KOLMOGOROV EQUATIONS ON THE SPACE OF PROBABILITY MEASURES ASSOCIATED TO THE NONLINEAR FILTERING EQUATION: THE VISCOSITY APPROACH

MATTIA MARTINI

Abstract. We study the backward Kolmogorov equation on the space of probability measures associated to the Kushner-Stratonovich equation of nonlinear filtering. We prove existence and uniqueness in the viscosity sense and, in particular, we provide a comparison theorem.

In the context of stochastic filtering it is natural to consider measure-valued processes that satisfy stochastic differential equations. In the literature, a classical way to address this problem is by assuming that these measure-valued processes admit a density. Our approach is different and we work directly with measures. Thus, the backward Kolmogorov equation we study is a second-order partial differential equation of parabolic type on the space of probability measures with compact support. In the literature only few results are available on viscosity solutions for this kind of problems and in particular the uniqueness is a very challenging issue. Here we find a viscosity solution and then we prove that it is unique via comparison theorem.

1. Introduction

The main goal of this paper is to discuss existence and uniqueness for viscosity solutions to a class of backward Kolmogorov equations on spaces of probability measures, arising in the context of stochastic filtering.

The study of partial differential equations on spaces of probability measures is a topic of increasing research interest. For instance, in [AG08; GT19; BCP18; PW18; CCP20] and references therein one can find many results both for the first and second-order case. A main issue is how to define the derivatives on this spaces and many approaches are available in the literature (see [AGS05] for a more geometric approach related to optimal transport or [Lio07; CD18] for a more functional analytic and probabilistic approach connected with the theory of mean field games).

Here we continue the study started in [Mar21]. In particular, we consider a stochastic differential equation for a probability measure-valued process arising in the context of nonlinear filtering, and then we study the associated backward Kolmogorov equation, that is a second-order linear partial differential equation of parabolic type on the space of probability measures. We prove the existence and uniqueness of a viscosity solution to this equation.

The stochastic filtering problem is about finding a way to approximate a signal (described by a stochastic process) that can not be directly observed. In the last decades a lot of research interest has focused in this direction, see for instance [BC09; Xio08] and references therein for an exhaustive introduction to the topic. A key object for the theory is the so-called filter, a probability measure-valued stochastic process defined, at any time, as the conditional law of the signal process given the available information. One can prove that this process is Markovian and satisfies a stochastic differential equation (in a way we will point out later), called Kushner-Stratonovich equation. Then, it is natural to study the associated backward Kolmogorov equation.

A classical way to approach this problem (see for instance [Roz90]), is to show that the filter admits a density which belongs to a suitable Hilbert space, and then exploit the tools of stochastic calculus available in this case. However, this requires the introduction of unnecessary
hypotheses. In this paper we avoid these conditions and rather follow the approach of [BC09; HL01; Szp78], where the filtering process is studied as genuine measure-valued process.

Let us fix a final time \( T > 0 \), a \( d \)-dimensional Brownian motion \( I = \{I_t, t \in [0,T]\} \) and let us denote with \( \pi = \{\pi_t, t \in [0,T]\} \) the filter, which takes values in \( \mathcal{P}(\mathbb{R}^d) \), the space of probability measures over \( \mathbb{R}^d \). The Kushner-Stratonovich equation reads as follow: for every test function \( \varphi \), it holds
\[
d\langle \pi_t, \varphi \rangle = \langle \pi_t, A \varphi \rangle \, dt + \left( \langle \pi_t, h \varphi + B \varphi \rangle^\top - \langle \pi_t, \varphi \rangle \langle \pi_t, h \rangle^\top \right) \, dI_t.
\]
Here \( \langle \mu, \varphi \rangle := \int \varphi(x) \mu(dx) \), \( A \) and \( B \) are respectively a second and first-order differential operators whilst \( h : \mathbb{R}^d \to \mathbb{R}^d \). The complete and rigorous framework is provided in Section 2.2. Under hypotheses that guarantee the existence and uniqueness of a Markovian solution to (1.1), it has been proved in [Mar21] that the Kolmogorov operator associated to the Kushner-Stratonovich equation has the following form:
\[
(\mathcal{L}u)(\mu) = \langle \mu, A \delta_\mu u(\mu) \rangle + \frac{1}{2} \langle \mu \otimes \mu, (h + B - \langle \mu, h \rangle) \rangle (h + B - \langle \mu, h \rangle) \delta_\mu^2 u(\mu),
\]
where \( \delta_\mu u \) and \( \delta_\mu^2 u \) are suitably defined derivatives of real-valued functions over \( \mathcal{P}(\mathbb{R}^d) \) (see Section 2.1 for the precise definitions). Thus, given a final condition \( \Phi : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R} \), the backward Kolmogorov equation reads as
\[
\begin{cases}
- \partial_t u(\mu, t) - \mathcal{L}u(\mu, t) = 0, & (\mu, t) \in \mathcal{P}(\mathbb{R}^d) \times [0,T), \\
u(\mu, T) = \Phi(\mu), & \mu \in \mathcal{P}(\mathbb{R}^d).
\end{cases}
\]
In [Mar21], the author proved that if \( \Phi \) is twice continuously differentiable, then there exists a unique classical solution to (1.2).

The aim of this paper is to study the case of a less regular final condition, namely just continuous and bounded. In this case, the notion of classical solution turns out to be too strong, thus a weaker notion is necessary. In particular, the notion of viscosity solution seems to be the natural one for this kind of problems.

The theory of viscosity solutions is well-established in the finite dimensional case (see for instance [CIL92]) and in the infinite dimensional Hilbert case (see for instance [Swi94]). Recently some developments have been obtained for more complicate spaces, where it is not easy to provide a notion of derivative. Two examples are the space of continuous paths (see for instance [CR22] and references therein) and the space of probability measures, to which the present work refers. In particular, in [BCP18; PW17; PW18; Cos+20; Ban+19] the existence of viscosity solutions to second-order equations on the space of probability measures is addressed. On the other hand, uniqueness results are very hard to achieve, and only few papers are available for second-order equations ([Bur+20; WZ20; Cos+21]).

