Existence of probability measure valued jump-diffusions in generalized Wasserstein spaces[*][†]

Martin Larsson[‡]   Sara Svaluto-Ferro[§]

Abstract

We study existence of probability measure valued jump-diffusions described by martingale problems. We develop a simple device that allows us to embed Wasserstein spaces and other similar spaces of probability measures into locally compact spaces where classical existence theory for martingale problems can be applied. The method allows for general dynamics including drift, diffusion, and possibly infinite-activity jumps. We also develop tools for verifying the required conditions on the generator, including the positive maximum principle and certain continuity and growth conditions. To illustrate the abstract results, we consider large particle systems with mean-field interaction and common noise.

Keywords: probability measure valued processes; martingale problem; Wasserstein spaces; positive maximum principle; McKean–Vlasov equations.
MSC2020 subject classifications: 60J60; 60J75; 60G57.
Submitted to EJP on August 22, 2019, final version accepted on November 29, 2020.
Supersedes arXiv:1908.08080v2.

Contents

[‡]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA.
E-mail: martinl@andrew.cmu.edu
[§]Faculty of Mathematics, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria.
E-mail: sara.svaluto-ferro@univie.ac.at

1 Introduction

In this paper we study existence of probability measure valued jump-diffusions, whose dynamics is specified by means of a martingale problem. Processes taking values in spaces of probability measures play an important role in a number of applied contexts. This includes population genetics (see Etheridge (2011) for an overview), stochastic partial differential equations (see e.g. Florchinger and Le Gland (1992) and Kurtz and Xiong (1999) among many others), statistical physics (see Huang (1987) for an overview), optimal transport (see Villani (2008) for an overview), and mathematical finance, in particular stochastic optimal control, McKean–Vlasov equations, and mean field games (see e.g. Carmona and Delarue (2017) and the references given there) and stochastic portfolio theory (see e.g. Fernholz (2002), Fernholz and Karatzas (2009), and Cuchiero (2019)).

The mathematical theory of probability measure valued processes has a long history going back to Watanabe (1968), Dawson (1977, 1978), and Fleming and Viot (1979). We refrain from a full literature review, but only mention the remarkable collection of St. Flour lecture notes of Sznitman (1991), Dawson (1993), and Perkins (2002), as well as the work of Ethier and Kurtz (1987, 1993, 2005).

Much of the classical literature on measure valued processes works with the weak topology on the space $M_1(\mathbb{R}^d)$ of all probability measures on $\mathbb{R}^d$ (or some other relevant underlying spaces). There are however other interesting topologies that one can place on spaces of probability measures, that are more appropriate in certain situations. Prominent examples are topologies induced by Wasserstein metrics on the spaces $\mathcal{P}_p(\mathbb{R}^d)$ of probability measures with finite p-th moments. A basic reason for considering such stronger topologies is to ensure that the quantities which are naturally associated to the current state of the model are continuous functions on the state space.

The price to pay is that the classical existence theory for martingale problems becomes more difficult to apply. As a result, most proofs of existence of measure valued processes proceed instead via interacting particle systems and a passage to the large-population limit (see for instance the approach presented by Dawson and Vaillancourt (1995)). In this paper we prove existence for the limiting system directly, without passing through particle systems.

A key difficulty in using the martingale problem is related to the fact that (say) the Wasserstein space $\mathcal{P}_p(\mathbb{R}^d)$ is not straightforward to compactify. To illustrate this, consider first $M_1(\mathbb{R}^d)$ with the topology of weak convergence. This space fails to be locally compact, and hence does not admit a standard one-point compactification. However, the space $M_1((\mathbb{R}^d)^\Delta)$ of probability measures on the one-point compactification of $\mathbb{R}^d$ *is* compact, and thus fits naturally with classical machinery. This simple procedure does not work for $\mathcal{P}_p(\mathbb{R}^d)$.

In this paper we develop a simple device for embedding $\mathcal{P}_p(\mathbb{R}^d)$, and other similar spaces, into compact spaces where the classical existence theory of martingale problems can be applied. This allows us to establish existence of solutions for martingale problems

in spaces of this kind. The operators for which the martingale problem is solved can be very general, including both drift, diffusion, and jumps which can be of infinite activity and even non-summable.

We start in Section 2 by reviewing some facts about martingale problems. The core of the paper is Section 3, where we state and prove our main abstract result, Theorem 3.4. There we consider a linear operator L on a carefully chosen domain of test functions. A key assumption on L is, as one would expect, that it satisfy the positive maximum principle. Since L acts on functions of probability measures, it may not be obvious how to verify the positive maximum principle in practice. To remedy this, we develop necessary conditions for optimality, see Theorem 5.1, that can be used to verify the positive maximum principle for operators of *Lévy type*, introduced in Section 4. This extends results in Cuchiero et al. (2019). Furthermore, in addition to the positive maximum principle, we impose certain continuity and growth conditions on L. In Section 6 we develop tools to aid the verification of these conditions. Finally, in Sections 7 and 8, we discuss some applications that illustrate the scope of the abstract theory. These applications are primarily related to large particle systems with mean-field interaction, where the particles are subject to common noise. In such systems, the limiting empirical distribution of the particles evolves as a probability measure valued stochastic process, whose dynamics can often be described in terms of a martingale problem of the type considered here.

The following notation is used throughout the paper. For a locally compact Polish space E, we let $M_+(E)$ denote the Polish space of positive measures on E, and $M_1(E)$ the subspace of probability measures. We also write $M(E) = M_+(E) - M_+(E)$ for the space of signed measures on E of bounded variation. These spaces are sometimes considered with the topology of weak convergence (defined using bounded continuous functions and denoted $\mu_n \Rightarrow \mu$) or vague convergence (defined using continuous functions vanishing at infinity). We remark that if E is compact, then $M_1(E)$ is compact and $M_+(E)$ is locally compact. However, if E is noncompact, $M_1(E)$ is not even locally compact. See for instance Remark 13.14(iii) and Corollary 13.30 in Klenke (2013) for more details. For a Polish space $\mathcal{X}$, we let $C(\mathcal{X})$ denote the space of all continuous functions $f\colon \mathcal{X} \to \mathbb{R}$. Subscripts 0 and c indicate that the functions are also vanishing at infinity and have compact support, respectively. If present, a superscript indicates their degree of continuous differentiability.

2 Martingale problems and the positive maximum principle

Let $\mathcal{X}$ be a Polish space, $\mathcal{D} \subseteq C(\mathcal{X})$ a linear subspace, and consider a linear operator

$$L\colon \mathcal{D} \to C(\mathcal{X}). \tag{2.1}$$

In this paper, $\mathcal{X}$ will be a subset of $M(E)$ for some closed subset $E \subseteq \mathbb{R}^d$, or of $M(E^\Delta)$ where E^Δ is the one-point compactification of E. The topology on $\mathcal{X}$ will however not always be the subspace topology (i.e. the topology of weak convergence). Moreover, the functions in $\mathcal{D}$ will usually be defined on a larger subset of $M(E)$ than $\mathcal{X}$, in which case the condition $\mathcal{D} \subseteq C(\mathcal{X})$ just means that $f|_{\mathcal{X}} \in C(\mathcal{X})$ for every $f \in \mathcal{D}$.

Definition 2.1. *An $\mathcal{X}$-valued càdlàg process X, defined on some filtered probability space, is called a solution to the martingale problem for $(L, \mathcal{D}, \mathcal{X})$ with initial condition $\mu \in \mathcal{X}$ if $X_0 = \mu$ and*

$$f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds, \quad t \geq 0,$$

is a local martingale for every $f \in \mathcal{D}$.

It is convenient to allow solutions to the martingale problem to leave the state space. If $\mathcal{X}$ is locally compact, this is formalized via a one-point compactification of $\mathcal{X}$. A similar procedure works more generally. Fix a cemetery state $\dagger \notin \mathcal{X}$. Define $\mathcal{X}^\dagger = \mathcal{X} \cup \{\dagger\}$, and let $\mathcal{D}^\dagger$ consist of all $f \colon \mathcal{X}^\dagger \to \mathbb{R}$ such that $(f - f(\dagger))|_{\mathcal{X}} \in \mathcal{D}$. For every $f \in \mathcal{D}^\dagger$, define a function $L^\dagger f \colon \mathcal{X}^\dagger \to \mathbb{R}$ by $L^\dagger f|_{\mathcal{X}} = L((f - f(\dagger))|_{\mathcal{X}})$ and $L^\dagger f(\dagger) = 0$. Assume that the given Polish topology on $\mathcal{X}$ can be extended to a Polish topology on $\mathcal{X}^\dagger$ in such a way that both $\mathcal{D}^\dagger$ and $L^\dagger(\mathcal{D}^\dagger)$ are contained in $C(\mathcal{X}^\dagger)$. For example, this is the case if $\mathcal{X}$ is locally compact, $\mathcal{X}^\dagger$ is the one-point compactification, and both $\mathcal{D}$ and $L(\mathcal{D})$ are contained in $C_0(\mathcal{X})$. Observe that $\dagger$ may or may not be an isolated point. If $\mathcal{X}$ is not locally compact, then the one-point compactification is not available, and other constructions must be used. This situation arises, for instance, when $\mathcal{X}$ is a Wasserstein space of probability measures.

Definition 2.2. *A solution X to the martingale problem for $(L^\dagger, \mathcal{D}^\dagger, \mathcal{X}^\dagger)$ with initial condition $\mu \in \mathcal{X}$ is called a possibly killed solution to the martingale problem for $(L, \mathcal{D}, \mathcal{X})$ with initial condition μ.*[1]

For definiteness, we now suppose that $\mathcal{X}$ is a subset of $M(E)$. We also suppose that for $f \in \mathcal{D}$, both f and Lf are defined on all of $M(E)$. The following classical definition is useful because it can very often be checked in practice.

Definition 2.3. *L satisfies the positive maximum principle on $\mathcal{X}$ at $\mu \in M(E)$ if*

$$f \in \mathcal{D} \text{ and } f(\mu) = \sup_{\mathcal{X}} f \geq 0 \quad \implies \quad Lf(\mu) \leq 0.$$

If this holds for all $\mu \in \mathcal{X}$, then L is said to satisfy the positive maximum principle on $\mathcal{X}$.

The positive maximum principle directly implies that $Lf|_{\mathcal{X}}$ only depends on $f|_{\mathcal{X}}$ and not on the values f takes outside $\mathcal{X}$. Thus, if L satisfies the positive maximum principle on $\mathcal{X}$, it can be regarded as an operator sending functions on $\mathcal{X}$ to functions on $\mathcal{X}$, consistent with (2.1). The positive maximum principle is linked to existence of solutions to the martingale problem. The following classical result deals with the locally compact case. The nontrivial part is the forward implication, whose proof can be found, e.g., in (Ethier and Kurtz, 2005, Theorem 4.5.4).

Theorem 2.4. *Assume $\mathcal{X}$ is locally compact, $\mathcal{D} \subseteq C_0(\mathcal{X})$ is dense, and $L(\mathcal{D}) \subseteq C_0(\mathcal{X})$. Then L satisfies the positive maximum principle on $\mathcal{X}$ if and only if there exists a possibly killed solution to the martingale problem for $(L, \mathcal{D}, \mathcal{X})$ for every initial condition $\mu \in \mathcal{X}$.*

One is often interested in solutions that are not killed. A general condition for this is that there exist functions $f_n \in \mathcal{D}$ such that $f_n \to 1$ and $(Lf_n)^- \to 0$ in the bounded pointwise sense. This follows from a slight modification of (Ethier and Kurtz, 2005, Theorem 4.3.8 and Remark 4.5.5).

Since $M_1(E)$ is compact whenever $E \subset \mathbb{R}^d$ is compact, we obtain the following result as a direct application of Theorem 2.4.

Corollary 2.5. *Let $\mathcal{X} = M_1(E)$ with $E \subset \mathbb{R}^d$ compact. Assume $\mathcal{D} \subseteq C(\mathcal{X})$ is a dense subset containing the constant function 1, and $L(\mathcal{D}) \subseteq C(\mathcal{X})$. Then L satisfies the positive maximum principle on $\mathcal{X}$ if and only if there exists a possibly killed solution to the martingale problem for $(L, \mathcal{D}, \mathcal{X})$ for every initial condition $\mu \in \mathcal{X}$. If additionally $L1 = 0$, then every such solution X satisfies $X_t \in M_1(E)$ for all $t \geq 0$, and is thus a solution to the martingale problem for $(L, \mathcal{D}, \mathcal{X})$.*

---

[1]In our terminology, a solution can be killed either by jumping to the cemetery state $\dagger$, or by reaching it continuously by means of an "explosion".

3 Main result

Let $w \colon \mathbb{R}^d \to [1, \infty)$ be a C^∞ function such that

$$\lim_{|x| \to \infty} w(x) = \infty, \tag{3.1}$$

and fix a closed subset $E \subseteq \mathbb{R}^d$. Define the set of probability measures on E with finite w-moment,

$$\mathcal{P}_w = \mathcal{P}_w(E) = \{\mu \in M_1(E) \colon \langle w, \mu \rangle < \infty\},$$

topologized by the following notion of convergence: $\mu_n \to \mu$ if and only if $\mu_n \Rightarrow \mu$ and $\langle w, \mu_n \rangle \to \langle w, \mu \rangle$. This turns $\mathcal{P}_w$ into a Polish space. A possible choice of metric is

$$d_w(\mu_1, \mu_2) = d(w\mu_1, w\mu_2), \tag{3.2}$$

where $d(\,\cdot\,, \,\cdot\,)$ is the Prokhorov metric on $M_+(E)$, and the measures $w\mu_i$ are given by $(w\mu_i)(dx) = w(x)\mu_i(dx)$. The Prokhorov metric is discussed in detail in Section 3.1 of Ethier and Kurtz (2005). See also the discussion after Example A.42 in Föllmer and Schied (2004).

Example 3.1. If $w(x) = |x|^p$ outside some ball around the origin, then $\mathcal{P}_w$ is the set of probability measures on E with finite p-th moments, and (3.2) generates the same topology as the Wasserstein p-distance $\mathcal{W}_p$. $\diamond$

We will use the following class of test functions:

$$\mathcal{D}_w = \text{algebra generated by all } \mu \mapsto \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle} \text{ with } \varphi \in C_c^\infty(\mathbb{R}^d).^2 \tag{3.3}$$

With the convention $\exp(-\langle w, \mu \rangle) = 0$ if $\langle w, |\mu| \rangle = \infty$, functions in $\mathcal{D}_w$ can be evaluated at any $\mu \in M(E)$. We will obtain possibly killed solutions to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$, where L is an operator satisfying suitable assumptions. In order to do so, fix a cemetery state $\dagger$ and define $\mathcal{P}_w^\dagger = \mathcal{P}_w \cup \{\dagger\}$.[3] The topology is extended to $\mathcal{P}_w^\dagger$ by declaring that a sequence of measures $\mu_n \in \mathcal{P}_w$ converges to $\dagger$ if $\langle w, \mu_n \rangle \to \infty$. Thus $\lim_{\mu \to \dagger} f(\mu) = 0$ for any $f \in \mathcal{D}_w$, so that $\mathcal{D}_w^\dagger$ as defined in Section 2 is indeed contained in $C(\mathcal{P}_w^\dagger)$. If one assumes that $\lim_{\mu \to \dagger} Lf(\mu) = 0$ for every $f \in \mathcal{D}_w$, which we shall, it follows that $L^\dagger(\mathcal{D}^\dagger) \subseteq C(\mathcal{P}_w^\dagger)$ as well, where $L^\dagger$ is defined as in Section 2. This allows us to speak about possibly killed solutions to the martingale problem.

