Solving a Class of Mean-Field LQG Problems

Yun Li, Qingshuo Song, Fuke Wu, and George Yin

***Abstract*—In this article, we study a class of mean-field linear quadratic Gaussian problems. Under suitable conditions, explicit solutions of the distribution-dependent optimal control problems are obtained. Riccati systems are derived by directly solving the associated master equations. Some extensions on controls with partial observations are also considered.**

***Index Terms*—Controlled diffusion, linear quadratic Gaussian (LQG) control, McKean–Vlasov equation, partially observable system.**

I. Introduction

We consider a one-dimensional LQG problem. Suppose the controlled process $X_t \in \mathbb{R}$ is the solution of a stochastic differential equation

$$dX_t = (A_t X_t + B_t u_t)dt + \sigma_t dW_t \tag{1}$$

where A_t, B_t, and σ_t are suitable functions of t, W_t is a standard real-valued Brownian motion, and u_t is the control. The objective is to minimize an expected cost function of the form

$$J(x, u) = \mathbb{E}_x \left[\int_0^T (R_t X_t^2 + Q_t u_t^2) dt \right] + \hat{g}(X_T)$$

where $R_t X_t^2 + Q_t u_t^2$ is the running cost rate and $\hat{g}(X_T)$ is the terminal cost.

If the terminal cost is $\hat{g}(X_T) = \mathbb{E}[X_T^2]$, it is the classical LQG problem; see, for example, Fleming and Rishel [4] and Yong and Zhou [19], among others. There is a vast literature for LQG control problems under complete observations as well as partial observations; see, for example, [4], [7], [8], [19], and related works in [3], [9], [11], and [12], among others. It is now standard that the associated Hamilton–Jacobi–Bellman (HJB) equations can be solved by the associated Riccati equations provided that the cost function is quadratic in the states and controls.

In this article, we study the control problem with the terminal cost given by a function not of the state but the distribution μ_T of the terminal state X_T. For instance, consider $\hat{g}(X_T) = g(\mu_T) = (\mathbb{E}[X_T])^2$. Then, it does not belong to traditional setup of the LQG problem. As noted in [1] and [17], this problem belongs to the class of time-inconsistent control problems. Indeed, in such a problem, the dynamic programming principle is not applicable.

An extensive literature is devoted to time-inconsistent control problems; see [1], [2], [16], [18], [21], and the references therein. It is worth mentioning because of no time-consistent optimal controls, the focus in the aforementioned references is to find "locally optimal" time-consistent controls, which are referred to as "equilibrium solution."

We emphasize that the optimal solutions are strictly different from the equilibrium solution discussed in the aforementioned references. For the optimal solution, Yong [17] provides the Riccati system based on the decoupling technique for FBSDE; see also [18, Example 1.2], [6, Sec. 6.7], [13], and [20].

In contract to the aforementioned works, our aim is to obtain explicit solutions by solving its associated master equation directly in Section III. The solution will provide us with insight on the dependence of the solution on the associated distribution. The key is to identify the time-inconsistent problem as an LQ control problem in a suitable sense, where linear and quadratic structure is referred to the functions with domains being suitable measure spaces. Similar to the approach of traditional LQG problems, we also guess the solution of the master equation as a quadratic function of the associate measure. This approach successfully reduces originally infinite-dimensional master equation to a finite-dimensional Riccati system after explicit computations using L-derivatives; see Section II, [5], and [6] for a brief introduction of L-derivatives. Using our new approach, Example 2 in this article recovers [18, Example 1.2]. As a result, the optimal trajectory is a Gaussian process, which justifies the underlying LQ problem being linear quadratic Gaussian.

Section IV is concerned with an extension of mean-field LQG problems in which the system is only partially observable. The optimal control can be obtained by a separation principle to covert the partially observed system to a fully observed one. Finally, we conclude this article with a brief discussion in Section V.

II. Preliminaries

A. Polynomials and Derivatives on Measure Space

Suppose μ is a distribution on Borel sets $\mathcal{B}(\mathbb{R})$ and $f : \mathbb{R} \mapsto \mathbb{R}$ is a real-valued function. We write

$$\langle f, \mu \rangle := \int_{\mathbb{R}} f(x) \mu(dx)$$

if the integral exists. We denote by

$$[\mu]_m := \langle x^m, \mu \rangle$$

the mth moment for any $m \geq 1$. If a distribution μ has a finite mth moment $[\mu]_m$, then we write it as $\mu \in \mathcal{P}_m$. For instance, for any $x \in \mathbb{R}$, a Dirac measure δ_x belongs to $\mathcal{P}_m$ for any $m \geq 1$, since $[\delta_x]_m = x^m$ holds.

Polynomials on $\mathcal{P}_2$ are defined as a linear combination of the monomials defined in this following.

Manuscript received July 31, 2020; revised December 22, 2020; accepted April 8, 2021. Date of publication April 14, 2021; date of current version February 28, 2022. The work of Yun Li and Fuke Wu was supported by the National Natural Science Foundation of China under Grant 61873320. The work of Qingshuo Song was supported in part by the RGC of Hong Kong CityU under Grant 11201518 and in part by the WPI TRIAD under Grant 2019/2020. The work of George Yin was supported by the Air Force Office of Scientific Research under Grant FA9550-18-1-0268. Recommended by Associate Editor P. G. Mehta. *(Corresponding author: George Yin.)*

Yun Li and Fuke Wu are with the School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: li_yun@hust.edu.cn; wufuke@mail.hust.edu.cn).