Our work gives a contribution in this direction. In Proposition 3.7 we show that there exists a viscosity solution to (1.2). Then, we are also able to prove a uniqueness result via comparison principle (Theorem 3.9 and Corollary 3.10). The technique is based on the fact that we have a candidate solution (given by a probabilistic representation formula of the usual type) that can be used to obtain partial comparison principles. An assumption we need in order to achieve this last result, is to restrict ourselves to the compact subspace \( \mathcal{P}(K) \subset \mathcal{P}(\mathbb{R}^d) \), the space of probabilities supported on a compact set \( K \subset \mathbb{R}^d \). This is for instance the case when the filtering equation is associated to a signal with trajectories confined on a compact subset or on a compact manifold like a torus. However, a remarkable feature of this approach is to avoid the lifting of the problem on a Hilbert space, which is a procedure often used to solve this type of second-order problems. In that case, for viscosity solutions it is not yet clear what is the relation between the lifted problem and the original one.
2. Notations and preliminaries

Let $E \subseteq \mathbb{R}^d$. We denote with $C(E)$ the space of continuous real-valued functions over $E$, with $C_b(E)$ the continuous and bounded real-valued functions over $E$ and with $C_c(E)$ the continuous functions over $E$ with compact support. Similarly, for $k \in \mathbb{N}$, we denote with $C^k(E)$ (resp. $C^k_b(E)$ and $C^k_c(E)$) the space of $k$-time continuously differentiable (resp. continuously differentiable with bounded derivatives up to order $k$-th and continuously differentiable with compact support) functions over $E$. If $E$ is compact, we say that a function is continuously differentiable over $E$ if it has bounded and uniformly continuous derivative over Int$(E)$.

Let $u$ be a real-valued (or $\mathbb{R}^d$-valued) function over a suitable space $X$. We denote with $\|u\|_\infty := \sup_{x \in X} |u(x)|$ its supremum norm.

We denote with $\mathcal{P}(E)$ the space of probability measures over $E$ and with $\langle \mu, \varphi \rangle$ the integral of $\varphi \in C_b(E)$ with respect to $\mu \in \mathcal{P}(E)$. We say that a sequence $\{\mu_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(E)$ converges weakly to $\mu$ if $\langle \mu_n, \varphi \rangle \to \langle \mu, \varphi \rangle$ as $n \to +\infty$, for every $\varphi \in C_b(E)$. If we take $E$ compact, we have that the sets $\mathcal{P}(E)$ coincides with the set of probability measures with finite moment of any order. Moreover, one can prove that if $E$ is compact, then $\mathcal{P}(E)$ is also compact in the weak topology (see for instance [Vil08]).

Let us consider a family $\{f_k\}_{k \in \mathbb{N}}$ of functions dense in $C^\infty_c(E)$, with $\sup_{x \in E} |f_k(x)| =: \|f_k\|_\infty \leq 1$, and let us assume that the function identically equal to 1 belongs to this family. Let us introduce a sequence $\{q_k\}_{k \in \mathbb{N}} \subset [1, \infty)$ that we will make it explicit in Appendix A. We define $d_2: \mathcal{P}(E) \times \mathcal{P}(E) \to [0, +\infty)$ as

$$d_2(\mu, \nu) := \left( \sum_{k=1}^{+\infty} \frac{1}{2^k q_k} \|\mu - \nu, f_k\|^2 \right)^{\frac{1}{2}}. \quad (2.1)$$

It is easy to show that $d_2$ is a distance over $\mathcal{P}(E)$ and, if we take $E$ compact, one can also prove that it is complete (see [Bor95, Theorem 2.4.1]). Moreover, $d_2$ metrize the weak convergence ([Bor95, Theorem 2.1.1]).

For the rest of the paper, we will denote with $K$ a compact subset of $\mathbb{R}^d$.

2.1. Notions of derivative on $\mathcal{P}(K)$. Let $K \subset \mathbb{R}^d$ be compact. We need to introduce a notion of derivative for maps from $\mathcal{P}(K)$ to $\mathbb{R}$. Among the many possible notions available in literature, we follow [CCP20]. We start from the notion of linear functional derivative (see [CD18]) and then we introduce the definition of $L$-derivative (which turns out to coincide with the one introduced by P.-L. Lions in [Lio07], see for instance [CD18]).

**Definition 2.1** (Linear functional derivative). We say that a function $u: \mathcal{P}(K) \to \mathbb{R}$ has linear functional derivative if there exists a map

$$\delta_\mu u: \mathcal{P}(K) \times K \to \mathbb{R},$$

which is jointly continuous and such that for every $\mu, \nu \in \mathcal{P}(K)$ it holds

$$u(\mu) - u(\nu) = \int_0^1 \int_K \delta_\mu u(\theta \mu + (1 - \theta)\nu, x)[\mu - \nu](dx) d\theta. \quad (2.2)$$
We denote by $C^1(\mathcal{P}(K))$ the space of functions that are differentiable in linear functional sense.

**Remark 2.2.** On $\mathcal{P}(K)$ the linear functional derivative is defined up to an additive constant. However, we can choose the version such that, for every $\mu \in \mathcal{P}(K)$,

$$\int_K \delta_\mu u(\mu, x) \mu(dx) = 0.$$  

**Remark 2.3.** Iterating Definition 2.1 (keeping the extra spatial variable in $\delta_\mu u$ fixed), one can give a notion of second-order derivative. In particular $\delta^2_\mu u$ will be a continuous mapping from $\mathcal{P}(K) \times K \times K$ to $\mathbb{R}$. Thus, we can introduce the spaces $C^2(\mathcal{P}(K))$. Analogously, this definition can be extended up to the $k$-th order, $k \in \mathbb{N}$.  

**Definition 2.4** ($L$-derivative). We say that a function $u$ is in $C^2_L(\mathcal{P}(K))$ if:

i. $u \in C^2(\mathcal{P}(K))$;

ii. for every $\mu \in \mathcal{P}(K)$, the maps $K \ni x \mapsto \delta_\mu u(\mu, x)$ and $K \times K \ni (x, y) \mapsto \delta^2_\mu u(\mu, x, y)$ are twice continuously differentiable, with jointly continuous derivatives over $\mathcal{P}(K) \times K$ and $\mathcal{P}(K) \times K \times K$.

We define the first and second-order $L$-derivatives as

$$D_\mu u(\mu, x) := D_x \delta_\mu u(\mu, x) \in \mathbb{R}^d, \quad D^2_\mu u(\mu, x, y) := D_x D^\top_y \delta^2_\mu u(\mu, x, y) \in \mathbb{R}^{d \times d},$$

for every $\mu \in \mathcal{P}(K)$ and $x, y \in K$.