If E is compact, then $\mathcal{P}_w = M_1(E)$ is also compact, and Corollary 2.5 yields a satisfactory existence theory for the martingale problem. From now on we consider the opposite situation, and assume that

$$E \text{ is not compact.}$$

In this case $\mathcal{P}_w$ is not even locally compact, and the classical results are not directly applicable. Instead, we will embed $\mathcal{P}_w$ into a space that is locally compact, where Theorem 2.4 can be applied. To describe this embedding, let $E^\Delta = E \cup \{\Delta\}$ be the one-point compactification of E, for some $\Delta \notin E$. The space $M_+(E^\Delta)$ is equipped with the weak topology. Define a map

$$T \colon \mathcal{P}_w \to M_+(E^\Delta), \quad T(\mu)(dx) = w(x)\mu(dx \cap E), \tag{3.4}$$

which is a topological embedding of $\mathcal{P}_w$ into $M_+(E^\Delta)$. Recall that $w \geq 1$ and observe that

$$T(\mathcal{P}_w) = \{\nu \in M_+(E^\Delta) \colon \langle w^{-1}, \nu \rangle = 1, \ \nu(\{\Delta\}) = 0\},$$

---

[2] That is, $\mathcal{D}_w$ consists of all sums of products of functions of the given form $\langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle}$. It does not contain the constant function 1.

[3] We may take any $\dagger \notin M(E^\Delta)$, where E^Δ is the one-point compactification of E.

where $w^{-1}(x) = 1/w(x)$, which is well defined everywhere on E^Δ with the convention $w^{-1}(\Delta) = \lim_{|x| \to \infty} w^{-1}(x)$. Let $\mathcal{X}$ denote the weak closure of $T(\mathcal{P}_w)$; this will serve as state space for an auxiliary martingale problem. Since $\mathcal{X}$ is a closed subset of the locally compact Polish space $M_+(E^\Delta)$, it is itself locally compact Polish. This places us in the framework of Theorem 2.4. Note that we have the explicit description

$$\mathcal{X} = \{\nu \in M_+(E^\Delta) \colon \langle w^{-1}, \nu \rangle = 1\}. \tag{3.5}$$

In particular, a measure $\nu \in \mathcal{X}$ lies in $T(\mathcal{P}_w)$ if and only if it does not charge Δ.

Using T, any martingale problem with state space $\mathcal{P}_w$ and operator $f \mapsto Lf$ can be regarded as a martingale problem with state space $T(\mathcal{P}_w)$ and operator $\widetilde{f} \mapsto L(\widetilde{f} \circ T) \circ T^{-1}$. Our strategy is to extend this to a martingale problem with state space $\mathcal{X}$ and, then, show that the solution does not charge Δ and thus actually lies in $T(\mathcal{P}_w)$. This gives a solution to the original martingale problem.

These steps depend in a somewhat delicate way on the particular choice (3.3) of test functions. In particular, in order to apply Theorem 2.4, the function $f \circ T^{-1}$ obtained by pushing forward a function $f \in \mathcal{D}_w$ using T needs to be extendible to a function in $C_0(\mathcal{X})$. This is captured by the following definition.

Definition 3.2. *A function $f \colon \mathcal{P}_w \to \mathbb{R}$ is of C_0 type if $f \circ T^{-1} \colon T(\mathcal{P}_w) \to \mathbb{R}$ extends to a C_0 function on $\mathcal{X}$. This extension is again denoted by $f \circ T^{-1}$.*

It is clear that sums and products of functions of C_0 type are again of C_0 type; these functions thus form an algebra. Since $\nu \mapsto \langle \varphi, T^{-1}(\nu) \rangle = \langle \varphi w^{-1}, \nu \rangle$ is continuous on $\mathcal{X}$ for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, the product of $\langle \varphi, \mu \rangle$ and a function of C_0 type is again of C_0 type. Also, $\mu \mapsto e^{-\langle w, \mu \rangle}$ is certainly of C_0 type. We deduce in particular that every $f \in \mathcal{D}_w$ is of C_0 type.

Example 3.3. Suppose $E = \mathbb{R}$ and let $f(\mu) = \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle}$ for some $\varphi \in C(\mathbb{R})$. When is f of C_0 type? Set $\mu_n = w(n)^{-1}\delta_n + (1 - w(n)^{-1})\delta_1 \in \mathcal{P}_w$. Then μ_n does not converge to any element of $\mathcal{P}_w$, but $\nu_n = T(\mu_n) = \delta_n + w(1)(1 - w(n)^{-1})\delta_1$ converges to $\delta_\Delta + w(1)\delta_1$ in $M_+(\mathbb{R}^\Delta)$. On the other hand, $\lim_n f \circ T^{-1}(\nu_n) = (\lim_n \varphi(n)/w(n) + \varphi(1))e^{-1-w(1)}$ only exists if $\varphi(n)/w(n)$ has a finite limit. By considering similar sequences μ_n, one sees that f is of C_0 type if and only if $\varphi(x)/w(x)$ has a finite limit as $x \to \Delta$. $\diamond$

The following is the main result of this paper. To state it, we define the compact subset $\mathcal{X}_c = \{\nu \in \mathcal{X} \colon \langle 1, \nu \rangle \leq c\}$ for any constant $c \geq 1$. The meaning of its conditions, and examples of how they can be verified, are discussed in later sections.

Theorem 3.4. *Consider a linear operator $L \colon \mathcal{D}_w \to C(\mathcal{P}_w)$, and assume the following conditions are satisfied:*

(i) *L satisfies the positive maximum principle on $\mathcal{P}_w$,*

(ii) *Lf is of C_0 type for every $f \in \mathcal{D}_w$,*

(iii) *for every constant $c \geq 1$, there exist a function $\widetilde{f} \colon \mathcal{X} \to \mathbb{R}$ and pairs $(\widetilde{f}_m, \widetilde{g}_m)$ in the bp-closure of the restricted graph*

$$\{(\widetilde{f}, \widetilde{g}) \in C_0(\mathcal{X}) \times C(\mathcal{X}_c) \colon \widetilde{f} \circ T \in \mathcal{D}_w, \ \widetilde{g} = L(\widetilde{f} \circ T) \circ T^{-1}|_{\mathcal{X}_c}\} \tag{3.6}$$

such that $(\widetilde{f}_m, \widetilde{g}_m^+)$ are uniformly bounded in m, and

(a) *$\widetilde{f}_m \to \widetilde{f}$ pointwise, $\widetilde{f} \geq 0$, and $\widetilde{f}(\nu) = 0$ for $\nu \in \mathcal{X}_c$ if and only if $\nu(\{\Delta\}) = 0$,*

(b) *$\limsup_{m \to \infty} \widetilde{g}_m^+ \leq c' \widetilde{f}|_{\mathcal{X}_c}$ pointwise for some constant c'.*

Then there exists a possibly killed solution to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for every initial condition $\mu \in \mathcal{P}_w$. Furthermore, assume that

(iv) *there exist pairs (f_n, g_n) in the bp-closure of the graph $\{(f, Lf) \colon f \in \mathcal{D}_w\}$ of L such that $(f_n, g_n^-) \to (1, 0)$ in the bounded pointwise sense.*

Then every possibly killed solution X to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ satisfies $X_t \in \mathcal{P}_w$ for all $t \geq 0$, and is thus a solution to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$.

Remark 3.5. While our focus is the case where (3.1) holds, one could also take $w \equiv 1$. In this case $\mathcal{P}_w = M_1(E)$ is the set of all probability measures on E with the topology of weak convergence. A slight modification of our main result holds also for this case. Specifically, letting $\mathcal{D}_w$ denote the algebra generated by $\langle \varphi, \mu \rangle$ with $\varphi \in \mathbb{R} + C_c^\infty(\mathbb{R}^d)$, Theorem 3.4 remains true as stated. Note that condition (iii) only needs to be verified for $c = 1$. The proof remains unchanged, apart from slightly different arguments in Lemma 3.6(i)–(ii) below.

The rest of this section is devoted to the proof of Theorem 3.4, so we now assume that its conditions are satisfied. As discussed above, the proof uses the embedding T in (3.4) to transform the original martingale problem into an auxiliary martingale problem on the state space $\mathcal{X}$ in (3.5). The domain of test functions for the auxiliary martingale problem is

$$\mathcal{D} = \text{algebra generated by all } \nu \mapsto \langle \varphi, \nu \rangle e^{-\langle 1, \nu \rangle} \text{ with } \varphi \in C_c^\infty(\mathbb{R}^d). \qquad (3.7)$$

The elements of $\mathcal{D}$ can be evaluated at any $\nu \in M(E^\Delta)$, with the conventions $\varphi(\Delta) = 0$ and $1(\Delta) = 1$. Note also that $\mathcal{D} \subset C_0(\mathcal{X})$, and that its elements $\widetilde{f}$ satisfy $\widetilde{f} \circ T \in \mathcal{D}_w$. Due to Theorem 3.4(ii), we can then define a linear operator $\widetilde{L} \colon \mathcal{D} \to C_0(\mathcal{X})$ by the formula

$$\widetilde{L}\widetilde{f} = L(\widetilde{f} \circ T) \circ T^{-1}. \qquad (3.8)$$

Lemma 3.6. *We have the following properties.*

(i) *$\mathcal{D}$ is dense in $C_0(\mathcal{X})$,*

(ii) *for every $\widetilde{f} \in \mathcal{D}$ and every $\nu^* \in \mathcal{X}$, there exist measures $\nu_n \in T(\mathcal{P}_w)$ with $\nu_n \Rightarrow \nu^*$ and $\widetilde{f}(\nu_n) = \widetilde{f}(\nu^*)$ for all $n \in \mathbb{N}$.*

Proof. (i): This follows from the Stone–Weierstrass theorem once we show that $\mathcal{D}$ separates points and vanishes nowhere on $\mathcal{X}$. Any $\nu \in \mathcal{X}$ satisfies $\langle w^{-1}, \nu \rangle = 1$, which implies that $\nu(E) > 0$. It is then clear that some element of $\mathcal{D}$ is nonzero at ν. Thus $\mathcal{D}$ vanishes nowhere. Next, take $\nu_1, \nu_2 \in \mathcal{X}$ such that $\langle \varphi, \nu_1 \rangle e^{-\langle 1, \nu_1 \rangle} = \langle \varphi, \nu_2 \rangle e^{-\langle 1, \nu_2 \rangle}$ for all $\varphi \in C_c^\infty(\mathbb{R}^d)$. By considering a sequence $\varphi_n \uparrow w^{-1}$ and using that $\langle w^{-1}, \nu_i \rangle = 1$, $i = 1, 2$, we deduce that $\langle 1, \nu_1 \rangle = \langle 1, \nu_2 \rangle$. Thus $\langle \varphi, \nu_1 \rangle = \langle \varphi, \nu_2 \rangle$ for all $\varphi \in C_c^\infty(\mathbb{R}^d)$, which implies that $\nu_1(\cdot \cap E) = \nu_2(\cdot \cap E)$. It follows that $\nu_1 = \nu_2$, so that $\mathcal{D}$ separates points as required.

(ii): If ν^* itself lies in $T(\mathcal{P}_w)$, simply take $\nu_n = \nu^*$ for all n. We thus assume that this is not the case, which means that $\nu^* = \nu_0 + \lambda \delta_\Delta$ for some $\nu_0 \in M_+(E)$ and $\lambda > 0$. Fix now a sequence $(x_n)_{n \in \mathbb{N}} \subseteq E$ with $|x_n| \to \infty$, or equivalently, $x_n \to \Delta$. Since $w^{-1}(\Delta) = 0$, we have $\langle w^{-1}, \nu_0 \rangle = \langle w^{-1}, \nu^* \rangle = 1$. For numbers $t_n \in (1, \infty)$ to be determined later, define the measures

$$\nu_n = (1 - t_n^{-1})\nu_0 + t_n^{-1} w(x_n)\delta_{x_n}. \qquad (3.9)$$

Since $\langle w^{-1}, \nu_n \rangle = 1$, these measures lie in $T(\mathcal{P}_w)$. Fix any $\widetilde{f} \in \mathcal{D}$ and observe that we have

$$\widetilde{f}(\nu) = p(\langle \varphi_1, \nu \rangle e^{-\langle 1, \nu \rangle}, \dots, \langle \varphi_m, \nu \rangle e^{-\langle 1, \nu \rangle})$$

for some polynomial p on $\mathbb{R}^m$ and some $\varphi_1, \dots, \varphi_m \in C_c^\infty(\mathbb{R}^d)$. For all sufficiently large n, x_n lies outside the supports of all the φ_i. For all such n, we have

$$\langle \varphi_i, \nu_n \rangle e^{-\langle 1, \nu_n \rangle} = \langle \varphi_i, \nu_0 \rangle (1 - t_n^{-1}) e^{-(1 - t_n^{-1})\langle 1, \nu_0 \rangle - t_n^{-1} w(x_n)}, \quad i = 1, \dots, m.$$

On the other hand, we have

$$\langle \varphi_i, \nu^* \rangle e^{-\langle 1, \nu^* \rangle} = \langle \varphi_i, \nu_0 \rangle e^{-\langle 1, \nu_0 \rangle - \lambda}, \quad i = 1, \dots, m.$$

Therefore, if t_n is chosen so that

$$w(x_n) = t_n \log(1 - t_n^{-1}) + \langle 1, \nu_0 \rangle + t_n \lambda, \qquad (3.10)$$

it follows that $\widetilde{f}(\nu_n) = \widetilde{f}(\nu^*)$. To see that this is possible, let $\alpha(t)$ denote the right-hand side of (3.10), with t_n replaced by t. Then $t \mapsto \alpha(t)$ is continuous and strictly increasing on $(1, \infty)$ with $\lim_{t \to \infty} \alpha(t) = \infty$ and $\lim_{t \to -\infty} \alpha(t) = -\infty$. Therefore α has a continuous inverse $\alpha^{-1}(s)$ which satisfies $\lim_{s \to \infty} \alpha^{-1}(s) = \infty$. We now define

$$t_n = \alpha^{-1}(w(x_n)).$$

Since $\lim_{|x| \to \infty} w(x) = \infty$, we have $t_n \to \infty$, and since (3.10) holds, we have $t_n^{-1} w(x_n) \to \lambda$. It is then clear from (3.9) that $\nu_n \Rightarrow \nu^*$. Therefore, after discarding the finitely many ν_n for which x_n lies in the support of some φ_i, the measures ν_n satisfy the desired properties. $\qquad\square$

Lemma 3.7. *The operator $\widetilde{L}$ satisfies the positive maximum principle on $\mathcal{X}$.*

Proof. Let $\widetilde{f} \in \mathcal{D}$ and $\nu^* \in \mathcal{X}$ be such that $\widetilde{f}(\nu^*) = \max_{\mathcal{X}} \widetilde{f} \geq 0$. By Lemma 3.6(ii), there exist measures $\nu_n \in T(\mathcal{P}_w)$ with $\nu_n \Rightarrow \nu^*$ and $\widetilde{f}(\nu_n) = \widetilde{f}(\nu^*)$ for all $n \in \mathbb{N}$. In particular, we have $\widetilde{f}(\nu_n) = \max_{T(\mathcal{P}_w)} \widetilde{f} \geq 0$ for all n. Thus, the function $f = \widetilde{f} \circ T \in \mathcal{D}_w$ attains a nonnegative maximum over $\mathcal{P}_w$ at the point $\mu_n = T^{-1}(\nu_n)$. Since L satisfies the positive maximum principle on $\mathcal{P}_w$, we get $\widetilde{L}\widetilde{f}(\nu_n) = Lf(\mu_n) \leq 0$. Sending n to infinity and using that $\widetilde{L}\widetilde{f}$ is continuous on $\mathcal{X}$ yields $\widetilde{L}\widetilde{f}(\nu^*) \leq 0$. This shows that $\widetilde{L}$ satisfies the positive maximum principle on $\mathcal{X}$, as claimed. $\qquad\square$

Proof of Theorem 3.4. We have established that $\mathcal{X}$ is locally compact, that $\mathcal{D} \subseteq C_0(\mathcal{X})$ is dense, and that $\widetilde{L}(\mathcal{D}) \subseteq C_0(\mathcal{X})$. Since $\widetilde{L}$ satisfies the positive maximum principle on $\mathcal{X}$, Theorem 2.4 yields a possibly killed solution to the martingale problem for $(\widetilde{L}, \mathcal{D}, \mathcal{X})$ for any initial condition $\nu \in \mathcal{X}$. The state space $\mathcal{X}^\dagger$ for the possibly killed solution is the one-point compactification of $\mathcal{X}$, and $\widetilde{L}^\dagger$ and $\mathcal{D}^\dagger$ are as in Section 2.

