ESAIM: Control, Optimisation and Calculus of Variations Will be set by the publisher
URL: http://www.emath.fr/cocv/

STOCHASTIC LINEAR QUADRATIC STACKELBERG DIFFERENTIAL GAME WITH OVERLAPPING INFORMATION *

Jingtao Shi[1], Guangchen Wang[2] and Jie Xiong[3]

Abstract. This paper is concerned with the stochastic linear quadratic Stackelberg differential game with overlapping information, where the diffusion terms contain the control and state variables. Here the term "overlapping" means that there are common part between the follower's and the leader's information, while they have no inclusion relation. Optimal controls of the follower and the leader are obtained by the stochastic maximum principle, the direct calculation of the derivative of the cost functional and stochastic filtering. A new system of Riccati equations is introduced to give the state estimate feedback representation of the Stackelberg equilibrium strategy, while its solvability is a rather difficult open problem. A special case is then studied and is applied to the continuous-time principal-agent problem.

1991 Mathematics Subject Classification. 60H10, 91A23, 93E20, 93E11.

The dates will be set by the publisher.

1. Introduction

Throughout this paper, we denote by $\mathbb{R}^n$ the Euclidean space of n-dimensional vectors, by $\mathbb{R}^{n\times d}$ the space of $n\times d$ matrices, by $\mathcal{S}^n$ the space of $n\times n$ symmetric matrices. $\langle\cdot,\cdot\rangle$ and $|\cdot|$ denote the scalar product and norm in the Euclidean space, respectively. $\top$ appearing in the superscripts denotes the transpose of a matrix. f_x, f_{xx} denote the partial derivative and twice partial derivative with respect to x for a differentiable function f, respectively.

1.1. Motivation

First, we present the following example which motivates us to study the problem in this paper.

Keywords and phrases: Stackelberg differential game, stochastic linear quadratic optimal control, overlapping information, maximum principle, stochastic filtering

* *This work is financially supported by the National Key R&D Program of China (Grant No. 2018YFB1305400), the National Natural Science Funds of China (Grant No. 11971266, 11831010, 11571205, 61821004, 61873325, 61925306), and the Southern University of Science and Technology Start-Up Fund (Grant No. Y01286220).*

[1] School of Mathematics, Shandong University, Jinan 250100, P. R. China, E-mail: shijingtao@sdu.edu.cn

[2] Corresponding author. School of Control Science and Engineering, Shandong University, Jinan 250061, P. R. China, E-mail: wguangchen@sdu.edu.cn

[3] Department of Mathematics and SUSTech International Center for Mathematics, Southern University of Science and Technology, Shenzhen 518055, P. R. China, E-mail: xiongj@sustc.edu.cn

Example 1.1: (Continuous time principal-agent problem) The principal contracts with the agent to manage a production process, whose output $Y(\cdot)$ evolves as

$$\begin{cases} dY(t) = Be(t)dt + \sigma_1 dW_1(t) + \sigma_2 dW_2(t) + \sigma_3 dW_3(t), \ t \in [0,T], \\ Y(0) = Y_0 \in \mathbb{R}, \end{cases} \tag{1}$$

where $e(\cdot) \in A \subset \mathbb{R}$ is the agent's effort choice, B represents the productivity of effort, and there are three additive shocks (due to the three independent Brownian motions $W_1(\cdot), W_2(\cdot), W_3(\cdot)$) to the output. The output of the production adds to the principal's asset $y(\cdot)$, which earns a risk free return r, and out of which he pays the agent $s(\cdot) \in S \subset \mathbb{R}$ and withdraws his own consumption $d(\cdot) \in \mathbb{R}$. Thus the principal's asset evolves as

$$\begin{cases} dy(t) = \big[ry(t) + Be(t) - s(t) - d(t)\big]dt + \sigma_1 dW_1(t) + \sigma_2 dW_2(t) + \sigma_3 dW_3(t), \ t \in [0,T], \\ y(0) = y_0 \in \mathbb{R}, \end{cases} \tag{2}$$

where y_0 is the initial asset. In addition, the agent has his own wealth $m(\cdot)$, out of which he consumes $c(\cdot)$, then

$$\begin{cases} dm(t) = \big[rm(t) + s(t) - c(t)\big]dt + \bar{\sigma}_1 dW_1(t) + \bar{\sigma}_2 dW_2(t) + \bar{\sigma}_3 dW_3(t), \ t \in [0,T], \\ m(0) = m_0 \in \mathbb{R}. \end{cases} \tag{3}$$

The agent earns the same rate of return r on his savings, gets income flows due to his payment $s(\cdot)$, and draws down wealth to consume. Here $\sigma_i, \bar{\sigma}_i, i = 1,2,3$ are all constants. At the terminal time T, the principal makes a final payment $s(T)$ and the agent chooses consumption based on this payment and his terminal wealth $m(T)$.

We consider an optimal contract problem in the following information structure which is different from those in Williams [30, 31]. In this problem, the principal can observe his asset $y(\cdot)$ and the agent's initial wealth m_0, but cannot monitor the agent's effort $e(\cdot)$, consumption $c(\cdot)$ and wealth $m(\cdot)$. The principal must provide incentives for the agent to put forth the desired amount of the effort. For any $s(\cdot), d(\cdot)$, the agent first chooses his effort $e^*(\cdot)$ and consumption $c^*(\cdot)$ such that his preference

$$J_1\big(e(\cdot), c(\cdot), s(\cdot), d(\cdot)\big) = \frac{1}{2}\mathbb{E}\left[\int_0^T \big[c^2(t) - e^2(t) + m^2(t)\big]dt + m^2(T)\right] \tag{4}$$

is maximized. The above $(e^*(\cdot), c^*(\cdot))$ is called an implementable contract if it meets the recommended actions of the principal, which is based on the principal's observable wealth $y(\cdot)$. Then, the principal selects his payment $s^*(\cdot)$ and consumption $d^*(\cdot)$, to maximize his preference

$$J_2\big(e^*(\cdot), c^*(\cdot), s(\cdot), d(\cdot)\big) = \frac{1}{2}\mathbb{E}\left[\int_0^T \big[d^2(t) - s^2(t) + y^2(t)\big]dt + y^2(T)\right]. \tag{5}$$

Note that in [30, 31], exponential preferences are considered. For $t > 0$, let

$$\mathcal{F}_t := \sigma\big\{W_1(s), W_2(s), W_3(s), 0 \leq s \leq t\big\}$$

which contains all the information up to time t. Let

$$\mathcal{G}_t^1 := \sigma\big\{W_1(s), W_3(s), 0 \leq s \leq t\big\}$$

contain the information available to the agent, and

$$\mathcal{G}_t^2 := \sigma\big\{W_2(s), W_3(s); 0 \leq s \leq t\big\}$$

contain the information available to the principal, up to time t respectively. Obviously, the information available to them at time t is asymmetric while possesses the overlapping part. In the problem, for any $s(\cdot), d(\cdot)$, first the agent solves the following optimization problem:

$$J_1\big(e^*(\cdot), c^*(\cdot), s(\cdot), d(\cdot)\big) = \max_{e,c} J_1\big(e(\cdot), c(\cdot), s(\cdot), d(\cdot)\big), \tag{6}$$

where $(e^*(\cdot), c^*(\cdot))$ is a $\mathcal{G}_t^1$-adapted process pair. Then the principal solves the following optimization problem:

$$J_2\big(e^*(\cdot), c^*(\cdot), s^*(\cdot), d^*(\cdot)\big) = \max_{s,d} J_2\big(e^*(\cdot), c^*(\cdot), s(\cdot), d(\cdot)\big), \tag{7}$$

where $(s^*(\cdot), d^*(\cdot))$ is a $\mathcal{G}_t^2$-adapted process pair. This formulates a *stochastic linear quadratic (LQ) Stackelberg differential game with overlapping information*. In this setting, the agent is the follower and the principal is the leader. Any process quadruple $(e^*(\cdot), c^*(\cdot), s^*(\cdot), d^*(\cdot))$ satisfying the above two equalities is called a *Stackelberg equilibrium strategy*. In [31], a solvable continuous time principal-agent model is considered under three information structures (full information, hidden actions and hidden savings) and the corresponding optimal contract problems are solved explicitly. But it can not cover our model. For more information for the principal-agent problem, please refer to Schattler and Sung [21], Cvitanić et al. [7], Sannikov [20], Williams [30, 31], Cvitanić and Zhang [8], Liu et al. [16], Cvitanić et al. [6] and the references therein.

Other examples which motivate us to study the problem in this paper can be found in the insider trading model (Øksendal [18]), the cooperative advertising and pricing problem (He et al. [10]), the continuous time manufacturer-newsvendor problem (Øksendal et al. [19]), the LQ Nash differential game with asymmetric information (Chang and Xiao [4]), leader-follower stochastic differential game with asymmetric information (Shi et al. [23]), and optimal reinsurance arrangement problem between the insurer and the reinsurer (Chen and Shen [5]), etc. We will not give their detail statement for the space limitation.

1.2. **Problem formulation**

Inspired by the motivational examples above, we study the stochastic LQ Stackelberg differential game with overlapping information in this paper.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, on which a standard three-dimensional Brownian motion $\{W_1(t), W_2(t), W_3(t)\}_{0\le t\le T}$ is defined, where $T > 0$ is a finite time duration. Let $\{\mathcal{F}_t\}_{0\le t\le T}$ be the natural filtration generated by $(W_1(\cdot), W_2(\cdot), W_3(\cdot))$, which satisfies the usual conditions and $\mathcal{F}_T = \mathcal{F}$. In this paper, $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ will denote the set of $\mathbb{R}^n$-valued, $\{\mathcal{F}_t\}$-adapted, square integrable processes on $[0,T]$.