**Example 2.5.** Let $n \in \mathbb{N}$, $g: \mathbb{R}^n \to \mathbb{R}$ be in $C^2_\mathcal{L}(\mathbb{R}^d)$ and let $\varphi_1: K \to \mathbb{R}$ be in $C^1(K)$ for every $i = 1, \ldots, n$. We define $f: \mathcal{P}(K) \to \mathbb{R}$ by

$$f(\mu) := g(\langle \mu, \varphi_1 \rangle, \ldots, \langle \mu, \varphi_n \rangle) = g(\langle \mu, \varphi \rangle).$$

Then, $f \in C^2_L(\mathcal{P}(K))$ and for every $\mu \in \mathcal{P}(K)$ and $x, y \in K$ it holds

$$\begin{align*}
\delta_\mu f(\mu, x) &= \sum_{i=1}^n \partial_i g(\langle \mu, \varphi \rangle) \varphi_i(x), \\
\delta^2_\mu f(\mu, x, y) &= \sum_{i,j=1}^n \partial^2_{ij} g(\langle \mu, \varphi \rangle) \varphi_i(x) \varphi_j(y),
\end{align*}$$

$$D_\mu f(\mu, x) = \sum_{i=1}^n \partial_i g(\langle \mu, \varphi \rangle) D_x \varphi_i(x), \quad D^2_\mu f(\mu, x, y) = \sum_{i,j=1}^n \partial^2_{ij} g(\langle \mu, \varphi \rangle) D_x \varphi_i(x) D_y \varphi_j(y).$$

### 2.2 Stochastic filtering framework.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $\{\mathcal{F}_t\}_{t \in [0,T]}$ that satisfies the usual conditions and let $I = \{I_t, t \in [0,T]\}$ be a $d$-dimensional $\{\mathcal{F}_t\}$-Brownian motion. Let $b, h: \mathbb{R}^d \to \mathbb{R}^d, \sigma: \mathbb{R}^d \to \mathbb{R}^{d \times d}$ and let us define two differential operators $A: C^2(\mathcal{P}(K)) \to C(K)$ and $B: C^2(\mathcal{P}(K)) \to C(K; \mathbb{R}^d)$ by

$$A\varphi := b^\top D_x \varphi + \frac{1}{2} \text{tr} \left\{ \sigma \sigma^\top D^2_x \varphi \right\} + \frac{1}{2} \text{tr} \left\{ \bar{\sigma} \bar{\sigma}^\top D^2_x \varphi \right\}, \quad B\varphi := \bar{\sigma}^\top (D_x \varphi).$$

We consider the Kushner-Stratonovich equation in weak form:

$$d\langle \pi_t, \varphi \rangle = \langle \pi_t, A\varphi \rangle dt + \left( \langle \pi_t, h\varphi + B\varphi \rangle^\top - \langle \pi_t, \varphi \rangle \langle \pi_t, h \rangle^\top \right) dI_t, \quad \varphi \in C^2(\mathcal{P}(K)).$$

**Remark 2.6.** This equation can be related to a stochastic filtering problem with signal $X = \{X_t, t \in [0,T]\}$ and observation $Y = \{Y_t, t \in [0,T]\}$ such that:

$$\begin{align*}
\text{d}X_t &= b(X_t) \text{d}t + \sigma(X_t) \text{d}B_t + \bar{\sigma}(X_t) \text{d}W_t, \\
\text{d}Y_t &= h(X_t) \text{d}t + W_t,
\end{align*}$$

where $W = \{W_t, t \in [0,T]\}$ and $B = \{B_t, t \in [0,T]\}$ are two independent and $d$-dimensional Brownian motions. If we take $I$ as the innovation process associate to (2.5) (see for instance [BC09, Chapter 3] for more details), then the equation (2.4) is the one solved by the optimal filter $\hat{\pi}_t = \mathcal{L}(X_t, I^T_t)$, where $I^T_t := \sigma(Y_s, 0 \leq s \leq t)$. However, we will look at (2.4) in the spirit of [Szp78], without relying directly on the filtering problem.
Following [HL01; Szp78], we say that the pair \( \{(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}), (\pi, I)\} \) is a weak solution to the equation (2.4) if

i. \( (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}) \) is a complete filtered probability space;

ii. \( I \) is a \( d \)-dimensional \( \{\mathcal{F}_t\} \)-Brownian motion;

iii. \( \pi = \{\pi_t, t \in [0, T]\} \) is a probability measure-valued continuous \( \{\mathcal{F}\} \)-adapted process such that

\[
\mathbb{P}\left( \int_0^T \sum_{i=1}^d \pi_s(h_i)^2 \, ds < +\infty \right) = 1,
\]

and equation (2.4) holds for every \( t \in [0, T] \), almost surely.

Now we give some conditions that guarantee the existence, uniqueness (in law) and Markov property of a weak solution to the filtering equation (2.4). A complete discussion can be found in [Szp78; HL01].

**Hypothesis 2.7.** The following conditions hold:

a. \( b, \sigma, \bar{\sigma}, h \) are Borel-measurable and bounded;

b. \( b, \sigma, \bar{\sigma} \) are Lipschitz continuous;

c. for every \( x \in K \) the matrix \( \sigma(x)\sigma(x)\top \) is a positive definite.

Moreover, we will ask for the following invariance property:

**Hypothesis 2.8.** The probability measures-valued process \( \pi \) has trajectories confined in the subset \( \mathcal{P}(K), K \subset \mathbb{R}^d \) compact.

For instance, this is the case when \( \pi \) is the filter associated to a signal process \( X = \{X_t, t \in [0, T]\} \) with trajectories confined in a compact subset \( K \subset \mathbb{R}^d \). Some conditions on \( b, \sigma, \bar{\sigma} \) that guarantee this property are stated in Remark 2.9.

**Remark 2.9.** Let \( G \) be a subset of \( \mathbb{R}^d \). The invariance of a stochastic process with respect to \( G \) has been intensively studied and many criteria that guarantee this property are available in the literature. For instance, following [Fri75, Section 12.2], let \( G \subset \mathbb{R}^d \) be closed and with \( C^3 \) connected boundary \( \partial G \). Let \( \nu \) be the outward normal to \( \partial G \) and let us define \( \rho(x) := \inf_{y \in G} |x - y| \) the distance from \( x \in \mathbb{R}^d \) to \( G \). If the conditions

\[
\sum_{i,j=1}^d a_{ij} \nu_i \nu_j = 0, \quad \sum_{i=1}^d b_i \nu_i + \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ij}^2 \rho \geq 0 \quad \text{on} \ \partial G,
\]

hold, where \( a = \sigma \sigma\top + \bar{\sigma} \bar{\sigma}\top \), then \( G \) is invariant for \( X \). Other less restrictive conditions can be found in [ADP90; DPF04; DPF07] and in the references therein.

**Remark 2.10.** All the following results do not change when \( K \) is a compact manifold as long as we consider the right differential operators. An example is when the signal process lives on the \( d \)-dimensional torus \( \mathbb{T}^d \).