Fix $\mu \in \mathcal{P}_w$ and let Y be a solution with initial condition $\nu_0 = T(\mu)$. We may suppose that $\dagger$ is an absorbing state, that is, $Y_t = \dagger$ for all $t \geq \inf\{t \geq 0 : Y_t = \dagger \text{ or } Y_{t-} = \dagger\}$. Assume for the moment that Y actually takes values in $T(\mathcal{P}_w) \cup \{\dagger\}$. We can then define $X = T^{-1}(Y)$, with the convention $T^{-1}(\dagger) = \dagger$. For any $f \in \mathcal{D}_w$, we have $\widetilde{f} = f \circ T^{-1} \in \mathcal{D}$ as well as $Lf = \widetilde{L}\widetilde{f} \circ T$. These identities hold on $\mathcal{P}_w \cup \{\dagger\}$. We thus obtain that

$$f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds = \widetilde{f}(Y_t) - \widetilde{f}(Y_0) - \int_0^t \widetilde{L}\widetilde{f}(Y_s)ds, \quad t \geq 0.$$

Since the right-hand side is a local martingale, it follows that X is a possibly killed solution to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ with initial condition μ.

We must still argue that Y takes values in $T(\mathcal{P}_w) \cup \{\dagger\}$. Fix any $c \geq \max\{1, \langle 1, \nu_0 \rangle\}$, and define the stopping time $\tau = \inf\{t \geq 0 : \langle 1, Y_t \rangle > c\}$, with the convention $\langle 1, \dagger \rangle = \infty$. Since Y is a possibly killed solution to the martingale problem, an application of the optional stopping theorem yields

$$\mathbb{E}[\widetilde{f}(Y_{t \wedge \tau})] = \widetilde{f}(\nu_0) + \int_0^t \mathbb{E}[\widetilde{L}^\dagger \widetilde{f}(Y_s) 1_{\{s < \tau\}}]ds$$

for every $t \geq 0$ and $\widetilde{f} \in \mathcal{D}$. Since $Y_s \in \mathcal{X}_c$ for $s < \tau$, we obtain

$$\mathbb{E}[\widetilde{f}(Y_{t \wedge \tau}) 1_{\{Y_{t \wedge \tau} \neq \dagger\}}] = \widetilde{f}(\nu_0) + \int_0^t \mathbb{E}[\widetilde{g}(Y_s) 1_{\{s < \tau\}}] ds \qquad (3.11)$$

for every $t \geq 0$ and $(\widetilde{f}, \widetilde{g})$ in the restricted graph (3.6). Since $\widetilde{f}(\dagger) = 0$, the indicator on the left-hand side of (3.11) is redundant. By dominated convergence, (3.11) remains true for all $(\widetilde{f}, \widetilde{g})$ in the bp-closure of the restricted graph (3.6). Now the indicator is needed, since these functions are not defined at $\dagger$.

Let now $\widetilde{f}$ and $(\widetilde{f}_m, \widetilde{g}_m)$ be as given in Theorem 3.4(iii). By dominated convergence, (3.11), and the conditions in Theorem 3.4(iii), we obtain

$$
\begin{aligned}
\mathbb{E}[\widetilde{f}(Y_{t \wedge \tau}) 1_{\{Y_{t \wedge \tau} \neq \dagger\}}] &= \lim_{m \to \infty} \mathbb{E}[\widetilde{f}_m(Y_{t \wedge \tau}) 1_{\{Y_{t \wedge \tau} \neq \dagger\}}] \\
&= \lim_{m \to \infty} \left(\widetilde{f}_m(\nu_0) + \int_0^t \mathbb{E}[\widetilde{g}_m(Y_s) 1_{\{s < \tau\}}] ds \right) \\
&\leq \widetilde{f}(\nu_0) + \int_0^t \mathbb{E}[\limsup_{m \to \infty} \widetilde{g}_m(Y_s)^+ 1_{\{s < \tau\}}] ds \\
&\leq c' \int_0^t \mathbb{E}[\widetilde{f}(Y_s) 1_{\{s < \tau\}}] ds \\
&\leq c' \int_0^t \mathbb{E}[\widetilde{f}(Y_{s \wedge \tau}) 1_{\{Y_{s \wedge \tau} \neq \dagger\}}] ds.
\end{aligned}
$$

By the Gronwall inequality (see Theorem 5.1 in the Appendixes of Ethier and Kurtz (2005)) we conclude that $\mathbb{E}[\widetilde{f}(Y_{t \wedge \tau}) 1_{\{Y_{t \wedge \tau} \neq \dagger\}}] = 0$. Together with the properties of $\widetilde{f}$ this implies that for $t < \tau$, $Y_t(\{\Delta\}) = 0$. Thus for $t < \tau$, Y_t takes values in $T(\mathcal{P}_w)$. Since the constant c was arbitrarily large, and since $\dagger$ is an absorbing state, it follows that Y takes values in $T(\mathcal{P}_w) \cup \{\dagger\}$, as desired.

Finally, suppose condition (iv) in Theorem 3.4 is in force, and consider the functions f_n given there. Let X be a possibly killed solution to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ with initial condition $\mu \in \mathcal{P}_w$. We then get

$$\mathbb{E}[1_{\{X_t \neq \dagger\}}] = \lim_{n \to \infty} \mathbb{E}[f_n(X_t)] = \lim_{n \to \infty} \mathbb{E}\left[f_n(\mu) + \int_0^t L^\dagger f_n(X_s) ds \right] \geq 1_{\{\mu \neq \dagger\}} = 1$$

for every fixed t. As a result, $X_t \in \mathcal{P}_w$ for all $t \geq 0$, as claimed. $\qquad \square$

4 Lévy type operators

Operators $L: \mathcal{D}_w \to C(\mathcal{P}_w)$ that satisfy the positive maximum principle are integro-differential operators of *Lévy type*, which we now introduce. Such operators involve derivatives of functions f of measure arguments, and we define

$$\partial_x f(\mu) = \lim_{\varepsilon \to 0} \frac{f(\mu + \varepsilon \delta_x) - f(\mu)}{\varepsilon}$$

for all $(x, \mu) \in \mathbb{R}^d \times M(\mathbb{R}^d)$ for which the limit is well-defined. We write $\partial f(\mu)$ for the map $x \mapsto \partial_x f(\mu)$. Iterated derivatives are written $\partial_{x_1, \ldots, x_k}^k f(\mu) = \partial_{x_1} \cdots \partial_{x_k} f(\mu)$ whenever they exist, and we write $\partial^k f(\mu)$ for the corresponding map. Define also the function space

$$C_w^\infty = \text{linear span of } w \text{ and } C_c^\infty(\mathbb{R}^d).$$

Example 4.1. If $f(\mu) = \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle}$ for some function $\varphi: \mathbb{R}^d \to \mathbb{R}$, then

$$\partial_x f(\mu) = (\varphi(x) - \langle \varphi, \mu \rangle w(x)) e^{-\langle w, \mu \rangle}$$

for any $x \in \mathbb{R}^d$ and any $\mu \in M(\mathbb{R}^d)$ such that φ and w are μ-integrable. One also has the product rule $\partial(fg) = f\partial g + g\partial f$. In particular, every test function in $f \in \mathcal{D}_w$ is infinitely many times differentiable and, for each k, the k-th derivative $\partial^k_{x_1,\dots,x_k} f(\mu)$ is jointly continuous in $(x_1, \dots, x_k, \mu) \in \mathbb{R}^k \times \mathcal{P}_w$, and the map $\partial^k f(\mu)$ lies in $(C_w^\infty)^{\otimes k}$ for every fixed μ. $\diamond$

We say that L is of *Lévy type* if it acts on test functions $f \in \mathcal{D}_w$ by

$$
\begin{aligned}
Lf(\mu) = -\kappa_\mu f(\mu) + \langle B_\mu(\partial f(\mu)), \mu \rangle + \frac{1}{2} \langle Q_\mu(\partial^2 f(\mu)), \mu^2 \rangle \\
+ \int_{\mathcal{P}_w} (f(\nu) - f(\mu) - \langle \partial f(\mu), \chi(\nu - \mu) \rangle) N(\mu, d\nu),
\end{aligned}
\tag{4.1}
$$

where, for each $\mu \in \mathcal{P}_w$, the following conditions are imposed to ensure that the right-hand side is well-defined:

- $\kappa_\mu \in \mathbb{R}_+$.

- B_μ is a linear operator from C_w^∞ to $L^1(E, \mu)$.

- Q_μ is a linear operator from $C_w^\infty \otimes C_w^\infty$ to $L^1(E \times E, \mu \otimes \mu)$ with $\langle Q_\mu(\varphi \otimes \varphi), \mu^2 \rangle \geq 0$ for all $\varphi \in C_w^\infty$.

- $N(\mu, d\nu)$ is a measure on $\mathcal{P}_w$ with $\int_{\mathcal{P}_w} 1 \wedge \langle \varphi, \mu - \nu \rangle^2 N(\mu, d\nu) < \infty$ for all $\varphi \in C_w^\infty$. In (4.1), $\chi(\nu - \mu) = (\nu - \mu)\rho(\langle w, \nu - \mu \rangle)$, with $\rho \colon \mathbb{R} \to [0,1]$ being a smooth function supported on $[-2, 2]$ and equal to one on $[-1, 1]$, acts as a truncation function for the large jumps.

If the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ has a possibly killed solution X for any initial conditions μ, then these objects govern the killing, drift, diffusion, and jump behavior of X, as one would expect.

5 Verifying the positive maximum principle

The positive maximum principle is convenient because it is often easy to verify in practice. The key tool for doing so when the operator is of Lévy type are the following optimality conditions for functions in $\mathcal{D}_w$.

Theorem 5.1. *Fix $f \in \mathcal{D}_w$ and $\overline{\mu} \in \mathcal{P}_w$ such that $f(\overline{\mu}) = \max_{\mathcal{P}_w} f$.*

(i) *$\langle \partial f(\overline{\mu}), \mu \rangle = \sup_E \partial f(\overline{\mu})$ for all $\mu \in \mathcal{P}_w$ such that $\operatorname{supp}(\mu) \subseteq \operatorname{supp}(\overline{\mu})$. In particular, $\partial_x f(\overline{\mu}) = \sup_E \partial f(\overline{\mu})$ for all $x \in \operatorname{supp}(\overline{\mu})$.*

(ii) *$\langle \partial^2 f(\overline{\mu}), \mu^2 \rangle \leq 0$ for all $\mu \in \mathcal{P}_w - \mathcal{P}_w$ such that $\langle 1, \mu \rangle = 0$ and $\operatorname{supp}(|\mu|) \subseteq \operatorname{supp}(\overline{\mu})$. In particular,*

$$
\partial^2_{xx} f(\overline{\mu}) + \partial^2_{yy} f(\overline{\mu}) - 2\partial^2_{xy} f(\overline{\mu}) = \langle \partial^2 f(\overline{\mu}), (\delta_x - \delta_y)^2 \rangle \leq 0, \quad x, y \in \operatorname{supp}(\overline{\mu}).
$$

(iii) *Let $\tau \colon \mathbb{R}^d \to \mathbb{R}^{d \times d}$ be C^1 and satisfy the condition*

$$
x \in E, \varphi \in C_c^\infty(\mathbb{R}^d), \varphi(x) = \max_E \varphi \implies \tau(x)^\top \nabla \varphi(x) = 0.
\tag{5.1}
$$

Define $A_\tau(\varphi) = \sum_{j=1}^d \tau_j^\top \nabla(\tau_j^\top \nabla \varphi)$ and $\aleph_\tau(\varphi \otimes \varphi) = \operatorname{Tr}((\tau^\top \nabla \varphi) \otimes (\tau^\top \nabla \varphi)^\top)$ for each $\varphi \in C_w^\infty$, where τ_j denotes the j-th column of τ. Assume the induced linear operators A_τ and $\aleph_\tau$ map C_w^∞ and $C_w^\infty \otimes C_w^\infty$ to $L^1(\mathbb{R}^d, \overline{\mu})$ and $L^1(\mathbb{R}^d \times \mathbb{R}^d, \overline{\mu} \otimes \overline{\mu})$, respectively. Then

$$
\langle A_\tau(\partial f(\overline{\mu})), \overline{\mu} \rangle + \langle \aleph_\tau(\partial^2 f(\overline{\mu})), \overline{\mu}^2 \rangle \leq 0.
\tag{5.2}
$$

If $E = \mathbb{R}^d$ one has the following improvement of (iii):

(iv) *Let $E = \mathbb{R}^d$. Then* (iii) *remains true if τ is assumed to be continuous but not C^1. Note that in this case* (5.1) *is vacuous and $A_\tau(\varphi) = \mathrm{Tr}(\tau\tau^\top \nabla^2 \varphi)$.*

Proof. Before we begin, we need a technical result. Fix $f \in \mathcal{D}_w$ and $\varphi_1, \ldots, \varphi_n \in C_c^\infty(\mathbb{R}^d)$ such that $f(\mu) = \Phi(\langle \varphi_1, \mu \rangle, \ldots, \langle \varphi_n, \mu \rangle, \langle w, \mu \rangle)$ for some $\Phi \in C^\infty(\mathbb{R}^{m+1})$. Applying the classical Taylor approximation theorem to Φ then yields

$$f(\mu + \nu_t) = f(\mu) + \langle \partial f(\mu), \nu_t \rangle + \frac{1}{2}\langle \partial^2 f(\mu), \nu_t^2 \rangle + o(t^2) \tag{5.3}$$

for all $\mu, \nu_t \in M(E)$ such that $w \in L^1(E, |\mu|) \cap L^1(E, |\nu_t|)$, $\langle w, \nu_t \rangle = O(t)$, and $\langle \varphi_i, \nu_t \rangle = O(t)$ for $i = 1, \ldots, n$.