Suppose that the $\mathbb{R}^n$-valued state process $x^{u_1,u_2}(\cdot)$ satisfies the linear *stochastic differential equation* (SDE)

$$\begin{cases} dx^{u_1,u_2}(t) = \big[A_0(t)x^{u_1,u_2}(t) + B_0(t)u_1(t) + C_0(t)u_2(t)\big]dt \\ \qquad\qquad\qquad + \displaystyle\sum_{i=1}^{3}\big[A_i(t)x^{u_1,u_2}(t) + B_i(t)u_1(t) + C_i(t)u_2(t)\big]dW_i(t),\ t\in[0,T], \\ x^{u_1,u_2}(0) = x_0. \end{cases} \tag{8}$$

Here $u_1(\cdot)$ is the follower's control process and $u_2(\cdot)$ is the leader's control process, which are $\mathbb{R}^{k_1}$ and $\mathbb{R}^{k_2}$-valued, respectively. For $i = 0,1,2,3,\ A_i(\cdot)\in\mathbb{R}^{n\times d_i}$, $B_i(\cdot)\in\mathbb{R}^{n\times k_1}$ and $C_i(\cdot)\in\mathbb{R}^{n\times k_2}$ are all deterministic, bounded, matrix-valued functions and $x_0\in\mathbb{R}^n$. We define the admissible control sets of the follower and the leader, as follows:

$$\mathcal{U}_i := \Big\{u_i(\cdot)\Big|u_i(\cdot):\Omega\times[0,T]\to\mathbb{R}^{k_i} \text{ is } \mathcal{G}_t^i\text{-adapted and square integrable}\Big\},\ i=1,2. \tag{9}$$

Here $\mathcal{G}_t^i := \sigma\{W_i(s), W_3(s); 0 \leq s \leq t\}, i = 1, 2$ denotes the information of the follower and the leader, respectively. It is standard that for any $(x_0, u_1(\cdot), u_2(\cdot)) \in \mathbb{R}^n \times \mathcal{U}_1 \times \mathcal{U}_2$, there exists a unique solution $x^{u_1,u_2}(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)$.

We now formulate the problem by the following two steps. In step 1, the follower chooses a $u_1^*(\cdot) \in \mathcal{U}_1$, which depends on the control $u_2(\cdot)$ of the leader, to minimize the cost functional

$$\begin{aligned} J_1(u_1(\cdot), u_2(\cdot)) = \frac{1}{2}\mathbb{E}\Big[\int_0^T \Big(\big\langle Q_1(t)x^{u_1,u_2}(t), x^{u_1,u_2}(t)\big\rangle + \big\langle N_1(t)u_1(t), u_1(t)\big\rangle\Big)dt \\ + \big\langle G_1 x^{u_1,u_2}(T), x^{u_1,u_2}(T)\big\rangle\Big]. \end{aligned} \tag{10}$$

Here $Q_1(\cdot) \in \mathbb{R}^{n\times n}$ is a bounded, matrix-valued function, $N_1(\cdot) \in \mathbb{R}^{k_1\times k_1}$ is a bounded, matrix-valued function and G_1 is an $\mathbb{R}^{n\times n}$-valued matrix. In step 2, the leader takes into account the follower's optimal control $u_1^*(\cdot)$ in his cost functional, and selects an optimal control $u_2^*(\cdot) \in \mathcal{U}_2$ which will minimize

$$\begin{aligned} J_2(u_1^*(\cdot), u_2(\cdot)) = \frac{1}{2}\mathbb{E}\Big[\int_0^T \Big[\big\langle Q_2(t)x^{u_1^*,u_2}(t), x^{u_1^*,u_2}(t)\big\rangle + \big\langle N_2(t)u_2(t), u_2(t)\big\rangle\Big]dt \\ + \big\langle G_2 x^{u_1^*,u_2}(T), x^{u_1^*,u_2}(T)\big\rangle\Big], \end{aligned} \tag{11}$$

where $x^{u_1^*,u_2}(\cdot)$ denotes the optimal state of the follower which is the solution to (8) with respect to $u_1^*(\cdot)$. Here $Q_2(\cdot) \in \mathbb{R}^{n\times n}$ is a bounded, matrix-valued function, $N_2(\cdot) \in \mathbb{R}^{k_2\times k_2}$ is a bounded, matrix-valued function and G_2 is an $\mathbb{R}^{n\times n}$-valued matrix.

In this general LQ model, the information of the leader and the follower have overlapping part, due to the structure of the admissible control sets. The target of this paper is to give the conditions of its Stackelberg equilibrium strategy $(u_1^*(\cdot), u_2^*(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2$.

Note that some Stackelberg differential games with partially observable information can be put into the above LQ model by the Girsanov transformation. See, for example, Shi et al. [23], Wang et al. [27].

1.3. **Literature review and the contribution of this paper**

In recent years, Stackelberg (also known as leader-follower) game has been an active topic, in the research of nonzero-sum games. Compared with its Nash counterpart, Stackelberg game has many appealing properties, which are useful both in theory and applications. The Stackelberg solution to the game is obtained when one of the players is forced to wait until the other player announces his decision, before making his own decision. Problems of this nature arise frequently in economics, where decisions must be made by two parties and one of them is subordinated to the other, and hence must wait for the other party's decision before formulating its own. The research of Stackelberg game can be traced back to the pioneering work by Stackelberg [26] in static competitive economics. Simann and Cruz [25] studied the dynamic LQ Stackelberg differential game, and the Stackelberg strategy was expressed in terms of Riccati-like differential equations. Bagchi and Başar [1] investigated the stochastic LQ Stackelberg differential game, where the diffusion term of the Ito-type state equation does not contain the state and control variables. Existence and uniqueness of its Stackelberg solution are established, and the leader's optimal strategy is solved as a nonstandard stochastic control problem and is shown to satisfy a particular integral equation. Yong [35] extended the stochastic LQ Stackelberg differential game to a rather general framework, where the coefficients could be random matrices, the control variables could enter the diffusion term of the state equation and the weight matrices for the controls in the cost functionals need not to be positive definite. The problem of the leader is first described as a stochastic control problem of a *forward-backward stochastic differential equation* (FBSDE). Moreover, it is shown that the open-loop solution admits a state feedback representation if a new stochastic Riccati equation is solvable. Øksendal et al. [19] proved a maximum principle for the Stackelberg differential game when the noise is described as an Ito-Lévy process, and found applications to a continuous time manufacturer-newsvendor model. Bensoussan et al. [2] proposed several solution concepts in terms of the players' information sets, for the stochastic Stackelberg differential game with the control-independent diffusion term, and derived the maximum principle under the adapted closed-loop memoryless information structure. Xu and Zhang [33] studied both discrete- and continuous-time stochastic Stackelberg differential games with time delay. By introducing a new costate, a necessary and sufficient condition for the existence and uniqueness of the Stackelberg equilibrium was presented and was designed in terms of three decoupled and symmetric Riccati equations. Moon and Başar [17] considered mean-field Stackelberg differential games with one leader and a large number of heterogeneous followers with distinct types. In Li and Yu [13], a kind of coupled FBSDE with a multilevel self-similar domination-monotonicity structure is introduced to characterize the unique equilibrium of an LQ generalized Stackelberg game with multilevel hierarchy in a closed form. Lin et al. [15] considered the open-loop LQ Stackelberg game of the mean-field stochastic systems in finite horizon. A sufficient condition for the existence and uniqueness of the Stackelberg strategy in terms of the solvability of some Riccati equations is presented. Furthermore, it was shown that the open-loop Stackelberg equilibrium admits a feedback representation involving the new state and its mean. Some recent progress about Stackelberg games can be seen in a review paper by Li and Sethi [14] and the references therein.

However, the above literatures do not consider the feature of asymmetric information in Stackelberg differential game, which we believe, to our best knowledge, that it is a nature and important feature from the point of view of theory and applications. In fact, there are some literatures about asymmetric information game theory. For example, Simann and Cruz [25] considered a non-zero sum velocity-controlled pursuit-evasion game, where the pursuer's information is always later in time than that of the evader's, which is in some sense of time asymmetry. Øksendal [18] solved a universal optimal consumption rate problem with insider trading, where the consumer is called an insider when he has more information than what can be obtained by observing the driving process. That is a kind of information asymmetry with respect to the driving process. Cardaliaguet and Rainer [3] investigated a two-player zero-sum stochastic differential game in which the players have an asymmetric information on the random payoff. Lempa and Matomäki [12] studied a Dynkin game with asymmetric information. The players have asymmetric information on the random expiry time, namely only one of the players is able to observe its occurrence. Chang and Xiao [4] studied an LQ nonzero sum differential game problem with asymmetric information, where different σ-algebra generated by different Brownian motions are introduced to represent the asymmetric information of the two players. Nash equilibrium points are obtained for several classes of asymmetric information by stochastic maximum principle and technique of completion of squares. Shi et el. [23] solved a stochastic leader-follower differential game with asymmetric information, where the information available to the follower is based on some sub-σ-algebra of that available to the leader. Stochastic maximum principles and verification theorems with partial information were obtained. An LQ stochastic leader-follower differential game with noisy observation was solved via measure transformation, stochastic filtering, where not all the diffusion coefficients contain the state and control variables. In a companion paper by Shi et al. [24], an LQ stochastic Stackelberg differential game with asymmetric information was researched, where the control variables enter both diffusion coefficients of the state equation, via some *forward-backward stochastic differential filtering equations* (FBSDFEs). Shi and Wang [22] considered another kind of LQ leader-follower stochastic differential game, where the information available to the leader is a sub-σ-algebra of the filtration generated by the underlying Brownian motion. Wang et. al. [27] focused on an LQ non-zero sum differential game problem driven by the BSDE with asymmetric information. Three classes of observable filtrations are described to classify the information available to the two players. Using the filters of FBSDEs, feedback Nash equilibrium points with observable information generated by Brownian motions were obtained.

In this paper, we consider the stochastic LQ Stackelberg differential game with overlapping information. The LQ problems constitute an extremely important class of optimal control or differential game problems, since they can model many problems in applications, and also reasonably approximate nonlinear control or game problems. The novelty of the formulation and the contribution in this paper is the following.

(i) In our framework, both information filtration available to the leader and the follower could be sub-σ-algebras of the complete information filtration naturally generated by the random noise source. Specifically, the system noise is described by three independent Brownian motions $W_1(t), W_2(t), W_3(t)$, from which the filtration generated denotes the complete information up to time t. The information of the follower comes from the filtration generated by $W_1(t), W_3(t)$, while the information of the leader comes from the filtration generated by $W_2(t), W_3(t)$. This framework is more suitable and interesting to illustrate some game problems in reality.

(ii) The general case that the diffusion terms contain the control and state variables is considered. As is well known in stochastic control and differential game theory, this brings us rather intrinsic mathematical difficulty and technical demanding, especially for the problem of the leader. We overcome the difficulty by the maximum principle approach, the direct calculation of the derivative of the cost functional, and stochastic filtering technique. A new system of high-dimensional Riccati equations is introduced to give the state estimate feedback representation of the Stackelberg equilibrium strategy, while its solvability is a difficult open problem.

(iii) A special case when the diffusion terms are control independent is considered. In this case the system of Riccati equations possesses a kind of decoupling structure.

(iv) A continuous-time principal-agent problem is solved by applying the theoretical results. The Stackelberg equilibrium strategy of the principal and the agent are represented explicitly.

The rest of this paper is organized as follows. In Section 2, the problem formulated in Section 1.2 is solved in two subsections. In subsection 2.1, the follower's problem is considered, while the leader's problem is studied in Subsection 2.2. The Stackelberg equilibrium strategy is derived. A special case with control independent diffusion terms is discussed in Section 3. In Section 4, the results in the previous sections are applied to a continuous-time principal-agent problem which was introduced in Subsection 1.1. Some concluding remarks are given in Section 5. The proof of Theorem 2.2 is left in the Appendix.

2. Main results

In this section, we deal with the problems of the follower and the leader in two subsections, respectively. First, we introduce the following lemma, which belongs to Xiong [32] and plays a fundamental role in this paper.

Lemma 2.1 *Let $f(\cdot), g(\cdot)$ be $\mathcal{F}_t$-adapted processes, satisfying $\mathbb{E}\int_0^T |f(s)|ds + \mathbb{E}\int_0^T |g(s)|^2 ds < \infty$. Then*

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t f(s)ds\bigg|\mathcal{G}_t^i\right] = \int_0^t \mathbb{E}\big[f(s)|\mathcal{G}_t^i\big]ds, \quad \mathbb{E}\left[\int_0^t g(s)dW_3(s)\bigg|\mathcal{G}_t^i\right] = \int_0^t \mathbb{E}\big[g(s)|\mathcal{G}_t^i\big]dW_3(s),\ i=1,2,\\
&\mathbb{E}\left[\int_0^t g(s)dW_1(s)\bigg|\mathcal{G}_t^1\right] = \int_0^t \mathbb{E}\big[g(s)|\mathcal{G}_t^1\big]dW_1(s), \quad \mathbb{E}\left[\int_0^t g(s)dW_1(s)\bigg|\mathcal{G}_t^2\right] = 0,\\
&\mathbb{E}\left[\int_0^t g(s)dW_2(s)\bigg|\mathcal{G}_t^1\right] = 0, \quad \mathbb{E}\left[\int_0^t g(s)dW_2(s)\bigg|\mathcal{G}_t^2\right] = \int_0^t \mathbb{E}\big[g(s)|\mathcal{G}_t^2\big]dW_2(s).
\end{aligned}
\tag{12}
$$

Proof. The proof is quite similar to that of Lemma 5.4 in [32]. □

For any $\mathcal{F}_t$-adapted process $\xi(\cdot)$, we denote by

$$\hat{\xi}(t) := \mathbb{E}[\xi(t)|\mathcal{G}_t^1],\ \check{\xi}(t) := \mathbb{E}[\xi(t)|\mathcal{G}_t^2]$$

and

$$\check{\hat{\xi}}(t) := \mathbb{E}\big[\mathbb{E}[\xi(t)|\mathcal{G}_t^1]\big|\mathcal{G}_t^2\big] \equiv \mathbb{E}\big[\mathbb{E}[\xi(t)|\mathcal{G}_t^2]\big|\mathcal{G}_t^1\big]$$

its optimal filtering estimates.

2.1. Problem of the follower

In this subsection, we try to find the necessary and sufficient conditions for the observable optimal control of the follower. First, we introduce the following assumptions.

(A2.1) *$Q_1(t) \geq 0$, $\forall t \in [0,T]$, and $G_1 \geq 0$.*

For given leader's control $u_2(\cdot)$, assume that there exists a $\mathcal{G}_t^1$-adapted optimal control $u_1^*(\cdot)$ of the follower, and the corresponding optimal state is $x^{u_1^*,u_2}(\cdot)$ as before. We define the follower's Hamiltonian function as

$$\begin{aligned} &H_1\big(t,x,u_1,u_2,q,k_1,k_2,k_3\big) := \big\langle q, A_0(t)x+B_0(t)u_1+C_0(t)u_2\big\rangle \\ &\quad+\sum_{i=1}^{3}\big\langle k_i, A_i(t)x+B_i(t)u_1+C_i(t)u_2\big\rangle-\frac{1}{2}\big\langle Q_1(t)x,x\big\rangle-\frac{1}{2}\big\langle N_1(t)u_1,u_1\big\rangle. \end{aligned} \tag{13}$$

The maximum principle (See, for example, [23]) yields that

$$N_1(t)u_1^*(t)=B_0^\top(t)\hat{q}(t)+\sum_{i=1}^{3}B_i^\top(t)\hat{k}_i(t), \tag{14}$$

where the $\mathcal{F}_t$-adapted process quadruple $(q(\cdot),k_1(\cdot),k_2(\cdot),k_3(\cdot))\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n$ satisfies the adjoint *backward SDE* (BSDE)

$$\begin{cases} -dq(t)=\Big[A_0(t)q(t)+\sum_{i=1}^{3}A_i(t)k_i(t)-Q_1(t)x^{u_1^*,u_2}(t)\Big]dt-\sum_{i=1}^{3}k_i(t)dW_i(t),\ t\in[0,T],\\ q(T)=-\,G_1x^{u_1^*,u_2}(T). \end{cases} \tag{15}$$

Taking clue from the terminal condition, we try to find

$$q(t)=-P_1(t)x^{u_1^*,u_2}(t)-\phi(t), \tag{16}$$

for some $\mathbb{R}^{n\times n}$-valued, deterministic, differentiable function $P_1(\cdot)$ with $P_1(T)=G_1$, and $\mathbb{R}^n$-valued, $\mathcal{F}_t$-adapted process $\phi(\cdot)$ which satisfies the BSDE

$$\begin{cases} d\phi(t)=\alpha(t)dt+\beta_1(t)dW_1(t)+\beta_3(t)dW_3(t),\ t\in[0,T],\\ \phi(T)=0. \end{cases} \tag{17}$$

In the above equation, $\alpha(\cdot),\beta_1(\cdot),\beta_3(\cdot)$ are all $\mathbb{R}^n$-valued, $\mathcal{F}_t$-adapted processes. Applying Itô's formula to (16), we get

$$\begin{aligned} -dq(t)=&\ \big[\dot{P}_1(t)x^{u_1^*,u_2}(t)+P_1(t)A_0(t)x^{u_1^*,u_2}(t)+P_1(t)B_0(t)u_1^*(t)+P^1(t)C_0(t)u_2(t)+\alpha(t)\big]dt\\ &+\sum_{i=1,3}\Big\{P_1(t)\big[A^i(t)x^{u_1^*,u_2}(t)+B^i(t)u_1^*(t)+C^i(t)u_2(t)\big]+\beta^i(t)\Big\}dW_i(t)\\ &+P_1(t)\big[A^2(t)x^{u_1^*,u_2}(t)+B^2(t)u_1^*(t)+C^2(t)u_2(t)\big]dW_2(t). \end{aligned} \tag{18}$$