### 3. Viscosity solutions to the backward Kolmogorov equations

The aim of this section is to continue the study of the backward Kolmogorov equation associated to the Kushner-Stratonovich equation (2.4). This partial differential equation of parabolic type over \( \mathcal{P}(K) \times [0, T] \) has been introduced in [Mar21]. Here we present a well posedness result for viscosity solutions, which is different from the one available in [Mar21], where existence and uniqueness of classical solutions have been proved under more restrictive requirements on the final condition.
Before writing down the equation, we introduce the operator \( \mathcal{L}: C^2_\mathcal{P}(\mathcal{P}(K)) \to C(\mathcal{P}(K)) \), which is the infinitesimal generator associated to the Kushner-Stratonovich equation. In particular, it is defined by

\[
(\mathcal{L}u)(\mu) = \langle \mu, A\delta u(\mu) \rangle + \frac{1}{2} \langle \mu \otimes \mu, \langle h + B - \langle \mu, h \rangle \rangle \rangle \phi + \frac{1}{2} \langle \mu \otimes \mu, \delta^2 u(\mu) \rangle + \frac{1}{2} \langle \mu \otimes \mu, \delta^2 u(\mu) \rangle \phi + \frac{1}{2} \langle \mu \otimes \mu, \delta^2 u(\mu) \rangle \phi^2 + \langle \mu \otimes \mu, \phi \rangle D_\mu u(\mu) + \langle \mu \otimes \mu, \phi \rangle D_\mu u(\mu) \phi(\mu, h)^2
\]

where \( A, B \) are the operators defined in (2.3) and \( b, \sigma, \bar{\sigma}, h \) are the mappings introduced in Section 2.2. For the second order terms we used the notation

\[
\langle (\mu \otimes \mu, (f_1, f_2) \rangle := \int_{K \times K} f_1(x)^\tau f_2(y) g(x, y) \mu(dx) \mu(dy),
\]

where \( f_1, f_2: K \to \mathbb{R} \) and \( g: K \times K \to \mathbb{R} \).

We can now introduce the backward Kolmogorov equation associated to (2.4) with final condition \( \Phi \in C(\mathcal{P}(K)) \):

\[
\begin{aligned}
&-\partial_t u(\mu, t) - \mathcal{L} u(\mu, t) = 0, \quad (\mu, t) \in \mathcal{P}(K) \times [0, T), \\
&u(\mu, T) = \Phi(\mu), \quad \mu \in \mathcal{P}(K).
\end{aligned}
\]

A first notion of solution is the classical one, in which we ask the solution to be in \( C^2_\mathcal{P}(\mathcal{P}(K)) \) with respect to the measure argument and in \( C^1([0, T]) \) with respect to the time. We will denote this class of functions by \( C^{2,1}_L(\mathcal{P}(K) \times [0, T]) \).

**Definition 3.1.** We say that \( u: \mathcal{P}(K) \times [0, T] \to \mathbb{R} \) is a classical solution to (3.1) if it is in \( C^{2,1}_L(\mathcal{P}(K) \times [0, T]) \), with all the bounds on the derivatives that are uniform, and if it satisfies the backward equation (3.1).

Assuming more regularity on the final condition \( \Phi \), it has been proved in [Mar21] that there exists a unique classical solution to (3.1).

**Theorem 3.2.** Let \( \{\Omega, \Gamma, \{\mathcal{F}_t\}, \mathcal{P}, (\pi^{\mu,s}, I)\} \) be a weak solution to (2.4), starting at time \( s \in [0, T] \) from \( \mu \in \mathcal{P}(K) \) and that satisfies the invariance Hypothesis 2.8. Moreover, let \( \Phi \in C^2_\mathcal{P}(\mathcal{P}(K)) \) and let Hypothesis 2.7 holds. Then, there exists a unique classical solution to (3.1) given by

\[
u(\mu, t) = \mathbb{E} \left[ \Phi(\pi_T^{\mu}) \right], \quad (\mu, t) \in \mathcal{P}(K) \times [0, T].
\]

**Remark 3.3.** In [Mar21] the process \( \pi^{\mu,s} \) is built starting from another measure-valued process, namely a weak solution to the Zakai equation (see for instance [HL01; Szp78] for more details). However, here we avoid to emphasize this to keep a simpler notation.

We want to investigate the well-posedness of (3.1) when the final condition has lower regularity, namely \( \Phi \in C(\mathcal{P}(K)) \). In this case classical solutions may not exist, and a natural way to face the problem is to consider a generalized notion of solution. In particular we will focus on the notion of viscosity solution:

**Definition 3.4.** A upper semicontinuous (resp. lower semicontinuous) function \( u: \mathcal{P}(K) \times [0, T] \to \mathbb{R} \) is a viscosity subsolution (resp. supersolution) to equation (3.1) if:

i. \( u(\mu, T) \leq (\text{resp. } \geq) \Phi(\mu) \), for every \( \mu \in \mathcal{P}(K) \).
ii. for every \((\mu, t) \in \mathcal{P}(K) \times [0, T]\) and any \(\varphi \in C^{2,1}_{\mathcal{L}}(\mathcal{P}(K) \times [0, T])\) such that \(u - \varphi\) has a maximum (resp. minimum) at \((\mu, t)\) with value 0, then (3.1) holds for \(\varphi\) with the inequality \(\leq\) (resp. \(\geq\)) instead of the equality.

We say that \(u\) is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

To conclude this preliminary part, we state an Itô formula for the composition of a \(C^2_{\mathcal{L}}(\mathcal{P}(K))\) function with a weak solution \(\pi\) of the Kushner-Stratonovich equation starting at \(s \in [0, T]\) from \(\mu \in \mathcal{P}(K)\).

**Proposition 3.5.** Let \(\{\mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}\}, (\pi^{\mu,s, t}, I)\) be a weak solution to (3.1) that satisfies the invariance Hypothesis 2.8 and let Hypothesis 2.7 hold. Moreover, let \(u \in C^2_{\mathcal{L}}(\mathcal{P}(K))\). Then, for every \(t \in [s, T]\)

\[
u(\pi_t^{\mu,s}) = u(\mu) + \int_s^t \langle \pi_t^{\mu,s}, A\delta u(\pi_t^{\mu,s}) \rangle \, dt + \int_s^t \langle \pi_t^{\mu,s}, (h + B - \langle \pi_t^{\mu,s}, h \rangle)^\top \delta u(\pi_t^{\mu,s}) \rangle \, dI_t + \frac{1}{2} \int_s^t \langle \pi_t^{\mu,s} \otimes \pi_t^{\mu,s}, (h + B - \langle \pi_t^{\mu,s}, h \rangle)^\top (h + B - \langle \pi_t^{\mu,s}, h \rangle) \delta^2 u(\pi_t^{\mu,s}) \rangle \, dt,
\]

almost surely.