(i) and (ii): The proof is a slight modification of the proof of Theorem 3.1 in Cuchiero et al. (2019). Pick any $x \in \mathrm{supp}(\overline{\mu})$, $y \in E$, and let A_n be the ball of radius $1/n$ centered at x, intersected with $\mathrm{supp}(\overline{\mu})$. Note that $\partial f(\overline{\mu}) \in C_w^\infty$ and that setting $\mu_n := \overline{\mu}(\,\cdot \cap A_n)/\overline{\mu}(A_n)$ we get that $\overline{\mu} + t(\delta_y - \mu_n) \in \mathcal{P}_w$ for all $t \in (0, \overline{\mu}(A_n))$ and that $(\mu_n)_{n \in \mathbb{N}}$ converges to δ_x in $\mathcal{P}_w$. Following the proof of Theorem 3.1 in Cuchiero et al. (2019) yields the result.

(iii): Assume that τ is compactly supported. We follow the idea of the proof of Proposition 4.1 in Abi Jaber et al. (2019). Fix $x \in \mathbb{R}^d$, $j \in \{1, \ldots, d\}$, and let $z_x : \mathbb{R}_+ \to \mathbb{R}^d$ be the solution of the ODE

$$z_x' = \tau_j(z_x) \qquad \text{and} \qquad z_x(0) = x.$$

By Proposition 2.5 in Da Prato and Frankowska (2004) we know that $z_x(t) \in E$ for all $x \in E$ and $t \geq 0$. Observe that for all $\varphi \in C_w^\infty$ and with $\psi = \varphi \otimes \varphi$ we have that

$$\lim_{t \to 0} \frac{\varphi(z_x(t)) - \varphi(x) - (z_x'(0)^\top \nabla \varphi(x))t}{t^2} = \frac{1}{2}\tau_j(x)^\top \nabla(\tau_j^\top \nabla \varphi)(x),$$

$$\lim_{t \to 0} \frac{\psi(x, y) - \psi(z_x(t), y) - \psi(x, z_y(t)) + \psi(z_x(t), z_y(t))}{t^2} = e_j^\top (\tau(x)^\top \nabla \varphi(x) \nabla \varphi(y)^\top \tau(y))e_j.$$

Define $\mu_t := F_*^t \overline{\mu}$, the pushforward of $\overline{\mu}$ under F^t, where $F^t(y) := z_y(t)$ for $y \in E$. Clearly $\mu_t \in M_1(E)$, and since τ_j is compactly supported, $\mu_t \in \mathcal{P}_w$. Moreover, the dominated convergence theorem yields $\lim_{t \to 0} \frac{1}{t}\langle \varphi, \mu_t - \overline{\mu} \rangle = \langle \tau_j^\top \nabla \varphi, \overline{\mu} \rangle$ for all $\varphi \in C_w^\infty$. We thus get that (5.3) holds true for $\mu := \overline{\mu}$ and $\nu_t := \mu_t - \overline{\mu}$. Since $\overline{\mu}$ maximizes f over $\mathcal{P}_w$, we then obtain

$$0 \geq f(\mu_t) - f(\overline{\mu})$$

$$= \langle \partial f(\overline{\mu}), \mu_t - \overline{\mu} \rangle + \frac{1}{2}\langle \partial^2 f(\overline{\mu}), (\mu_t - \overline{\mu})^2 \rangle + o(t^2)$$

$$= \langle \partial_{F^t} f(\overline{\mu}) - \partial f(\overline{\mu}), \overline{\mu} \rangle + \frac{1}{2}\langle \partial^2 f(\overline{\mu}) - 2\partial^2_{\cdot, F^t} f(\overline{\mu}) + \partial^2_{F^t, F^t} f(\overline{\mu}), \overline{\mu}^2 \rangle + o(t^2).$$

Recall that by Theorem 5.1(i) we know that $\partial_y f(\overline{\mu}) = \sup_E \partial f(\overline{\mu})$ for all $y \in \mathrm{supp}(\overline{\mu})$ and hence $\tau(y)^\top \nabla(\partial p(\overline{\mu}))(y) = 0$ by (5.1). As a result, dividing the above expression by t^2, letting t go to 0, applying the dominated convergence theorem, and summing over $1 \leq j \leq d$ yields (5.2). A further application of dominated convergence allows to remove the assumption that τ is compactly supported.

(iv): The proof follows the proof of (iii) using $F^t(x) := x + t\tau_j(x)$ for all $x \in \mathbb{R}^d$. $\qquad \square$

The property (5.1) in Condition (iii) can be understood as requiring that τ is a possible diffusion matrix for a (possibly killed) E-valued diffusion process. Under slightly more regularity on τ, it is a known result in the stochastic invariance literature that this holds

if and only if (5.1) is satisfied; see for instance Theorem 4.1 in Da Prato and Frankowska (2004). The extension to the current generality, where τ is merely C^1, can be proven using Theorem 2.4 and proceeding as in Abi Jaber et al. (2019). We do not elaborate on the details, as we have no need for this result in the present paper.

Condition (iii) is in fact an application of a more general condition, which we report here for the case $w \equiv 1$. See Theorem 3.4 and Remark 5.7(i) in Cuchiero et al. (2019) for more details and a proof. An analogous result can be proved for w as in (3.1).

Lemma 5.2. *Fix $w \equiv 1$, $f \in \mathcal{D}_w$, and $\overline{\mu} \in \mathcal{P}_w$ such that $f(\overline{\mu}) = \max_{\mathcal{P}_w} f$. Let A be the generator of a strongly continuous group of positive isometries of $\mathbb{R} + C_0(E)$, and assume the domain of A and the domain of A^2 both contain C_w^∞. Then*

$$\langle A^2(\partial f(\overline{\mu})), \overline{\mu} \rangle + \langle (A \otimes A)(\partial^2 f(\overline{\mu})), \overline{\mu}^2 \rangle \leq 0.$$

6 Verifying the technical conditions

We now turn to the technical conditions (ii)–(iv) in Theorem 3.4. At the end of the section, we follow up on Remark 3.5 and consider the case $w \equiv 1$. Recall the embedding $T \colon \mathcal{P}_w \to M_+(E^\Delta)$ defined in (3.4).

We now start to work towards concrete ways of checking these assumptions.

Lemma 6.1. *Suppose L is of Lévy type (4.1). Assume that the functions*

$$\mu \mapsto \kappa_\mu e^{-\langle w, \mu \rangle}, \qquad \mu \mapsto \langle B_\mu(\varphi), \mu \rangle e^{-\langle w, \mu \rangle}, \qquad \mu \mapsto \langle Q_\mu(\varphi \otimes \varphi), \mu^2 \rangle e^{-\langle w, \mu \rangle},$$

$$\mu \mapsto e^{-\langle w, \mu \rangle} \int_{\mathcal{P}_w} \left(\langle \varphi, \nu \rangle e^{-\langle w, \nu - \mu \rangle} - \langle \varphi, \mu \rangle - \langle \varphi - w \langle \varphi, \mu \rangle, \chi(\nu - \mu) \rangle \right) N(\mu, d\nu),$$

and

$$\mu \mapsto e^{-k\langle w, \mu \rangle} \int_{\mathcal{P}_w} \left(\langle \varphi, \nu \rangle e^{-\langle w, \nu - \mu \rangle} - \langle \varphi, \mu \rangle \right)^k N(\mu, d\nu), \qquad k \geq 2,$$

are of C_0 type for every $\varphi \in C_w^\infty$. Then so is Lf for every $f \in \mathcal{D}_w$, that is, condition (ii) in Theorem 3.4 is satisfied.

Proof. Note that one can check that each term of Lf in (4.1) is of C_0 type separately. Recall that $\mu \mapsto e^{-\langle w, \mu \rangle}$ is of C_0 type, as are all the elements of $\mathcal{D}_w$.

Killing: The result for $\mu \mapsto \kappa_\mu f(\mu)$ follows by noting that $f(\mu) = \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle} g(\mu)$ for some g constant or in $\mathcal{D}_w$, for some $\varphi \in C_c^\infty(\mathbb{R}^d)$, and for all $\mu \in \mathcal{P}_w$.

Drift: Observe that for

$$f(\mu) = \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle} \text{ with } \varphi \in C_c^\infty(\mathbb{R}^d) \tag{6.1}$$

we have $\partial f(\mu) = (\varphi - \langle \varphi, \mu \rangle w) e^{-\langle w, \mu \rangle}$. One then sees that the given conditions ensure that $\mu \mapsto \langle B_\mu(\partial f(\mu)), \mu \rangle$ is of C_0 type. By the product rule and linearity in f, this result extends to each $f \in \mathcal{D}_w$.

Diffusion: For f as in (6.1) we have $\partial^2 f(\mu) = (\langle \varphi, \mu \rangle w \otimes w - 2w \otimes \varphi) e^{-\langle w, \mu \rangle}$. The polarization identity $w \otimes \varphi = \frac{1}{4}((w + \varphi) \otimes (w + \varphi) - (w - \varphi) \otimes (w - \varphi))$ and the given conditions yield that $\mu \mapsto \langle Q_\mu(\partial^2 f(\mu)), \mu \rangle$ and $\mu \mapsto \langle Q_\mu(\partial f(\mu) \otimes \partial f(\mu)), \mu \rangle$ are of C_0 type. Applying the product rule twice we get $\partial^2(fg) = f \partial^2 g + 2 \partial g \otimes \partial f + g \partial^2 f$ for each $f, g \in \mathcal{D}_w$. Combined with linearity in f this yields that $\mu \mapsto \langle Q_\mu(\partial^2 f(\mu)), \mu \rangle$ is of C_0 type for each $f \in \mathcal{D}_w$.

Jumps: Finally, consider $g(\mu) := p(f(\mu))$ for some polynomial $p \colon \mathbb{R} \to \mathbb{R}$ and some f as in (6.1). Since p is a polynomial and $\langle \partial g(\mu), \chi(\nu - \mu) \rangle = p'(f(\mu)) \langle \partial f(\mu), \chi(\nu - \mu) \rangle$, an

application of the classical Taylor approximation theorem to p yields

$$g(\nu) - g(\mu) - \langle \partial g(\mu), \chi(\nu - \mu) \rangle$$
$$= p'(f(\mu))\Big(f(\nu) - f(\mu) - \langle \partial f(\mu), \chi(\nu - \mu)\rangle\Big) + \sum_{\ell=2}^{k} \frac{p^{(\ell)}(f(\mu))}{\ell!}(f(\nu) - f(\mu))^{\ell},$$

where k denotes the degree of p. Thus the given conditions ensure that the last term of Lg in (4.1) is of C_0 type. By polarization, this is also true for all $f \in \mathcal{D}_w$. $\qquad\square$

We focus now on condition (iii) in Theorem 3.4, assuming that (i)–(ii) are satisfied. This ensures that the operator $\widetilde{L} \colon \mathcal{D} \to C_0(\mathcal{X})$ in (3.7)–(3.8) is well-defined.

To verify these conditions it is useful to first extend L to a larger class of functions than $\mathcal{D}_w$, and then search for appropriate sequences $\{f_m\}_{m\in\mathbb{N}}$ in this larger class. The precise notion of extension is given in the following definition.

For any compact subset $\mathcal{K} \subset \mathcal{X}$ we define the following restricted graph of $\widetilde{L}$:

$$\mathrm{gph}_{\mathcal{K}}(\widetilde{L}) = \{(\widetilde{f}, \widetilde{L}\widetilde{f}|_{\mathcal{K}}) \colon \widetilde{f} \in \mathcal{D}\} \subset C_0(\mathcal{X}) \times C(\mathcal{K}).$$

Definition 6.2. *We say that $\widetilde{L}$ can be extended to a function $\widetilde{f} \colon \mathcal{X} \to \mathbb{R}$ if there is another function $\widetilde{g} \colon \mathcal{X} \to \mathbb{R}$ such that $(\widetilde{f}, \widetilde{g}|_{\mathcal{K}})$ lies in the bp-closure of $\mathrm{gph}_{\mathcal{K}}(\widetilde{L})$ for every compact subset $\mathcal{K} \subset \mathcal{X}$. We say that L can be extended to a function $f \colon \mathcal{P}_w \to \mathbb{R}$ if $\widetilde{L}$ can be extended to a function $\widetilde{f} \colon \mathcal{X} \to \mathbb{R}$ that satisfies $f = \widetilde{f} \circ T$.*

If $f = \widetilde{f} \circ T$, and if $(\widetilde{f}, \widetilde{g}|_{\mathcal{K}})$ lies in the bp-closure of $\mathrm{gph}_{\mathcal{K}}(\widetilde{L})$ for every compact subset $\mathcal{K} \subset \mathcal{X}$, we write Lf for $\widetilde{g} \circ T$. Sometimes, it happens that the expression given in (4.1) is well defined for f and coincides with $\widetilde{g} \circ T$. Those cases will be particularly important for our purposes.

Lemma 6.3. *Suppose L is of Lévy type (4.1) and satisfies conditions (i)–(ii) of Theorem 3.4. Assume L can be extended to all functions f in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w,\mu\rangle}$, and that Lf is given by (4.1). Assume also there exist $[0,1]$-valued functions $\psi_m \in C_c^\infty(\mathbb{R}^d)$ with the following properties:*

- *$\psi_m \to 1$ pointwise.*

- *The functions $h_m \colon \mathcal{P}_w \to \mathbb{R}$ given by $h_m(\mu) = \langle B_\mu((1-\psi_m)w), \mu\rangle^+$, which are of C_0 type, satisfy $\limsup_{m\to\infty} h_m \circ T^{-1} \le c'\langle 1_{\{\Delta\}}, \cdot \rangle$ in the bounded pointwise sense on every compact subset of $\mathcal{X}$, for some constant c'.*

Then condition (iii) in Theorem 3.4 is satisfied.

Proof. We claim that L can be extended to all maps $f \colon \mathcal{P}_w \to \mathbb{R}$ of the form

$$f(\mu) := p\big(\langle \varphi_1, \mu\rangle e^{-\langle w,\mu\rangle}, \dots, \langle \varphi_n, \mu\rangle e^{-\langle w,\mu\rangle}\big) \tag{6.2}$$

for some $p \in C^2(\mathbb{R}^n)$ with $p(0) = 0$ and $\varphi_1, \dots, \varphi_n \in \mathbb{R} + C_c^\infty(\mathbb{R}^d)$, and that Lf is given by (4.1). To see this, define functions

$$f_m(\mu) := p_m\big(\langle \varphi_1, \mu\rangle e^{-\langle w,\mu\rangle}, \dots, \langle \varphi_n, \mu\rangle e^{-\langle w,\mu\rangle}\big)$$

for some polynomials p_m on $\mathbb{R}^n$ with $p_m(0) = 0$ such that $p_m \to p$, $\nabla p_m \to \nabla p$, and $\nabla^2 p_m \to \nabla^2 p$ uniformly on $[-R, R]^n$, where $R = \max_{i=1,\dots,n} \sup_{\mu\in\mathcal{P}_w} |\langle \varphi_i, \mu\rangle| e^{-\langle w,\mu\rangle}$. Then $f_m \to f$ and $Lf_m \to Lf$ uniformly on $\mathcal{P}_w$, where Lf is defined by (4.1). Therefore the functions $\widetilde{f}_m = f_m \circ T^{-1}$ and $\widetilde{L}\widetilde{f}_m = (Lf_m) \circ T^{-1}$ converge in the bounded pointwise sense (even uniformly) to $\widetilde{f} = f \circ T^{-1}$ and $Lf \circ T^{-1}$. This shows that L can be extended to f, and that Lf is given by (4.1).

