Solving A Class of McKean-Vlasov LQG Control Problems with Complete and Partial Observations

Yun Li,* Qingshuo Song,† Fuke Wu‡ George Yin§

Abstract

In this work, we develop an approach for solving a class of linear quadratic Gaussian (LQG) problems, in which the terminal cost is quadratic in the measure space. The motivation for treating such problems stems from a wide variety of applications. The main ingredient of the solution depends on the use of the associated measure flow. Three related problems are considered. The first one is a time-inconsistent control problem. We solve the underlying optimal control problem by use of the corresponding McKean-Vlasov equations and the associated Ricatti systems. The second one is a controlled system under partial observations. Using filtering techniques and a separation principle, we provide solutions of the control problems with partially observable noise. As a ramification of the approach presented in the first two parts, the third problem generalizes the methods of solution of McKean-Vlasov LQG to Kalman-Bucy filtering.

Key Words. controlled diffusion, LQG control, McKean-Vlasov equation, partially observable system, Kalman-Bucy filter.

Mathematics Subject Classification. 60J60, 93E20, 93E11.

*Huazhong University of Science and Technology. li_yun@hust.edu.cn. The research of this author was supported in part by the National Natural Science Foundation of China (Grant No.61873320).

†Worcester Polytechnic Institute, City University of Hong Kong. qsong@wpi.edu. The research of this author was supported in part by the RGC of Hong Kong CityU (11201518).

‡Huazhong University of Science and Technology. wufuke@hust.edu.cn. The research of this author was supported in part by the National Natural Science Foundation of China (Grant No.61873320).

§Wayne State University. gyin@math.wayne.edu. The research of this author was supported in part by the Air Force Office of Scientific Research under grant FA9550-18-1-0268.
1 Introduction

One of the most classical problems in stochastic control is the control of linear quadratic Gaussian regulator problem, in short LQG control problem. Such a problem can be found in any well-known text in control theory; see for example, [4] and [15], among others. In the formulation of such a problem, one aims to minimize an expected quadratic cost function subject to the constraint of a dynamic system given as a linear stochastic differential equation. Because the appealing nature of the simple structure of the problems, the class of problems has been used in a wide range of applications and enjoyed much success in recent years. Some of the recent applications include, for example, multi-agent systems, portfolio optimization, wireless communications, and cyber-physical systems, among others.

In view of the success in the past study and discovery, this work is also devoted to a class of LQG control problems. However, in contrast to the work in the literature, the problem that we consider is not “classical,” in the sense that the terminal cost is quadratic in the distribution of the terminal state $X_T$, i.e., $E[g_1(X_T)]E[g_2(X_T)]$. Such a problem has a wide range of applications. For instance, a specific quadratic terminal cost of the form $R_1E[X_T^2] + R_2(E[X_T])^2 + R_3E[X_T]$ is used to characterize the efficient frontier in the well-known Markowitz mean-variance portfolio optimization; see [1], [8], [16], and references therein. If $R_2 = 0$ in the cost, then it belongs to a standard linear-quadratic-Gaussian (LQG) problem. An explicit solution can be obtained using the solution of the associated Ricatti equation; see [3, 4, 15]. Nevertheless, due to the presence of $(E[X_T])^2$ in the cost, this problem does not satisfy dynamic programming principle (DPP), and the traditional approach using the dynamic programming approach or Hamilton-Jacobi-Bellman (HJB) equation (or Ricatti equation) fails.

As noted in [14], such a problem turns out to belong to the class of time-inconsistent control problems. In the aforementioned reference, Yong gave an explicit solution in Example 1.2, in which the terminal cost is of the form $(E[X_T])^2$. The idea is to compute the value function in an enlarged space by increasing one dimension on the process $t \mapsto E[X_t]$. Alternatively, one may use equivalent game-theoretic framework to solve the problem, which is related to solve the Markowitz portfolio optimization in [1].

Considering the class of time-inconsistent LQG problems, and using the McKean-Vlasov (MV) dynamics as a starting point, this paper focuses on three related problems. The first problem is the aforementioned LQG problem under complete observation; see brief discussions in Section 6.7 of [5]. As alluded to, we no longer have the HJB equation in the usual sense. Unlike [5] using the FBSDE approach, we solve its Master equation (extended HJB on its lifted space) with L-Differentials in a measure space with a judicious guess on the structure of the optimal control. We will derive results leading optimal controls using a Ricatti equation and recover Example 1.2 of [14]. Furthermore, we justify the key feature of LQG control as the first time in the literature, that is, the optimal control and the value function are linear and quadratic in its distribution of the initial state, respectively, and the optimal trajectory is a Gaussian process.