**Remark 3.6.** The results in Theorem 3.2 and in Proposition 3.5 hold more generally when \(\pi\) takes values in \(\mathcal{P}_2(\mathbb{R}^d)\), the space of probability measures over \(\mathbb{R}^d\) with finite second moment.

### 3.1. Existence of viscosity solutions

Let us introduce a function \(u\) by the usual probabilistic representation formula (3.3), where \(\Phi \in C(\mathcal{P}(K))\) is the final condition of (3.1) and \(\pi^{\mu,s}\) is a weak solution to the Kushner-Stratonovich equation starting at \(s \in [0, T]\) from \(\mu \in \mathcal{P}(K)\).

**Proposition 3.7.** Let \(\Phi \in C(\mathcal{P}(K))\), let Hypothesis 2.7 holds and let \(\pi^{\mu,s}\) a solution to (2.4) satisfying the invariance Hypothesis (2.8). Then, the function

\[
u(\mu, t) := \mathbb{E} \left[ \Phi(\pi_t^{\mu,s}) \right], \quad (\mu, t) \in \mathcal{P}(K) \times [0, T],
\]

is a viscosity solution to the backward Kolmogorov equation (3.1).

**Proof.** We only prove that \(u\) is a subsolution, since proving that it is a supersolution is analogue.

First, thanks to the continuity of \(\Phi\) and \(\pi^{\mu,s}\), we notice that \(u\) is continuous and bounded. Moreover \(u(\mu, T) = \mathbb{E} \left[ \Phi(\pi_T^{\mu,T}) \right] = \Phi(\mu)\) for every \(\mu \in \mathcal{P}(K)\).

Now, let us fix \((\mu, t) \in \mathcal{P}(K) \times [0, T]\) and let us consider \(\varphi \in C^{2,1}_{\mathcal{L}}(\mathcal{P}(K) \times [0, T])\) such that \(\varphi(\mu, t) = u(\mu, t)\) and \(\varphi(\nu, s) \geq u(\nu, s)\) for every \((\nu, s) \in \mathcal{P}(K) \times [0, T]\). Then, for \(h > 0\) it holds that

\[
0 \leq \frac{\varphi(\pi_{t+h}^{\mu,t}, t+h) - u(\pi_{t+h}^{\mu,t}, t+h)}{h},
\]

and, by taking the expectation combined with Itô formula and Markov property of \(\pi^{\mu,t}\), we obtain

\[
0 \leq \frac{1}{h} \mathbb{E} \left[ \int_t^{t+h} \mathcal{L} \varphi(\pi_t^{\mu,t}, t+h) \, d\tau \right] + \frac{\varphi(\mu, t+h) - \varphi(\mu, t)}{h}.
\]

Finally, thanks to the regularity of \(\varphi\) and its derivatives, we can pass to the limit as \(h \rightarrow 0\) and obtain

\[
0 \leq \mathcal{L} \varphi(\mu, t) + \partial_t \varphi(\mu, t).
\]

**Remark 3.8.** This existence result is still valid even if we drop Hypothesis 2.8 and we look at the equation over \(\mathcal{P}_2(\mathbb{R}^d)\).
3.2. Comparison theorem and uniqueness. Here we state and prove a comparison principle for viscosity solutions to (3.1), which is a crucial tool to prove the uniqueness. The technique for the proof is inspired by the one used for Theorem II.1 in [Lio83], and by its recent refinement used in [Cos+21]. In particular, we use the solution $u$ introduced by the representation formula (3.3) as a reference to obtain partial comparison results.

**Theorem 3.9.** Let $u_1, u_2 : \mathcal{P}(K) \times [0, T] \to \mathbb{R}$ be respectively a viscosity subsolution and supersolution of (3.1). Moreover, let Hypothesis 2.7 and Hypothesis 2.8 hold. Then, $u_1(\mu, t) \leq u_2(\mu, t)$ for every $(\mu, t) \in \mathcal{P}(K) \times [0, T]$.

**Proof.** We want to prove that $u_1 \leq u \leq u_2$, where $u$ is given by (3.3). The proof is divided in five steps and we show only that $u_1 \leq u$. Indeed we can prove $u \leq u_2$ by noticing that $u_2$ is a subsolution of (3.1) with final condition $-\Phi$ and by taking $-u$ instead of $u$.

By contradiction, let us assume that there exists $(\mu_0, t_0) \in \mathcal{P}(K) \times [0, T]$ such that

$$u_1(\mu_0, t_0) - u(\mu_0, t_0) > 0. \quad (3.4)$$

First, we notice that $t_0 < T$, indeed $u_1(\mu, T) \leq \Phi(\mu) = u(\mu, T)$ for every $\mu \in \mathcal{P}(K)$, thanks to the subsolution property of $u_1$.

**Step 1.** Let $\{\Phi_n\}_{n \in \mathbb{N}} \subset C^2_b(\mathcal{P}(K))$ be the sequence that converges uniformly to $\Phi \in C(\mathcal{P}(K))$ given by Proposition A.5. For every $n \in \mathbb{N}$, thanks to Theorem 3.2 we can set

$$u_n(\mu, t) := \mathbb{E}\left[\Phi_n(\pi^{\mu,t}_n)\right], \quad (\mu, t) \in \mathcal{P}(K) \times [0, T],$$

which is the unique classical solution to equation (3.1) with final condition $\Phi_n$. Since $\{\Phi_n\}_{n \in \mathbb{N}}$ converges uniformly to $\Phi$, the sequence is bounded by a positive constant $M = M(\Phi)$. Then, dominated convergence allows us to conclude that

$$\lim_{n \to +\infty} u_n(\mu, t) = u(\mu, t), \quad (\mu, t) \in \mathcal{P}(K) \times [0, T]. \quad (3.5)$$

**Step 2.** Let us introduce

$$\bar{u}_1(\mu, t) := e^{t-t_0}u_1(\mu, t), \quad \bar{u}_n(\mu, t) := e^{t-t_0}u_n(\mu, t), \quad (\mu, t) \in \mathcal{P}(K) \times [0, T), \quad (3.6)$$

and

$$\bar{\Phi}(\mu) := e^{T-t_0}\Phi(\mu), \quad \bar{\Phi}_n(\mu) := e^{T-t_0}\Phi_n(\mu), \quad \mu \in \mathcal{P}(K).$$