Comparing (18) with (15), we have

$$\begin{cases} k_1(t)=-P_1(t)\big[A_1(t)x^{u_1^*,u_2}(t)+B_1(t)u_1^*(t)+C_1(t)u_2(t)\big]-\beta_1(t),\\ k_2(t)=-P_1(t)\big[A_2(t)x^{u_1^*,u_2}(t)+B_2(t)u_1^*(t)+C_2(t)u_2(t)\big],\\ k_3(t)=-P_1(t)\big[A_3(t)x^{u_1^*,u_2}(t)+B_3(t)u_1^*(t)+C_3(t)u_2(t)\big]-\beta_3(t), \end{cases} \tag{19}$$

and

$$\begin{aligned} &A_0(t)q(t)+\sum_{i=1}^{3}A_i(t)k_i(t)-Q_1(t)x^{u_1^*,u_2}(t)\\ &=\dot{P}_1(t)x^{u_1^*,u_2}(t)+P_1(t)A_0(t)x^{u_1^*,u_2}(t)+P_1(t)B_0(t)u_1^*(t)+P_1(t)C_0(t)u_2(t)+\alpha(t). \end{aligned} \tag{20}$$

Taking $\mathbb{E}[\cdot|\mathcal{G}_t^1]$ on both sides of (16), (19) and (20), we get

$$\hat{q}(t) = -P_1(t)\hat{x}^{u^{1*},u^2}(t) - \hat{\phi}(t), \tag{21}$$

$$\begin{cases} \hat{k}_1(t) = -P_1(t)\big[A_1(t)\hat{x}^{u^{1*},u^2}(t) + B_1(t)u_1^*(t) + C_1(t)\hat{u}_2(t)\big] - \hat{\beta}_1(t), \\ \hat{k}_2(t) = -P_1(t)\big[A_2(t)\hat{x}^{u^{1*},u^2}(t) + B_2(t)u_1^*(t) + C_2(t)\hat{u}_2(t)\big], \\ \hat{k}_3(t) = -P_1(t)\big[A_3(t)\hat{x}^{u^{1*},u^2}(t) + B_3(t)u_1^*(t) + C_3(t)\hat{u}_2(t)\big] - \hat{\beta}_3(t), \end{cases} \tag{22}$$

and

$$\begin{aligned} &A_0(t)\hat{q}(t) + \sum_{i=1}^{3} A_i(t)\hat{k}_i(t) - Q_1(t)\hat{x}^{u_1^*,u_2}(t) \\ &= \dot{P}_1(t)\hat{x}^{u_1^*,u_2}(t) + P_1(t)A_0(t)\hat{x}^{u_1^*,u_2}(t) + P_1(t)B_0(t)u_1^*(t) + P_1(t)C_0(t)\hat{u}_2(t) + \hat{\alpha}(t). \end{aligned} \tag{23}$$

Applying Lemma 2.1 to (8) corresponding to $u_1^*(\cdot)$ and (15) with $\mathbb{E}[\cdot|\mathcal{G}_t^1]$, we derive the follower's optimal filtering equation

$$\begin{cases} d\hat{x}^{u_1^*,u_2}(t) = \big[A_0(t)\hat{x}^{u_1^*,u_2}(t) + B_0(t)u_1^*(t) + C_0(t)\hat{u}_2(t)\big]dt \\ \qquad\qquad + \sum_{i=1,3} \Big[A_i(t)\hat{x}^{u_1^*,u_2}(t) + B_i(t)u_1^*(t) + C_i(t)\hat{u}_2(t)\Big]dW_i(t), \\ -d\hat{q}(t) = \Big[A_0(t)\hat{q}(t) + \sum_{i=1}^{3} A_i(t)\hat{k}_i(t) - Q_1(t)\hat{x}^{u_1^*,u_2}(t)\Big]dt - \hat{k}_1(t)dW_1(t) - \hat{k}_3(t)dW_3(t),\ t \in [0,T], \\ \hat{x}^{u_1^*,u_2}(0) = x_0, \quad \hat{q}(T) = -G_1\hat{x}^{u_1^*,u_2}(T). \end{cases} \tag{24}$$

Putting (21), (22) into (14), we get

$$\begin{aligned} u_1^*(t) = &-\Big[N_1(t) + \sum_{i=1}^{3} B_i^\top(t)P_1(t)B_i(t)\Big]^{-1}\Big\{\Big[B_0^\top(t)P_1(t) + \sum_{i=1}^{3} B_i^\top(t)P_1(t)A_i(t)\Big]\hat{x}^{u_1^*,u_2}(t) \\ &+ B_0^\top(t)P_1(t)\hat{\phi}(t) + B_1^\top(t)\hat{\beta}_1(t) + B_3^\top(t)\hat{\beta}_3(t) + \Big(\sum_{i=1}^{3} B_i^\top(t)P_1(t)C_i(t)\Big)\hat{u}_2(t)\Big\}, \end{aligned} \tag{25}$$

where we have assumed that

(A2.2) $\overline{N}_1(t) := N_1(t) + \sum\limits_{i=1}^{3} B_i^\top(t)P_1(t)B_i(t) > 0$, $\forall t \in [0,T]$.

Substituting (21), (22) and (25) into (23), we obtain the Riccati's type equation

$$\begin{cases} \dot{P}_1(t) + P_1(t)A_0(t) + A_0^\top(t)P_1(t) + \sum_{i=1}^{3} A_i^\top(t)P_1(t)A_i(t) + Q_1(t) \\ -\Big[P_1(t)B_0(t) + \sum_{i=1}^{3} A_i^\top(t)P_1(t)B_i(t)\Big]\overline{N}_1^{-1}(t)\Big[B_0^\top(t)P_1(t) + \sum_{i=1}^{3} B_i^\top(t)P_1(t)A_i(t)\Big] = 0, \\ P_1(T) = G_1, \end{cases} \tag{26}$$

which is similar to (7.60) of Chapter 6 in Yong and Zhou [36]. The solvability of it can be guaranteed under some additional conditions or in some standard cases, and we will not discuss the detail in this paper.

From the above, we then have

$$\hat{\alpha}(t) = -L_0(t)\hat{\phi}(t) - L_1(t)\hat{\beta}_1(t) - L_3(t)\hat{\beta}_3(t) - L_4(t)\hat{u}_2(t), \tag{27}$$

where (t is omitted for simplification)

$$\begin{cases} L_0 := \overline{N}_1^{-1}\Big(P_1B_0 + \sum_{i=1}^{3} A_i^\top P_1 B_i\Big)B_0^\top P - A_0, \\ L_j := \overline{N}_1^{-1}\Big(P_1B_0 + \sum_{i=1}^{3} A_i^\top P_1 B_i\Big)B_j^\top - A_j,\ j = 1, 3, \\ L_4 := \overline{N}_1^{-1}\Big(P_1B_0 + \sum_{i=1}^{3} A_i^\top P_1 B_i\Big)\Big(\sum_{i=1}^{3} B_i^\top P_1 C_i\Big) - P_1C_0 - \sum_{i=1}^{3} A_i^\top P_1 C_i. \end{cases} \tag{28}$$

Applying Lemma 2.1 again to BSDE (17), we have

$$\begin{cases} -d\hat{\phi}(t) = \big[L_0(t)\hat{\phi}(t) + L_1(t)\hat{\beta}_1(t) + L_3(t)\hat{\beta}_3(t) + L_4(t)\hat{u}_2(t)\big]dt \\ \qquad\qquad - \hat{\beta}_1(t)dW_1(t) - \hat{\beta}_3(t)dW_3(t),\ t \in [0, T], \\ \hat{\phi}(T) = 0. \end{cases} \tag{29}$$

For given $u_2(\cdot)$, (29) admits a unique $\mathcal{G}_t^1$-adapted solution triple $(\hat{\phi}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_3(\cdot))$ by the standard BSDE theory (See, for example, El Karoui et al. [9]). Putting (25) into the forward equation in (24), we get

$$\begin{cases} d\hat{x}^{u_1^*,u_2}(t) = \bigg\{ \Big[A_0(t) - B_0(t)\overline{N}_1^{-1}(t)\Big(B_0^\top(t)P_1(t) + \sum_{i=1}^{3} B_i^\top(t)P_1(t)A_i(t)\Big)\Big]\hat{x}^{u_1^*,u_2}(t) \\ \qquad - B_0(t)\overline{N}_1^{-1}(t)B_0^\top(t)P_1(t)\hat{\phi}(t) - B_0(t)\overline{N}_1^{-1}(t)B_1^\top(t)\hat{\beta}_1(t) - B_0(t)\overline{N}_1^{-1}(t)B_3^\top(t)\hat{\beta}_3(t) \\ \qquad + \Big[C_0(t) - B_0(t)\overline{N}_1^{-1}(t)\Big(\sum_{i=1}^{3} B_i^\top(t)P_1(t)C_i(t)\Big)\Big]\hat{u}_2(t)\bigg\}dt \\ \quad + \sum_{i=1,3}\bigg\{\Big[A_i(t) - B_i(t)\overline{N}_1^{-1}(t)\Big(B_0^\top(t)P_1(t) + \sum_{i=1}^{3} B_i^\top(t)P_1(t)A_i(t)\Big)\Big]\hat{x}^{u_1^*,u_2}(t) \\ \qquad - B_i(t)\overline{N}_1^{-1}(t)B_0^\top(t)P_1(t)\hat{\phi}(t) - B_i(t)\overline{N}_1^{-1}(t)B_1^\top(t)\hat{\beta}_1(t) \\ \qquad - B_i(t)\overline{N}_1^{-1}(t)B_3^\top(t)\hat{\beta}_3(t)\bigg\}dW_i(t),\ t \in [0, T], \\ \hat{x}^{u_1^*,u_2}(0) = x_0, \end{cases} \tag{30}$$

which admits a unique $\mathcal{G}_t^1$-adapted solution $\hat{x}^{u_1^*,u_2}(\cdot)$, from (29). In fact, for given $u_2(\cdot)$, we can verify the solvability of (24). The optimal control $u_1^*(\cdot)$ is expressed by (25).