The second problem is concerned with McKean-Vlasov LQG problems in which the driven noise is only partially observable. We show how the optimal control can be obtained by a
A separation principle combined with the approach used in the fully observable system. As a byproduct, it provides the optimal deterministic policy, which answers a question considered by [13]. Moreover, the same approach can be generalized to more complex quadratic terminal cost in the state distribution, which extends the result of MV-LQG problem [10]. The third problem departs from the two problems above by concentrating on a Kalman filtering problem. It should be noted that in all of these problems, instead of starting from a very general setup, the main contribution of the paper is to provide the explicit solution for a class of McKean-Vlasov LQG with or without filtering using its associated Ricatti equations. As alluded to in the last section, this in fact reveals some properties that cannot be seen in the general setup.

The rest of the paper is arranged as follows. To ease the readability of the paper, Section 2 presents the polynomial structure of a measure space and provides a tailor-made verification theorem for MV control problems. Then, we discuss three related problems in the subsequent three sections separately. Section 3 obtains the solution of MV-LQG problem using the Ricatti system. Section 4 provides the solution of MV-LQG under partial observations. Section 5 extends the consideration of MV-LQG to Kalman-Bucy filtering.

Because of our concern is to present the main features of the LQG control problem with the dependence on the associated measure, we decided to use a relative simple setup, namely, the state $x$ is a scalar. All the processes considered are real valued as well. Multidimensional systems can be treated. It does not present essential difficulties. However, the notation will be more complex.

2 Preliminaries

2.1 Polynomial and Derivatives in A Lifted Space

A notion of derivative on the functions of probability measures can be used in a lifted space. The derivative is termed an L-Differential. We briefly recall some properties of L-Differentials and refer more details to the book [5] and the paper [2] for further reading.

Suppose $\mu$ 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 $m$th moment for any $m \geq 1$. If a distribution $\mu$ has a finite $m$th moment $[\mu]_m$, then we write it as $\mu \in \mathcal{P}_m$. For instance, a Dirac measure $\delta_x$ for any $x \in \mathbb{R}$ belongs to $\mathcal{P}_m$ for any $m \geq 1$.

The functions under consideration have the form $f : \mathcal{P}_2(\mathbb{R}) \mapsto \mathbb{R}$. An example of such a
function is, for some $n \in \mathbb{N}$,
\[
f(\mu) = [\mu]_2^n, \quad \forall \mu \in \mathcal{P}_2, x \in \mathbb{R}.
\]
(1)
The function $f$ above is indeed a polynomial of degree $n$. The partial derivative of the function $f$ above in the variable $\mu$, computed in the lifted space, becomes a polynomial of degree $n - 1$ in that
\[
\partial_\mu f(\mu)(x) = 2n[\mu]_2^{n-1}x.
\]
In what follows, we provide a brief explanation of the above computations of the derivatives in the lifted space. First, let us define polynomials in the variable $\mu$.

• Suppose that $p_1(\mu)$ is given in the form of
\[
p_1(\mu) = \langle \phi, \mu \rangle + c
\]
for a constant $c$ and for some function $\phi : \mathbb{R} \mapsto \mathbb{R}$ such that $\phi$ belongs to $C^2_2(\mathbb{R})$. \(^1\) Then $p_1$ is defined as a polynomial of degree one in $\mu$ and the partial derivative in $\mu$ is
\[
\partial_\mu p_1(\mu, x) = \phi'(x).
\]

• Now, we can define polynomials of any degree inductively. Given an $n$-degree polynomial $p_n(\mu)$ and 1-degree polynomial $p_1(\mu)$, the product $p_{n+1} = p_np_1$ is said to be an $(n+1)$-degree polynomial. By the product rule, one can write
\[
\partial_\mu p_{n+1}(\mu, x) = \partial_\mu p_n \cdot p_1 + p_n \cdot \partial_\mu p_1.
\]
In particular, the $n$-degree polynomial $p_1^n(\mu) = (\phi, \mu)^n$ has its $\mu$-derivative as
\[
\partial_\mu p_1^n(\mu)(x) = np_1^{n-1}(\mu)\phi'(x).
\]

By the above definition, any function of the form $f(\mu) = \prod_{i=1}^n \langle \phi_i, \mu \rangle$ is a polynomial of degree $n$ in $\mu$, where $\phi_i \in C^2_2(\mathbb{R})$.

### 2.2 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 a $\mathbb{F}$-adapted Brownian motion $W$. Given a controlled SDE
\[
dX_t = b(t, X_t, \rho_t)dt + \sigma_t dW_t,
\]
(2)
we denote by $\mu_t$ the probability law of $X_t$ and consider the cost function
\[
J(x, \rho(\cdot)) = \mathbb{E}^x \left[ \int_0^T \ell(t, X_t, \rho_t)dt \right] + g(\mu_T).
\]
(3)
\(^1\)We use $C^k_p(\mathbb{R})$ to denote the collection of functions $f$ with continuous $k$th derivative and $p$th-order polynomial growth in value. Clearly, $\langle f, \mu \rangle$ is well defined if $f \in C^2_2(\mathbb{R})$ and $\mu \in \mathcal{P}_2(\mathbb{R})$. 