Fix any $c \geq 1$. Let $p_c \in C^2(\mathbb{R}_+)$ be such that $1_{\{x \leq c\}} \leq p_c(x) \leq 1_{\{x \leq 2c\}}$, and set

$$f_m(\mu) := p_c(\langle w, \mu \rangle) \langle (1 - \psi_m) w, \mu \rangle.$$

The function f_m is of the form (6.2). Indeed, we can write f_m as

$$f_m(\mu) = p_1\big(\langle 1, \mu \rangle e^{-\langle w, \mu \rangle}\big) - p_2\big(\langle 1, \mu \rangle e^{-\langle w, \mu \rangle}\big) \langle \psi_m w, \mu \rangle e^{-\langle w, \mu \rangle},$$

where $p_1(x) = p_c(-\log(x))(-\log(x))$ and $p_2(x) = p_c(-\log(x))/x$ for $x > 0$, and $p_1(x) = p_2(x) = 0$ for $x \leq 0$.

We now define $\widetilde{f}_m = f_m \circ T^{-1}$, $\widetilde{g}_m = Lf_m \circ T^{-1}|_{\mathcal{X}_c}$, and $\widetilde{f}(\nu) = p_c(\langle 1, \nu \rangle)\nu(\{\Delta\})$, and prove that these functions satisfy the properties in Theorem 3.4(iii). Since L can be extended to f_m, the pair $(\widetilde{f}_m, \widetilde{g}_m)$ lies in the bp-closure of $\mathrm{gph}_{\mathcal{X}_c}(\widetilde{L})$. Moreover, the $\widetilde{f}_m$ are uniformly bounded in m because $0 \leq \widetilde{f}_m(\nu) = p_c(\langle 1, \nu \rangle)\langle (1 - \psi_m), \nu \rangle \leq 2c$. Next, for all $\mu \in \mathcal{P}_w$ such that $\langle w, \mu \rangle \leq c$, we have $f_m(\mu) = \langle (1 - \psi_m) w, \mu \rangle$, $\partial f_m(\mu) = (1 - \psi_m) w$, and $\partial^2 f_m(\mu) = 0$. Since also κ_μ is nonnegative we have

$$Lf_m(\mu) \leq \langle B_\mu((1 - \psi_m)w), \mu \rangle + \int_{\mathcal{P}_w} \big(p_c(\langle w, \nu \rangle) - 1\big)\langle (1 - \psi_m)w, \nu \rangle N(\mu, d\nu)$$

$$\leq \langle B_\mu((1 - \psi_m)w), \mu \rangle.$$

Therefore $\widetilde{g}_m^+ \leq h_m \circ T^{-1}|_{\mathcal{X}_c}$, and our hypotheses imply that $\widetilde{g}_m^+$ is uniformly bounded in m, and that $\limsup_{m \to \infty} \widetilde{g}_m^+ \leq c'\widetilde{f}|_{\mathcal{X}_c}$ pointwise. The dominated convergence theorem implies that $\widetilde{f}_m \to \widetilde{f}$ pointwise. Finally, it follows by inspection that $\widetilde{f} \geq 0$, and that $\widetilde{f}(\nu) = 0$ for $\nu \in \mathcal{X}_c$ if and only if $\nu(\{\Delta\}) = 0$. $\qquad\square$

The last condition to analyze is condition (iv).

Lemma 6.4. *Suppose L is of Lévy type (4.1) and satisfies the conditions* (i)–(ii) *of Theorem 3.4. Assume that for every function f in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w, \mu \rangle}$, (f, Lf) with Lf is given by (4.1) lies in the bp-closure of the graph $\{(h, Lh): h \in \mathcal{D}_w\}$. Assume also that $\kappa_\mu = 0$ for all $\mu \in \mathcal{P}_w$, and that we have the linear growth condition*

$$\langle w \otimes B_\mu(w), \mu^2 \rangle^+ + \langle Q_\mu(w \otimes w), \mu^2 \rangle + \int_{\mathcal{P}_w} \langle w, \nu - \mu \rangle^2 \wedge \langle w, \mu \rangle^2 N(\mu, d\nu) \leq c\langle w, \mu \rangle^2$$

for all $\mu \in \mathcal{P}_w$ and some constant c. Then condition (iv) *in Theorem 3.4 is satisfied.*

Proof. We showed in the proof of Lemma 6.3 that L can be extended to all maps $f \colon \mathcal{P}_w \to \mathbb{R}$ of the form (6.2), and that Lf is given by (4.1). In fact, under our current assumptions, the argument shows that (f, Lf) lies in the bp-closure (even the uniform closure) of the graph $\{(h, Lh): h \in \mathcal{D}_w\}$. In particular, these facts apply to the maps $f_n(\mu) = q(\langle w, \mu \rangle/n)$, where $q \in C_c^\infty(\mathbb{R}_+)$ is nonincreasing and satisfies $1_{[0,1]} \leq q \leq 1_{[0,2]}$. It is clear that $f_n \to 1$ in the bounded pointwise sense. We must argue that $(Lf_n)^- \to 0$ in the same sense. A direct computation gives

$$Lf_n(\mu) = q'(\langle w, \mu \rangle/n)\frac{1}{n}\langle B_\mu(w), \mu \rangle + \frac{1}{2}q''(\langle w, \mu \rangle/n)\frac{1}{n^2}\langle Q_\mu(w \otimes w), \mu^2 \rangle$$

$$+ \int_{\mathcal{P}_w} \Big(q(\langle w, \nu \rangle/n) - q(\langle w, \mu \rangle/n) - q'(\langle w, \mu \rangle/n)\frac{1}{n}\langle w, \chi(\nu - \mu) \rangle\Big) N(\mu, d\nu).$$

Using the properties of q and, in the last step, the assumed linear growth condition, we get for some constant c' the lower bound,

$$
\begin{aligned}
&- c'\Big(\frac{1}{n}\langle B_\mu(w),\mu\rangle^+ + \frac{1}{n^2}\langle Q_\mu(w\otimes w),\mu^2\rangle + \int_{\mathcal{P}_w}\frac{\langle w,\nu-\mu\rangle^2}{n^2}\wedge 1\, N(\mu,d\nu)\Big)1_{\{\langle w,\mu\rangle\leq 2n\}}\\
&\geq -\frac{4c'}{\langle w,\mu\rangle^2}\Big(\langle w\otimes B_\mu(w),\mu^2\rangle^+ + \langle Q_\mu(w\otimes w),\mu^2\rangle + \int_{\mathcal{P}_w}\langle w,\nu-\mu\rangle^2\wedge\langle w,\mu\rangle^2 N(\mu,d\nu)\Big)\\
&\geq -4c'c.
\end{aligned}
$$

We deduce that $(Lf_n)^-$ is uniformly bounded in n, and it is clear from the first line that $(Lf_n)^- \to 0$ pointwise. $\hfill\square$

Remark 6.5. We now comment on the case $w\equiv 1$, and consider the setting of Remark 3.5. Lemma 6.1 would then be replaced by the requirement that Lf can be extended to a function in $C(M_1(E^\Delta))$ for all $f\in\mathcal{D}_w$. Concrete conditions when L is of Lévy type are that the maps $\mu\mapsto\kappa_\mu$, $\mu\mapsto B_\mu(\varphi)$, and $\mu\mapsto Q_\mu(\varphi\otimes\varphi)$ are continuous from $\mathcal{P}_w$ to $\mathbb{R}$, $\mathbb{R}+C_0(\mathbb{R}^d)$, and $\mathbb{R}+C_0(\mathbb{R}^d)\otimes C_0(\mathbb{R}^d)$, respectively, and that $\mu\mapsto\int\langle\varphi,\nu-\mu\rangle^\ell N(\mu,d\nu)$ can be extended to a function in $C(M_1(E^\Delta))$ for every $\ell\geq 2$ and $\varphi\in C_w^\infty$. Next, Lemma 6.3 holds without the assumption that L can be extended to all functions in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w,\mu\rangle}$. In the proof, one simply takes $f_m(\mu)=\langle(1-\psi_m),\mu\rangle$.

In the case of condition (iv), the result for $w\equiv 1$ is quite different from Lemma 6.4. The reason is that, in contrast to the cases where condition (3.1) is in force, for $w\equiv 1$ the cemetery state $\dagger$ is an isolated point. Thus a solution to the martingale problem can reach $\dagger$ only by means of a jump. This has the consequence that (iv) can essentially be derived from (iii). We now report a precise formulation of this statement.

Lemma 6.6. *Suppose that $w\equiv 1$, L is of Lévy type (4.1) and satisfies conditions* (i)–(iii) *of Theorem 3.4. Then, if $\kappa_\mu=0$ for all $\mu\in\mathcal{P}_w$, condition* (iv) *of Theorem 3.4 is satisfied.*

Proof. Choose $(\widetilde{f}_m,\widetilde{g}_m)_{m\in\mathbb{N}}$ as in (iii) for $c=1$. Define $f_m=1-\widetilde{f}_m\circ T$ and $g_m=-\widetilde{g}_m\circ T$. Since $\kappa_\mu=0$, (f_m,g_m) lies in the bp-closure of the graph of L. Moreover, $(f_m,g_m^-)\to(1,0)$ in the bounded pointwise sense. $\hfill\square$

7 Applications of the main result

Take $E=\mathbb{R}^d$ and assume that $w(x)=|x|^p$ for $|x|>2$, where $p\in(0,\infty)$. Consider a linear operator $L\colon\mathcal{D}_w\to C(\mathcal{P}_w)$ of Lévy type (4.1) with $\kappa=0$, $N=0$, and B and Q given by

$$
B_\mu(\varphi)=b_\mu^\top\nabla\varphi+\frac{1}{2}\operatorname{Tr}((\sigma_\mu^2+\tau_\mu^2)\nabla^2\varphi),
$$
$$
Q_\mu(\varphi\otimes\varphi)=(\tau_\mu\nabla\varphi)\otimes(\tau_\mu\nabla\varphi),
$$

for some maps $b\colon\mathcal{P}_w\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma,\tau\colon\mathcal{P}_w\times\mathbb{R}^d\to\mathbb{S}^d$. Here we use the notation $u\otimes v=u_1\otimes v_1+\cdots+u_d\otimes v_d$ whenever $u,v\colon\mathbb{R}^d\to\mathbb{R}^d$.

Theorem 7.1. *Assume that $b_\mu(x)=\widetilde{b}_{w\mu}(x)$, $\sigma_\mu(x)=\widetilde{\sigma}_{w\mu}(x)$, $\tau_\mu(x)=\widetilde{\tau}_{w\mu}(x)$ for some continuous maps $\widetilde{b}\colon\mathcal{X}\times\mathbb{R}^d\to\mathbb{R}^d$ and $\widetilde{\sigma},\widetilde{\tau}\colon\mathcal{X}\times\mathbb{R}^d\to\mathbb{S}^d$ such that*

$$
\frac{\widetilde{b}_{\nu,i}(x)}{1+|x|},\quad\frac{\widetilde{\sigma}_{\nu,ij}(x)^2}{1+|x|^2},\quad\text{and}\quad\frac{\widetilde{\tau}_{\nu,ij}(x)}{1+|x|},\qquad i,j\in\{1,\ldots,d\},\tag{7.1}
$$

are continuous as functions from $\mathcal{X}$ to $\mathbb{R}+C_0(\mathbb{R}^d)$. Assume also that

$$
\sup_{x\in\mathbb{R}^d}\frac{|x|\,|\widetilde{b}_\nu(x)|+|\widetilde{\sigma}_\nu(x)|^2+|\widetilde{\tau}_\nu(x)|^2}{1+|x|^2}\leq c\langle 1,\nu\rangle^\gamma,\quad\nu\in\mathcal{X},\tag{7.2}
$$

for some constants $c, \gamma \geq 0$. Then conditions (i)–(iii) of Theorem 3.4 are satisfied, and thus there exists a possibly killed solution X to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for every initial condition $\mu \in \mathcal{P}_w$. If one can take $\gamma = 0$ in (7.2), then condition (iv) of Theorem 3.4 holds, and thus $X_t \in \mathcal{P}_w$ for all $t \geq 0$.

Proof. For notational simplicity we only prove the case $d = 1$. For all $\varphi \in C_c^\infty(\mathbb{R})$ set $f_\varphi(\mu) = \langle \varphi, \mu \rangle e^{-\langle w, \mu \rangle}$ and recall that $\partial f_\varphi(\mu) = (\varphi - \langle \varphi, \mu \rangle w) e^{-\langle w, \mu \rangle}$ and $\partial^2 f_\varphi(\mu) = (\langle \varphi, \mu \rangle w \otimes w - 2w \otimes \varphi) e^{-\langle w, \mu \rangle}$. Define also $\widetilde{f}_\varphi = f_\varphi \circ T^{-1}$, and set $\widetilde{B}_\nu(\varphi) = \widetilde{b}_\nu \varphi' + \frac{1}{2}(\widetilde{\sigma}_\nu^2 + \widetilde{\tau}_\nu^2)\varphi''$ and $\widetilde{\Sigma}_\nu(\varphi) = \widetilde{\tau}_\nu \varphi'$ for all $\nu \in \mathcal{X}$ and $\varphi \in C_w^\infty$.

Next, observe that φ, φ', and φ'' are continuous and $|\varphi'(x)x|/w(x)$ and $|\varphi''(x)x^2|/w(x)$ are bounded for all $\varphi \in C_w^\infty$. Condition (7.1) then yields that for all such φ,

$$\frac{\widetilde{b}_\nu(x)\varphi'(x)}{w(x)}, \quad \frac{(\widetilde{\sigma}_\nu(x)^2 + \tau_\nu(x)^2)\varphi''(x)}{w(x)}, \quad \text{and} \quad \frac{\widetilde{\tau}_\nu(x)\varphi'(x)}{w(x)} \tag{7.3}$$

are continuous as functions from $\mathcal{X}$ to $\mathbb{R} + C_0(\mathbb{R})$, and condition (7.2) yields that

$$\sup_{x \in \mathbb{R}} \frac{|\widetilde{b}_\nu(x)\varphi'(x)| + (\widetilde{\sigma}_\nu(x)^2 + \widetilde{\tau}_\nu(x)^2)|\varphi''(x)| + |\widetilde{\tau}_\nu(x)\varphi'(x)|}{w(x)} \leq c_\varphi \langle 1, \nu \rangle^\gamma, \quad \nu \in \mathcal{X}, \tag{7.4}$$

for some constant c_φ depending on φ. We are now ready to verify the conditions of Theorem 3.4.

