SIAM J. CONTROL OPTIM.
Vol. 57, No. 1, pp. 693–717

STOCHASTIC CONTROL WITH DELAYED INFORMATION AND RELATED NONLINEAR MASTER EQUATION*

YURI F. SAPORITO† AND JIANFENG ZHANG‡

Abstract. In this paper we study stochastic control problems with delayed information, that is, the control at time t can depend only on the information observed before time $t-h$ for some delay parameter H. Such delay occurs frequently in practice and can be viewed as a special case of partial observation. When the time duration T is smaller than H, the problem becomes a deterministic control problem in the stochastic setting. While seemingly simple, the problem involves certain time inconsistency issues, and the value function naturally relies on the distribution of the state process and thus is a solution to a nonlinear master equation. Consequently, the optimal state process solves a McKean–Vlasov SDE. In the general case that T is larger than H, the master equation becomes path-dependent and the corresponding McKean–Vlasov SDE involves the conditional distribution of the state process. We shall build these connections rigorously and obtain the existence of a classical solution of these nonlinear (path-dependent) master equations in some special cases.

Key words. information delay, partial observation, master equation, McKean–Vlasov SDE, functional Itô formula

AMS subject classifications. 60H30, 93E20

DOI. 10.1137/17M1154746

1. Introduction.

Consider a stochastic control problem

$$
\begin{aligned}
V_0 &= \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[g(X_T^\alpha) + \int_0^T f(t, X_t^\alpha, \alpha_t)dt\Big],\\
\text{where}\quad X_t^\alpha &= x + \int_0^t b(s, X_s^\alpha, \alpha_s)ds + \int_0^t \sigma(s, X_s^\alpha, \alpha_s)dW_s,
\end{aligned}
\tag{1.1}
$$

and $\mathcal{A}$ is an appropriate set of A-valued admissible controls. It is well known that, under mild conditions, $V_0 = u(0,x)$, where u is the solution of an HJB equation. One standard but crucial condition in the literature is that the admissible control is $\mathbb{F}$-progressively measurable, where $\mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}$ is a filtration under which W is a Brownian motion.

Our paper is mainly motivated by the following practical consideration. Note that $\mathcal{F}_t$ stands for the information the player observes over time period $[0,t]$. In many practical situations, the player needs some time to collect and/or to analyze the information, including numerical computations. Thus, the control α_t the player needs to act at time t may not be able to utilize the most recent information, or, say, there is some information delay. To be precise, let $\mathrm{H} > 0$ be a fixed constant standing for the delay parameter. In this paper we shall study the control problem (1.1) by restricting the admissible control α in $\mathcal{A}_0^{\mathrm{H}}$, i.e., such that $\alpha_t \in \mathcal{F}_{(t-\mathrm{H})^+}$, for all $t\in[0,T]$. This can be viewed as a special case of stochastic controls with *partial observation*. For a literature review, see section 1.1.

*Received by the editors October 31, 2017; accepted for publication (in revised form) December 27, 2018; published electronically February 19, 2019.
http://www.siam.org/journals/sicon/57-1/M115474.html
Funding: The second author's research was supported in part by NSF grant DMS 1413717.
†Escola de Matemática Aplicada, Fundação Getulio Vargas, Rio de Janeiro, Brazil (yuri.saporito@gmail.com).
‡Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532 (jianfenz@usc.edu).

We first consider the simple case that $T \le \text{H}$. Then $\alpha_t \in \mathbb{L}^0(\mathcal{F}_0)$ for all $t \in [0,T]$, and thus this is a deterministic control problem (assuming $\mathcal{F}_0$ is degenerate), but in a stochastic framework. While seemingly simpler, the constraint that the control is deterministic actually makes the problem more involving. The main reason is that such a problem is time inconsistent if one follows the standard approach. Intuitively, for problem (1.1) the optimal control $\alpha_t^* = \alpha^*(t, X_t^*)$ may typically depend on the corresponding state process X_t^* and thus is random. If the control is deterministic, the optimal control (assuming its existence) $\alpha_t^* = \alpha^*(t,x)$ should typically depend only on the initial value x for all $t \in [0,T]$. When one considers a dynamic problem over time period $[t_0, T]$, the new "deterministic" optimal control will become $\tilde{\alpha}_t^* = \tilde{\alpha}^*(t, X_{t_0}^*)$ for all $t \in [t_0, T]$, which will be $\mathcal{F}_{t_0}$-measurable rather than $\mathcal{F}_0$-measurable, and thus typically $\tilde{\alpha}_t^* \neq \alpha_t^*$ for $t \ge t_0$. That is, the problem is time inconsistent.

We aim to solve the problem in a time consistent way. Note again that, in a standard control problem, the optimal control reacts to the state process X_t^*. If the control is deterministic, and we still want the optimal one to react to the state process in some way, the most natural choice would be that α_t^* reacts to the law of X_t^*. This is indeed true. At time t_0, instead of specifying a value of X_{t_0}, we shall specify the distribution μ of X_{t_0} and define the value $V(t_0, \mu)$ for optimization over $[t_0, T]$ with deterministic control. It turns out that this dynamic problem is time consistent. The function V satisfies an appropriate dynamic programming principle and is the solution of a so-called master equation. Moreover, $V_0 = V(0, \delta_x)$, where x is the initial value X_0 and δ_x is the Dirac-measure of x.

To understand the master equation, we remark that $V : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a deterministic function, where $\mathcal{P}_2(\mathbb{R}^d)$ is the set of square-integrable probability measures on $\mathbb{R}^d$. It is known that the derivative of V in terms of $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ takes the form $\partial_\mu V : (t, \mu, x) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$. Denote by $\partial_x \partial_\mu V$ the standard derivative of $\partial_\mu V$ with respect to x. Then the optimization problem (1.1) with deterministic control is associated with the following HJB type of master equation:

$$\partial_t V(t,\mu) + H(t, \mu, \partial_\mu V, \partial_x \partial_\mu V) = 0, \quad V(T, \mu) = \mathbb{E}[g(\xi)], \tag{1.2}$$

where, for $p : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ and $q : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$,

$$\begin{aligned} H(t,\mu,p,q) &:= \sup_{a \in A} h(t,\mu,p,q,a), \\ h(t,\mu,p,q,a) &:= \mathbb{E}\Big[b(t,\xi,a) \cdot p(t,\mu,\xi) + \frac{1}{2}\sigma\sigma^\top(t,\xi,a) : q(t,\mu,\xi) + f(t,\xi,a)\Big]. \end{aligned} \tag{1.3}$$

Here ξ is a random variable with law μ. We shall prove the existence of classical solutions for a special case of (1.2), which to our best knowledge is new in the literature.

Assume further that the Hamiltonian H has optimal argument $a^* = I(t,\mu)$ for some function $I : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to A$; then the optimal control is $\alpha_t^* = I(t, \mathcal{L}_{X_t^*})$, where $\mathcal{L}_\xi$ is the law of the random variable ξ and X^* solves the following McKean–Vlasov SDE (assuming its well-posedness):

$$X_t^* = x + \int_0^t b\big(s, X_s^*, I(s, \mathcal{L}_{X_s^*})\big) ds + \int_0^t \sigma\big(s, X_s^*, I(s, \mathcal{L}_{X_s^*})\big) dW_s. \tag{1.4}$$

We shall carry out the verification theorem rigorously when I is continuous.

We finally consider the general case $T > \text{H}$. In this case, for $t > \text{H}$, the control α_t is required to be $\mathcal{F}_{t-\text{H}}$-measurable. Motivated by both theoretical and practical considerations, we shall use closed-loop controls, namely, $\alpha_t = \alpha_t(X_{[0,t-\text{H}]})$ is $\mathcal{F}_{t-\text{H}}$-measurable. Then the value function, $V(t, \mu_{[0,t]})$, will be path-dependent in the sense that $\mu_{[0,t]}$ denotes the law of the stopped process $X_{[0,t]}$, and the master equation (1.2) becomes a path-dependent equation. Consequently, the McKean–Vlasov SDE (1.4) will involve the conditional law of X_t^*.

We finish this section with a thorough comparison of different problems and methods that relate to the ones proposed here, which is further developed in Appendix B. The rest of the paper will be organized as follows. We discuss the deterministic case in section 2. Moreover, a special case is fully developed in section 3, and the general theory is presented in section 4.

1.1. Comparison to similar control problems and methods. As mentioned above, problem (1.1) with $\alpha \in \mathcal{A}_0^{\text{H}}$ might be seen as a special case of stochastic controls with partial observation. Generally, these stochastic control problems assume that the admissible controls are adapted to a smaller filtration $\mathbb{G}$, i.e., $\mathcal{G}_t \subset \mathcal{F}_t$, for all $t \in [0, T]$. Few papers have tackled this problem under this generality; see, for instance, Christopeit (1980), where the existence of the optimal control is studied, and Baghery and Øksendal (2007), where a maximum principle was derived under the more general Lévy processes.

Additionally to the delayed case studied in our paper, a very important example of the partial observation problem is the case of noisy observation. This situation has drawn significantly more attention than the other types of partially observed systems. For references, see Bensoussan (1992), Fleming and Pardoux (1982), Fleming (1980), (1982), Bismut (1982), Tang (1998), and the more recent Bandini et al. (2018), (2016). In these references, a *separated* control problem is proposed and studied. The optimal control problem of the partially observed system is connected to this separated problem, which is completely observed, using stochastic nonlinear filtering. It is worth noticing that, similarly to what we have found in the control with delayed information, the state variable of the separated control problem is an unnormalized conditional distribution measure and the class of admissible controls is a set of probability measures. Moreover, the dynamics of the aforementioned unnormalized conditional distribution measure is given by the so-called Zakai's equation.

Moreover, in Bandini et al. (2016), the authors have derived, in this context of noisy observation, the dynamic programming principle with flow of probability measures as state variable and the verification theorem of their master equation. Since the deterministic control problem studied in section 2 is a particular case of the noisy observation problem, our master equation and the dynamic programming principle in this section could be seen as a special case of theirs. However, our arguments here are much simpler, due to our special setting, and will be important for the general case in section 4, so we decided to report our proofs in detail so that the readers can easily grasp the main ideas.

In a different direction, although analyzing the same control problem as in the references in the paragraph above, Mortensen (1966) and Beneš and Karatzas (1983) have studied the value of the control problem as a function of the initial conditional probability density and an HJB equation analogous to our master equation (1.2) was derived. Moreover, an Itô formula for functions of density-valued processes was proved; cf. Lemma 2.7. Under the assumption that the agent observes pure independent noise, it turns out that their control problem is equivalent to our deterministic control problem in section 2. Moreover, when restricting to only those measures with density, our master equation (1.2) is equivalent to Mortensen's HJB equation. In order to verify this, one needs to understand the relation between Gâteaux derivatives with respect to the probability densities and $\partial_\mu V$; see Bensoussan et al. (2017). For more details, see Appendix B.

Furthermore, in the direction of applications of the delayed information setting to mathematical finance, Ichiba and Mousavi (2017) have proposed a discrete-time binomial model with delayed information for the price of asset. They studied the superreplication of derivatives with convex payoffs and also the convergence of their model to a continuous-time one (without delay). On the other hand, Bayraktar and Zhou (2016) studied an optimal stopping problem where the player has inside information instead of delayed information. In the contexts of stochastic control, this amounts to, say, α_t can be $\mathcal{F}_{t+\text{H}}$-measurable, and will be left for our future research.

A different aspect of delay in control problems is when the control chosen in a previous time, for instance, at $t-\text{H}$, influences the dynamics and/or the cost function at time t. In the literature, this is usually called stochastic control problems with *delay in the control*; see, for example, Gozzi and Marinelli (2006), Gozzi and Masiero (2017), Alekal et al. (1971), and Chen and Wu (2011). More generally, path dependence in the control was studied in Saporito (2017) in the framework of functional Itô calculus. This type of delay in the control is fundamentally different from the one we study here. Notice that although the control acts with delay, the agent has full information at time t to choose α_t. This departs completely from the setting we are proposing in this paper. Moreover, as one could easily notice from the aforesaid references, the value function V is not seen as a function of probability measures but as a function of the history of the state process. This type of delay in the control was recently applied to the study of systemic risk of a system of banks in Carmona et al. (2018).

We remark that the McKean–Vlasov SDE (for the forward state process) and the master equation (for the backward value function) have received very strong attention in recent years, mainly due to their application in mean field games and systemic risk; see Caines et al. (2006) and Larsy and Lions (2007), as well as Cardaliaguet (2013), Bensoussan et al. (2013), Carmona and Delarue (2017a), (2017b), and the references therein. In particular, Pham and Wei (2018), Bayraktar et al. (2018), and Wu and Zhang (2018) studied stochastic control problems for McKean–Vlasov dynamics under various type of controls and derived the dynamic programming principle as well as the master equations. In those applications, a large number of players are involved and the measure μ is introduced to characterize the aggregate behavior of the players. Our motivation here is quite different. We also remark that our paper deals with control problems and the master equation is nonlinear in $\partial_\mu V$ (and/or $\partial_x\partial_\mu V$). For mean field game problems, the master equation involves $V(t,x,\mu)$ and has a quite different nature. On one hand those master equations are nonlocal, and on the other hand they are typically nonlinear in $\partial_x V$ but linear in $\partial_\mu V$. In fact, in some literature, master equations refer to only those for mean field games, while the equations for control problems are called HJB equations in Wasserstein space. We nevertheless call both master equations since they share many features. In a special case, we will prove the existence of classical solutions for the nonlinear master equation (1.2). In general it is difficult to obtain classical solutions for master equations; some positive results include Buckdahn et al. (2017), Cardaliaguet et al. (2015), and Chassagneux, Crisan, and Delarue (2014), where the equations are linear in $\partial_\mu V$ and $\partial_x\partial_\mu V$, and Gangbo and Swiech (2015) and Bensoussan and Yam (2018), where the equations are of first order (without involving $\partial_x\partial_\mu V$). We also refer to Pham and Wei (2018) and Wu and Zhang (2018) for viscosity solutions of master equations.