4
In the above, $\rho(\cdot)$ is the control process, $\ell(\cdot, \cdot, \cdot)$ is the running cost function, and $g(\cdot)$ is the terminal cost. Our objective is to minimize the above cost function among all adapted square integrable control $\rho_t$. By a solution, we mean an optimal pair $(\rho^*, v)$ satisfying

$$v(x) = J(x, \rho^*) \leq J(x, \rho), \forall \rho \in L^2_x.$$ (4)

The process $\rho^*$ and the $v$, if they exist, are referred to as the optimal control and the optimal value, respectively.

To proceed, we obtain the verification theorem in terms of the following master equation:

$$\inf_{a \in C^2_{\mathbb{R}}} \{ \langle b(t, \cdot, a)\partial_\mu u(t, \mu)(\cdot) + \ell(t, \cdot, a), \mu \rangle + \frac{1}{2}\sigma_t^2 \langle \partial_x u(t, \mu), \mu \rangle + \partial_t u(t, \mu) = 0 \}$$ (5)

with terminal condition

$$u(T, \mu) = g(\mu).$$ (6)

The solution of the above master equation is the pair $u : (0, T) \times \mathcal{P}_2 \mapsto \mathbb{R}$ and $a^* : (0, T) \times \mathcal{P}_2 \mapsto C^2_{\mathbb{R}}$ satisfying (5)-(6) together with

$$\langle b(t, \cdot, a^*(t, \mu, \cdot))\partial_\mu u(t, \mu)(\cdot) + \ell(t, \cdot, a^*(t, \mu, \cdot)), \mu \rangle = \inf_{a \in C^2_{\mathbb{R}}} \{ \langle b(t, \cdot, a)\partial_\mu u(t, \mu)(\cdot) + \ell(t, \cdot, a), \mu \rangle \}.$$

In what follows, we shall use the convention $a^*(t, \mu)(x) = a^*(t, \mu, x)$, similarly, $\partial_\mu u(t, \mu)(x) = \partial_\mu u(t, \mu, x)$. For later use, we remark the difference on the range of infimum over $a$ using the following example. Let $f$ be $f(x, a) = a^2 - 2ax$. Then,

$$\inf_{a \in C^2_{\mathbb{R}}} \langle f(\cdot, a), \mu \rangle = -[\mu]_2, \text{ with } a^*(x) = x,$$

and

$$\inf_{a \in \mathbb{R}} \langle f(\cdot, a), \mu \rangle = -[\mu]_1^2, \text{ with } a^* = [\mu]_1.$$

**Proposition 1** If there exists a pair $u : (0, T) \times \mathcal{P}_2 \mapsto \mathbb{R}$ and $a^* : (0, T) \times \mathcal{P}_2 \mapsto C^2_{\mathbb{R}}$ satisfying the master equation (5)-(6), then the optimal value is

$$v(x) = u(0, \delta_x)$$

and the optimal control is

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

provided that $v(x)$ under the control $\rho$ is well defined and $u$ is sufficiently smooth.

**Remark 2** To ensure $v$ under $\rho^*$ being well defined and sufficiently smooth, one may impose assumptions (like Lipschitz continuity and linear growth) on the functions $(b, \sigma, \ell, g)$ and $a^*$. To better focus on LQG problems, we do not pursue further sufficient conditions for the assumptions of Proposition 1. Indeed, one can easily check that all examples in this paper satisfy these assumptions. □
This optimal control is indeed a Markovian strategy in an extended sense since it is a function of \((t, \mu_t, X_t)\). In this case, (5)-(6) can be explicitly solved. The main ingredient is the Itô formula given by Theorem 5.9.2 of [5]. For a sufficiently smooth function \(f\),

\[
f(\mu_t) = f(\mu_0) + \int_0^t \mathbb{E}[(\partial_x f(\mu_s)(X_s)b(s, X_s, \rho_s)]ds + \frac{1}{2} \int_0^t \mathbb{E}[\sigma_x^2 \partial_x f(\mu_s)(X_s)]ds.
\]

**Proof:** (Proposition 1) By the Itô formula, for any control \(\rho\), we have

\[
u(t, \mu_t) = g(\mu_T) - \int_t^T \partial_t u(s, \mu_s)ds
\]

\[-\int_t^T \mathbb{E}[\partial_x u(s, \mu_s)(X_s)b(s, X_s, \rho_s)]ds
\]

\[-\frac{1}{2} \int_t^T \sigma_x^2 \mathbb{E}[(\partial_x b(s, \mu_s)(X_s))]ds
\]

\[\leq g(\mu_T) + \left[\int_t^T \mathbb{E} \ell(s, X_s, \rho_s)ds\right].
\]