Then, it is easy to see that $\bar{u}_n$ is a classical solution of the equation

$$\begin{cases}
-\partial_t v(\mu, t) - \mathcal{L}v(\mu, t) = -v(\mu, t), & (\mu, t) \in \mathcal{P}(K) \times [0, T), \\
v(\mu, T) = \bar{\Phi}_n(\mu), & \mu \in \mathcal{P}(K),
\end{cases} \quad (3.7)$$

and, analogously, that $\bar{u}_1$ is a viscosity subsolution of

$$\begin{cases}
-\partial_t v(\mu, t) - \mathcal{L}v(\mu, t) = -v(\mu, t), & (\mu, t) \in \mathcal{P}(K) \times [0, T), \\
v(\mu, T) = \bar{\Phi}(\mu), & \mu \in \mathcal{P}(K).
\end{cases}$$

**Step 3.** For every $\lambda > 0$, let us define the mapping $\Psi_\lambda : \mathcal{P}(K) \times [0, T] \to \mathbb{R}$ as

$$\Psi_\lambda(\mu, t) := \bar{u}_1(\mu, t) - \bar{u}_n(\mu, t) - \lambda(t-t_0)^2 - \lambda \mathcal{d}_2(\mu_0, \mu)^2. \quad (3.8)$$

Since $\bar{u}_1$ is upper semicontinuous over a compact space, and $\bar{u}_n$ is continuous and bounded uniformly in $n$, we have that $(\mu_\lambda, t_\lambda) \in \mathcal{P}(K) \times [0, T]$ is a global maximum for $\Psi_\lambda$. Moreover, from (3.8) it follows that

$$\bar{u}_1(\mu_0, t_0) - \bar{u}_n(\mu_0, t_0) = \Psi_\lambda(\mu_0, t_0) \leq \Psi(\mu_\lambda, t_\lambda) \leq \bar{u}_1(\mu_\lambda, t_\lambda) - \bar{u}_n(\mu_\lambda, t_\lambda), \quad (3.9)$$
and so
\[
\lambda(t_\lambda - t_0)^2 + \lambda d_2(\mu_0, \mu_\lambda)^2 \leq (\bar{u}_1(\mu_\lambda, t_\lambda) - \bar{u}_n(\mu_\lambda, t_\lambda)) - (\bar{u}_1(\mu_0, t_0) - \bar{u}_n(\mu_0, t_0)).
\] (3.10)

Thanks to (3.4) and (3.5), for \( n \) large enough we have that \((\bar{u}_1(\mu_0, t_0) - \bar{u}_n(\mu_0, t_0)) > 0\). Combining this with the fact that \(\bar{u}_1\) is bounded from above and \(\bar{u}_n\) is bounded uniformly in \( n \), from (3.10) we obtain that there exists a constant \( M > 0 \) such that
\[
|t_\lambda - t_0| \leq \sqrt{\frac{M}{\lambda}}, \quad d_2(\mu_0, \mu_\lambda) \leq \sqrt{\frac{M}{\lambda}}.
\] (3.11)

**Step 4.** Now we show that \( t_\lambda < T \). By contradiction, let \( t_\lambda = T \). From (3.9),
\[
u_1(\mu_0, t_0) - u_n(\mu_0, t_0) = \bar{u}_1(\mu_0, t_0) - \bar{u}_n(\mu_0, t_0) \leq \bar{u}_1(\mu_\lambda, T) - \bar{u}_n(\mu_\lambda, T),
\]
and then
\[
u_1(\mu_0, t_0) - u_n(\mu_0, t_0) \leq e^{T-t_0} (\Phi(\mu_\lambda) - \Phi_n(\mu_\lambda)).
\] (3.12)

Letting \( n \to \infty \), since \( \Phi_n \to \Phi \) uniformly (see Proposition A.5) and \( u(\mu_0, t_0) \to u_n(\mu_0, t_0) \), we obtain
\[
u_1(\mu_0, t_0) - u(\mu_0, t_0) \leq 0,
\]
which contradicts (3.4).

**Step 5.** Let us define \( \mathcal{P}(K) \times [0, T] \ni (\mu, t) \mapsto \varphi(\mu, t) := \lambda(t - t_0)^2 + \lambda d_2(\mu, \mu_0)^2 \) and let us notice that \( \varphi \in C^{2,1}_L(\mathcal{P}(K) \times [0, T]) \) thanks to Lemma A.1. We want to use \( \bar{u}_n + \varphi \) as test function for the viscosity subsolution property of \( \bar{u}_1 \). First, \( \bar{u}_n + \varphi \) is in \( C^{2,1}_L(\mathcal{P}(K) \times [0, T]) \) since \( \bar{u}_n, \varphi \in C^{2,1}_L(\mathcal{P}(K) \times [0, T]) \). Moreover, from **Step 3** we have that \( \bar{u}_1 - (\bar{u}_n + \varphi) \) has a maximum in \((\mu_\lambda, t_\lambda)\) and from **Step 4** it holds that \( t_\lambda < T \).

Thus, we can actually use \( \bar{u}_n + \varphi \) as a test function in \((\mu_\lambda, t_\lambda)\) and, by the viscosity subsolution property of \( \bar{u}_1 \), obtain
\[
\bar{u}_1(\mu_\lambda, t_\lambda) \leq \partial_t(\bar{u}_n + \varphi)(\mu_\lambda, t_\lambda) + \mathcal{L}(\bar{u}_n + \varphi)(\mu_\lambda, t_\lambda).
\]
Since \( \bar{u}_n \) is a classical solution of (3.7), this entails
\[
\bar{u}_1(\mu_\lambda, t_\lambda) - \bar{u}_n(\mu_\lambda, t_\lambda) \leq \partial_t(\varphi(\mu_\lambda, t_\lambda) + \mathcal{L}(\varphi(\mu_\lambda, t_\lambda)
\]
Now we can use the estimates (3.11) and the bound in Remark A.2 to obtain, for a positive constant \( C > 0 \),
\[
\bar{u}_1(\mu_\lambda, t_\lambda) - \bar{u}_n(\mu_\lambda, t_\lambda) \leq 2\lambda(t_\lambda - t_0) + \lambda(\mathcal{L}(\rho_\mu))(\mu_\lambda) \leq 2\sqrt{\lambda M} + \lambda C,
\]
where \( \rho_\mu \) is defined as in (A.1). Finally, by (3.9) and (3.6), it follows that
\[
u_1(\mu_0, t_0) - u_n(\mu_0, t_0) = \bar{u}_1(\mu_0, t_0) - \bar{u}_n(\mu_0, t_0) \leq 2\sqrt{\lambda M} + \lambda C(\mu_0),
\] (3.13)
which contradicts (3.4) if we let \( \lambda \to 0 \) and \( n \to \infty \).