Furthermore, although we are considering *control* problems, the delayed observation aspect of our setting is present in Bensoussan et al. (2015), (2017), which study

Stackelberg stochastic *games* with delayed information. A simple version of these games can described by two players: a leader and a follower. The leader has full information of both players and the follower has delayed information of the leader state variable (and full information of him/herself). In these references, the authors studied the convergence of the system of N-players to its mean field counterpart. Moreover, in the linear quadratic case, they were able to analyze and derive exact formulas for the mean field game.

2. The deterministic control problem. We remark that this case is the intersection of several related works. For example, Hu and Tang (2017) studied the linear quadratic case by using the stochastic maximum principle (see Appendix A); Beneš and Karatzas (1983) derived a similar equation when the measures have a density (see Appendix B); in particular, our master equation (2.12)–(2.13) and the dynamic programming principle theorem, Theorem 2.3 below, are already covered by Bandini et al. (2016) as a special case. However, since the arguments here are much simpler due to the special structure, which could be helpful for readers to grasp the main ideas, and more importantly since these arguments will be important for the general case in section 4, we still provide the details.

Let $T > 0$ be a fixed time horizon, $(\Omega, \mathbb{F}, \mathbb{P})$ a filtered probability space on $[0, T]$, and W an $\mathbb{F}$-Brownian motion under $\mathbb{P}$. In this section we assume $T \leq \text{H}$ and thus the controls are deterministic. Denote by $\mathcal{P}_2(\mathbb{R}^d)$ the set of square-integrable measures on $\mathbb{R}^d$, and for each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, denote $\mathbb{L}^2_\mu(\mathcal{F}_t) := \{\xi \in \mathbb{L}^2(\mathcal{F}_t) : \mathcal{L}_\xi = \mu\}$, where $\mathcal{L}_\xi$ denotes the law of ξ and $\mathbb{L}^2(\mathcal{F}_t)$ is the space of $\mathcal{F}_t$-measurable square-integrable random variables. For technical convenience, we shall assume $\mathcal{F}_0$ is rich enough such that $\mathbb{L}^2_\mu(\mathcal{F}_0) \neq \emptyset$ for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. However, in this and the next section we nevertheless assume the controls are deterministic, rather than $\mathcal{F}_0$-measurable. Finally, denote $\Theta := [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\overline{\Theta} := \{(t, \xi) : t \in [0, T], \xi \in \mathbb{L}^2(\mathcal{F}_t)\}$.

2.1. The control problem. Let A be an (arbitrary) measurable set in certain Euclidian space and $\mathcal{A}_t$ the set of all Borel measurable functions $\alpha : [t, T] \to A$. For any $(t, \xi) \in \overline{\Theta}$ and $\alpha \in \mathcal{A}_t$, define

$$
\begin{aligned}
X^{t,\xi,\alpha}_s &= \xi + \int_t^s b(r, X^{t,\xi,\alpha}_r, \alpha_r)dr + \int_t^s \sigma(r, X^{t,\xi,\alpha}_r, \alpha_r)dW_r, \ \ s \in [t, T], \\
J(t, \xi, \alpha) &:= \mathbb{E}\left[g(X^{t,\xi,\alpha}_T) + \int_t^T f(s, X^{t,\xi,\alpha}_s, \alpha_s)ds\right],
\end{aligned}
\tag{2.1}
$$

where b, σ, f, g are deterministic functions with appropriate dimensions.

Assumption 2.1. (i) b, σ, f, g are measurable in all their variables, and $b(t, 0, a)$, $\sigma(t, 0, a)$, $f(t, 0, a)$ are bounded;

(ii) b, σ are uniformly Lipschitz continuous in x and uniformly continuous in t;

(iii) f is uniformly continuous in (t, x), and g is uniformly continuous in x.

Under Assumption 2.1, clearly the SDE in (2.1) is well-posed and

$$
|J(t, \xi, \alpha)| \leq C[1 + \|\xi\|_{\mathbb{L}^2}].
\tag{2.2}
$$

Moreover, the following result is obvious.

LEMMA 2.2. *Under Assumption* 2.1, *the mapping* $\xi \mapsto J(t, \xi, \alpha)$ *is law invariant. That is, if* $\mathcal{L}_\xi = \mathcal{L}_{\xi'}$, *then* $J(t, \xi, \alpha) = J(t, \xi', \alpha)$.

We are now ready to introduce the optimization problem:

$$V(t,\mu) := \sup_{\alpha\in\mathcal{A}_t} J(t,\xi,\alpha), \quad (t,\mu)\in\Theta \text{ and } \xi\in\mathbb{L}^2_\mu(\mathcal{F}_t). \tag{2.3}$$

By (2.2), $V(t,\mu)$ is finite. We emphasize that V does not depend on the choice of ξ, thanks to Lemma 2.2. Throughout this section, when there is no confusion, for a given $(t,\mu)\in\Theta$ we shall always use ξ to denote some random variable in $\mathbb{L}^2_\mu(\mathcal{F}_t)$, and the claimed results will not depend on the choice of ξ.

We next establish the dynamic programming principle for V.

THEOREM 2.3. *Let Assumption* 2.1 *hold. Then, for any* $(t_1,\mu)\in\Theta$, $t_2\in(t_1,T]$,

$$V(t_1,\mu) = \sup_{\alpha\in\mathcal{A}_{t_1}} \Big[V\big(t_2,\mathcal{L}_{X^{t_1,\xi,\alpha}_{t_2}}\big) + \int_{t_1}^{t_2}\mathbb{E}[f(s,X^{t_1,\xi,\alpha}_s,\alpha_s)]ds\Big]. \tag{2.4}$$

Proof. For notational simplicity, we assume $t_1=0$ and $t_2=t$; then (2.4) becomes

$$V(0,\mu) = \widetilde{V}(0,\mu) := \sup_{\alpha\in\mathcal{A}_0}\Big[V\big(t,\mathcal{L}_{X^{0,\xi,\alpha}_t}\big) + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha}_s,\alpha_s)]ds\Big]. \tag{2.5}$$

First, for any $\alpha\in\mathcal{A}_0$, by the flow property for the SDE we have

$$X^{0,\xi,\alpha}_s = X^{t,X^{0,\xi,\alpha}_t,\alpha'}_s, \quad s\in[t,T],$$

where $\alpha' := \alpha\big|_{[t,T]}\in\mathcal{A}_t$. Then,

$$\begin{aligned} J(0,\xi,\alpha) &= \mathbb{E}\Big[g(X^{t,X^{0,\xi,\alpha}_t,\alpha'}_T) + \int_t^T f(s,X^{t,X^{0,\xi,\alpha}_t,\alpha'}_s,\alpha'_s)ds + \int_0^t f(s,X^{0,\xi,\alpha}_s,\alpha_s)ds\Big] \\ &= J(t,X^{0,\xi,\alpha}_t,\alpha') + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha}_s,\alpha_s)]ds \qquad (2.6)\\ &\le V\big(t,\mathcal{L}_{X^{0,\xi,\alpha}_t}\big) + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha}_s,\alpha_s)]ds \le \widetilde{V}(0,\mu). \end{aligned}$$

By the arbitrariness of α, we obtain $V(0,\mu)\le\widetilde{V}(0,\mu)$.

Next, for any $\varepsilon>0$, by the definition of $\widetilde{V}(0,\mu)$, there exists $\alpha^\varepsilon\in\mathcal{A}_0$ such that

$$V\big(t,\mathcal{L}_{X^{0,\xi,\alpha^\varepsilon}_t}\big) + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha^\varepsilon}_s,\alpha^\varepsilon_s)]ds \ge \widetilde{V}(0,\mu) - \frac{\varepsilon}{2}.$$

Moreover, by the definition of $V\big(t,\mathcal{L}_{X^{0,\xi,\alpha^\varepsilon}_t}\big)$ there exists $\widetilde{\alpha}^\varepsilon\in\mathcal{A}_t$ such that

$$J(t,X^{0,\xi,\alpha^\varepsilon}_t,\widetilde{\alpha}^\varepsilon) \ge V\big(t,\mathcal{L}_{X^{0,\xi,\alpha^\varepsilon}_t}\big) - \frac{\varepsilon}{2}.$$

Note that $\hat{\alpha}^\varepsilon := \alpha\mathbf{1}_{[0,t)} + \alpha'\mathbf{1}_{[t,T]}\in\mathcal{A}_0$. Then, by the middle line of (2.6),

$$\begin{aligned} V(0,\mu) \ge J(0,\xi,\hat{\alpha}^\varepsilon) &= J(t,X^{0,\xi,\alpha^\varepsilon}_t,\widetilde{\alpha}^\varepsilon) + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha^\varepsilon}_s,\alpha^\varepsilon_s)]ds \\ &\ge V\big(t,\mathcal{L}_{X^{0,\xi,\alpha^\varepsilon}_t}\big) + \int_0^t \mathbb{E}[f(s,X^{0,\xi,\alpha^\varepsilon}_s,\alpha^\varepsilon_s)]ds - \frac{\varepsilon}{2} \ge \widetilde{V}(0,\mu) - \varepsilon. \end{aligned}$$

Because $\varepsilon>0$ is arbitrary, we obtain $V(0,\mu)\ge\widetilde{V}(0,\mu)$. ∎

Remark 2.4. Since we are in the simple setting of deterministic control, no regularity or even measurability of V in terms of (t,μ) is needed in the above result.

2.2. The master equation. In this subsection we derive the master equation associated with the value function V. For this purpose, we first introduce the 2-Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$: for $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\mathcal{W}_2(\mu, \mu') := \inf\big\{\|\xi - \xi'\|_{\mathbb{L}^2} : \xi \in \mathbb{L}^2_\mu(\mathcal{F}_T), \xi' \in \mathbb{L}^2_\mu(\mathcal{F}_T)\big\}. \tag{2.7}$$

Let $V : \Theta \to \mathbb{R}$. The time derivative of V is defined in the standard way:

$$\partial_t V(t, \mu) := \lim_{\delta \downarrow 0} \frac{V(t+\delta, \mu) - V(t, \mu)}{\delta}, \tag{2.8}$$

provided the limit exists. Notice that the above is actually the right time derivative. The derivative in terms of μ is much more involved. We first lift the function V:

$$U(t, \xi) := V(t, \mathcal{L}_\xi), \quad \xi \in \mathbb{L}^2(\mathcal{F}_t). \tag{2.9}$$

Assume U is continuously Fréchet differentiable in ξ, then the Fréchet derivative $DU(t, \xi)$ can be identified as an element in $\mathbb{L}^2(\mathcal{F}_t)$. By Cardaliaguet (2013) (based on Lions' lecture), there exists a deterministic function $\partial_\mu V : \Theta \times \mathbb{R}^d \to \mathbb{R}^d$ such that $DU(t, \xi) = \partial_\mu V(t, \mu, \xi)$. See also Wu and Zhang (2017) for an elementary proof. This function $\partial_\mu V$ is our spatial derivative, which is called the L-derivative or Wasserstein gradient. In particular, the L-derivative is also a Gatêaux derivative:

$$\mathbb{E}\Big[\partial_\mu V(t, \mu, \xi) \cdot \xi'\Big] = \lim_{\varepsilon \to 0} \frac{V(t, \mathcal{L}_{\xi + \varepsilon \xi'}) - V(t, \mu)}{\varepsilon} \tag{2.10}$$

for all $\xi \in \mathbb{L}^2_\mu(\mathcal{F}_t)$ and $\xi' \in \mathbb{L}^2(\mathcal{F}_t)$.

Remark 2.5. We shall remark that $\partial_\mu V(t, \mu, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is unique only in the support of μ. Assume $\partial_\mu V$ exists and can be extended to $\mathbb{R}^d$ continuously; then we may define $\partial_x \partial_\mu V$ as the standard derivative of $\partial_\mu V$ in terms of the third variable. Obviously, $\partial_x \partial_\mu V(t, \mu, \cdot)$ is also well defined only in the support of μ. In this paper we shall always understand $\partial_\mu V$ in this way. In particular, we emphasize that the possible nonuniqueness of $\partial_\mu V(t, \mu, \cdot)$ outside of the support of μ does not affect the Itô formula (2.11) below, which is what we will actually need in the paper.

DEFINITION 2.6. (i) *Let $C^1_{Lip,b}(\mathcal{P}_2(\mathbb{R}^d))$ denote the space of functions $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ such that $\partial_\mu f$ exists everywhere and $\partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded and Lipschitz continuous.*

(ii) *Let $C^2_{Lip,b}(\mathcal{P}_2(\mathbb{R}^d))$ denote the subset of $C^1_{Lip,b}(\mathcal{P}_2(\mathbb{R}^d))$ such that*
- *for each $x \in \mathbb{R}$, all components of $\partial_\mu f(\cdot, x)$ belongs to $C^1_{Lip,b}(\mathcal{P}_2(\mathbb{R}^d))$;*
- *$\partial^2_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is bounded and Lipschitz continuous;*
- *$\partial_x \partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ exists and it is bounded and Lipschitz continuous.*

(iii) *Let $C^{1,2}(\Theta) := C^{1,2}_{Lip,b}(\Theta)$ denote the space of $V : \Theta \to \mathbb{R}$ such that*
- *$V(\cdot, \mu) \in C^1([0, T])$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$;*
- *$V(t, \cdot) \in C^2_{Lip,b}(\mathcal{P}_2(\mathbb{R}^d))$ for any $t \in [0, T]$.*

The following Itô formula is crucial for the results developed here; see, e.g., Buckdahn et al. (2017) and Chassagneux, Crisan, and Delarue (2014).

LEMMA 2.7. *Let $V \in C^{1,2}(\Theta)$ and $dX_t = b_t dt + \sigma_t dW_t$ for some $\mathbb{F}$-progressively measurable processes b and σ such that $\mathbb{E}[\int_0^T [|b_t|^2 + |\sigma_t|^4] dt] < \infty$. Then*

$$\frac{d}{dt} V(t, \mathcal{L}_{X_t}) = \partial_t V(t, \mathcal{L}_{X_t}) + \mathbb{E}\left[\left[b_t \cdot \partial_\mu V + \frac{1}{2}\sigma\sigma_t^\top : \partial_x \partial_\mu V\right](t, \mathcal{L}_{X_t}, X_t)\right]. \tag{2.11}$$

The main result of this section is the following verification theorem.

