us to the conclusion that X is a solution of the martingale problem for $(\{(F_f, \mathscr{L}F_f)\colon f \in \mathbf{D}\}, \delta_\nu)$.

THEOREM A.4.1. *Assume that Assumption* U *is in effect, let* $\lambda \geq 0$, $\sigma > 0$, $\xi = \lambda\gamma$ *and* $\delta = \lambda\sigma^2$ *and define for all monomials* F *on* $M_F(\overline{\mathbb{R}})$,

$$\mathscr{L}F(\mu) = \int_{\overline{\mathbb{R}}} (L+\xi)\left(\frac{\delta F(\mu)}{\delta\mu(x)}\right)\mu(dx) + \frac{1}{2}\delta \int_{\overline{\mathbb{R}}} \frac{\delta^2 F(\mu)}{\delta\mu(x)^2}\mu(dx)$$
$$+ \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_{\overline{\mathbb{R}}} \int_{\overline{\mathbb{R}}} a_{ij}^{(m)}(x, y) \frac{\partial^2}{\partial x_i \partial y_j}\left(\frac{\delta^2 F(\mu)}{\delta\mu(x)\delta\mu(y)}\right)\mu(dx)\mu(dy),$$

where L *and* $a_{ij}^{(m)}$ *are defined by* (2.15) *and* (2.17), *respectively. Then the martingale problem for* $(\{(F_f, \mathscr{L}F_f)\colon f \in \mathbf{D}\}, \delta_\nu)$ *is well posed.*

We will apply Theorem 4.4 in [7] under their Hypotheses 4.3. In order to do that, we will need two lemmas which we first state and prove. Then we return to the proof of the theorem.

LEMMA A.4.2. *For any* $N \geq 1$ *and* $f \in C_b^2(E^N)$ *define*

$$G_N f(x_1, \dots, x_N)$$
$$= \sum_{p=1}^N \sum_{i=1}^d b_i(x_p) \frac{\partial f}{\partial x_{p,i}}(x_1, x_2, \dots, x_N)$$
$$+ \frac{1}{2}\sum_{p=1}^N \sum_{i,j=1}^d d_{ij}(x_p) \frac{\partial^2 f}{\partial x_{p,i}\partial x_{p,j}}(x_1, x_2, \dots, x_N)$$
$$+ \frac{1}{2}\sum_{\substack{p,q=1 \\ p \neq q}}^N \sum_{i,j=1}^d a_{ij}^{(m)}(x_p, x_q) \frac{\partial^2 f}{\partial x_{p,i}\partial x_{q,j}}(x_1, x_2, \dots, x_N),$$

where d_{ij} *and* $a_{ij}^{(m)}$ *are defined by* (2.16) *and* (2.17), *respectively. Then, under Assumption* U, *we have that*

(a) *The closure of* $\{f, G_N f)\colon f \in C_K^\infty(E^N)\}$ *is single-valued and generates a Feller semigroup* S_t^N *on* $C_0(E^N)$.

(b) *Define, for* $f \in C_l(E^N)$,

$$\overline{G}_N \bar{f}(x) = \begin{cases} G_N f(x), & \text{if } x \in E^N, \\ 0, & \text{otherwise,} \end{cases} \tag{A.21}$$

where $\bar{f} \in C(\overline{\mathbb{R}}^N)$ *is the continuous extension to* $\overline{\mathbb{R}}^N$ *of* f. *Then the closure of* $\{(\bar{f}, \overline{G}_N \bar{f}): f \in C_K^\infty(E^N)\}$ *is single-valued and generates a strongly continuous semigroup* $\overline{S}_t^N$ *on* $C(\overline{\mathbb{R}}^N)$. *The space* $D(\overline{\mathbb{R}}^N)$ *is contained in the domain of the generator just mentioned and also is invariant under* $\overline{S}_t^N$.

PROOF. Part (a) follows directly from by Theorem 8.2.5, [11], since all the conditions there are easily seen to follow from our Assumption U. The first conclusion in part (b) follows trivially if we just define, for $f \in C_l(E^N)$,

$$\overline{S}_t^N \bar{f}(x) = \begin{cases} f(\infty) + S_t^N f_0(x), & \text{if } x \in E^N, \\ f(\infty), & \text{otherwise,} \end{cases}$$

where $f(\infty) = \lim_{|x|\to\infty} f(x)$ and f_0 is defined by $f_0(x) = f(x) - f(\infty)$. It follows by Assumption U that the semigroup S_t^N has a transition density (see the Appendix in [8] for the uniformly elliptic case and Section 5.6 in [15] for the linear case). Then one can use the estimates for the transition density stated in $0.24.C_2$ in Section 6 in the Appendix in [8] to prove that $D(\overline{\mathbb{R}}^N)$ is invariant under $\overline{S}_t^N$ in the uniformly elliptic case. We have, by direct calculation, the same conclusion in the linear case, since in this case the transition density is Gaussian and has an explicit form (see Section 5.6 in [15]). □

LEMMA A.4.3. *Let* X *be a solution of the martingale problem for* $(\{(F_f, \mathscr{L} F_f): f \in \mathbf{D}\}, \delta_\nu)$. *Then the moment problem for* $X_t(1)$ *is well posed for* $t \geq 0$.