This implies that

\[
u(0, \mu_0) \leq g(\mu_T) + \left[\int_0^T \mathbb{E} \ell(t, X_t, \rho_t)dt\right]
\]

for any control \(\rho\) and initial distribution \(\mu_0\). The result follows if we replace \(\mu_0\) by \(\delta_x\). \(\square\)

### 3 MV-LQG: Fully Observable Case

#### 3.1 Setup

For simplicity, we consider MV-LQG problem with linear function \(b\) and quadratic running cost function \(\ell\) and terminal cost \(g\) so that in general setup (2), (3), and (4), the coefficients or the functions become

\[
b(t, x, a) = A_t x + C_t a, \quad \ell(t, x, a) = Q_t a^2,
\]

and

\[
g(\mu_T) = R_1 [\mu_T]_2 + R_2 [\mu_T]^2_1 = R_1 \mathbb{E}[X_T^2] + R_2 (\mathbb{E}[X_T])^2.
\]

Note that the terminal cost only depends on the probability distribution \(\mu_T = \mathbb{P}(X_T)\). Furthermore, the function \(g\) is a quadratic function on \(\mu_T\).

#### 3.2 Examples

**Example 1** (A standard LQG.) If

\[
A \equiv 0, C \equiv 1, \sigma \equiv 1, Q \equiv 1, R_1 = 1, R_2 = 0,
\]

(9)
then the problem is a standard LQG problem. The terminal cost is \( g(\mu_T) = \mathbb{E}[X_T^2] \) (a linear function in \( \mu_T \)), which can be replaced by a quadratic function in \( X_T \) of the form

\[
g(X_T) = \mathbb{E}X_T^2.
\]

In this case, the dynamic programming principle is applicable and one can write its HJB as

\[
\inf_{a \in \mathbb{R}} \{ b(t, x, a) \partial_x u(t, x) + \ell(t, x, a) \} + \frac{1}{2} \sigma_t^2 \partial_{xx} u(t, x) + \partial_t u(t, x) = 0, \forall (t, x) \in (0, T) \times \mathbb{R},
\]

with its terminal condition

\[
u(T, x) = g(x), \forall x \in \mathbb{R}.
\]

The solution of the above HJB is the pair \((u, a^*) : (0, T) \times \mathbb{R} \mapsto \mathbb{R}^2\) satisfying (10)-(11) and

\[
b(t, x, a^*(t, x)) \partial_x u(t, x) + \ell(t, x, a^*(t, x)) = \inf_{a \in \mathbb{R}} \{ b(t, x, a) \partial_x u(t, x) + \ell(t, x, a) \}
\]

for all \((t, x) \in (0, T) \times \mathbb{R}\). Then, the optimal value corresponds to

\[
v(x) = u(0, x)
\]

and the optimal control is of the form

\[
r_t = a^*(t, X_t).
\]

This optimal control is termed a Markovian strategy since it is a function of \((t, X_t)\). Equation (10)-(11) can be explicitly solved. □

**Example 2** This problem is taken from [14]. Let

\[
A \equiv 0, C \equiv 1, \sigma \equiv 1, Q \equiv 1, R_2 = 1, R_1 = 0.
\]

In contrast to Example 1, the terminal cost is not a liner function but a proper quadratic function in \( \mu_T \). This problem has been studied in the context of time inconsistent control problems, and the dynamic programming does not apply and the HJB does not hold in general. □

### 3.3 Semi-Explicit Solution in Terms of Ricatti Equations

In this section, we solve explicitly the MV-LQG problem with parameters given by (7)-(8) by treating (5)-(6). We impose the following assumptions.

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

With parameters given by (7) in equation (5) note that the inimum is quadratic in \( a \), which can be rewritten as

\[
\langle b(t, \cdot, a) \partial_{\mu} u(t, \mu, \cdot) + \ell(t, \cdot, a), \mu \rangle = \int_{\mathbb{R}} ((A_t x + C_t a) \partial_{\mu} u(t, \mu, x) + Q_t a^2) \mu(dx)
\]
Since $\mu$ is a non-negative measure and $Q_t > 0$, the infimum over $C^2_2(\mathbb{R})$ is attained at

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

with its value

$$\inf_{a \in C^2(\mathbb{R})} \{\langle b(t, \cdot, a) \partial_\mu u(t, \mu, \cdot) + \ell(t, \cdot, a), \mu \rangle \} = \left\langle A_t x \partial_\mu u - \frac{C_t^2}{4Q_t} |\partial_\mu u|^2, \mu \right\rangle(t, \mu, x)$$

provided that $\partial_\mu u(t, \mu) \in C^2_2(\mathbb{R})$ for all $\mu \in \mathcal{P}_2(\mathbb{R})$ which confirms later in (14). Therefore, the master equation (5) becomes

$$\langle L_0 u(t, \mu, \cdot), \mu \rangle + \partial_t u(t, \mu) = 0, \quad \text{(13)}$$