THEOREM 2.8. *Let Assumption* 2.1 *hold and* $V \in C^{1,2}(\Theta)$. *Then* V *is the value function defined by* (2.3) *if and only if* V *is a classical solution to the master equation:*

$$\partial_t V(t,\mu) + H(t,\mu,\partial_\mu V, \partial_x\partial_\mu V) = 0, \quad V(T,\mu) = \mathbb{E}[g(\xi)], \tag{2.12}$$

where, for $p : \Theta \times \mathbb{R}^d \to \mathbb{R}^d$ *and* $q : \Theta \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$,

$$H(t,\mu,p,q) := \sup_{a\in A} h(t,\mu,p,q,a), \tag{2.13}$$

$$h(t,\mu,p,q,a) := \mathbb{E}\Big[b(t,\xi,a)\cdot p(t,\mu,\xi) + \frac{1}{2}\sigma\sigma^\top(t,\xi,a) : q(t,\mu,\xi) + f(t,\xi,a)\Big].$$

Consequently, the above master equation has at most one classical solution in $C^{1,2}(\Theta)$.

Proof. We first assume $V \in C^{1,2}(\Theta)$ is defined by (2.3). Then clearly V satisfies the terminal condition in (2.12). Now fix $(t,\mu) \in \Theta$ and $\xi \in \mathbb{L}^2_\mu(\mathcal{F}_t)$. Recall (2.13) and apply Itô formula (2.11) on (2.4) with $t_1 = t, t_2 = t+\delta$; we have

$$\sup_{\alpha\in\mathcal{A}_t} \int_t^{t+\delta} \big[\partial_t V(s, \mathcal{L}_{X_s^{t,\xi,\alpha}}) + h(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, \partial_\mu V, \partial_x\partial_\mu V, \alpha_s)\big] ds = 0. \tag{2.14}$$

Under Assumption 2.1, $\mathcal{W}_2(\mathcal{L}_{X_s^{t,\xi,\alpha}}, \mu) \le \|X_s^{t,\xi,\alpha} - \xi\|_{\mathbb{L}^2} \le C\sqrt{\delta}$, for $s \in [t, t+\delta]$, where C may depend on $\|\xi\|_{\mathbb{L}^2}$. By the required regularity on V, there exists a modulus of continuity function ρ such that, again for $\alpha \in \mathcal{A}_t$ and $s \in [t, t+\delta]$,

$$\begin{aligned}
&|\partial_t V(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}) - \partial_t V(t,\mu)| + |\partial_\mu V(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, X_s^{t,\xi,\alpha}) - \partial_\mu V(t,\mu,\xi)| \\
&\quad + |\partial_x\partial_\mu V(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, X_s^{t,\xi,\alpha}) - \partial_x\partial_\mu V(t,\mu,\xi)| \le \rho\big(C\sqrt{\delta} + |X_s^{t,\xi,\alpha} - \xi|\big), \\
&|b(s, X_s^{t,\xi,\alpha}, \alpha_s) - b(t,\xi,\alpha_s)| + |\sigma(s, X_s^{t,\xi,\alpha},\alpha_s) - \sigma(t,\xi,\alpha_s)| \\
&\quad + |f(s, X_s^{t,\xi,\alpha}, \alpha_s) - f(t,\xi,\alpha_s)| \le \rho\big(\delta + |X_s^{t,\xi,\alpha} - \xi|\big).
\end{aligned} \tag{2.15}$$

These lead to, for a possibly different modulus of continuity function ρ',

$$\big|h(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, \partial_\mu V, \partial_x\partial_\mu V, \alpha_s) - h(t,\mu,\partial_\mu V, \partial_x\partial_\mu V, \alpha_s)\big| \le \rho'(\delta). \tag{2.16}$$

Then, by (2.14) we have, when $\delta \to 0$,

$$\partial_t V(t,\mu) + \sup_{\alpha\in\mathcal{A}_t} \frac{1}{\delta}\int_t^{t+\delta} h(t,\mu,\partial_\mu V, \partial_x\partial_\mu V, \alpha_s) ds = o(1).$$

On one hand, this clearly implies $\partial_t V(t,\mu) + H(t,\mu,\partial_\mu V, \partial_x\partial_\mu V) \ge 0$. On the other hand, by restricting the above α to constant functions we obtain $\partial_t V(t,\mu) + H(t,\mu,\partial_\mu V, \partial_x\partial_\mu V) \le 0$. That is, V satisfies (2.12).

We now assume $V \in C^{1,2}(\Theta)$ is a classical solution of (2.12) and want to verify (2.3). Fix $(t,\mu) \in \Theta$ and $\xi \in \mathbb{L}^2_\mu(\mathcal{F}_t)$. For any $\alpha \in \mathcal{A}_t$, by Itô formula (2.11) we have

$$\begin{aligned}
J(t,\xi,\alpha) &= \mathbb{E}[g(X_T^{t,\xi,\alpha})] + \int_t^T \mathbb{E}\big[f(s, X_s^{t,\xi,\alpha}, \alpha_s)\big] ds \\
&= V(T, \mathcal{L}_{X_T^{t,\xi,\alpha}}) + \int_t^T \mathbb{E}\big[f(s, X_s^{t,\xi,\alpha}, \alpha_s)\big] ds \\
&= V(t,\mu) + \int_t^T \big[\partial_t V(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}) + h(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, \partial_\mu V, \partial_x\partial_\mu V, \alpha_s)\big] ds \\
&\le V(t,\mu) + \int_t^T \big[\partial_t V(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}) + H(s,\mathcal{L}_{X_s^{t,\xi,\alpha}}, \partial_\mu V, \partial_x\partial_\mu V)\big] ds = V(t,\mu).
\end{aligned} \tag{2.17}$$

On the other hand, fix $\varepsilon > 0$ and $n \geq 1$, and denote $t_i := t + \frac{i}{n}[T-t]$, $i = 0, \cdots, n$. We construct an $\alpha^{n,\varepsilon} \in \mathcal{A}_t$ as follows. First, there exists $a_0^\varepsilon \in A$ such that

$$h(t_0, \mu, \partial_\mu V, \partial_x \partial_\mu V, a_0^\varepsilon) \geq H(t_0, \mu, \partial_\mu V, \partial_x \partial_\mu V) - \frac{\varepsilon}{T-t}.$$

Define $\alpha_s^{n,\varepsilon} := a_0^\varepsilon$ for $s \in [t_0, t_1)$. Next, there exists $a_1^\varepsilon \in A$ such that

$$h(t_1, \mathcal{L}_{X_{t_1}^{t,\xi,\alpha^{n,\varepsilon}}}, \partial_\mu V, \partial_x \partial_\mu V, a_1^\varepsilon) \geq H(t_1, \mathcal{L}_{X_{t_1}^{t,\xi,\alpha^{n,\varepsilon}}}, \partial_\mu V, \partial_x \partial_\mu V) - \frac{\varepsilon}{T-t}.$$

Define $\alpha_s^{n,\varepsilon} := a_1^\varepsilon$ for $s \in [t_1, t_2)$. Repeat the procedure and define $\alpha_s^{n,\varepsilon}$ for $s \in [t_i, t_{i+1})$ for $i = 1, \ldots, n-1$. Clearly $\alpha^{n,\varepsilon} \in \mathcal{A}_t$. Now, by the second equality of (2.17) and then by (2.15) and (2.16), as $n \to \infty$, we have

$$\begin{aligned}
&J(t, \xi, \alpha^{n,\varepsilon}) - V(t, \mu)\\
&= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \big[\partial_t V(s, \mathcal{L}_{X_s^{t,\xi,\alpha^{n,\varepsilon}}}) + h(s, \mathcal{L}_{X_s^{t,\xi,\alpha^{n,\varepsilon}}}, \partial_\mu V, \partial_x \partial_\mu V, \alpha_s^{n,\varepsilon})\big] ds\\
&= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \big[\partial_t V(t_i, \mathcal{L}_{X_{t_i}^{t,\xi,\alpha^{n,\varepsilon}}}) + h(t_i, \mathcal{L}_{X_{t_i}^{t,\xi,\alpha^{n,\varepsilon}}}, \partial_\mu V, \partial_x \partial_\mu V, \alpha_s^{n,\varepsilon})\big] ds + o(1)\\
&\geq \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \big[\partial_t V(t_i, \mathcal{L}_{X_{t_i}^{t,\xi,\alpha^{n,\varepsilon}}}) + H(t_i, \mathcal{L}_{X_{t_i}^{t,\xi,\alpha^{n,\varepsilon}}}, \partial_\mu V, \partial_x \partial_\mu V) - \varepsilon\big] ds + o(1)\\
&= o(1) - \varepsilon.
\end{aligned}$$

Here the $o(1)$ may depend on $\|\xi\|_{\mathbb{L}^2}$ and we have used the fact that

$$\sup_{t \leq s \leq T} \|X_s^{t,\xi,\alpha^{n,\varepsilon}}\|_{\mathbb{L}^2} \leq C[1 + \|\xi\|_{\mathbb{L}^2}].$$

By first sending $n \to \infty$ and then $\varepsilon \to 0$, we obtain $\sup_{\alpha \in \mathcal{A}_t} J(t, \xi, \alpha) \geq V(t, \mu)$, hence V is indeed the value function defined by (2.3). □

Remark 2.9. While we shall provide some positive results in the next section, in general it is difficult to expect classical solutions for nonlinear master equations. There have been some studies on viscosity solutions to such master equations. For example, Pham and Wei (2018) proposed a notion of viscosity solution by first lifting the function V to U in the sense of (2.9) and then studying the viscosity property of U in the Hilbert space $\mathbb{L}^2(\mathcal{F}_T)$. More recently, Wu and Zhang (2018) proposed an intrinsic notion of viscosity solutions for path-dependent master equations in the Wasserstein space directly, which, in particular, is consistent with the classical solution in Theorem 2.8 when V is smooth.

2.3. The optimal control. We now turn to the optimal control.

THEOREM 2.10. *Let Assumption* 2.1 *hold and* $V \in C^{1,2}(\Theta)$ *be the classical solution to the master equation* (2.12)–(2.13). *Assume further that*

(i) *the Hamiltonian* $H(t, \mu, \partial_\mu V, \partial_x \partial_\mu V)$ *defined by* (2.13) *has an optimal control* $a^* = I(t, \mu) \in A$, *for any* $(t, \mu) \in \Theta$, *where* $I : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \to A$ *is measurable;*

(ii) *for any fixed* $(t, \mu) \in \Theta$ *and* $\xi \in \mathbb{L}^2_\mu(\mathcal{F}_t)$, *the McKean–Vlasov SDE*

$$X_s^* = \xi + \int_t^s b\big(r, X_r^*, I(r, \mathcal{L}_{X_r^*})\big) dr + \int_t^s \sigma\big(r, X_r^*, I(r, \mathcal{L}_{X_r^*})\big) dW_r \tag{2.18}$$

has a (strong) solution X^*.

Then $\alpha^*_s := I(s, \mathcal{L}_{X^*_s})$, $s \in [t, T]$, *is an optimal control for the optimization problem* (2.3) *with this fixed* (t, μ).

Proof. Note that $X^* = X^{t,\xi,\alpha^*}$. Set $\alpha = \alpha^*$ in (2.17). By optimality condition (i) we see that equality holds for (2.17), namely, $J(t, \xi, \alpha^*) = V(t, \mu)$, implying that α^* is optimal. ∎

As in standard control theory, in general, the existence of the classical solution V is not sufficient for the existence of the optimal control. In particular, the McKean–Vlasov SDE (2.18) may not have a solution, even if I exists. At below we provide a sufficient condition.

THEOREM 2.11. *Let all the conditions in Theorem* 2.10 *hold true, except possibly the* (ii) *there. Assume further* b, σ *are bounded and continuous in* a, *and* $I : \Theta \to A$ *is continuous. Then the McKean–Vlasov SDE* (2.18) *has a strong solution for any* (t, μ), *and hence the optimization problem* (2.3) *has an optimal control.*

Proof. Without loss of generality, we prove the result only at $(0, \mu)$. Fix $\xi \in \mathbb{L}^2_0(\mu)$. For any $\alpha \in \mathcal{A}_0$, denote

$$X^\alpha_t = \xi + \int_0^t b(s, X^\alpha_s, \alpha_s)ds + \int_0^t \sigma(s, X^\alpha_s, \alpha_s)dW_s.$$

Under Assumption 2.1, it is clear that

$$\mathbb{E}[|X^\alpha_t - X^\alpha_s|^2] \le C_\mu |t - s|, \quad \text{and thus} \quad \mathcal{W}_2(\mathcal{L}_{X^\alpha_t}, \mathcal{L}_{X^\alpha_s}) \le C_\mu \sqrt{|t - s|}, \tag{2.19}$$

where the constant C_μ may depend on μ but does not depend on α. Moreover, assume $|b|, |\sigma| \le L$. Let $\mathcal{D}_L(\mu)$ denote the set of $\mathcal{L}_{\tilde{X}_t}$, where $t \in [0, T]$, $\tilde{X}_t = \tilde{X}_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s$ in some arbitrary probability space with $\mathcal{L}_{\tilde{X}_0} = \mu$ and $|\tilde{b}|, |\tilde{\sigma}| \le L$. As in Wu and Zhang (2018, Lemma 3.1), one can easily show that $\mathcal{D}_L(\mu)$ is compact under $\mathcal{W}_2$. Since I is continuous in Θ, it is uniformly continuous on $[0, T] \times \mathcal{D}_L(\mu)$ with certain modulus of continuity function ρ_μ, which may depend on μ. Clearly $\mathcal{L}_{X^\alpha_t} \in \mathcal{D}_L(\mu)$ for all $\alpha \in \mathcal{A}_0$ and $t \in [0, T]$. Then we have

$$\Big| I(t, \mathcal{L}_{X^\alpha_t}) - I(s, \mathcal{L}_{X^\alpha_s}) \Big| \le \rho_\mu(|t - s|) \quad \text{for all } \alpha \in \mathcal{A}_0. \tag{2.20}$$

Denote

$$\mathcal{A}_0(\rho_\mu) := \Big\{ \alpha \in \mathcal{A}_0 : |\alpha_t - \alpha_s| \le \rho_\mu(t - s), \quad 0 \le s < t \le T \Big\}. \tag{2.21}$$

We now define a mapping $\Phi : \mathcal{A}_0(\rho_\mu) \to \mathcal{A}_0(\rho_\mu)$ by $\Phi_t(\alpha) := I(t, \mathcal{L}_{X^\alpha_t})$, where (2.20) ensures that $\Phi(\alpha) \in \mathcal{A}_0(\rho_\mu)$ for all $\alpha \in \mathcal{A}_0(\rho_\mu)$. One can easily show that $\mathcal{A}_0(\rho_\mu)$ is convex and compact under the uniform norm, and Φ is continuous. Then, applying the Schauder's fixed point theorem, Φ has a fixed point $\alpha^* \in \mathcal{A}_0(\rho_\mu)$: $\Phi(\alpha^*) = \alpha^*$. Now it is clear that $X^* := X^{\alpha^*}$ satisfies (2.18), and hence α^* is an optimal control. ∎

Remark 2.12. In this section, we used the dynamic programming principle. Since the control here is deterministic and thus falls in strong formulation, one may also use the stochastic maximum principle, provided the optimal control exists. We will present heuristic arguments in Appendix A to show how the McKean–Vlasov SDEs come into play naturally.

3. Classical solution of a nonlinear master equation. The existence of classical solutions for nonlinear master equations is a very challenging problem. We shall leave the general case to future research. In this section we study a special type of master equations. Consider the equation (2.12)–(2.13) with

$$\sigma = I_d, \quad b = b(t,a), \quad f = f(t,a).$$

Then (2.12) becomes

$$\partial_t V(t,\mu) + \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,(\partial_x \partial_\mu V(t,\mu,\xi))\Big] + \sup_a \Big[b(t,a)\cdot \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big] + f(t,a)\Big] = 0.$$

This is a special case of the following nonlinear master equation:

$$\begin{aligned} \partial_t V(t,\mu) + \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,(\partial_x \partial_\mu V(t,\mu,\xi))\Big] + F\big(t, \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]\big) = 0,\\ V(T,\mu) = \mathbb{E}\big[g(\xi)\big]. \end{aligned} \tag{3.1}$$

Theorem 3.1. *Let F and g be smooth enough with bounded derivatives. Assume one of the following two conditions holds true:*

(i) *T is sufficiently small;*

(ii) *$d=1$, and either $\partial_{xx} g > 0 > \partial_{xx} F$ or $\partial_{xx} g < 0 < \partial_{xx} F$.*

Then the master equation (3.1) *has a classical solution $V \in C^{1,2}(\Theta)$.*

Proof. We shall proceed in two steps.

Step 1. Consider the following master equation which is linear in $\partial_\mu \widetilde{V}$:

$$\begin{aligned} \partial_t \widetilde{V}(t,\mu) + \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,(\partial_x \partial_\mu \widetilde{V}(t,\mu,\xi))\Big] + \partial_x F\big(t, \widetilde{V}(t,\mu)\big)\mathbb{E}\big[\partial_\mu \widetilde{V}(t,\mu,\xi)\big] = 0,\\ \widetilde{V}(T,\mu) = \mathbb{E}\big[\partial_x g(\xi)\big]. \end{aligned} \tag{3.2}$$

We shall prove in Step 2 below that under (i) or (ii) the above master equation has a unique classical solution $\widetilde{V}$. We next consider the linear master equation:

$$\begin{aligned} \partial_t V(t,\mu) + \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,(\partial_x \partial_\mu V(t,\mu,\xi))\Big] + F\big(t, \widetilde{V}(t,\mu)\big) = 0,\\ V(T,\mu) = \mathbb{E}\big[g(\xi)\big], \end{aligned} \tag{3.3}$$

Then clearly V is also smooth. It remains to verify that the above V satisfies (3.1). Indeed, by (3.3) we see that

$$V(t,\mu) = \mathbb{E}\big[g(X_T^{t,\xi})\big] + \int_t^T F(s, \widetilde{V}(s, \mathcal{L}_{X_s^{t,\xi}}))ds, \tag{3.4}$$

where $X_s^{t,\xi} := \xi + W_s - W_t$. Differentiating with respect to μ, we obtain

$$\begin{aligned} \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big] &= \mathbb{E}\Big[\partial_x g(X_T^{t,\xi})\Big]\\ &\quad + \mathbb{E}\Big[\int_t^T \partial_x F(s, \widetilde{V}(s, \mathcal{L}_{X_s^{t,\xi}}))\cdot \partial_\mu \widetilde{V}\big(s, \mathcal{L}_{X_s^{t,\xi}}, X_s^{t,\xi}\big)ds\Big]. \end{aligned} \tag{3.5}$$

That is, $\overline{V}(t,\mu) := \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]$ satisfies the following linear master equation:

$$\begin{aligned} \partial_t \overline{V}(t,\mu) + \frac{1}{2}\mathbb{E}\Big[\mathrm{tr}\,(\partial_x \partial_\mu \overline{V}(t,\mu,\xi))\Big] + \partial_x F\big(t, \widetilde{V}(t,\mu)) \cdot \mathbb{E}\big[\partial_\mu \widetilde{V}(t,\mu,\xi)\big]\big) = 0,\\ \overline{V}(T,\mu) = \mathbb{E}\big[\partial_x g(\xi)\big]. \end{aligned}$$

However, by (3.2), $\widetilde{V}$ also satisfies the above master equation. Then by the uniqueness of classical solutions, we have $\widetilde{V}(t,\mu) = \overline{V}(t,\mu) = \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]$. Plugging this into (3.3), we see that V satisfies (3.1).

Step 2. We now prove the well-posedness of (3.2) under (i) or (ii). When T is small, the arguments are rather standard; see, e.g., Chassagneux, Crisan, and Delarue (2014). We now assume (ii) holds true. Without loss of generality, we assume F is convex in x and g is concave. For any $y \in \mathbb{R}$, define

$$\begin{aligned}\Phi(y;t,\mu) &:= \mathbb{E}\Big[\partial_x g\big(\xi + W_T - W_t + \int_t^T \partial_x F(s,y)ds\big)\Big],\\ \Psi(y,t,\mu) &:= \Phi(y;t,\mu) - y,\end{aligned} \tag{3.6}$$

where $\mathcal{L}_\xi = \mu$ and $W_T - W_t$ is independent of ξ. It is straightforward to show that Φ is smooth in (y,t,μ) and, for any y, $\Phi(y;\cdot)$ solves the following linear master equation:

$$\partial_t \Phi(y;t,\mu) + \frac{1}{2}\mathbb{E}\big[\partial_x\partial_\mu\Phi(y;t,\mu,\xi))\big] + \partial_x F(t,y)\mathbb{E}\big[\partial_\mu\Phi(y;t,\mu,\xi)\big] = 0. \tag{3.7}$$

Under our conditions, $\partial_x g$ is decreasing and $\partial_x F$ is increasing in y; then by (3.6) Φ is decreasing in y and thus $\partial_y \Psi \le -1$, so $y \mapsto \Psi(y,t,\mu)$ has an inverse function Ψ^{-1}, which is also smooth. Since $\partial_x g$ is bounded by some constant C_0, then $|\Phi(y;t,\mu)| \le C_0$, and thus $\Psi(C_0,t,\mu) \le 0 \le \Psi(-C_0,t,\mu)$ for any fixed (t,μ). In particular, 0 is in the range of $\Psi(\cdot;t,\mu)$ for any fixed (t,μ). Define $U(t,\mu) := \Psi^{-1}(0,t,\mu)$; then U is smooth. Note that $U(t,\mu) = \Phi(U(t,\mu);t,\mu)$. Applying the chain rule (which is obvious from the definitions), we have

$$\partial_t U = \partial_t\Phi + \partial_y\Phi\partial_t U, \quad \partial_\mu U = \partial_\mu\Phi + \partial_y\Phi\partial_\mu U, \quad \partial_x\partial_\mu U = \partial_x\partial_\mu\Phi + \partial_y\Phi\partial_x\partial_\mu U.$$

Namely, denoting $c := 1 - \partial_y\Phi(U(t,\mu);t,\mu) \ge 1$,

$$\begin{aligned}\partial_t\Phi(U(t,\mu);t,\mu) = c\ \partial_t U(t,\mu), \quad \partial_\mu\Phi(U(t,\mu);t,\mu,\cdot) = c\ \partial_\mu U(t,\mu,\cdot),\\ \partial_x\partial_\mu\Phi(U(t,\mu);t,\mu,\cdot) = c\ \partial_x\partial_\mu U(t,\mu,\cdot).\end{aligned}$$

Plugging these into (3.7) with $y = U(t,\mu)$, we obtain that U satisfies (3.2). $\square$

3.1. An example. We now consider a special case. For some $R > 0$, which will be specified later, set

$$d = 1, \quad A = [-R,R], \quad b(t,x,a) = a, \quad \sigma = 1, \quad f(t,x,a) = -\frac{1}{2}a^2. \tag{3.8}$$

Then

$$\begin{aligned}h(t,\mu,p,q,a) &= \frac{1}{2}\mathbb{E}[q(t,\mu,\xi)] + a\mathbb{E}[p(t,\mu,\xi)] - \frac{1}{2}a^2,\\ H(t,\mu,p,q) &= \frac{1}{2}\mathbb{E}[q(t,\mu,\xi)] + F(\mathbb{E}[p(t,\mu,\xi)]),\\ \text{where } F(x) &= \frac{1}{2}|x|^2\mathbf{1}_{\{|x|\le R\}} + [R|x| - \frac{1}{2}R^2]\mathbf{1}_{\{|x|>R\}},\end{aligned} \tag{3.9}$$

and thus (2.12) becomes

$$\partial_t V + \frac{1}{2}\mathbb{E}\big[\partial_x\partial_\mu V(t,\mu,\xi)\big] + F\big(\mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big)\big) = 0, \quad V(T,\mu) = \mathbb{E}[g(\xi)]. \tag{3.10}$$

Notice that F is convex; however, it is in $C^1(\mathbb{R})$ but not in $C^2(\mathbb{R})$.

THEOREM 3.2. *Assume g is smooth enough with bounded derivatives, and in particular $|\partial_x g| \le C_0 < R$. Then, either for T small enough or $d = 1$ and g is concave,*

(i) *the master equation* (3.10) *has a unique classical solution V such that*

$$\Big|\mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]\Big| \le C_0, \quad (t,\mu) \in \Theta, \xi \in \mathbb{L}^2_\mu(\mathcal{F}_t); \tag{3.11}$$

(ii) *for any $(t,\mu) \in \Theta$ and $\xi \in \mathbb{L}^2_\mu(\mathcal{F}_t)$, the McKean–Vlasov SDE* (2.18) *with $I(t,\mu) := \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]$ has a solution X^*;*

(iii) *for any $(t,\mu) \in \Theta$, the optimization problem* (2.3) *has an optimal control: $\alpha^*_s := I(s, \mathcal{L}_{X^*_s})$.*

Proof. Let $\widetilde{F} : \mathbb{R} \to \mathbb{R}$ be a smooth function such that

$$\widetilde{F} \text{ is convex and } \widetilde{F}(x) = F(x) \text{ for } |x| \le C_0 \text{ or } |x| \ge R.$$

Applying Theorem 3.1, the master equation (3.10) corresponding to $\widetilde{F}$ has a classical solution V. Introduce the conjugate of $\widetilde{F}$: $\widetilde{f}(a) := \sup_{x\in\mathbb{R}}[ax - \widetilde{F}(x)]$, $a \in A$. By the convexity of $\widetilde{F}$, we have $\widetilde{F}(x) = \sup_{a\in A}[ax - \widetilde{f}(a)]$. Then by Theorem 2.8 we see that

$$V(t,\mu) = \sup_{\alpha\in\mathcal{A}_t} \widetilde{J}(t,\xi,\alpha), \quad \text{where} \tag{3.12}$$

$$\widetilde{X}^{t,\xi,\alpha}_s := \xi + \int_t^s \alpha_r dr + W_s - W_t, \quad \widetilde{J}(t,\xi,\alpha) := \mathbb{E}\big[g(\widetilde{X}^{t,\xi,\alpha}_T)\big] - \int_t^T \widetilde{f}(\alpha_s) ds.$$

For any $t \in [0,T]$, $\xi, \xi' \in \mathbb{L}^2(\mathcal{F}_t)$, and $\alpha \in \mathcal{A}_t$, under our conditions it is clear that

$$|\widetilde{J}(t,\xi,\alpha) - \widetilde{J}(t,\xi',\alpha)| \le C_0 \mathbb{E}[|\xi - \xi'|].$$

Since ξ, ξ' are arbitrary, then it follows from (3.12) that

$$|V(t,\mu) - V(t,\mu')| \le C_0 \mathcal{W}_2(\mu,\mu'),$$

which implies (3.11) immediately. Since $\widetilde{F}(x) = F(x) = \frac{1}{2}x^2$ for $|x| \le C_0$, then (3.11) implies further that V is a classical solution to master equation (3.10) corresponding to F.

(ii) Clearly in this case the optimal argument of the Hamiltonian F leads to $I(t,\mu) = \mathbb{E}\big[\partial_\mu V(t,\mu,\xi)\big]$, which is continuous. Then (ii) follows from Theorem 2.11.

Finally, (iii) follows directly from Theorem 2.10. ∎

We remark that in this example it is more natural to set $A = \mathbb{R}$ and all the results still hold true. The constraint $A = [-R, R]$ is to ensure the uniform requirement in Assumption 2.1(i), which is more convenient for establishing the general theory but can be relaxed.

4. The general case.

In this section we investigate the general case $T > \textsc{h}$.

4.1. Strong formulation with closed loop controls.

In this subsection we illustrate how the information delay naturally leads to the path dependence of the value function, even if the coefficients b, σ, f, g in (2.1) depend only on the current state of X. It is easier to show the idea in strong formulation, namely, we fix a probability space and the state process X^α is controlled, but we emphasize that we shall use closed loop controls, both for practical and for theoretical reasons.

As in section 2, let $(\Omega, \mathbb{F}, \mathbb{P})$ be a filtered probability space on $[0,T]$ and W an $\mathbb{F}$-Brownian motion under $\mathbb{P}$. For simplicity, in this subsection we assume $T \le 2\textsc{h}$, which will not be required in later subsections. Let $t \in (\textsc{h}, T]$, and ξ be an $\mathbb{F}$-progressively measurable process on $[0,t]$. Consider the following counterpart of (2.1):

$$
\begin{aligned}
X_s^{t,\xi,\alpha} &= \xi_t + \int_t^s b\big(r, X_r^{t,\xi,\alpha}, \alpha_r(\xi_{[0,r-\textsc{h}]})\big)dr + \int_t^s \sigma\big(r, X_r^{t,\xi,\alpha}, \alpha_r(\xi_{[0,r-\textsc{h}]})\big)dW_r;\\
J(t,\xi,\alpha) &:= \mathbb{E}\left[g(X_T^{t,\xi,\alpha}) + \int_t^T f\big(s, X_s^{t,\xi,\alpha}, \alpha_s(\xi_{[0,s-\textsc{h}]})\big)ds\right].
\end{aligned}
\tag{4.1}
$$

Similar to Lemma 2.2, $J(t,\xi,\alpha)$ depends on ξ only through the law of the stopped process $\xi_{[0,t]}$. That is, if ξ' is another process such that $\mathcal{L}_{\xi_{[0,t]}} = \mathcal{L}_{\xi'_{[0,t]}}$, then $J(t,\xi,\alpha) = J(t,\xi',\alpha)$. Consequently, the following value function is also law invariant:

$$
\widetilde{V}(t,\xi) := \sup_{\alpha\in\mathcal{A}_t} J(t,\xi,\alpha). \tag{4.2}
$$

We emphasize that the above law invariant property relies on the law of the stopped process $\xi_{[0,t]}$, rather than the law of the current state ξ_t.

Example 4.1. Let $d=1$, $A=[-1,1]$, $b(t,x,a)=a$, $\sigma(t,x,a)=1$, $f(t,x,a)=0$, $g(x)=x^2$, $T=2\textsc{h}$, $t=\frac{3}{2}\textsc{h}$. Set

$$
\xi_s = W_s, 0\le s\le t, \quad \xi'_s := W_{3(s-\textsc{h})}\mathbf{1}_{[\textsc{h},t]}(s). \tag{4.3}
$$

Then $\xi_t = \xi'_t = W_t$ but in general $\widetilde{V}(t,\xi) \neq \widetilde{V}(t,\xi')$.

Proof. First, since $\xi'_s = 0$, $s \le \textsc{h}$, then $\alpha_r(\xi'_{[0,r-\textsc{h}]}) = \alpha_r(0)$ is deterministic. Thus

$$
J(t,\xi',\alpha) = \mathbb{E}\left[\left|W_t + \int_t^T \alpha_r(0)dr + W_T - W_t\right|^2\right] = \left|\int_t^T \alpha_r(0)dr\right|^2 + T.
$$

This implies

$$
\widetilde{V}(t,\xi') = T + (T-t)^2 = 2\textsc{h} + \frac{1}{4}\textsc{h}^2.
$$

On the other hand, denote $\beta_r := \alpha_{r+\textsc{h}}(W_{[0,r]})$ which is $\mathcal{F}_r$-measurable; then

$$
\begin{aligned}
J(t,\xi,\alpha) &= \mathbb{E}\left[\left|W_t + \int_{t-\textsc{h}}^{T-\textsc{h}} \beta_r dr + W_T - W_t\right|^2\right]\\
&= \mathbb{E}\left[\left|\int_{t-\textsc{h}}^{T-\textsc{h}} \beta_r dr\right|^2 + 2\int_{t-\textsc{h}}^{T-\textsc{h}} W_r\beta_r dr\right] + T \ge 2\mathbb{E}\left[\int_{\frac{\textsc{h}}{2}}^{\textsc{h}} W_r\beta_r dr\right] + 2\textsc{h}.
\end{aligned}
$$

By choosing $\beta_r = \operatorname{sign}(W_r)$, we have

$$
\widetilde{V}(t,\xi) \ge 2\mathbb{E}\Big[\int_{\frac{\textsc{h}}{2}}^{\textsc{h}} |W_r|dr\Big] + 2\textsc{h} = 2\textsc{h} + c\textsc{h}^{\frac{3}{2}},
$$

where $c>0$ is a generic constant independent of $\textsc{h}$. Then clearly $\widetilde{V}(t,\xi) > \widetilde{V}(t,\xi')$, when $\textsc{h}$ is small enough. ∎

We also remark that it is crucial to use closed loop controls. If we use open loop controls with delay, namely, $\alpha_s = \alpha_s(W_{[0,s-\mathrm{H}]})$, then for each α, obviously $J(t,\xi,\alpha)$ would depend on the joint law of (ξ, W) on $[0,t]$. The following example shows that the corresponding value function $\widetilde{V}(t,\xi)$ may also violate the law invariant property.

Example 4.2. Consider the same setting in Example 4.1, but replace (4.3) with

$$\xi_s = [W_s - W_{\mathrm{H}}]\mathbf{1}_{[\mathrm{H},t]}(s), \quad \xi'_s = W_{s-\mathrm{H}}\mathbf{1}_{[\mathrm{H},t]}(s), \quad 0 \le s \le t.$$

Then $\mathcal{L}_{\xi_{[0,t]}} = \mathcal{L}_{\xi'_{[0,t]}}$. However, if we use open loop controls but still denote the value function as $\widetilde{V}$, then $\widetilde{V}(t,\xi) \neq \widetilde{V}(t,\xi')$.

Proof. First, note that $\alpha_s = \alpha_s(W_{[0,s-\mathrm{H}]})$ is $\mathcal{F}_{\mathrm{H}}$-measurable. Then

$$J(t,\xi,\alpha) = \mathbb{E}\left[\left|W_t - W_{\mathrm{H}} + \int_t^T \alpha_s ds + W_T - W_t\right|^2\right] = \mathbb{E}\left[\left|\int_t^T \alpha_s ds\right|^2\right] + T - \mathrm{H}.$$

This implies

$$\widetilde{V}(t,\xi) = T - \mathrm{H} + (T-t)^2 = \mathrm{H} + \frac{1}{4}\mathrm{H}^2.$$

On the other hand, denote $\beta_r := \alpha_{r+\mathrm{H}}(W_{[0,r]})$ which is $\mathcal{F}_r$-measurable; then

$$J(t,\xi',\alpha) = \mathbb{E}\left[\left|W_{\frac{\mathrm{H}}{2}} + \int_{t-\mathrm{H}}^{T-\mathrm{H}} \beta_r dr + W_T - W_t\right|^2\right] = \mathbb{E}\left[\left|W_{\frac{\mathrm{H}}{2}} + \int_{\frac{\mathrm{H}}{2}}^{\mathrm{H}} \beta_r dr\right|^2\right] + \frac{\mathrm{H}}{2}.$$

By choosing $\beta_r = \mathrm{sign}(W_{\frac{\mathrm{H}}{2}})$, we have

$$\widetilde{V}(t,\xi) \ge \mathbb{E}\left[\left[\left|W_{\frac{\mathrm{H}}{2}}\right| + \frac{\mathrm{H}}{2}\right]^2\right] + \frac{\mathrm{H}}{2} = \mathrm{H} + \frac{1}{4}\mathrm{H}^2 + \mathrm{H}\mathbb{E}[|W_{\frac{\mathrm{H}}{2}}|] > \widetilde{V}(t,\xi).$$

This completes the proof. ∎

4.2. Weak formulation in path-dependent setting. Both for closed loop controls and for path-dependent problems, it is a lot more convenient to use the weak formulation on canonical space. We shall follow the setting of Wu and Zhang (2018).

Let $\Omega := C([0,T];\mathbb{R}^d)$ be the canonical space equipped with the metric $\|\omega\|_T := \sup_{0\le t\le T}|\omega_t|$, X the canonical process, and $\mathbb{F} = \mathbb{F}^X = \{\mathcal{F}_t\}_{t\in[0,T]}$ the natural filtration generated by X. Denote by $\mathcal{P}_2(\mathcal{F}_T)$ the set of probability measures $\mathbb{P}$ on $\mathcal{F}_T$ such that $\mathbb{E}^{\mathbb{P}}[\|X\|_T^2] < \infty$ and $\Theta_T := [0,T]\times\mathcal{P}_2(\mathcal{F}_T)$. Quite often we will also use μ to denote elements of $\mathcal{P}_2(\mathcal{F}_T)$. We equip $\mathcal{P}_2(\mathcal{F}_T)$ with the 2-Wasserstein distance $\mathcal{W}_2$ which extends (2.7):

$$\mathcal{W}_2(\mu,\mu') := \inf\left\{\left(\mathbb{E}[\sup_{0\le t\le T}|\eta_t - \eta'_t|^2]\right)^{\frac{1}{2}} : \mathcal{L}_\eta = \mu, \mathcal{L}_{\eta'} = \mu'\right\} \tag{4.4}$$

for $\mu,\mu' \in \mathcal{P}_2(\mathcal{F}_T)$, where $\mathcal{L}_\eta$ is the law of the process η.

Given $\mu \in \mathcal{P}_2(\mathcal{F}_T)$, let $\mu_{[0,t]}$ denote the μ-distribution of the stopped process $X_{[0,t]}$. For a function $V : \Theta_T \to \mathbb{R}$, we say V is $\mathbb{F}$-adapted if $V(t,\mu) = V(t,\mu_{[0,t]})$ for any $(t,\mu)\in\Theta_T$. For such V, we define its time derivative as

$$\partial_t V(t,\mu) := \lim_{\delta\downarrow 0}\frac{V(t+\delta,\mu_{[0,t]}) - V(t,\mu_{[0,t]})}{\delta}, \tag{4.5}$$

where we are freezing the law of X from t to $t+\delta$. The spatial derivative takes the form $\partial_\mu V : \Theta_T \times \Omega \to \mathbb{R}^d$ and is $\mathbb{F}$-progressively measurable, namely, measurable in all variables and $\mathbb{F}$-adapted. We emphasize that, as in Dupire (2009), $\partial_\mu V$ is not a Fréchet derivative with respect to the law of the whole stopped process $X_{[0,t]}$ but is a derivative with respect to $\mathcal{L}_{X_t}$ only. Roughly speaking, by extending the whole setting to the space of càdlàg paths, let ξ be a process on $[0,t]$ such that $\mathcal{L}_\xi = \mu_{[0,t]}$, and let ξ'_t be an $\mathcal{F}_t$-measurable random variable. Then

$$\mathbb{E}^\mu\Big[\partial_\mu V(t,\mu,\xi)\cdot\xi'_t\Big] := \lim_{\varepsilon\downarrow 0}\frac{V(t,\mathcal{L}_{\xi+\varepsilon\xi'_t\mathbf{1}_{\{t\}}}) - V(t,\mathcal{L}_\xi)}{\varepsilon}. \tag{4.6}$$

Moreover, for the process $\partial_\mu V(t,\mu,X.)$, we may introduce the path derivative $\partial_\omega\partial_\mu V$ in the spirit of Dupire (2009). When V is smooth enough in these senses, the functional Itô formula (4.7) below holds true. We refer to Wu and Zhang (2018) for details. In this section, to avoid the technical details, we take the approach in Ekren, Touzi, and Zhang (2016) and use the functional Itô formula directly to define the smoothness of V.

DEFINITION 4.3. *Let $C^{1,2}(\Theta_T)$ denote the space of functions $V : \Theta_T \to \mathbb{R}$ such that there exist functions $\partial_\mu V : \Theta_T \times \Omega \to \mathbb{R}^d$ and $\partial_\omega\partial_\mu V : \Theta_T \times \Omega \to \mathbb{R}^{d\times d}$ satisfying that*

(i) *the $\partial_t V$ defined by* (4.5) *exists, and V, $\partial_t V$, $\partial_\mu V$, $\partial_\omega\partial_\mu V$ are all $\mathbb{F}$-adapted and uniformly continuous;*

(ii) *for any semimartingale measure $\mathbb{P}$, namely, X is a semimartingale under $\mathbb{P}$, the following functional Itô formula holds:*

$$dV(t,\mathbb{P}) = \partial_t V(t,\mathbb{P})dt + \mathbb{E}^{\mathbb{P}}\Big[\partial_\mu V(t,\mathbb{P},X.)\cdot dX_t + \frac{1}{2}\partial_\omega\partial_\mu V(t,\mathbb{P},X.) : d\langle X\rangle_t\Big]. \tag{4.7}$$

By Lemma 2.7 and Wu and Zhang (2018), the spatial derivatives there coincide with the above $\partial_\mu V, \partial_\omega\partial_\mu V$ (with $\partial_\omega\partial_\mu V = \partial_x\partial_\mu V$ in the Markovian case). We remark that, for the purpose of viscosity solutions, in Wu and Zhang (2018), (4.7) is required only for semimartingale measures whose drift and diffusion characteristics are bounded. In that case, the regularity requirements on V are weaker than the corresponding conditions in Lemma 2.7. It is not difficult to extend the functional Itô formula in Wu and Zhang (2018) to allow for more general semimartingale measures. Nevertheless, it is more convenient to define the derivatives through the functional Itô formula directly as we do here.