Qingshuo Song is with the Department of Mathematics, Worcester Polytechnic Institute, Hopkinton, MA 01748 USA (e-mail: qsong@wpi.edu).

George Yin is with the University of Connecticut, Storrs, CT 06269-1009 USA (e-mail: gyin@wayne.edu).

Digital Object Identifier 10.1109/TAC.2021.3073204

0018-9286 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.

1) A 1-monomial is given by a function in the form of

$$f(\mu) = \langle \phi, \mu \rangle$$

for some appropriate function $\phi : \mathbb{R} \mapsto \mathbb{R}$.

2) An n-monomial is a product of n many 1-monomials

$$f(\mu) = \Pi_{i=1}^{n} \langle \phi_i, \mu \rangle$$

for some coefficients ϕ_i.

We use a notion of L-derivative on the functions of probability measures in a lifted space. We summarize a few useful results to be used in this article, which are as follows.

1) The derivative of 1-monomial becomes μ-invariant

$$\partial_\mu \langle \phi, \mu \rangle = \phi'(x).$$

2) Chain rule and product rules can be used as usual, which yields that the derivative of n-monomial becomes $(n-1)$-monomial. For instance, we have

$$\partial_\mu([\mu]_m)^n = n[\mu]_m^{n-1} m x^{m-1}.$$

Note that the notion of L-derivative $\partial_\mu f$ is taken from [6], which is equivalent to the intrinsic derivative $D_\mu f$ introduced in [5], i.e.,

$$\partial_\mu f(\mu, x) = D_\mu f(\mu, x) = \partial_x \frac{\delta f}{\delta \mu}(\mu, x).$$

B. Verification Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a complete filtered probability space satisfying the usual conditions, where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is the filtration on which there exists an $\mathbb{F}$-adapted Brownian motion W. Given a controlled SDE

$$X_t = x + \int_0^t b(s, X_s, u_s)ds + \int_0^t \sigma_s dW_s \tag{2}$$

we denote by μ_t the probability law of X_t and consider the cost function

$$J(u) = \mathbb{E}\left[\int_0^T \ell(t, X_t, u_t)dt\right] + g(\mu_T) \tag{3}$$

where u is an $\mathcal{F}_t$ progressively measurable control process, $\ell(\cdot,\cdot,\cdot)$ is the running cost function, and $g(\cdot)$ is the terminal cost. Our objective is to minimize the cost function J over an admissible control space $\mathcal{U}$, i.e.,

$$V^* = J(u^*) \leq J(u)\ \forall u \in \mathcal{U}. \tag{4}$$

Definition 1: A random process $u : [0,T] \times \Omega \mapsto \mathbb{R}$ is said to be admissible if u together with (X, J) satisfies (2)–(3) and

$$u_t = a(t, \mu_t, X_t) \text{ for all } t \in [0,T] \tag{5}$$

for some controller a in the feedback form of (t, μ_t, X_t). The collection of all such admissible controls is denoted by $\mathcal{U}$.

Since the terminal cost is a function of a measure, we lift the optimal value V^* to a value function of the form $V(t,\mu)$ such that $V^* = V(0, \delta_x)$ accordingly. The verification theorem says that under sufficient regularity, the value function $V(t,\mu)$ solves the following master equation:

$$\begin{aligned}&\inf_{a\in\mathcal{M}(\mathbb{R})} \langle H(t, \cdot, \mu, v, a(\cdot)), \mu \rangle \\ &\quad + \tfrac{1}{2}\sigma_t^2 \langle \partial_{x\mu} v(t, \mu, \cdot), \mu\rangle + \partial_t v(t,\mu) = 0\end{aligned} \tag{6}$$

with the terminal condition

$$v(T, \mu) = g(\mu) \tag{7}$$

where $\mathcal{M}(D)$ is the collections of all real-valued measurable mappings on a metric space D, and H is given as

$$H(t, x, \mu, v, a) = b(t,x,a)\partial_\mu v(t, \mu, x) + \ell(t, x, a).$$

Throughout the rest of this article, we use the convention $f(t,\mu)(x) = f(t,\mu,x)$. To proceed, we say a function $f : [0,T] \times \mathcal{P}_2 \mapsto \mathbb{R}$ is partial $\mathcal{C}^{1,2}$ if there exists continuous derivatives $\partial_t f, \partial_\mu f, \partial_{x\mu} f : [0,T] \times \mathcal{P}_2 \times \mathbb{R} \mapsto \mathbb{R}$. For convenience, we denote by $\mathcal{C}_I$ all partial $\mathcal{C}^{1,2}$ functions f satisfying a growth condition $\langle |\partial_{x\mu} f|^2, \mu\rangle \leq C(1 + [\mu]_2^m)$ for some $C, m > 0$. Recalling the chain rule [6, Proposition 5.102], a function $f \in \mathcal{C}_I$ satisfies

$$\begin{aligned}f(t, \mu_t) = f(0, \mu_0) &+ \textstyle\int_0^t \mathbb{E}[\partial_\mu f(s, \mu_s, X_s) b(s, X_s, u_s)]ds \\ &+ \tfrac{1}{2}\textstyle\int_0^t \mathbb{E}[\sigma_s^2 \partial_{x\mu} f(s,\mu_s, X_s)]ds + \int_0^t \partial_t f(s, \mu_s)ds.\end{aligned}$$