We summarize the above argument in the following theorem.

Theorem 2.1 *Let* $(\mathbf{A2.1}), (\mathbf{A2.2})$ *hold and* $P_1(\cdot)$ *satisfy (26). For chosen* $u_2(\cdot)$ *of the leader, let* $u_1^*(\cdot)$ *be an optimal control of the follower, then it has the feedback representation of (25), where* $(\hat{x}^{u_1^*,u_2}(\cdot), \hat{\phi}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_3(\cdot))$ *is determined by (29) and (30).*

Remark 2.1 In Yong [35] (also Lin et al. [15]), the explicit optimal cost and the sufficient conditions of the optimality (verification theorem) was proved via the completion of squares method. However, in our overlapping information (including partial and asymmetric information) problem, we need to take the conditional expectation $\mathbb{E}[\cdot|\mathcal{G}_t^1]$ during the computation of the optimal cost of the follower. This approach is invalid when the conditional expectation is taken with respect to the term such as $\langle x^{u_1^*,u_2}(\cdot), x^{u_1^*,u_2}(\cdot)\rangle$, since the variance estimates of the state process $x^{u_1^*,u_2}(\cdot)$ has to be investigated. Up to now, there are no explicit results for this.

However, since the problem in this paper is the LQ case, if we assume additionally that

(A2.3) $N_1(t) \geq 0$, $\forall t \in [0,T]$,

then it is easy to check that the concavity/convexity conditions in the verification theorem of Proposition 2.2 of Shi et al. [23] hold, and $u_1^*(\cdot)$ given by (25) is really optimal.

2.2. Problem of the leader

In this subsection, since the follower's optimal control $u_1^*(\cdot)$ by (25) is a linear functional of $\hat{x}^{u_1^*,u_2}(\cdot), \hat{\phi}(\cdot)$, $\hat{\beta}_1(\cdot), \hat{\beta}_3(\cdot)$ and $\hat{u}_2(\cdot)$, the leader's state equation now writes

$$\begin{cases} dx^{u_2}(t) = \big[A_0(t)x^{u_2}(t) + L_{01}(t)\hat{x}^{u_2}(t) + L_{02}(t)\hat{\phi}(t) + L_{03}(t)\hat{\beta}_1(t) + L_{04}(t)\hat{\beta}_3(t) \\ \qquad\qquad + C_0(t)u_2(t) + L_{05}(t)\hat{u}_2(t)\big]dt + \sum_{i=1}^{3}\Big[A_i(t)x^{u_2}(t) + L_{i1}(t)\hat{x}^{u_2}(t) \\ \qquad\qquad + L_{i2}(t)\hat{\phi}(t) + L_{i3}(t)\hat{\beta}_1(t) + L_{i4}(t)\hat{\beta}_3(t) + C_i(t)u_2(t) + L_{i5}(t)\hat{u}_2(t)\Big]dW_i(t), \\ -d\hat{\phi}(t) = \big[L_0(t)\hat{\phi}(t) + L_1(t)\hat{\beta}_1(t) + L_3(t)\hat{\beta}_3(t) + L_4(t)\hat{u}_2(t)\big]dt \\ \qquad\qquad - \hat{\beta}_1(t)dW_1(t) - \hat{\beta}_3(t)dW_3(t),\ t \in [0,T], \\ x^{u_2}(0) = x_0, \quad \hat{\phi}(T) = 0, \end{cases} \tag{31}$$

where we denote $x^{u_2} \equiv x^{u_1^*,u_2}$, $\hat{x}^{u_2} \equiv \hat{x}^{u_1^*,u_2}$ and for $j = 0,1,2,3$,

$$\begin{cases} L_{j1} := -B_j\overline{N}_1^{-1}\Big(B_0^\top P_1 + \sum_{i=1}^{3} B_i^\top P_1 A_i\Big),\ L_{j2} := -B_j\overline{N}_1^{-1}B_0^\top P_1,\ L_{j3} := -B_j\overline{N}_1^{-1}B_1^\top, \\ L_{j4} := -B_j\overline{N}_1^{-1}B_3^\top,\ L_{j5} := -B_j\overline{N}_1^{-1}\Big(\sum_{i=1}^{3} B_i^\top P_1 C_i\Big). \end{cases} \tag{32}$$

The problem of the leader is to select a $\mathcal{G}_t^2$-adapted optimal control $u_2^*(\cdot)$ such that the cost functional

$$\begin{aligned} &J_2(u_2(\cdot)) \equiv J_2(u_1^*(\cdot), u_2(\cdot)) \\ &= \frac{1}{2}\mathbb{E}\bigg[\int_0^T \Big(\big\langle Q_2(t)x^{u_2}(t), x^{u_2}(t)\big\rangle + \big\langle N_2(t)u_2(t), u_2(t)\big\rangle\Big)dt + \big\langle G_2 x^{u_2}(T), x^{u_2}(T)\big\rangle\bigg] \end{aligned} \tag{33}$$

is minimized.

This subsection is devoted to the necessary and sufficient conditions for the observable optimal control of the leader. We introduce the following assumptions.

(A2.4) $Q_2(t) \geq 0, N_2(t) > 0$, $\forall t \in [0,T]$, *and* $G_2 \geq 0$.

Now, suppose that there exists a $\mathcal{G}_t^2$-adapted optimal control $u_2^*(\cdot)$ of the leader, and his optimal state is $(x^*(\cdot), \hat{\phi}^*(\cdot), \hat{\beta}_1^*(\cdot), \hat{\beta}_3^*(\cdot)) \equiv (x^{u_2^*}(\cdot), \hat{\phi}^*(\cdot), \hat{\beta}_1^*(\cdot), \hat{\beta}_3^*(\cdot))$. Next, we will derive the necessary condition for $u_2^*(\cdot)$, by a direct calculation of the derivative of the cost functional. We define the leader's Hamiltonian function

$$
\begin{aligned}
&H_2\big(t, x^{u_2}, u_2, \phi, \beta_1, \beta_3; p, y, z_1, z_2, z_3\big)\\
:=&\ \big\langle y, A_0(t)x^{u_2} + L_{01}(t)\hat{x}^{u_2} + L_{02}(t)\hat{\phi} + L_{03}(t)\hat{\beta}_1 + L_{04}(t)\hat{\beta}_3 + C_0(t)u_2 + L_{05}(t)\hat{u}_2\big\rangle\\
&+ \big\langle p, L_0(t)\hat{\phi} + L_1(t)\hat{\beta}_1 + L_3(t)\hat{\beta}_3 + L_4(t)\hat{u}_2\big\rangle + \frac{1}{2}\big\langle Q_2(t)x^{u_2}, x^{u_2}\big\rangle + \frac{1}{2}\big\langle N_2(t)u_2, u_2\big\rangle\\
&+ \sum_{i=1}^{3}\big\langle z_i, A_i(t)x^{u_2} + L_{i1}(t)\hat{x}^{u_2} + L_{i2}(t)\hat{\phi} + L_{i3}(t)\hat{\beta}_1 + L_{i4}(t)\hat{\beta}_3 + C_0(t)u_2 + L_{i5}(t)\hat{u}_2\big\rangle,
\end{aligned}
\tag{34}
$$

where the $\mathcal{F}_t$-adapted process quintuple $(p(\cdot), y(\cdot), z_1(\cdot), z_2(\cdot), z_3(\cdot)) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ satisfies the adjoint equation

$$
\begin{cases}
dp(t) = \Big[L_{02}(t)y(t) + L_0(t)p(t) + \displaystyle\sum_{i=1}^{3} L_{i2}(t)z_i(t)\Big]dt\\
\qquad\qquad + \Big[L_{03}(t)y(t) + L_1(t)p(t) + \displaystyle\sum_{i=1}^{3} L_{i3}(t)z_i(t)\Big]dW_1(t)\\
\qquad\qquad + \Big[L_{04}(t)y(t) + L_3(t)p(t) + \displaystyle\sum_{i=1}^{3} L_{i4}(t)z_i(t)\Big]dW_3(t),\\
-dy(t) = \Big[A_0(t)y(t) + L_{01}(t)\hat{y}(t) + \displaystyle\sum_{i=1}^{3} A_i(t)z_i(t) + \sum_{i=1}^{3} L_{i1}(t)\hat{z}_i(t) + Q_2(t)x^*(t)\Big]dt\\
\qquad\qquad - z_1(t)dW_1(t) - z_2(t)dW_2(t) - z_3(t)dW_3(t),\ t \in [0, T],\\
p(0) = 0, \quad y(T) = G_2x^*(T).
\end{cases}
\tag{35}
$$

Without loss of generality, let $x_0 \equiv 0$, and define the perturbed optimal control $u_2^*(\cdot) + \epsilon u_2(\cdot)$ for sufficiently small $\epsilon > 0$, with any $u_2(\cdot)$. From the linearity of (31), the solution to it is $x^*(\cdot) + \epsilon x^{u_2}(\cdot)$. First we have

$$
\begin{aligned}
\widetilde{J}(\epsilon) := J_2\big(u_2^*(\cdot) + \epsilon u_2(\cdot)\big) = \frac{1}{2}\mathbb{E}\int_0^T \Big[\big\langle Q_2(t)(x^*(t) + \epsilon x^{u_2}(t)), x^*(t) + \epsilon x^{u_2}(t)\big\rangle\\
+ \big\langle N_2(t)(u_2^*(t) + \epsilon u_2(t)), u_2^*(t) + \epsilon u_2(t)\big\rangle\Big]dt + \frac{1}{2}\mathbb{E}\big\langle G_2(x^*(T) + \epsilon x^{u_2}(T)), x^*(T) + \epsilon x^{u_2}(T)\big\rangle.
\end{aligned}
$$

