HAL
open science

Bellman equation and viscosity solutions for mean-field stochastic control problem

Huyên Pham, Xiaoli Wei

To cite this version:

Huyên Pham, Xiaoli Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. 2015. hal-01248317v1

HAL Id: hal-01248317

https://hal.science/hal-01248317v1

Preprint submitted on 24 Dec 2015 (v1), last revised 7 Mar 2017 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bellman equation and viscosity solutions for mean-field stochastic control problem

Huyên PHAM*
Laboratoire de Probabilités et
Modèles Aléatoires, CNRS, UMR 7599
Université Paris Diderot
pham at math.univ-paris-diderot.fr
and CREST-ENSAE

Xiaoli WEI
Laboratoire de Probabilités et
Modèles Aléatoires, CNRS, UMR 7599
Université Paris Diderot
tyswxl at gmail.com

December 24, 2015

Abstract

We consider the stochastic optimal control problem of McKean-Vlasov stochastic differential equation. By using feedback controls, we reformulate the problem into a deterministic control problem with only the marginal distribution as controlled state variable, and prove that dynamic programming principle holds in its general form. Then, by relying on the notion of differentiability with respect to probability measures recently introduced by P.L. Lions in [30], and a special Itô formula for flows of probability measures, we derive the (dynamic programming) Bellman equation for mean-field stochastic control problem. This Bellman equation in the Wassertein space of probability measures reduces to the classical finite dimensional partial differential equation in the case of no mean-field interaction. We prove a verification theorem in our McKean-Vlasov framework, and give explicit solutions to the Bellman equation for the linear quadratic mean-field control problem, with applications to the mean-variance portfolio selection and a systemic risk model. Finally, we consider a notion of lifted viscosity solutions for the Bellman equation, and show the viscosity property and uniqueness of the value function to the McKean-Vlasov control problem.

MSC Classification: 93E20, 60H30, 60K35.

Keywords: McKean-Vlasov SDEs, dynamic programming, Bellman equation, Wasserstein space, viscosity solutions.

*Research supported in part by FiME (Finance for Energy Market Research Centre) and the "Finance et Développement Durable - Approches Quantitatives" EDF - CACIB Chair.

1 Introduction

The problem studied in this paper concerns the optimal control of mean-field stochastic differential equations (SDEs), also known as McKean-Vlasov equations. This topic is closely related to the mean-field game (MFG) problem as originally formulated by Lasry and Lions in [24] and simultaneously by Huang, Caines and Malhamé in [21]. It aims at describing equilibrium states of large population of symmetric players (particles) with mutual interactions of mean-field type, and we refer to [12] for a discussion pointing out the subtle differences between the notions of equilibrium in MFG and optimal control of McKean-Vlasov dynamics.

While the analysis of McKean-Vlasov SDEs has a long history with the pioneering works by Kac [23] and H. McKean [31], and later on with papers in the general framework of propagation of chaos, see e.g. [35], [22], the optimal control of McKean-Vlasov dynamics is a rather new problem, which attracts an increasing interest since the emergence of the MFG theory and its numerous applications in several areas outside physics, like economics and finance, biology, social interactions, networks. Actually, it has been first studied in [1] by functional analysis method with a value function expressed in terms of the Nisio semigroup of operators. More recently, several papers have adopted the stochastic maximum (also called Pontryagin) principle for characterizing solutions to the controlled McKean-Vlasov systems in terms of an adjoint backward stochastic differential equation (BSDE) coupled with a forward SDE: see [3], [7], [37] with a state dynamics depending upon moments of the distribution, and [10] for a deep investigation in a more general setting. Alternatively, and although the dynamics of mean-field SDEs is non-Markovian, it is tempting to use dynamic programming (DP) method (also called Bellman principle), which is known to be a powerful tool for standard Markovian stochastic control problem, see e.g. [18], [32], and does not require any convexity assumption usually imposed in Pontryagin principle. Indeed, mean-field type control problem was tackled by DP in [25] and [4] for specific McKean-Vlasov SDE and cost functional, typically depending only upon statistics like its mean value or with uncontrolled diffusion coefficient, and especially by assuming the existence at all times of a density for the marginal distribution of the state process. The key idea in both papers [25] and [4] is to reformulate the stochastic control problem with feedback strategy as a deterministic control problem involving the density of the marginal distribution, and then to derive a dynamic programming equation in the space of density functions.

Inspired by the works [4] and [25], the objective of this paper is to analyze in detail the dynamic programming method for the optimal control of mean-field SDEs where the drift and diffusion coefficients may depend both upon the marginal distribution of the state and of the control. This additional dependence related to the mean-field interaction on control is natural in the context of McKean-Vlasov control problem, but has been few considered in the literature, see however [37] for a dependence only through the moments of the control. Our paper can be viewed as the continuous time version of the discrete time problem studied recently in [33]. By using closed-loop (also called feedback) controls, we first convert the stochastic optimal control problem into a deterministic control problem where the marginal distribution is the sole controlled state variable, and we prove that

dynamic programming holds in its general form. The next step for exploiting the DP is to differentiate functions defined on the space of probability measures. There are various notions of derivatives with respect to measures which have been developed in connection with the theory of optimal transport and using Wasserstein metric on the space of probability measures, see e.g. the monographs [2], [36]. For our purpose, we shall use the notion of differentiability introduced by P.L. Lions in his lectures at the Collège de France [30], see also the helpful redacted notes [9]. This notion of derivative is based on the lifting of functions defined on the space of square integrable probability measures into functions defined on the Hilbert space of square integrable random variables distributed according to the "lifted" probability measure. It has been used in [10] for differentiating the Hamiltonian function appearing in stochastic Pontryagin principle for controlled McKean-Vlasov dynamics. As usual in continuous time control problem, we need a dynamic differential calculus for deriving the infinitesimal version of the DP, and shall rely on a special Itô's chain rule for flows of probability measures as recently developed in [8] and [14], and used in [11] for deriving the so-called Master equation in MFG. We are then able to derive the dynamic programming Bellman equation for mean-field stochastic control problem. This infinite dimensional fully nonlinear partial differential equation (PDE) of second order in the Wassertein space of probability measures seems to be new in the literature: it reduces to the Bellman equation in the space of density functions derived by Bensoussan, Frehse and Yam [5] when the marginal distribution admits a density, and on the other hand, we notice that it differs from the Master equation for McKean-Vlasov control problem obtained by Carmona and Delarue in [11] where the value function is a function of both the state and its marginal distribution, and so with associated PDE in the state space comprising probability measures but also Euclidian vectors. Following the traditional approach for stochastic control problem, we prove a verification theorem for the Bellman equation of the McKean-Vlasov control problem, which reduces to the classical Bellman equation in the case of no mean-field interaction. We apply our verification theorem to the important class of linear quadratic (LQ) McKean-Vlasov control problems where it turns out that derivations in the space of probability measures are quite tractable and lead to explicit classical solutions for the Bellman equation. We illustrate these results with two examples arising from finance: the mean-variance portfolio selection and an inter-bank systemic risk model. Finally, since there are in general no classical solutions to the Bellman equation, we introduce a notion of viscosity solutions for the Bellman equation in the Wasserstein space of probability measures. There are several definitions of viscosity solutions for Hamilton Jacobi equations of first order in Wasserstein space and more generally in metric spaces, see e.g. [2], [19], [17] or [20]. We adopt the approach in [30], and detailed in [9], which consists, after the lifting identification between measures and random variables, in working in the Hilbert space of square integrable random variables instead of working in the Wasserstein space of probability measures, in order to use the various tools developed for viscosity solutions in separable Hilbert spaces, in particular in our context, for second order Hamilton-Jacobi equations, see [27], [28], [29], and the recent monograph [16]. We then prove the viscosity property of the value function and a comparison principle, hence uniqueness result, for our Bellman equation associated to the McKean-Vlasov control problem.

The rest of the paper is organized as follows. Section 2 describes the McKean-Vlasov control problem and fix the standing assumptions. In Section 3, we state the dynamic programming principle after the reformulation into a deterministic control problem. Section 4 is devoted to the derivation of the Bellman equation and the proof of the verification theorem. We present in Section 5 the applications to the LQ framework where explicit solutions are provided with two examples arising from financial models. The last part of this paper in Section 6 deals with viscosity solutions for the Bellman equation.

2 McKean-Vlasov control problem

Let us fix some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a n-dimensional Brownian motion $B = (B_t)_{0\le t\le T}$, and denote by $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ its natural filtration, augmented with an independent σ-algebra $\mathcal{F}_0 \subset \mathcal{F}$. For each random variable X, we denote by $\mathbb{P}_X$ its probability law (also called distribution) under $\mathbb{P}$ (which is deterministic), and by δ_X the Dirac measure on X. Given a normed space $(E, |.|)$, we denote by $\mathcal{P}_2(E)$ the set of probability measures μ on E, which are square integrable, i.e. $\|\mu\|_2^2 := \int_E |x|^2\mu(dx) < \infty$, and by $L^2(\mathcal{F}_0; E)$ $(= L^2(\Omega, \mathcal{F}_0, \mathbb{P}; E))$ the set of square integrable random variables on $(\Omega, \mathcal{F}_0, \mathbb{P})$. In the sequel, E will be either $\mathbb{R}^d$, the state space, or A, the control space, a subset of $\mathbb{R}^m$. We shall assume without loss of generality (see Remark 2.1 below) that $\mathcal{F}_0$ is rich enough to carry E-valued random variables with any arbitrary square integrable distribution, i.e. $\mathcal{P}_2(E) = \{\mathbb{P}_\xi, \xi \in L^2(\mathcal{F}_0; E)\}$.

Remark 2.1 A possible construction of a probability space, which is rich enough to satisfy the above conditions is the following. We consider a Polish space Ω_0, its Borel σ-algebra $\mathcal{F}_0$ and let $\mathbb{P}_0$ be an atomless probability measure on $(\Omega_0, \mathcal{F}_0)$. We consider another probability space $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ supporting a n-dimensional Brownian motion B and denote by $\mathbb{F}^B = (\mathcal{F}_t^B)$ its natural filtration. By defining $\Omega = \Omega_0 \times \Omega_1$, $\mathcal{F} = \mathcal{F}_0 \vee \mathcal{F}_1$, $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1$, and $\mathbb{F} = (\mathcal{F}_t)$ with $\mathcal{F}_t = \mathcal{F}_t^B \vee \mathcal{F}_0$, $0 \le t \le T$, we then obtain that the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfies the required condition in the above framework. □

We also denote by W_2 the 2-Wasserstein distance defined on $\mathcal{P}_2(E)$ by

$$\begin{aligned} W_2(\mu, \mu') &:= \inf\Big\{\Big(\int_{E\times E} |x-y|^2\pi(dx,dy)\Big)^{\frac{1}{2}} : \pi \in \mathcal{P}_2(E\times E) \text{ with marginals } \mu \text{ and } \mu'\Big\} \\ &= \inf\Big\{\Big(\mathbb{E}|\xi-\xi'|^2\Big)^{\frac{1}{2}} : \ \xi, \xi' \in L^2(\mathcal{F}_0; E) \text{ with } \mathbb{P}_\xi = \mu,\ \mathbb{P}_{\xi'} = \mu'\Big\}. \end{aligned}$$

We consider a controlled stochastic dynamics of McKean-Vlasov type for the process $X^\alpha = (X_t^\alpha)_{0\le t\le T}$ valued in $\mathbb{R}^d$:

$$dX_t^\alpha = b(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, \alpha_t, \mathbb{P}_{\alpha_t})dt + \sigma(t, X_t^\alpha, \mathbb{P}_{X_t^\alpha}, \alpha_t, \mathbb{P}_{\alpha_t})dB_t, \ \ X_0^\alpha = X_0, \qquad (2.1)$$

where $X_0 \in L^2(\mathcal{F}_0, \mathbb{R}^d)$, and the control process $\alpha = (\alpha_t)_{0\le t\le T}$ is progressively measurable with values in a subset A of $\mathbb{R}^m$. The coefficients b and σ are deterministic measurable functions from $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \times \mathcal{P}_2(A)$ into $\mathbb{R}^d$ and $\mathbb{R}^{d\times n}$ respectively. Notice here

that the drift and diffusion coefficients b, σ of the controlled state process do not depend only on the marginal distribution of the state process X_t at time t but also on the marginal distribution of the control α_t at time t, which represents an additional mean-field feature with respect to classical McKean-Vlasov equations. We make the following assumption:

(H1) There exists some constant $C_{b,\sigma} > 0$ s.t. for all $t \in [0,T]$, $x, x' \in \mathbb{R}^d$, $a, a' \in A$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $\lambda, \lambda' \in \mathcal{P}_2(A)$,

$$\begin{aligned} & |b(t,x,\mu,a,\lambda) - b(t,x',\mu',a',\lambda')| + |\sigma(t,x,\mu,a,\lambda) - \sigma(t,x',\mu',a',\lambda')| \\ \leq \quad & C_{b,\sigma}\big[|x-x'| + |a-a'| + W_2(\mu,\mu') + W_2(\lambda,\lambda')\big], \end{aligned}$$

and

$$\int_0^T |b(t,0,\delta_0,0,\delta_0)|^2 + |\sigma(t,0,\delta_0,0,\delta_0)|^2 dt \quad < \quad \infty.$$

Condition **(H1)** ensures that for any control process α, which is square integrable, i.e. $\mathbb{E}[\int_0^T |\alpha_t|^2 dt] < \infty$, there exists a unique solution X^α to (2.1), and moreover this solution satisfies (see e.g. [35] or [22]):

$$\mathbb{E}\big[\sup_{0\leq t\leq T} |X_t^\alpha|^2\big] \quad \leq \quad C\Big(1 + \mathbb{E}|X_0|^2 + \mathbb{E}\big[\int_0^T |\alpha_t|^2 dt\big]\Big) \; < \; \infty. \tag{2.2}$$

In the sequel of the paper, we stress the dependence of X^α on α if needed, but most often, we shall omit this dependence and simply write $X = X^\alpha$ when there is no ambiguity.

The cost functional associated to the McKean-Vlasov equation (2.1) is

$$J(\alpha) \quad := \quad \mathbb{E}\Big[\int_0^T f(t, X_t, \mathbb{P}_{X_t}, \alpha_t, \mathbb{P}_{\alpha_t})dt + g(X_T, \mathbb{P}_{X_T})\Big] \tag{2.3}$$

for a square integrable control process α. The running cost function f is a deterministic real-valued function on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \times \mathcal{P}_2(A)$ and the terminal gain function g is a deterministic real-valued function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. We shall assume the following quadratic condition on f, g:

(H2) There exists some constant $C_{f,g} > 0$ s.t. for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $a \in A$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\lambda \in \mathcal{P}_2(A)$,

$$|f(t,x,\mu,a,\lambda)| + |g(x,\mu)| \quad \leq \quad C_{f,g}\big(1 + |x|^2 + |a|^2 + \|\mu\|_2^2 + \|\lambda\|_2^2\big).$$

Under Condition **(H2)**, and from (2.2), we see that $J(\alpha)$ is well-defined and finite for any square integrable control process α. The stochastic control problem of interest in this paper is to minimize the cost functional:

$$V_0 \quad := \quad \inf_{\alpha\in\mathcal{A}} J(\alpha), \tag{2.4}$$

over a set of admissible controls $\mathcal{A}$ to be precised later.

Notations: We denote by $x.y$ the scalar product of two Euclidian vectors x and y, and by $M^\intercal$ the transpose of a matrix or vector M. For any $\mu \in \mathcal{P}_2(E)$, F Euclidian space, we denote by $L^2_\mu(F)$ the set of measurable functions $\varphi : E \to F$ which are square integrable with respect to μ, and we set

$$< \varphi, \mu > \quad := \quad \int_E \varphi(x)\mu(dx).$$

We also denote by $L^\infty_\mu(F)$ the set of measurable functions $\varphi : E \to F$ which are bounded μ a.e., and $\|\varphi\|_\infty$ denotes the essential supremum of $\varphi \in L^\infty_\mu(F)$.

3 Dynamic programming

In this section, we make the standing assumptions **(H1)-(H2)**, and our purpose is to show that dynamic programming principle holds for problem (2.4), which we would like to combine with some Markov property of the controlled state process. However, notice that the McKean-Vlasov type dependence on the dynamics of the state process rules out the standard Markov property of the controlled process $(X_t)_t$. Actually, this Markov property can be restored by considering its probability law $(\mathbb{P}_{X_t})_t$. To be more precise and for the sake of definiteness, we shall restrict ourselves to controls $\alpha = (\alpha_t)_{0\le t\le T}$ given in closed loop (or feedback) form:

$$\alpha_t \quad = \quad \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}), \qquad 0 \le t \le T, \tag{3.1}$$

for some deterministic measurable function $\tilde{\alpha}(t, x, \mu)$ defined on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. We denote by $Lip([0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); A)$ the set of deterministic measurable functions $\tilde{\alpha}$ on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, valued in A, which are Lipschitz in (x, μ), and satisfy a linear growth condition on (x, μ), uniformly on $t \in [0, T]$, i.e. there exists some positive constant $C_{\tilde{\alpha}}$ s.t. for all $t \in [0, T]$, $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\begin{aligned} |\tilde{\alpha}(t, x, \mu) - \tilde{\alpha}(t, x', \mu')| \quad &\le \quad C_{\tilde{\alpha}}\big(|x - x'| + W_2(\mu, \mu')\big), \\ \int_0^T |\tilde{\alpha}(t, 0, \delta_0)|^2 dt \quad &< \quad \infty. \end{aligned}$$

Notice that for any $\tilde{\alpha} \in Lip([0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); A)$, and under the Lipschitz condition in **(H1)**, there exists a unique solution to the SDE:

$$\begin{aligned} dX_t \quad = \quad & b(t, X_t, \mathbb{P}_{X_t}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}), \mathbb{P}_{\tilde{\alpha}(t, X_t, \mathbb{P}_{X_t})})dt \\ & + \sigma(t, X_t, \mathbb{P}_{X_t}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}), \mathbb{P}_{\tilde{\alpha}(t, X_t, \mathbb{P}_{X_t})})dB_t, \end{aligned} \tag{3.2}$$

starting from some square integrable random variable, and this solution satisfies the square integrability condition (2.2). The set $\mathcal{A}$ of so-called admissible controls α is then defined as the set of control processes α of feedback form (3.1) with $\tilde{\alpha} \in Lip([0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); A)$. We shall often identify $\alpha \in \mathcal{A}$ with $\tilde{\alpha}$ in $Lip([0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d); A)$ via (3.1), and we see that any α in $\mathcal{A}$ is square-integrable: $\mathbb{E}[\int_0^T |\alpha_t|^2 dt] < \infty$, by (2.2) and Gronwall's lemma.

Let us now check the flow property of the marginal distribution process $\mathbb{P}_{X_t} = \mathbb{P}_{X_t^\alpha}$ for any admissible control α in $\mathcal{A}$. First, observe that the marginal distribution $\mathbb{P}_{\alpha_t}$ of the feedback control process $\alpha \in \mathcal{A}$ is equal to the image by $\tilde{\alpha}(t,.,\mathbb{P}_{X_t})$ of the marginal distribution $\mathbb{P}_{X_t}$ of the controlled state process X, i.e. $\mathbb{P}_{\alpha_t} = \tilde{\alpha}(t,.,\mathbb{P}_{X_t}) \star \mathbb{P}_{X_t}$, where we set for any $\tilde{\alpha} \in L(\mathbb{R}^d; A)$, the set of measurable functions from $\mathbb{R}^d$ into A with linear growth condition, and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$(\tilde{\alpha} \star \mu)(B) \;=\; \mu\big(\tilde{\alpha}^{-1}(B)\big), \quad \forall B \in \mathcal{B}(A).$$

We consider the dynamic version of (3.2) starting at time $t \in [0,T]$ from $\xi \in L^2(\mathcal{F}_t; \mathbb{R}^d)$, which is then written as:

$$\begin{aligned} X_s^{t,\xi} \;=\; \xi \;+\; & \int_t^s b(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}}, \tilde{\alpha}(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}}), \tilde{\alpha}(r,.,\mathbb{P}_{X_r^{t,\xi}}) \star \mathbb{P}_{X_r^{t,\xi}})dr \\ & + \int_t^s \sigma(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}}, \tilde{\alpha}(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}}), \tilde{\alpha}(r,.,\mathbb{P}_{X_r^{t,\xi}}) \star \mathbb{P}_{X_r^{t,\xi}})dB_r, \quad t \le s \le T. \end{aligned} \tag{3.3}$$

Existence and uniqueness of a solution to (3.3) implies the flow property:

$$X_s^{t,\xi} \;=\; X_s^{\theta, X_\theta^{t,\xi}}, \qquad \forall\, 0 \le t \le \theta \le s \le T,\; \xi \in L^2(\mathcal{F}_t; \mathbb{R}^d). \tag{3.4}$$

Moreover, as pointed out in Remark 3.1 in [8] (see also the remark following (2.3) in [14]), the solution to (3.3) is also unique in law from which it follows that the law of $X^{t,\xi}$ depends on ξ only through its law $\mathbb{P}_\xi$. Therefore, we can define

$$\mathbb{P}_s^{t,\mu} \;:=\; \mathbb{P}_{X_s^{t,\xi}}, \quad \text{for} \quad 0 \le t \le s \le T,\;\; \mu \;=\; \mathbb{P}_\xi \;\in\; \mathcal{P}_2(\mathbb{R}^d), \tag{3.5}$$

As a consequence of the flow property (3.4), and recalling that $\mathcal{P}_2(\mathbb{R}^d) = \{\mathbb{P}_\xi, \xi \in L^2(\mathcal{F}_0; \mathbb{R}^d)\}$, it is clear that we also get the flow property for the marginal distribution process:

$$\mathbb{P}_s^{t,\mu} \;=\; \mathbb{P}_s^{\theta, \mathbb{P}_\theta^{t,\mu}}, \qquad \forall\, 0 \le t \le \theta \le s \le T,\; \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.6}$$

Recall that the process $X^{t,\xi}$, hence also the law process $\mathbb{P}^{t,\mu}$ depends on the feedback control $\alpha \in \mathcal{A}$, and if needed, we shall stress the dependence on α by writing $\mathbb{P}^{t,\mu,\alpha}$.

We next show that the initial stochastic control problem can be reduced to a deterministic control problem. Indeed, by definition of the marginal distribution $\mathbb{P}_{X_t}$, and Fubini's theorem, we see that the cost functional can be written for any admissible control $\alpha \in \mathcal{A}$ as:

$$J(\alpha) \;=\; \int_0^T \hat{f}(t, \mathbb{P}_{X_t}, \tilde{\alpha}(t,.,\mathbb{P}_{X_t}))dt + \hat{g}(\mathbb{P}_{X_T}),$$

where the function $\hat{f}$ is defined on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times L(\mathbb{R}^d; A)$ and $\hat{g}$ is defined on $\mathcal{P}_2(\mathbb{R}^d)$ by

$$\hat{f}(t,\mu,\tilde{\alpha}) \;:=\; < f(t,.,\mu,\tilde{\alpha}(.), \tilde{\alpha} \star \mu), \mu >, \qquad \hat{g}(\mu) \;:=\; < g(.,\mu), \mu > . \tag{3.7}$$

We have thus transformed the initial control problem (2.4) into a deterministic control problem involving the infinite dimensional controlled marginal distribution process valued in $\mathcal{P}_2(\mathbb{R}^d)$. In view of the flow property (3.6), it is then natural to define the value function

$$v(t,\mu) \;:=\; \inf_{\alpha \in \mathcal{A}} \Big[\int_t^T \hat{f}(s, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}))ds + \hat{g}(\mathbb{P}_T^{t,\mu}) \Big], \;\; t \in [0,T],\; \mu \in \mathcal{P}_2(\mathbb{R}^d), \tag{3.8}$$

so that the initial control problem in (2.4) is given by: $V_0 = v(0, \mathbb{P}_{X_0})$. It is clear that $v(t,\mu) < \infty$, and we shall assume that

$$v(t,\mu) \quad > \quad -\infty, \qquad \forall\, t \in [0,T],\ \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.9}$$

Remark 3.1 The finiteness condition (3.9) can be checked a priori directly from the assumptions on the model. For example, when f, g, hence $\hat{f}$, $\hat{g}$, are lower-bounded functions, condition (3.9) clearly holds. Another example is the case when $f(t,x,\mu,a,\lambda)$, and $g(x,\mu)$ are lower bounded by a quadratic function in x, μ, and λ (uniformly in (t,a)) so that

$$\hat{f}(t,\mu,\tilde{\alpha}) + \hat{g}(x,\mu) \quad \geq \quad -C\big(1 + \|\mu\|_2\big), \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\ \tilde{\alpha} \in L(\mathbb{R}^d; A),$$

and we are able to derive moment estimates on the controlled process X, uniformly in α: $\big\|\mathbb{P}_s^{t,\mu}\big\|_2^2 = \mathbb{E}[|X_s^{t,\xi}|^2] \leq C(1 + \|\mu\|_2^2)$, (for $\mu = \mathbb{P}_\xi$) which arises typically from (2.2) when A is bounded. Then, it is clear that (3.9) holds true. Otherwise, this finiteness condition can be checked a posteriori from a verification theorem, see Theorem 4.1. □

The dynamic programming principle (DPP) for the deterministic control problem (3.8) takes the following formulation:

Theorem 3.1 *(Dynamic Programming Principle)*
Under (3.9)*, we have for all* $0 \leq t \leq \theta \leq T$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*:*

$$v(t,\mu) \quad = \quad \inf_{\alpha \in \mathcal{A}} \Big[\int_t^\theta \hat{f}(s, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}))ds \ + \ v(\theta, \mathbb{P}_\theta^{t,\mu}) \Big]. \tag{3.10}$$

Proof. In the context of deterministic control problem, the proof of the DPP is elementary and does not require any measurable selection arguments. For sake of completeness, we provide it. Denote by $J(t,\mu,\alpha)$ the cost functional:

$$J(t,\mu,\alpha) \quad := \quad \int_t^T \hat{f}(s, \mathbb{P}_s^{t,\mu,\alpha}, \tilde{\alpha}(s,.)) + \hat{g}(\mathbb{P}_T^{t,\mu,\alpha}), \quad 0 \leq t \leq T,\ \mu \in \mathcal{P}_2(\mathbb{R}^d), \alpha \in \mathcal{A},$$

so that $v(t,\mu) = \inf_{\alpha \in \mathcal{A}} J(t,x,\alpha)$, and by $w(t,\mu)$ the r.h.s. of (3.10) (here we stress the dependence of the controlled marginal distribution process $\mathbb{P}^{t,\mu,\alpha}$ on α). Then,

$$\begin{aligned}
w(t,\mu) \quad &= \quad \inf_{\alpha \in \mathcal{A}} \Big[\int_t^\theta \hat{f}(s, \mathbb{P}_s^{t,\mu,\alpha}, \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu,\alpha}))ds + \inf_{\beta \in \mathcal{A}} J(\theta, P_\theta^{t,\mu,\alpha}, \beta) \Big] \\
&= \quad \inf_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{A}} \Big[\int_t^\theta \hat{f}(s, \mathbb{P}_s^{t,\mu,\alpha}, \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu,\alpha}))ds + J(\theta, P_\theta^{t,\mu,\alpha}, \beta) \Big] \\
&= \quad \inf_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{A}} \Big[\int_t^\theta \hat{f}(s, \mathbb{P}_s^{t,\mu,\gamma[\alpha,\beta]}, \tilde{\gamma}[\alpha,\beta](s,.,\mathbb{P}_s^{t,\mu,\gamma[\alpha,\beta]}))ds + J(\theta, P_\theta^{t,\mu,\gamma[\alpha,\beta]}, \gamma[\alpha,\beta]) \Big]
\end{aligned}$$

where we define $\gamma[\alpha,\beta] \in \mathcal{A}$ by: $\tilde{\gamma}[\alpha,\beta](s,.) = \tilde{\alpha}(s,.)1_{0 \leq s \leq \theta} + \tilde{\beta}(s,.)1_{\theta < s \leq T}$. Now, it is clear

that when α, β run over $\mathcal{A}$, then $\gamma[\alpha,\beta]$ also runs over $\mathcal{A}$, and so:

$$\begin{aligned} w(t,\mu) &= \inf_{\gamma\in\mathcal{A}} \big[\int_t^\theta \hat{f}(s,\mathbb{P}_s^{t,\mu,\gamma},\tilde{\gamma}(s,.,\mathbb{P}_s^{t,\mu,\gamma}))ds + J(\theta,P_\theta^{t,\mu,\gamma},\gamma)\big] \\ &= \inf_{\gamma\in\mathcal{A}} \big[\int_t^\theta \hat{f}(s,\mathbb{P}_s^{t,\mu},\tilde{\gamma}(s,.,\mathbb{P}_s^{t,\mu}))ds + \int_\theta^T \hat{f}(s,\mathbb{P}_s^{\theta,\mathbb{P}_\theta^{t,\mu}},\tilde{\gamma}(s,.,\mathbb{P}_s^{\theta,\mathbb{P}_\theta^{t,\mu}})) + \hat{g}(\mathbb{P}_T^{\theta,\mathbb{P}_\theta^{t,\mu}})\big] \\ &= \inf_{\gamma\in\mathcal{A}} \big[\int_t^\theta \hat{f}(s,\mathbb{P}_s^{t,\mu},\tilde{\gamma}(s,.,\mathbb{P}_s^{t,\mu}))ds + \int_\theta^T \hat{f}(s,\mathbb{P}_s^{t,\mu},\tilde{\gamma}(s,.,\mathbb{P}_s^{t,\mu})) + \hat{g}(\mathbb{P}_T^{t,\mu})\big], \end{aligned}$$

by the flow property (3.6) (here we have omitted in the second and third line the dependence of $\mathbb{P}_s$ in γ). This proves the required equality: $w(t,\mu) = v(t,\mu)$. □

Remark 3.2 Problem (2.4) includes the case where the cost functional in (2.3) is a non-linear function of the expected value of the state process, i.e. the running cost functions and the terminal gain function are in the form: $f(t,X_t,\mathbb{P}_{X_t},\alpha_t) = \bar{f}(t,X_t,\mathbb{E}[X_t],\alpha_t)$, $t \in [0,T]$, $g(X_T,\mathbb{P}_{X_T}) = \bar{g}(X_T,\mathbb{E}[X_T])$, which arises for example in mean-variance problem (see Section 5). It is claimed in [6] and [37] that Bellman optimality principle does not hold, and therefore the problem is time-inconsistent. This is correct when one takes into account only the state process X (that is its realization), since it is not Markovian, but as shown in this section, dynamic programming principle holds true whenever we consider the marginal distribution as state variable. This gives more information and the price to paid is the infinite-dimensional feature of the marginal distribution state variable. □

4 Bellman equation

The purpose of this section is to derive from the dynamic programming principle (3.10), a partial differential equation (PDE) for the value function $v(t,\mu)$, called Bellman equation. We shall rely on the notion of derivative with respect to a probability measure, as introduced by P.L. Lions in his course at Collège de France, and detailed in the lecture notes [9].

This notion is based on the lifting of functions $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ into functions U defined on $L^2(\mathcal{F}_0;\mathbb{R}^d)$ by $U(X) = u(\mathbb{P}_X)$. We say that u is differentiable (resp. $\mathcal{C}^1$) on $\mathcal{P}_2(\mathbb{R}^d)$ if the lift U is Fréchet differentiable (resp. Fréchet differentiable with continuous derivatives) on $L^2(\mathcal{F}_0;\mathbb{R}^d)$. In this case, the Fréchet derivative $[DU](X)$, viewed as an element $DU(X)$ of $L^2(\mathcal{F}_0;\mathbb{R}^d)$ by Riesz' theorem: $[DU](X)(Y) = \mathbb{E}[DU(X).Y]$, can be represented as

$$DU(X) = \partial_\mu u(\mathbb{P}_X)(X), \tag{4.1}$$

for some function $\partial_\mu u(\mathbb{P}_X) : \mathbb{R}^d \to \mathbb{R}^d$, which is called derivative of u at $\mu = \mathbb{P}_X$. Moreover, $\partial_\mu u(\mu) \in L^2_\mu(\mathbb{R}^d)$ for $\mu \in \mathcal{P}_2(\mathbb{R}^d) = \{\mathbb{P}_X, X \in L^2(\mathcal{F}_0;\mathbb{R}^d)\}$. Following [14], we say that u is partially $\mathcal{C}^2$ if it is $\mathcal{C}^1$, the mapping $(\mu,x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_\mu u(\mu)(x)$ is continuous at any point (μ,x) such that $x \in \mathrm{Supp}(\mu)$, and if for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the mapping $x \in \mathbb{R}^d \mapsto \partial_\mu u(\mu)(x)$ is differentiable, its derivative being jointly continuous at any point (μ,x) such that $x \in \mathrm{Supp}(\mu)$. The gradient is then denoted by $\partial_x\partial_\mu u(\mu)(x) \in \mathbb{R}^{d\times d}$. We say that u

$\in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$ if it is partially $\mathcal{C}^2$, $\partial_x\partial_\mu u(\mu) \in L^\infty_\mu(\mathbb{R}^{d\times d})$, and for any compact set $\mathcal{K}$ of $\mathcal{P}_2(\mathbb{R}^d)$, we have

$$\sup_{\mu\in\mathcal{K}}\Big[\int_{\mathbb{R}^d}|\partial_\mu u(\mu)(x)|^2\mu(dx) + \big\|\partial_x\partial_\mu u(\mu)\big\|_\infty\Big] \quad < \quad \infty.$$

As shown in [14], if the lifted function U is twice continuously Fréchet differentiable on $L^2(\mathcal{F}_0;\mathbb{R}^d)$, then u lies in $\mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$. In this case, the second Fréchet derivative $D^2U(X)$ is identified indifferently by Riesz' theorem as a bilinear form on $L^2(\mathcal{F}_0;\mathbb{R}^d)$ or as a symmetric operator (hence bounded) on $L^2(\mathcal{F}_0;\mathbb{R}^d)$, denoted by $D^2U(X) \in S(L^2(\mathcal{F}_0;\mathbb{R}^d))$, and we have the relation (see Appendix A.2 in [11]):

$$\mathbb{E}\Big[D^2U(X)(YN).YN\Big] \quad = \quad \mathbb{E}\Big[\mathrm{tr}\big(\partial_x\partial_\mu u(\mathbb{P}_X)(X)YY^\intercal\big)\Big], \tag{4.2}$$

for any $X \in L^2(\mathcal{F}_0;\mathbb{R}^d)$, $Y \in L^2(\mathcal{F}_0;\mathbb{R}^{d\times q})$, and where $N \in L^2(\mathcal{F}_0;\mathbb{R}^q)$ is independent of (X,Y) with zero mean $\mathbb{E}[N] = 0$.

We shall need a chain rule (or Itô's formula) for functions defined on $\mathcal{P}_2(\mathbb{R}^d)$, proved independently in [8] and [14], see also the Appendix in [11], and that we recall here. Let us consider an $\mathbb{R}^d$-valued Itô process

$$dX_t \quad = \quad b_t dt + \sigma_t dB_t, \quad X_0 \in L^2(\mathcal{F}_0;\mathbb{R}^d),$$

where (b_t) and (σ_t) are progressively measurable processes with respect to the filtration generated by the n-dimensional Brownian motion B, valued respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d\times n}$, and satisfying the integrability condition:

$$\mathbb{E}\Big[\int_0^T |b_t|^2 + |\sigma_t|^2 dt\Big] \quad < \quad \infty. \tag{4.3}$$

Let $u \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$. Then, for all $t \in [0,T]$,

$$u(\mathbb{P}_{X_t}) \quad = \quad u(\mathbb{P}_{X_0}) \; + \; \int_0^t \mathbb{E}\big[\partial_\mu u(\mathbb{P}_{X_s})(X_s).b_s + \frac{1}{2}\mathrm{tr}\big(\partial_x\partial_\mu u(\mathbb{P}_{X_s})(X_s)\sigma_s\sigma_s^\intercal\big)\big]ds. \tag{4.4}$$

We have now the ingredients for deriving the Bellman equation associated to the DPP (3.10), and it turns out that it takes the following form:

$$\left\{\begin{array}{rcl} \partial_t v \; + \; \inf\limits_{\tilde\alpha\in L(\mathbb{R}^d;A)}\Big[\hat f(t,\mu,\tilde\alpha) + \; < \mathcal{L}_t^{\tilde\alpha}v(t,\mu),\mu > \Big] & = & 0, \;\; \text{on} \;\; [0,T)\times\mathcal{P}_2(\mathbb{R}^d), \\ v(T,.) & = & \hat g, \;\; \text{on} \;\; \mathcal{P}_2(\mathbb{R}^d) \end{array}\right. \tag{4.5}$$

where for $\tilde\alpha \in L(\mathbb{R}^d;A)$, $\varphi \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$ and $(t,\mu) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, $\mathcal{L}_t^{\tilde\alpha}\varphi(\mu) \in L^2_\mu(\mathbb{R})$ is the function: $\mathbb{R}^d \to \mathbb{R}$ defined by

$$\begin{aligned} \mathcal{L}_t^{\tilde\alpha}\varphi(\mu)(x) \quad := \quad & \partial_\mu\varphi(\mu)(x).b(t,x,\mu,\tilde\alpha(x),\tilde\alpha\star\mu) \\ & + \frac{1}{2}\mathrm{tr}\big(\partial_x\partial_\mu\varphi(\mu)(x)\sigma\sigma^\intercal(t,x,\mu,\tilde\alpha(x),\tilde\alpha\star\mu)\big). \end{aligned} \tag{4.6}$$

In the spirit of classical verification theorem for stochastic control of diffusion processes, we prove the following result in our McKean-Vlasov control framework, which shall rely on Itô's formula.

Theorem 4.1 *(Verification theorem)*
Let $w : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function in $\mathcal{C}_b^{1,2}([0,T] \times \mathcal{P}_2(\mathbb{R}^d))$, i.e. w is continuous on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, $w(t,.) \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$, for all $t \in [0,T]$, and $w(.,\mu) \in \mathcal{C}^1([0,T))$. Suppose that w is solution to (4.5), and there exists for all $(t,\mu) \in [0,T) \times \mathcal{P}_2(\mathbb{R}^d)$ an element $\tilde{\alpha}^*(t,.,\mu) \in L(\mathbb{R}^d;A)$ attaining the infimum in (4.5) s.t. the mapping $(t,x,\mu) \mapsto \tilde{\alpha}^*(t,x,\mu) \in Lip([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);A)$. Then, $w = v$, and the feedback control $\alpha^* \in \mathcal{A}$ defined by

$$\alpha_t^* \;=\; \tilde{\alpha}^*(t, X_t, \mathbb{P}_{X_t}), \quad 0 \le t < T,$$

is an optimal control, i.e. $V_0 = J(\alpha^*)$.

Proof. Fix $(t, \mu = \mathbb{P}_\xi) \in [0,T) \times \mathcal{P}_2(\mathbb{R}^d)$, and consider some arbitrary feedback control $\alpha \in \mathcal{A}$ associated to $X^{t,\xi}$ the solution to the controlled SDE (3.3). Under condition **(H1)**, we have the standard estimate

$$\mathbb{E}\big[\sup_{t \le s \le T} |X_s^{t,\xi}|^2\big] \;\le\; C\big(1 + \mathbb{E}|\xi|^2\big) \;<\; \infty,$$

which implies that

$$\begin{aligned}\mathbb{E}\Big[\int_t^T & \big|b(s, X_s^{t,\xi}, \mathbb{P}_{X_s^{t,\xi}}, \tilde{\alpha}(s, X_s^{t,\xi}, \mathbb{P}_{X_s^{t,\xi}}), \tilde{\alpha}(s,.,\mathbb{P}_{X_s^{t,\xi}}) \star \mathbb{P}_{X_s^{t,\xi}})\big|^2 \\ & + \big|\sigma(s, X_s^{t,\xi}, \mathbb{P}_{X_s^{t,\xi}}, \tilde{\alpha}(s, X_s^{t,\xi}, \mathbb{P}_{X_s^{t,\xi}}), \tilde{\alpha}(s,.,\mathbb{P}_{X_s^{t,\xi}}) \star \mathbb{P}_{X_s^{t,\xi}})\big|^2 ds\Big] \;<\; \infty.\end{aligned}$$

One can then apply the Itô's formula (4.4) to $w(s, \mathbb{P}_{X_s^{t,\xi}}) = w(s, \mathbb{P}_s^{t,\mu})$ (with the definition (3.5)) between $s = t$ and $s = T$, and obtain

$$\begin{aligned} w(T, \mathbb{P}_T^{t,\mu}) \;&=\; w(t,\mu) \;+\; \int_t^T \frac{\partial w}{\partial t}(s, \mathbb{P}_s^{t,\mu}) + \\ & \quad \mathbb{E}\Big[\partial_\mu w(s, \mathbb{P}_s^{t,\mu})(X_s^{t,\xi}).b(s, X_s^{t,\xi}, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s, X_s^{t,\xi}, \mathbb{P}_s^{t,\mu}), \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}) \star \mathbb{P}_s^{t,\mu}) \\ & \quad + \frac{1}{2}\mathrm{tr}\big(\partial_x \partial_\mu w(s, \mathbb{P}_s^{t,\mu})(X_s^{t,\xi})\sigma_s \sigma_s^\intercal(s, X_s^{t,\xi}, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s, X_s^{t,\xi}, \mathbb{P}_s^{t,\mu}), \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}) \star \mathbb{P}_s^{t,\mu})\big)\Big] ds \\ &=\; w(t,\mu) \;+\; \int_t^T \frac{\partial w}{\partial t}(s, \mathbb{P}_s^{t,\mu}) \;+\; < \mathcal{L}_s^{\tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu})} w(s, \mathbb{P}_s^{t,\mu}), \mathbb{P}_s^{t,\mu} > ds, \end{aligned} \tag{4.7}$$

where we used in the second equality the fact that $\mathbb{P}_s^{t,\mu}$ is the distribution of $X_s^{t,\xi}$ for $s \in [t,T]$. Since $x \mapsto \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}) \in L(\mathbb{R}^d;A)$ for $s \in [t,T]$, we deduce from the Bellman equation satisfied by w and (4.7) that

$$\hat{g}(\mathbb{P}_T^{t,\mu}) \;\ge\; w(t,\mu) - \int_t^T \hat{f}(s, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s,.,\mathbb{P}_s^{t,\mu}))ds.$$

Since α is arbitrary in $\mathcal{A}$, this shows that $w(t,\mu) \le v(t,\mu)$.

In the final step, let us apply the same Itô's argument (4.7) with the feedback control $\alpha^* \in \mathcal{A}$ associated with the fonction $\tilde{\alpha}^* \in Lip([0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d);A)$. Since $\tilde{\alpha}$ attains the infimum in (4.5), we thus get

$$\hat{g}(\mathbb{P}_T^{t,\mu}) \;=\; w(t,\mu) - \int_t^T \hat{f}(s, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}^*(s,.,\mathbb{P}_s^{t,\mu}))ds,$$

which shows that $w(t,\mu) = J(t,\mu,\alpha^*)$ $(\geq v(t,\mu))$, and therefore gives the required result: $v(t,\mu) = w(t,\mu) = J(t,\mu,\alpha^*)$. $\square$

We shall apply the verification theorem in the next section, where we can derive explicit (smooth) solutions to the Bellman equation (4.5) in some class of examples.

5 Applications

5.1 No mean-field interaction

We consider the classical case of stochastic control where there is no mean-field interaction in the dynamics of the state process, i.e. $b(x,a)$ and $\sigma(x,a)$ do not depend on μ,λ, as well as in the cost functions $f(x,a)$ and $g(x)$. In this special case, let us show how the verification Theorem 4.1 is reduced to the classical verification result for smooth functions on $[0,T]\times\mathbb{R}^d$, see e.g. [18] or [32].

Suppose that there exists a function u in $\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ solution to the standard HJB equation

$$\left\{\begin{array}{rcl} \partial_t u + \inf\limits_{a\in A}\big[f(t,x,a) + L_t^a u(t,x)\big] & = & 0, \quad \text{on} \quad [0,T)\times\mathbb{R}^d, \\ u(T,.) & = & g \quad \text{on} \quad \mathbb{R}^d. \end{array}\right. \tag{5.1}$$

where L_t^a is the second-order differential operator

$$L_t^a u(t,x) \;=\; \partial_x u(t,x).b(t,x,a) + \frac{1}{2}\mathrm{tr}\big(\partial^2_{xx}u(t,x)\sigma\sigma^\intercal(t,x,a)\big),$$

and that for all $(t,x)\in[0,T)\times\mathbb{R}^d$, there exists $\hat{a}(t,x)$ attaining the argmin in (5.1), s.t. the map $x\mapsto\hat{a}(t,x)$ is Lipschitz on $\mathbb{R}^d$.

Let us then consider the function defined on $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ by

$$w(t,\mu) \;=\; <u(t,.),\mu> = \int_{\mathbb{R}^d} u(t,x)\mu(dx).$$

The lifted function of w is thus equal to $\mathcal{W}(t,X) = \mathbb{E}[u(t,X)]$ with Fréchet derivative (with respect to $X\in L^2(\mathcal{F}_0,\mathbb{P})$): $[D\mathcal{W}](t,X)(Y) = \mathbb{E}[\partial_x u(t,X).Y]$. Assuming that the time derivative of u w.r.t. t satisfies a quadratic growth condition in x, the first derivative of u w.r.t. x satisfies a linear growth condition, and the second derivative of u w.r.t. x is bounded, this shows that w lies in $\mathcal{C}^{1,2}_b([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ with

$$\partial_t w(t,\mu) \;=<\partial_t u(t,.),\mu>, \qquad \partial_\mu w(t,\mu) \;=\; \partial_x u(t,.), \quad \partial_x\partial_\mu v(t,\mu) \;=\; \partial^2_{xx}u(t,.).$$

Recalling the definition (4.6) of $\mathcal{L}_t^{\tilde{\alpha}}w(t,\mu)$, we then get for any fixed $(t,\mu)\in[0,T)\times\mathcal{P}_2(\mathbb{R}^d)$:

$$\begin{aligned} &\partial_t w(t,\mu) \;+\; \inf_{\tilde{\alpha}\in L(\mathbb{R}^d;A)}\Big[\hat{f}(t,\mu,\tilde{\alpha}) + \;<\mathcal{L}_t^{\tilde{\alpha}}w(t,\mu),\mu>\Big] \\ =\;& \inf_{\tilde{\alpha}\in L(\mathbb{R}^d;A)}\int_{\mathbb{R}^d}\big[\partial_t u(t,x) + f(t,x,\tilde{\alpha}(x)) + L_t^{\tilde{\alpha}(x)}u(t,x)\big]\mu(dx) \\ =\;& \int_{\mathbb{R}^d}\inf_{a\in A}\big[\partial_t u(t,x) + f(t,x,a) + L_t^a u(t,x)\big]\mu(dx). \end{aligned} \tag{5.2}$$

Indeed, the inequality $\geq$ in (5.2) is clear since $\tilde{\alpha}(x)$ lies in A for all $x \in \mathbb{R}^d$, and $\tilde{\alpha} \in L(\mathbb{R}^d; A)$. Conversely, by taking $\hat{a}(t,x)$ which attains the infimum in (5.1), and since the map $x \in \mathbb{R}^d \mapsto \hat{a}(t,x)$ is Lipschitz, we then have

$$\begin{aligned}
& \int_{\mathbb{R}^d} \inf_{a \in A} \big[\partial_t u(t,x) + f(t,x,a) + L_t^a u(t,x)\big] \mu(dx) \\
= & \int_{\mathbb{R}^d} \big[\partial_t u(t,x) + f(t,x,\hat{a}(t,x)) + L_t^{\hat{a}(t,x)} u(t,x)\big] \mu(dx) \\
\geq & \inf_{\tilde{\alpha} \in L(\mathbb{R}^d;A)} \int_{\mathbb{R}^d} \big[\partial_t u(t,x) + f(t,x,\tilde{\alpha}(x)) + L_t^{\tilde{\alpha}(x)} u(t,x)\big] \mu(dx),
\end{aligned}$$

which thus shows the equality (5.2). Since u is solution to (5.1), this proves that w is solution to the Bellman equation (4.5), $\tilde{\alpha}^*(t,x) = \hat{a}(t,x)$ is an optimal feedback control, and therefore, the value function is equal to $v(t,\mu) = < u(t,.), \mu >$.

5.2 Linear-quadratic McKean-Vlasov control problem

We consider a multivariate linear McKean-Vlasov controlled dynamics with coefficients given by

$$\begin{aligned}
b(t,x,\mu,a,\lambda) &= b_0(t) + B(t)x + \bar{B}(t)\bar{\mu} + C(t)a + \bar{C}(t)\bar{\lambda}, \\
\sigma(t,x,\mu,a,\lambda) &= \sigma_0(t) + D(t)x + \bar{D}(t)\bar{\mu} + F(t)a + \bar{F}(t)\bar{\lambda},
\end{aligned} \tag{5.3}$$

for $(t,x,\mu,a,\lambda) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^m \times \mathcal{P}_2(\mathbb{R}^m)$, where we set

$$\bar{\mu} := \int_{\mathbb{R}^d} x\mu(dx), \qquad \bar{\lambda} := \int_{\mathbb{R}^m} a\lambda(da),$$

Here B, $\bar{B}$, D, $\bar{D}$ are deterministic continuous functions valued in $\mathbb{R}^{d\times d}$, and C, $\bar{C}$, F, $\bar{F}$ are deterministic continuous functions valued in $\mathbb{R}^{d\times m}$, and b_0, σ_0 are deterministic continuous function valued in $\mathbb{R}^d$. The quadratic cost functions are given by

$$\begin{aligned}
f(t,x,\mu,a,\lambda) &= x^\intercal Q_2(t)x + \bar{\mu}^\intercal \bar{Q}_2(t)\bar{\mu} + a^\intercal R_2(t)a + \bar{\lambda}^\intercal \bar{R}_2(t)\bar{\lambda} + 2x^\intercal M_2(t)a \\
&\quad + 2\bar{\mu}^\intercal \bar{M}_2(t)\bar{\lambda} + q_1(t).x + \bar{q}_1(t).\bar{\mu} + r_1(t).a + \bar{r}_1(t).\bar{\lambda}, \\
g(x,\mu) &= x^\intercal P_2 x + \bar{\mu}^\intercal \bar{P}_2 \bar{\mu} + p_1.x + \bar{p}_1.\bar{\mu},
\end{aligned} \tag{5.4}$$

where Q_2, $\bar{Q}_2$ are deterministic continuous functions, P_2, $\bar{P}_2$ are constants valued in $\mathbb{R}^{d\times d}$, R_2, $\bar{R}_2$ are deterministic continuous functions valued in $\mathbb{R}^{m\times m}$, M_2, $\bar{M}_2$ are deterministic continuous functions valued in $\mathbb{R}^{d\times m}$, q_1, $\bar{q}_1$ are deterministic continuous functions, p_1, $\bar{p}_1$ are constants valued in $\mathbb{R}^d$, and r_1, $\bar{r}_1$ are deterministic continuous functions valued in $\mathbb{R}^m$. Since f and g are real-valued, we may assume w.l.o.g. that all the matrices Q_2, $\bar{Q}_2$, R_2, $\bar{R}_2$, P_2, $\bar{P}_2$ are symmetric. We denote by $\mathbb{S}^d$ the set of symmetric matrices in $\mathbb{R}^{d\times d}$, by $\mathbb{S}^d_+$ the subset of nonnegative symmetric matrices, by $\mathbb{S}^d_{>+}$ the subset of symmetric positive definite matrices, and similarly for $\mathbb{S}^m$, $\mathbb{S}^m_+$, $\mathbb{S}^m_{>+}$.

The functions $\hat{f}$ and $\hat{g}$ defined in (3.7) are then given by

$$\left\{\begin{aligned}
\hat{f}(t,\mu,\tilde{\alpha}) &= \mathrm{Var}(\mu)(Q_2(t)) + \bar{\mu}^\intercal (Q_2(t) + \bar{Q}_2(t))\bar{\mu} \\
&\quad + \mathrm{Var}(\tilde{\alpha} \star \mu)(R_2(t)) + \overline{\tilde{\alpha} \star \mu}^\intercal (R_2(t) + \bar{R}_2(t))\overline{\tilde{\alpha} \star \mu} \\
&\quad + 2\bar{\mu}^\intercal (M_2(t) + \bar{M}_2(t))\overline{\tilde{\alpha} \star \mu} + 2\int_{\mathbb{R}^d} (x - \bar{\mu})^\intercal M_2(t)\tilde{\alpha}(x)\mu(dx) \\
&\quad + \big(q_1(t) + \bar{q}_1(t)\big).\bar{\mu} + \big(r_1(t) + \bar{r}_1(t)\big).\overline{\tilde{\alpha} \star \mu} \\
\hat{g}(\mu) &= \mathrm{Var}(\mu)(P_2) + \bar{\mu}^\intercal (P_2 + \bar{P}_2)\bar{\mu} + (p_1 + \bar{p}_1).\bar{\mu},
\end{aligned}\right. \tag{5.5}$$

for any $(t,\mu) \in [0,T) \times \mathcal{P}_2(\mathbb{R}^d)$, $\tilde{\alpha} \in L(\mathbb{R}^d;A)$ (here with $A = \mathbb{R}^m$), where we set for any Λ in $\mathbb{S}^d$ (resp. in $\mathbb{S}^m$), and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ (resp. $\mathcal{P}_2(\mathbb{R}^m)$):

$$\bar{\mu}_2(\Lambda) \ := \ \int x^\intercal \Lambda x \mu(dx), \qquad \mathrm{Var}(\mu)(\Lambda) \ := \ \bar{\mu}_2(\Lambda) - \bar{\mu}^\intercal \Lambda \bar{\mu}.$$

We look for a value function solution to the Bellman equation (4.5) in the form

$$w(t,\mu) \ = \ \mathrm{Var}(\mu)(\Lambda(t)) + \bar{\mu}^\intercal \Gamma(t) \bar{\mu} + \gamma(t).\bar{\mu} + \chi(t), \tag{5.6}$$

for some functions Λ, $\Gamma \in \mathcal{C}^1([0,T];\mathbb{S}^d)$, $\gamma \in \mathcal{C}^1([0,T];\mathbb{R}^d)$, and $\chi \in \mathcal{C}^1([0,T];\mathbb{R})$. The lifted function of w in (5.6) is given by

$$\mathcal{W}(t,X) \ = \ \mathbb{E}[X^\intercal \Lambda(t) X] + \mathbb{E}[X]^\intercal (\Gamma(t) - \Lambda(t)) \mathbb{E}[X] + \gamma(t).\mathbb{E}[X] + \chi(t),$$

for $X \in L^2(\mathcal{F}_0;\mathbb{R}^d)$. By computing for all $Y \in L^2(\mathcal{F}_0;\mathbb{R}^d)$ the difference

$$\mathcal{W}(t,X+Y) - \mathcal{W}(t,X) \ = \ \mathbb{E}\Big[\big(2X^\intercal \Lambda(t) + 2\mathbb{E}[X]^\intercal (\Gamma(t) - \Lambda(t)) + \gamma(t)\big).Y\Big] \ + \ o(\|Y\|_{L^2}),$$

we see that $\mathcal{W}$ is Fréchet differentiable (w.r.t. X) with $[D\mathcal{W}](t,X)(Y) = \mathbb{E}\big[\big(2X^\intercal \Lambda(t) + 2\mathbb{E}[X]^\intercal (\Gamma(t) - \Lambda(t)) + \gamma(t)\big).Y\big]$. This shows that w lies in $\mathcal{C}^{1,2}_b([0,T] \times \mathcal{P}_2(\mathbb{R}^d))$ with

$$\begin{aligned}
\partial_t w(t,\mu) \ &= \ \mathrm{Var}(\mu)(\Lambda'(t)) + \bar{\mu}^\intercal \Gamma'(t) \bar{\mu} + \gamma'(t) \bar{\mu} + \chi'(t), \\
\partial_\mu w(t,\mu)(x) \ &= \ 2x^\intercal \Lambda(t) + 2\bar{\mu}^\intercal (\Gamma(t) - \Lambda(t)) + \gamma(t), \\
\partial_x \partial_\mu w(t,\mu)(x) \ &= \ 2\Lambda(t).
\end{aligned}$$

Together with the quadratic expression (5.5) of $\hat{f}$, $\hat{g}$, we then see that w satisfies the Bellman equation (4.5) iff

$$\begin{aligned}
& \mathrm{Var}(\mu)(\Lambda(T)) + \bar{\mu}^\intercal \Gamma(T) \bar{\mu} + \gamma(T).\bar{\mu} + \chi(T) \\
= \ & \mathrm{Var}(\mu)(P_2) + \bar{\mu}^\intercal (P_2 + \bar{P}_2) \bar{\mu} + (p_1 + \bar{p}_1).\bar{\mu},
\end{aligned} \tag{5.7}$$

holds for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and

$$\begin{aligned}
& \mathrm{Var}(\mu)\big(\Lambda'(t) + Q_2(t) + D(t)^\intercal \Lambda(t) D(t) + \Lambda(t) B(t) + B(t)^\intercal \Lambda(t)\big) + \inf_{\tilde{\alpha} \in L(\mathbb{R}^d,A)} G^\mu_t(\tilde{\alpha}) \\
& + \bar{\mu}^\intercal \Big(\Gamma'(t) + Q_2(t) + \bar{Q}_2(t) + (D(t) + \bar{D}(t))^\intercal \Lambda(t) (D(t) + \bar{D}(t)) \\
& \qquad + \Gamma(t)(B(t) + \bar{B}(t)) + (B(t) \ + \ \bar{B}(t))^\intercal \Gamma(t)\Big) \bar{\mu} \\
& + \big(q_1(t) + \bar{q}_1(t) + \gamma(t)(B(t) + \bar{B}(t)) + 2\sigma_0^\intercal \Lambda(t)(D(t) + \bar{D}(t)) + 2b_0(t)^\intercal \Gamma(t)\big) \bar{\mu} \\
& + \chi'(t) \ + \ \gamma(t).b_0(t) + \sigma_0(t)^\intercal \Lambda(t) b_0(t) \\
= \ & 0,
\end{aligned} \tag{5.8}$$

holds for all $t \in [0,T)$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, where the function $G^\mu_t : L^2_\mu(A) \supset L(\mathbb{R}^d;A) \to \mathbb{R}$ is defined by

$$\begin{aligned}
G^\mu_t(\tilde{\alpha}) \ = \ & \mathrm{Var}(\tilde{\alpha} \star \mu)(U_t) \ + \ \overline{\tilde{\alpha} \star \mu}^\intercal V_t \overline{\tilde{\alpha} \star \mu} \ + \ 2 \int_{\mathbb{R}^d} (x - \bar{\mu})^\intercal S_t \tilde{\alpha}(x) \mu(dx) \\
& + 2\bar{\mu}^\intercal Z_t \overline{\tilde{\alpha} \star \mu} \ + \ Y_t.\overline{\tilde{\alpha} \star \mu},
\end{aligned} \tag{5.9}$$

and we set $U_t = U(t,\Lambda(t))$, $V_t = V(t,\Lambda(t))$, $S_t = S(t,\Lambda(t))$, $Z_t = Z(t,\Lambda(t),\Gamma(t))$, $Y_t = Y(t,\Gamma(t),\gamma(t))$ with

$$\left\{\begin{array}{rcl} U(t,\Lambda(t)) &=& F(t)^\intercal\Lambda(t)F(t)+R_2(t),\\ V(t,\Lambda(t)) &=& (F(t)+\bar{F}(t))^\intercal\Lambda(t)(F(t)+\bar{F}(t))+R_2(t)+\bar{R}_2(t),\\ S(t,\Lambda(t)) &=& D(t)^\intercal\Lambda(t)F(t)+\Lambda(t)C(t)+M_2(t),\\ Z(t,\Lambda(t),\Gamma(t)) &=& (D(t)+\bar{D}(t))^\intercal\Lambda(t)(F(t)+\bar{F}(t))+\Gamma(t)(C(t)+\bar{C}(t))+M_2(t)+\bar{M}_2(t)\\ Y(t,\Gamma(t),\gamma(t)) &=& \big(C(t)+\bar{C}(t)\big)^\intercal\gamma(t)+r_1(t)+\bar{r}_1(t)+2\big(F(t)+\bar{F}(t)\big)^\intercal\Lambda(t)\sigma_0(t). \end{array}\right. \tag{5.10}$$

We now search for the infimum of the function G_t^μ. After some straightforward calculation, we derive the Gateaux derivative of G_t^μ at $\tilde{\alpha}$ in the direction $\beta \in L^2_\mu(A)$, which is given by:

$$DG_t^\mu(\tilde{\alpha},\beta) \;:=\; \lim_{\varepsilon\to 0}\frac{G_t^\mu(\tilde{\alpha}+\varepsilon\beta)-G_t^\mu(\tilde{\alpha})}{\varepsilon} \;=\; \int_{\mathbb{R}^d}\dot{g}_t^\mu(x,\tilde{\alpha}).\beta(x)\mu(dx)$$

with

$$\dot{g}_t^\mu(x,\tilde{\alpha}) \;=\; 2U_t\tilde{\alpha}+2(V_t-U_t)\overline{\tilde{\alpha}\star\mu}+2S_t^\intercal(x-\bar{\mu})+2Z_t^\intercal\bar{\mu}+Y_t.$$

Suppose that the symmetric matrices U_t and V_t in (5.10) are positive, hence invertible (this will be discussed later on). Then, the function G_t^μ is convex and coercive on the Hilbert space $L^2_\mu(A)$, and attains its infimum at some $\tilde{\alpha} = \tilde{\alpha}^*(t,.,\mu)$ s.t. $DG_t^\mu(\tilde{\alpha};.)$ vanishes, i.e. $\dot{g}_t^\mu(x,\tilde{\alpha}^*(t,.,\mu)) = 0$ for all $x \in \mathbb{R}^d$, which gives:

$$\tilde{\alpha}^*(t,x,\mu) \;=\; -U_t^{-1}S_t^\intercal(x-\bar{\mu}) \;-\; V_t^{-1}Z_t^\intercal\bar{\mu} \;-\; \frac{1}{2}V_t^{-1}Y_t. \tag{5.11}$$

It is clear that $\tilde{\alpha}^*(t,.,\mu)$ lies in $L(\mathbb{R}^d;A)$, and so after some straightforward caculation:

$$\begin{aligned}\inf_{\tilde{\alpha}\in L(R^d,A)} G_t^\mu(\tilde{\alpha}) \;=\; G_t^\mu(\tilde{\alpha}^*(t,.,\mu)) \;&=\; -\mathrm{Var}(\mu)\big(S_tU_t^{-1}S_t^\intercal\big) \;-\; \bar{\mu}^\intercal\big(Z_tV_t^{-1}Z_t^\intercal\big)\bar{\mu}\\ &\quad -\; Y_t^\intercal V_t^{-1}Z_t^\intercal\bar{\mu} \;-\; \frac{1}{4}Y_t^\intercal V_t^{-1}Y_t.\end{aligned}$$

Plugging the above expression in (5.8), we observe that the relation (5.7)-(5.8), hence the Bellman equation, is satisfied by identifying the terms in $\mathrm{Var}(\mu)(.)$, $\bar{\mu}^\intercal(.)\bar{\mu}$, $\bar{\mu}$, which leads to the system of ordinary differential equations (ODEs) for $(\Lambda,\Gamma,\gamma,\chi)$:

$$\left\{\begin{array}{rcl} \Lambda'(t)+Q_2(t)+D(t)^\intercal\Lambda(t)D(t)+\Lambda(t)B(t)+B(t)^\intercal\Lambda(t)&&\\ -S(t,\Lambda(t))U(t,\Lambda(t))^{-1}S(t,\Lambda(t))^\intercal &=& 0,\\ \Lambda(T) &=& P_2, \end{array}\right. \tag{5.12}$$

$$\left\{\begin{array}{rcl} \Gamma'(t)+Q_2(t)+\bar{Q}_2(t)+(D(t)+\bar{D}(t))^\intercal\Lambda(t)(D(t)+\bar{D}(t))+\Gamma(t)(B(t)+\bar{B}(t))&&\\ +(B(t)+\bar{B}(t))^\intercal\Gamma(t)-Z(t,\Lambda(t),\Gamma(t))V(t,\Lambda(t))^{-1}Z(t,\Lambda(t),\Gamma(t))^\intercal &=& 0,\\ \Gamma(T) \;=\; P_2+\bar{P}_2,&& \end{array}\right. \tag{5.13}$$

$$\left\{\begin{array}{rcl} \gamma'(t)+\big(B(t)+\bar{B}(t)\big)\gamma(t)-Z(t,\Lambda(t),\Gamma(t))V(t,\Lambda(t))^{-1}Y(t,\Gamma(t),\gamma(t))&&\\ +q_1(t)+\bar{q}_1(t)+2\big(D(t)+\bar{D}(t)\big)^\intercal\Lambda(t)\sigma_0(t)+2\Gamma(t)b_0(t) &=& 0,\\ \gamma(T) \;=\; p_1+\bar{p}_1&& \end{array}\right. \tag{5.14}$$

$$\left\{\begin{array}{rcl} \chi'(t) - \frac{1}{4}Y(t,\Gamma(t),\gamma(t))^\intercal V(t,\Lambda(t))^{-1}Y(t,\Gamma(t),\gamma(t)) & & \\ + \gamma(t).b_0(t) + \sigma_0(t)^\intercal\Lambda(t)\sigma_0(t) & = & 0, \\ \chi(T) & = & 0. \end{array}\right. \tag{5.15}$$

Therefore, the resolution of the Bellman equation in the LQ framework is reduced to the resolution of the Riccati equations (5.12) and (5.13) for Λ and Γ, and then given (Λ,Γ), to the resolution of the linear ODEs (5.14) and (5.15) for γ and χ. Suppose that there exists a solution $(\Lambda,\Gamma) \in \mathcal{C}^1([0,T];\mathbb{S}^d) \times \mathcal{C}^1([0,T];\mathbb{S}^d)$ to (5.12)-(5.13) s.t. (U_t,V_t) in (5.10) lies in $\mathbb{S}^m_{>+} \times \mathbb{S}^m_{>+}$ for all $t \in [0,T]$ (see Remark 5.1). Then, the above calculations are justified a posteriori, and by noting also that the mapping $(t,x,\mu) \mapsto \tilde{\alpha}^*(t,x,\mu) \in Lip([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);A)$, we deduce by the verification theorem that the value function v is equal to w in (5.6) with $(\Lambda,\Gamma,\gamma,\chi)$ solution to (5.12)-(5.13)-(5.14)-(5.15). Moreover, the optimal control is given in feedback form from (5.11) by

$$\alpha^*_t \;=\; \tilde{\alpha}^*(t,X^*_t,\mathbb{P}_{X^*_t}) = -U_t^{-1}S_t^\intercal(X^*_t - \mathbb{E}[X^*_t]) \;-\; V_t^{-1}Z_t^\intercal\mathbb{E}[X^*_t] \;-\; \frac{1}{2}V_t^{-1}Y_t, \tag{5.16}$$

where X^* is the state process controlled by α^*.

Remark 5.1 In the case where $M_2 = \bar{M}_2 = 0$ (i.e. no crossing term between the state and the control in the quadratic cost function f), it is shown in Proposition 3.1 and 3.2 in [37] that under the condition

$$\begin{array}{ll} P_2 \;\geq\; 0,\; P_2 + \bar{P}_2 \;\geq\; 0, & Q_2(t) \;\geq\; 0,\; Q_2(t) + \bar{Q}_2(t) \;\geq\; 0, \\ & R_2(t) \;\geq \delta I_m,\; R_2(t) + \bar{R}_2(t) \;\geq\; \delta I_m \end{array} \tag{5.17}$$

for some $\delta > 0$, the Riccati equations (5.12)-(5.13) admit unique solutions $(\Lambda,\Gamma) \in \mathcal{C}^1([0,T];\mathbb{S}^d_+) \times \mathcal{C}^1([0,T];\mathbb{S}^d_+)$, and then U_t, V_t in (5.10) are symmetric positive definite matrices, i.e. lie in $\mathbb{S}^m_{>+}$ for all $t \in [0,T]$. In this case, we retrieve the expressions (5.16) of the optimal control in feedback form obtained in [37].

We shall see in the next two paragraphs, some other examples arising from finance with explicit solutions where condition (5.17) is not satisfied. □

5.2.1 Mean-variance portfolio selection

The mean-variance discrete-time problem consists in minimizing a cost functional of the form:

$$\begin{array}{rcl} J(\alpha) & = & \frac{\eta}{2}\mathrm{Var}(X_T) - \mathbb{E}[X_T] \\ & = & \mathbb{E}\big[\frac{\eta}{2}\big(X_T\big)^2 - X_T\big] \;-\; \frac{\eta}{2}\Big(\mathbb{E}[X_T]\Big)^2 \end{array}$$

for some $\eta > 0$, with a dynamics for the wealth process $X = (X^\alpha)$ controlled by the amount α_t valued in $A = \mathbb{R}$ invested in one risky stock at time t:

$$dX_t \;=\; r(t)X_t dt + \alpha_t\big(\rho(t)dt + \vartheta(t)dB_t\big), \quad X_0 \;= x_0 \;\in\; \mathbb{R},$$

where r is the interest rate, ρ and $\vartheta > 0$ are the excess rate of return (w.r.t. the interest rate) and volatility of the stock price, and these deterministic functions are assumed to be continuous. This model fits into the LQ framework (5.3)-(5.4) of the McKean-Vlasov problem, with a linear controlled dynamics that does not have mean-field interaction:

$$\begin{aligned} b_0 = 0,\ B(t) = r(t),\ \bar{B} = 0,\ C(t) = \rho(t),\ \bar{C} = 0, \\ \sigma_0 = D = \bar{D} = 0,\ F(t) = \vartheta(t),\ \bar{F} = 0, \\ Q_2 = \bar{Q}_2 = M_2 = \bar{M}_2 = R_2 = \bar{R}_2 = 0, \\ q_1 = \bar{q}_1 = r_1 = \bar{r}_1 = 0,\ P_2 = \frac{\eta}{2},\ \bar{P}_2 = -\frac{\eta}{2},\ p_1 = 0,\ \bar{p}_1 = -1. \end{aligned}$$

The Riccati system (5.12)-(5.13)-(5.14)-(5.15) for $(\Lambda(t), \Gamma(t), \gamma(t), \chi(t))$ is written in this case as

$$\left\{ \begin{array}{rcl} \Lambda'(t) - (\frac{\rho^2(t)}{\vartheta^2(t)} - 2r(t))\Lambda(t) & = & 0, \qquad \Lambda(T) = \frac{\eta}{2}, \\ \Gamma'(t) - \frac{\rho^2(t)\Gamma^2(t)}{\vartheta^2(t)\Lambda(t)} + 2r(t)\Gamma(t) & = & 0, \qquad \Gamma(T) = 0, \\ \gamma'(t) + r(t)\gamma(t) - \gamma(t)\frac{\rho^2(t)\Gamma(t)}{\vartheta^2(t)\Lambda(t)} & = & 0, \qquad \gamma(T) = -1, \\ \chi'(t) - \frac{\rho^2(t)\gamma^2(t)}{4\vartheta^2(t)\Lambda(t)} & = & 0, \qquad \chi(T) = 0, \end{array} \right. \tag{5.18}$$

whose explicit solution is given by

$$\left\{ \begin{array}{rcl} \Lambda(t) & = & \frac{\eta}{2}\exp\big(\int_t^T 2r(s) - \frac{\rho^2(s)}{\vartheta^2(s)}\,ds\big), \\ \Gamma(t) & = & 0, \\ \gamma(t) & = & -\exp\big(\int_t^T r(s)ds\big) \\ \chi(t) & = & \frac{1}{4}\exp\big(\int_t^T \frac{\rho^2(s)}{\vartheta^2(s)}ds\big) - \frac{1}{4}. \end{array} \right. \tag{5.19}$$

Although the condition (5.17) is not satisfied, we see that (U_t, V_t) in (5.10), which are here explicitly given by $U_t = V_t = \vartheta(t)^2\Lambda(t)$, are positive, and this validates our calculations for the verification theorem. Notice also that the functions (Z_t, Y_t) in (5.10) are explicitly given by $Z_t = 0$, $Y_t = \rho(t)\gamma(t)$. Therefore, the optimal control is given in feedback form from (5.16) by

$$\begin{aligned} \alpha_t^* &= \tilde{\alpha}^*(t, X_t^*, \mathbb{P}_{X_t^*}) \\ &= -\frac{\rho(t)}{\vartheta^2(t)}(X_t^* - \mathbb{E}[X_t^*]) + \frac{\rho(t)}{\eta\vartheta^2(t)}\exp\Big(\int_t^T \frac{\rho^2(s)}{\vartheta^2(s)} - r(s)\,ds\Big), \end{aligned} \tag{5.20}$$

where X^* is the optimal wealth process with portfolio strategy α^*, hence with mean process governed by

$$d\mathbb{E}[X_t^*] = r(t)\mathbb{E}[X_t^*]dt + \frac{\rho^2(t)}{\eta\vartheta^2(t)}\exp\Big(\int_t^T \frac{\rho^2(s)}{\vartheta^2(s)} - r(s)\,ds\Big)dt,$$

and explicitly given by

$$\mathbb{E}[X_t^*] = x_0 e^{\int_0^t r(s)ds} + \frac{1}{\eta}\exp\Big(\int_t^T \frac{\rho^2(s)}{\vartheta^2(s)} - r(s)\,ds\Big)\Big(\exp\big(\int_0^t \frac{\rho^2(s)}{\vartheta^2(s)}ds\big) - 1\Big).$$

Plugging into (5.20), we get the optimal control for the mean-variance portfolio problem

$$\alpha_t^* = \frac{\rho(t)}{\vartheta^2(t)}\Big[x_0 e^{\int_0^t r(s)ds} + \frac{1}{\eta}\exp\Big(\int_0^T \frac{\rho^2(s)}{\vartheta^2(s)}ds - \int_t^T r(s)ds\Big) - X_t^*\Big],$$

and retrieve the closed-form expression of the optimal control found in [26] and [3].

5.2.2 Inter-bank systemic risk model

We consider a model of inter-bank borrowing and lending studied in [13] where the log-monetary reserve of each bank in the asymptotics when the number of banks tend to infinity, is governed by the McKean-Vlasov equation:

$$dX_t \;=\; \big[\kappa(\mathbb{E}[X_t]-X_t)+\alpha_t\big]dt+\sigma dB_t, \qquad X_0 \;=\; x_0\in\mathbb{R}. \tag{5.21}$$

Here, $\kappa\geq 0$ is the rate of mean-reversion in the interaction from borrowing and lending between the banks, and $\sigma>0$ is the volatility coefficient of the bank reserve, assumed to be constant. Moreover, all banks can control their rate of borrowing/lending to a central bank with the same policy α in order to minimize a cost functional of the form

$$J(\alpha) \;=\; \mathbb{E}\Big[\int_0^T\Big(\frac{1}{2}\alpha_t^2-q\alpha_t(\mathbb{E}[X_t]-X_t)+\frac{\eta}{2}(\mathbb{E}[X_t]-X_t)^2\Big)dt+\frac{c}{2}(\mathbb{E}[X_T]-X_T)^2\Big],$$

where $q>0$ is a positive parameter for the incentive to borrowing ($\alpha_t>0$) or lending ($\alpha_t<0$), and $\eta>0$, $c>0$ are positive parameters for penalizing departure from the average. This model fits into the LQ McKean-Vlasov framework (5.3)-(5.4) with $d=m=1$ and

$$\begin{aligned} b_0=0,\ B=-\kappa,\ \bar{B}=\kappa,\ C=1,\ \bar{C}=0,\\ \sigma_0=\sigma,\ D=\bar{D}=F=\bar{F}=0,\\ Q_2=\frac{\eta}{2},\ \bar{Q}_2=-\frac{\eta}{2},\ R_2=\frac{1}{2},\ \bar{R}_2=0,\ M_2=\frac{q}{2},\ \bar{M}_2=-\frac{q}{2},\\ q_1=\bar{q}_1=r_1=\bar{r}_1=0,\quad P_2=\frac{c}{2},\ \bar{P}_2=-\frac{c}{2},\ p_1=\bar{p}_1=0. \end{aligned}$$

The Riccati system (5.12)-(5.13)-(5.14)-(5.15) for $(\Lambda(t),\Gamma(t),\gamma(t),\chi(t))$ is written in this case as

$$\left\{\begin{array}{rcl@{\qquad}rcl} \Lambda'(t)-2(\kappa+q)\Lambda(t)-2\Lambda^2(t)-\frac{1}{2}(q^2-\eta) &=& 0, & \Lambda(T) &=& \frac{c}{2},\\ \Gamma'(t)-2\Gamma^2(t) &=& 0, & \Gamma(T) &=& 0,\\ \gamma'(t)-2\gamma(t)\Gamma(t) &=& 0, & \gamma(T) &=& 0,\\ \chi'(t)+\sigma^2\Lambda(t)-\frac{1}{2}\gamma^2(t) &=& 0, & \chi(T) &=& 0, \end{array}\right. \tag{5.22}$$

whose explicit solution is given by $\Gamma=\gamma=0$, and

$$\begin{aligned} \chi(t) &= \sigma^2\int_t^T\Lambda(s)ds,\\ \Lambda(t) &= \frac{1}{2}\frac{(q-\eta^2)\big(e^{(\delta^+-\delta^-)(T-t)}-1\big)-c\big(\delta^+e^{(\delta^+-\delta^-)(T-t)}-\delta^-\big)}{\delta^-e^{(\delta^+-\delta^-)(T-t)}-\delta^+\big)-ce^{(\delta^+-\delta^-)(T-t)}-1}, \end{aligned}$$

where we set

$$\delta^\pm \;=\; -(\kappa+q)\pm\sqrt{(\kappa+q)^2+(\eta-q^2)}.$$

Moreover, the functions (U_t,V_t,Z_t,Y_t) in (5.10) are explicitly given by: $U_t=V_t=\frac{1}{2}$ (hence >0), $S_t=\Lambda(t)+\frac{q}{2}$, $Z_t=\Gamma(t)=0$, $Y_t=\gamma(t)=0$. Therefore, the optimal control is given in feedback form from (5.16) by

$$\alpha_t^* \;=\; \tilde{\alpha}^*(t,X_t^*,\mathbb{P}_{X_t^*}) \;=\; -(2\Lambda(t)+q)(X_t^*-\mathbb{E}[X_t^*]), \tag{5.23}$$

where X^* is the optimal log-monetary reserve controlled by the rate of borrowing/lending α^*. We then retrieve the expression found in [13] by sending the number of banks N to infinity in their formula for the optimal control. Actually, from (5.21), we have $d\mathbb{E}[X_t^*] = \mathbb{E}[\alpha_t^*]dt$, while $\mathbb{E}[\alpha_t^*] = 0$ from (5.23). We conclude that the optimal rate of borrowing/lending is equal to

$$\alpha_t^* \;=\; -(2\Lambda(t)+q)(X_t^*-x_0), \quad 0\le t\le T.$$

6 Viscosity solutions

In general, there are no smooth solutions to the HJB equation, and in the spirit of HJB equation for standard stochastic control, we shall introduce in this section a notion of viscosity solutions for the Bellman equation (4.5) in the Wasserstein space of probability measures $\mathcal{P}_2(\mathbb{R}^d)$. We adopt the approach in [30], and detailed in [9], which consists, after the lifting identification between measures and random variables, in working in the Hilbert space $L^2(\mathcal{F}_0;\mathbb{R}^d)$ instead of working in the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$, in order to use the various tools developed for viscosity solutions in Hilbert spaces, in particular in our context, for second order Hamilton-Jacobi equations.

Let us rewrite the the Bellman equation (4.5) in the "Hamiltonian" form:

$$\left\{\begin{array}{rcll} -\dfrac{\partial v}{\partial t} + H(t,\mu,\partial_\mu v(t,\mu),\partial_x\partial_\mu v(t,\mu)) &=& 0 & \text{on } [0,T)\times\mathcal{P}_2(\mathbb{R}^d),\\ v(T,.) &=& \hat{g} & \text{on } \mathcal{P}_2(\mathbb{R}^d)\end{array}\right. \tag{6.1}$$

where H is the function defined by

$$\begin{aligned} H(t,\mu,p,\Gamma) \;=\; -\inf_{\tilde\alpha\in L(\mathbb{R}^d;A)}\Big[&< f(t,.,\mu,\tilde\alpha(.),\tilde\alpha\star\mu) + p(.).b(t,.,\mu,\tilde\alpha(.),\tilde\alpha\star\mu)\\ &+\frac{1}{2}\mathrm{tr}\big(\Gamma(.)\sigma\sigma^\intercal(t,.,\mu,\tilde\alpha(.),\tilde\alpha\star\mu)\big),\mu>\Big],\end{aligned}$$

for $(t,\mu)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, $(p,\Gamma)\in L^2_\mu(\mathbb{R}^d)\times L^\infty_\mu(\mathbb{R}^{d\times d})$.

We then consider the "lifted" Bellman equation in $[0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d)$:

$$\left\{\begin{array}{rcll} -\dfrac{\partial V}{\partial t} + \mathcal{H}(t,\xi,DV(t,\xi),D^2V(t,\xi)) &=& 0 & \text{on } [0,T)\times L^2(\mathcal{F}_0;\mathbb{R}^d),\\ V(T,\xi) &=& \hat{G}(\xi) \;:=\; \mathbb{E}[g(\xi,\mathbb{P}_\xi)], & \xi\in L^2(\mathcal{F}_0;\mathbb{R}^d),\end{array}\right. \tag{6.2}$$

where $\mathcal{H}:[0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d)\times L^2(\mathcal{F}_0;\mathbb{R}^d)\times S(L^2(\mathcal{F}_0;\mathbb{R}^d))\to\mathbb{R}$ is defined by

$$\begin{aligned}\mathcal{H}(t,\xi,P,Q) \;=\; -\inf_{\tilde\alpha\in L(\mathbb{R}^d;A)}\Big\{\mathbb{E}\Big[&f(t,\xi,\mathbb{P}_\xi,\tilde\alpha(\xi),\tilde\alpha\star\mathbb{P}_\xi) + P.b(t,\xi,\mathbb{P}_\xi,\tilde\alpha\star\mathbb{P}_\xi)\\ &+\frac{1}{2}Q\big(\sigma(t,\xi,\mathbb{P}_\xi,\tilde\alpha(\xi),\tilde\alpha\star\mathbb{P}_\xi)N\big).\big(\sigma(t,\xi,\mathbb{P}_\xi,\tilde\alpha(\xi),\tilde\alpha\star\mathbb{P}_\xi)N\big)\Big]\Big\},\end{aligned} \tag{6.3}$$

with $N\in L^2(\mathcal{F}_0,\mathbb{R}^n)$ of zero mean $\mathbb{E}[N]=0$, and independent of ξ. Observe that when v and V are smooth functions respectively in $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ and $[0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d)$, linked by the lifting relation $V(t,\xi)=v(t,\mathbb{P}_\xi)$, then from (4.1)-(4.2), v is solution to the Bellman

equation (6.1) iff V is solution to the Bellman equation (6.2). Let us mention that the lifted Bellman equation was also derived in [4] in the case where $\sigma = \sigma(x)$ is not controlled and does not depend on the distribution of the state process, and there is no dependence on the marginal distribution of the control process on the coefficients b and f.

It is then natural to define viscosity solutions for the Bellman equation (6.1) (hence (4.5)) from viscosity solutions to (6.2). As usual, we say that a function u (resp. U) is locally bounded in $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ (resp. on $[0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d)$) if it is bounded on bounded subsets of $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ (resp. of $[0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d)$), and we denote by u^* (resp. U^*) its upper semicontinuous envelope, and by u_* (resp. U_*) its lower semicontinuous envelope. We denote by $\mathcal{C}^2([0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d))$ the set of real-valued continuous functions on $[0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d)$ which are continuously differentiable in $t \in [0,T)$ and twice continuously Fréchet differentiable on $L^2(\mathcal{F}_0;\mathbb{R}^d)$.

Definition 6.1 *We say that locally bounded function $u : [0,T]\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a viscosity (sub, super) solution to* (6.1) *if the lifted function $U : [0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d) \to \mathbb{R}$ defined by*

$$U(t,\xi) \;=\; u(t,\mathbb{P}_\xi), \qquad (t,\xi) \in [0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d),$$

is a viscosity (sub, super) solution to the lifted Bellman equation (6.2)*, that is:*
(i) $U^*(T,.) \leq \hat{G}$, and for any test function $\Phi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d))$ such that $U^* - \Phi$ has a maximum at $(t_0,\xi_0) \in [0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$, one has

$$-\frac{\partial\Phi}{\partial t}(t_0,\xi_0) + \mathcal{H}(t_0,\xi_0,D\Phi(t_0,\xi_0),D^2\Phi(t_0,\xi_0)) \;\leq\; 0.$$

(ii) $U_*(T,.) \geq \hat{G}$, and for any test function $\Phi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d))$ such that $U_* - \Phi$ has a minimum at $(t_0,\xi_0) \in [0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$, one has

$$-\frac{\partial\Phi}{\partial t}(t_0,\xi_0) + \mathcal{H}(t_0,\xi_0,D\Phi(t_0,\xi_0),D^2\Phi(t_0,\xi_0)) \;\geq\; 0.$$

The main goal of this section is to prove the viscosity characterization of the value function v in (3.8) to the Bellman equation (4.5), hence equivalently the viscosity characterization of the lifted value function $V : [0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d)$ defined by

$$V(t,\xi) \;=\; v(t,\mathbb{P}_\xi), \quad \xi \in L^2(\mathcal{F}_0;\mathbb{R}^d),$$

to the lifted Bellman equation (6.2). We shall strenghten condition **(H1)** by assuming in addition that b,σ are uniformly continuous in t, and bounded in (a,λ):

(H1') There exists some constant $C_{b,\sigma} > 0$ s.t. for all $t,t' \in [0,T]$, $x,x' \in \mathbb{R}^d$, $a,a' \in A$, $\mu,\mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $\lambda,\lambda' \in \mathcal{P}_2(A)$,

$$\begin{aligned} & |b(t,x,\mu,a,\lambda) - b(t',x',\mu',a',\lambda')| + |\sigma(t,x,\mu,a,\lambda) - \sigma(t',x',\mu',a',\lambda')| \\ \leq\; & C_{b,\sigma}\big[m_{b,\sigma}(|t-t'|) + |x-x'| + |a-a'| + W_2(\mu,\mu') + W_2(\lambda,\lambda')\big], \end{aligned}$$

for some modulus $m_{b,\sigma}$ (i.e. $m_{b,\sigma}(\tau) \to 0$ when τ goes to zero) and

$$|b(t,0,\delta_0,a,\lambda)| + |\sigma(t,0,\delta_0,a,\lambda)| \;\leq\; C_{b,\sigma}.$$

We also strenghten condition **(H2)** by making additional (uniform) continuity assumptions on the running and terminal cost functions, and boundedness conditions in (a, λ):

(H2') (i) g is continuous on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and there exists some constant $C_g > 0$ s.t. for all $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$|g(x,\mu)| \quad \leq \quad C_g\big(1 + |x|^2 + \|\mu\|_2^2\big).$$

(ii) There exists some constant $C_f > 0$ s.t. for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $a \in A$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\lambda \in \mathcal{P}_2(A)$,

$$|f(t,x,\mu,a,\lambda)| \quad \leq \quad C_f\big(1 + |x|^2 + \|\mu\|_2^2\big),$$

and some modulus m_f (i.e. $m_f(\tau) \to 0$ when τ goes to zero) s.t. for all $t, t' \in [0,T]$, $x, x' \in \mathbb{R}^d$, $a \in A$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $\lambda, \lambda' \in \mathcal{P}_2(A)$,

$$|f(t,x,\mu,a,\lambda) - f(t',x',\mu',a,\lambda')| \quad \leq \quad m_f\big(|t-t'| + |x-x'| + W_2(\mu,\mu') + W_2(\lambda,\lambda')\big).$$

The boundedness condition in **(H1')**-**(H2')** of b, σ, f w.r.t. $(a,\lambda) \in A \times \mathcal{P}_2(A)$ is typically satisfied when A is bounded. Under **(H1')**, it is easily seen by standard arguments that

$$\sup_{\alpha \in \mathcal{A}} \mathbb{E}\big[\sup_{t \leq s \leq T} |X_s^{t,\xi}|^2\big] \quad < \quad \infty,$$

for any $t \in [0,T]$, $\xi \in L^2(\mathcal{F}_t; \mathbb{R}^d)$, which shows under the quadratic growth condition of g and f in **(H2')** (uniformly in a) that v and V also satisfy a quadratic growth condition: there exists some positive constant C s.t.

$$\left\{ \begin{array}{lcll} |v(t,\mu)| & \leq & C\big(1 + \|\mu\|_2^2\big), & (t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d), \\ |V(t,\xi)| & \leq & C\big(1 + \mathbb{E}|\xi|^2\big), & (t,\xi) \in [0,T] \times L^2(\mathcal{F}_0; \mathbb{R}^d), \end{array} \right. \tag{6.4}$$

and are thus in particular locally bounded.

We first state a flow continuity property of the marginal distribution of the controlled state process. Indeed, from standard estimates on the state process under **(H1')**, one easily checks (see also Lemma 3.1 in [8]) that there exists some positive constant C, such that for all $\alpha \in \mathcal{A}$, $t, t' \in [0,T$, $t \leq s \leq T, t' \leq s' \leq T$, $\mu = \mathbb{P}_\xi, \mu' = \mathbb{P}_{\xi'} \in \mathcal{P}_2(\mathbb{R}^d)$:

$$\mathbb{E}\big|X_s^{t,\xi} - X_{s'}^{t',\xi'}\big|^2 \quad \leq \quad C\big(1 + \mathbb{E}|\xi|^2 + \mathbb{E}|\xi'|^2\big)\big(|t-t'| + |s-s'| + \mathbb{E}|\xi - \xi'|^2\big),$$

and so from the definition of the 2-Wasserstein distance

$$W_2(\mathbb{P}_s^{t,\mu}, \mathbb{P}_{s'}^{t',\mu'}) \quad \leq \quad C\big(1 + \|\mu\|_2 + \|\mu'\|_2\big)\big(|t-t'|^{\frac{1}{2}} + |s-s'|^{\frac{1}{2}} + W_2(\mu,\mu')\big). \tag{6.5}$$

The next result states the viscosity property of the value function to the Bellman equation as a consequence of the dynamic programming principle (3.10).

Proposition 6.1 *The value function v is a viscosity solution to the Bellman equation* (4.5).

Proof. We first show the continuity of $t \mapsto v(t,.)$ and $V(t,.)$ at $t = T$. For any $(t, \mu = \mathbb{P}_\xi) \in [0,T) \times \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, we have from (6.5)

$$W_2(\mathbb{P}_T^{t,\mu}, \mu) \;\leq\; \Big(\mathbb{E}|X_T^{t,\xi} - \xi|^2\Big)^{\frac{1}{2}} \;\leq\; C(1 + \|\mu\|_2)|T-t|^{\frac{1}{2}}, \tag{6.6}$$

for some positive constant C (independent of t, μ, α). This means that $\mathbb{P}_T^{t,\mu}$ converges to μ in $\mathcal{P}_2(\mathbb{R}^d)$ when $t \nearrow T$, uniformly in $\alpha \in \mathcal{A}$. Now, from the definition of v in (3.8), we have

$$\begin{aligned} |v(t,\mu) - \hat{g}(\mu)| \;&\leq\; \sup_{\alpha\in\mathcal{A}} \int_t^T \big|\hat{f}(s, \mathbb{P}_s^{t,\mu}, \tilde{\alpha}(s,., \mathbb{P}_s^{t,\mu}))\big| ds + \big|\hat{g}(\mathbb{P}_T^{t,\mu}) - \hat{g}(\mu)\big| \\ &\leq\; C(1 + \|\mu\|_2^2)|T-t| \;+\; \sup_{\alpha\in\mathcal{A}} \big|\hat{g}(\mathbb{P}_T^{t,\mu}) - \hat{g}(\mu)\big|, \end{aligned} \tag{6.7}$$

from the growth condition on f in **(H2')**. By the continuity assumption on g together with the growth condition on g in **(H2')**, which allows to use dominated convergence theorem, we deduce from (6.6) that $\hat{g}(\mathbb{P}_T^{t,\mu})$ converges to $\hat{g}(\mu)$ when $t \nearrow T$, uniformly in $\alpha \in \mathcal{A}$. This proves by (6.7) that $v(t,\mu)$ converges to $\hat{g}(\mu)$ when $t \nearrow T$, i.e. $v^*(T,\mu) = v_*(T,\mu) = \hat{g}(\mu) = v(T,\mu)$, and equivalently that $V(T,\xi)$ converges to $\hat{G}(\xi)$ when $t \nearrow T$, i.e. $V^*(T,\xi) = V_*(T,\mu) = \hat{G}(\xi) = V(T,\xi)$,

Let us now prove the viscosity subsolution property of V on $[0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$. Fix $(t_0,\xi_0) \in [0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$, and consider some test function $\Phi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{F}_0;\mathbb{R}^d))$ such that $V^* - \Phi$ has a maximum at (t_0,ξ_0), and w.l.o.g. $V^*(t_0,\xi_0) = \Phi(t_0,\xi_0)$, so that $V^* \leq \Phi$. By definition of $V^*(t_0,\xi_0)$, there exists a sequence $(t_n,\xi_n)_n$ in $[0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$ s.t.

$$(t_n,\xi_n) \;\longrightarrow\; (t_0,\xi_0), \qquad V(t_n,\xi_n) \;\longrightarrow\; V^*(t_0,\xi_0),$$

as n goes to infinity. By continuity of Φ, we have

$$\gamma_n \;:=\; (V - \Phi)(t_n,\xi_n) \;\longrightarrow\; (V^* - \Phi)(t_0,\xi_0) \;=\; 0,$$

and let (h_n) be a strictly positive sequence s.t. $h_n \to 0$ and $\gamma_n/h_n \to 0$. Consider the inverse-lifted function of Φ, namely $\varphi : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ defined by $\varphi(t,\mu) = \Phi(t,\xi)$ for $t \in [0,T]$ and $\mu = \mathbb{P}_\xi \in \mathcal{P}_2(\mathbb{R}^d)$, and recall that $\varphi \in \mathcal{C}_b^{1,2}([0,T] \times \mathcal{P}_2(\mathbb{R}^d))$. Let $\tilde{\alpha}$ be an arbitrary element in $L(\mathbb{R}^d;A)$, and consider the time-independent feedback control $\alpha \in \mathcal{A}$ associated with $\tilde{\alpha}$. From the DPP (3.10) applied to $v(t_n,\mu_n)$, with $\mu_n = \mathbb{P}_{\xi_n}$, we have

$$v(t_n,\mu_n) \;\leq\; \int_{t_n}^{t_n+h_n} \hat{f}(s, \mathbb{P}_s^{t_n,\mu_n}, \tilde{\alpha}) ds \;+\; v(t_n + h_n, \mathbb{P}_{t_n+h_n}^{t_n,\mu_n}).$$

Since $v(t,\mu) = V(t,\xi) \leq V^*(t,\xi) \leq \Phi(t,\xi) = \varphi(t,\mu)$ for all $(t, \mu = \mathbb{P}_\xi) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, this implies

$$\frac{\gamma_n}{h_n} \;\leq\; \frac{1}{h_n} \int_{t_n}^{t_n+h_n} \hat{f}(s, \mathbb{P}_s^{t_n,\mu_n}, \tilde{\alpha}) ds \;+\; \frac{\varphi(t_n + h_n, \mathbb{P}_{t_n+h_n}^{t_n,\mu_n}) - \varphi(t_n,\mu_n)}{h_n}.$$

Applying Itô's formula (4.4) (similarly as in the verification theorem 4.1) to $\varphi(s, \mathbb{P}_s^{t_n,\mu_n})$ between t_n and $t_n + h_n$, we get

$$\frac{\gamma_n}{h_n} \;\leq\; \frac{1}{h_n} \int_{t_n}^{t_n+h_n} \big[\hat{f}(s, \mathbb{P}_s^{t_n,\mu_n}, \tilde{\alpha}) + \frac{\partial\varphi}{\partial t}(s, \mathbb{P}_s^{t_n,\mu_n}) + < \mathcal{L}^{\tilde{\alpha}}\varphi(s, \mathbb{P}_s^{t_n,\mu_n}), \mathbb{P}_s^{t_n,\mu_n} > \big] ds$$

Recall that $W_2(\mu_n,\mu_0) \leq \big(\mathbb{E}|\xi_n-\xi_0|^2\big)^{\frac{1}{2}}$, where $\mu_0 = \mathbb{P}_{\xi_0}$, which shows that $\mu_n \to \mu_0$ in $\mathcal{P}_2(\mathbb{R}^d)$ as n goes to infinity. By the continuity of b,σ,f,φ on their respective domains, the flow continuity property (6.5), we then obtain by sending n to infinity in the above inequality:

$$0 \;\leq\; \hat{f}(t_0,\mu_0,\tilde{\alpha}) + \frac{\partial\varphi}{\partial t}(t_0,\mu_0) + < \mathcal{L}^{\tilde{\alpha}}\varphi(t_0,\mu_0),\mu_0 >,$$

Since $\tilde{\alpha}$ is arbitrary in $L(\mathbb{R}^d;A)$, this shows

$$-\frac{\partial\varphi}{\partial t}(t_0,\mu_0) + H(t_0,\mu_0,\partial_\mu\varphi(t_0,\mu_0),\partial_x\partial_\mu\varphi(t_0,\mu_0)) \;\leq\; 0,$$

and thus at the lifted level:

$$-\frac{\partial\Phi}{\partial t}(t_0,\xi_0) + \mathcal{H}(t_0,\xi_0,D\Phi(t_0,\xi_0),D^2\Phi(t_0,\xi_0)) \;\leq\; 0,$$

which is the required viscosity subsolution property.

We proceed finally with the viscosity supersolution property. Fix $(t_0,\xi_0) \in [0,T) \times L^2(\mathcal{F}_0;\mathbb{R}^d)$, and consider some test function $\Phi \in \mathcal{C}^2([0,T]\times L^2(\mathcal{F}_0;\mathbb{R}^d))$ such that $V_* - \Phi$ has a minimum at (t_0,ξ_0), and w.l.o.g. $V_*(t_0,\xi_0) = \Phi(t_0,\xi_0)$, so that $V_* \geq \Phi$. Again, by definition of $V_*((t_0,\xi_0)$, there exists a sequence $(t_n,\xi_n)_n$ in $[0,T)\times L^2(\mathcal{F};\mathbb{R}^d)$ s.t. $(t_n,\xi_n) \longrightarrow (t_0,\xi_0)$, and $V(t_n,\xi_n) \longrightarrow V_*(t_0,\xi_0)$ as n goes to infinity. We set $\gamma_n := (V-\Phi)(t_n,\xi_n)$, which converges to zero, and we consider a strictly positive sequence (h_n) converging to zero and s.t. γ_n/h_n also converges to zero. Consider the inverse-lifted function of Φ, namely $\varphi \in \mathcal{C}_b^{1,2}([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ defined by $\varphi(t,\mu) = \Phi(t,\xi)$ for $t\in[0,T]$ and $\mu = \mathbb{P}_\xi \in \mathcal{P}_2(\mathbb{R}^d)$. From the DPP (3.10), for each n, and denoting by $\mu_n = \mathbb{P}_{\xi_n} \in \mathcal{P}_2(\mathbb{R}^d)$, there exists $\alpha^n \in \mathcal{A}$ associated to a feedback control $\tilde{\alpha}^n \in Lip([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d);A)$ s.t.

$$v(t_n,\mu_n) + h_n^2 \;\geq\; \int_{t_n}^{t_n+h_n} \hat{f}(s,\mathbb{P}_s^{t_n,\mu_n},\tilde{\alpha}^n)ds \;+\; v(t_n+h_n,\mathbb{P}_{t_n+h_n}^{t_n,\mu_n}).$$

Since $v(t,\mu) = V(t,\xi) \geq V_*(t,\xi) \geq \Phi(t,\xi) = \varphi(t,\mu)$ for all $(t,\mu=\mathbb{P}_\xi) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, this implies

$$\frac{\gamma_n}{h_n} + h_n \;\geq\; \frac{1}{h_n}\int_{t_n}^{t_n+h_n} \hat{f}(s,\mathbb{P}_s^{t_n,\mu_n},\tilde{\alpha}^n)ds \;+\; \frac{\varphi(t_n+h_n,\mathbb{P}_{t_n+h_n}^{t_n,\mu_n}) - \varphi(t_n,\mu_n)}{h_n}.$$

Applying Itô's formula (4.4) to $\varphi(s,\mathbb{P}_s^{t_n,\mu_n})$, we then get

$$\begin{aligned}
\frac{\gamma_n}{h_n} + h_n \;&\geq\; \frac{1}{h_n}\int_{t_n}^{t_n+h_n} \big[\frac{\partial\varphi}{\partial t}(s,\mathbb{P}_s^{t_n,\mu_n}) \;+\; \hat{f}(s,\mathbb{P}_s^{t_n,\mu_n},\tilde{\alpha}^n(s,.,\mathbb{P}_s^{t_n,\mu_n})) \\
&\qquad + \;< \mathcal{L}^{\tilde{\alpha}^n(s,.,\mathbb{P}_s^{t_n,\mu_n})}\varphi(s,\mathbb{P}_s^{t_n,\mu_n}),\mathbb{P}_s^{t_n,\mu_n} > \big]ds \\
&\geq\; \frac{1}{h_n}\int_{t_n}^{t_n+h_n} \Big(\frac{\partial\varphi}{\partial t}(s,\mathbb{P}_s^{t_n,\mu_n}) \\
&\qquad + \inf_{\tilde{\alpha}\in L(\mathbb{R}^d;A)} \big[\hat{f}(s,\mathbb{P}_s^{t_n,\mu_n},\tilde{\alpha}) + < \mathcal{L}^{\tilde{\alpha}}\varphi(s,\mathbb{P}_s^{t_n,\mu_n}),\mathbb{P}_s^{t_n,\mu_n} > \big]\Big)ds.
\end{aligned}$$

By sending n to infinity together with the continuity assumption in **(H1')-(H2')** of b, σ, f, φ, uniformly in $a \in A$, and the flow continuity property (6.5) (which is uniform in $\alpha \in \mathcal{A}$), we get

$$-\frac{\partial \varphi}{\partial t}(t_0, \mu_0) + H(t_0, \mu_0, \partial_\mu \varphi(t_0, \mu_0), \partial_x \partial_\mu \varphi(t_0, \mu_0)) \quad \geq \quad 0,$$

which gives the required viscosity supersolution property of V_*, and ends the proof. □

We finally turn to comparison principle (hence uniqueness result) for the Bellman equation (4.5) (or (6.1)), hence equivalently for the lifted Bellman equation (6.2), which shall follow from comparison results for second order Hamilton-Jacobi equations in separable Hilbert space stated in [29], see also [16]. We shall assume that the σ-algebra $\mathcal{F}_0$ is countably generated upto null sets, which ensures that the Hilbert space $L^2(\mathcal{F}_0; \mathbb{R}^d)$ is separable, see [15], p. 92. This is satisfied for example when $\mathcal{F}_0$ is the Borel σ-algebra of a canonical space Ω_0 of continuous functions on $\mathbb{R}_+$, in which case, $\mathcal{F}_0 = \vee_{s \geq 0} \mathcal{F}_s^{B^0}$, where $(\mathcal{F}_s^{B^0})$ is the canonical filtration on Ω_0, and it is then known that $\mathcal{F}_0$ is countably generated, see for instance Exercise 4.21 in Chapter 1 of [34].

Proposition 6.2 *Let u and w be two functions defined on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ satisfying a quadratic growth condition such that u (resp. w) is an upper (resp. lower) semicontinuous viscosity subsolution (resp. supersolution) to* (4.5)*. Then $u \leq w$. Consequently, the value function v is the unique viscosity solution to the Bellman equation* (4.5) *satisfying a quadratic growth condition* (6.4)*.*

Proof. In view of our definition 6.1 of viscosity solution, we have to show a comparison principle for viscosity solutions to the lifted Bellman equation (6.2). We use the comparison principle proved in Theorem 3.50 in [16] and only need to check that the hypotheses of this theorem are satisfied in our context for the lifted Hamiltonian $\mathcal{H}$ defined in (6.3). Notice that the lifted Bellman equation (6.2) is a bounded equation in the terminology of [16] (see their section 3.3.1) meaning that there is no linear dissipative operator on $L^2(\mathcal{F}_0; \mathbb{R}^d)$ in the equation. Therefore, the notion of B-continuity reduces to the standard notion of continuity in $L^2(\mathcal{F}_0; \mathbb{R}^d)$ since one can take for B the identity operator. Their Hypothesis 3.44 follows from the uniform continuity of b, σ, and f in **(H1')-(H2')**. Hypothesis 3.45 is immediately satisfied since there is no discount factor in our equation, i.e. $\mathcal{H}$ does not depend on V but only on its derivatives. The monotonicity condition in $Q \in S(L^2(\mathcal{F}_0; \mathbb{R}^d))$ of $\mathcal{H}$ in Hypothesis 3.46 is clearly satisfied. Hypothesis 3.47 holds directly when dealing with bounded equations. Hypothesis 3.48 is obtained from the Lipschitz condition of b, σ in **(H1')**, and the uniform continuity condition on f in **(H2')**, while Hypothesis 3.49 follows from the quadratic growth condition of σ in **(H1')**. One can then apply Theorem 3.50 in [16] and conclude that comparison principle holds for the Bellman equation (6.2), hence for the Bellman equation (4.5). □

References

[1] Ahmed N.U. and X. Ding (2001): "Controlled McKean-Vlasov equation", *Communications on Applied Analysis*, **5**, 183-206.

[2] Ambrosio, L., Gigli N. and G. Savaré (2005): "Gradient Flows in Metric Spaces and in the Space of Probability Measures", Lectures in Mathematics, Birkhäuser Verlag, Basel.

[3] Andersson D. and B. Djehiche (2010): "A maximum principle for SDEs of mean-field type", *Applied Mathematics and Optimization*, **63**, 341-356.

[4] Bensoussan A., Frehse J. and P. Yam (2015): "The Master equation in mean-field theory", *Journal de Mathématiques Pures et Appliquées*, **103**(6), 1441-1474.

[5] Bensoussan A., Frehse J. and P. Yam (2015): "On the interpretation of the Master equation", http://arxiv.org/abs/1503.07754

[6] Björk T. and A. Murgoci (2008): "A general theory of Markovian time inconsistent stochastic control problems", Preprint.

[7] Buckdahn R., Djehiche B. and J. Li (2011): "A general maximum principle for SDEs of mean-field type", *Applied Mathematics and Optimization*, **64**(2), 197-216.

[8] Buckdahn R., Li J., Peng S. and C. Rainer (2014): "Mean-field stochastic differential equations and associated PDEs", http://arxiv.org/abs/1407.1215, to appear in the *Annals of Probability*.

[9] Cardaliaguet P. (2012): "Notes on mean field games", Notes from P.L. Lions lectures at Collège de France, https://www.ceremade.dauphine.fr/cardalia/MFG100629.pdf

[10] Carmona R. and F. Delarue (2014): "Forward-backward Stochastic Differential Equations and Controlled McKean Vlasov Dynamics", to appear in the *Annals of Probability*.

[11] Carmona R. and F. Delarue (2014): "The Master equation for large population equilibriums", D. Crisan et al. (eds.), Stochastic Analysis and Applications 2014, Springer Proceedings in Mathematics & Statistics 100.

[12] Carmona R., Delarue F. and A. Lachapelle (2013): "Control of McKean-Vlasov dynamics versus mean field games", *Mathematics and Financial Economics*, **7**, 131-166.

[13] Carmona R., Fouque J.P. and L. Sun (2014): "Mean field games and systemic risk", to appear in *Communications in Mathematical Sciences*.

[14] Chassagneux J.F., Crisan D. and F. Delarue (2015): "A probabilistic approach to classical solutions of the master equation for large population equilibria", http://arxiv.org/pdf/1411.3009.pdf

[15] Doob J.L. (1994): *Measure Theory*, Berlin Heidelberg New York, Springer.

[16] Fabbri G., Gozzi F. and A. Swiech (2015): *Stochastic optimal control in infinite dimension: dynamic programming and HJB equations* (with Chapter 6 by M. Fuhrman and G. Tessitore), http://people.math.gatech.edu/ swiech/book.version11-10-2015.pdf

[17] Feng J. and M. Katsoulakis (2009): "A comparison principle for Hamilton-Jacobi equations related to controlled gradient flows in infinite dimensions", *Arch. Ration. Mech. Anal.*, **192**(2), 275-310.

[18] Fleming W. H. and H. M. Soner (2006): *Controlled Markov Processes and Viscosity Solutions*, 2nd edition, Springer-Verlag.

[19] Gangbo W, Nguyen T. and A. Tudorascu (2008): "Hamilton-Jacobi equations in the Wasserstein space, Methods and Applications of Analysis", **15**(2), 155-184.

[20] Gangbo W. and A. Swiech (2015): "Metric viscosity solutions of Hamilton-Jacobi equations", to appear in *Calculus of Variations and Partial Differential Equations*.

[21] Huang M., Caines P., and R. Malhamé (2006): "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle", *Communication in Information and Systems*, **6**(3), 221-252.

[22] Jourdain B., Méléard S. and W. Woyczynski (2008): "Nonlinear SDEs driven by Lévy processes and related PDEs", *ALEA, Latin American Journal of Probability*, **4**, 1-29.

[23] Kac M. (1956): "Foundations of kinetic theory", in Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, **3**, 171-197.

[24] Lasry J.M. and P.L. Lions (2007): "Mean-field games", *Japanese Journal of Mathematics*, **2**, 229-260.

[25] Laurière M. and O. Pironneau (2014): "Dynamic programming for mean-field type control", CRAS, **352**(9), 707-713.

[26] Li D. and X.Y. Zhou (2000): "Continuous-time mean-variance portfolio selection: a stochastic LQ framework", *Applied Mathematics and Optimization*, **42**, 19-33.

[27] Lions P.L. (1988): "Viscosity solutions of fully nonlinear second-order equations and optimal control in infinite dimension. Part I: the case of bounded stochastic evolutions", *Acta Mathematica*, **161**, 243-278.

[28] Lions P.L. (1989): "Viscosity solutions of fully nonlinear second-order equations and optimal control in infinite dimension. Part II: Optimal control of Zakai's equation", *Stochastic partial differential equations and applications, Lect. Notes in Math*, **1390**, Springer.

[29] Lions P.L. (1989): "Viscosity solutions of fully nonlinear second-order equations and optimal control in infinite dimension. Part III: Uniqueness of viscosity solutions for general second-order equations", *Journal of Functional Analysis*, **86**(1), 1-18.

[30] Lions P.L. (2012): *Cours au Collège de France: Théorie des jeux à champ moyens*, audio conference 2006-2012.

[31] McKean H.P. (1967): "Propagation of chaos for a class of nonlinear parabolic equations", Lecture Series in Differential Equations, **7**, 41-57.

[32] Pham H. (2009): *Continuous-time stochastic control and applications with financial applications*, Series Stochastic Modeling and Applied Probability, **Vol. 61**, Springer.

[33] Pham H. and X. Wei (2015): "Discrete time McKean-Vlasov control problem: a dynamic programming approach", http://arxiv.org/pdf/1511.09273v1.pdf

[34] Revuz D. and M. Yor (1999): *Continuous Martingales and Brownian Motion*, 3rd edition. New York, Berlin: Springer.

[35] Sznitman A.S. (1989): Topics in propagation of chaos, in *Lecture Notes in Mathematics*, Springer, **1464**, 165-251.

[36] Villani C. (2009): *Optimal transport, old and new*. Springer Verlag.

[37] Yong J. (2013): "A linear-quadratic optimal control problem for mean-field stochastic differential equations", *SIAM Journal on Control and Optimization*, **51**(4), 2809-2838.