(i): The positive maximum principle follows directly form Theorem 5.1(i) and (iv).

(ii): We verify the conditions of Lemma 6.1. That is, in the current notation, we must check that the functions

$$\nu \mapsto \langle \widetilde{B}_\nu(\varphi)w^{-1}, \nu \rangle e^{-\langle 1, \nu \rangle} \quad \text{and} \quad \nu \mapsto \langle \widetilde{\Sigma}_\nu(\varphi)w^{-1}, \nu \rangle^2 e^{-\langle 1, \nu \rangle} \tag{7.5}$$

are C_0 functions on $\mathcal{X}$ for every $\varphi \in C_w^\infty$. To see that the function involving $\widetilde{B}_\nu$ is continuous, we write

$$\left| \langle \widetilde{B}_\nu(\varphi)w^{-1}, \nu \rangle - \langle \widetilde{B}_{\nu_n}(\varphi)w^{-1}, \nu_n \rangle \right| \leq \left| \langle \widetilde{B}_\nu(\varphi)w^{-1}, \nu - \nu_n \rangle \right|$$
$$+ \sup_{x \in \mathbb{R}} \left| (\widetilde{B}_\nu(\varphi)(x) - \widetilde{B}_{\nu_n}(\varphi)(x))w^{-1}(x) \right| \langle 1, \nu_n \rangle.$$

If $\nu_n \Rightarrow \nu$, then the first term tends to zero since $\widetilde{B}_\nu(\varphi)w^{-1} \in \mathbb{R} + C_0(\mathbb{R})$. The second term tends to zero due to (7.3). A similar calculation shows that the function in (7.5) involving $\widetilde{\Sigma}_\nu$ is continuous as well.

It remains to show that the functions in (7.5) vanish at infinity. To see this, note that

$$|\langle \widetilde{B}_\nu(\varphi)w^{-1}, \nu \rangle| \leq \sup_{x \in \mathbb{R}} \frac{|\widetilde{b}_\nu(x)\varphi'(x)| + \frac{1}{2}(\widetilde{\sigma}_\nu(x)^2 + \widetilde{\tau}_\nu(x)^2)|\varphi''(x)|}{w(x)} \langle 1, \nu \rangle,$$
$$|\langle \widetilde{Q}_\nu(\varphi \otimes \varphi)w^{-2}, \nu^2 \rangle| \leq \sup_{x \in \mathbb{R}} \frac{|\widetilde{\tau}_\nu(x)\varphi'(x)|^2}{w(x)^2} \langle 1, \nu \rangle^2 \tag{7.6}$$

for all $\varphi \in C_w^\infty$. Since the suprema in (7.6) grow at most polynomially in $\langle 1, \nu \rangle$ due to (7.4), the functions in (7.5) vanish at infinity.

(iii): We verify the conditions of Lemma 6.3. We have already shown that L satisfies conditions (i)–(ii) of Theorem 3.4. Let us show that L can be extended to all functions f in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w, \mu \rangle}$, and that Lf is given by (4.1).

Fix $\psi_n(x) := \psi(x/n)$ for some $\psi \in C_c^\infty(\mathbb{R})$ such that $1_{[-1,1]} \leq \psi \leq 1_{[-2,2]}$ and set $\widetilde{f}_n := p(\widetilde{f}_{\varphi_1}, \ldots, \widetilde{f}_{\varphi_k}, \widetilde{f}_{\psi_n})$ and $\widetilde{f} := p(\widetilde{f}_{\varphi_1}, \ldots, \widetilde{f}_{\varphi_k}, e^{-\langle 1, \cdot \rangle})$ for an arbitrary polynomial

$p: \mathbb{R}^{k+1} \to \mathbb{R}$ with $p(0) = 0$ and some $\varphi_1, \ldots, \varphi_k \in C_c^\infty(\mathbb{R})$. Note that $\widetilde{f}_n$ converges to $\widetilde{f}$ on $\mathcal{X}$ bounded pointwise. Next, observe that condition (7.2) yields

$$
\begin{aligned}
|\langle \widetilde{B}_\nu(\psi_n) w^{-1}, \nu \rangle| &\leq 2 \sup_{x \in \mathbb{R}} \left| \frac{\widetilde{b}_\nu(x)}{n} 1_{[-2n,2n]}(x) + \frac{\widetilde{\sigma}_\nu(x)^2 + \widetilde{\tau}_\nu(x)^2}{2n^2} 1_{[-2n,2n]}(x) \right| \langle w^{-1}, \nu \rangle \\
&\leq c'' \sup_{x \in \mathbb{R}} \frac{|x| |\widetilde{b}_\nu(x)| + \widetilde{\sigma}_\nu(x)^2 + \widetilde{\tau}_\nu(x)^2}{1 + x^2} \\
&\leq c'' c \langle 1, \nu \rangle^\gamma,
\end{aligned}
\tag{7.7}
$$

for some constant c''. Similarly, one can bound

$$
|\langle \widetilde{Q}_\nu(\psi_n \otimes \psi_n)/w^2, \nu^2 \rangle| \quad \text{and} \quad |\langle \widetilde{Q}_\nu(\varphi \otimes \psi_n)/w^2, \nu^2 \rangle|
$$

by $c''' c \langle 1, \nu \rangle^\gamma$, for some constant c''' and all $\varphi \in \{\varphi_1, \ldots, \varphi_k, w\}$. Since both bounds grow at most polynomially in $\langle 1, \nu \rangle$, the product rule and the polarization identity yield that $(\widetilde{L} \widetilde{f}_n)_{n \in \mathbb{N}}$ is a bounded sequence in $C_0(\mathcal{X})$. Moreover, the dominated convergence theorem shows that $\widetilde{L} \widetilde{f}_n(\nu) \to \widetilde{g}(\nu)$ for all $\nu \in \mathcal{X}$, where $\widetilde{g} \circ T(\mu) = Lf(\mu)$ as given by (4.1) for $f = \widetilde{f} \circ T = p(f_{\varphi_1}, \ldots, f_{\varphi_k}, e^{-\langle w, \cdot \rangle})$. This proves that $(\widetilde{f}, \widetilde{g})$ lies in the bp-closure of the graph $\{(\widetilde{f}, \widetilde{L}\widetilde{f}): \widetilde{f} \in \mathcal{D}\}$, and thus that $\widetilde{L}$ can be extended to $\widetilde{f}$. By definition, this implies that L can be extended to f with Lf given by (4.1).

Let now h_m be as in Lemma 6.3 and note that

$$
h_m \circ T^{-1}(\nu) = \langle \widetilde{B}_\nu((1 - \psi_m)w)w^{-1}, \nu \rangle^+.
$$

The first inequality in (7.6), condition (7.4), and the reasoning in (7.7) imply that $(h_m|_{\mathcal{K}})_{m \in \mathbb{N}}$ is a bounded sequence in $C(\mathcal{K})$ for every compact set $\mathcal{K}$. Moreover,

$$
\lim_{m \to \infty} \frac{\widetilde{B}_\nu((1 - \psi_m)w)}{w} = \left(\widetilde{b}_\nu(\Delta) \frac{w'(\Delta)}{w(\Delta)} + \frac{1}{2} (\widetilde{\sigma}_\nu^2(\Delta) + \widetilde{\tau}_\nu^2(\Delta)) \frac{w''(\Delta)}{w(\Delta)} \right) 1_{\{\Delta\}},
$$

which is well defined by (7.3). Write the right-hand side as $c' 1_{\{\Delta\}}$ for a constant c'. By the dominate convergence theorem we can conclude that $h_m \circ T^{-1} \to c' \langle 1_{\{\Delta\}}, \cdot \rangle$ in the bounded pointwise sense on every compact subset of $\mathcal{X}$. Thus the conditions of Lemma 6.3 hold, as required.

(iv): We verify the conditions of Lemma 6.4 under the additional assumption that one can take $\gamma = 0$ in (7.2). We already know that L satisfies the conditions (i)–(ii) of Theorem 3.4 and that $\kappa_\mu = 0$ for all $\mu \in \mathcal{P}_w$. We show now that for every function f in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w, \mu \rangle}$, the pair (f, Lf) with Lf given by (4.1) lies in the bp-closure of the graph $\{(h, Lh): h \in \mathcal{D}_w\}$. To do this, we just need to follow the first part of the proof of (iii). Indeed, setting $f_n = \widetilde{f}_n \circ T$ we get that $(f_n, Lf_n) \in \{(h, Lh): h \in \mathcal{D}_w\}$ converges bounded pointwise to (f, Lf) for Lf as given in (4.1).

The last condition of Lemma 6.4 to be verified is the linear growth. By (7.4) with $\varphi = w$ and $\gamma = 0$ we compute

$$
\begin{aligned}
\langle w \otimes B_\mu(w), \mu^2 \rangle &= \langle w, \mu \rangle \langle b_\mu w' + \frac{1}{2}(\sigma_\mu^2 + \tau_\mu^2) w'', \mu \rangle \leq c_w \langle w, \mu \rangle^2 \\
\langle Q_\mu(w \otimes w), \mu^2 \rangle &= \langle \tau_\mu w', \mu \rangle^2 \leq c_w \langle w, \mu \rangle^2.
\end{aligned}
$$

The claim follows. $\qquad\square$

Remark 7.2. As will be explored further in Section 8, the linear operator L introduced at the beginning of the section coincides with the generator of the conditional distribution $X_t = \mathbb{P}(Z_t \in \cdot \mid \mathcal{F}_t^0)$ of a solution of a McKean–Vlasov equation with common noise,

$$
dZ_t = b_{X_t}(Z_t)dt + \sigma_{X_t}(Z_t)dW_t + \tau_{X_t}(Z_t)dW_t^0, \quad X_t = \mathbb{P}(Z_t \in \cdot \mid \mathcal{F}_t^0),
$$

where $\mathcal{F}_t^0 := \sigma(W_s^0, s \leq t)$. The same result provided by Theorem 7.1 can be obtained when the common noise is replaced by a common jump mechanism. For example, consider a poisson random measure $\mathcal{P}^0(dt, dy)$ with compensator $F(dy)dt$ for some probability measure F supported on $\mathbb{R}$, and let (X, Z) satisfy the McKean–Vlasov equation

$$dZ_t = b_{X_t}(Z_t)dt + \sigma_{X_t}(Z_t)dW_t + \int \ell_{X_{t-}}(Z_{t-}, y)\mathcal{P}^0(dt, dy), \quad X_t = \mathbb{P}(Z_t \in \cdot \mid \mathcal{F}_t^0),$$

where $\mathcal{F}_t^0 := \sigma(\mathcal{P}^0([0, s], dy), s \leq t)$. Here $\ell_\mu(x, y)$ describes the sizes of the common jumps, which we assume are confined to a cube $[0, c]^d$ for some $c > 0$. The generator of the probability measure valued process X is then the linear operator of Lévy type (4.1) given by

$$Lf(\mu) = \langle B_\mu(\partial f(\mu)), \mu \rangle + \int_{\mathbb{R}} f(\gamma(\mu, y)) - f(\mu)F(dy)$$

for $B_\mu(\varphi) = b_\mu^\top \nabla \varphi + \frac{1}{2} \operatorname{Tr}(\sigma_\mu^2 \nabla^2 \varphi) + \int (\varphi(\cdot + \ell_\mu(\cdot, y)) - \varphi)F(dy)$, where

$$\gamma(\mu, y) := (\cdot + \ell_\mu(\cdot, y))_* \mu \in \mathcal{P}_w.$$

Note that since F is a probability measure, we are free to choose $\chi \equiv 0$ as truncation function. Suppose now that b and σ satisfy the conditions of Theorem 7.1 with $\gamma = 0$. Assume also that $\ell_\mu(x, y) = \widetilde{\ell}_{w\mu}(x, y)$ for some continuous map $(\nu, x) \mapsto \widetilde{\ell}_\nu(x, y)$ from $\mathcal{X} \times \mathbb{R}^d$ to $[0, c]^d \setminus \{0\}$ such that $\widetilde{\ell}_\nu(x, y)$ is a continuous map from $\mathcal{X}$ to the space $C(\mathbb{R}^d, [0, c]^d)$ of continuous functions from $\mathbb{R}^d$ to $[0, c]^d$. Then there exists a solution X to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for every initial condition $\mu \in \mathcal{P}_w$. This result can be proved following the proof of Theorem 7.1.

The next corollary follows directly from Theorem 7.1.

Corollary 7.3. *Suppose that b, σ, and τ do not depend on μ and*

$$\frac{b_i(x)}{1 + |x|}, \quad \frac{\sigma_{ij}(x)^2}{1 + |x|^2}, \quad \frac{\tau_{ij}(x)}{1 + |x|} \in \mathbb{R} + C_0(\mathbb{R}^d), \qquad i, j \in \{1, \ldots, d\}.$$

Then there exists a solution X to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for every initial condition $\mu \in \mathcal{P}_w$.

We consider now a different linear operator $L \colon \mathcal{D}_w \to C(\mathcal{P}_w)$ of Lévy type (4.1) with $\kappa = 0$, $N = 0$, and B and Q given by

$$B_\mu(\varphi) = b_\mu^\top \nabla \varphi + \frac{1}{2} \operatorname{Tr}(\sigma_\mu^2 \nabla^2 \varphi), \quad Q_\mu(\varphi \otimes \varphi) = \alpha_\mu \Psi(\varphi \otimes \varphi) \tag{7.8}$$

for some maps $b \colon \mathcal{P}_w \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma \colon \mathcal{P}_w \times \mathbb{R}^d \to \mathbb{S}^d$, and $\alpha \colon \mathcal{P}_w \times \mathbb{R}^{2d} \to \mathbb{R}$. Here we use the notation $\Psi(\varphi \otimes \varphi)(x, y) := (\varphi(x) - \varphi(y))^2$ for all $\varphi \in C_w^\infty$ and $x \in \mathbb{R}$.

Theorem 7.4. *Assume that $b_\mu(x) = \widetilde{b}_{w\mu}(x)$, $\sigma_\mu(x) = \widetilde{\sigma}_{w\mu}(x)$, $\alpha_\mu(x, y) = \widetilde{\alpha}_{w\mu}(x, y)$ for some continuous maps $\widetilde{b} \colon \mathcal{X} \times \mathbb{R}^d \to \mathbb{R}^d$, $\widetilde{\sigma} \colon \mathcal{X} \times \mathbb{R}^d \to \mathbb{S}^d$ and $\widetilde{\alpha} \colon \mathcal{X} \times \mathbb{R}^{2d} \to \mathbb{R}$. Assume that*

$$\frac{\widetilde{b}_{\nu,i}(x)}{1 + |x|} \quad \text{and} \quad \frac{\widetilde{\sigma}_{\nu,ij}(x)^2}{1 + |x|^2}, \qquad i, j \in \{1, \ldots, d\},$$

are continuous as functions from $\mathcal{X}$ to $\mathbb{R} + C_0(\mathbb{R}^d)$, and that $\alpha_\mu(x, y)$ is continuous as a function from $\mathcal{X}$ to $\mathbb{R} + C_0(\mathbb{R}^{2d})$. Assume also that

$$\sup_{x \in \mathbb{R}^d} \frac{|x| |\widetilde{b}_\nu(x)| + |\widetilde{\sigma}_\nu(x)|^2}{1 + |x|^2} + \sup_{x, y \in \mathbb{R}^d} |\widetilde{\alpha}_\nu(x, y)| \leq c\langle 1, \nu \rangle^\gamma, \quad \nu \in \mathcal{X},$$

for some constants $c, \gamma \geq 0$. Then conditions (i)–(iii) of Theorem 3.4 are satisfied, and thus there exists a possibly killed solution X to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for every initial condition $\mu \in \mathcal{P}_w$. If one can take $\gamma = 0$, then condition (iv) of Theorem 3.4 holds, and thus $X_t \in \mathcal{P}_w$ for all $t \geq 0$.