LEMMA 4.4. *For any $V \in C^{1,2}(\Theta_T)$, the derivatives $\partial_\mu V$ and $\partial_\omega\partial_\mu V$ are unique in the sense that $\partial_\mu V(t,\mu,X.)$ and $\frac{1}{2}[\partial_\omega\partial_\mu V + (\partial_\omega\partial_\mu V)^\top](t,\mu,X.)$ are μ-a.s. unique for any $(t,\mu)\in\Theta_T$.*

We remark that since $\langle X\rangle$ is symmetric, so the uniqueness of $\frac{1}{2}[\partial_\omega\partial_\mu V+(\partial_\omega\partial_\mu V)^\top]$ $(t,\mu,X.)$ implies that uniqueness of $\partial_\omega\partial_\mu V(t,\mathbb{P},X.) : d\langle X\rangle_t$ in (4.7).

Proof. First let $\mu\in\mathcal{P}_2(\mathcal{F}_T)$ be a semimartingale measure. For any $t\in[0,T]$ and any $\mathcal{F}_t$-measurable and bounded random variables b_t and $\sigma_t > 0$, let $\mathbb{P}\in\mathcal{P}_2(\mathcal{F}_T)$ be such that

$$\mathbb{P}_{[0,t]} = \mu_{[0,t]} \quad \text{and} \quad X_s - X_t = b_t[s-t] + \sigma_t[W_s - W_t],\ t\le s\le T,\ \mathbb{P}\text{-a.s.},$$

for some $\mathbb{P}$-Brownian motion W. Then, by (4.7), we see that

$$\mathbb{E}^{\mathbb{P}}\Big[b_t\cdot\int_t^s \partial_\mu V(r,\mathbb{P},X.)dr + \frac{1}{2}\sigma_t\sigma_t^\top : \int_t^s \partial_\omega\partial_\mu V(r,\mathbb{P},X.)dr\Big]$$

is unique. By the uniform continuity of $\partial_\mu V$ and $\partial_\omega \partial_\mu V$, this implies that

$$\mathbb{E}^\mu\Big[b_t \cdot \partial_\mu V(t,\mu,X.) + \frac{1}{2}\sigma_t\sigma_t^\top : \partial_\omega\partial_\mu V(t,\mu,X.)\Big]$$

is unique. Here we rewrite $\mathbb{P}$ as μ since $\mathbb{P}_{[0,t]} = \mu_{[0,t]}$ and the integrand above is $\mathcal{F}_t$-measurable. Since b_t and σ_t are arbitrary, we obtain the desired uniqueness.

Now assume $\mu \in \mathcal{P}_2(\mathcal{F}_T)$ is arbitrary. For any $\varepsilon > 0$, denote $X^\varepsilon_t := \frac{1}{\varepsilon}\int_{(t-\varepsilon)^+}^t X_s ds$ and $\mu^\varepsilon := \mu \circ (X^\varepsilon)^{-1}$. Then

$$\mathcal{W}_2^2(\mu,\mu^\varepsilon) \le \mathbb{E}^\mu\Big[\|X - X^\varepsilon\|_T^2\Big] \to 0 \quad \text{as } \varepsilon \to 0,$$

which implies that $\mu^\varepsilon \to \mu$ weakly. Clearly X^ε is a μ-semimartinagle, and then μ^ε is a semimartingale measure. Thus $\partial_\mu V(t,\mu^\varepsilon_t, X.)$ is μ^ε-a.s. unique. Let η_t be $\mathcal{F}_t$-measurable, bounded, and continuous in ω (under $\|\cdot\|_T$). Note that, denoting by ρ the modulus of continuity function of $\partial_\mu V$,

$$\begin{aligned}
&\Big|\mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu^\varepsilon,X.)\eta_t\Big] - \mathbb{E}^{\mu}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big]\Big| \\
&\le \Big|\mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu^\varepsilon,X.)\eta_t\Big] - \mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big]\Big| \\
&+ \Big|\mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big] - \mathbb{E}^{\mu}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big]\Big| \\
&\le C\rho(\mathcal{W}_2(\mu,\mu^\varepsilon)) + \Big|\mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big] - \mathbb{E}^{\mu}\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big]\Big|.
\end{aligned}$$

Sending $\varepsilon \to 0$, by the weak convergence of $\mu^\varepsilon \to \mu$, we see that

$$\mathbb{E}^\mu\Big[\partial_\mu V(t,\mu,X.)\eta_t\Big] = \lim_{\varepsilon\to 0}\mathbb{E}^{\mu_\varepsilon}\Big[\partial_\mu V(t,\mu^\varepsilon,X.)\eta_t\Big]$$

is unique. Since η_t is arbitrary, we obtain the desired uniqueness of $\partial_\mu V(t,\mu,X.)$. Similarly we have the uniqueness of $\partial_\omega\partial_\mu V$. ☐

4.3. The control problem in weak formulation. Since the value function depends on the path of the state process X anyway, we shall work on the path-dependent setting directly, i.e., we will allow b, σ, f, and g to depend on the paths of X, namely, b,σ,f are functions on $[0,T]\times\Omega\times A$ and g is a function on Ω, so as to have a more general result. Let $\mathcal{A}^{\mathrm{H}}_t$ denote the set of $\mathbb{F}$-progressively measurable A-valued processes α on $[t,T]$ such that α_s is $\mathcal{F}_{(s-\mathrm{H})^+}$-measurable, namely, $\alpha_s = \alpha_s(X_{[0,(s-\mathrm{H}^+]})$. Given $(t,\mu) \in \Theta_T$ and $\alpha \in \mathcal{A}^{\mathrm{H}}_t$, denote by $\mathbb{P}^{t,\mu,\alpha}$ the unique probability measure $\mathbb{P} \in \mathcal{P}_2(\mathcal{F}_T)$ such that $\mathbb{P}_{[0,t]} = \mu_{[0,t]}$ and $\mathbb{P}$ is the strong solution of the following SDE on $[t,T]$:

$$dX_s = b(s,X.,\alpha_s)ds + \sigma(s,X.,\alpha_s)dW_s,\ t\le s\le T, \mathbb{P}\text{-a.s.} \tag{4.8}$$

We emphasize again that above $\alpha_s = \alpha_s(X_{[0,(s-\mathrm{H}^+]})$. We then define

$$V(t,\mu) := \sup_{\alpha\in\mathcal{A}^{\mathrm{H}}_t} J(t,\mu,\alpha) := \sup_{\alpha\in\mathcal{A}^{\mathrm{H}}_t} \mathbb{E}^{\mathbb{P}^{t,\mu,\alpha}}\Big[g(X.) + \int_t^T f(s,X.,\alpha_s)ds\Big]. \tag{4.9}$$

Remark 4.5. When $T \le \mathrm{H}$, α_t is $\mathcal{F}_0$-measurable for $t \in [0,T]$. Since $\mathcal{F}_0$ is not degenerate here, in general α may not be deterministic, and thus rigorously speaking the formulation here is slightly different from that in sections 2 and 3. However, they are equivalent when μ_0 is degenerate, namely, X_0 is a constant, μ-a.s.

Alternatively, following the rationale of information delay, one may require α_t to be $\mathcal{F}_{0-} := \{\emptyset, \Omega\}$-measurable for $t < \text{H}$, and thus it is deterministic. One minor disadvantage of this reformulation is that the information flow will have a jump at $t = \text{H}$. Again, this discontinuity disappears when μ_0 is degenerate.

Similar to Assumption 2.1, we shall assume the following.

Assumption 4.6. (i) b, σ, f are $\mathbb{F}$-adapted, and $b(t,0,a)$, $\sigma(t,0,a)$, and $f(t,0,a)$ are bounded;

(ii) b and σ are uniformly Lipschitz continuous in ω, uniformly continuous in t, and continuous in a;

(iii) f is uniformly continuous in (t,ω) and continuous in a, and g is uniformly continuous in ω.

Under the above assumptions, it is clear that (4.8) is well-posed, V is $\mathbb{F}$-adapted, and analogous to Theorem 2.3 one can easily prove

$$V(t,\mu) = \sup_{\alpha \in \mathcal{A}_t^{\text{H}}} \Big[V(t+\delta, \mathbb{P}^{t,\mu,\alpha}) + \int_t^{t+\delta} \mathbb{E}^{\mathbb{P}^{t,\mu,\alpha}} \big[f(s, X_\cdot, \alpha_s) \big] ds \Big]. \tag{4.10}$$

Now assume $V \in C^{1,2}(\Theta_T)$ in the sense of Definition 4.3. By (4.10), similar to Theorem 2.8 one can easily derive

$$\partial_t V(t,\mu) + H(t,\mu,\partial_\mu V, \partial_\omega \partial_\mu V) = 0, \tag{4.11}$$

where, for $p : \Theta_T \times \Omega \to \mathbb{R}^d$ and $q : \Theta_T \times \Omega \to \mathbb{R}^{d\times d}$,

$$\begin{aligned} H(t,\mu,p,q) &:= \sup_{\alpha \in \mathcal{A}_t^{\text{H}}} h(t,\mu,p,q,\alpha_t), \\ h(t,\mu,p,q,\alpha_t) &:= \mathbb{E}^\mu \Big[\big[b(\cdot)\cdot p(t,\mu,X_\cdot) + \frac{1}{2}\sigma\sigma^\top(\cdot) : q(t,\mu,X_\cdot) + f(\cdot) \big](t, X_\cdot, \alpha_t) \Big]. \end{aligned} \tag{4.12}$$

Note that α_t is $\mathcal{F}_{(t-\text{H})^+}$-measurable. Denote $\bar t := (t-\text{H})^+$ and let $\mu^{\bar t,\omega}$ denote the regular conditional probability distribution of μ given $\mathcal{F}_{\bar t}$, i.e., $\mu^{\bar t,\omega}(E) = \mathbb{E}^\mu[1_E(X_{[0,t]}) \mid \mathcal{F}_{\bar t}](\omega)$ for μ-a.e. $\omega \in \Omega$. Then,

$$\begin{aligned} h(t,\mu,p,q,\alpha_t) &:= \mathbb{E}^\mu \Big[\overline{h}\big(t, X_{[0,\bar t]}, \mu, p, q, \alpha_t(X_{[0,\bar t]})\big) \Big], \quad \text{where} \\ \overline{h}(t,\omega,\mu,p,q,a) &:= \mathbb{E}^{\mu^{\bar t,\omega}} \Big[\big[b(\cdot)\cdot p(t,\mu,X_\cdot) + \frac{1}{2}\sigma\sigma^\top(\cdot) : q(t,\mu,X_\cdot) + f(\cdot) \big](t, X_\cdot, a) \Big]. \end{aligned} \tag{4.13}$$

We remark that in (4.12) h depends on the whole random variable α_t, while in (4.13) $\overline{h}$ depends on the realized value $a \in A$. We have the following result.

THEOREM 4.7. *Let Assumption* 4.6 *hold.*

(i) *For any* $p : \Theta_T \times \mathbb{R}^d \to \mathbb{R}^d$, $q : \Theta_T \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ *uniformly continuous, the Hamiltonian* H *in* (4.12) *becomes*

$$H(t,\mu,p,q) = \mathbb{E}^\mu \Big[\sup_{a\in A} \overline{h}(t, X_{[0,\bar t]}, \mu, p, q, a) \Big]; \tag{4.14}$$

(ii) *Assume* $V \in C^{1,2}(\Theta_T)$. *Then* V *is the value function in* (4.9) *if and only if* V *satisfies the following path-dependent master equation:*

$$\partial_t V(t,\mu) + \mathbb{E}^\mu \Big[\sup_{a\in A} \overline{h}(t, X_{[0,\bar t]}, \mu, \partial_\mu V, \partial_\omega \partial_\mu V, a) \Big] = 0, \ \ V(T,\mu) = \mathbb{E}^\mu[g(X_\cdot)]. \tag{4.15}$$

Proof. (i) Define

$$\widetilde{H}(t,\mu,p,q) := \mathbb{E}^{\mu}\Big[\sup_{a\in A}\overline{h}(t,X_{[0,\overline{t}]},\mu,p,q,a)\Big].$$

It is clear that $H \le \widetilde{H}$. To see the opposite inequality, fix (t,μ,p,q) as specified in (i). By our conditions, it is obvious that $\omega \mapsto \overline{h}(t,\omega,\mu,p,q,a)$ is $\mathcal{F}_{\overline{t}}$-measurable for each a, and $a \mapsto \overline{h}(t,\omega,\mu,p,q,a)$ is continuous for each ω. Then $(\omega,a) \mapsto \overline{h}(t,\omega,\mu,p,q,a)$ is $\mathcal{F}_{\overline{t}} \times \mathcal{B}(A)$-measurable. By the standard measurable selection theorem (see, e.g., El Karoui and Tan (2013, Proposition 2.21)) for any $\varepsilon > 0$, there exists an $\mathcal{F}^{\mu}_{\overline{t}}$-measurable random variable a^{ε} such that

$$\overline{h}(t,X_{[0,\overline{t}]},\mu,p,q,a^{\varepsilon}) \ge \sup_{a\in A}\overline{h}(t,X_{[0,\overline{t}]},\mu,p,q,a) - \varepsilon, \quad \mu\text{-a.s},$$

where $\mathcal{F}^{\mu}_{\overline{t}}$ denotes the μ-augmentation of $\mathcal{F}_{\overline{t}}$. By Zhang (2017, Proposition 1.2.2), there exists $\mathcal{F}_{\overline{t}}$-measurable α^{ε}_t such that $\alpha^{\varepsilon}_t = a^{\varepsilon}$, μ-a.s. Then

$$\widetilde{H}(t,\mu,p,q) \le \mathbb{E}^{\mu}\big[\overline{h}(t,X_{[0,\overline{t}]},\mu,p,q,\alpha^{\varepsilon}_t)\big] + \varepsilon \le H(t,\mu,p,q) + \varepsilon.$$

By the arbitrariness of ε, we obtain $\widetilde{H} \le H$, and thus the equality holds.

(ii) This follows from similar arguments as in Theorem 2.8. ∎

Assume further that the following Hamiltonian $\overline{H}$ has an optimal argument a^*:

$$\overline{H}(t,\omega,\mu,\partial_\mu V,\partial_\omega\partial_\mu V) := \sup_{a\in A}\overline{h}(t,\omega,\mu,\partial_\mu V,\partial_\omega\partial_\mu V,a). \tag{4.16}$$

By (4.13), we see that a^* takes the form $I(t,\mu^{\overline{t},\omega},\omega_{[0,\overline{t}]})$. Then (4.8) becomes a McKean–Vlasov SDE again:

$$dX^*_s = b\big(s,X^*_\cdot,I(s,\mathbb{P}^{\overline{s},X^*},X^*_{[0,\overline{s}]})\big)ds + \sigma\big(s,X^*_\cdot,I(s,\mathbb{P}^{\overline{s},X^*},X^*_{[0,\overline{s}]})\big)dW_s, \mathbb{P}\text{-a.s.} \tag{4.17}$$

Similar to Theorem 2.10, one can easily prove the following.

THEOREM 4.8. *Let Assumption* 4.6 *hold and $V \in C^{1,2}(\Theta_T)$ be the classical solution to the master equation* (4.15)*. Assume further that*

(i) *the Hamiltonian $\overline{H}$ defined by* (4.16) *has an optimal control $a^* = I(t,\mu^{\overline{t},\omega},\omega_{[0,\overline{t}]})$ for any $(t,\mu)\in\Theta_T$, where $I:\Theta_T\times\Omega\to A$ is measurable;*

(ii) *for a fixed $(t,\mu)\in\Theta_T$, the McKean–Vlasov SDE* (4.17) *on $[t,T]$ has a (strong) solution $\mathbb{P}^*$ such that $\mathbb{P}^*_{[0,t]} = \mu_{[0,t]}$.*

Then $\alpha^*_s := I(s,(\mathbb{P}^*)^{\overline{s},\omega},\omega_{[0,\overline{s}]})$, $s\in[t,T]$, is an optimal control for the optimization problem (4.9) *with this fixed (t,μ).*

It will be interesting to extend Theorems 2.11 and 3.1 to this case. This requires the measurability and/or regularity in terms of the paths and is more challenging. We shall leave a more systematic study on these issues to future research. In the subsection below, we shall solve the linear quadratic case which extends the example in subsection 3.1.

Finally, consider a special case where b,σ,f do not depend on X. Then

$$\overline{h}(t,\omega,\mu,p,q,a) = \frac{1}{2}\sigma\sigma^\top(t,a) : \mathbb{E}^{\mu}[q(t,\mu,X_\cdot)|\mathcal{F}_{\overline{t}}] + b(t,a)\cdot\mathbb{E}^{\mu}[p(t,\mu,X_\cdot)|\mathcal{F}_{\overline{t}}] + f(t,a),$$

and thus a^* takes the form $a^* = I\big(t, \mathbb{E}^\mu[p(t,\mu,X.)|\mathcal{F}_{\bar t}], \mathbb{E}^\mu[q(t,\mu,X.)|\mathcal{F}_{\bar t}]\big)$. Therefore, (4.17) becomes

$$\begin{aligned} dX_s^* &= b\big(s, X_\cdot^*, I\big(s, \mathbb{E}[\partial_\mu V|\mathcal{F}_{\bar s}], \mathbb{E}[\partial_\omega\partial_\mu V|\mathcal{F}_{\bar s}]\big)\big)ds \\ &\quad + \sigma\big(s, X_\cdot^*, I\big(s, \mathbb{E}[\partial_\mu V|\mathcal{F}_{\bar s}], \mathbb{E}[\partial_\omega\partial_\mu V|\mathcal{F}_{\bar s}]\big)\big)dW_s, \quad \mathbb{P}\text{-a.s.} \end{aligned} \tag{4.18}$$

where $\partial_\mu V$ and $\partial_\omega\partial_\mu V$ are computed at $(s, \mathcal{L}_{X^*_{[0,s]}}, X_s)$.

4.4. The linear quadratic example. Consider the path-dependent setting of the example in section 3:

$$d=1,\ A=\mathbb{R},\ b(t,x,a)=a,\ \sigma=1,\ f(t,x,a)=-\frac{1}{2}a^2,\ g(x)=x^2,\ T=2\textsc{h}. \tag{4.19}$$

In this case (4.11) becomes as follows: recalling $\bar t := (t-\textsc{h})^+$,

$$\begin{aligned} &\partial_t V(t,\mu) + \frac{1}{2}\mathbb{E}^\mu\Big[\partial_\omega\partial_\mu V(t,\mu,X.)\Big] + \frac{1}{2}\mathbb{E}^\mu\Big[\Big|\mathbb{E}^\mu\big[\partial_\mu V(t,\mu,X.)|\mathcal{F}_{\bar t}\big]\Big|^2\Big] = 0, \\ &V(T,\mu) = \mathbb{E}^\mu[|X_T|^2]. \end{aligned} \tag{4.20}$$

Moreover, provided (4.20) has a classical solution, then (4.18) reduces to

$$dX_s^* = \mathbb{E}\big[\partial_\mu V(s, \mathcal{L}_{X^*_{[0,s]}}, X_s^*)|\mathcal{F}_{\bar s}\big]ds + dW_s, \quad \mathbb{P}\text{-a.s.} \tag{4.21}$$

THEOREM 4.9. *Let* (4.19) *hold. Assume* $\textsc{h} < \frac{1}{4}$ *and denote* $\overline{\textsc{h}} := \frac{1}{2} - \textsc{h}$.

(i) *The V defined by* (4.9) *is equal to*

$$V(t,\mu) = \begin{cases} \mathbb{E}^\mu\Big[|X_t|^2 + \displaystyle\int_{t-\textsc{h}}^{\textsc{h}} \frac{|\mathbb{E}_s^\mu[X_t]|^2}{2(\overline{\textsc{h}}+s)^2}ds\Big] + T - t, & t\in[\textsc{h}, 2\textsc{h}]; \\ \dfrac{\mathbb{E}^\mu[|X_t|^2]}{2(\overline{\textsc{h}}+t)} + \displaystyle\int_0^t \frac{\mathbb{E}^\mu[|\mathbb{E}_s^\mu[X_t]|^2]}{2(\overline{\textsc{h}}+s)^2}ds + \textsc{h} + \frac{1}{2}\ln\frac{1}{2(\overline{\textsc{h}}+t)} & \\ \qquad + \dfrac{\textsc{h}-t}{2\overline{\textsc{h}}(\frac{1}{2}-2\textsc{h}+t)}\mathbb{E}^\mu\Big[\big|\mathbb{E}_0^\mu[X_t]\big|^2\Big], & t\in[0,\textsc{h}). \end{cases} \tag{4.22}$$

It is in $C^{1,2}(\Theta_T)$ and is a classical solution to the path-dependent master equation (4.20).

(ii) *For any $(t,\mu)\in\Theta_T$, the SDE* (4.21) *on $[t,T]$ with initial condition $\mathbb{P}\circ(X^*_{[0,t]})^{-1} = \mu_{[0,t]}$ has a strong solution X^*, and the optimal control takes the form*

$$\alpha_t^* = \mathbb{E}^{\mathbb{P}^*}\big[\partial_\mu V(t,\mathbb{P}^*,X.)|\mathcal{F}_{(t-\textsc{h})^+}\big], \quad \textit{where} \quad \mathbb{P}^* = \mathbb{P}\circ(X^*)^{-1}. \tag{4.23}$$

The proof of this theorem is lengthy but quite standard, so we omit it here and refer readers to the arXiv version of this paper (Saporito and Zhang (2017)).

Remark 4.10. In this remark we investigate the special setting with $t=0$ and $\mu=\delta_{x_0}$. We consider three cases, always with $T=2\textsc{h}$:

(i) the case discussed in this subsection, with delay parameter $\textsc{h}$;

(ii) the case with delay parameter $2\textsc{h}$, but still with $T=2\textsc{h}$.

(iii) the case without delay, which is the case studied in the standard literature.

Denote the optimization values as V_0^{H}, $V_0^{2\mathrm{H}}$, and V_0^0, respectively. Then we have

$$V_0^{\mathrm{H}} = \frac{x_0^2}{1-4\mathrm{H}} + \mathrm{H} - \frac{1}{2}\ln(1-2\mathrm{H}),\ V_0^{2\mathrm{H}} = \frac{x_0^2}{1-4\mathrm{H}} + 2\mathrm{H};\ V_0^0 = \frac{x_0^2}{1-4\mathrm{H}} - \frac{1}{2}\ln(1-4\mathrm{H}).$$

Again we refer to the details in the arXiv version. Recalling $0 < \mathrm{H} < \frac{1}{4}$, one can easily see that $V_0^{2\mathrm{H}} < V_0^{\mathrm{H}} < V_0^0$. This indicates that the information delay indeed decreases the value function, consistent with our intuition.

Appendix A. In this appendix, we show heuristically how the stochastic maximum principle leads to the same structure as in section 2. We remark that this approach has also been used by Hu and Tang (2017) recently for a mixture of deterministic and stochastic controls in a linear quadratic setting. To focus on the main idea and simplify the presentation, we consider the following simple case with deterministic controls $\alpha \in \mathcal{A}_0$:

$$\text{(A.1)}\qquad \begin{aligned} V_0 := \sup_{\alpha\in\mathcal{A}_0} J(\alpha) := \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[g(X_T^\alpha) + \int_0^T f(t,\alpha_t)dt\Big],\\ \text{where } X_t^\alpha = x + \int_0^t b(s,\alpha_s)ds + W_t, \end{aligned}$$

where $\mathcal{A}_0$ is the set of all Borel measurable functions $\alpha : [0,T] \to A$.

Since $\mathcal{A}_0$ is convex, namely, for $\alpha, \alpha' \in \mathcal{A}$, we have $\alpha + \varepsilon(\alpha' - \alpha) \in \mathcal{A}$ for all $\varepsilon \in (0,1)$. Fix $\alpha, \alpha' \in \mathcal{A}$ and denote $\Delta\alpha := \alpha' - \alpha$, $\alpha^\varepsilon := \alpha + \varepsilon\Delta\alpha$. Assume b and f are continuously differentiable in a and g is continuously differentiable in x. Then

$$\begin{aligned} \nabla X_t &:= \lim_{\varepsilon\to 0} \frac{X_t^{\alpha^\varepsilon} - X_t^\alpha}{\varepsilon} = \int_0^t \partial_a b(s,\alpha_s)\Delta\alpha_s ds,\\ \nabla J &:= \lim_{\varepsilon\to 0} \frac{J(\alpha^\varepsilon) - J(\alpha)}{\varepsilon} = \mathbb{E}\Big[\partial_x g(X_T^\alpha)\nabla X_T + \int_0^t \partial_a f(s,\alpha_s)\Delta\alpha_s ds\Big]. \end{aligned}$$

Let $(\widetilde{Y}^\alpha, \widetilde{Z}^\alpha)$ be the solution to the following BSDE:

$$\widetilde{Y}_t^\alpha = \partial_x g(X_T^\alpha) - \int_t^T \widetilde{Z}_s^\alpha dW_s.$$

We emphasize that $(\widetilde{Y}^\alpha, \widetilde{Z}^\alpha)$ depend on α but not on $\Delta\alpha$. Then

$$\text{(A.2)}\qquad \begin{aligned} \nabla J &= \mathbb{E}\Big[\int_t^T \big[\widetilde{Y}_s^\alpha \partial_a b(s,\alpha_s) + \partial_a f(s,\alpha_s)\big]\Delta\alpha_s ds\Big]\\ &= \int_t^T \Big[\mathbb{E}[\widetilde{Y}_s^\alpha]\partial_a b(s,\alpha_s) + \partial_a f(s,\alpha_s)\Big]\Delta\alpha_s ds\Big], \end{aligned}$$

where the second equality relies on the fact that α and $\Delta\alpha$ are deterministic. Now assume $\alpha^* \in \mathcal{A}_0$ is an optimal argument, then $\Delta J \le 0$ for all possible $\Delta\alpha$. Assume further that α^* is an inner point of $\mathcal{A}$ in the sense that one may choose $\Delta\alpha$ in all directions. Then

$$\text{(A.3)}\qquad \mathbb{E}[\widetilde{Y}_t^{\alpha^*}]\partial_a b(t,\alpha_t^*) + \partial_a f(t,\alpha_t^*) = 0.$$

Assume b and f are such that the above equation determines a function $\widetilde{I}(t,x)$ such that $\alpha_t^* = \widetilde{I}(t, \mathbb{E}[\widetilde{Y}_t^{\alpha^*}])$. Then, denoting $X^* := X^{\alpha^*}, \widetilde{Y}^* := \widetilde{Y}^{\alpha^*}, \widetilde{Z}^* := \widetilde{Z}^{\alpha^*}$, we obtain the following coupled forward backward SDE (FBSDE):

$$X_t^* = x + \int_0^t b(s, \widetilde{I}(s, \mathbb{E}[\widetilde{Y}_s^*]))ds + W_t, \qquad \widetilde{Y}_t^* = \partial_x g(X_T^*) - \int_t^T \widetilde{Z}_s^* dW_s. \tag{A.4}$$

We emphasize that the above FBSDE is of McKean–Vlasov type because the forward one includes $\mathbb{E}[\widetilde{Y}_s^*]$, which is determined by the law of $\widetilde{Y}_s^*$ rather than the value of Y_s^*. Assume the above FBSDE is well-posed and we have the decoupling field: $\widetilde{Y}_t^* = \widetilde{V}(t, \mathcal{L}_{X_t^*}, X_t^*)$, which without surprise involves the law of X^*. Denote $I(t,\mu) := \widetilde{I}\big(t, \mathbb{E}[\widetilde{V}(t,\mu,\xi)]\big)$, where as usual $\mathcal{L}_\xi = \mu$. Then $\widetilde{I}(t, \mathbb{E}[\widetilde{Y}_t^*]) = I(t, \mathcal{L}_{X_t^*})$, and thus

$$X_t^* = x + \int_0^t b(s, I(s, \mathcal{L}_{X_s^*}))ds + W_t, \tag{A.5}$$

which is consistent with (2.18).

Remark A.1. When the control α_t is $\mathcal{F}_t$-measurable, the first equality of (A.2) still holds but the second fails. Due to the arbitrariness of $\Delta\alpha$, in this case the first order condition (A.3) becomes

$$\widetilde{Y}_t^{\alpha^*} \partial_a b(t, \alpha_t^*) + \partial_a f(t, \alpha_t^*) = 0. \tag{A.6}$$

This leads to $\alpha_t^* = \widetilde{I}(t, \widetilde{Y}_t^{\alpha^*})$, which in turn leads to a standard FBSDE. These are very standard arguments in the literature. Again, here due to our constraint of deterministic control, the optimal control α_t^* depends on $\mathbb{E}[\widetilde{Y}_t^*]$ instead of $\widetilde{Y}_t^*$ and hence depends on the law of X_t^*.

Appendix B. In this appendix, we will show some mathematical details of the discussion outlined in section 1.1. Specifically, we will describe some aspects of the noisy observation case and how it compares to ours. The state variable X is governed by the dynamics (2.1). For simplicity of notation, we assume $d = 1$ and $\sigma \equiv 1$ in what follows. Then,

$$\begin{aligned} X_s^{t,\xi,\alpha} &= \xi + \int_t^s b(r, X_r^{t,\xi,\alpha}, \alpha_r)dr + W_s - W_t, \ s \in [t,T], \\ J(t,p,\alpha) &:= \mathbb{E}^p\left[g(X_T^{t,\xi,\alpha}) + \int_t^T f(s, X_s^{t,\xi,\alpha}, \alpha_s)ds \right], \end{aligned} \tag{B.1}$$

for ξ with probability density p. Differently from the previous control problem, the agent observes a nonlinear noisy process given by

$$Y_s = \int_t^s h(r, X_r^{t,\xi,\alpha})dr + \widetilde{W}_s,$$

where $\widetilde{W}$ is a Brownian motion independent of W. Thus, an admissible control α has to be $\mathbb{F}^Y$-progressively measurable. We will denote this space by $\widetilde{\mathcal{A}}_{[t,T]}$. Hence, the value function is given by

$$V(t,p) := \sup_{\alpha \in \widetilde{\mathcal{A}}_{[t,T]}} J(t,p,\alpha). \tag{B.2}$$

We will follow closely the approach of Beneš and Karatzas (1983). First we introduce some notation. Given two functions $\varphi, \psi : \mathbb{R}^d \to \mathbb{R}$, denote $\langle \varphi, \psi \rangle := \int_{\mathbb{R}^d} \varphi(z)\psi(z)dz$. Given a function $F : L^2(\mathbb{R}) \longrightarrow \mathbb{R}$, its derivative with respect to p is a function $\partial_p F(p) : L^2(\mathbb{R}) \longrightarrow \mathbb{R}$, defined in the Gâteaux sense,

$$\frac{d}{d\varepsilon}F(p+\varepsilon\varphi)\bigg|_{\varepsilon=0} = \langle \partial_p F(p), \varphi\rangle,$$

for appropriate test function $\varphi : \mathbb{R}^d \to \mathbb{R}$. The second order derivative $\partial_{pp}F(p)$ is defined similarly through $\langle \partial_{pp}F, [\varphi, \psi]\rangle$ and can be viewed as a bilinear mapping. Moreover, ∂_μ and ∂_p are related through the equation (see, e.g., Bensoussan, Chau, and Yam (2017)): for measure μ with density p,

$$\partial_\mu F(\mu, x) = \partial_x \partial_p F(p)(x). \tag{B.3}$$

Beneš and Karatzas (1983) show that the dynamics of a proper unnormalized density of the distribution of X_s^α given $\mathcal{F}_s^Y$, denoted by ρ_s, is given by

$$d\rho_s^{t,p}(x) = \mathcal{L}_s^{*\alpha_s}\rho_s^{t,p}(x)dt + h(s,x)\rho_s^{t,p}(x)dY_s,$$

with $\rho_t^{t,p} = p$, which is the unnormalized density of X_t, and

$$\mathcal{L}_s^{*a} = \frac{1}{2}\partial_{xx} - b(s,x,a)\partial_x - \partial_x b(s,x,a).$$

Moreover, one may write

$$J(t,p,\alpha) = \mathbb{E}\left[\frac{\langle g, \rho_T^{t,p}\rangle}{\langle 1, \rho_T^{t,p}\rangle} + \int_t^T \frac{\langle f(s,\cdot,\alpha_s), \rho_s^{t,p}\rangle}{\langle 1, \rho_s^{t,p}\rangle} ds\right].$$

Under certain conditions, V satisfies the following HJB equation (see Beneš and Karatzas (1983, equations (2.14)–(2.15))) with terminal condition $V(T,p) = \langle g, p\rangle$:

$$\begin{aligned}\partial_t V(t,p) &+ \frac{1}{2}\big\langle \partial_{pp}V(t,p), [h(t,\cdot)p,\ h(t,\cdot)p]\big\rangle \\ &+ \sup_{a\in A}\Big[\big\langle \partial_p V(t,p), \mathcal{L}_t^{*a}p\big\rangle + \big\langle f(t,\cdot,a), p\big\rangle\Big] = 0.\end{aligned} \tag{B.4}$$

We would like to point out that the deterministic control problem studied in section 2 is equivalent to the noisy observation control problem with $h \equiv 0$, i.e., the pure noise case. Under this situation, we will now show that the master equation (2.12) is the HJB equation (B.4), when restricted to those measures with density. In fact, in this case, the HJB equation (B.4) becomes

$$\partial_t V(t,p) + \sup_{a\in A}\Big[\big\langle \partial_p V(t,p), \mathcal{L}_t^{*a}p\big\rangle + \big\langle f(t,\cdot,a), p\big\rangle\Big] = 0. \tag{B.5}$$

By using the integrating by parts formula, we have

$$\begin{aligned}\langle \partial_p V(t,p), \partial_{xx}p\rangle &= \langle \partial_{xx}\partial_p V(t,p), p\rangle; \\ \langle \partial_p V(t,p), b(t,\cdot,a)\partial_x p\rangle &= -\big\langle \partial_x\partial_p V(t,p) b(t,\cdot,a) + \partial_p V(t,p)\partial_x b(t,\cdot,a),\ p\big\rangle.\end{aligned}$$

Then, for measure μ with density and by using (B.3),

$$\begin{aligned}\langle \partial_p V(t,p), \mathcal{L}_t^{*a}p\rangle &= \Big\langle \partial_p V(t,p),\ \frac{1}{2}\partial_{xx}p - b(t,\cdot,a)\partial_x p - \partial_x b(t,\cdot,a)p\Big\rangle \\ &= \Big\langle \frac{1}{2}\partial_{xx}\partial_p V(t,p) + b(t,\cdot,a)\partial_x\partial_p V(t,p),\ p\Big\rangle \\ &= \Big\langle \frac{1}{2}\partial_x\partial_\mu V(t,\mu) + b(t,\cdot,a)\partial_\mu V(t,\mu),\ p\Big\rangle.\end{aligned}$$

Plugging this into (B.4) we obtain our master equation (2.12) immediately.

REFERENCES

Y. Alekal, P. Brunovsky, D. Chyung, and E. Lee (1971), *The quadratic problem for systems with time delays*, IEEE Trans. Automat. Control, 16, pp. 673–687.

F. Baghery and B. Øksendal (2007), *A maximum principle for stochastic control with partial information*, Stoch. Anal. Appl., 25, pp. 705–717.

E. Bandini, A. Cosso, M. Fuhrman, and H. Pham (2018), *Randomization method and backward SDEs for optimal ccontrol of partially observed path-dependent stochastic systems*, Ann. Appl. Probab., 28, pp. 1634–1678.

E. Bandini, A. Cosso, M. Fuhrman, and H. Pham (2019), *Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem*, Stochastic Process. Appl., 129, pp. 674–711.

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

E. Bayraktar and Z. Zhou (2017), *On an optimal stopping problem of an insider*, Theory Probab. Appl., 61, pp. 133–139.

V. Beneš and I. Karatzas (1983), *On the relation of Zakai's and Mortensen's equation*, SIAM J. Control Optim., 21, pp. 472–489.

A. Bensoussan (1992), *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, UK.

A. Bensoussan, M. H. M. Chau, and S. C. P. Yam (2015), *Mean field Stackelberg games: Aggregation of delayed instructions*, SIAM J. Control Optim., 53, pp. 2237–2266.

A. Bensoussan, M. H. M. Chau, and S. C. P. Yam (2017), *Linear-quadratic mean field Stackelberg games with state and control delays*, SIAM J. Control Optim., 55, pp. 2748–2781.

A. Bensoussan, J. Frehse, and S. C. P. Yam (2013), *Mean Field Games and Mean Field Type Control Theory*, Springer Briefs Math., Springer, New York.

A. Bensoussan, J. Frehse, and S. C. P. Yam (2017), *On the interpretation of the master equation*, Stoch. Proc. Appl., 127, pp. 2093–2137.

A. Bensoussan and S. C. P. Yam, *Control problem on space of random variables and master equation*, ESAIM Control Optim. Calc. Var., to appear.

J.-M. Bismut, *Partially observed diffusions and their control* (1982), SIAM J. Control Optim., 20, pp. 302–309.

R. Buckdahn, J. Li, S. Peng, and C. Rainer (2017), *Mean-field stochastic differential equations and associated PDEs*, Ann. Probab., 45, pp. 824–878.

P. Caines, M. Huang, and R. Malhame (2006), *Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6, pp. 221–252.

P. Cardaliaguet (2013), *Notes on Mean Field Games (from P.-L. Lions lectures at College de France)*, preprint, https://www.ceremade.dauphine.fr/~cardaliaguet/MFG100629.pdf.

P. Cardaliaguet, F. Delarue, J. M. Lasry, and P. L. Lions (2015), *The Master Equation and the Convergence Problem in Mean Field Games*, preprint, arXiv:1509.02505.

R. Carmona and F. Delarue (2017a), *Probabilistic Theory of Mean Field Games* I*: Mean Field FBSDEs, Control, and Games*, Springer, New York.

R. Carmona and F. Delarue (2017b), *Probabilistic Theory of Mean Field Games* II*: Mean Field Games with Common Noise and Master Equations*, Springer, New York.

R. Carmona, J.-P. Fouque, S. M. Mousavi, and L.-H. Sun (2018), *Systemic risk and stochastic games with delay*, J. Optim. Appl., 179, pp. 366–399.

J.-F. Chassagneux, D. Crisan, and F. Delarue (2014), *A Probabilistic Approach to Classical Solutions of the Master Equation for Large Population Equilibria*, preprint, arXiv:1411.3009.

L. Chen and Z. Wu (2011), *The quadratic problem for stochastic linear control systems with delay*, in Proceedings of the 30th Chinese Control Conference, pp. 1344–1349.

N. Christopeit (1980), *Existence of Optimal Stochastic Controls under Partial Observation*, Probab. Theory Related Fields, 51, pp. 201–213.

B. Dupire (2009), *Functional Itô Calculus*, http://ssrn.com/abstract=1435551.

N. El Karoui and X. Tan (2013), *Capacities, Measurable Selection and Dynamic Programming. Part* I*: Abstract Framework*, preprint, arXiv:1310.3363.

I. Ekren, N. Touzi, and J. Zhang (2016), *Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part* I, Ann. Probab., 44, pp. 1212–1253.

W. Fleming (1980), *Measure-valued processes in the control of partially-observable stochastic systems*, Appl. Math. Optim., 6, pp. 271–285.

W. Fleming (1982), *Nonlinear semigroup for controlled partially observed diffusions*, SIAM J. Control Optim., 20, pp. 286–301.

W. Fleming and E. Pardoux (1982), *Optimal control for partially observed diffusions*, SIAM J. Control Optim., 20, pp. 261–285.

W. Gangbo and A. Swiech (2015), *Existence of a solution to an equation arising from the theory of mean field games*, J. Differential Equations, 259, pp. 6573–6643.

F. Gozzi and C. Marinelli (2006), *Stochastic optimal control of delay equations arising in advertising models*, in Stochastic Partial Differential Equations and Applications VII, G. D. Prato and L. Tubaro, eds., Chapman & Hall/CRC, Press, Boca Raton, FL, pp. 133–148.

F. Gozzi and F. Masiero (2017), *Stochastic optimal control with delay in the control* I: *Solving solution through partial smoothing*, SIAM J. Control Optim., 55, pp. 2981–3012.

Y. Hu and S. Tang (2017), *Mixed Deterministic and Random Optimal Control of Linear Stochastic Systems with Quadratic Costs*, preprint, arXiv:1708.06547.

T. Ichiba and S. M. Mousavi (2017), *Option Pricing with Delayed Information*, preprint, arXiv:1707.01600.

J. Larsy and P. Lions, *Mean field games* (2007), Jpn. J. Math., 2, pp. 229–260.

R. Mortensen (1996), *Stochastic optimal control with noisy observations*, Internat. J. Control., 4, pp. 455–464.

H. Pham and X. Wei (2018), *Bellman equation and viscosity solutions for mean-field stochastic control problem*, ESAIM Control Optim. Calc. Var., 24, pp. 437–461.

Y. F. Saporito (2017), *Stochastic Control and Differential Games with Path-Dependent Controls*, preprint, arXiv:1611.00589.

Y. F. Saporito and J. Zhang (2017), *Stochastic Control with Delayed Information and Related Nonlinear Master Equation*, arXiv:1710.05609.

S. Tang (1998), *The maximum principle for partially observed optimal control of stochastic differential equations*, SIAM J. Control Optim., 36, pp. 1596–1617.

C. Wu and J. Zhang (2017), *An Elementary Proof for the Structure of Wasserstein Derivatives*, preprint, arXiv:1705.08046.

C. Wu and J. Zhang (2018), *Viscosity Solutions to Parabolic Master Equations and McKean-Vlasov SDEs with Closed-Loop Controls*, preprint, arXiv:1805.02639.

J. Zhang (2017), *Backward Stochastic Differential Equations—From Linear to Fully Nonlinear Theory*, Springer, New York.