Hence

$$
\begin{aligned}
0 = \frac{\partial \widetilde{J}(\epsilon)}{\partial \epsilon}\bigg|_{\epsilon=0} &= \mathbb{E}\int_0^T \Big[\big\langle Q_2(t)x^*(t), x^{u_2}(t)\big\rangle + \big\langle N_2(t)u_2^*(t), u_2(t)\big\rangle\Big]dt + \mathbb{E}\big\langle G_2x^*(T), x^{u_2}(T)\big\rangle\\
&= \mathbb{E}\int_0^T \Big[\big\langle Q_2(t)x^*(t), x^{u_2}(t)\big\rangle + \big\langle N_2(t)u_2^*(t), u_2(t)\big\rangle\Big]dt + \mathbb{E}\big\langle y(T), x^{u_2}(T)\big\rangle.
\end{aligned}
\tag{36}
$$

Applying Itô's formula to $\langle y(t), x^{u_2}(t)\rangle - \langle p(t), \hat{\phi}(t)\rangle$, noticing (31) and (35), we derive

$$
\begin{aligned}
&d\langle y(t), x^{u_2}(t)\rangle - d\langle p(t), \hat{\phi}(t)\rangle\\
&= \Big\langle y(t), \big[L_{01}(t)\hat{x}^{u_2}(t) + C_0(t)u_2(t) + L_{05}(t)\hat{u}_2(t)\big]dt + \sum_{i=1}^{3}\Big[A_i(t)x^{u_2}(t) + L_{i1}(t)\hat{x}^{u_2}(t)\\
&\qquad + L_{i2}(t)\hat{\phi}(t) + L_{i3}(t)\hat{\beta}_1(t) + L_{i4}(t)\hat{\beta}_3(t) + C_i(t)u_2(t) + L_{i5}(t)\hat{u}_2(t)\Big]dW_i(t)\Big\rangle
\end{aligned}
$$

$$- \Big\langle x^{u_2}(t), \Big[L_{01}(t)\hat{y}(t) + \sum_{i=1}^{3} L_{i1}(t)\hat{z}_i(t) + Q_2(t)x^*(t)\Big]dt - \sum_{i=1}^{3} z_i(t)dW_i(t)\Big\rangle$$
$$+ \sum_{i=1}^{3} \Big\langle z_i(t), \big[L_{i1}(t)\hat{x}^{u_2}(t) + C_i(t)u_2(t) + L_{i5}(t)\hat{u}_2(t)\big]dt\Big\rangle + \big\langle p(t), L_4(t)\hat{u}_2(t)dt\big\rangle$$
$$+ \hat{\beta}_1(t)dW_1(t) + \hat{\beta}_3(t)dW_3(t) - \Big\langle \hat{\phi}(t), \Big[L_{03}(t)y(t) + L_1(t)p(t) + \sum_{i=1}^{3} L_{i3}(t)z_i(t)\Big]dW_1(t)$$
$$+ \Big[L_{04}(t)y(t) + L_3(t)p(t) + \sum_{i=1}^{3} L_{i4}(t)z_i(t)\Big]dW_3(t)\Big\rangle.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\big\langle y(T), x^{u_2}(T)\big\rangle &= -\mathbb{E}\int_0^T \Big[\big\langle Q_2(t)x^*(t), x^{u_2}(t)\big\rangle + \big\langle N_2(t)u_2^*(t), u_2(t)\big\rangle\Big]dt \\
&= \mathbb{E}\int_0^T \Big\langle y(t), \big[L_{01}(t)\hat{x}^{u_2}(t) + C_0(t)u_2(t) + L_{05}(t)\hat{u}_2(t)\big]dt\Big\rangle + \mathbb{E}\int_0^T \big\langle p(t), L_4(t)\hat{u}_2(t)dt\big\rangle \\
&\quad - \mathbb{E}\int_0^T \Big\langle x^{u_2}(t), \Big[L_{01}(t)\hat{y}(t) + \sum_{i=1}^{3} L_{i1}(t)\hat{z}_i(t) + Q_2(t)x^*(t)\Big]dt\Big\rangle \\
&\quad + \sum_{i=1}^{3} \mathbb{E}\int_0^T \big\langle z_i(t), \big(L_{i1}(t)\hat{x}^{u_2}(t) + C_i(t)u_2(t) + L_{i5}(t)\hat{u}_2(t)\big)dt\big\rangle.
\end{aligned}$$

Noticing that

$$\mathbb{E}\int_0^T \langle \mathbb{E}[\xi|\mathcal{G}_t^1], \eta\rangle dt = \mathbb{E}\int_0^T \langle \xi, \mathbb{E}[\eta|\mathcal{G}_t^1]\rangle dt, \quad \mathbb{E}\int_0^T \langle \mathbb{E}[\xi|\mathcal{G}_t^2], \eta\rangle dt = \mathbb{E}\int_0^T \langle \xi, \mathbb{E}[\eta|\mathcal{G}_t^2]\rangle dt$$

for any $\mathcal{F}_t$-adapted random variables ξ, η, we have

$$\begin{aligned}
0 &= \mathbb{E}\int_0^T \big\langle N_2(t)u_2^*(t), u_2(t)\big\rangle dt + \mathbb{E}\int_0^T \Big\langle y(t), \big[C_0(t)u_2(t) + L_{05}(t)\hat{u}_2(t)\big]dt\Big\rangle \\
&\quad + \sum_{i=1}^{3} \mathbb{E}\int_0^T \Big\langle z_i(t), \big[C_i(t)u_2(t) + L_{i5}(t)\hat{u}_2(t)\big]dt\Big\rangle + \mathbb{E}\int_0^T \big\langle p(t), L_4(t)\hat{u}_2(t)dt\big\rangle \\
&= \mathbb{E}\int_0^T \big\langle N_2(t)u_2^*(t), u_2(t)\big\rangle dt + \mathbb{E}\int_0^T \Big\langle L_4^\top(t)\hat{p}(t) + C_0^\top(t)y(t) + L_{05}^\top(t)\hat{y}(t) \\
&\qquad + \sum_{i=1}^{3} C_i^\top(t)z_i(t) + \sum_{i=1}^{3} L_{i5}^\top(t)\hat{z}_i(t), u_2(t)\Big\rangle dt.
\end{aligned}$$

This implies that

$$u_2^*(t) = -N_2^{-1}(t)\Big[L_4^\top(t)\check{\hat{p}}(t) + C_0^\top(t)\check{y}(t) + L_{05}^\top(t)\check{\hat{y}}(t) + \sum_{i=1}^{3} C_i^\top(t)\check{z}_i(t) + \sum_{i=1}^{3} L_{i5}^\top(t)\check{\hat{z}}_i(t)\Big], \tag{37}$$

where we have used that

$$\mathbb{E}\int_0^T \langle \xi, \eta\rangle dt = \mathbb{E}\int_0^T \mathbb{E}\big[\langle \xi, \eta\rangle\big|\mathcal{G}_t^2\big]dt = \mathbb{E}\int_0^T \langle \mathbb{E}[\xi|\mathcal{G}_t^2], \eta\rangle dt$$

for any $\mathcal{F}_t$-adapted random variable ξ and $\mathcal{G}_t^2$-adapted random variable η.

In the following, we need to derive the filtering equations for $\check{p}(\cdot),\check{y}(\cdot),\check{z}_i(\cdot)$ and $\check{\hat{p}}(\cdot),\check{\hat{y}}(\cdot),\check{\hat{z}}_i(\cdot)$. Applying again Lemma 2.1 to (35) and (31) corresponding to $u_2^*(\cdot)$ with $\mathbb{E}[\cdot|\mathcal{G}_t^2]$, and putting (37) into it, we obtain