Proof. The proof follows the proof of Theorem 7.1, applying Theorem 5.1(ii) to verify the positive maximum principle instead of Theorem 5.1(iv). □

Example 7.5. The Fleming–Viot diffusion was introduced by Fleming and Viot (1979) and subsequently studied by many other authors. Its generator is of the form (7.8) for $b \equiv 0$, σ constant, and $\alpha \equiv 1$. Here $E = \mathbb{R}$. A generalization of this model allows for so-called *weighted sampling* by letting the coefficient α be non-constant. The interpretation is that the *sampling-replacement rate* depends on the *types* x and y: $\alpha(x, y)$ denotes the rate at which an individual of type x is replaced by one of type y. See Section 5.7.8 of Dawson (1993) for further details. Theorem 7.4 permits us to deal with the case where α in addition depends on μ. Thus, the sampling-replacement rate depends not only on x and y, but on the entire distribution μ of types. As a concrete example, choosing $w(x) := x^2$, we can let the sampling-replacement rate be monotonic in the variance of the distribution. Setting $\mathrm{Var}(\mu) := \langle (\,\cdot\,)^2, \mu \rangle - \langle (\,\cdot\,), \mu \rangle^2$, the corresponding operators are then given by

$$B_\mu(\varphi) = \frac{1}{2}\sigma^2 \varphi'', \qquad Q_\mu(\varphi \otimes \varphi) = f\big(\mathrm{Var}(\mu)\big)\Psi(\varphi \otimes \varphi),$$

for some nonnegative increasing bounded function $f \in C^\infty(\mathbb{R})$. ◇

Similarly to standard SDEs with linearly growing coefficients, the linear growth properties (implicit in Lemma 6.4) imply that all moments of $\langle w, X_t \rangle$ are finite.

Proposition 7.6. *Let L be as in Theorem 7.1 and assume it satisfies the assumptions given there for $\gamma = 0$. Let X be a solution to the martingale problem for $(L, \mathcal{D}_w, \mathcal{P}_w)$ for some initial condition $\overline{\mu} \in \mathcal{P}_w$. Then the following conditions hold.*

(i) $\mathbb{E}[\langle w, X_t \rangle^k] \le \langle w, \overline{\mu} \rangle^k e^{C_k t}$ *for all $k \in \mathbb{N}$, where $C_k := k(k+1)C$ for some $C > 0$.*

(ii) *X solves the martingale problem for $(L, \mathcal{E}_w, \mathcal{P}_w)$ where $\mathcal{E}_w$ denotes the algebra generated by $\{\mu \mapsto \langle \varphi, \mu \rangle \colon \varphi \in C_w^\infty\}$ and Lf is given by (4.1) for all $f \in \mathcal{E}_w$. Moreover, $f(X) - f(X_0) - \int_0^\cdot Lf(X_s)ds$ is a k-integrable martingale for all $f \in \mathcal{E}_w$.*

Proof. To simplify the notation we just prove the case $d = 1$. Before to start observe that by the dominated convergence theorem the process $f(X) - f(\overline{\mu}) - \int_0^\cdot Lf(X_s)ds$ denotes a bounded martingale for each map f such that

$$\text{the pair } (f, Lf) \text{ with } Lf \text{ given by (4.1) lies in} \atop \text{the bp-closure of the graph } \{(h, Lh)\colon h \in \mathcal{D}_w\}. \tag{7.9}$$

We already shown in the proof of condition (iv) during the proof of Theorem 7.1 that (7.9) is satisfied by every function f in the algebra generated by $\mathcal{D}_w$ and $e^{-\langle w, \mu \rangle}$. Proceeding as in the proof of Lemma 6.3 we can extend this result to all maps $f : \mathcal{P}_w \to \mathbb{R}$ of the form $f(\mu) := p(f_1(\mu), \dots, f_n(\mu))$ for $p \in C^2(\mathbb{R}^n)$ with $p(0) = 0$ and $f_1, \dots, f_n$ being elements of that algebra.

(i): Fix now $c > \langle w, \overline{\mu} \rangle$ and set $f_c(\mu) := p_c(\langle w, \mu \rangle)\langle w, \mu \rangle^k$ for $p_c(x) = p(x/c)$ where $p \in C^\infty(\mathbb{R}_+)$ satisfies $1_{\{x \le 1\}} \le p(x) \le 1_{\{x \le 2\}}$. Note that f_c satisfies (7.9). Setting $T_c := \inf\{t \ge 0 \ : \ \langle w, X_t \rangle \ge c\}$ we have, due to (7.4) for $\gamma = 0$,

$$\mathbb{E}[f_c(X_{t \wedge T_c})] \le \langle w, \overline{\mu} \rangle^k + \int_0^t C_k \mathbb{E}[\langle w, X_s \rangle^k 1_{\{s < T_c\}}]ds$$

$$\le \langle w, \overline{\mu} \rangle^k + \int_0^t C_k \mathbb{E}[f_c(X_{s \wedge T_c})]ds.$$

The Gronwall inequality implies that $\mathbb{E}[f_c(X_{t\wedge T_c})] \leq \langle w, \overline{\mu}\rangle e^{C_k t}$ and Fatou's lemma yields

$$\mathbb{E}[\langle w, X_t\rangle^k] \leq \liminf_{c\to\infty} \mathbb{E}[p_c(\langle w, X_{t\wedge T_c}\rangle)]\langle w, X_{t\wedge T_c}\rangle^k] \leq \langle w, \overline{\mu}\rangle^k e^{C_k t}.$$

(ii): Fix now $g \in \mathcal{E}_w$ and set $g_c(\mu) := p_c(\langle w, \mu\rangle)g(\mu)$. Observe that each g_c satisfies condition (7.9), $|g_c(\mu)| \leq C\langle w, \mu\rangle^k$, and $|Lg_c(\mu)| \leq C\langle w, \mu\rangle^k$ for some k big enough and some constant C not depending on c and μ. Since by (i) we get

$$\mathbb{E}[\langle w, X_t\rangle^k - \langle w, \overline{\mu}\rangle^k - \int_0^t \langle w, X_s\rangle^k ds] \leq \langle w, \overline{\mu}\rangle^k(e^{C_k t} + 2 + C_k^{-1}(e^{C_k t} - 1)) < \infty$$

and $g_c \to g$ pointwise the claim follows by the dominated convergence theorem.

This completes the proof. $\square$

8 McKean–Vlasov equations with common noise

We continue to consider the setting of Section 7: $E = \mathbb{R}^d$, $w(x) = |x|^p$ for $|x| > 2$ and some $p \in (0, \infty)$, $L\colon \mathcal{D}_w \to C(\mathcal{P}_w)$ is a linear operator of Lévy type (4.1) with $\kappa = 0$, $N = 0$, and B and Q given by

$$B_\mu(\varphi) = b_\mu^\top \nabla\varphi + \frac{1}{2}\operatorname{Tr}((\sigma_\mu^2 + \tau_\mu^2)\nabla^2\varphi),$$
$$Q_\mu(\varphi \otimes \varphi) = (\tau_\mu\nabla\varphi) \otimes (\tau_\mu\nabla\varphi),$$

for some maps $b\colon \mathcal{P}_w \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma, \tau\colon \mathcal{P}_w \times \mathbb{R}^d \to \mathbb{S}^d$.

Definition 8.1. *A weak solution of the McKean–Vlasov equation specified by (b, σ, τ) is a tuple (X, Z, W, W^0), defined on some filtered probability space, where X and Z are adapted with values in $\mathcal{P}_w$ and $\mathbb{R}^d$, W and W^0 are independent d-dimensional Brownian motions, and such that*

$$dZ_t = b_{X_t}(Z_t)dt + \sigma_{X_t}(Z_t)dW_t + \tau_{X_t}(Z_t)dW_t^0 \tag{8.1}$$

and

$$X_t = \mathbb{P}(Z_t \in \cdot \mid \mathcal{G}_t)$$

for some filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ to which W^0 is adapted, and of which W is independent.

Our aim in this section is to give an existence result for the McKean–Vlasov equation specified by (b, σ, τ) by solving a martingale problem satisfied by the solution (X, Z, W, W^0). The state space for this martingale problem is the product space $\mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$. We will show below that the solution satisfies

$$d\langle\varphi, X_t\rangle = \langle B_{X_t}\varphi, X_t\rangle dt + \langle\tau_{X_t}\nabla\varphi, X_t\rangle^\top dW_t^0 \tag{8.2}$$

for each $\varphi \in C_c^\infty(\mathbb{R}^d)$. The corresponding generator H has domain $D(H)$ consisting of all algebraic combinations of functions $f(\mu)$, $\varphi(z)$, $\psi(x)$, $\theta(x^0)$ with $f \in \mathcal{D}_w$ and $\varphi, \psi, \theta \in C_c^\infty(\mathbb{R}^d)$. In view of (8.1) and (8.2), H acts on functions of the form $h(\mu, z, x, x^0) = f(\mu)\varphi(z)\psi(x)\theta(x^0)$ by the somewhat cumbersome expression

$$Hh(\mu, z, x, x^0) = Lf(\mu)\varphi(z)\psi(x)\theta(x^0) + f(\mu)B_\mu\varphi(z)\psi(x)\theta(x^0)$$
$$+ \frac{1}{2}f(\mu)\varphi(z)\Delta\psi(x)\theta(x^0) + \frac{1}{2}f(\mu)\varphi(z)\psi(x)\Delta\theta(x^0)$$
$$+ \nabla\varphi(z)^\top\tau_\mu(z)\langle\tau_\mu\nabla(\partial f(\mu)), \mu\rangle\psi(x)\theta(x^0)$$
$$+ \nabla\theta(x^0)^\top\langle\tau_\mu\nabla(\partial f(\mu)), \mu\rangle\psi(x)\varphi(z)$$
$$+ f(\mu)\theta(x^0)\nabla\varphi(z)^\top\sigma_\mu(z)\nabla\psi(x)$$
$$+ f(\mu)\psi(x)\nabla\varphi(z)^\top\tau_\mu(z)\nabla\theta(x^0).$$

Theorem 8.2. *Fix $\bar{z} \in \mathbb{R}^d$ and assume b, σ, τ satisfy the conditions of Theorem 7.1 for $\gamma = 0$. Then the martingale problem for $(H, D(H), \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ with initial condition $(\delta_{\bar{z}}, \bar{z}, 0, 0)$ has a solution (X, Z, W, W^0), where W and W^0 are independent d-dimensional Brownian motions. Moreover, the linear equation*

$$\langle \varphi, Y_t \rangle = \varphi(\bar{z}) + \int_0^t \langle B_{X_s}\varphi, Y_s \rangle ds + \int_0^t \langle \tau_{X_s} \nabla \varphi, Y_s \rangle^\top dW_s^0, \qquad \varphi \in C_c^\infty(\mathbb{R}^d) \qquad (8.3)$$

is satisfied for $Y = X$. If one has the compatibility conditions that W is independent of the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ generated by (X, W^0), and for all $s \leq t$, $\mathcal{F}_s$ and $\mathcal{G}_t$ are conditionally independent given $\mathcal{G}_s$, then (8.3) is satisfied for $Y_t = \mathbb{P}(Z_t \in \cdot \mid \mathcal{G}_t)$ as well. In particular, if in addition uniqueness holds for (8.3), then (X, Z, W, W^0) is a weak solution of the McKean–Vlasov equation specified by (b, σ, τ).

The compatibility conditions on the filtrations $\mathbb{F}$ and $\mathbb{G}$ are rather implicit. However, similar conditions are known to be required elsewhere in the literature; see for instance page 114 in Kurtz and Xiong (1999), and the conditions of Theorem 2 in Kailath et al. (1978). See also the remark at the beginning of page 142 in Kailath et al. (1978). Let us also mention (without proof) that whenever (X, W, W^0) solves the corresponding martingale problem, one can construct a process $\widetilde{W}$ such that $(X, \widetilde{W}, W^0)$ solves the same martingale problem and $\widetilde{W}$ is independent of the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ generated by (X, W^0).

Remark 8.3. Let $f(\nu) := p(\langle \varphi_1, \nu \rangle, \ldots, \langle \varphi_n, \nu \rangle)$ for some nonnegative map $p \colon \mathbb{R}^n \to \mathbb{R}$ satisfying $p(0) = 0$, some $\varphi_1, \ldots, \varphi_n \in C_c^\infty(\mathbb{R}^d)$, and all $\nu \in \mathcal{P}_w$. Note that setting $\langle \varphi_i, \nu - \tilde{\nu} \rangle := \langle \varphi_i, \nu \rangle - \langle \varphi_i, \tilde{\nu} \rangle$ we can naturally extend f to $\mathcal{P}_w - \mathcal{P}_w$. Consider now two solutions Y and $\widetilde{Y}$ of (8.3). An application of Itô's formula yields

$$\mathbb{E}[f(Y_t - \widetilde{Y}_t)] = \int_0^t \mathbb{E}[L_{X_s}f(Y_s - \widetilde{Y}_s)]ds,$$

where $L_\mu f(\nu) = \langle B_\mu(\partial f(\nu)), \nu \rangle + \frac{1}{2}\langle Q_\mu(\partial^2 f(\nu)), \nu^2 \rangle$. If f additionally satisfies

$$|L_\mu f(\nu)| \leq C f(\nu) \text{ for all } \nu \in \mathcal{P}_w - \mathcal{P}_w \text{ and } \mu \in \mathcal{P}_w, \qquad (8.4)$$

an application of the Gronwall inequality yields $\mathbb{E}[f(Y_t - \widetilde{Y}_t)] = 0$, and thus that $Y_t - \widetilde{Y}_t \in \{f = 0\}$. If this condition holds for sufficiently many f, we would be able to conclude that $Y_t = \widetilde{Y}_t$ almost surely and that uniqueness holds for (8.3). We illustrate a situation where this is the case in the following example.