where the operator $L_0$ is defined by

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

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

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

Then we use the method of undetermined “coefficients” to determine $\phi = (\phi_1, \phi_2, \phi_3)$. One can directly write the derivative as

$$\partial_\mu u(t, \mu, x) = 2\phi_1(t)x + 2\phi_2(t)[\mu]_1, \quad \text{(14)}$$

which is a polynomial in $x$ and confirms $\partial_\mu u(t, \mu) \in C^2_2(\mathbb{R})$. Moreover, we have

$$\partial_t u(t, \mu, x) = \phi'_1(t)[\mu]_2 + \phi'_2(t)[\mu]_1^2 + \phi'_3(t),$$

and

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

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

$$0 = [\mu]_2 L_1 \phi(t) + [\mu]_1^2 L_2 \phi(t) + L_3 \phi(t), \quad \text{(15)}$$

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

$$L_1 \phi(t) = \phi'_1(t) - \frac{C_t^2}{Q_t^2} \phi^2_1(t) + 2A_1 \phi_1(t),$$
$$L_2 \phi(t) = \phi'_2(t) - \frac{C_t^2}{Q_t^2} \phi^2_2(t) - \frac{2C_t^2}{Q_t^2} \phi_1(t) \phi_2(t) + 2A_t \phi_2(t),$$
$$L_3 \phi(t) = \phi'_3(t) + \sigma_t^2 \phi_1(t). \quad \text{(16)}$$

Since (15) holds for all $\mu$, together with terminal condition, we have the following system of ODEs with the first order differential operator $L$ of (16)

$$L \phi(t) = 0, \forall t \in (0, T), \quad \text{with } \phi(T) = (R_1, R_2, 0). \quad \text{(17)}$$

Note that each equation is a linear combination of $\phi$’s and quadratic functions in $\phi$. Such a system is called Ricatti equations. It is straightforward to carry out the derivation using verification theorem in a rigorous way and conclude the following result.
Theorem 3 Suppose \((A1)\) holds and there exists \(\phi \in C^1((0, T), \mathbb{R}^3)\) for Ricatti system (17). Then the pair \((u, a^*)\) given by

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

and

\[
a^*(t, \mu, x) = -\frac{C_t}{Q_t}(\phi_1(t)x + \phi_2(t)[\mu]_1)
\]
solves the master equation (5)-(6). Moreover, if \(X^x,\rho\) and \(v(x)\) with parameter sets (7)-(8) and the strategy

\[
\rho_t = a^*(t, \mu_t, X_t)
\]

are well defined, then \(\rho\) is the optimal control and \(v(x) = u(0, \delta_x)\).

3.3.1 Explicit Solution for Example 1: Traditional LQG

We give explicit solution for Example 1. Recall that the parameter sets are

\[
A \equiv 0, C \equiv 1, \sigma \equiv 1, Q \equiv 1, R_1 = 1, R_2 = 0.
\]

In this case, the Ricatti system (17) becomes

\[
\begin{align*}
\phi_1' &= \phi_2^2, \\
\phi_2' &= \phi_3^2 + 2\phi_1\phi_2, \\
\phi_3' &= -\phi_1,
\end{align*}
\]

with terminal condition

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

The solution for this Ricatti system can be written by, for all \(t \in (0, T)\)

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

By Theorem 3, one can verify that the optimal strategy is a traditional Markovian control of the form

\[
\rho_t = -\frac{X_t}{1 + T - t}
\]

and the optimal value is

\[
v(x) = \frac{x^2}{1 + T} + \ln(1 + T).
\]

3.3.2 Explicit Solution for Example 2: Time Inconsistent Control

We give explicit solution for Example 2. Recall that the parameter sets are

\[
A \equiv 0, C \equiv 1, \sigma \equiv 1, Q \equiv 1, R_1 = 1, R_2 = 0.
\]
In this case, the Ricatti system \((17)\) becomes the same as \((18)\) but with different terminal conditions

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

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

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

By Theorem 3, one can verify that the optimal strategy is

\[
\rho_t = -\frac{\mathbb{E}[X_t]}{1 + T - t}
\]

and the value function is

\[
u(t, \mu) = \frac{1}{1 + T - t} \left( \int x \mu(dx) \right)^2,
\]

which implies the optimal value

\[
v(x) = \frac{x^2}{1 + T}.
\]

Note that the above optimal strategy is Markovian only in the extended sense due to its dependence on \(\mathbb{E}[X_t]\). In fact, one can verify that any traditional form of Markovian strategy as a function of \((t, X_t)\) cannot be optimal by Jensen’s inequality.

4 MV-LQG: A Controlled System under Partial Observations

The following interesting question considered in \([13]\) motivates our second problem:

- 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 \(\rho\) 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 this below, the underlying problem ensembles that of Section 3 in that it preserves the structure of MV LQG. The difference is that the Brownian motion \(\tilde{W}_t = \sigma W_t + \sqrt{1 - \sigma^2} B_t\) is partially observable via \(W_t\). Note that the deterministic control problem raised by [13] can be recovered by the case of \(\sigma = 0\), while the fully observable control problem (7) - (8) corresponds to the case with \(\sigma = 1\).
4.1 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 $\mathbb{F}$-adapted Brownian motions $W$ and $B$, respectively. For simplicity, we assume $\mathbb{F} = \mathbb{F}^W \times \mathbb{F}^B$ and $\mathcal{F} = \mathcal{F}^W \times \mathcal{F}^B$, where $\mathbb{F}^W = (\mathcal{F}^W_t)_{0 \leq t \leq T}$ and $\mathbb{F}^B = (\mathcal{F}^B_t)_{0 \leq t \leq T}$ are the filtrations generated by $W$ and $B$, respectively.

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

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

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

$$X_t = X_s + \int_s^t \rho_r dr + \int_s^t \sigma dW_r + \int_s^t \dot{\sigma} dB_r.$$  

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

$$X_s = x + \eta W_s + \dot{\eta} B_s.$$  

The cost functional to be minimized is given by

$$J(x, \rho) = \mathbb{E} \left[ \int_s^T \rho_r^2 dr \right] + R_1[\mu_T]^2 + R_2[\mu_T]^2.$$  

The distinction of the current problem compared with the previous control problem is the following crucial point. Though the player wants to minimize the cost functional, he or she can only observe his or her own generic noise $W$, but not the common noise $B$. In other words, the optimal value is defined as

$$v(x) = \inf_{\rho \in L^2_\mathbb{F}} J(x, \rho).$$  

In the above, we abbreviate the dependence of $v(x)$ on starting time $s$ for simplicity.

4.2 Semi-Explicit Solution: Separation Principle

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

- Step 1: Let $\hat{X}$ be the prediction of $X$. That is,

$$\hat{X}_t = \mathbb{E}[X_t|\mathcal{F}^W_t],$$

and $\mathcal{E}$ and $P$ are

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

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

$$\dot{\hat{X}}_t = x + \eta W_s + \int_s^t \rho_r dr + \int_s^t \sigma dW_r.$$  

11
\[ E_t = \hat{\eta} B_s + \hat{\sigma} (B_t - B_s), \]

and

\[ P_t = \hat{\eta}^2 s + \hat{\sigma}^2 (t - s). \]

Let us use \( \hat{\mu}_t \) to denote the distribution of \( \hat{X}_t \). Owing to

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

we can rewrite the cost by

\[ J(x, \rho) = \hat{J}(x, \rho) + R_1 P_T, \]

where

\[ \hat{J}(x, \rho) = \mathbb{E} \left[ \int_s^T \rho_r^2 dr \right] + R_1 [\hat{\mu}_T]_2 + R_2 [\hat{\mu}_T]^2. \]  

(22)

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

\[ A \equiv 0, \quad C \equiv 1, \quad \sigma_t = \sigma, \quad Q \equiv 1 \]

for

\[ \hat{v}(x) = \inf_{\rho \in L^2} \hat{J}(x, \rho) \]

with \( \hat{J} \) of (22) subject to the process \( \hat{X} \) of (21). This yields the Ricatti system

\[ \begin{align*}
\phi'_1 &= \phi_1^2, \\
\phi'_2 &= \phi_2^2 + 2\phi_1 \phi_2, \\
\phi'_3 &= -\sigma^2 \phi_1,
\end{align*} \]

(23)

Now we summarize the result in the following proposition.

**Proposition 4** Suppose \( \phi = (\phi_1, \phi_2, \phi_3) \in C^1([0, T], \mathbb{R}^3) \) solves Ricatti system (23). Then, the optimal strategy for the control problem (20) is

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

and the value is

\[ v(x) = \phi_1(s)(x^2 + \eta^2 s) + \phi_2(s)x^2 + \phi_3(s) + R_1(\eta^2 s + \sigma^2(T - s)). \]

**Proof:** By Theorem 3, the pair \( (\hat{u}, \hat{a}^*) \) given by

\[ \hat{u}(t, \mu) = \phi_1(t)[\mu]_2 + \phi_2(t)[\mu]^2 + \phi_3(t) \]

and

\[ \hat{a}^*(t, \mu, x) = -\phi_1(t)x - \phi_2(t)[\mu]_1 \]
solves the master equation corresponding to \((\hat{X}, \hat{v})\). Moreover, the strategy
\[
\rho_t = \hat{a}^*(t, \hat{\mu}_t, \hat{X}_t) = -\phi_1(t) \hat{X}_t - \phi_2(t) \mathbb{E}[\hat{X}_t], \quad \forall t \in (s, T)
\]
makes \(\hat{X}\) of (21) well defined as an OU process. Therefore, \(\rho\) given above is optimal and the corresponding value for (22) is given by
\[
v(x) = \hat{u}(s, \hat{\mu}_s) + R_1 P_T = \phi_1(s)(x^2 + \eta^2 s) + \phi_2(s)x^2 + \phi_3(s) + R_1(\eta^2 s + \hat{\sigma}^2(T - s)).
\]

It is worthwhile to mention that for Proposition 4, the value function \(v\) of (20) may vary for the same initial distributions according to the dependency on \(F^{\mathbb{W}}\). In other words, there is no function \(\tilde{v}\) such that \(v = \tilde{v}(s, \mu_s)\).

### 4.3 Two Examples

We consider two cases with
- \((R_1, R_2) = (1, 0)\): linear terminal cost in measure
- and \((R_1, R_2) = (0, 1)\): quadratic terminal cost in measure.

As a result, it agrees with our intuition that the value is non-increasing with respect to \(\sigma\). Interestingly, as \(\sigma\) is increasing, the value is strictly decreasing for \((R_1, R_2) = (1, 0)\), while stays constant for \((R_1, R_2) = (0, 1)\). With that being said, observation of the noise does not help in minimization for the proper quadratic terminal cost.

#### 4.3.1 When \(R_1 = 1\) and \(R_2 = 0\)

In this part, we compute the control problem (20), when \(R_1 = 1\) and \(R_2 = 0\). Solving the Ricatti system (23), we have
\[
\phi_1(t) = \frac{1}{1 + T - t}, \quad \phi_2 \equiv 0, \quad \phi_3(t) = \sigma^2 \ln(1 + T - t).
\]
Then, the optimal strategy is
\[
\rho_t = -\frac{\hat{X}_t}{1 + T - t}, \quad \forall t \in (s, T)
\]
and the value is
\[
v(x) = \frac{1}{1 + T - s}(x^2 + \eta^2 s) + \sigma^2 \ln(1 + T - s) + \hat{\eta}^2 s + \hat{\sigma}^2(T - s).
\]
It is noted that the above value with \(s = 0\) is
\[
v(x) = \frac{x^2}{1 + T} + \sigma^2 \ln(1 + T),
\]
which recovers the solution of Example 1 in Section 3.3.1 for \(\sigma = 1\).
4.3.2 When $R_1 = 0$ and $R_2 = 1$

In this part, we consider control problem (20) when $R_1 = 0$ and $R_2 = 1$. Solving the Ricatti system (23), we have

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

Then, the optimal strategy is given by

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

and the value is

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

It is noted that, the above value with $s = 0$ and $\sigma = \eta = 1$ recovers the solution of Example 2; see Section 3.3.2.

5 Ramification of MV-LQG: Kalman-Bucy Filtering

Let $X$ be the underlying process of

$$dX_t = A_t X_t dt + C_t \rho_t dt + \sigma_t dB_t$$

and $Y$ be the observed process of

$$dY_t = H_t X_t dt + K_t dW_t.$$  

Our objective is to minimize the cost

$$J(x, \rho) = E[\int_0^T Q_t \rho_t^2 dt] + R_1[\mu_T]_2 + R_2[\mu_T]_1^2$$

over all $\mathbb{F}(Y)$-adapted processes. Let us denote the value as

$$v(x) = \inf_{\rho \in L^2(\mathbb{F}(Y))} J(x, \rho).$$

We assume

$$K_t > 0, Q_t > 0, \forall t \geq 0.$$  

Let

$$\hat{X}_t = E[X_t|\mathcal{F}_t(Y)], \ E_t = X_t - \hat{X}_t, P_t = E[\mathcal{E}_t].$$

By Kalman-Bucy filter (see [12]), we have

$$d\hat{X}_t = A_t \hat{X}_t dt + C_t \rho_t dt + P_t K_t^{-1} H_t dW_t,$$

$$\frac{dP_t}{dt} = 2A_t P_t - P_t^2 H_t^2 K_t^{-2} + \sigma_t^2,$$
where $\hat{W}$ is an $\mathbb{F}(Y)$-Brownian motion, which is usually called an innovation process in the filtering theory given by

$$d\hat{W}_t = K_t^{-1}(dY_t - H_t\hat{X}_t dt).$$

Next, we can rewrite the cost by

$$J(x, \rho) = \hat{J}(x, \rho) + R_1 P_T,$$

where

$$\hat{J}(x, \rho) = \mathbb{E}\left[ \int_0^T Q_r \rho_r^2 dr \right] + R_1 [\hat{\mu}_T]_2 + R_2 [\hat{\mu}_T]_1^2. \quad (25)$$

Now, we only need to minimize $\hat{J}$ of (25) based on $\hat{X}$, which is fully observable control problem.

6 Summary

With the terminal cost $(\mathbb{E}[X_T])^2$ on the fully observable system, we list the following notable features extracted from Theorem 3 and Section 3.3.2, which justifies MV-LQG in our paper in contrast to the classical LQG problems.