$$
\begin{cases}
d\check{x}^*(t)=\Big\{A_0\check{x}^*+L_{01}\check{\hat{x}}^*+L_{02}\check{\hat{\phi}}^*+L_{03}\check{\hat{\beta}}_1^*+L_{04}\check{\hat{\beta}}_3^*-C_0N_2^{-1}\Big[L_4^\top\check{p}+C_0^\top\check{y}+L_{05}^\top\check{\hat{y}}+\sum_{j=1}^3C_j^\top(t)\check{z}_j(t)\\
\qquad+\sum_{j=1}^3L_{j5}^\top(t)\check{\hat{z}}_j(t)\Big]-L_{05}N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}+\sum_{j=1}^3\big(C_j+L_{j5}\big)^\top\check{\hat{z}}_j\Big]\Big\}dt\\
\qquad+\sum_{i=2,3}\Big\{A_i\check{x}^*+L_{i1}\check{\hat{x}}^*+L_{i2}\check{\hat{\phi}}^*+L_{i3}\check{\hat{\beta}}_1^*+L_{i4}\check{\hat{\beta}}_3^*-C_iN_2^{-1}\Big[L_4^\top\check{p}+C_0^\top\check{y}+L_{05}^\top\check{\hat{y}}\\
\qquad+\sum_{j=1}^3C_j^\top\check{z}_j+\sum_{j=1}^3L_{j5}^\top\check{\hat{z}}_j\Big]-L_{i5}N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}+\sum_{j=1}^3\big(C_j+L_{j5}\big)^\top\check{\hat{z}}_j\Big]\Big\}dW_i(t),\\
d\check{p}(t)=\Big[L_{02}\check{y}+L_0\check{p}+\sum_{i=1}^3L_{i2}\check{z}_i\Big]dt+\Big[L_{04}\check{y}+L_3\check{p}+\sum_{i=1}^3L_{i4}\check{z}_i\Big]dW_3(t),\\
-d\check{\hat{\phi}}^*(t)=\Big\{L_0\check{\hat{\phi}}^*+L_1\check{\hat{\beta}}_1^*+L_3\check{\hat{\beta}}_3^*-L_4N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}+\sum_{j=1}^3\big(C_j+L_{j5}\big)^\top\check{\hat{z}}_j\Big]\Big\}dt-\check{\hat{\beta}}_3dW_3(t),\\
-d\check{y}(t)=\Big[\big(A_0+L_{01}\big)\check{y}+\sum_{j=1}^3\big(A_j+L_{j1}\big)\check{z}_j+Q_2\check{x}^*\Big]dt-\check{z}_2dW_2(t)-\check{z}_3dW_3(t),\ t\in[0,T],\\
\check{x}^*(0)=x_0,\quad \check{p}(0)=0,\quad \check{\hat{\phi}}^*(T)=0,\quad \check{y}(T)=G_2\check{x}^*(T),
\end{cases}
\tag{38}
$$

where the nine tuple $(\check{\hat{x}}^*(\cdot),\check{\hat{\phi}}^*(\cdot),\check{\hat{\beta}}_1^*(\cdot),\check{\hat{\beta}}_3^*(\cdot),\check{\hat{p}}^*(\cdot),\check{\hat{y}}^*(\cdot),\check{\hat{z}}_1^*(\cdot),\check{\hat{z}}_2^*(\cdot),\check{\hat{z}}_3^*(\cdot))$ satisfies

$$
\begin{cases}
d\check{\hat{x}}^*(t)=\Big\{\big(A_0+L_{01}\big)\check{\hat{x}}^*+L_{02}\check{\hat{\phi}}^*+L_{03}\check{\hat{\beta}}_1^*+L_{04}\check{\hat{\beta}}_3^*-\big(C_0+L_{05}\big)N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}\\
\qquad+\sum_{i=1}^3\big(C_i+L_{i5}\big)^\top\check{\hat{z}}_i\Big]\Big\}dt+\Big\{\big(A_3+L_{31}\big)\check{\hat{x}}^*+L_{32}\check{\hat{\phi}}^*+L_{33}\check{\hat{\beta}}^{1*}+L_{34}\check{\hat{\beta}}_3^*\\
\qquad-\big(C_3+L_{35}\big)N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}+\sum_{i=1}^3\big(C_3+L_{35}\big)^\top\check{\hat{z}}_3\Big]\Big\}dW_3(t),\\
d\check{\hat{p}}(t)=\Big[L_{02}\check{\hat{y}}+L_0\check{\hat{p}}+\sum_{i=1}^3L_{i2}\check{\hat{z}}_i\Big]dt+\Big[L_{04}\check{\hat{y}}+L_3\check{\hat{p}}+\sum_{i=1}^3L_{i4}\check{\hat{z}}_i\Big]dW_3(t),\\
-d\check{\hat{\phi}}^*(t)=\Big\{L_0\check{\hat{\phi}}^*+L_1\check{\hat{\beta}}_1^*+L_3\check{\hat{\beta}}_3^*-L_4N_2^{-1}\Big[L_4^\top\check{\hat{p}}+\big(C_0+L_{05}\big)^\top\check{\hat{y}}+\sum_{i=1}^3\big(C_i+L_{i5}\big)^\top\check{\hat{z}}_i\Big]\Big\}dt-\check{\hat{\beta}}_3dW_3(t),\\
-d\check{\hat{y}}(t)=\Big[\big(A_0+L_{01}\big)\check{\hat{y}}+\sum_{i=1}^3\big(A_i+L_{i1}\big)\check{\hat{z}}_i+Q_2\check{\hat{x}}^*\Big]dt-\check{\hat{z}}_3dW_3(t),\ t\in[0,T],\\
\check{\hat{x}}^*(0)=x_0,\quad \check{\hat{p}}(0)=0,\quad \check{\hat{\phi}}^*(T)=0,\quad \check{\hat{y}}(T)=G_2\check{\hat{x}}^*(T).
\end{cases}
\tag{39}
$$

Up to now, we have obtained the optimal control $u_2^*(\cdot)$ of the leader by (37). However, this representation relies on the solvability of filtering equations (38) and (39). In the following, we will derive the state estimate feedback representation of (37), via some Riccati type equations. And the solvability of the above filtering equations will be solved as a corollary.

For this target, first we rewrite the optimal state of the leader as

$$
\begin{cases}
dx^*(t) = \Big\{ A_0 x^* + L_{01}\hat{x}^* + L_{02}\hat{\phi}^* + L_{03}\hat{\beta}_1^* + L_{04}\hat{\beta}_3^* - C^0 N_2^{-1}\Big[L_4^\top \check{\hat{p}} + C_0^\top \check{y} + L_{05}^\top \check{\hat{y}} + \sum_{j=1}^3 C_j^\top \check{z}_j \\
\qquad + \sum_{j=1}^3 L_{j5}^\top \check{\hat{z}}_j \Big] - L_{05} N_2^{-1}\Big[L_4^\top \check{\hat{p}} + \big(C_0 + L_{05}\big)^\top \check{\hat{y}} + \sum_{i=1}^3 \Big\{ A_i x^* + L_{i1}\hat{x}^* + L_{i2}\hat{\phi}^* + L_{i3}\hat{\beta}_1^* \\
\qquad + L_{i4}\hat{\beta}_3^* - C_i N_2^{-1}\Big[L_4^\top \check{\hat{p}} + C_0^\top \check{y} + L_{05}^\top \check{\hat{y}} + \sum_{j=1}^3 C_j^\top \check{z}_j + \sum_{j=1}^3 L_{j5}^\top \check{\hat{z}}_j \Big] \\
\qquad - L_{i5} N_2^{-1}\Big[L_4^\top \check{\hat{p}} + \big(C_0 + L_{05}\big)^\top \check{\hat{y}} + \sum_{j=1}^3 \big(C_j + L_{j5}\big)^\top \check{\hat{z}}_j \Big] \Big\} dW_i(t), \\
-d\hat{\phi}^*(t) = \Big\{ L_0\hat{\phi}^* + L_1\hat{\beta}_1^* + L_3\hat{\beta}_3^* - L_4 N_2^{-1}\Big[L_4^\top \check{\hat{p}} + \big(C_0 + L_{05}\big)^\top \check{\hat{y}} + \sum_{j=1}^3 \big(C_j + L_{j5}\big)^\top \check{\hat{z}}_j \Big] \Big\} dt \\
\qquad - \hat{\beta}_1^* dW_1(t) - \hat{\beta}_3^* dW_3(t), \ t \in [0,T], \\
x^*(0) = x_0, \quad \hat{\phi}^*(T) = 0.
\end{cases}
\tag{40}
$$

Now, inspired by [35], let

$$
X := \begin{pmatrix} x^* \\ p \end{pmatrix}, \ Y := \begin{pmatrix} y \\ \hat{\phi}^* \end{pmatrix}, \ Z_1 := \begin{pmatrix} z_1 \\ \hat{\beta}_1^* \end{pmatrix}, \ Z_2 := \begin{pmatrix} z_2 \\ 0 \end{pmatrix}, \ Z_3 := \begin{pmatrix} z_3 \\ \hat{\beta}_3^* \end{pmatrix},
\tag{41}
$$

then (40) and (35) can be rewritten as

$$
\begin{cases}
dX(t) = \big(\mathcal{A}_0 X + \widehat{\mathcal{A}}_0 \hat{X} + \overline{\mathcal{A}}_0 \check{\hat{X}} + \mathcal{B}_0 Y + \mathcal{C}_0 \check{Y} + \widetilde{\mathcal{C}}_0 \check{\hat{Y}} + \mathcal{B}_1^\top Z_1 + \widetilde{\mathcal{B}}_1^\top \check{Z}_1 + \overline{\mathcal{C}}_0 \check{\hat{Z}}_1 + \mathcal{B}_2^\top Z_2 + \widetilde{\mathcal{B}}_2^\top \check{Z}_2 \\
\qquad + \overline{\mathcal{D}}_0 \check{\hat{Z}}_2 + \mathcal{B}_3^\top Z_3 + \widetilde{\mathcal{B}}_3^\top \check{Z}_3 + \overline{\mathcal{E}}_0 \check{\hat{Z}}_3 \big) dt + \sum_{i=1}^3 \Big(\mathcal{A}_i X + \widehat{\mathcal{A}}_i \hat{X} + \overline{\mathcal{A}}_i \check{\hat{X}} + \mathcal{B}_i Y + \widetilde{\mathcal{B}}_i \check{Y} + \overline{\mathcal{B}}_i \check{\hat{Y}} \\
\qquad + \mathcal{B}_i Z_1 + \widetilde{\mathcal{C}}_i \check{Z}_1 + \overline{\mathcal{C}}_i \check{\hat{Z}}_1 + \mathcal{D}_i Z_2 + \widetilde{\mathcal{D}}_i \check{Z}_2 + \overline{\mathcal{D}}_i \check{\hat{Z}}_2 + \mathcal{E}_i Z_3 + \widetilde{\mathcal{E}}_i \check{Z}_3 + \overline{\mathcal{E}}_i \check{\hat{Z}}_3 \Big) dW_i(t), \\
-dY(t) = \big(\mathcal{Q}_2 X + \mathcal{H}_1 \check{\hat{X}} + \mathcal{A}_0 Y + \mathcal{H}_2 \hat{Y} + \overline{\mathcal{A}}_0^\top \check{\hat{Y}} + \mathcal{A}_1 Z_1 + \mathcal{H}_3 \hat{Z}_1 + \overline{\mathcal{A}}_1^\top \check{\hat{Z}}_1 + \mathcal{A}_2 Z_2 + \hat{\mathcal{A}}_2 \hat{Z}_2 + \overline{\mathcal{A}}_2^\top \check{\hat{Z}}_2 \\
\qquad + \mathcal{A}_3 Z_3 + \hat{\mathcal{A}}_3 \hat{Z}_3 + \overline{\mathcal{A}}_3^\top \check{\hat{Z}}_3 \big) dt - Z_1 dW_1(t) - Z_2 dW_2(t) - Z_3 dW_3(t), \ t \in [0,T], \\
X(0) = X_0, \quad Y(T) = \mathcal{G}_2 X(T),
\end{cases}
\tag{42}
$$