**Corollary 3.10.** Let \( \Phi \in C(\mathcal{P}(K)) \) and let Hypothesis 2.7 and Hypothesis 2.8 hold. Then
\[
u(\mu, t) = \mathbb{E} \left[ \Phi(\pi_{\tau_\mu}^{\mu,f}) \right], \quad (\mu, t) \in \mathcal{P}(K) \times [0, T],
\]
is the unique viscosity solution to the backward Kolmogorov equation (3.1).

**Proof.** Let \( \nu \) be the solution to (3.1) defined by (3.3) and let \( v \) be another viscosity solution. Since in particular \( \nu \) is a subsolution and \( v \) is a supersolution, Theorem 3.9 tells us that \( \nu \leq v \). Changing the role of \( \nu \) and \( v \) we obtain that \( v \leq \nu \) and so the two solutions coincide. \( \square \)
Appendix A. Some analytical results

Here we collect some analytical tools necessary in the proof of the comparison principle (Theorem 3.9). In particular we need to study the differentiability of the distance $d_2$ introduced in (2.1). Moreover, we also need to introduce a way to approximate continuous functions over $\mathcal{P}(K)$ with functions in $C^2_{L}(\mathcal{P}(K))$.

For a family $\{f_k\}_{k \in \mathbb{N}}$ dense in $C^\infty_0(K)$, with $\|f_k\|_\infty \leq 1$ and containing the constant function equal to one, and for a sequence $\{q_k\}_{k \in \mathbb{N}} \subset [1, +\infty)$, we recall that

$$d_2^2(\mu, \nu) = \sum_{k=1}^{+\infty} \frac{1}{2^k q_k} (\mu - \nu, f_k)^2, \quad \mu, \nu \in \mathcal{P}(K).$$

For what follows, we need to choose

$$q_k := \max \{1, \|D_x f_k\|_\infty, \|D_x^2 f_k\|_\infty, \|D_x f_k\|^2_\infty\}, \quad k \in \mathbb{N}.$$

Let $\mu_0 \in \mathcal{P}(K)$ be fixed and let $\rho_{\mu_0} : \mathcal{P}(K) \to [0, +\infty)$ be defined as

$$\rho_{\mu_0}(\mu) := d_2(\mu, \mu_0)^2, \quad \mu \in \mathcal{P}(K). \quad (A.1)$$

Lemma A.1. Let $\mu_0 \in \mathcal{P}(K)$ be fixed. Then, the mapping $\rho_{\mu_0}$ is in $C^2_{L}(\mathcal{P}(K))$ with derivatives uniformly bounded with respect to $\mu_0$.

Proof. First, let us fix $k \in \mathbb{N}$ and show that the mapping $\mu \mapsto (\mu - \mu_0, f_k)^2 =: g_k(\mu)$, where $f_k$ is in the family appearing in (2.1), is in $C^2_{L}(\mathcal{P}(K))$. By standard computations we get

\[
\delta_\mu g_k(\mu, x) = 2(\mu - \mu_0, f_k) f_k(x), \quad \mu \in \mathcal{P}(K), x \in K,
\]

\[
\delta^2_\mu g_k(\mu, x, y) = 2 f_k(x) f_k(y), \quad \mu \in \mathcal{P}(K), x, y \in K,
\]

with $\|\delta_\mu g_k\|_\infty \leq 4$ and $\|\delta^2_\mu g_k\|_\infty \leq 2$ for every $k \in \mathbb{N}$. Moreover, we have that

\[
D_\mu g_k(\mu, x) = 2(\mu - \mu_0, f_k) D_x f_k(x), \quad \mu \in \mathcal{P}(2)(K), x \in K,
\]

\[
D_x D_\mu g_k(\mu, x) = 2(\mu - \mu_0, f_k) D^2_x f_k(x), \quad \mu \in \mathcal{P}(2)(K), x \in K,
\]

\[
D^2_x g_k(\mu, x, y) = 2D_x f_k(x) D^T_y f_k(y), \quad \mu \in \mathcal{P}(2)(K), x, y \in K,
\]

\[
\delta_\mu D_x g_k(\mu, x, y) = 2D_x f_k(x) f_k(y), \quad \mu \in \mathcal{P}(2)(K), x, y \in K,
\]

with

$$\|D_\mu g_k\|_\infty \leq 4\|D_x f_k\|_\infty, \quad \|D_x D_\mu g_k\|_\infty \leq 4\|D^2_x f_k\|_\infty, \quad \|D^2_x g_k\|_\infty \leq 2\|D_x f_k\|^2_\infty, \quad \|\delta_\mu D_x g_k\|_\infty \leq 2\|D_x f_k\|_\infty.$$

To conclude, by dominated convergence we can bring the derivative inside the series and get

$$\delta_\mu \rho_{\mu_0}(\mu, x) = \sum_{k=1}^{+\infty} \frac{1}{2^k q_k} \delta_\mu g_k(\mu, x), \quad \mu \in \mathcal{P}(K), x \in K,$n

which is also jointly continuous. Moreover, recalling the definition of $\{q_k\}_{k \in \mathbb{N}}$, we have

$$\|\delta_\mu \rho_{\mu_0}\|_\infty \leq \sum_{k=1}^{+\infty} \frac{1}{2^k q_k} \|\delta_\mu g_k\|_\infty \leq 4.$$n

We can proceed in the same way for the other derivatives and the bounds follow easily. \qed
Remark A.2. Let $\mathcal{L}$ be the differential operator introduced in Section 3. Under Hypothesis 2.7, we have that from Lemma A.1 follows that there exists a constant (independent of $\mu_0$) $C > 0$ such that $\|L\rho_\mu\|_\infty = \sup_{\mu \in \mathcal{P}(K)} |(L\rho_\mu)(\mu)| \leq C$. Indeed

$$\|L\rho_\mu\|_\infty \leq \|A\delta_\mu u\|_\infty + \frac{1}{2}\|(h + B - \langle \mu, h \rangle)^\top (h + B - \langle \mu, h \rangle)\delta_\mu^2 u\|_\infty \leq 4\|b\|_\infty + 2\|\sigma\|_{C^2}^2 + 3\|\bar{\sigma}\|_{C^2}^2 + 4\|h\|_{C^2}^2 =: C.$$  (A.3)

To conclude this auxiliary section, we provide a class of polynomials on the space of probability measures that can be used to approximate functions in $C(\mathcal{P}(K))$, which is the set of continuous real-valued functions over $\mathcal{P}(K)$. This kind of approximation has been discussed for instance in [CLSF19].