PROOF. The total mass process $X_t(1)$ is a Feller diffusion, that is, a continuous state branching process. By the corresponding theory it follows (see equations (2.3) and (2.14) in [19]) that the Laplace transform of its transition distribution function $P_t(x, \cdot)$ can be written in the form

$$\int_0^\infty e^{-\rho y} P_t(x, dy) = \exp(-x\Psi_t(\rho)) \quad \text{for } \rho \geq 0,$$

where

$$\Psi_t(\rho) = \frac{\rho e^{\lambda\gamma t}}{1 + (\lambda\sigma^2\rho/2)c(t, \lambda\gamma)}$$

and

$$c(t, a) = \begin{cases} t, & \text{if } a = 0, \\ (e^{at} - 1)/a, & \text{if } a > 0. \end{cases}$$

Since the complex variable function $g(z) = \exp(-(\kappa_1 z/(1+\kappa_2 z))$, (where κ_1, κ_2 are real constants and $\kappa_2 > 0$) defined on an appropriate neighborhood of 0 is analytic, we have that $\exp(-x\Psi_t(\rho))$ has derivatives of any order with respect to ρ at 0. Therefore, $X_t(1)$ has moments of any order given by

$$m_k = E(X_t(1))^k = (-1)^k \frac{\partial^k}{\partial \rho^k} \exp(-x\Psi_t(\rho))|_{\rho=0}, \qquad k = 1, 2, \dots,$$

where $x = X_0(1) = \nu(1)$ and by the Taylor series expansion we have

$$\sum_{k=1}^{\infty} m_k \frac{r^k}{k!} = \exp(-x\Psi_t(-r)) < \infty \qquad \text{for } |r| < r_0, \text{ for some } r_0 > 0.$$

Finally, application of Theorem 30.1 in [3] completes the proof. □

PROOF OF THEOREM. Let $f \in \mathbf{D}$ and $\mu \in M_F(\overline{\mathbb{R}})$. A few lines of calculations show that

$$\int_{\overline{\mathbb{R}}} \frac{\delta F_f(\mu)}{\delta \mu(x)} \mu(dx) = N(f) F_f(\mu),$$

$$\frac{1}{2}\delta \int_{\overline{\mathbb{R}}} \frac{\delta^2 F_f(\mu)}{\delta \mu(x)^2} \mu(dx) = \frac{1}{2} \sum_{\substack{p,q=1\\ p\neq q}}^{N(f)} \Big(F_{B_0^{pq} f}(\mu) - F_f(\mu) \Big) + \frac{1}{2} N(f)(N(f)-1) F_f(\mu),$$

where $B_0\colon D(\overline{\mathbb{R}}^2) \mapsto D(\overline{\mathbb{R}})$ is defined by $B_0 g(x, y) = \delta g(x, x)$ and $B_0^{pq} f$ denotes the action of B_0 acting on f as a function of the pth and qth variables and

$$\int_{\overline{\mathbb{R}}} L\left(\frac{\delta F_f(\mu)}{\delta \mu(x)}\right) \mu(dx) + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \int_{\overline{\mathbb{R}}} \int_{\overline{\mathbb{R}}} a_{ij}^{(m)}(x, y) \frac{\partial^2}{\partial x_i \partial y_j} \left(\frac{\delta^2 F_f(\mu)}{\delta \mu(x) \delta \mu(y)} \right) \times \mu(dx)\mu(dy) = F_{\overline{G}_{N(f)} f}(\mu),$$

where $\overline{G}_N$ are the operators defined by (3.9). Therefore the operator $\mathscr{L}$ can be written in the form

$$\mathscr{L} F_f(\mu) = F_{\overline{G}_{N(f)} f}(\mu) + \tfrac{1}{2} \sum_{\substack{p,q=1\\ p\neq q}} \Big(F_{B_0^{pq} f}(\mu) - F_f(\mu) \Big) + V(N(f)) F_f(\mu)$$

for $f \in \mathbf{D}$, where $V(n) = \xi n + \frac{1}{2} n(n-1)$. Note that for every $N \geq 1$, $1_{(N)}$ (the function of N variables which assumes the constant value 1) is in $D(\overline{\mathbb{R}}^N)$, $\overline{G}_N 1_{(N)} = 0$ and $\mathscr{L} F_{1_{(N)}}(\mu) = N(\mu(1))^{N-1}(\xi\mu(1) + \frac{1}{2}\delta(N-1)) \geq 0$ for all $\mu \in M_F(\overline{\mathbb{R}})$. Finally, we see, using Lemmas A.4.2 and A.4.3, that the Hypotheses 4.3 in Theorem 4.4 in [7] is satisfied. Therefore we conclude that the martingale problem for $(\{(F_f, \mathscr{L} F_f)\colon f \in \mathbf{D}\}, \delta_\nu)$ has a unique solution. □

REFERENCES

[1] ADLER, R. J. (1992). Superprocess local and intersection local times and their corresponding particle pictures. In *Seminar on Stochastic Processes* (E. Çinlar, K. L. Chung, and M. J. Sharpe, eds.) 1–42. Birkhäuser, Boston.