1. The optimal control has a linear (feedback) form with respect to its measure $\mu_t$. It is well known that a time-inconsistent problem has no feedback (in the state variable) form for its optimal control. However, the optimal control process is a feedback form via probability measure $\mu_t$ given by

$$\rho^*(t) = -\frac{\mathbb{E}[X_t]}{1 + T - t} = -\frac{1}{1 + T - t} \int x\mu_t(dx).$$

2. The value function is a quadratic function in its initial probability measure. The optimal value $v(x)$ is given by $u(0, \delta_x)$, where the function $u(t, \mu)$ is the solution of its master equation explicitly given by

$$u(t, \mu) = \frac{1}{1 + T - t} \left( \int x\mu(dx) \right)^2.$$

3. The optimal trajectory turns out to be a Gaussian process.

In lieu of $(\mathbb{E}[X_T])^2$ considered in this paper, it is conceivable with more detailed calculations, all the above features are preserved for the quadratic terminal cost of the form $\mathbb{E}[g_1(X_T)]\mathbb{E}[g_2(X_T)]$ for sufficiently smooth $g_1$ and $g_2$. The main approach is that after converting the underlying problem to MV control problem, the key step is to obtain Ricatti equation from the master equation by taking the advantage of the quadratic cost structure in the measure space; see Theorem 3.

For the partially observable MV control, we summarize the following features by Proposition 4 and two related examples:
1. The optimal control/value/trajectory depends on the distribution of the prediction process.

2. Suppose the initial state is given by a random variable $X_s$ at initial time $s$, the value function cannot be simply represented by a function of the initial distribution of $X_s$, but is a function of joint distribution $(X_s - \hat{X}_s, \hat{X}_s)$.

3. Deterministic control problem yields its optimal value no less than fully observable counterpart. However, they are equal with the proper quadratic terminal cost $(\mathbb{E}[X_T])^2$.

Moreover, if the separation principle works out, all the above features can be extended to the more complex quadratic terminal cost of the form $\mathbb{E}[g_1(X_T, \hat{X}_T)]\mathbb{E}[g_2(X_T, \hat{X}_T)]$ rather than $\mathbb{E}[g_1(X_T, \hat{X}_T)]$ for sufficiently smooth functions $g_1$ and $g_2$, which can be seen as the extension of [10].

This paper only develops one dimensional MV-LQG optimal control and filtering problems. The result can be extended to multidimensional problems with no essentially technical difficulty but much more complex notation. Moreover, for simplicity, the parameters we use for the LQG problem is chosen specifically as (7) and (8). However, one can also extend it to more general LQG form, for instance, the terminal cost could be set as $g(\mu) = \langle g_1, \mu \rangle \langle g_2, \mu \rangle$ with appropriate $g_1$ and $g_2$. Analogous to the traditional LQG, one may also attempt MV-LQG with bounded domain or infinite time horizon with a proper discount.

The solution of time-inconsistent problem may be defined differently using Nash equilibrium, which is termed as time-consistent solution; see [14]. The current paper covers only time-inconsistent solution and similar approach in time-consistent solution will be explored in our future work.

References

[1] Tomas Björk, Agatha Murgoci, and Xun Yu Zhou. Mean–variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 24(1):1–24, 2014.

[2] Rainer Buckdahn, Juan Li, Shige Peng, Catherine Rainer, et al. Mean-field stochastic differential equations and associated pdes. *The Annals of Probability*, 45(2):824–878, 2017.

[3] Peter E Caines. *Linear stochastic systems*, volume 77. SIAM, 2018.

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

[5] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field FBSDEs, control, and games.
[6] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field games with common noise and master equations.

[7] M. Huang. Large-population lqg games involving a major player: the nash certainty equivalence principle. *SIAM J Control Optim*, 48:3318–3353, 2010.

[8] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[9] Son Luu Nguyen and Minyi Huang. Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players. *SIAM J. Control Optim.*, 50(5):2907–2937, 2012.

[10] Huyën Pham. Linear quadratic optimal control of conditional mckean-vlasov equation with random coefficients and applications. *Probability, Uncertainty and Quantitative Risk*, 1(1):7, 2016.

[11] Huyën Pham and Xiaoli Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55(2):1069–1101, 2017.

[12] Ramon Van Handel. Stochastic calculus, filtering, and stochastic control. *Course notes.*, URL http://www.princeton.edu/rvan/acm217/ACM217.pdf, 2007.

[13] Cong Wu and Jianfeng Zhang. Viscosity solutions to parabolic master equations and Mckean-Vlasov SDEs with closed-loop controls. *arXiv*, 2019.

[14] Jiongmin Yong. Linear-quadratic optimal control problems for mean-field stochastic differential equations — time-consistent solutions. *Transactions of the American Mathematical Society*, 2015.

[15] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.

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