Definition A.3. We define the set of polynomials over $\mathcal{P}(K)$:

$$\mathcal{P}(\mathcal{P}(K)) := \{\phi: \mathcal{P}_2 \to \mathbb{R} \text{ of the form } \phi(\mu) = f(\langle \mu, \psi_1 \rangle, \ldots, \langle \mu, \psi_n \rangle), \text{ for some } n \in \mathbb{N}\},$$  (A.4)

with $f: \mathbb{R}^n \to \mathbb{R}$ and $\psi_i: \mathbb{R}^d \to \mathbb{R}$ polynomials, $i = 1, \ldots, n$.

Remark A.4. It holds that $\mathcal{P}(\mathcal{P}(K)) \subset C^2_b(\mathcal{P}(K))$. Indeed, every $\phi \in \mathcal{P}(\mathcal{P}(K))$ is of the form discussed in Example 2.5.

One can prove (see for instance [CLSF19, Section 2.3]) that the family $\mathcal{P}(\mathcal{P}(K))$ is an algebra in $\mathcal{P}(K)$ that separates the points and which contains the constants. Thus, by Stone-Weiestrass theorem one can conclude that it is dense in $C(\mathcal{P}(K))$:

Proposition A.5. The family $\mathcal{P}(\mathcal{P}(K))$ is dense in $C^b(\mathcal{P}(K))$ with respect to the supremum norm over $\mathcal{P}(K)$.

Acknowledgements. The author is grateful to Prof. Andrea Cosso for the helpful discussions and suggestions.

References

[ADP90] J.-P. Aubin and G. Da Prato. “Stochastic viability and invariance”. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* Ser. 4, 17.4 (1990), 595–613.

[AG08] L. Ambrosio and W. Gangbo. “Hamiltonian ODEs in the Wasserstein space of probability measures”. *Comm. Pure Appl. Math.* 61.1 (2008), 18–53.

[AGS05] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows In Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhäuser Basel, 2005.

[Ban+19] E. Bandini, A. Cosso, M. Fuhrman, and H. Pham. “Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem”. *Stochastic Processes and their Applications* 129.2 (2019), 674–711.

[BC09] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Stochastic Modelling and Applied Probability 60. Springer-Verlag New York, 2009.

[BCP18] E. Bayraktar, A. Cosso, and H. Pham. “Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics”. *Trans. Amer. Math. Soc.* 370 (2018), 2115–2160.

[Bor95] V. S. Borkar. *Probability Theory: An Advanced Course*. Springer, 1995.

[Bur+20] M. Burzoni, V. Ignazio, M. Reppen, and M. Soner. “Viscosity solutions for controlled McKean–Vlasov jump-diffusions”. *SIAM J. Control Optim.* 58.3 (2020), 1676–1699.

[CCP20] P Cardaliaguet, M Cirant, and A Porretta. “Splitting methods and short time existence for the master equations in mean field games”. *arXiv preprint* (2020).
[CD18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications*. Vol. 1. Probability Theory and Stochastic Modelling 83. Springer International Publishing, 2018.

[CIL92] M. G. Crandall, H. Ishii, and P.-L. Lions. “User’s Guide to Viscosity Solutions of Second-Order Partial Differential Equations”. *Bulletin of the AMS* 27 (1992), 1–67.

[CLSF19] C. Cuchiero, M. Larsson, and S. Svaluto-Ferro. “Probability measure-valued polynomial diffusions”. *Electronic Journal of Probability* 24.30 (2019), 1–32.

[Cos+20] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, and M. Rosestolato. “Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension”. arxiv: 2012.14772 (2020).

[Cos+21] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, and M. Rosestolato. “Master Bellman equation in the Wasserstein space: Uniqueness of viscosity solutions”. arXiv: 2107.10535 (2021).

[CR22] A. Cosso and F. Russo. “Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation”. *Bernoulli* 28.1 (2022), 481–503.

[DPF04] G. Da Prato and H Frankowska. “Invariance of stochastic control systems with deterministic arguments”. *Journal of Differential Equations* 200.1 (2004), 18–52.

[DPF07] G. Da Prato and H Frankowska. “Stochastic viability of convex sets”. *Journal of Mathematical Analysis and Applications* 333.1 (2007), 151–163.

[Fri75] A. Friedman. *Stochastic Differential Equations and Applications*. Vol. 2. Academic Press, 1975.

[GT19] W. Gangbo and A. Tudorascu. “On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations”. *Journal de Mathématiques Pures et Appliquées* 125 (2019), 119–174.

[HL01] A. J. Heunis and V. M. Lucic. “On uniqueness of solutions for the stochastic differential equations of nonlinear filtering”. *Ann. Appl. Probab.* 11.1 (2001), 182–209.

[Lio07] P.-L. Lions. *Cours au Collège de France*. 2007-2013.

[Lio83] P.-L. Lions. “Optimal control of diffusion processes and hamilton–jacobi–bellman equations part 2 : viscosity solutions and uniqueness”. *Communications in Partial Differential Equations* 8.11 (1983), 1229–1276.

[Mar21] M. Martini. “Kolmogorov equations on spaces of measures associated to nonlinear filtering processes”. arXiv: 2107.11865 (2021).

[PW17] H. Pham and X. Wei. “Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics”. *SIAM J. Control Optim.* 55.2 (2017), 1069–1101.

[PW18] H. Pham and X. Wei. “Bellman equation and viscosity solutions for mean-field stochastic control problem”. *ESAIM Control Optim. Calc. Var.* 24.1 (2018), 437–461.

[Roz90] B. L. Rozovsky. *Stochastic Evolution Systems*. Mathematics and its Applications 35. Springer Netherlands, 1990.

[Swi94] A. Swiech. “Unbounded second order partial differential equations in infinite dimensional Hilbert spaces”. *Communications in Partial Differential Equations* 19 (1994), 1999–2036.

[Szp78] J. Szpirglas. “Sur l’équivalence d’équations différentielles stochastiques à valeurs mesures intervenant dans le filtrage Markovien non linéaire”. *Ann. Inst. H. Poincaré Sect. B (N.S.)* 14.1 (1978), 33–59.

[Vil08] C. Villani. *Optimal transport - Old and new*. 338. Springer, 2008.

[WZ20] C. Wu and J. Zhang. “Viscosity solutions to parabolic master equations and McKean–Vlasov SDEs with closed-loop controls”. *The Annals of Applied Probability* 30.2 (2020), 936–986.

[Xio08] J. Xiong. *An Introduction to Stochastic Filtering Theory*. Oxford University Press, 2008.
Mattia Martini: Dipartimento di Matematica "Federigo Enriques", Università degli Studi di Milano, Via Saldini 50, 20133 Milano (Italy)

Email address: mattia.martini@unimi.it