[2] ADLER, R. J. (1997). Superprocesses and plankton dynamics. *Aha Huliko'a* 121–128.

[3] BILLINGSLEY, P. (1979). *Probability and Measure*. Wiley, New York.

[4] BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42.

[5] DA PRATO, G. and TUBARO, L. (1996). Fully nonlinear stochastic partial differential equations. *SIAM J. Math. Anal.* **27** 40–55.

[6] DAWSON, D. (1993). Measure-valued Markov processes. *Lecture Notes in Math.* **1541** 1–260. Springer, Berlin.

[7] DAWSON, D. and KURTZ, T. G. (1982). Applications of duality to measure-valued diffusion processes. *Lecture Notes in Control and Inform. Sci.* **42** 91–105. Springer, Berlin.

[8] DYNKIN, E. B. (1965). *Markov Processes* **2**. Springer, Berlin.

[9] DYNKIN, E. B. (1988). Representation for functionals of superprocesses by multiple stochastic integrals, with application to self-intersection local times. *Astérisque* **157–158** 147–171.

[10] DYNKIN, E. B. (1989). Superprocesses and their linear additive functional. *Trans. Amer. Math. Soc.* **314** 255–282.

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

[12] FINGER, C. C. (1996). *Branching particle systems on a flow*. Ph.D. dissertation, Princeton Univ.

[13] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, New York.

[14] JACOD, J. and SHIRYAEV, A. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.

[15] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer, New York.

[16] KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations*. Cambridge Univ. Press.

[17] LAHA, R. G. and ROHATGI, V. K. (1979). *Probability Theory*. Wiley, New York.

[18] LALLI, C. M. and PARSON, T. R. (1993). *Biological Oceanography: An Introduction*. Butterworth-Heinemann, Oxford.

[19] LAMPERTI, J. (1967). The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete* **7** 271–288.

[20] PERKINS, E. (2000). Dawson–Watanabe superprocesses and measure valued diffusions, *Lecture Notes in Math.* Springer, Berlin. To appear.

[21] ROELLY-COPPOLETTA, S. (1986). A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics* **17** 43–65.

[22] ROGERS, L. C. G. and WILLIAMS, D. (1994). *Diffusions, Markov Processes and Martingales* **2***: Itô Calculus*. Wiley, New York.

[23] ROZOVSKII, B. (1991). *Stochastic Evolution Systems*. Kluwer, Amsterdam.

[24] SKOULAKIS, G. (1999). Superprocesses over a stochastic flow. Ph.D. disseration, Univ. North Carolina.

[25] TWARDOWSKA, K. (1996). Wong–Zakai approximations for stochastic differential equations. *Acta Appl. Math.* **43** 317–359.

[26] WANG, H. (1997). State classification for a class of measure-valued branching diffusions in a Brownian medium. *Probab. Theory Related Fields* **109** 39–55.

[27] WANG, H. (1998). A class of measure-valued branching diffusions in a random medium. *Stochastic Anal. Appl*. **16** 753–786.

[28] ZIRBEL, C. L. (1993). *Stochastic Flows: Dispersion of a mass distribution and Lagrangian observations of a random field*. Ph.D. dissertation, Princeton Univ.

[29] ZIRBEL, C. L. (1997). Mean occupation times of continuous one-dimensional Markov processes. *Stochastic Processes Appl*. **69** 161–178.

[30] ZIRBEL, C. L. (1997). Translation and dispersion of mass by isotropic Brownian flows. *Stochastic Process. Appl*. **70** 1–29.

[31] ZIRBEL, C. L. and ÇINLAR, E. (1997). Dispersion of particle systems in Brownian flows. *Adv. Appl. Probab.* **28** 53–74.

[32] ZIRBEL, C. L. and ÇINLAR, E. (1997). Mass transport by Brownian flows. In *Stochastic Models in Geosystems* (S. A. Molchanov and W. A. Woyczynski, eds.) 459–492. Springer, Berlin.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27599-3260
E-MAIL: georgios@email.unc.edu

FACULTY OF INDUSTRIAL ENGINEERING
AND MANAGEMENT
TECHNION
HAIFA 32000
ISRAEL
E-MAIL: robert@ieadler.technion.ac.il