Proposition 2: Let b and ℓ be Lipschitz continuous in (t,x). Suppose there exists a solution $v \in \mathcal{C}_I$ of the master equation (6)–(7) and a feedback form $a^* : (0,T) \times \mathcal{P}_2 \times \mathbb{R} \mapsto \mathbb{R}$ satisfying the optimality condition

$$H(t, x, \mu, v, a^*(t, \mu, x)) = \inf_{a\in\mathbb{R}} H(t, x, \mu, v, a) \tag{8}$$

for all $(t, \mu, x) \in (0,T) \times \mathcal{P}_2 \times \mathbb{R}$. In addition, if there exists an optimal pair (X^*, u^*) of state trajectory and admissible control satisfying

$$u_t^* = a^*(t, \mu_t^*, X_t^*)$$

then the optimal value is

$$V^* = v(0, \delta_x).$$

Proof: Applying the chain rule to the solution v of the master equation, for any control $u \in \mathcal{U}$, we have

$$\begin{aligned}v(t,\mu_t) = g(\mu_T) &- \int_t^T \partial_t v(s, \mu_s)ds \\ &- \int_t^T \mathbb{E}[\partial_\mu v(s, \mu_s, X_s) b(s, X_s, u_s)]ds \\ &- \frac{1}{2}\int_t^T \sigma_s^2 \mathbb{E}[\partial_{x\mu} v(s,\mu_s, X_s)]ds\end{aligned}$$

Since $u \in \mathcal{U}$ and v solves (6), there exists feedback form $u_t = a(t, \mu_t, X_t)$ and we can write

$$\begin{aligned}&\{\langle H(s, \cdot, \mu_s, v, a(s, \mu_s, \cdot)),\ \mu_s\rangle\} \\ &\quad + \frac{1}{2}\sigma_s^2 \langle \partial_{x\mu} v(s, \mu_s, \cdot), \mu_s\rangle + \partial_t v(s, \mu_s) \geq 0.\end{aligned}$$

Therefore, with the definition of $H(\cdot)$, we obtain

$$\begin{aligned}&\mathbb{E}[\partial_\mu v(s, \mu_s, X_s) b(s, X_s, u_s)] \\ &\quad + \frac{1}{2}\sigma_s^2 \mathbb{E}[\partial_{x\mu} v(s, \mu_s, X_s)] + \partial_t v(s, \mu_s) \geq -\mathbb{E}[\ell(s, X_s, u_s)].\end{aligned}$$

This implies that

$$v(0, \mu_0) \leq g(\mu_T) + \int_0^T \mathbb{E}[\ell(t, X_t, u_t)]dt = J(u)$$

for any control $u \in \mathcal{U}$ and initial distribution μ_0. The other direction $V^* = J(u^*) \leq J(u)$ is straightforward. ■

The verification theorem has been studied in various forms for McKean–Vlasov control problems, for instance, [6, Proposition 6.32]. Proposition 2 is tailor-made for our calculation compared to [6, Proposition 6.32] in that Proposition 2 characterizes $v(t,\mu)$ while the latter does the verification of its kernel $V(t, x, \mu)$. In this sense, Proposition 2 can be considered as a generalization of [6, Proposition 5.108] with

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.

general cost structure. It is also worth mentioning that the $\inf_{a\in\mathbb{R}} H(\cdot)$ is used for the optimality condition (8) to simplify our calculation, but it can be replaced by $\inf_{a\in\mathcal{M}(\mathbb{R})}\langle H(\cdot),\mu\rangle$ for a general purpose. As mentioned, our main objective in this article is to obtain explicit solutions of the control problems.

III. LQG: Fully Observable Case

A. Setup

We consider the following simplified version of mean-field LQG problem. It appears to be more instructive to choose a simpler formulation so that we can bring out the main feature of the underlying problem. For general setup (2)–(4), the coefficients or the functions are given as

$$b(t,x,u) = A_t x + B_t u,\ \ell(t,x,u) = Q_t u^2 \tag{9}$$

and

$$\begin{aligned} g(\mu_T) &= D_1[\mu_T]_2 + D_2[\mu_T]_1^2 \\ &= D_1\mathbb{E}[X_T^2] + D_2(\mathbb{E}[X_T])^2 \end{aligned} \tag{10}$$

for some continuous and bounded A_t, B_t, Q_t and constants D_1, D_2. Note that g is polynomial of degree 2 in μ.

Example 1 (A standard LQG): If

$$A\equiv 0, B\equiv 1, \sigma\equiv 1, Q\equiv 1, D_1 = 1, D_2 = 0 \tag{11}$$

then the problem is a standard LQG problem. Note that the terminal cost $g(\mu_T) = [\mu_T]_2$ is linear in measure. In this case, the dynamic programming principle is applicable and its HJB can be explicitly solved.