Let $d = 1$, $\bar{z} \in [0, 1]$, and assume that $Y_t([0,1]) = \widetilde{Y}_t([0,1]) = 1$ for each $t \geq 0$. Assume that the maps $x \mapsto b_\mu(x)$ and $x \mapsto \tau_\mu(x)$ are polynomials of degree at most 1 and the map $x \mapsto \sigma_\mu(x)^2$ is a polynomial of degree at most 2. This in particular implies that B_μ and Q_μ are polynomial operators in the sense of Filipović and Larsson (2020), meaning that they map any polynomial to a polynomial of the same or lower degree. Fix then $H_0, \ldots, H_m \in C_c^\infty(\mathbb{R})$ such that $H_i(x) = x^i$ for each $x \in [0, 1]$ and set $p_m(\nu) = \sum_{i=0}^m \langle H_i, \nu \rangle^2$. Note that for each $\nu \in \mathcal{P}_w - \mathcal{P}_w$ such that $\operatorname{supp}(\nu) \subseteq [0, 1]$ we have

$$|L_\mu p_m(\nu)| = |\sum_{i=0}^m 2\langle H_i, \nu \rangle\langle B_\mu H_i, \nu \rangle + \langle Q_\mu(H_i \otimes H_i), \nu^2 \rangle|$$

$$= |\sum_{i,j=0}^m \alpha_{ij}^\mu \langle H_i, \nu \rangle\langle H_j, \nu \rangle|$$

$$\leq (m+1)\sup_{ij} |\alpha_{ij}^\mu| p_m(\nu),$$

for some $\alpha_{ij}^\mu \in \mathbb{R}$. This implies that if $\sup_{\mu \in \mathcal{P}_w} |\alpha_{ij}^\mu| < \infty$, then condition (8.4) is satisfied, and $\langle H_i, Y_t \rangle = \langle H_i, \widetilde{Y}_t \rangle$ for each $i \in \{1, \ldots, m\}$. Since m was arbitrary the same conclusion holds for each $i \in \mathbb{N}$. Since two measures on $[0, 1]$ have the same moments if and only if they are the same, it follows that $Y_t = \widetilde{Y}_t$ almost surely and uniqueness holds for (8.3).

The rest of the section is devoted to the proof of Theorem 8.2. We start with a corollary of Theorem 7.1.

Corollary 8.4. *Assume that b, σ, and τ satisfy the conditions of Theorem 7.1 for $\gamma = 0$. Then there exists a solution (X, Z, W, W^0) to the martingale problem for $(H, D(H), \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ for every initial condition $(\mu, z, x, x^0) \in \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$.*

Proof. We first observe that H satisfies the positive maximum principle on $\mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$. This can be proven by the classical optimality conditions on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$, Theorem 5.1(i), and a slightly modification of the argument in the proof of Theorem 5.1(iii).

Observe then that the concepts introduced in Section 3 can be generalized by setting

$$\mathcal{T} : \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d, \qquad \mathcal{T}(\mu, z, x, x^0) := (T(\mu), z, x, x^0)$$

and calling $f : \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ of C_0 type if $f \circ \mathcal{T}^{-1} : \mathcal{T}(\mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d) \to \mathbb{R}$ extends to a C_0 function on $\mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$. Since L satisfies the conditions of Theorem 7.1 we know that Hh is of C_0 type for every $h \in D(H)$. Following the proof of Theorem 3.4 we can conclude that there exists a possibly killed solution of the martingale problem for $(\widetilde{H}, D(H) \circ \mathcal{T}^{-1}, \mathcal{X} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ where

$$\widetilde{H}\widetilde{h} = H(\widetilde{h} \circ \mathcal{T}) \circ \mathcal{T}^{-1}.$$

Using that by Theorem 7.1 the operator L satisfies conditions (iii)–(iv) of Theorem 3.4, we can conclude the proof by following the proof of Theorem 3.4. $\square$

The fact that a solution of the martingale problem is also a weak solution of the corresponding SDE is due, in the classical case, to Stroock and Varadhan (1972).

Lemma 8.5. *Assume that the conditions of Corollary 8.4 are satisfied and consider a solution (X, Z, W, W^0) to the martingale problem for $(H, D(H), \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ with initial condition $(\delta_{\overline{z}}, \overline{z}, 0, 0)$. Then W, W^0 are independent Brownian motions, and (8.1) and (8.2) hold for each $\varphi \in C_c^\infty(\mathbb{R}^d)$.*

Proof. For $h(\mu, z, x, x^0) = \psi(x, x^0)$, we have that $Hh(\mu, z, x, x^0) = \frac{1}{2}\Delta\psi(x, x^0)$ is the Laplacian. Thus W and W^0 are independent Brownian motions. To prove (8.2), we must show that the process $\langle \varphi, X_t \rangle - \int_0^t \langle B_{X_s}\varphi, X_s \rangle ds - \int_0^t \langle \tau_{X_s}\nabla\varphi(X_s), X_s \rangle^\top dW_s^0$, which is known to be a martingale due to Proposition 7.6, is constant. This is done by verifying that its quadratic variation is zero; we omit the details. The proof of (8.1) is similar. $\square$

Lemma 8.6. *Assume that the conditions of Corollary 8.4 are satisfied and consider a solution (X, Z, W, W^0) to the martingale problem for $(H, D(H), \mathcal{P}_w \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d)$ with initial condition $(\delta_{\overline{z}}, \overline{z}, 0, 0)$. If one has the compatibility conditions that W is independent of the filtration $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ generated by (X, W^0), and for all $s \leq t$, $\mathcal{F}_s$ and $\mathcal{G}_t$ are conditionally independent given $\mathcal{G}_s$, then the conditional law process $Y_t := \mathbb{P}(Z_t \in \, \cdot \, | \, \mathcal{G}_t)$ satisfies (8.3).*

Before we start the proof, observe that condition (7.2) for $\gamma = 0$ implies

$$\sup_{x \in \mathbb{R}, \, \mu \in \mathcal{P}_w} |B_\mu(x)\varphi(x)| + |\tau_\mu(x)\nabla\varphi(x)| < \infty \tag{8.5}$$

for each $\varphi \in C_c^\infty(\mathbb{R}^d)$.

Proof. Observe that an application of the Itô formula yields

$$\varphi(Z_t) = \varphi(\overline{z}) + \int_0^t B_{X_s}\varphi(Z_s)ds + \int_0^t (\sigma_{X_s}(Z_s)\nabla\varphi(Z_s))^\top dW_s + \int_0^t (\tau_{X_s}(Z_s)\nabla\varphi(Z_s))^\top dW_s^0,$$

for each $\varphi \in C_c^\infty(\mathbb{R}^d)$. Note that condition (8.5) yields that $\mathbb{E}[|\tau_{X_s}(Z_s)\nabla\varphi(Z_s)|^2]$ and $\mathbb{E}[|\sigma_{X_s}(Z_s)\nabla\varphi(Z_s)|^2]$ are almost surely bounded in s for each $\varphi \in C_c^\infty(\mathbb{R}^d)$. Combining Lemma A.1, Fubini theorem for conditional expectation, and the $\mathcal{G}_s$-measurability of X_s we then get

$$\mathbb{E}[\varphi(Z_t) \mid \mathcal{G}_t] = \varphi(\overline{z}) + \int_0^t \mathbb{E}[B_{X_s}(Z_s) \mid \mathcal{G}_s]ds + \int_0^t \mathbb{E}[(\tau_{X_s}(Z_s)\nabla\varphi(Z_s)) \mid \mathcal{G}_s]^\top dW_s^0$$

$$= \varphi(\overline{z}) + \int_0^t \langle B_{X_s}\varphi, Y_s \rangle ds + \int_0^t \langle \tau_{X_s}\nabla\varphi, Y_s \rangle^\top dW_s^0,$$

for each $\varphi \in C_c^\infty(\mathbb{R}^d)$. □

Proof of Theorem 8.2. The result follows directly from Corollary 8.4 and Lemmas 8.5 and 8.6. □

A A Fubini type result

The result presented in this section is based on Theorem 2 in Kailath et al. (1978) and its proof. Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space endowed with two d-dimensional Brownian motions W and W^0. Consider then a second filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ to which W^0 is adapted, and of which W is independent.

Lemma A.1. *If $\mathcal{F}_s$ and $\mathcal{G}_t$ are conditionally independent given $\mathcal{G}_s$, then*

$$\mathbb{E}[\int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t] = \int_0^t \mathbb{E}[H_s^\top \mid \mathcal{G}_s]dW_s^0 \qquad \text{and} \qquad \mathbb{E}[\int_0^t H_s^\top dW_s \mid \mathcal{G}_t] = 0,$$

for each square integrable continuous process H satisfying $\int_0^t \mathbb{E}[|H_s|^2]ds < \infty$.

Observe that the conditional independence assumption implies that each $\mathbb{G}$-martingale is also an $\mathbb{F}$-martingale. This condition is automatically satisfied if $\mathcal{G}_t := \sigma(W_s^0, s \leq t)$.

Proof. Set $\xi_t := \mathbb{E}[\int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t]$ and note that ξ and W_0 are square integrable martingales with respect to $\mathbb{G}$ and thus with respect to $\mathbb{F}$. Moreover, since $W_t^0 \xi_t = \mathbb{E}[W_t^0 \int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t]$, for each $A \in \mathcal{G}_u$ we can compute

$$\mathbb{E}[(W_t^0\xi_t - W_u^0\xi_u)1_A] = \mathbb{E}\Big[\Big(W_t^0\mathbb{E}[\int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t] - W_u^0\mathbb{E}[\int_0^u H_s^\top dW_s^0 \mid \mathcal{G}_u]\Big)1_A\Big]$$

$$= \mathbb{E}\Big[\Big(W_t^0\int_0^t H_s^\top dW_s^0 - W_u^0\int_0^u H_s^\top dW_s^0\Big)1_A\Big]$$

$$= \mathbb{E}[\int_u^t (H_s^\top 1)ds1_A]$$

$$= \mathbb{E}\Big[\int_u^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top 1ds1_A\Big],$$

and thus conclude that $W_t^0\xi_t - \int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top 1ds$ is an $\mathbb{F}$-martingale. This in particular implies that $\int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top 1ds$ is the predictable quadratic covariation of ξ and W^0

with respect to $\mathbb{F}$. Since $\int_0^t H_s^\top dW_s^0$ is a square integrable $\mathbb{F}$-martingale and $\xi_t^2 = \mathbb{E}[\xi_t \int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t]$ we get

$$
\begin{aligned}
\mathbb{E}[\xi_t^2] &= \mathbb{E}[\xi_t \mathbb{E}[\int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t]] \\
&= \mathbb{E}[\xi_t \int_0^t H_s^\top dW_s^0] \\
&= \mathbb{E}[\int_0^t H_s^\top \mathbb{E}[H_s \mid \mathcal{G}_s] ds] \\
&= \mathbb{E}[\int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top \mathbb{E}[H_s \mid \mathcal{G}_s] ds].
\end{aligned}
$$

Similarly we also get that $\mathbb{E}[\xi_t \int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top dW_s^0] = \mathbb{E}[\int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top \mathbb{E}[H_s \mid \mathcal{G}_s] ds]$. Using that $\mathbb{E}[(\int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top dW_s^0)^2] = \mathbb{E}[\int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top \mathbb{E}[H_s \mid \mathcal{G}_s] ds]$ we can thus conclude that

$$
\mathbb{E}[(\xi_t - \int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top dW_s^0)^2] = 0
$$

proving that $\mathbb{E}[\int_0^t H_s^\top dW_s^0 \mid \mathcal{G}_t] = \int_0^t \mathbb{E}[H_s \mid \mathcal{G}_s]^\top dW_s^0$.

For the second part set $\eta_t := \mathbb{E}[\int_0^t H_s^\top dW_s \mid \mathcal{G}_t]$. Since η and W are two independent continuous $\mathbb{F}$-martingales we already know that $(\eta_t W_t)_{t \geq 0}$ defines a square integrable $\mathbb{F}$-martingale. Proceeding as in the first part we can thus conclude that $\mathbb{E}[\eta_t^2] = \mathbb{E}[\eta_t \int_0^t H_s^\top dW_t] = 0$. $\qquad\square$

References

E. Abi Jaber, B. Bouchard, and C. Illand. Stochastic invariance of closed sets with non-lipschitz coefficients. *Stochastic Processes and their Applications*, 129(5):1726–1748, 2019. MR-3944782

R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2017. MR-3753660

C. Cuchiero. Polynomial processes in stochastic portfolio theory. *Stochastic processes and their applications*, 129(5):1829–1872, 2019. MR-3944786

C. Cuchiero, M. Larsson, and S. Svaluto-Ferro. Probability measure-valued polynomial diffusions. *Electronic Journal of Probability*, 24, 2019. MR-3933209

G. Da Prato and H. Frankowska. Invariance of stochastic control systems with deterministic arguments. *Journal of Differential Equations*, 200(1):18–52, 2004. MR-2046316

D. Dawson. The critical measure diffusion process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 40(2):125–145, 1977. MR-0478374

D. Dawson. Geostochastic calculus. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 6(2):143–168, 1978. MR-0532855

D. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI–1991*, Lecture Notes in Mathematics, pages 1–260. Springer, 1993. MR-1242575

D. Dawson and J. Vaillancourt. Stochastic McKean-Vlasov equations. *Nonlinear Differential Equations and Applications NoDEA*, 2(2):199–229, 1995. MR-1328577

A. Etheridge. Some Mathematical Models from Population Genetics. In *École d'Été de Probabilités de Saint-Flour XXXIX-2009*, Lecture Notes in Mathematics. Springer, 2011. MR-2759587

S. N. Ethier and T. G. Kurtz. *The Infinitely-Many-Alleles Model with Selection as a Measure-Valued Diffusion*, pages 72–86. Springer, Berlin, 1987. MR-0893637

S. N. Ethier and T. G. Kurtz. Fleming–Viot processes in population genetics. *SIAM Journal on Control and Optimization*, 31(2):345–386, 1993. MR-1205982

S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2 edition, 2005. MR-0838085

R. Fernholz. *Stochastic Portfolio Theory*. Applications of Mathematics. Springer-Verlag, New York, 2002. MR-1894767

R. Fernholz and I. Karatzas. Stochastic portfolio theory: an overview. *Handbook of numerical analysis*, 15:89–167, 2009. MR-1894767

D. Filipović and M. Larsson. Polynomial Jump-Diffusion Models. *Stochastic Systems*, 10(1):71–97, 2020. MR-4084227

W. H. Fleming and M. Viot. Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.*, 28(5):817–843, 1979. MR-0542340

P. Florchinger and F. Le Gland. Particle approximation for first order stochastic partial differential equations. In *Applied stochastic analysis*, pages 121–133. Springer, 1992. MR-1169922

H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. De Gruyter studies in mathematics. Walter de Gruyter, 2 edition, 2004. MR-1925197

K. Huang. *Statistical mechanics*. Wiley, 1987. MR-1042093

T. Kailath, A. Segall, and M. Zakai. Fubini-Type Theorems for Stochastic Integrals. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 40(2):138–143, 1978. MR-0546404

A. Klenke. *Probability Theory: A Comprehensive Course*. Universitext. Springer London, 2 edition, 2013. MR-3112259

T. Kurtz and J. Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications*, 83, 1999. MR-1705602

E. Perkins. Dawson-Watanabe Superprocesses and Measure-valued Diffusions. In *Ècole d'Ètè de Probabilitès de Saint-Flour XXIX–1999*, Lecture Notes in Mathematics, pages 125–329. Springer, 2002. MR-1915445

D. W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, volume 3, pages 333–359, 1972. MR-0400425

A. S. Sznitman. Topics in Propagation of Chaos. In *École d'Été de Probabilités de Saint-Flour XIX–1989*, Lecture Notes in Mathematics, pages 165–251. Springer, 1991. MR-1108185

C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008. MR-2459454

S. Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.*, 8(1):141–167, 1968. MR-0237008

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm