HAL
open science

Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics

Huyên Pham, Xiaoli Wei

To cite this version:

Huyên Pham, Xiaoli Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics . 2016. hal-01302289v1

HAL Id: hal-01302289

https://hal.science/hal-01302289v1

Preprint submitted on 13 Apr 2016 (v1), last revised 4 Jan 2017 (v2)

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.

Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics *

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

April 13, 2016

Abstract

We study the optimal control of general stochastic McKean-Vlasov equation. Such problem is motivated originally from the asymptotic formulation of cooperative equilibrium for a large population of particles (players) in mean-field interaction under common noise. Our first main result is to state a dynamic programming principle for the value function in the Wasserstein space of probability measures, which is proved from a flow property of the conditional law of the controlled state process. Next, by relying on the notion of differentiability with respect to probability measures due to P.L. Lions [32], and Itô's formula along a flow of conditional measures, we derive the dynamic programming Hamilton-Jacobi-Bellman equation, and prove the viscosity property together with a uniqueness result for the value function. Finally, we solve explicitly the linear-quadratic stochastic McKean-Vlasov control problem and give an application to an interbank systemic risk model with common noise.

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

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

*This work is part of the ANR project CAESARS (ANR-15-CE05-0024).

1 Introduction

Let us consider the controlled McKean-Vlasov dynamics in $\mathbb{R}^d$ given by

$$dX_t = b(X_t, \mathbb{P}_{X_t}^{W^0}, \alpha_t)dt + \sigma(X_t, \mathbb{P}_{X_t}^{W^0}, \alpha_t)dB_t + \sigma_0(X_t, \mathbb{P}_{X_t}^{W^0}, \alpha_t)dW_t^0, \tag{1.1}$$

where B, W^0 are two independent Brownian motions on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{P}_{X_t}^{W^0}$ denotes the conditional distribution of X_t given W^0 (or equivalently given $\mathcal{F}_t^0$ where $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t\geq 0}$ is the natural filtration generated by W^0), valued in $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$, and the control α is an $\mathbb{F}^0$-progressive process valued in some Polish space $\mathbf{A}$. When there is no control, the dynamics (1.1) is sometimes called stochastic McKean-Vlasov equation (see [19]), where the term "stochastic" refers to the presence of the random noise caused by the Brownian motion W^0 w.r.t. a McKean-Vlasov equation when $\sigma_0 = 0$, and for which coefficients depend on the (deterministic) marginal distribution $\mathbb{P}_{X_t}$. One also uses the terminology conditional mean-field stochastic differential equation (CMFSDE) to emphasize the dependence of the coefficients on the conditional law with respect to the random noise, and such CMFSDE was studied in [16], and more generally in [8]. In this context, the control problem is to minimize over α a cost functional of the form:

$$J(\alpha) = \mathbb{E}\Big[\int_0^T f(X_t, \mathbb{P}_{X_t}^{W^0}, \alpha_t)dt + g(X_T, \mathbb{P}_{X_T}^{W^0})\Big]. \tag{1.2}$$

The motivation and applications for the study of such stochastic control problem, refereed alternatively to as control of stochastic McKean-Vlasov dynamics, or stochastic control of conditional McKean-Vlasov equation, comes mainly from the *McKean-Vlasov control problem with common noise*, that we briefly describe now: we consider a system of controlled individuals (referred also to as particles or players) in mutual interaction, where the dynamics of the state process X^i of player $i \in \{1, \ldots, N\}$ is governed by

$$dX_t^i = \tilde{b}(X_t^i, \bar{\rho}_t^N, \tilde{\alpha}_t^i)dt + \tilde{\sigma}(X_t^i, \bar{\rho}_t^N, \tilde{\alpha}_t^i)dB_t^i + \tilde{\sigma}_0(X_t^i, \bar{\rho}_t^N, \tilde{\alpha}_t^i)dW_t^0.$$

Here, the Wiener process W^0 accounts for the common random environment in which all the individuals evolve, called common noise, and $B^1, \ldots, B^N$ are independent Brownian motions, independent of W^0, called idiosyncratic noises. The particles are in interaction of mean-field type in the sense that any any time t, the coefficients $\tilde{b}$, $\tilde{\sigma}$, $\tilde{\sigma}_0$ of their state process depend on the empirical distribution of all individual states

$$\bar{\rho}_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}.$$

The processes $(\tilde{\alpha}_t^i)_{t\geq 0}$, $i = 1, \ldots, N$, are in general progressively measurable w.r.t. the filtration generated by $B^1, \ldots, B^N, W^0$, valued in some subset A of a Euclidian space, and represent the control processes of the players with cost functionals:

$$J^i(\tilde{\alpha}^1, \ldots, \tilde{\alpha}^n) = \mathbb{E}\Big[\int_0^T \tilde{f}(X_t^i, \bar{\rho}_t^N, \tilde{\alpha}_t^i)dt + g(X_T^i, \bar{\rho}_T^N)\Big].$$

For this N-player stochastic differential game, one looks for equilibriums, and different notions may be considered. Classically, the search for a consensus among the players leads

to the concept of Nash equilibrium where each player minimizes its own cost functional, and the goal is to find a N-tuple control strategy for which there is no interest for any player to leave from this consensus state. The asymptotic formulation of this Nash equilibrium when the number of players N goes to infinity leads to the (now well-known) theory of mean-field games (MFG) pioneered in the works by Lasry and Lions [28], and Huang, Malhamé and Caines [25]. In this framework, the analysis is reduced to the problem of a single representative player in interaction with the theoretical distribution of the whole population by the propagation of chaos phenomenon, who first solves a control problem by freezing a probability law in the coefficients of her/his state process and cost function, and then has to find a fixed point probability measure that matches the distribution of her/his optimal state process. The case of MFG with common noise has been recently studied in [1] and [14]. Alternatively, one may take the point of view of a center of decision (or social planner), which decides of the strategies for all players, with the goal of minimizing the global cost to the collectivity. This leads to the concept of Pareto or cooperative equilibrium whose asymptotic formulation is reduced to the optimal control of McKean-Vlasov dynamics for a representative player. More precisely, given the symmetry of the set-up, when the social planner chooses the same control policy for all the players in feedback form: $\tilde{\alpha}_t^i = \tilde{\alpha}(t, X_t^i, \bar{\rho}_t^N)$, $i = 1, \ldots, N$, for some deterministic function $\tilde{\alpha}$ depending upon time, private state of player, and the empirical distribution of all players, then the theory of propagation of chaos implies that, in the limit $N \to \infty$, the particles X^i become asymptotically independent conditionally on the random environment W^0, and the empirical measure $\bar{\rho}_t^N$ converge to the distribution $\mathbb{P}_{X_t}^{W^0}$ of X_t given W^0, and X is governed by the (stochastic) McKean-Vlasov equation:

$$\begin{aligned} dX_t &= \tilde{b}(X_t, \mathbb{P}_{X_t}^{W^0}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}^{W^0}))dt + \tilde{\sigma}(X_t, \mathbb{P}_{X_t}^{W^0}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}^{W^0}))dB_t \\ &\quad + \tilde{\sigma}_0(X_t, \mathbb{P}_{X_t}^{W^0}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}^{W^0}))dW_t^0, \end{aligned}$$

for some Brownian motion B independent of W^0. The objective of the representative player for the Pareto equilibrium becomes the minimization of the functional

$$J(\tilde{\alpha}) = \mathbb{E}\Big[\int_0^T \tilde{f}(X_t, \mathbb{P}_{X_t}^{W^0}, \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t}^{W^0}))dt + g(X_T, \mathbb{P}_{X_T}^{W^0})\Big]$$

over the class of feedback controls $\tilde{\alpha}$. We refer to [13] for a detailed discussion of the differences between the nature and solutions to the MFG and optimal control of McKean-Vlasov dynamics related respectively to the notions of Nash and Pareto equilibrium. Notice that in this McKean-Vlasov control formulation, the control $\tilde{\alpha}$ is of feedback (also called closed-loop) form both w.r.t. the state process X_t, and its conditional law process $\mathbb{P}_{X_t}^{W^0}$, which is $\mathbb{F}^0$-adapted. More generally, we can consider semi-feedback control $\alpha(t, x, \omega^0)$, in the sense that it is of closed-loop form w.r.t. the state process X_t, but of open-loop form w.r.t. the common noise W^0. In other words, one can consider random field control $\mathbb{F}^0$-progressive control process $\alpha = \{\alpha_t(x), x \in \mathbb{R}^d\}$, which may be viewed equivalently as processes valued in some functional space $\mathbf{A}$ on $\mathbb{R}^d$, typically a closed subset of the Polish space $C(\mathbb{R}^d, A)$, of continuous functions from $\mathbb{R}^d$ into some Euclidian space A. In this case, we are in the framework (1.1)-(1.2) with $b(x, \mu, a) = \tilde{b}(x, \mu, a(x))$, $\sigma(x, \mu, a) = \tilde{\sigma}(x, \mu, a(x))$, $\sigma_0(x, \mu, a) = \tilde{\sigma}_0(x, \mu, a(x))$, $f(x, \mu, a) = \tilde{f}(x, \mu, a(x))$, for $(x, \mu, a) \in \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbf{A}$.

We also mention that partial observation control problem arises as a particular case of our stochastic control framework (1.1)-(1.2): Indeed, let us consider a controlled process with dynamics

$$d\bar{X}_t \;=\; \bar{b}(\bar{X}_t,\alpha_t)dt + \bar{\sigma}(\bar{X}_t,\alpha_t)dB_t + \bar{\sigma}_0(\bar{X}_t,\alpha_t)dB^0_t,$$

where B, B^0 are two independent Brownian motions on some physical probability space $(\Omega,\mathcal{F},\mathbb{Q})$, and the signal control process can only be observed through W^0 given by

$$dW^0_t \;=\; h(\bar{X}_t)dt + dB^0_t.$$

The control process α is progressively measurable w.r.t. the observation filtration $\mathbb{F}^0$ generated by W^0, valued typically in some Euclidian space A, and the cost functional to minimize over α is

$$J(\alpha) \;=\; \mathbb{E}^{\mathbb{Q}}\Big[\int_0^T \bar{f}(\bar{X}_t,\alpha_t)dt + \bar{g}(\bar{X}_T)\Big].$$

By considering the process Z via

$$Z_t^{-1} \;=\; \exp\big(-\int_0^t h(\bar{X}_s)dB^0_s - \frac{1}{2}\int_0^t |h(\bar{X}_s)|^2 ds\big),\; 0\le t\le T,$$

the process Z^{-1} is (under suitable integrability conditions on h) a martingale under $\mathbb{Q}$, and by Girsanov's theorem, this defines a probability measure $\mathbb{P}(d\omega) = Z_T^{-1}(\omega)\mathbb{Q}(d\omega)$, called reference probability measure, under which the pair (B, W^0) is a Brownian motion. We then see that the partial observation control problem can be recast into the framework (1.1)-(1.2) of a particular stochastic McKean-Vlasov control problem with $X = (\bar{X}, Z)$ governed by

$$\begin{aligned} d\bar{X}_t &= \big(\bar{b}(\bar{X}_t,\alpha_t) - \bar{\sigma}_0(\bar{X}_t,\alpha_t)h(\bar{X}_t)\big)dt + \bar{\sigma}(\bar{X}_t,\alpha_t)dB_t + \bar{\sigma}_0(\bar{X}_t,\alpha_t)dW^0_t,\\ dZ_t &= Z_t h(\bar{X}_t)dW^0_t, \end{aligned}$$

and a cost functional rewritten under the reference probability measure from Bayes formula as

$$J(\alpha) \;=\; \mathbb{E}\Big[\int_0^T Z_t\bar{f}(\bar{X}_t,\alpha_t)dt + Z_T\bar{g}(\bar{X}_T)\Big].$$

The optimal control of McKean-Vlasov dynamics is a rather new problem with an increasing interest in the field of stochastic control problem. It has been studied by maximum principle methods in [3], [7], [11] for state dynamics depending upon marginal distribution, and in [16], [8] for conditional McKean-Vlasov dynamics. This leads to a characterization of the solution in terms of an adjoint backward stochastic differential equation (BSDE) coupled with a forward SDE, and we refer to [17] for a theory of BSDE of McKean-Vlasov type. Alternatively, dynamic programming approach for the control of McKean-Vlasov dynamics has been considered in [5], [6], [29] for specific McKean-Vlasov dynamics and under a density assumption on the probability law of the state process, and then analyzed in a

general framework in [33] (without noise W^0), where the problem is reformulated into a deterministic control problem involving the marginal distribution process.

The aim of this paper is to develop the dynamic programming method for stochastic McKean-Vlasov equation in a general setting. For this purpose, a key step is to show the flow property of the conditional distribution $\mathbb{P}_{X_t}^{W^0}$ of the controlled state process X_t given the noise W^0. Then, by reformulating the original control problem into a stochastic control problem where the conditional law $\mathbb{P}_{X_t}^{W^0}$ is the sole controlled state variable driven by the random noise W^0, and by showing the continuity of the value function in the Wasserstein space of probability measures, we are able to prove a dynamic programming principle (DPP) for our stochastic McKean-Vlasov control problem. Next, for exploiting the DPP, we use a notion of differentiability with respect to probability measures introduced by P.L. Lions in his lectures at the Collège de France [32], and detailed in the notes [10]. This notion of derivative is based on the lifting of functions defined on the Hilbert space of square integrable random variables distributed according to the "lifted" probability measure. By combining with a special Itô's chain rule for flows of conditional distributions, we derive the dynamic programming Bellman equation for stochastic McKean-Vlasov control problem, which is a fully nonlinear second order partial differential equation (PDE) in the infinite dimensional Wasserstein space of probability measures. By adapting standard arguments to our context, we prove the viscosity property of the value function to the Bellman equation from the dynamic programming principle. To complete our PDE characterization of the value function with a uniqueness result, it is convenient to work in the lifted Hilbert space of square integrable random variables instead of the Wasserstein metric space of probability measures, in order to rely on the general results for viscosity solutions of second order Hamilton-Jacobi-Bellman equations in separable Hilbert spaces, see [30], [31], [21]. We also state a verification theorem which is useful for getting an analytic feedback form of the optimal control when there is a smooth solution to the Bellman equation. Finally, we apply our results to the class of linear-quadratic (LQ) stochastic McKean-Vlasov control problem for which one can obtain explicit solutions, and we illustrate with an example arising from an interbank systemic risk model.

The outline of the paper is organized as follows. Section 2 formulates the stochastic McKean-Vlasov control problem, and fix the standing assumptions. Section 3 is devoted to the proof and statement of the dynamic programming principle. We prove in Section 4 the viscosity characterization of the value function to the Bellman equation, and the last Section 5 presents the application to the LQ framework with explicit solutions.

2 Conditional McKean-Vlasov control problem

Let us fix some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ assumed of the form $(\Omega^0 \times \Omega^1, \mathcal{F}^0 \otimes \mathcal{F}^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$, where $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$ supports a m-dimensional Brownian motion W^0, and $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$ supports a n-dimensional Brownian motion B. So an element $\omega \in \Omega$ is written as $\omega = (\omega^0, \omega^1) \in \Omega^0 \times \Omega^1$, and we extend canonically W^0 and W on Ω by setting $W^0(\omega^0, \omega^1) := W^0(\omega^0)$, $W(\omega^0, \omega^1) := W(\omega^1)$, and extend similarly on Ω any random variable on Ω^0 or Ω^1. We assume that $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$ is in the form $\Omega^1 = \tilde{\Omega}^1 \times \Omega'^1$, $\mathcal{F}^1 = \mathcal{G} \otimes \mathcal{F}'^1$, $\mathbb{P}^1 = \tilde{\mathbb{P}}^1 \otimes \mathbb{P}'^1$, where

$\tilde{\Omega}^1$ is a Polish space, $\mathcal{G}$ its Borel σ-algebra, $\tilde{\mathbb{P}}^1$ an atomless probability measure on $(\tilde{\Omega}^1, \mathcal{G})$, while $(\Omega'^1, \mathcal{F}'^1, \mathbb{P}'^1)$ supports B. We denote by $\mathbb{E}^0$ (resp. $\mathbb{E}^1$ and $\tilde{\mathbb{E}}^1$) the expectation under $\mathbb{P}^0$ (resp. $\mathbb{P}^1$ and $\tilde{\mathbb{P}}^1$), by $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ the natural filtration generated by W^0 (and w.l.o.g. we assume that $\mathcal{F}^0 = \mathcal{F}^0_\infty$), and by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ the natural filtration generated by W^0, B, augmented with the independent σ-algebra $\mathcal{G}$. We denote by $\mathcal{P}_2(\mathbb{R}^d)$ the set probability measures μ on $\mathbb{R}^d$, which are square integrable, i.e. $\|\mu\|_2^2 := \int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty$. For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we denote by $L^2_\mu(\mathbb{R}^q)$ the set of measurable functions $\varphi : \mathbb{R}^d \to \mathbb{R}^q$, which are square integrable with respect to μ, by $L^2_{\mu\otimes\mu}(\mathbb{R}^q)$ the set of measurable functions $\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^q$, which are square integrable with respect to the product measure $\mu \otimes \mu$, and we set

$$\mu(\varphi) \ := \ \int_{\mathbb{R}^d} \varphi(x)\, \mu(dx), \qquad \mu \otimes \mu(\psi) \ := \ \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x, x')\mu(dx)\mu(dx').$$

We also define $L^\infty_\mu(\mathbb{R}^q)$ (resp. $L^\infty_{\mu\otimes\mu}(\mathbb{R}^q)$) as the subset of elements $\varphi \in L^2_\mu(\mathbb{R}^q)$ (resp. $L^2_{\mu\otimes\mu}(\mathbb{R}^q)$) which are bounded μ (resp. $\mu \otimes \mu$) a.e., and $\|\varphi\|_\infty$ is their essential supremum. We denote by $L^2(\mathcal{G}; \mathbb{R}^d)$ (resp. $L^2(\mathcal{F}_t; \mathbb{R}^d)$) the set of $\mathbb{R}^d$-valued square integrable random variables on $(\tilde{\Omega}^1, \mathcal{G}, \tilde{\mathbb{P}}^1)$ (resp. on $(\Omega, \mathcal{F}_t, \mathbb{P})$). For any random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$, we denote by $\mathbb{P}_X$ its probability law (or distribution) under $\mathbb{P}$, and we know that $\mathcal{P}_2(\mathbb{R}^d) = \{\mathbb{P}_\xi = \tilde{\mathbb{P}}^1_\xi : \xi \in L^2(\mathcal{G}; \mathbb{R}^d)\}$ since $(\tilde{\Omega}^1, \mathcal{G}, \tilde{\mathbb{P}}^1)$ is Polish and atomless (we say that $\mathcal{G}$ is rich enough). We often write $\mathcal{L}(\xi) = \mathbb{P}_\xi = \tilde{\mathbb{P}}^1_\xi$ for the law of $\xi \in L^2(\mathcal{G}; \mathbb{R}^d)$. The space $\mathcal{P}_2(\mathbb{R}^d)$ is a metric space equipped with the 2-Wasserstein distance

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

• *Admissible controls.* We are given a Polish set $\mathbf{A}$ equipped with the distance $d_{_A}$, satisfying w.l.o.g. $d_{_A} < 1$, representing the control set, and we denote by $\mathcal{A}$ the set of $\mathbb{F}^0$-progressive processes α valued in $\mathbf{A}$. Notice that $\mathcal{A}$ is a separable metric space endowed with the Krylov distance $\Delta(\alpha, \beta) = \mathbb{E}^0[\int_0^T d_{_A}(\alpha_t, \beta_t)dt]$. We denote by $\mathcal{B}_\mathcal{A}$ the Borel σ-algebra of $\mathcal{A}$.

• *Controlled stochastic McKean-Vlasov dynamics.* For $(t, \xi) \in [0, T] \times L^2(\mathcal{F}_t; \mathbb{R}^d)$, and given $\alpha \in \mathcal{A}$, we consider the stochastic McKean-Vlasov equation:

$$\left\{ \begin{array}{rcl} dX_s &=& b(X_s, \mathbb{P}^{W^0}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}^{W^0}_{X_s}, \alpha_s)dB_s \\ && \qquad + \sigma_0(X_s, \mathbb{P}^{W^0}_{X_s}, \alpha_s)dW^0_s, \quad t \leq s \leq T, \\ X_t &=& \xi. \end{array} \right. \tag{2.1}$$

Here, $\mathbb{P}^{W^0}_{X_s}$ denotes the regular conditional distribution of X_s given $\mathcal{F}^0$, and its realization at some $\omega^0 \in \Omega^0$ also reads as the law under $\mathbb{P}^1$ of the random variable $X_s(\omega^0, .)$ on $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$, i.e. $\mathbb{P}^{W^0}_{X_s}(\omega^0) = \mathbb{P}^1_{X_s(\omega^0,.)}$. The coefficients b, σ, σ_0 are measurable functions from $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbf{A}$ into $\mathbb{R}^d$, respectively $\mathbb{R}^{d\times n}$, $\mathbb{R}^{d\times m}$, and satisfy the condition:

(H1)

(i) There exists some positive constant C s.t. for all $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \times\mathcal{P}_2(\mathbb{R}^d)$, and $a \in \mathbf{A}$,

$$\begin{aligned} & |b(x,\mu,a) - b(x',\mu',a)| + |\sigma(x,\mu,a) - \sigma(x',\mu',a)| + |\sigma_0(x,\mu,a) - \sigma_0(x',\mu',a)| \\ \leq \; & C\Big(|x-x'| + \mathcal{W}_2(\mu,\mu')\Big), \end{aligned}$$

and

$$|b(0,\delta_0,a)| + |\sigma(0,\delta_0,a)| + |\sigma_0(0,\delta_0,a)| \;\leq\; C.$$

(ii) For all $(x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the functions $a \mapsto b(x,\mu,a)$, $\sigma(x,\mu,a)$, $\sigma_0(x,\mu,a)$ are continuous on $\mathbf{A}$.

Under **(H1)**(i), there exists a unique solution to (2.1) (see e.g. [27]), denoted by $\{X_s^{t,\xi,\alpha}, t \leq s \leq T\}$, which is $\mathbb{F}$-adapted, and satisfies the square-integrability condition:

$$\mathbb{E}\Big[\sup_{t\leq s\leq T} |X_s^{t,\xi,\alpha}|^2\Big] \;\leq\; C\Big(1 + \mathbb{E}|\xi|^2\Big) \;<\; \infty, \tag{2.2}$$

for some positive constant C independent of α. We shall sometimes omit the dependence of $X^{t,\xi} = X^{t,\xi,\alpha}$ on α when there is no ambiguity. Since $\{X_s^{t,\xi}, t \leq s \leq T\}$ is $\mathbb{F}$-adapted, and W^0 is a $(\mathbb{P},\mathbb{F})$-Wiener process, we notice that $\mathbb{P}^{W^0}_{X_s^{t,\xi}}(dx) = \mathbb{P}[X_s^{t,\xi} \in dx|\mathcal{F}^0] = \mathbb{P}[X_s^{t,\xi} \in dx|\mathcal{F}_s^0]$, and thus $\{\mathbb{P}^{W^0}_{X_s^{t,\xi}}, t \leq s \leq T\}$ admits an $\mathbb{F}^0$-progressive modification (see e.g. Theorem 2.24 in [4]), that will be identified with itself in the sequel, valued in $\mathcal{P}_2(\mathbb{R}^d)$, and we have for any $\varphi \in \mathcal{B}_2(\mathbb{R}^d)$, the set of Borel-measurable functions on $\mathbb{R}^d$ with quadratic growth:

$$\mathbb{P}^{W^0}_{X_s^{t,\xi}}(\varphi) \;=\; \mathbb{E}\Big[\varphi(X_s^{t,\xi})\big|\mathcal{F}^0\Big] \;=\; \mathbb{E}\Big[\varphi(X_s^{t,\xi})\big|\mathcal{F}_s^0\Big], \quad t \leq s \leq T. \tag{2.3}$$

• *Cost functional and value function.* We are given a running cost function f defined on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbf{A}$, and a terminal cost function g defined on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, assumed to satisfy the condition

(H2)

(i) There exists some positive constant C s.t. for all $(x,\mu,a) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbf{A}$,

$$|f(x,\mu,a)| + |g(x,\mu)| \;\leq\; C\big(1 + |x|^2 + \|\mu\|_2^2\big).$$

(ii) The functions f, g are continuous on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbf{A}$, resp. on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, and satisfy the local Lipschitz condition, uniformly w.r.t. $\mathbf{A}$: there exists some positive constant C s.t. for all $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in \mathbf{A}$,

$$\begin{aligned} & |f(x,\mu,a) - f(x',\mu',a)| + |g(x,\mu) - g(x',\mu')| \\ \leq \; & C(1 + |x| + |x'| + \|\mu\|_2 + \|\mu'\|_2)\big(|x-x'| + \mathcal{W}_2(\mu,\mu')\big). \end{aligned}$$

We then consider the cost functional:

$$J(t,\xi,\alpha) \;:=\; \mathbb{E}\Big[\int_t^T f\big(X_s^{t,\xi}, \mathbb{P}^{W^0}_{X_s^{t,\xi}}, \alpha_s\big)ds + g\big(X_T^{t,\xi}, \mathbb{P}^{W^0}_{X_T^{t,\xi}}\big)\Big],$$

which is well-defined and finite for all $(t,\xi,\alpha) \in [0,T] \times L^2(\mathcal{G};\mathbb{R}^d) \times \mathcal{A}$, and we define the value function of the conditional McKean-Vlasov control problem as

$$v(t,\xi) \quad := \quad \inf_{\alpha\in\mathcal{A}} J(t,\xi,\alpha), \qquad (t,\xi) \in [0,T] \times L^2(\mathcal{G};\mathbb{R}^d). \tag{2.4}$$

From the estimate (2.2) and the growth condition in **(H2)**(i), it is clear that v also satisfies a quadratic growth condition:

$$|v(t,\xi)| \quad \leq \quad C\big(1+\mathbb{E}|\xi|^2\big), \quad \forall \xi \in L^2(\mathcal{G};\mathbb{R}^d). \tag{2.5}$$

Our goal is to characterize the value function v as solution of a partial differential equation by means of a dynamic programming approach.

3 Dynamic programming

The aim of this section is to prove the dynamic programming principle (DPP) for the value function v in (2.4) of the conditional McKean-Vlasov control problem.

3.1 Flow properties

We shall assume that $(\Omega^0, W^0, \mathbb{P}^0)$ is the canonical space, i.e. $\Omega^0 = C(\mathbb{R}_+,\mathbb{R}^m)$, the set of continuous functions from $\mathbb{R}_+$ into $\mathbb{R}^m$, W^0 is the canonical process, and $\mathbb{P}^0$ the Wiener measure. Following [18], we introduce the class of shifted control processes constructed by concatenation of paths: for $\alpha \in \mathcal{A}$, $(t,\bar{\omega}^0) \in [0,T] \times \Omega^0$, we set

$$\alpha_s^{t,\bar{\omega}^0}(\omega^0) \quad := \quad \alpha_s(\bar{\omega}^0 \otimes_t \omega^0), \quad (s,\omega^0) \in [0,T] \times \Omega^0,$$

where $\bar{\omega}^0 \otimes_t \omega^0$ is the element in Ω^0 defined by

$$\bar{\omega}^0 \otimes_t \omega^0(s) \quad := \quad \bar{\omega}^0(s)1_{s<t} + \big(\bar{\omega}^0(t) + \omega^0(s) - \omega^0(t)\big)1_{s\geq t}.$$

We notice that for fixed $(t,\bar{\omega}^0)$, the process $\alpha^{t,\bar{\omega}^0}$ lies in $\mathcal{A}_t$, the set of elements in $\mathcal{A}$ which are independent of $\mathcal{F}_t^0$ under $\mathbb{P}^0$. For any $\alpha \in \mathcal{A}$, and $\mathbb{F}^0$-stopping time θ, we denote by α^θ the map

$$\begin{array}{rcl} \alpha^\theta : (\Omega^0, \mathcal{F}_\theta^0) & \to & (\mathcal{A}, \mathcal{B}_\mathcal{A}) \\ \omega^0 & \mapsto & \alpha^{\theta(\omega^0),\omega^0}. \end{array}$$

The key step in the proof the DPP is to obtain a flow property on the controlled conditional distribution $\mathbb{F}^0$-progressive process $\{\mathbb{P}^{W^0}_{X_s^{t,\xi}}, t \leq s \leq T\}$, for $(t,\xi) \in [0,T] \times L^2(\mathcal{F}_t;\mathbb{R}^d)$, and $\alpha \in \mathcal{A}$.

Lemma 3.1 *For any $t \in [0,T]$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, the relation given by*

$$\rho_s^{t,\mu,\alpha} \quad := \quad \mathbb{P}^{W^0}_{X_s^{t,\xi,\alpha}}, \quad t \leq s \leq T, \quad \textit{for} \;\; \xi \in L^2(\mathcal{F}_t;\mathbb{R}^d) \;\; \textit{s.t.} \;\; \mathbb{P}^{W^0}_\xi = \mu, \tag{3.1}$$

defines a square integrable $\mathbb{F}^0$-progressive continuous process in $\mathcal{P}_2(\mathbb{R}^d)$, which satisfies the flow property: $\rho_s^{t,\mu,\alpha} = \rho_s^{\theta,\rho_\theta^{t,\mu,\alpha},\alpha^\theta}$, $\mathbb{P}^0$-a.s., i.e.

$$\rho_s^{t,\mu,\alpha}(\omega^0) \;=\; \rho_s^{\theta(\omega^0),\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0),\alpha^{\theta(\omega^0),\omega^0}}(\omega^0), \quad s\in[\theta,T],\ \mathbb{P}^0(d\omega^0)-a.s \tag{3.2}$$

for all $\theta\in\mathcal{T}^0_{t,T}$, the set of $\mathbb{F}^0$-stopping times valued in $[t,T]$.

Proof. 1. First observe that for any $t\in[0,T]$, $\xi\in L^2(\mathcal{F}_t;\mathbb{R}^d)$, $\alpha\in\mathcal{A}$, we have: $\mathbb{E}^0[\|\mathbb{P}^{W^0}_{X_s^{t,\xi,\alpha}}\|_2^2]=\mathbb{E}[|X_s^{t,\xi,\alpha}|^2]<\infty$, which means that the $\mathbb{F}^0$-progressive process $\{\mathbb{P}^{W^0}_{X_s^{t,\xi,\alpha}}, t\le s\le T\}$ is square integrable.

(i) Notice that for $\mathbb{P}^0$-a.s $\omega^0\in\Omega^0$, the law of the solution $\{X_s^{t,\xi,\alpha}(\omega^0,.), t\le s\le T\}$ to (2.1) on $(\Omega^1,\mathcal{F}^1,\mathbb{P}^1)$ is unique in law, which implies that $\mathbb{P}^{W^0}_{X_s^{t,\xi,\alpha}}(\omega^0)=\mathbb{P}^1_{X_s^{t,\xi,\alpha}(\omega^0,.)}$, $t\le s\le T$, depends on ξ only through $\mathbb{P}^{W^0}_\xi(\omega^0)=\mathbb{P}^1_{\xi(\omega^0,.)}$. In other words, for any ξ_1, $\xi_2\in L^2(\mathcal{F}_t;\mathbb{R}^d)$ s.t. $\mathbb{P}^{W^0}_{\xi_1}=\mathbb{P}^{W^0}_{\xi_2}$, the processes $\{\mathbb{P}^{W^0}_{X_s^{t,\xi_1,\alpha}}, t\le s\le T\}$ and $\{\mathbb{P}^{W^0}_{X_s^{t,\xi_2,\alpha}}, t\le s\le T\}$ are indistinguishable.

(ii) Let us now check that for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, one can find $\xi\in L^2(\mathcal{F}_t;\mathbb{R}^d)$ s.t. $\mathbb{P}^{W^0}_\xi=\mu$. Indeed, recalling that $\mathcal{G}$ is rich enough, one can find $\xi\in L^2(\mathcal{G};\mathbb{R}^d)\subset L^2(\mathcal{F}_t;\mathbb{R}^d)$ s.t. $\mathcal{L}(\xi)=\mu$. Since $\mathcal{G}$ is independent of W^0, this also means that $\mathbb{P}^{W^0}_\xi=\mu$.

In view of the uniqueness result in (i), and the representation result in (ii), one can define the process $\{\rho_s^{t,\mu,\alpha}, t\le s\le T\}$ by the relation (3.1), and this process is a square integrable $\mathbb{F}^0$-progressive process in $\mathcal{P}_2(\mathbb{R}^d)$. It has also continuous trajectories as $\rho^{t,\mu,\alpha}$ is valued in $\mathcal{P}_2(C([t,T];\mathbb{R}^d))$, the set of square integrable probability measures on the space $C([t,T];\mathbb{R}^d)$ of continuous functions from $[t,T]$ into $\mathbb{R}^d$.

2. Let us finally check the flow property (3.2). From pathwise uniqueness of the solution $\{X_s(\omega^0,.), t\le s\le T\}$ to (2.1) on $(\Omega,\mathcal{F}^1,\mathbb{P}^1)$ for $\mathbb{P}^0$-a.s. $\omega^0\in\Omega^0$, and recalling the definition of the shifted control process, we have the flow property: for $t\in[0,T]$, $\xi\in L^2(\mathcal{F}_t.\mathbb{R}^d)$, $\alpha\in\mathcal{A}$, and $\mathbb{P}^0$-a.s. $\omega^0\in\Omega^0$,

$$X_s^{t,\xi,\alpha}(\omega^0,.) \;=\; X_s^{\theta(\omega^0),X^{t,\xi,\alpha}_{\theta(\omega^0)}(\omega^0,.),\alpha^{\theta(\omega^0),\omega^0}}(\omega^0,.), \quad \mathbb{P}^1-\text{a.s.}$$

for all $\mathbb{F}^0$-stopping time θ valued in $[t,T]$. It follows that for any Borel-measurable bounded function φ on $\mathbb{R}^d$, and for $\mathbb{P}^0$-a.s $\omega^0\in\Omega^0$,

$$\begin{aligned}\rho_s^{t,\mu,\alpha}(\omega^0)(\varphi) \;=\; \mathbb{E}^1\Big[\varphi\big(X_s^{t,\xi,\alpha}(\omega^0,.)\big)\Big] \;&=\; \mathbb{E}^1\Big[\varphi\big(X_s^{\theta(\omega^0),X^{t,\xi,\alpha}_{\theta(\omega^0)}(\omega^0,.),\alpha^{\theta(\omega^0),\omega^0}}(\omega^0,.)\big)\Big]\\ &=\; \rho_s^{\theta(\omega^0),\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0),\alpha^{\theta(\omega^0),\omega^0}}(\omega^0)(\varphi),\end{aligned}$$

where the last equality is obtained by noting that $\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0)=\mathbb{P}^{W^0}_{X^{t,\xi,\alpha}_{\theta(\omega^0)}(\omega^0,.)}$, and the definition of $\rho_s^{t,\mu,\alpha}$. This shows the required flow property (3.2). □

Now, by the law of iterated conditional expectations, from (2.3), (3.1), and recalling that $\alpha \in \mathcal{A}$ is $\mathbb{F}^0$-progressive, we can rewrite the cost functional as

$$\begin{aligned} J(t,\xi,\alpha) &= \mathbb{E}\Big[\int_t^T \mathbb{E}\big[f\big(X_s^{t,\xi},\mathbb{P}^{W^0}_{X_s^{t,\xi}},\alpha_s\big)\big|\mathcal{F}_s^0\big]ds + \mathbb{E}\big[g\big(X_T^{t,\xi},\mathbb{P}^{W^0}_{X_T^{t,\xi}}\big)\big|\mathcal{F}_T^0\big]\Big] \\ &= \mathbb{E}\Big[\int_t^T \rho_s^{t,\mu}\big(f(.,\rho_s^{t,\mu},\alpha_s)\big)ds + \rho_T^{t,\mu}\big(g(.,\rho_T^{t,\mu})\big)\Big] \\ &= \mathbb{E}\Big[\int_t^T \hat{f}(\rho_s^{t,\mu},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu})\Big], \end{aligned} \tag{3.3}$$

for $t \in [0,T]$, $\xi \in L^2(\mathcal{G};\mathbb{R}^d)$ with law $\mu = \mathcal{L}(\xi) = \mathbb{P}^{W^0}_\xi \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, and with the functions $\hat{f} : \mathcal{P}_2(\mathbb{R}^d) \times \mathbf{A} \to \mathbb{R}$, and $\hat{g} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, defined by

$$\left\{ \begin{array}{rcrcl} \hat{f}(\mu,a) &:=& \mu\big(f(.,\mu,a)\big) &=& \int_{\mathbb{R}^d} f(x,\mu,a)\mu(dx) \\ \hat{g}(\mu) &:=& \mu\big(g(.,\mu)\big) &=& \int_{\mathbb{R}^d} g(x,\mu)\mu(dx). \end{array} \right. \tag{3.4}$$

(To alleviate notations, we have omitted here the dependence of $\rho_s^{t,\mu} = \rho_s^{t,\mu,\alpha}$ on α). Relation (3.3) means that the cost functional depends on ξ only through its distribution $\mu = \mathcal{L}(\xi)$, and by misuse of notation, we set:

$$J(t,\mu,\alpha) \;:=\; J(t,\xi,\alpha) \;=\; \mathbb{E}^0\Big[\int_t^T \hat{f}(\rho_s^{t,\mu},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu})\Big],$$

for $(t,\mu) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, $\xi \in L^2(\mathcal{G};\mathbb{R}^d)$ with $\mathcal{L}(\xi) = \mu$, and the expectation is taken under $\mathbb{P}^0$ since $\{\rho_s^{t,\mu}, t \le s \le T\}$ is $\mathbb{F}^0$-progressive, and the control $\alpha \in \mathcal{A}$ is an $\mathbb{F}^0$-progressive process. Therefore, the value function can be identified with a function defined on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, equal to (we keep the same notation $v(t,\mu) = v(t,\xi)$):

$$v(t,\mu) \;=\; \inf_{\alpha\in\mathcal{A}} \mathbb{E}^0\Big[\int_t^T \hat{f}(\rho_s^{t,\mu},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu})\Big],$$

and satisfying from (2.5) the quadratic growth condition

$$|v(t,\mu)| \;\le\; C(1+\|\mu\|_2^2), \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.5}$$

As a consequence of the flow property in Lemma 3.1, we obtain the following conditioning lemma, also called pseudo-Markov property in the terminology of [18], for the controlled conditional distribution $\mathbb{F}^0$-progressive process $\{\rho_s^{t,\mu,\alpha}, t \le s \le T\}$.

Lemma 3.2 *For any $(t,\mu,\alpha) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{A}$, and $\theta \in \mathcal{T}^0_{t,T}$, we have*

$$J(\theta,\rho_\theta^{t,\mu,\alpha},\alpha^\theta) \;=\; \mathbb{E}^0\Big[\int_\theta^T \hat{f}(\rho_s^{t,\mu,\alpha},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu,\alpha})\big|\mathcal{F}_\theta^0\Big], \quad \mathbb{P}^0 - a.s \tag{3.6}$$

In particular,

$$v(t,\mu) \;=\; \inf_{\alpha\in\mathcal{A}_t} \mathbb{E}^0\Big[\int_t^T \hat{f}(\rho_s^{t,\mu},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu})\Big]. \tag{3.7}$$

Proof. By the flow property (3.2), and since $\rho_\theta^{t,\mu,\alpha}$ is $\mathcal{F}_\theta^0$-measurable for θ $\mathbb{F}^0$-stopping time, we have for $\mathbb{P}^0$-a.s $\omega^0 \in \Omega^0$,

$$\begin{aligned}
& \mathbb{E}^0\Big[\int_\theta^T \hat{f}(\rho_s^{t,\mu,\alpha},\alpha_s)ds + \hat{g}(\rho_T^{t,\mu,\alpha})\big|\mathcal{F}_\theta^0\Big](\omega^0) \\
= \quad & \mathbb{E}^0\Big[\int_r^T \hat{f}(\rho_s^{r,\pi,\beta},\beta_s) + \hat{g}(\rho_T^{r,\pi,\beta})\big|\mathcal{F}_r^0\Big](\omega^0)\Big|_{r=\theta(\omega^0),\pi=\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0),\beta=\alpha^{r,\omega^0}} \\
= \quad & \mathbb{E}^0\Big[\int_r^T \hat{f}(\rho_s^{r,\pi,\beta},\beta_s) + \hat{g}(\rho_T^{r,\pi,\beta})\Big]\Big|_{r=\theta(\omega^0),\pi=\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0),\beta=\alpha^{r,\omega^0}},
\end{aligned}$$

where we used in the second equality the fact that for fixed ω^0, $r \in [t,T]$, $\pi \in \mathcal{P}_2(\mathbb{R}^d)$ represented by $\eta \in L^2(\mathcal{G};\mathbb{R}^d)$ s.t. $\mathcal{L}(\xi) = \pi$, the process α^{r,ω^0} lies in $\mathcal{A}_r$, hence is independent of $\mathcal{F}_r^0$, which implies that $X_s^{r,\eta,\alpha^{r,\omega^0}}$ is independent of $\mathcal{F}_r$, and thus $\rho_s^{r,\pi,\alpha^{r,\omega^0}}$ is also independent of $\mathcal{F}_r^0$ for $r \le s$. This shows the conditioning relation (3.6).

Let us denote by $\tilde{v}(t,\mu)$ the r.h.s. of (3.7). Since $\mathcal{A}_t \subset \mathcal{A}$, it is clear that $v(t,\mu) \le \tilde{v}(t,\mu)$. To prove the reverse inequality, we apply the conditioning relation (3.6) for $\theta = t$, and get in particular for all $\alpha \in \mathcal{A}$:

$$\int_{\Omega^0} J(t,\mu,\alpha^{t,\omega^0})\mathbb{P}^0(d\omega^0) \quad = \quad J(t,\mu,\alpha). \tag{3.8}$$

Now, recalling that for any fixed $\omega^0 \in \Omega^0$, α^{t,ω^0} lies in $\mathcal{A}_t$, we have $J(t,\mu,\alpha^{t,\omega^0}) \ge \tilde{v}(t,\mu)$, which proves the required result since α is arbitrary in (3.8). □

3.2 Continuity of the value function and dynamic programming principle

In this paragraph, we show the continuity of the value function, which is helpful for proving next the dynamic programming principle. We mainly follow arguments from [26] for the continuity result that we extend to our McKean-Vlasov framework.

Lemma 3.3 *The function $(t,\mu) \mapsto J(t,\mu,\alpha)$ is continuous on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, uniformly with respect to $\alpha \in \mathcal{A}$, and the function $\alpha \mapsto J(t,x,\alpha)$ is continuous on $\mathcal{A}$ for any $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$. Consequently, the cost functional J is continuous on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{A}$, and the value function v is continuous on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$.*

Proof. *(1)* For any $0 \le t \le s \le T$, $\mu,\pi \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, recall that $\mathbb{P}^0$-a.s. $\omega^0 \in \Omega^0$, we have $\mathbb{P}^1_{X_r^{t,\xi,\alpha}(\omega^0,.)} = \rho_r^{t,\mu,\alpha}(\omega^0)$, $\mathbb{P}^1_{X_r^{s,\zeta,\alpha}(\omega^0,.)} = \rho_r^{s,\pi,\alpha}(\omega^0)$ for $r \in [s,T]$, and any $\xi,\zeta \in L^2(\mathcal{G};\mathbb{R}^d)$ s.t. $\mathcal{L}(\xi) = \mu$, $\mathcal{L}(\zeta) = \pi$. By definition of $\|.\|_2$ and the Wasserstein distance in $\mathcal{P}_2(\mathbb{R}^d)$, we then have: $\|\rho_r^{t,\mu,\alpha}(\omega^0)\|_2 = \mathbb{E}^1|X_r^{t,\xi,\alpha}(\omega^0,.)|^2$, and $\mathcal{W}_2^2\big(\rho_r^{t,\mu,\alpha}(\omega^0),\rho_r^{s,\pi,\alpha}(\omega^0)\big) \le \mathbb{E}^1|X_r^{t,\xi,\alpha}(\omega^0,.) - X_r^{s,\zeta,\alpha}(\omega^0,.)|^2$, so that

$$\mathbb{E}^0\Big[\sup_{s\le r\le T}\|\rho_r^{t,\mu,\alpha}\|_2^2\Big] \quad \le \quad \mathbb{E}\Big[\sup_{s\le r\le T}|X_r^{t,\xi,\alpha}|^2\Big], \tag{3.9}$$

$$\mathbb{E}^0\Big[\sup_{s\le r\le T}\mathcal{W}_2^2(\rho_r^{t,\mu,\alpha},\rho_r^{s,\pi,\alpha})\Big] \quad \le \quad \mathbb{E}\Big[\sup_{s\le r\le T}|X_r^{t,\xi,\alpha} - X_r^{s,\zeta,\alpha}|^2\Big]. \tag{3.10}$$

From the state equation (2.1), and using standard arguments involving Burkholder-Davis-Gundy inequalities, (3.9), (3.10), and Gronwall lemma, under the Lipschitz condition in **(H1)**(i), we obtain the following estimates similar to the ones for controlled diffusion processes (see [26], Chap.2, Thm.5.9, Cor.5.10): there exists some positive constant C s.t. for all $t \in [0,T]$, $\xi, \zeta \in L^2(\mathcal{G};\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, $h \in [0, T-t]$,

$$\begin{aligned}
\mathbb{E}\big[\sup_{t\le s\le t+h} |X_s^{t,\xi,\alpha} - \xi|^2\big] &\le C(1+\mathbb{E}|\xi|^2)h, \\
\mathbb{E}\big[\sup_{t\le s\le T} |X_s^{t,\xi,\alpha} - X_s^{t,\zeta,\alpha}|^2\big] &\le C\mathbb{E}[|\xi-\zeta|^2],
\end{aligned}$$

from which we easily deduce that for all $0 \le t \le s \le T$, $\xi, \zeta \in L^2(\mathcal{G};\mathbb{R}^d)$, $\alpha \in \mathcal{A}$

$$\mathbb{E}\Big[\sup_{s\le r\le T} |X_r^{t,\xi,\alpha} - X_r^{s,\zeta,\alpha}|^2\Big] \le C\big(\mathbb{E}|\xi-\zeta|^2 + (1+\mathbb{E}|\xi|^2+\mathbb{E}|\zeta|^2)|s-t|\big). \quad (3.11)$$

Together with the estimates (2.2), and by definition of $\mathcal{W}_2(\mu,\pi)$, $\|\mu\|_2$, $\|\pi\|_2$, we then get from (3.9), (3.10):

$$\mathbb{E}^0\Big[\sup_{s\le r\le T} \|\rho_r^{t,\mu,\alpha}\|_2^2\Big] \le C(1+\|\mu\|_2^2), \quad (3.12)$$

$$\mathbb{E}^0\Big[\sup_{s\le r\le T} \mathcal{W}_2^2(\rho_r^{t,\mu,\alpha}, \rho_r^{s,\pi,\alpha})\Big] \le C\big(\mathcal{W}_2^2(\mu,\pi) + (1+\|\mu\|_2^2+\|\pi\|_2^2)|s-t|\big). \quad (3.13)$$

(2) Let us now show the continuity of the cost functional J in (t,μ), uniformly w.r.t. $\alpha \in \mathcal{A}$. First, we notice from the growth condition in **(H2)**(i) and the local Lipschitz condition in **(H2)**(ii) that there exists some positive constant C s.t. for all $\mu, \pi \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$,

$$\begin{aligned}
|\hat{f}(\mu,\alpha)| &\le C\big(1+\|\mu\|_2^2\big), \\
|\hat{f}(\mu,\alpha) - \hat{f}(\pi,\alpha)| + |\hat{g}(\mu) - \hat{g}(\pi)| &\le C(1+\|\mu\|_2+\|\pi\|_2)\mathcal{W}_2(\mu,\pi).
\end{aligned}$$

Then, we have for all $0 \le t \le s \le T$, $\mu, \pi \in \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$

$$\begin{aligned}
\big|J(t,\mu,\alpha) - J(s,\pi,\alpha)\big| &\le \mathbb{E}^0\Big[\int_t^s |\hat{f}(\rho_r^{t,\mu,\alpha}|dr\Big] \\
&\quad + \mathbb{E}^0\Big[\int_s^T \big|\hat{f}(\rho_r^{t,\mu,\alpha},\alpha_r) - \hat{f}(\rho_r^{s,\pi,\alpha},\alpha_r)\big|dr + \big|\hat{g}(\rho_T^{t,\mu,\alpha}) - \hat{g}(\rho_T^{s,\pi,\alpha})\big|\Big] \\
&\le C\mathbb{E}^0\Big[(1+\sup_{t\le r\le s}(\|\rho_r^{t,\mu,\alpha}\|_2)|s-t|\Big] \\
&\quad + C\mathbb{E}^0\Big[\big(1+\sup_{s\le r\le T}(\|\rho_r^{t,\mu,\alpha}\|_2 + \|\rho_r^{s,\pi,\alpha}\|_2)\big)\sup_{s\le r\le T}\mathcal{W}_2(\rho_r^{t,\mu,\alpha},\rho_r^{s,\pi,\alpha})\Big] \\
&\le C(1+\|\mu\|_2)|s-t| \\
&\quad + C(1+\|\mu\|_2+\|\pi\|_2)\big(\mathcal{W}_2(\mu,\pi) + (1+\|\mu\|_2+\|\pi\|_2)|s-t|^{\frac{1}{2}}\big),
\end{aligned}$$

by Cauchy Schwarz inequality and (3.12)-(3.13), which shows the desired continuity result.

(3) Let us show the continuity of the cost functional with respect to the control. Fix $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, and consider $\alpha \in \mathcal{A}$, a sequence $(\alpha^n)_n$ in $\mathcal{A}$ s.t. $\Delta(\alpha^n,\alpha) \to 0$, i.e. $d_A(\alpha_t^n, \alpha_t) \to 0$ in $dt \otimes d\mathbb{P}^0$-measure, as n goes to infinity. Denote by $\rho^n = \rho^{t,\mu,\alpha^n}$, $\rho =$

$\rho^{t,\mu,\alpha}$, $X^n = X^{t,\xi,\alpha^n}$, $X = X^{t,\xi,\alpha}$ for $\xi \in L^2(\mathcal{G};\mathbb{R}^d)$ s.t. $\mathcal{L}(\xi) = \mu$. By the same arguments as in (3.10), we have

$$\mathbb{E}^0\Big[\sup_{t\le s\le T} \mathcal{W}_2^2(\rho_s^n,\rho_s)\Big] \;\le\; \mathbb{E}\Big[\sup_{t\le s\le T} |X_s^n - X_s|^2\Big]. \tag{3.14}$$

Next, starting from the state equation (2.1), using standard arguments involving Burkholder-Davis-Gundy inequalities, (3.14), and Gronwall lemma, under the Lipschitz condition in **(H1)**(i), we arrive at:

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\le s\le T} |X_s^n - X_s|^2\Big] \;\le\; C\Big\{&\mathbb{E}\Big[\int_t^T |b(X_s,\rho_s,\alpha_s) - b(X_s,\rho_s,\alpha_s^n)|^2 ds \\ &+ \int_t^T |\sigma(X_s,\rho_s,\alpha_s) - \sigma(X_s,\rho_s,\alpha_s^n)|^2 ds \\ &+ \int_t^T |\sigma_0(X_s,\rho_s,\alpha_s) - \sigma_0(X_s,\rho_s,\alpha_s^n)|^2 ds\Big]\Big\},\end{aligned}$$

for some positive constant C independent of n. Recalling the bound (2.2), and (3.9), we deduce by the dominated convergence theorem under the linear growth condition in **(H1)**(i), and the continuity assumption in **(H1)**(ii) that $\mathbb{E}\big[\sup_{t\le s\le T} |X_s^n - X_s|^2\big] \to 0$, and thus by (3.14)

$$\mathbb{E}^0\Big[\sup_{t\le s\le T} \mathcal{W}_2^2(\rho_s^n,\rho_s)\Big] \;\to\; 0, \quad \text{as } n\to\infty. \tag{3.15}$$

Now, by writing

$$\begin{aligned}&\big|J(t,\mu,\alpha^n) - J(t,\mu,\alpha)\big| \\ \le\;& \mathbb{E}^0\Big[\int_t^T \big|\hat{f}(\rho_s^n,\alpha_s^n) - \hat{f}(\rho_s,\alpha_s)\big| ds + \big|\hat{g}(\rho_T^n) - \hat{g}(\rho_T)\big|\Big],\end{aligned} \tag{3.16}$$

and noting that $\hat{f}$ and $\hat{g}$ are continuous on $\mathcal{P}_2(\mathbb{R}^d)\times\mathbf{A}$, resp. on $\mathcal{P}_2(\mathbb{R}^d)$, under the continuity assumption in **(H2)**(ii), we conclude by the same arguments as in [26] using (3.15) (see Chapter 3, Sec. 2, or also Lemma 4.1 in [23]) that the r.h.s. of (3.16) tends to zero as n goes to infinity, which proves the continuity of $J(t,\mu,.)$ on $\mathcal{A}$.

(4) Finally, the global continuity of the cost functional J on $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)\times\mathcal{A}$ is a direct consequence of the continuity of $J(.,.,\alpha)$ on $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ uniformly w.r.t. $\alpha\in\mathcal{A}$, and the continuity of $J(t,\mu,.)$ on $\mathcal{A}$, while the continuity of the value function v on $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ follows immediately from the fact that

$$|v(t,\mu) - v(s,\pi)| \;\le\; \sup_{\alpha\in\mathcal{A}} |J(t,\mu,\alpha) - J(s,\pi,\alpha)|, \quad t,s\in[0,T],\ \mu,\pi\in\mathcal{P}_2(\mathbb{R}^d),$$

and again from the continuity of $J(.,.,\alpha)$ on $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ uniformly w.r.t. $\alpha\in\mathcal{A}$. □

We can now state the dynamic programming principle (DPP) for the value function to the stochastic McKean-Vlasov control problem.

Proposition 3.1 *(Dynamic Programming Principle)*
We have for all $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$,

$$\begin{aligned} v(t,\mu) &= \inf_{\alpha\in\mathcal{A}} \inf_{\theta\in\mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big] \\ &= \inf_{\alpha\in\mathcal{A}} \sup_{\theta\in\mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big], \end{aligned}$$

which means equivalently that
(i) for all $\alpha \in \mathcal{A}$, $\theta \in \mathcal{T}^0_{t,T}$,

$$v(t,\mu) \ \leq \ \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big], \tag{3.17}$$

(ii) for all $\varepsilon > 0$, *there exists* $\alpha \in \mathcal{A}$, *such that for all* $\theta \in \mathcal{T}^0_{t,T}$,

$$v(t,\mu)+\varepsilon \ \geq \ \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big]. \tag{3.18}$$

Remark 3.1 The above formulation of the DPP implies in particular that for all $\theta \in \mathcal{T}^0_{t,T}$,

$$v(t,\mu) \ = \ \inf_{\alpha\in\mathcal{A}} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big],$$

which is the usual formulation of the DPP. The formulation in Proposition 3.1 is stronger, and the difference relies on the fact that in the inequality (3.18), the ε-optimal control $\alpha = \alpha^\varepsilon$ does not depend on θ. This condition will be useful to show later the viscosity supersolution property of the value function. □

Proof. 1. Fix $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$. From the conditioning relation (3.6), we have for all $\theta \in \mathcal{T}^0_{t,T}$, $\alpha \in \mathcal{A}$,

$$J(t,\mu,\alpha) \ = \ \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds + J(\theta,\rho^{t,\mu,\alpha}_\theta,\alpha^\theta)\Big]. \tag{3.19}$$

Since $J(.,.,\alpha^\theta) \geq v(.,.)$, and θ is arbitrary in $\mathcal{T}^0_{t,T}$, we have

$$J(t,\mu,\alpha) \ \geq \ \sup_{\theta\in\mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds + v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big],$$

and since α is arbitrary in $\mathcal{A}$, it follows that

$$v(t,\mu) \ \geq \ \inf_{\alpha\in\mathcal{A}} \sup_{\theta\in\mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho^{t,\mu,\alpha}_s,\alpha_s)ds \ + \ v(\theta,\rho^{t,\mu,\alpha}_\theta)\Big] \tag{3.20}$$

2. Fix $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, $\alpha \in \mathcal{A}$ and $\theta \in \mathcal{T}^0_{t,T}$. For any $\varepsilon > 0$, $\omega^0 \in \Omega^0$, one can find from (3.7) some $\alpha^{(\varepsilon,\omega^0)} \in \mathcal{A}_{\theta(\omega^0)}$ s.t.

$$v(\theta(\omega^0),\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0))+\varepsilon \ \geq \ J(\theta(\omega^0),\rho^{t,\mu,\alpha}_{\theta(\omega^0)}(\omega^0),\alpha^{(\varepsilon,\omega^0)}). \tag{3.21}$$

Since J and v are continuous (by Lemma 3.3), one can invoke measurable selection arguments (see e.g. [36]), to claim that the map $\omega^0 \in (\Omega^0, \mathcal{F}^0) \mapsto \alpha^{(\varepsilon,\omega^0)} \in (\mathcal{A}, \mathcal{B}_{\mathcal{A}})$ can be chosen measurable. Let us now define the process $\bar{\alpha}$ on $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$ obtained by concatenation at θ of the processes α and $\alpha^{(\varepsilon,\omega^0)}$ in $\mathcal{A}$, namely:

$$\bar{\alpha}_s(\omega^0) \quad := \quad \alpha_s(\omega^0)1_{s<\theta(\omega^0)} + \alpha^{(\varepsilon,\omega^0)}(\omega^0)1_{s\geq\theta(\omega^0)}, \quad 0 \leq s \leq T.$$

By Lemma 2.1 in [35], and since $\mathcal{A}$ is a separable metric space, the process $\bar{\alpha}$ is $\mathbb{F}^0$-progressively measurable, and thus $\bar{\alpha} \in \mathcal{A}$. Notice with our notations of shifted control process that $\bar{\alpha}^{\theta(\omega^0),\omega^0} = \alpha^{(\varepsilon,\omega^0)}$ for all ω^0 in Ω^0, and then (3.21) reads as

$$v(\theta, \rho_\theta^{t,\mu,\alpha}) + \varepsilon \quad \geq \quad J(\theta, \rho_\theta^{t,\mu,\alpha}, \bar{\alpha}^\theta), \quad \mathbb{P}^0 - \text{a.s.}$$

Therefore, by using again (3.19) to $\bar{\alpha}$, and since $\rho_s^{t,\mu,\bar{\alpha}} = \rho_s^{t,\mu,\alpha}$ for $s \leq \theta$ (recall that $\bar{\alpha}_s = \alpha_s$ for $s < \theta$, and $\rho^{t,\mu,\alpha}$ has continuous trajectories), we get

$$\begin{aligned} v(t,\mu) \leq \; J(t,\mu,\bar{\alpha}) \quad &= \quad \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho_s^{t,\mu,\alpha}, \alpha_s)ds + J(\theta, \rho_\theta^{t,\mu,\alpha}, \bar{\alpha}^\theta)\Big] \\ &\leq \quad \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho_s^{t,\mu,\alpha}, \alpha_s)ds + v(\theta, \rho_\theta^{t,\mu,\alpha})\Big] + \varepsilon \end{aligned}$$

Since α, θ and ε are arbitrary, this gives the inequality

$$v(t,\mu) \quad \leq \quad \inf_{\alpha\in\mathcal{A}} \inf_{\theta\in\mathcal{T}_{t,T}^0} \mathbb{E}^0\Big[\int_t^\theta \hat{f}(\rho_s^{t,\mu,\alpha}, \alpha_s)ds \; + \; v(\theta, \rho_\theta^{t,\mu,\alpha})\Big],$$

which, combined with the first inequality (3.20), proves the DPP result. □

4 Bellman equation and viscosity solutions

4.1 Differentiability and Itô's formula in Wasserstein space

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 [32]. We provide a brief introduction to this concept and refer to the lecture notes [10] (see also [9], [17]) for the details.

This notion is based on the *lifting* of functions $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ into functions $\tilde{u}$ defined on $L^2(\mathcal{G};\mathbb{R}^d)$ $(= L^2(\tilde{\Omega}^1, \mathcal{G}, \tilde{\mathbb{P}}^1; \mathbb{R}^d))$ by setting $\tilde{u}(\xi) = u(\mathcal{L}(\xi))$ $(= u(\tilde{\mathbb{P}}_\xi^1))$. Conversely, given a function $\tilde{u}$ defined on $L^2(\mathcal{G};\mathbb{R}^d)$, we call *inverse-lifted* function of $\tilde{u}$ the function u defined on $\mathcal{P}_2(\mathbb{R}^d)$ by $u(\mu) = \tilde{u}(\xi)$ for $\mu = \mathcal{L}(\xi)$, and we notice that such u exists iff $\tilde{u}(\xi)$ depends only on the distribution of ξ for any $\xi \in L^2(\mathcal{G};\mathbb{R}^d)$. In this case, we shall often identify in the sequel the function u and its lifted version $\tilde{u}$, by using the same notation $u = \tilde{u}$.

We say that u is differentiable (resp. $\mathcal{C}^1$) on $\mathcal{P}_2(\mathbb{R}^d)$ if the lift $\tilde{u}$ is Fréchet differentiable (resp. Fréchet differentiable with continuous derivatives) on $L^2(\mathcal{G};\mathbb{R}^d)$. In this case, the Fréchet derivative $[D\tilde{u}](\xi)$, viewed as an element $D\tilde{u}(\xi)$ of $L^2(\mathcal{G};\mathbb{R}^d)$ by Riesz' theorem: $[D\tilde{u}](\xi)(Y) = \tilde{\mathbb{E}}[D\tilde{u}(\xi).Y]$, can be represented as

$$D\tilde{u}(\xi) \quad = \quad \partial_\mu u(\mathcal{L}(\xi))(\xi), \tag{4.1}$$

for some function $\partial_\mu u(\mathcal{L}(\xi)) : \mathbb{R}^d \to \mathbb{R}^d$, which is called derivative of u at $\mu = \mathcal{L}(\xi)$. Moreover, $\partial_\mu u(\mu) \in L^2_\mu(\mathbb{R}^d)$ for $\mu \in \mathcal{P}_2(\mathbb{R}^d) = \{\mathcal{L}(\xi) : \xi \in L^2(\mathcal{G};\mathbb{R}^d)\}$. Following [17], we say that u is fully $\mathcal{C}^2$ 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 and

(i) for each fixed $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the mapping $x \in \mathbb{R}^d \mapsto \partial_\mu u(\mu)(x)$ is differentiable in the standard sense, with a gradient denoted by $\partial_x \partial_\mu u(\mu)(x) \in \mathbb{R}^{d\times d}$, and s.t. the mapping $(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_x \partial_\mu u(\mu)(x)$ is continuous

(ii) for each fixed $x \in \mathbb{R}^d$, the mapping $\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_\mu u(\mu)(x)$ is differentiable in the above lifted sense. Its derivative, interpreted thus as a mapping $x' \in \mathbb{R}^d \mapsto \partial_\mu\big[\partial_\mu u(\mu)(x)\big](x') \in \mathbb{R}^{d\times d}$ in $L^2_\mu(\mathbb{R}^{d\times d})$, is denoted by $x' \in \mathbb{R}^d \mapsto \partial^2_\mu u(\mu)(x, x')$, and s.t. the mapping $(\mu, x, x') \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \mapsto \partial^2_\mu u(\mu)(x, x')$ is continuous.

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

$$\sup_{\mu\in\mathcal{K}} \Big[\int_{\mathbb{R}^d} \big|\partial_\mu u(\mu)(x)\big|^2 \mu(dx) + \big\|\partial_x \partial_\mu u(\mu)\big\|_\infty + \big\|\partial^2_\mu u(\mu)\big\|_\infty \Big] \quad < \quad \infty. \tag{4.2}$$

If u lies in $\mathcal{C}^2_b(\mathcal{P}_2(\mathbb{R}^d))$, then its lifted function $\tilde{u}$ is twice continuously Fréchet differentiable on $L^2(\mathcal{G};\mathbb{R}^d)$. In this case, the second Fréchet derivative $D^2\tilde{u}(\xi)$ is identified indifferently by Riesz' theorem as a bilinear form on $L^2(\mathcal{G};\mathbb{R}^d)$ or as a self-adjoint operator (hence bounded) on $L^2(\mathcal{G};\mathbb{R}^d)$, denoted by $D^2\tilde{u}(\xi) \in S(L^2(\mathcal{G};\mathbb{R}^d))$, and we have the relation (see Appendix A.2 in [12]):

$$\left\{ \begin{array}{rcl} D^2\tilde{u}(\xi)[Y,Y] & = & \tilde{\mathbb{E}}^1\Big[D^2\tilde{u}(\xi)(Y).Y\Big] \quad = \quad \tilde{\mathbb{E}}^1\Big[\tilde{\mathbb{E}}'^1\big[\mathrm{tr}\big(\partial^2_\mu u(\mathcal{L}(\xi))(\xi,\xi')Y(Y')^\intercal\big)\big]\Big] \\ & & \qquad\qquad\qquad\qquad\qquad\qquad + \ \tilde{\mathbb{E}}^1\Big[\mathrm{tr}\big(\partial_x\partial_\mu u(\mathcal{L}(\xi))(\xi)YY^\intercal\big)\Big], \\ D^2\tilde{u}(\xi)[ZN,ZN] & = & \tilde{\mathbb{E}}^1\Big[D^2\tilde{u}(\xi)(ZN).ZN\Big] \quad = \quad \tilde{\mathbb{E}}^1\Big[\mathrm{tr}\big(\partial_x\partial_\mu u(\mathcal{L}(\xi))(\xi)ZZ^\intercal\big)\Big], \end{array} \right. \tag{4.3}$$

for any $\xi \in L^2(\mathcal{G};\mathbb{R}^d)$, $Y \in L^2(\mathcal{G};\mathbb{R}^d)$, $Z \in L^2(\mathcal{G};\mathbb{R}^{d\times q})$, and where (ξ', Y') is a copy of (ξ, Y) on another Polish and atomless probability space $(\tilde{\Omega}'^1, \mathcal{G}', \tilde{\mathbb{P}}'^1)$, $N \in L^2(\mathcal{G};\mathbb{R}^q)$ is independent of (ξ, Z) with zero mean, and unit variance. Here $^\intercal$ denotes the transpose of any vector or matrix.

We next need an Itô's formula along a flow of conditional measures proved in [12]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space of the form $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega^0 \times \Omega^1, \mathcal{F}^0 \otimes \mathcal{F}^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$, where $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$ supports W^0 and $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$ supports B as in Section 2. Let us consider an Itô process in $\mathbb{R}^d$ of the form:

$$X_t \quad = \quad X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \int_0^t \sigma^0_s dW^0_s, \quad 0 \le t \le T, \tag{4.4}$$

where X_0 is independent of (B, W^0), b, σ, σ^0 are progressively measurable processes with respect to the natural filtration $\mathbb{F}$ generated by (X^0, B, W^0), and satisfying the square integrability condition: $\mathbb{E}\big[\int_0^T |b_t|^2 + |\sigma_t|^2 + |\sigma^0_t|^2 dt\big] < \infty$. Denote by $\mathbb{P}^{W^0}_{X_t}$ the conditional law of X_t, $t \in [0, T]$, given the σ-algebra $\mathcal{F}^0$ generated by the whole filtration of W^0, and

by $\mathbb{E}_{W^0} = \mathbb{E}^1$ the conditional expectation w.r.t. $\mathcal{F}^0$. Let $u \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$. Then, for all $t \in [0,T]$, we have (see Proposition 6.3 in [12]):

$$\begin{aligned} u(\mathbb{P}_{X_t}^{W^0}) &= u(\mathbb{P}_{X_0}) + \int_0^t \mathbb{E}_{W^0}\Big[\partial_\mu u(\mathbb{P}_{X_s}^{W^0})(X_s).b_s + \frac{1}{2}\mathrm{tr}\big(\partial_x\partial_\mu u(\mathbb{P}_{X_s}^{W^0})(X_s)(\sigma_s\sigma_s^\intercal + \sigma_s^0(\sigma_s^0)^\intercal)\big)\Big] \\ &\quad + \mathbb{E}_{W^0}\Big[\mathbb{E}'_{W^0}\big[\frac{1}{2}\mathrm{tr}\big(\partial_\mu^2 u(\mathbb{P}_{X_s}^{W^0})(X_s,X'_s)\sigma_s^0(\sigma_s'^0)^\intercal\big)\big]\Big]ds \\ &\quad + \int_0^t \mathbb{E}_{W^0}\Big[\partial_\mu u(\mathbb{P}_{X_s}^{W^0})(X_s)^\intercal\sigma_s^0\Big]dW_s^0, \end{aligned} \tag{4.5}$$

where X' and σ'^0 are copies of X and σ^0 on another probability space $(\Omega' = \Omega^0 \times \Omega'^1\, \mathcal{F}^0 \otimes \mathcal{F}'^1, \mathbb{P}^0 \times \mathbb{P}'^1)$, with $(\Omega'^1, \mathcal{F}'^1, \mathbb{P}'^1)$ supporting B' a copy of B, and $\mathbb{E}'_{W^0} = \mathbb{E}'^1$.

Alternatively, we can formulate Itô's formula for the lifted function $\tilde{u}$ on $L^2(\mathcal{G},\mathbb{R}^d)$ $(= L^2(\tilde{\Omega}^1,\mathcal{G},\tilde{\mathbb{P}}^1;\mathbb{R}^d))$. For this, consider a copy $\tilde{B}$ of B on the probability space $(\tilde{\Omega}^1,\mathcal{G},\tilde{\mathbb{P}}^1)$, denote by $\tilde{X}_0$, $\tilde{b}$, $\tilde{\sigma}$, $\tilde{\sigma}_0$ copies of X_0, b, σ, σ_0 on $(\tilde{\Omega} = \Omega^0 \times \tilde{\Omega}^1, \tilde{\mathcal{F}} = \mathcal{F}^0 \otimes \mathcal{G}, \tilde{\mathbb{P}} = \mathbb{P}^0 \otimes \tilde{\mathbb{P}}^1)$, and consider the Itô process $\tilde{X}$ on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ of the form

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s d\tilde{B}_s + \int_0^t \tilde{\sigma}_s^0 dW_s^0, \quad 0 \le t \le T,$$

which is then a copy of X in (4.4). The process $\check{X}$ defined by $\check{X}_t(\omega^0) = \tilde{X}_t(\omega^0,.)$, $0 \le t \le T$, is $\mathbb{F}^0$-progressive, and valued in $L^2(\mathcal{G};\mathbb{R}^d)$. Similarly, the processes defined by $\check{b}_t(\omega^0) = \tilde{b}_t(\omega^0,.)$, $\check{\sigma}_t(\omega^0) = \tilde{\sigma}_t(\omega^0,.)$, $\check{\sigma}_t^0(\omega^0) = \tilde{\sigma}_t^0(\omega^0,.)$, $0 \le t \le T$, are valued in $L^2(\mathcal{G};\mathbb{R}^d)$, $\mathbb{P}^0$-a.s. The Itô's formula (4.5) is then written for the lifted function $\tilde{u} \in \mathcal{C}^2(L^2(\mathcal{G};\mathbb{R}^d))$ from (4.1)-(4.3) as:

$$\begin{aligned} \tilde{u}(\check{X}_t) &= \tilde{u}(\check{X}_0) + \int_0^t \tilde{\mathbb{E}}^1\Big[D\tilde{u}(\check{X}_s).\check{b}_s + \frac{1}{2}D^2\tilde{u}(\check{X}_s)(\check{\sigma}_s N).\check{\sigma}_s N + \frac{1}{2}D^2\tilde{u}(\check{X}_s)(\check{\sigma}_s^0).\check{\sigma}_s^0\Big]ds \\ &\quad + \int_0^t \tilde{\mathbb{E}}^1\big[D\tilde{u}(\check{X}_s)^\intercal\check{\sigma}_s^0\big]dW_s^0, \qquad 0 \le t \le T,\ \mathbb{P}^0 - a.s. \end{aligned} \tag{4.6}$$

where $N \in L^2(\mathcal{G};\mathbb{R}^d)$ is independent of $(\tilde{B},\tilde{X}_0)$, with zero mean, and unit variance.

4.2 Dynamic programming equation

The dynamic programming Bellman equation associated to the value function of the stochastic McKean-Vlasov control problem takes the form:

$$\left\{ \begin{aligned} -\partial_t v - \inf_{a\in\mathbf{A}}\Big[\hat{f}(\mu,a) + \mu\big(\mathbb{L}^a v(t,\mu)\big) + \mu\otimes\mu\big(\mathbb{M}^a v(t,\mu)\big)\Big] &= 0, \quad (t,\mu) \in [0,T)\times\mathcal{P}_2(\mathbb{R}^d), \\ v(T,\mu) &= \hat{g}(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \end{aligned} \right. \tag{4.7}$$

where for $\phi \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$, $a \in \mathbf{A}$, and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mathbb{L}^a\phi(\mu) \in L^2_\mu(\mathbb{R})$ is the function $\mathbb{R}^d \to \mathbb{R}$ defined by

$$\mathbb{L}^a\phi(\mu)(x) := \partial_\mu\phi(\mu)(x).b(x,\mu,a) + \frac{1}{2}\mathrm{tr}\big(\partial_x\partial_\mu\phi(\mu)(x)(\sigma\sigma^\intercal + \sigma_0\sigma_0^\intercal)(x,\mu,a)\big), \tag{4.8}$$

and $\mathbb{M}^a\phi(\mu) \in L^2_{\mu\otimes\mu}(\mathbb{R})$ is the function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$\mathbb{M}^a\phi(\mu)(x,x') := \frac{1}{2}\mathrm{tr}\big(\partial_\mu^2\phi(\mu)(x,x')\sigma_0(x,\mu,a)\sigma_0^\intercal(x',\mu,a)\big). \tag{4.9}$$

Alternatively, by viewing the value function as a function on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ via the lifting identification, and keeping the same notation $v(t,\xi) = v(t,\mathcal{L}(\xi))$ (recall that v depends on ξ only via its distribution), we see from the connection (4.1)-(4.3) between derivatives in the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$ and in the Hilbert space $L^2(\mathcal{G};\mathbb{R}^d)$ that the Bellman equation (4.7) is written also in $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ as

$$\left\{ \begin{array}{rcl} -\partial_t v - H\big(\xi, Dv(t,\xi), D^2 v(t,\xi)\big) & = & 0, \quad (t,\xi) \in [0,T) \times L^2(\mathcal{G};\mathbb{R}^d), \\ v(T,\xi) & = & \tilde{\mathbb{E}}^1\big[g(\xi,\mathcal{L}(\xi))\big], \quad \xi \in L^2(\mathcal{G};\mathbb{R}^d), \end{array} \right. \tag{4.10}$$

where $H : L^2(\mathcal{G};\mathbb{R}^d) \times L^2(\mathcal{G};\mathbb{R}^d) \times S(L^2(\mathcal{G};\mathbb{R}^d)) \to \mathbb{R}$ is defined by

$$\begin{aligned} H(\xi,P,Q) \;=\; & \inf_{a\in\mathbf{A}} \tilde{\mathbb{E}}^1\Big[f(\xi,\mathcal{L}(\xi),a) + P.b(\xi,\mathcal{L}(\xi),a) \\ & + \frac{1}{2} Q(\sigma_0(\xi,\mathcal{L}(\xi),a)).\sigma_0(\xi,\mathcal{L}(\xi),a) + \frac{1}{2} Q(\sigma(\xi,\mathcal{L}(\xi),a)N).\sigma(\xi,\mathcal{L}(\xi),a)N \Big], \end{aligned} \tag{4.11}$$

with $N \in L^2(\mathcal{G};\mathbb{R}^n)$ of zero mean, and unit variance, and independent of ξ.

The purpose of this section is to prove an analytic characterization of the value function in terms of the dynamic programming Bellman equation. We shall adopt a notion of viscosity solutions following the approach in [32], which consists via the lifting identification in working in the Hilbert space $L^2(\mathcal{G};\mathbb{R}^d)$ instead of working in the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$. Indeed, comparison principles for viscosity solutions in the Wasserstein space, or more generally in metric spaces, are difficult to obtain as we have to deal with locally non compact spaces (see e.g. [2], [24], [22]), and instead by working in separable Hilbert spaces, one can essentially reduce to the case of Euclidian spaces by projection, and then take advantage of the results developed for viscosity solutions this context, in particular here, for second order Hamilton-Jacobi-Bellman equations, see [31], [21]. We shall assume that the σ-algebra $\mathcal{G}$ is countably generated upto null sets, which ensures that the Hilbert space $L^2(\mathcal{G};\mathbb{R}^d)$ is separable, see [20], p. 92. This is satisfied for example when $\mathcal{G}$ is the Borel σ-algebra of a canonical space $\tilde{\Omega}^1$ of continuous functions on $\mathbb{R}_+$ (see Exercise 4.21 in Chapter 1 of [34]).

Definition 4.1 *We say that a continuous function $u : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a viscosity (sub, super) solution to* (4.7) *if its lifted version $\tilde{u}$ on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ is a viscosity (sub, super) solution to* (4.10)*, that is:*
(i) $\tilde{u}(T,\xi) \leq \tilde{\mathbb{E}}^1\big[g(\xi,\mathcal{L}(\xi))\big]$, and for any test function $\varphi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ (the set of real-valued continuous functions on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ which are continuously differentiable in $t \in [0,T)$, and twice continuously Fréchet differentiable on $L^2(\mathcal{G};\mathbb{R}^d)$) s.t. $\tilde{u} - \varphi$ has a maximum at $(t,\xi) \in [0,T) \times L^2(\mathcal{G};\mathbb{R}^d)$, one has

$$-\partial_t \varphi(t,\xi) - H\big(\xi, D\varphi(t,\xi), D^2\varphi(t,\xi)\big) \;\leq\; 0.$$

(ii) $\tilde{u}(T,\xi) \geq \tilde{\mathbb{E}}^1\big[g(\xi,\mathcal{L}(\xi))\big]$, and for any test function $\varphi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ s.t. $\tilde{u} - \varphi$ has a minimum at $(t,\xi) \in [0,T) \times L^2(\mathcal{G};\mathbb{R}^d)$, one has

$$-\partial_t \varphi(t,\xi) - H\big(\xi, D\varphi(t,\xi), D^2\varphi(t,\xi)\big) \;\geq\; 0.$$

The main result of this section is the viscosity characterization of the value function for the stochastic McKean-Vlasov control problem (2.4) to the dynamic programming Bellman equation (4.7) (or (4.10)).

Theorem 4.1 *The value function v is the unique continuous viscosity solution to* (4.7) *satisfying a quadratic growth condition* (3.5).

Proof. *(1) Viscosity property.* Let us first reformulate the dynamic programming principle (DPP) of Proposition 3.1 for the value function viewed now as a function on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$. For this, we take a copy $\tilde{B}$ of B on the probability space $(\tilde{\Omega}^1, \mathcal{G}, \tilde{\mathbb{P}}^1)$, and given $(t,\xi) \in [0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$, $\alpha \in \mathcal{A}$, we consider on $(\tilde{\Omega} = \Omega^0 \times \tilde{\Omega}^1, \tilde{\mathcal{F}} = \mathcal{F}^0 \otimes \mathcal{G}, \tilde{\mathbb{P}} = \mathbb{P}^0 \otimes \tilde{\mathbb{P}}^1)$ the solution $\tilde{X}^{t,\xi,\alpha}$, $t \leq s \leq T$, to the McKean-Vlasov equation

$$\begin{aligned}\tilde{X}_s^{t,\xi,\alpha} &= \xi + \int_t^s b(\tilde{X}_r^{t,\xi,\alpha}, \tilde{\mathbb{P}}^{W^0}_{\tilde{X}_s^{t,\xi,\alpha}}, \alpha_r)dr + \int_t^s \sigma(\tilde{X}_r^{t,\xi,\alpha}, \tilde{\mathbb{P}}^{W^0}_{\tilde{X}_s^{t,\xi,\alpha}}, \alpha_r)d\tilde{B}_r \\ &\quad + \int_t^s \sigma_0(\tilde{X}_r^{t,\xi,\alpha}, \tilde{\mathbb{P}}^{W^0}_{\tilde{X}_s^{t,\xi,\alpha}}, \alpha_r)dW_r^0, \quad t \leq s \leq T,\end{aligned}$$

where $\tilde{\mathbb{P}}^{W^0}_{\tilde{X}_s^{t,\xi,\alpha}}$ denotes the regular conditional distribution of $\tilde{X}_s^{t,\xi,\alpha}$ given $\mathcal{F}^0$. In other words, $\tilde{X}^{t,\xi,\alpha}$ is a copy of $X^{t,\xi,\alpha}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, and denoting by $\check{X}_s^{t,\xi,\alpha}(\omega^0) = \tilde{X}_s^{t,\xi,\alpha}(\omega^0, .)$, $t \leq s \leq T$, we see that the process $\{\check{X}_s^{t,\xi,\alpha}, t \leq s \leq T\}$ is $\mathbb{F}^0$-progressive, valued in $L^2(\mathcal{G};\mathbb{R}^d)$, and $\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,\alpha}} = \rho_s^{t,\mu,\alpha}$ for $\mu = \mathcal{L}(\xi)$. Therefore, the lifted value function on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ identified with the value function on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ satisfies $v(s, \check{X}_s^{t,\xi,\alpha}) = v(s, \rho_s^{t,\mu,\alpha})$, $t \leq s \leq T$. By noting that $\hat{f}(\rho_s^{t,\mu,\alpha}, \alpha_s) = \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,\alpha}, \tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,\alpha}}, \alpha_s)\big]$, we obtain from Proposition 3.1 the lifted DPP: for all $(t,\xi) \in [0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$,

$$v(t,\xi) = \inf_{\alpha \in \mathcal{A}} \inf_{\theta \in \mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,\alpha}, \tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,\alpha}}, \alpha_s)\big]ds \ + \ v(\theta, \check{X}_\theta^{t,\xi,\alpha})\Big] \tag{4.12}$$

$$= \inf_{\alpha \in \mathcal{A}} \sup_{\theta \in \mathcal{T}^0_{t,T}} \mathbb{E}^0\Big[\int_t^\theta \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,\alpha}, \tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,\alpha}}, \alpha_s)\big]ds \ + \ v(\theta, \check{X}_\theta^{t,\xi,\alpha})\Big]. \tag{4.13}$$

We already know that v is continuous on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$, hence in particular at T, so that $v(T,\xi) = \tilde{\mathbb{E}}^1[g(\xi, \mathcal{L}(\xi))]$, and it remains to derive the viscosity property for the value function in $[0,T) \times L^2(\mathcal{G};\mathbb{R}^d)$ by following standard arguments that we adapt in our context.

(i) Subsolution property. Fix $(t,\xi) \in [0,T) \times L^2(\mathcal{G};\mathbb{R}^d)$, and consider some test function $\varphi \in \mathcal{C}^2([0,T] \times L^2(\mathcal{G};\mathbb{R}^d))$ s.t. $v - \varphi$ has a maximum at (t,ξ), and w.l.o.g. $v(t,\xi) = \varphi(t,\xi)$, so that $v \leq \varphi$. Let a be an arbitrary element in $\mathbf{A}$, $\alpha \equiv a$ the constant control in $\mathcal{A}$ equal to a, and consider the stopping time in $\mathcal{T}^0_{t,T}$: $\theta_h = \inf\{s \geq t : \tilde{\mathbb{E}}^1[|\check{X}_s^{t,\xi,a} - \xi|^2] \geq \delta\} \wedge (t+h)$, with δ some positive constant, and $h \in (0, T-t)$. From the first part (4.12) of the DPP, we get

$$\varphi(t,\xi) \ \leq \ \mathbb{E}^0\Big[\int_t^{\theta_h} \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,a}, \tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}}, a)\big]ds \ + \ \varphi(\theta_h, \check{X}_{\theta_h}^{t,\xi,a})\Big].$$

Applying Itô's formula (4.6) to $\varphi(s, \check{X}_s^{t,\xi,a})$, and noting that the stochastic integral w.r.t. W^0 vanishes under expectation $\mathbb{E}^0$ by the localization with the stopping time θ_h, we then

have

$$
\begin{aligned}
0 \;\leq\; & \mathbb{E}^0\Big[\frac{1}{h}\int_t^{\theta_h} \partial_t\varphi(s,\check{X}_s^{t,\xi,a}) + \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a) + D\varphi(s,\check{X}_s^{t,\xi,a}).b(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a) \\
& + \frac{1}{2}D^2\varphi(s,\check{X}_s^{t,\xi,a})(\sigma(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a)N).\sigma(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a)N \\
& + \frac{1}{2}D^2\varphi(s,\check{X}_s^{t,\xi,a})(\sigma_0(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a)).\sigma_0(\check{X}_s^{t,\xi,a},\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}},a)\big]ds\Big] \\
=:\; & \mathbb{E}^0\Big[\frac{1}{h}\int_t^{\theta_h} F_s(t,\xi,a)ds\Big], \qquad (4.14)
\end{aligned}
$$

with $N \in L^2(\mathcal{G};\mathbb{R}^n)$ of zero mean, and unit variance, and independent of $(\tilde{B},\xi)$. It is known that, for any continuous function ψ, the map $s \in [t,T] \mapsto \tilde{\mathbb{E}}^1[\psi(\check{X}_s^{t,\xi,a})] = \mathbb{E}[\psi(X_s^{t,\xi,a}|\mathcal{F}^0] = \rho_s^{t,\mu,a}(\psi)$ (for $\mu = \mathcal{L}(\xi)$) is càd-làg $\mathbb{P}^0$-a.s. (see e.g. Theorem 2.9 in [4]). Recalling also that $(\check{X}_s^{t,\xi,a})_{t\leq s\leq T}$ and $\{\tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,a}} = \rho_s^{t,\mu,a}, t \leq s \leq T\}$ have continuous trajectories, we deduce that the process $(F_s(t,\xi,a))_{t\leq s\leq T}$ has càd-làg paths $\mathbb{P}^0$ almost surely. Moreover, by (standard) Itô's formula, we have for all $t \leq s \leq T$,

$$
\begin{aligned}
\tilde{\mathbb{E}}^1\big[|\check{X}_s^{t,\xi,a}-\xi|^2\big] \;=\; \mathbb{E}\big[|X_s^{t,\xi,a}-\xi|^2|\mathcal{F}^0\big] \;=\; & \int_t^s \mathbb{E}\big[2(X_r^{t,\xi,a}-\xi).b_r + \sigma_r\sigma_r^\intercal + \sigma_r^0(\sigma_r^0)^\intercal|\mathcal{F}^0\big]dr \\
& + \int_t^s \mathbb{E}\big[2(X_r^{t,\xi,a}-\xi)^\intercal\sigma_r^0|\mathcal{F}^0\big]dW_r^0,
\end{aligned}
$$

where we set $b_s = b(X_s^{t,\xi,a},\mathbb{P}^{W^0}_{X_s^{t,\xi,a}},a)$, $\sigma_s = \sigma(X_s^{t,\xi,a},\mathbb{P}^{W^0}_{X_s^{t,\xi,a}},a)$, $\sigma_s^0 = \sigma_0(X_s^{t,\xi,a},\mathbb{P}^{W^0}_{X_s^{t,\xi,a}},a)$. This shows that the map $s \in [t,T] \mapsto \tilde{\mathbb{E}}^1[|\check{X}_s^{t,\xi,a}-\xi|^2]$ is continuous $\mathbb{P}^0$-a.s., and thus $\theta_h(\omega^0) = t+h$ for h small enough ($\leq \bar{h}(\omega^0)$), $\mathbb{P}^0(d\omega^0)$-a.s. By the mean-value theorem, we then get $\mathbb{P}^0$ almost surely, $\frac{1}{h}\int_t^{\theta_h} F_s(t,\xi,a)ds \to F_t(t,\xi,a)$, as h goes to zero, and so from the dominated convergence theorem in (4.14):

$$
\begin{aligned}
0 \;\leq\; F_t(t,\xi,a) \;=\; & \partial_t\varphi(t,\xi) + \tilde{\mathbb{E}}^1\big[f(\xi,\mathcal{L}(\xi),a) + D\varphi(t,\xi).b(\xi,\mathcal{L}(\xi),a) \\
& + \frac{1}{2}D^2\varphi(t,\xi)(\sigma(\xi,\mathcal{L}(\xi),a)N).\sigma(\xi,\mathcal{L}(\xi),a)N \\
& + \frac{1}{2}D^2\varphi(s,\xi)(\sigma_0(\xi,\mathcal{L}(\xi),a)).\sigma_0(\xi,\mathcal{L}(\xi),a)\big].
\end{aligned}
$$

Since a is arbitrary in $\mathbf{A}$, this shows the required viscosity subsolution property.

(ii) Supersolution property. Fix $(t,\xi) \in [0,T) \times L^2(\mathcal{G};\mathbb{R}^d)$, and consider some test function $\varphi \in \mathcal{C}^2([0,T]\times L^2(\mathcal{G};\mathbb{R}^d))$ s.t. $v-\varphi$ has a minimum at (t,ξ), and w.l.o.g. $v(t,\xi) = \varphi(t,\xi)$, so that $v \geq \varphi$. From the continuity assumptions in **(H1)-(H2)**, we observe that the function $\mathcal{H}$ defined on $[0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ by

$$
\mathcal{H}(s,\zeta) \;:=\; H(\zeta, D\varphi(s,\zeta), D^2\varphi(s,\zeta)),
$$

is continuous. Then, given an arbitrary $\varepsilon > 0$, there exists $\bar{h} \in (0,T-t)$, $\delta > 0$ s.t. for all $s \in [t,t+\bar{h}]$, and $\zeta \in L^2(\mathcal{G};\mathbb{R}^d)$ with $\tilde{\mathbb{E}}^1[|\zeta-\xi|^2] \leq \delta$,

$$
\Big|\big(\partial_t\varphi + \mathcal{H}\big)(s,\zeta) - \big(\partial_t\varphi + \mathcal{H}\big)(t,\xi)\Big| \;\leq\; \varepsilon.
$$

From the second part (4.13) of the DPP, for any $h \in (0, \bar{h})$, there exists $\alpha \in \mathcal{A}$ s.t.

$$\varphi(t,\xi) + \varepsilon h \;\; \geq \;\; \mathbb{E}^0\Big[\int_t^{\theta_h} \tilde{\mathbb{E}}^1\big[f(\check{X}_s^{t,\xi,\alpha}, \tilde{\mathbb{P}}^1_{\check{X}_s^{t,\xi,\alpha}}, \alpha_s)\big]ds \;+\; \varphi(\theta_h, \check{X}_{\theta_h}^{t,\xi,\alpha})\Big],$$

where we take $\theta_h = \inf\{s \geq t : \tilde{\mathbb{E}}^1[|\check{X}_s^{t,\xi,\alpha} - \xi|^2] \geq \delta\} \wedge (t+h)$. Applying again Itô's formula (4.6) to $\varphi(s, \check{X}_s^{t,\xi,\alpha})$, and by definition of $\mathcal{H}$, we get

$$\begin{aligned} \varepsilon \;\; &\geq \;\; \mathbb{E}^0\Big[\frac{1}{h}\int_t^{\theta_h} \big(\partial_t\varphi + \mathcal{H}\big)(s, \check{X}_s^{t,\xi,\alpha})ds\Big] \\ &\geq \;\; \Big[\big(\partial_t\varphi + \mathcal{H}\big)(t,\xi) - \varepsilon\Big]\frac{\mathbb{E}^0[\theta_h] - t}{h}, \end{aligned} \tag{4.15}$$

by the choice of h, δ, and θ_h. Now, by noting from Chebyshev's inequality that

$$\begin{aligned} \mathbb{P}^0[\theta_h < t+h] \;\; &\leq \;\; \mathbb{P}^0\big[\sup_{t\leq s\leq t+h} \tilde{\mathbb{E}}^1[|\check{X}_s^{t,\xi,\alpha} - \xi|^2] \geq \delta\big] \\ &\leq \;\; \frac{\mathbb{E}^0\Big[\sup_{t\leq s\leq t+h} \tilde{\mathbb{E}}^1[|\check{X}_s^{t,\xi,\alpha} - \xi|^2]\Big]}{\delta} \;\; \leq \;\; \frac{C(1+\tilde{\mathbb{E}}^1[|\xi|^2])h}{\delta} \end{aligned}$$

and using the obvious inequality: $1 - \mathbb{P}^0[\theta_h < t+h] = \mathbb{P}[\theta_h = t+h] \leq \frac{\mathbb{E}^0[\theta_h]-t}{h} \leq 1$, we see that $\frac{\mathbb{E}^0[\theta_h]-t}{h}$ converges to 1 when h goes to zero, and deduce from (4.15) that

$$2\varepsilon \;\; \geq \;\; \big(\partial_t\varphi + \mathcal{H}\big)(t,\xi).$$

We obtain the required viscosity supersolution property by sending ε to zero.

(2) Uniqueness property. In view of our definition of viscosity solution, we have to show a comparison principle for viscosity solutions to the lifted Bellman equation (4.10). We use the comparison principle proved in Theorem 3.50 in [21] and only need to check that the hypotheses of this theorem are satisfied in our context for the lifted Hamiltonian H defined in (4.11). Notice that the Bellman equation (4.10) is a bounded equation in the terminology of [21] (see their section 3.3.1) meaning that there is no linear dissipative operator on $L^2(\mathcal{G};\mathbb{R}^d)$ in the equation. Therefore, the notion of B-continuity reduces to the standard notion of continuity in $L^2(\mathcal{G};\mathbb{R}^d)$ since one can take for B the identity operator. Their Hypothesis 3.44 follows from the uniform continuity of b, σ, σ_0 and f in **(H1)**-**(H2)**. Hypothesis 3.45 is immediately satisfied since there is no discount factor in our equation, i.e. H does not depend on v but only on its derivatives. The monotonicity condition in $Q \in S(L^2(\mathcal{G};\mathbb{R}^d))$ of 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, σ_0 in **(H1)**, and the uniform continuity condition on f in **(H2)**, while Hypothesis 3.49 follows from the growth condition of σ, σ_0 in **(H1)**. One can then apply Theorem 3.50 in [21] and conclude that comparison principle holds for the Bellman equation (4.10). □

We conclude this section with a verification theorem, which gives an analytic feedback form of the optimal control when there is a smooth solution to the Bellman equation.

Theorem 4.2 *(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 $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, $w(t,.) \in \mathcal{C}_b^2(\mathcal{P}_2(\mathbb{R}^d))$, and $w(.,\mu) \in C^1([0,T))$, and satisfying a quadratic growth condition as in (3.5). Suppose that w is solution to the Bellman equation (4.7), and there exists for all $(t,\mu) \in [0,T) \times \mathcal{P}_2(\mathbb{R}^d)$ an element $\hat{a}(t,\mu) \in \mathbf{A}$ attaining the infimum in (4.7) s.t. the map $(t,\mu) \mapsto \hat{a}(t,\mu)$ is measurable, and the stochastic McKean-Vlasov equation

$$\begin{aligned} d\hat{X}_s &= b(\hat{X}_s, \mathbb{P}^{W^0}_{\hat{X}_s}, \hat{a}(s, \mathbb{P}^{W^0}_{\hat{X}_s}))ds + \sigma(\hat{X}_s, \mathbb{P}^{W^0}_{\hat{X}_s}, \hat{a}(s, \mathbb{P}^{W^0}_{\hat{X}_s}))dB_s \\ &\quad + \sigma(\hat{X}_s, \mathbb{P}^{W^0}_{\hat{X}_s}, \hat{a}(s, \mathbb{P}^{W^0}_{\hat{X}_s}))dW^0_s, \quad t \le s \le T, \ \hat{X}_t = \xi, \end{aligned}$$

admits a unique solution denoted $(\hat{X}^{t,\xi}_s)_{t \le s \le T}$, for any $(t,\xi) \in [0,T] \times L^2(\mathcal{G};\mathbb{R}^d)$ (This is satisfied e.g. when $\mu \mapsto \hat{a}(t,\mu)$ is Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$). Then, $w = v$, and the feedback control $\alpha^* \in \mathcal{A}$ defined by

$$\alpha^*_s = \hat{a}(s, \mathbb{P}^{W^0}_{\hat{X}^{t,\xi}_s}), \quad t \le s < T, \tag{4.16}$$

is an optimal control for $v(t,\mu)$, i.e. $v(t,\mu) = J(t,\mu,\alpha^*)$, with $\mu = \mathcal{L}(\xi)$.

Proof. Fix $(t, \mu = \mathcal{L}(\xi)) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$, and consider some arbitrary control $\alpha \in \mathcal{A}$ associated to $\rho^{t,\mu,\alpha}_s = \mathbb{P}^{W^0}_{X^{t,\xi,\alpha}_s}$, $t \le s \le T$. Denote by $X'^{t,\xi,\alpha}_s$ a copy of $X^{t,\xi,\alpha}_s$ on another probability space $(\Omega' = \Omega^0 \times \Omega'^1 \, \mathcal{F}^0 \otimes \mathcal{F}'^1, \mathbb{P}^0 \times \mathbb{P}'^1)$, with $(\Omega'^1, \mathcal{F}'^1, \mathbb{P}'^1)$ supporting B' a copy of B. Applying Itô's formula (4.5) to $w(s, \rho^{t,\mu,\alpha}_s)$ between t and the $\mathbb{F}^0$-stopping time $\theta^n_T = \inf\{s \ge t : \|\rho^{t,\mu,\alpha}_s\|_2 \ge n\} \wedge T$, we obtain

$$\begin{aligned} & w(\theta^n_T, \rho^{t,\mu,\alpha}_{\theta^n_T}) \\ = \ & w(t,\mu) + \int_t^{\theta^n_T} \Big\{ \frac{\partial w}{\partial t}(s, \rho^{t,\mu,\alpha}_s) + \mathbb{E}_{W^0}\Big[\partial_\mu w(s, \rho^{t,\mu,\alpha}_s)(X^{t,\xi,\alpha}_s).b(X^{t,\xi,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s) \\ & + \frac{1}{2}\mathrm{tr}\big[\partial_x \partial_\mu w(s, \rho^{t,\mu,\alpha}_s)(X^{t,\xi,\alpha}_s)(\sigma\sigma^\intercal(X^{t,\xi,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s) + \sigma_0\sigma_0^\intercal(X^{t,\xi,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s))\big]\Big] \\ & + \mathbb{E}_{W^0}\Big[\mathbb{E}'_{W^0}\big[\frac{1}{2}\mathrm{tr}\big(\partial^2_\mu w(s, \rho^{t,\mu,\alpha}_s)(X^{t,\xi,\alpha}_s, X'^{t,\xi,\alpha}_s)\sigma_0(X^{t,\xi,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s)\sigma_0^\intercal(X'^{t,\xi,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s)\big)\big]\Big]\Big\} ds \\ & + \int_t^{\theta^n_T} \mathbb{E}_{W^0}\big[\partial_\mu w(s, \rho^{t,\mu,\alpha}_s)(X^{t,\mu,\alpha}_s)^\intercal \sigma_0(X^{t,\mu,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s)\big] dW^0_s \\ = \ & w(t,\mu) + \int_t^{\theta^n_T} \Big[\frac{\partial w}{\partial t}(s, \rho^{t,\mu,\alpha}_s) + \rho^{t,\mu,\alpha}_s\big(\mathbb{L}^{\alpha_s} w(s, \rho^{t,\mu,\alpha}_s)\big) + \rho^{t,\mu,\alpha}_s \otimes \rho^{t,\mu,\alpha}_s\big(\mathbb{M}^{\alpha_s} w(s, \rho^{t,\mu,\alpha}_s)\big)\Big] ds \\ & + \int_t^{\theta^n_T} \mathbb{E}_{W^0}\big[\partial_\mu w(s, \rho^{t,\mu,\alpha}_s)(X^{t,\mu,\alpha}_s)^\intercal \sigma_0(X^{t,\mu,\alpha}_s, \rho^{t,\mu,\alpha}_s, \alpha_s)\big] dW^0_s, \end{aligned} \tag{4.17}$$

by definition of $\mathbb{L}^a$ and $\mathbb{M}^a$ in (4.8)-(4.9), and recalling again that $\rho^{t,\mu,\alpha}_s = \mathbb{P}^{W^0}_{X^{t,\xi,\alpha}_s}$. Now,

the integrand of the stochastic integral w.r.t. W^0 in (4.17) satisfies:

$$\begin{aligned}
&\Big|\mathbb{E}_{W^0}\big[\partial_\mu w(s,\rho_s^{t,\mu,\alpha})(X_s^{t,\mu,\alpha})^\intercal \sigma_0(X_s^{t,\mu,\alpha},\rho_s^{t,\mu,\alpha},\alpha_s)\big]\Big|^2 \\
\leq\ & \Big(\int_{\mathbb{R}^d}\big|\partial_\mu w(s,\rho_s^{t,\mu,\alpha})(x)^\intercal \sigma_0(x,\rho_s^{t,\mu,\alpha},\alpha_s)\big|\rho_s^{t,\mu,\alpha}(dx)\Big)^2 \\
\leq\ & \int_{\mathbb{R}^d}\big|\partial_\mu w(s,\rho_s^{t,\mu,\alpha})(x)\big|^2\rho_s^{t,\mu,\alpha}(dx)\int_{\mathbb{R}^d}\big|\sigma_0(x,\rho_s^{t,\mu,\alpha},\alpha_s)\big|^2\rho_s^{t,\mu,\alpha}(dx) \\
\leq\ & C(1+n)\sup_{\|\pi\|_2\leq n}\int_{\mathbb{R}^d}\big|\partial_\mu w(s,\pi)(x)\big|^2\pi(dx)\ <\infty, \qquad t\leq s\leq \theta_T^n,
\end{aligned}$$

from Cauchy-Schwarz inequality, the linear growth condition of σ_0 in **(H1)**, the choice of θ_T^n, and condition (4.2). Therefore, the stochastic integral in (4.17) vanishes in $\mathbb{E}^0$-expectation, and we get

$$\begin{aligned}
\mathbb{E}^0\big[w(\theta_T^n,\rho_{\theta_T^n}^{t,\mu,\alpha})\big] &= w(t,\mu)\ +\ \mathbb{E}^0\Big[\int_t^{\theta_T^n}\frac{\partial w}{\partial t}(s,\rho_s^{t,\mu,\alpha})+\rho_s^{t,\mu,\alpha}\big(\mathbb{L}^{\alpha_s}w(s,\rho_s^{t,\mu,\alpha})\big) \\
&\qquad + \rho_s^{t,\mu,\alpha}\otimes\rho_s^{t,\mu,\alpha}\big(\mathbb{M}^{\alpha_s}w(s,\rho_s^{t,\mu,\alpha})\big)ds\Big] \\
&\geq w(t,\mu)\ -\ \mathbb{E}^0\Big[\int_t^{\theta_T^n}\hat{f}(\rho_s^{t,\mu,\alpha},\alpha_s)ds\Big], \qquad (4.18)
\end{aligned}$$

since w satisfies the Bellman equation (4.7). By sending n to infinity into (4.18), and from the dominated convergence theorem (under the condition that w, f satisfy a quadratic growth condition and recalling the estimation (3.12)), we obtain:

$$w(t,\mu)\ \leq\ J(t,\mu,\alpha)\ =\ \mathbb{E}^0\Big[\int_t^T\hat{f}(\rho_s^{t,\mu,\alpha},\alpha_s)ds+\hat{g}(\rho_T^{t,\mu,\alpha})\Big].$$

Since α is arbitrary in $\mathcal{A}$, this shows that $w\leq v$.

Finally, by applying the same Itô's argument with the feedback control $\alpha^*\in\mathcal{A}$ in (4.16), and noting that $\hat{X}_s^{t,\xi}=X_s^{t,\xi,\alpha^*}$, $\mathbb{P}^{W^0}_{\hat{X}_s^{t,\xi}}=\rho_s^{t,\mu,\alpha^*}$, we have now equality in (4.18), hence $w(t,\mu)=J(t,\mu,\alpha^*)$ $(\geq v(t,\mu))$, and thus finally the required equality: $w(t,\mu)=v(t,\mu)=J(t,\mu,\alpha^*)$. □

5 Linear quadratic stochastic McKean-Vlasov control

We consider the linear-quadratic (LQ) stochastic McKean-Vlasov control problem where the control set **A** is a functional space, which corresponds to the McKean-Vlasov problem with common noise as presented in the introduction.

The control set **A** is the set $L(\mathbb{R}^d;\mathbb{R}^m)$ of Lipschitz functions from $\mathbb{R}^d$ into $A=\mathbb{R}^m$, and we consider a multivariate linear McKean-Vlasov controlled dynamics with coefficients given by

$$\begin{aligned}
b(x,\mu,a) &= b_0+Bx+\bar{B}\bar{\mu}+Ca(x), \\
\sigma(x,\mu,a) &= \vartheta+Dx+\bar{D}\bar{\mu}+Fa(x), \qquad (5.1)\\
\sigma_0(t,x,\mu,a) &= \vartheta_0+D_0x+\bar{D}_0\bar{\mu}+F_0a(x),
\end{aligned}$$

for $(x,\mu,a) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times L(\mathbb{R}^d;\mathbb{R}^m)$, where we set

$$\bar{\mu} \;:=\; \int_{\mathbb{R}^d} x\mu(dx).$$

Here B, $\bar{B}$, D, $\bar{D}$, D_0, $\bar{D}_0$, are constant matrices in $\mathbb{R}^{d\times d}$, C, F, F_0 are constant matrices in $\mathbb{R}^{d\times m}$, and b_0, ϑ, ϑ_0 are constant vectors in $\mathbb{R}^d$. The quadratic cost functions are given by

$$\begin{array}{rcl} f(x,\mu,a) & = & x^\intercal Q_2 x + \bar{\mu}^\intercal \bar{Q}_2 \bar{\mu} + a(x)^\intercal R_2 a(x) \\ g(x,\mu) & = & x^\intercal P_2 x + \bar{\mu}^\intercal \bar{P}_2 \bar{\mu}, \end{array} \tag{5.2}$$

where Q_2, $\bar{Q}_2$, P_2, $\bar{P}_2$ are constant matrices in $\mathbb{R}^{d\times d}$, R_2 is a constant matrix in $\mathbb{R}^{m\times m}$. 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, 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 (3.4) are then given by

$$\left\{ \begin{array}{rcl} \hat{f}(t,\mu,a) & = & \mathrm{Var}(\mu)(Q_2) + \bar{\mu}^\intercal (Q_2 + \bar{Q}_2)\bar{\mu} + \overline{a \star \mu_2}(R_2) \\ \hat{g}(\mu) & = & \mathrm{Var}(\mu)(P_2) + \bar{\mu}^\intercal (P_2 + \bar{P}_2)\bar{\mu} \end{array} \right. \tag{5.3}$$

for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in \mathbf{A} = L(\mathbb{R}^d;\mathbb{R}^m)$, 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. $\mathcal{P}_2(\mathbb{R}^m)$):

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

and $a \star \mu \in \mathcal{P}_2(\mathbb{R}^m)$ is the image by $a \in L(\mathbb{R}^d;\mathbb{R}^m)$ of the measure $\mu \in \mathbb{R}^m$, so that

$$\overline{a \star \mu} \;=\; \int_{\mathbb{R}^d} a(x)\mu(dx), \qquad \overline{a \star \mu}_2(\Lambda) \;:=\; \int a(x)^\intercal \Lambda a(x) \mu(dx).$$

We look for a value function solution to the Bellman equation (4.7) in the form

$$w(t,\mu) \;=\; \mathrm{Var}(\mu)(\Lambda(t)) + \bar{\mu}^\intercal \Gamma(t)\bar{\mu} + \bar{\mu}^\intercal \gamma(t) + \chi(t), \tag{5.4}$$

for some functions Λ, $\Gamma \in C^1([0,T];\mathbb{S}^d)$, $\gamma \in C^1([0,T];\mathbb{R}^d)$, and $\chi \in C^1([0,T];\mathbb{R})$. One easily checks that w lies in $\mathcal{C}^{1,2}_b([0,T] \times \mathcal{P}_2(\mathbb{R}^d))$ with

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

Together with the quadratic expression (5.3) of $\hat{f}$, $\hat{g}$, we then see after some tedious but direct calculations that w satisfies the Bellman equation (4.7) iff

$$\begin{array}{rl} & \mathrm{Var}(\mu)(\Lambda(T)) + \bar{\mu}^\intercal \Gamma(T)\bar{\mu} + \bar{\mu}^\intercal \gamma(T) + \chi(T) \\ = & \mathrm{Var}(\mu)(P_2) + \bar{\mu}^\intercal (P_2 + \bar{P}_2)\bar{\mu}, \end{array} \tag{5.5}$$

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

$$\begin{aligned}
&\mathrm{Var}(\mu)\big(\Lambda'(t) + Q_2 + D^\intercal\Lambda(t)D + D_0^\intercal\Lambda(t)D_0 + \Lambda(t)B + B^\intercal\Lambda(t)\big) + \inf_{a\in L(\mathbb{R}^d;\mathbb{R}^m)} G_t^\mu(a) \\
&\quad + \bar{\mu}^\intercal\Big(\Gamma'(t) + Q_2 + \bar{Q}_2 + (D+\bar{D})^\intercal\Lambda(t)(D+\bar{D}) \\
&\qquad + (D_0+\bar{D}_0)^\intercal\Gamma(t)(D_0+\bar{D}_0) + \Gamma(t)(B+\bar{B}) + (B + \bar{B})^\intercal\Gamma(t)\Big)\bar{\mu} \\
&\quad + \bar{\mu}^\intercal\big(\gamma'(t) + (B+\bar{B})^\intercal\gamma(t) + 2(D+\bar{D})^\intercal\Lambda(t)\vartheta + 2(D_0+\bar{D}_0)^\intercal\Gamma(t)\vartheta_0 + 2\Gamma(t)b_0\big) \\
&\quad + \chi'(t) + \gamma(t)^\intercal b_0 + \vartheta^\intercal\Lambda(t)\vartheta + \vartheta_0{}^\intercal\Gamma(t)\vartheta_0 \\
&= 0,
\end{aligned} \tag{5.6}$$

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

$$\begin{aligned}
G_t^\mu(a) &= \mathrm{Var}(a\star\mu)(U_t) + \overline{a\star\mu}^\intercal V_t\overline{a\star\mu} + 2\int_{\mathbb{R}^d}(x-\bar{\mu})^\intercal S_t a(x)\mu(dx) \\
&\quad + 2\bar{\mu}^\intercal Z_t\overline{a\star\mu} + Y_t.\overline{a\star\mu},
\end{aligned}$$

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

$$\left\{\begin{array}{rcl}
U(t,\Lambda(t)) &=& F^\intercal\Lambda(t)F + F_0^\intercal\Lambda(t)F_0 + R_2, \\
V(t,\Lambda(t),\Gamma(t)) &=& F^\intercal\Lambda(t)F + F_0^\intercal\Gamma(t)F_0 + R_2 \\
S(t,\Lambda(t)) &=& D^\intercal\Lambda(t)F + D_0^\intercal\Lambda(t)F_0 + \Lambda(t)C + M_2, \\
Z(t,\Lambda(t),\Gamma(t)) &=& (D+\bar{D})^\intercal\Lambda(t)F + (D_0+\bar{D}_0)^\intercal\Gamma(t)F + \Gamma(t)C + M_2 \\
Y(t,\Gamma(t),\gamma(t)) &=& C^\intercal\gamma(t) + 2F^\intercal\Lambda(t)\vartheta + 2F_0^\intercal\Gamma(t)\vartheta_0.
\end{array}\right. \tag{5.7}$$

Then, under the condition that the symmetric matrices U_t and V_t in (5.7) are positive, hence invertible (this will be discussed later on), we get after square completion:

$$\begin{aligned}
G_t^\mu(a) &= \mathrm{Var}((a - a^*(t,.,\mu))\star\mu)(U_t) + \overline{(a-a^*(t,.,\mu))\star\mu}^\intercal V_t\overline{(a-a^*(t,.,\mu))\star\mu} \\
&\quad - \mathrm{Var}(\mu)\big(S_tU_t^{-1}S_t^\intercal\big) - \bar{\mu}^\intercal\big(Z_tV_t^{-1}Z_t^\intercal\big)\bar{\mu} - Y_t^\intercal V_t^{-1}Z_t^\intercal\bar{\mu} - \frac{1}{4}Y_t^\intercal V_t^{-1}Y_t.
\end{aligned}$$

where $a(t,.,\mu) \in L^(\mathbb{R}^d;\mathbb{R}^m)$ is given by

$$a^*(t,x,\mu) = -U_t^{-1}S_t^\intercal(x-\bar{\mu}) - V_t^{-1}Z_t^\intercal\bar{\mu} - \frac{1}{2}V_t^{-1}Y_t. \tag{5.8}$$

This means that G_t^μ attains its infimum at $a^*(t,.,\mu)$, and plugging the above expression of $G_t^\mu(a^*(t,.,\mu))$ in (5.6), we observe that the relation (5.5)-(5.6), hence the Bellman equation, is satisfied by identifying the terms in $\mathrm{Var}(.)$, $\bar{\mu}^\intercal(.)\bar{\mu}$, $\bar{\mu}$, which leads to the system of ordinary differential equations (ODEs) for $(\Lambda,\Gamma,\gamma,\chi)$:

$$\left\{\begin{array}{rcl}
\Lambda'(t) + Q_2 + D^\intercal\Lambda(t)D + D_0^\intercal\Lambda(t)D_0 + \Lambda(t)B + B^\intercal\Lambda(t) & & \\
-S(t,\Lambda(t))U(t,\Lambda(t))^{-1}S(t,\Lambda(t))^\intercal &=& 0, \\
\Lambda(T) &=& P_2,
\end{array}\right. \tag{5.9}$$

$$\left\{\begin{array}{rcl}
\Gamma'(t) + Q_2 + \bar{Q}_2 + (D+\bar{D})^\intercal\Lambda(t)(D+\bar{D}) & & \\
+(D_0+\bar{D}_0)^\intercal\Gamma(t)(D_0+\bar{D}_0) + \Gamma(t)^\intercal(B+\bar{B}) & & \\
+ (B+\bar{B})^\intercal\Gamma(t) - Z(t,\Lambda(t),\Gamma(t))V(t,\Lambda(t),\Gamma(t))^{-1}Z(t,\Lambda(t),\Gamma(t))^\intercal &=& 0, \\
\Gamma(T) &=& P_2 + \bar{P}_2,
\end{array}\right. \tag{5.10}$$

$$\left\{\begin{array}{rcl} \gamma'(t) + \big(B+\bar{B}\big)^\intercal\gamma(t) - Z(t,\Lambda(t),\Gamma(t))V(t,\Lambda(t),\Gamma(t))^{-1}Y(t,\Gamma(t),\gamma(t)) & & \\ + 2\big(D+\bar{D}\big)^\intercal\Lambda(t)\vartheta + 2\big(D_0+\bar{D}_0\big)^\intercal\Gamma(t)\vartheta_0 + 2\Gamma(t)b_0 & = & 0, \\ \gamma(T) & = & 0 \end{array}\right. \quad (5.11)$$

$$\left\{\begin{array}{rcl} \chi'(t) - \frac{1}{4}Y(t,\Gamma(t),\gamma(t))^\intercal V(t,\Lambda(t),\Gamma(t))^{-1}Y(t,\Gamma(t),\gamma(t)) & & \\ + \gamma(t)^\intercal b_0 + \vartheta^\intercal\Lambda(t)\vartheta + \vartheta_0^\intercal\Gamma(t)\vartheta_0 & = & 0, \\ \chi(T) & = & 0. \end{array}\right. \quad (5.12)$$

Therefore, the resolution of the Bellman equation in the LQ framework is reduced to the resolution of the Riccati equations (5.9) and (5.10) for Λ and Γ, and then given (Λ,Γ), to the resolution of the linear ODEs (5.11) and (5.12) for γ and χ. Suppose that there exists a solution $(\Lambda,\Gamma) \in C^1([0,T];\mathbb{S}^d) \times C^1([0,T];\mathbb{S}^d)$ to (5.9)-(5.10) s.t. (U_t,V_t) in (5.7) 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 $(x,\mu) \mapsto a^*(t,x,\mu)$ is Lipschitz on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, we deduce by the verification theorem that the value function v is equal to w in (5.4) with $(\Lambda,\Gamma,\gamma,\chi)$ solution to (5.9)-(5.10)-(5.11)-(5.12). Moreover, the optimal control is given in feedback form from (5.8) by

$$\begin{array}{rcl} \alpha^*_t(X^*_t) & = & a^*(t,X^*_t,\mathbb{P}^{W^0}_{X^*_t}) \\ & = & -U_t^{-1}S_t^\intercal\big(X^*_t - \mathbb{E}[X^*_t|\mathcal{F}^0_t]\big) - V_t^{-1}Z_t^\intercal\mathbb{E}[X^*_t|\mathcal{F}^0_t] \; - \; \dfrac{1}{2}V_t^{-1}Y_t, \end{array} \quad (5.13)$$

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

Remark 5.1 It is known from [37] that under the condition

$$P_2 \;\geq\; 0,\; P_2+\bar{P}_2 \;\geq\; 0, \qquad Q_2 \;\geq\; 0,\; Q_2+\bar{Q}_2 \;\geq\; 0, \quad R_2 \;\geq \delta I_m, \quad (5.14)$$

for some $\delta > 0$, the matrix Riccati equations (5.9)-(5.10) 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 (5.7) are symmetric positive definite matrices, i.e. lie in $\mathbb{S}^m_{>+}$ for all $t \in [0,T]$. The expression in (5.13) of the optimal control extends then to the case of stochastic LQ McKean-Vlasov control problem the feedback form obtained in [38] for LQ McKean-Vlasov without common noise, i.e. $\sigma_0 = 0$. □

Example: Interbank systemic risk model

We consider a model of inter-bank borrowing and lending studied in [15] 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:

$$\begin{array}{rcl} dX_t & = & \big[\kappa(\mathbb{E}[X_t|W^0] - X_t) + \alpha_t(X_t)\big]dt \\ & & + (\sigma_0 + \sigma_1 X_t)(\sqrt{1-\rho^2}dB_t + \rho dW^0_t),\; X_0 \;=\; x_0 \in \mathbb{R}. \end{array} \quad (5.15)$$

Here, $\kappa \geq 0$ is the rate of mean-reversion in the interaction from borrowing and lending between the banks, $\sigma_0 > 0$, $\sigma_1 \in \mathbb{R}$ are the affine coefficients of the volatility of the bank reserve, and there is a common noise W^0 for all the banks. This is a slight extension of the model considered in [15] where $\sigma_1 = 0$. Moreover, all banks can control their rate of

borrowing/lending to a central bank with the same feedback policy α in order to minimize a cost functional of the form

$$\begin{aligned} J(\alpha) &= \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\alpha_t(X_t)^2 - q\alpha_t(X_t)(\mathbb{E}[X_t|W^0] - X_t) + \frac{\eta}{2}(\mathbb{E}[X_t|W^0] - X_t)^2\Big)dt \\ &\quad + \frac{c}{2}(\mathbb{E}[X_T|W^0] - X_T)^2\Big], \end{aligned}$$

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. After square completion, we can rewrite the cost functional as

$$J(\alpha) = \mathbb{E}\Big[\int_0^T \Big(\frac{1}{2}\tilde{\alpha}_t(X_t)^2 + \frac{\eta - q^2}{2}(\mathbb{E}[X_t|W^0] - X_t)^2\Big)dt + \frac{c}{2}(\mathbb{E}[X_T|W^0] - X_T)^2\Big],$$

with $\tilde{\alpha}_t(X_t) = \alpha_t(X_t) - q(\mathbb{E}[X_t|W^0] - X_t)$. This model fits into the framework of (5.1)-(5.2) of the LQ stochastic McKean-Vlasov problem with

$$\begin{aligned} b_0 = 0,\ B = -(\kappa + q),\ \bar{B} = \kappa + q,\ C = 1, \\ D = \sigma_1\sqrt{1-\rho^2}, D_0 = \sigma_1\rho,\ \bar{D} = F = \bar{D}^0 = F^0 = 0,\ \vartheta = \sigma_0\sqrt{1-\rho^2},\ \vartheta_0 = \sigma_0\rho, \\ Q_2 = \frac{\eta - q^2}{2},\ \bar{Q}_2 = -\frac{\eta - q^2}{2},\ R_2 = \frac{1}{2}, P_2 = \frac{c}{2},\ \bar{P}_2 = -\frac{c}{2}. \end{aligned}$$

The Riccati system (5.9)-(5.10)-(5.11)-(5.12) for $(\Lambda(t), \Gamma(t), \gamma(t), \chi(t))$ is written in this case as

$$\left\{ \begin{array}{rcl} \Lambda'(t) - 2(\kappa + q - \frac{\sigma_1^2}{2})\Lambda(t) - 2\Lambda^2(t) + \frac{1}{2}(\eta - q^2) &=& 0, \quad \Lambda(T) = \frac{c}{2}, \\ \Gamma'(t) - 2\Gamma^2(t) + \sigma_1^2\rho^2\Gamma(t) + \sigma_1^2(1-\rho^2)\Lambda(t) &=& 0, \quad \Gamma(T) = 0, \\ \gamma'(t) - 2\Gamma(t)\gamma(t) + 2\sigma_0\sigma_1\rho^2\Gamma(t) + 2\sigma_0\sigma_1(1-\rho^2)\Lambda(t) &=& 0, \quad \gamma(T) = 0, \\ \chi'(t) - \frac{1}{2}\gamma^2(t) + \sigma_0^2\rho^2\Gamma(t) + \sigma_0^2(1-\rho^2)\Lambda(t) &=& 0, \quad \chi(T) = 0. \end{array} \right. \tag{5.16}$$

Assuming that $q^2 \leq \eta$, the explicit solution to the Riccati equation for Λ is given by

$$\Lambda(t) = \frac{1}{2}\frac{(\eta - q^2)\big(e^{(\delta^+ - \delta^-)(T-t)} - 1\big) + c\big(\delta^+ e^{(\delta^+ - \delta^-)(T-t)} - \delta^-\big)}{c\big(e^{(\delta^+ - \delta^-)(T-t)} - 1\big) + \delta^+ - \delta^- e^{(\delta^+ - \delta^-)(T-t)}} > 0,$$

where we set

$$\delta^\pm = -\big(\kappa + q - \frac{\sigma_1^2}{2}\big) \pm \sqrt{\big(\kappa + q - \frac{\sigma_1^2}{2}\big)^2 + \eta - q^2}.$$

Since $\Lambda \geq 0$, there exists a unique solution to the Riccati equation for Γ, and then γ, and finally χ are determined the linear ordinary differential equations in (5.16). Moreover, the functions (U_t, V_t, Z_t, Y_t) in (5.7) are explicitly given by: $U_t = V_t = \frac{1}{2}$ (hence > 0), $S_t = \Lambda(t) + \frac{q}{2}$, $Z_t = \Gamma(t)$, $Y_t = \gamma(t)$. Therefore, the optimal control is given in feedback form from (5.13) by

$$\begin{aligned} \alpha_t^*(X_t^*) &= a^*(t, X_t^*, \mathbb{P}_{X_t^*}) \\ &= -(2\Lambda(t) + q)(X_t^* - \mathbb{E}[X_t^*|W^0]) - 2\Gamma(t)\mathbb{E}[X_t^*|W^0] - \gamma(t), \end{aligned} \tag{5.17}$$

where X^* is the optimal log-monetary reserve controlled by the rate of borrowing/lending α^*. Moreover, denoting by $\bar{X}_t^* = \mathbb{E}[X_t^*|W^0]$ the conditional mean of the optimal log monetary reserve, we see that $\mathbb{E}[\alpha_t^*(X_t^*)|W^0] = -2\Gamma(t)\bar{X}_t^* - \gamma(t)$, and thus $\bar{X}^*$ is given from (5.15) by

$$d\bar{X}_t^* \quad = \quad -\big(2\Gamma(t)\bar{X}_t^* + \gamma(t)\big)dt + (\sigma_1\bar{X}_t^* + \sigma_0)\rho dW_t^0, \quad \bar{X}_0^* = x_0.$$

When $\sigma_1 = 0$, we have $\Gamma(t) = \gamma(t) = 0$, hence $\bar{X}_t^* = x_0 + \sigma_0\rho W_t^0$, and we retrieve the expression found in [15] by sending the number of banks N to infinity in their formula for the optimal control of the borrowing/lending rate:

$$\alpha_t^*(X_t^*) \quad = \quad -(2\Lambda(t) + q)(X_t^* - x_0 - \sigma\rho W_t^0), \quad 0 \leq t \leq T.$$

References

[1] Ahuja S. (2016): "Wellposedness of mean field games with common noise under a weak monotonicity condition", *SIAM J. Cont. and Optimization*, **54**, 30-48.

[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] A. Bain and D. Crisan (2009): *Fundamentals of stochastic filtering*, Series Stochastic Modelling and Applied Probability, **60**, Springer, New York.

[5] 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.

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

[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. and J. Ma (2015): "A mean-field stochastic control problem with partial observations", preprint.

[9] 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*.

[10] 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

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

[12] 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.

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

[14] Carmona R., Delarue F. and D. Lacker (2014): "Mean field games with common noise", to appear in *Annals of Probability*.

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

[16] Carmona R. and X. Zhu (2014): "A probabilistic approach to mean field games with major and minor players", arXiv: 1409.7141v1, to appear in *Annals of Applied Probability*.

[17] 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

[18] Claisse J., Talay D. and X. Tan (2015): "A pseudo-Markov property for controlled diffusion processes", http://fr.arxiv.org/pdf/1501.03939

[19] Dawson D. and J. Vaillancourt (1995): "Stochastic McKean-Vlasov equations", *Nonlinear Differential Equations and Applications*, **2**(2), 199-229.

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

[21] 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

[22] 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.

[23] Fuhrman, M. and H. Pham (2015): "Randomized and backward SDE representation for optimal control of non-Markovian SDEs", *Annals of Applied Probability*, **25**, 2134-2167.

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

[25] 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.

[26] Krylov N. (1980): *Controlled diffusion processes*, Springer, Series Applications of Mathematics, **14**.

[27] Kurtz T. and J. Xiong (1999): "Particle representations for a class of nonlinear SPDEs", *Stochastic Processes and their Applications*, **83(1)**, 103-126.

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

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

[30] 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.

[31] 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.

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

[33] Pham H. and X. Wei (2015): "Bellman equation and viscosity solutions for mean-field stochastic control problem", http://arxiv.org/pdf/1512.07866.pdf

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

[35] Soner M. and N. Touzi (2002): "Dynamic programming for stochastic target problems and geometric flows", *J. Eur. Math. Soc.*, **4**, 201-236.

[36] Wagner D. (1980): "Survey of measurable selection theorems: an update", Lect. Notes in Math., 794, Springer Verlag.

[37] Wonham W. (1968): "On a matrix Riccati equation of stochastic control", *SIAM J. on Control and Optimization*, **6**, 681-697.

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