Example 2:

This problem is taken from [18]. Let

$$A\equiv 0, B\equiv 1, \sigma\equiv 1, Q\equiv 1, D_2 = 1, D_1 = 0. \tag{12}$$

Note that, the terminal cost $g(\mu_T) = [\mu_T]_1^2$ is a quadratic function in μ_T and the HJB does not hold.

B. Semiexplicit Solution in Terms of Riccati Equations

In this section, we solve explicitly the master equation (6)–(7) and apply Proposition 2 to the control problem.

(A1) $Q_t > 0$ for all t.

With parameters given by (9), the Hamiltonian in the optimality condition (8) is quadratic in action a

$$H(t,x,\mu,v,a) = (A_t x + B_t a)\partial_\mu v(t,\mu,x) + Q_t a^2.$$

Since $Q_t > 0$, the infimum over $a\in\mathbb{R}$ is attained at

$$a^*(t,\mu,x) = -\frac{B_t\partial_\mu v(t,\mu,x)}{2Q_t}$$

with its minimum

$$\inf_{a\in\mathbb{R}} H(t,x,\mu,v,a) = A_t x\partial_\mu v - \frac{B_t^2}{4Q_t}|\partial_\mu v|^2.$$

Therefore, master equation (6) becomes

$$\langle L_0 v(t,\mu,\cdot),\mu\rangle + \partial_t v(t,\mu) = 0 \tag{13}$$

where the operator L_0 is defined by

$$L_0 v := \left(A_t x\partial_\mu v - \frac{B_t^2}{4Q_t}|\partial_\mu v|^2 + \frac{1}{2}\sigma_t^2\partial_{x\mu}v\right).$$

Similar to the traditional approach in LQG, we start with a guess of the value function in a quadratic function form

$$v(t,\mu) = \phi_1(t)[\mu]_2 + \phi_2(t)[\mu]_1^2 + \phi_3(t).$$

Then, we use the method of undetermined "coefficients" to determine the three-dimensional vector function $\phi = (\phi_1,\phi_2,\phi_3)$. One can directly write the derivative as

$$\partial_\mu v(t,\mu,x) = 2\phi_1(t)x + 2\phi_2(t)[\mu]_1$$

which is a polynomial in x. Moreover, we have

$$\partial_t v(t,\mu) = \phi_1'(t)[\mu]_2 + \phi_2'(t)[\mu]_1^2 + \phi_3'(t)$$

and

$$\partial_{x\mu}v(t,\mu,x) = 2\phi_1(t).$$

By plugging the derivatives in (13) and combining the like terms, the master equation yields that

$$0 = [\mu]_2 L_1\phi(t) + [\mu]_1^2 L_2\phi(t) + L_3\phi(t) \tag{14}$$

where $L = [L_1, L_2, L_3] : C^1((0,T),\mathbb{R}^3)\mapsto C((0,T),\mathbb{R}^3)$ are operators acted on the vector function $\phi = (\phi_1,\phi_2,\phi_3)$ as

$$\begin{aligned} L_1\phi(t) &= \phi_1'(t) - \frac{B_t^2}{Q_t}\phi_1^2(t) + 2A_t\phi_1(t) \\ L_2\phi(t) &= \phi_2'(t) - \frac{B_t^2}{Q_t}\phi_2^2(t) - \frac{2B_t^2}{Q_t}\phi_1(t)\phi_2(t) + 2A_t\phi_2(t) \\ L_3\phi(t) &= \phi_3'(t) + \sigma_t^2\phi_1(t). \end{aligned}$$

Since (14) holds for all μ together with terminal condition, we have the following system of ODEs in terms of the first-order differential operator L:

$$L\phi(t) = 0 \quad \forall t\in(0,T), \text{ with } \phi(T) = (D_1, D_2, 0). \tag{15}$$

Note that $L\phi$ is a linear combination of $\phi'(\cdot)$ and quadratic functions in ϕ. Such a system $L\phi = 0$ is referred to as a system of Riccati equations. One can easily verify the growth condition for $\partial_{x\mu}v$, and carry out verification theorem to conclude the following result. Furthermore, one can readily verify that the optimal path follows the Gaussian process.

Theorem 3: Suppose $Q_t > 0$ for all t, and there exists $\phi\in C^1((0,T),\mathbb{R}^3)$ for Riccati system (15). Then, the pair (v, a^*) is given by

$$v(t,\mu) = \phi_1(t)[\mu]_2 + \phi_2(t)[\mu]_1^2 + \phi_3(t)$$

and

$$a^*(t,\mu,x) = -\frac{B_t}{Q_t}(\phi_1(t)x + \phi_2(t)[\mu]_1)$$

solves the master equation (6)–(7) and the optimality condition (8). Moreover, if $J(u^*)$ of (3) with parameter sets (9)–(10) is well defined via (X^*, u^*) satisfying (2)–(3) and

$$u_t^* = a^*(t,\mu_t^*,X_t^*)$$

then (X^*, u^*) are optimal trajectory and optimal control, respectively, and the optimal value is

$$V^* = v(0,\delta_x).$$

C. Examples: Explicit Solutions

We use Theorem 3 to solve both traditional LQG Example 1 and mean-field LQG Example 2. In both cases, the Riccati system (15) becomes

$$\begin{aligned} \phi_1' &= \phi_1^2 \\ \phi_2' &= \phi_2^2 + 2\phi_1\phi_2 \\ \phi_3' &= -\phi_1. \end{aligned} \tag{16}$$

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.

1) Solution of Example 1: This problem can be solved using the traditional LQG approach; see [19]. To use Theorem 3, one can solve (16) with a terminal condition

$$\phi_1(T) = 1, \phi_2(T) = \phi_3(T) = 0.$$

The solution for this Riccati system can be written as follows. For all $t \in (0, T)$

$$\phi_2(t) = 0$$

$$\phi_1(t) = \frac{1}{1+T-t}$$

$$\phi_3(t) = \ln(1+T-t)$$

which yields the optimal strategy

$$u_t^* = -\frac{X_t^*}{1+T-t}$$

and the value function

$$v(t,\mu) = \frac{[\mu]_2}{1+T-t} + \ln(1+T-t).$$

Thus, the optimal value is

$$V^* = v(0, \delta_x) = \frac{x^2}{1+T} + \ln(1+T).$$

2) Solution of Example 2: The solution given in [18] is attained by decoupling FBSDEs and we recover it using Theorem 3. We solve the Riccati system (16) but with different terminal conditions

$$\phi_2(T) = 1, \phi_1(T) = \phi_3(T) = 0.$$

The solution for this Riccati system can be written as: For all $t \in (0, T)$

$$\phi_1(t) = \phi_3(t) = 0, \text{ and } \phi_2(t) = \frac{1}{1+T-t}.$$

Hence, the optimal strategy is

$$u_t^* = -\frac{\mathbb{E}[X_t^*]}{1+T-t}$$

and the value function is

$$v(t,\mu) = \frac{1}{1+T-t}\left(\int_{\mathbb{R}} x\mu(dx)\right)^2$$

which implies the optimal value

$$V^* = \frac{x^2}{1+T}.$$

IV. Mean-Field LQG: Controlled Systems Under Partial Observations

The following interesting question considered in [15] motivates our second example. Given a $\mathbb{F} = \{\mathcal{F}_t : 0 \le t \le T\}$ progressively measurable process $u : [0,T] \times \Omega \mapsto \mathbb{R}$, we say $u \in L^2_{\mathbb{F}}$ if $\mathbb{E}[\int_0^T |u_s|^2 ds] < \infty$. A deterministic function $u : [0,T] \mapsto \mathbb{R}$ is said to be $u \in L^2([0,T])$, if $\int_0^T |u_s|^2 ds < \infty$. Note that both $L^2_{\mathbb{F}}$ and $L^2([0,T])$ are both Hilbert spaces. We ask the following question.

- How does the optimal value of (2)–(4) change if $L^2_{\mathbb{F}}$ is replaced by $L^2([0,T])$?

Roughly speaking, the question can be interpreted as: What is the infimum that can be achieved if the control u is only allowed to be a deterministic process instead of a random one? It is obvious that the optimal value achieved in the space of deterministic controls is no less than the value with random controls due to $L^2([0,T]) \subset L^2_{\mathbb{F}}$. In what follows, we consider more general questions.

A. Setup

Recall that we are working with $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. Suppose that on this filtered probability space, there exist two independent Brownian motions $\hat{W}$ and $\widetilde{W}$, respectively. For simplicity, we assume $\mathbb{F} = \mathbb{F}^{\hat{W}} \vee \mathbb{F}^{\widetilde{W}}$ and $\mathcal{F} = \mathcal{F}_T^{\hat{W}} \vee \mathcal{F}_T^{\widetilde{W}}$, where $\mathbb{F}^{\hat{W}} = (\mathcal{F}_t^{\hat{W}})_{0\le t\le T}$ and $\mathbb{F}^{\widetilde{W}} = (\mathcal{F}_t^{\widetilde{W}})_{0\le t\le T}$ are the filtrations generated by $\hat{W}$ and $\widetilde{W}$, respectively.

Let $\hat{\sigma}, \widetilde{\sigma}, \hat{\eta}, \widetilde{\eta}$ be nonnegative constants satisfying

$$\hat{\sigma}^2 + \widetilde{\sigma}^2 = 1, \ \hat{\eta}^2 + \widetilde{\eta}^2 = 1.$$

A generic player with its initial state X_s at time s has its evolution under control u in the form of

$$X_t = X_s + \int_s^t u_r dr + \int_s^t \hat{\sigma} d\hat{W}_r + \int_s^t \widetilde{\sigma} d\widetilde{W}_r. \tag{17}$$

For simplicity, we require X_s to have a normal distribution $\mathcal{N}(x, s)$ given by

$$X_s = x + \hat{\eta}\hat{W}_s + \widetilde{\eta}\widetilde{W}_s. \tag{18}$$

The cost functional to be minimized is given by

$$J(u) = \mathbb{E}\left[\int_s^T u_r^2 dr\right] + D_1[\mu_T]_2 + D_2[\mu_T]_1^2. \tag{19}$$

The distinction of the current problem compared with the previous control problem is the following crucial point. Although the player wants to minimize the cost functional, he or she cannot directly access to the state X_t due to the lack of the knowledge for $\widetilde{W}_t$ and, hence, for W_t. Instead, he or she is up to design a controller using the prediction process

$$\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^{\hat{W}}]. \tag{20}$$

We denote by $\hat{\mu}_t$ the distribution induced by $\hat{X}_t$, i.e., $\hat{\mu}_t = \mathbb{P}\hat{X}_t^{-1}$. Indeed, $\hat{X}_t$ can be written as

$$\hat{X}_t = x + \hat{\eta}\hat{W}_s + \int_s^t u_r dr + \int_s^t \hat{\sigma} d\hat{W}_r. \tag{21}$$

Definition 4: A random process $u : [0,T] \times \Omega \mapsto \mathbb{R}$ is said to be admissible if $u \in L^2_{\mathbb{F}^{\hat{W}}}$ together with (X, J) satisfies (17)–(19) and

$$u_t = a(t, \hat{\mu}_t, \hat{X}_t) \text{ for all } t \in [0,T] \tag{22}$$

for some controller a. The collection of all such admissible controls is denoted by $\hat{\mathcal{U}}$.

Now, we are ready to define the optimal value under partial observation by

$$V^* = \inf_{u \in \hat{\mathcal{U}}} J(u). \tag{23}$$

Note that if $u \in L^2([0,T])$, then one can verify with $a(t,\mu,x) = u_t$ that $u \in \hat{\mathcal{U}}$ by definition. We remark that if $s = 0$ and $\hat{\sigma} = 0$, then $\hat{X}_t$ of (21) is deterministic, $L^2([0,T]) = \hat{\mathcal{U}}$ holds.

B. Semiexplicit Solution: Separation Principle

We use the separation principle in filtering theory. The treatment of the problem is outlined ahead.

- *Step 1:* Let $\hat{X}$ be the prediction of X given by (20) and $\mathcal{E}$ and P are the error term and variance of the error term, respectively

$$\mathcal{E}_t = X_t - \hat{X}_t, \ P_t = \mathbb{E}[\mathcal{E}_t^2].$$

Then, $\mathcal{E}$ and P satisfy

$$\mathcal{E}_t = \widetilde{\eta}\widetilde{W}_s + \widetilde{\sigma}(\widetilde{W}_t - \widetilde{W}_s)$$

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.

and

$$P_t = \widetilde{\eta}^2 s + \widetilde{\sigma}^2 (t-s).$$

Recall that $\hat{\mu}_t$ to denote the distribution of $\hat{X}_t$. Owing to

$$[\mu_T]_1 = [\hat{\mu}_T]_1,\ [\mu_T]_2 = [\hat{\mu}_T]_2 + P_T$$

we can rewrite the cost by

$$J(u) = \hat{J}(u) + D_1 P_T$$

where

$$\hat{J}(u) = \mathbb{E}\left[\int_s^T u_r^2 dr\right] + D_1[\hat{\mu}_T]_2 + D_2[\hat{\mu}_T]_1^2. \tag{24}$$

- *Step 2:* Since P_T is independent to the control u, to minimize $J(u)$, it is sufficient to minimize $\hat{J}(u)$. Next, we can apply Theorem 3 with parameters

$$A \equiv 0,\ B \equiv 1,\ \sigma_t = \hat{\sigma},\ Q \equiv 1$$

for

$$\hat{V}^* = \inf_{u \in \hat{\mathcal{U}}} \hat{J}(u)$$

with $\hat{J}$ of (24) subject to the process $\hat{X}$ of (21). This yields the Riccati system

$$\begin{aligned}
\phi_1' &= \phi_1^2\\
\phi_2' &= \phi_2^2 + 2\phi_1\phi_2\\
\phi_3' &= -\hat{\sigma}^2\phi_1\\
\phi_1(T) = D_1,\ \phi_2(T) &= D_2,\ \phi_3(T) = 0.
\end{aligned} \tag{25}$$

Now, we summarize the result in the following proposition.

Proposition 5: Suppose $\phi = (\phi_1, \phi_2, \phi_3) \in C^1([0,T], \mathbb{R}^3)$ solves Riccati system (25). Then, the optimal strategy for the control problem (23) is

$$u_t^* = -\phi_1(t)\hat{X}_t^* - \phi_2(t)\mathbb{E}[\hat{X}_t^*] \quad \forall t \in (s,T)$$

and the value is

$$\begin{aligned}
V^* &= \phi_1(s)(x^2 + \hat{\eta}^2 s) + \phi_2(s)x^2\\
&\quad + \phi_3(s) + D_1(\widetilde{\eta}^2 s + \widetilde{\sigma}^2(T-s)).
\end{aligned}$$

Proof: By Theorem 3, the solution of the master equation $\hat{v}^*$ and the optimized controller $\hat{a}^*$ associated to $\hat{J}$ of (24) and the state prediction $\hat{X}$ of (21) are given by

$$\hat{v}^*(t,\hat{\mu}) = \phi_1(t)[\hat{\mu}]_2 + \phi_2(t)[\hat{\mu}]_1^2 + \phi_3(t)$$

and

$$\hat{a}^*(t,\hat{\mu},\hat{x}) = -\phi_1(t)\hat{x} - \phi_2(t)[\hat{\mu}]_1.$$

Moreover, the strategy

$$\begin{aligned}
u_t^* &= \hat{a}^*(t, \hat{\mu}_t, \hat{X}_t^*)\\
&= -\phi_1(t)\hat{X}_t^* - \phi_2(t)\mathbb{E}[\hat{X}_t^*] \quad \forall t \in (s,T)
\end{aligned}$$

makes $\hat{X}^*$ of (21) well defined as a Gaussian process. So, u^* given above is optimal and the corresponding value for (24) is given by $\hat{V}^* = \hat{v}^*(s, \hat{\mu}_s)$, and finally the value of (23) is

$$V^* = \hat{V}^* + D_1 P_T$$

which yields the desired conclusion. ■

C. Two Examples

Example 3 (Linear Terminal Cost in Measure): With $(D_1, D_2) = (1, 0)$, we solve the optimization of (23) defined through partially observed system (17)–(19). Solving the Riccati system (25), we have

$$\begin{aligned}
\phi_1(t) &= \frac{1}{1+T-t}\\
\phi_2 &\equiv 0\\
\phi_3(t) &= \hat{\sigma}^2 \ln(1+T-t).
\end{aligned}$$

Then, the optimal strategy is

$$u_t^* = -\frac{\hat{X}_t^*}{1+T-t} \quad \forall t \in (s,T)$$

and the value is

$$\begin{aligned}
V^* &= \frac{1}{1+T-s}(x^2 + \hat{\eta}^2 s)\\
&\quad + \hat{\sigma}^2 \ln(1+T-s) + \widetilde{\eta}^2 s + \widetilde{\sigma}^2(T-s).
\end{aligned}$$

It is noted that the above value with $s = 0$ is

$$V^*\Big|_{s=0} = \frac{x^2}{1+T} + \hat{\sigma}^2 \ln(1+T) + \widetilde{\sigma}^2 T.$$

Moreover, if $\hat{\sigma} = 1$ and $\widetilde{\sigma} = 0$, then the previous value recovers the solution of fully observable traditional LQG; see Example 1 in Section III-C1.

Example 4 (Quadratic Terminal Cost in Measure): With $(D_1, D_2) = (0, 1)$, we solve the optimization of (23) defined through (17)–(19). Solving the Riccati system (25), we have

$$\begin{aligned}
\phi_2(t) &= \frac{1}{1+T-t}\\
\phi_1 &\equiv 0\\
\phi_3 &\equiv 0.
\end{aligned}$$

Then, the optimal strategy is given by

$$u_t^* = -\frac{\mathbb{E}[\hat{X}_t^*]}{1+T-t} \quad \forall t \in (s,T)$$

and the value is

$$V^* = \frac{x^2}{1+T-s}.$$

Note that the previous value with $s = 0$ and $\hat{\sigma} = \hat{\eta} = 1$ recovers the solution of fully observable mean field LQG; see Example 2 of Section III-C2. Interestingly, the value is invariant with respect to the observability, i.e., $\partial_{\hat{\sigma}} V^* = 0$.

The computations above both agree with our intuition; the value is nonincreasing with respect to $\hat{\sigma}$. Interestingly, as $\hat{\sigma}$ increases, the value is strictly decreasing for Example 3 while stays constant for Example 4. With that being said, observation of the noise does not help in minimization for the proper quadratic terminal cost.

V. Summary

This article focuses on mean-field LQGs with some examples. These simplified frameworks make it possible to obtain some explicit solutions that provide us with valuable insight to a potentially complicated system. For instance, Proposition 5 along with Examples 3 and 4 clearly indicates that the value function of a partially observable system depends not only on the distribution μ_s of the initial state X_s, but also on its joint distribution of $(\hat{X}_s, X_s - \hat{X}_s)$ in the observable probability

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.

space and its orthogonal probability space. Thus, to characterize the value function in the form of $V(t,\mu)$ depending only on the time and initial distribution is not sufficient (cf. [14, (4.7)]).

The result can be extended to multidimensional problems with no essential difficulty but more complex notation. For instance, we consider the process $X_t \in \mathbb{R}^d$ and the cost given by

$$dX_t = (A_t X_t + B_t u_t)dt + \sigma_t dW_t$$

$$J(u) = \mathbb{E}\left[\int_0^T u_t^\top Q_t u_t dt\right] + g(\mu_T)$$

with

$$g(\mu_T) = \int_{\mathbb{R}^d} x^\top D_1 x \mu_T(dx) + [\mu_T]_1^\top D_2 [\mu_T]_1.$$

Solving the master equation yields the following Riccati system:

$$\phi_1'(t) - \phi_1^\top(t) B_t Q_t^{-1} B_t^\top \phi_1(t) + A_t^\top \phi_1(t) + \phi_1(t) A_t = 0$$

$$\phi_2'(t) - \phi_1(t) B_t Q_t^{-1} B_t^\top \phi_2(t) - \phi_2(t) B_t Q_t^{-1} B_t^\top \phi_1(t)$$

$$- \phi_2^\top(t) B_t Q_t^{-1} B_t^\top \phi_2(t) + A_t^\top \phi_2(t) + \phi_2(t) A_t = 0$$

$$\phi_3'(t) + tr[\sigma_t \sigma_t^\top \phi_1(t)] = 0$$

with the terminal condition

$$\phi_1(T) = D_1, \phi_2(T) = D_2, \phi_3(T) = 0$$

where D_1 and D_2 are symmetric matrices.

More challenging generalization is to consider more general cost. For instance, going back to 1-D problems (2)–(4) and (9) with the terminal cost

$$g(\mu_T) = \mathbb{E}[X_T^2] + (\mathbb{E}[\Psi(X_T)])^2$$

one shall solve the master equation with a guess

$$v = \phi_1 \langle \psi, \mu\rangle^2 + \phi_2 \langle x^2, \mu\rangle + \phi_3 + \phi_4 \langle \psi, \mu\rangle \langle x, \mu\rangle + \phi_5 \langle x, \mu\rangle^2.$$

Acknowledgment

The authors are grateful to Prof. Peter Caines for several stimulating discussions on this article during his visit to Hong Kong City University in summer 2019.

References

[1] T. Bjork, M. Khapko, and A. Murgoci, "Time inconsistent stochastic control in continuous time: Theory and examples," *Finance Stochastics*, vol. 21, no. 2, pp. 331–360, 2017.

[2] T. Björk, A. Murgoci, and X. Y. Zhou, "Mean-variance portfolio optimization with state-dependent risk aversion," *Math. Finance*, vol. 24, no. 1, pp. 1–24, 2014.

[3] P. E. Caines, *Linear Stochastic Systems*, vol. 77. Philadelphia, PA, USA: SIAM, 2018.

[4] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. New York, NY, USA: Springer-Verlag, 1975.

[5] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The Master Equation and the Convergence Problem in Mean Field Games*, vol. 201. Princeton, NJ, USA: Princeton Univ. Press, 2019.

[6] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games With Applications I (Probability Theory and Stochastic Modelling 83)*. Cham, Switzerland: Springer, 2018.

[7] V. Dragan, H. Mukaidani, and P. Shi, "The linear quadratic regulator problem for a class of controlled systems modeled by singularly perturbed Itô differential equations," *SIAM J. Control Optim.*, vol. 50, no. 1, pp. 448–470, 2012.

[8] K. Helmes and R. Rishel, "The solution of a partially observed stochastic optimal control problem in terms of predicted miss," *IEEE Trans. Autom. Control*, vol. 37, no. 9, pp. 1462–1464, Sep. 1992.

[9] M. Y. Huang, P. E. Caines, and R. P. Malhamé, "Social optima in mean field LQG control: Centralized and decentralized strategies," *IEEE Trans. Autom. Control*, vol. 57, no. 7, pp. 1736–1751, Jul. 2012.

[10] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[11] S. L. Nguyen and M. Huang, "Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players," *SIAM J. Control Optim.*, vol. 50, no. 5, pp. 2907–2937, 2012.

[12] S. Nguyen, D. Nguyen, and G. Yin, "A stochastic maximum principle for switching diffusions using conditional mean-fields with applications to control problems," *ESAIM: Control, Optim., Calculus Variations*, vol. 26, 2020, Art. no. 69.

[13] H. Pham, "Linear quadratic optimal control of conditional Mckean-Vlasov equation with random coefficients and applications," *Probability, Uncertainty, Quantitative Risk*, vol. 1, no. 1, 2016, Art. no. 7.

[14] H. Pham and X. Wei, "Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics," *SIAM J. Control Optim.*, vol. 55, no. 2, pp. 1069–1101, 2017.

[15] C. Wu and J. Zhang, "Viscosity solutions to parabolic master equations and Mckean-Vlasov SDEs with closed-loop controls," *Ann. Appl. Probability*, vol. 30, no. 2, pp. 936–986, 2019.

[16] J. Yong, "Time-inconsistent optimal control problems and the equilibrium HJB equation," *Math. Control Related Fields*, vol. 2, no. 3, pp. 271–329, 2012.

[17] J. Yong, "Linear-quadratic optimal control problems for mean-field stochastic differential equations," *SIAM J. Control Optim.*, vol. 51, no. 4, pp. 2809–2838, 2013.

[18] J. Yong, "Linear-quadratic optimal control problems for mean-field stochastic differential equations-time-consistent solutions," *Trans. Amer. Math. Soc.*, vol. 369, no. 8, pp. 5467–5523, 2017.

[19] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations* (Applications of Mathematics, 43). New York, NY, USA: Springer-Verlag, 1999.

[20] H. Zhang, "An optimal control problem for linear SDE of mean-field type with terminal constraint and partial information," *Adv. Difference Equ.*, vol. 2019, no. 1, 2019, Art. no. 160.

[21] X. Y. Zhou and G. Yin, "Markowitz mean-variance portfolio selection with regime switching: A continuous-time model," *SIAM J. Control Optim.*, vol. 42, pp. 1466–1482, 2003.

Authorized licensed use limited to: Gordon Library WPI. Downloaded on January 05,2025 at 15:21:07 UTC from IEEE Xplore. Restrictions apply.