where

$$
\begin{cases}
X_0 := \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \ \mathcal{Q}_2 := \begin{pmatrix} Q_2 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{G}_2 := \begin{pmatrix} G_2 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{H}_1 := \begin{pmatrix} 0 & 0 \\ 0 & -L_4 N_2^{-1} L_4^\top \end{pmatrix}, \\
\mathcal{H}_2 := \begin{pmatrix} L_{01} & 0 \\ 0 & L_0 \end{pmatrix}, \ \mathcal{H}_3 := \begin{pmatrix} L_{11} & 0 \\ 0 & L_1 \end{pmatrix}, \ \mathcal{A}_0 := \begin{pmatrix} A_0 & 0 \\ 0 & L_0 \end{pmatrix}, \ \widehat{\mathcal{A}}_0 := \begin{pmatrix} L_{01} & 0 \\ 0 & 0 \end{pmatrix}, \\
\overline{\mathcal{A}}_0 := \begin{pmatrix} 0 & -(C_0 + L_{05}) N_2^{-1} L_4^\top \\ 0 & 0 \end{pmatrix}, \ \mathcal{B}_0 := \begin{pmatrix} 0 & L_{02} \\ L_{02} & 0 \end{pmatrix}, \ \mathcal{C}_0 := \begin{pmatrix} -C_0 N_2^{-1} C_0^\top & 0 \\ 0 & 0 \end{pmatrix},
\end{cases}
$$

$$
\begin{cases}
\widetilde{\mathcal{C}}_0 := \begin{pmatrix} -(C_0+L_{05})N_2^{-1}(C_0+L_{05})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{C}}_0 := \begin{pmatrix} -C_0N_2^{-1}L_{15}^\top - L_{05}N_2^{-1}(C_1+L_{15})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\overline{\mathcal{D}}_0 := \begin{pmatrix} -C_0N_2^{-1}L_{25}^\top - L_{05}N_2^{-1}(C_2+L_{25})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{E}}_0 := \begin{pmatrix} -C_0N_2^{-1}L_{35}^\top - L_{05}N_2^{-1}(C_3+L_{35})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{A}_1 := \begin{pmatrix} A_1 & 0 \\ 0 & L_1 \end{pmatrix}, \ \widehat{\mathcal{A}}_1 := \begin{pmatrix} L_{11} & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{A}}_1 := \begin{pmatrix} 0 & -(C_1+L_{15})N_2^{-1}L_4^\top \\ 0 & 0 \end{pmatrix}, \ \mathcal{B}_1 := \begin{pmatrix} 0 & L_{12} \\ L_{03} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{B}}_1 := \begin{pmatrix} -C_1N_2^{-1}C_0^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{B}}_1 := \begin{pmatrix} -(C_1+L_{15})N_2^{-1}(C_0+L_{05})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{C}_1 := \begin{pmatrix} 0 & L_{13} \\ L_{13} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{C}}_1 := \begin{pmatrix} -C_1N_2^{-1}C_1^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{C}}_1 := \begin{pmatrix} -C_1N_2^{-1}L_{15} - L_{15}N_2^{-1}(C_1+L_{15})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{D}_1 := \begin{pmatrix} 0 & 0 \\ L_{23} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{D}}_1 := \begin{pmatrix} -C_1N_2^{-1}C_2^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{D}}_1 := \begin{pmatrix} -C_1N_2^{-1}L_{25} - L_{15}N_2^{-1}(C_2+L_{25})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{E}_1 := \begin{pmatrix} 0 & L_{14} \\ L_{33} & 0 \end{pmatrix}, \ \widetilde{\mathcal{E}}_1 := \begin{pmatrix} -C_1N_2^{-1}C_3^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{E}}_1 := \begin{pmatrix} -C_1N_2^{-1}L_{35} - L_{15}N_2^{-1}(C_3+L_{35})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{A}_2 := \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix}, \ \widehat{\mathcal{A}}_2 := \begin{pmatrix} L_{21} & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{A}}_2 := \begin{pmatrix} 0 & -(C_2+L_{25})N_2^{-1}L_4^\top \\ 0 & 0 \end{pmatrix}, \ \mathcal{B}_2 := \begin{pmatrix} 0 & L_{22} \\ 0 & 0 \end{pmatrix}, \\
\widetilde{\mathcal{B}}_2 := \begin{pmatrix} -C_2N_2^{-1}C_0^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{B}}_2 := \begin{pmatrix} -(C_2+L_{25})N_2^{-1}(C_0+L_{05})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{C}_2 := \begin{pmatrix} 0 & L_{23} \\ 0 & 0 \end{pmatrix}, \\
\widetilde{\mathcal{C}}_2 := \begin{pmatrix} -C_2N_2^{-1}C_1^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{C}}_2 := \begin{pmatrix} -C_2N_2^{-1}L_{15} - L_{25}N_2^{-1}(C_1+L_{15})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{D}_2 \equiv 0, \\
\widetilde{\mathcal{D}}_2 := \begin{pmatrix} -C_2N_2^{-1}C_2^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{D}}_2 := \begin{pmatrix} -C_2N_2^{-1}L_{25} - L_{25}N_2^{-1}(C_2+L_{25})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{E}_2 := \begin{pmatrix} 0 & L_{24} \\ 0 & 0 \end{pmatrix}, \ \widetilde{\mathcal{E}}_2 := \begin{pmatrix} -C_2N_2^{-1}C_3^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{E}}_2 := \begin{pmatrix} -C_2N_2^{-1}L_{35} - L_{25}N_2^{-1}(C_3+L_{35})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{A}_3 := \begin{pmatrix} A_3 & 0 \\ 0 & L_3 \end{pmatrix}, \ \widehat{\mathcal{A}}_3 := \begin{pmatrix} L_{31} & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{A}}_3 := \begin{pmatrix} 0 & -(C_3+L_{35})N_2^{-1}L_4^\top \\ 0 & 0 \end{pmatrix}, \ \mathcal{B}_3 := \begin{pmatrix} 0 & L_{32} \\ L_{04} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{B}}_3 := \begin{pmatrix} -C_3N_2^{-1}C_0^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{B}}_3 := \begin{pmatrix} -(C_3+L_{35})N_2^{-1}(C_0+L_{05})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{C}_3 := \begin{pmatrix} 0 & L_{33} \\ L_{14} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{C}}_3 := \begin{pmatrix} -C_3N_2^{-1}C_1^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{C}}_3 := \begin{pmatrix} -C_3N_2^{-1}L_{15} - L_{35}N_2^{-1}(C_1+L_{15})^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \mathcal{D}_3 := \begin{pmatrix} 0 & 0 \\ L_{24} & 0 \end{pmatrix}, \\
\widetilde{\mathcal{D}}_3 := \begin{pmatrix} -C_3N_2^{-1}C_2^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{D}}_3 := \begin{pmatrix} -C_3N_2^{-1}L_{25} - L_{35}N_2^{-1}(C_2+L_{25})^\top & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{E}_3 := \begin{pmatrix} 0 & L_{34} \\ L_{34} & 0 \end{pmatrix}, \ \widetilde{\mathcal{E}}_3 := \begin{pmatrix} -C_3N_2^{-1}C_3^\top & 0 \\ 0 & 0 \end{pmatrix}, \ \overline{\mathcal{E}}_3 := \begin{pmatrix} -C_3N_2^{-1}L_{35} - L_{35}N_2^{-1}(C_3+L_{35})^\top & 0 \\ 0 & 0 \end{pmatrix}.
\end{cases}
$$

And (37) can be rewritten as

$$
u_2^*(t) = -N_2^{-1}\Big[\mathcal{L}_4^\top \check{\hat{X}}(t) + \mathcal{C}_{05}^\top \check{Y}(t) + \mathcal{L}_{05}^\top \check{\hat{Y}}(t) + \sum_{i=1}^{3} \mathcal{C}_{i5}^\top \check{Z}_i(t) + \sum_{i=1}^{3} \mathcal{L}_{i5}^\top \check{\hat{Z}}_i(t)\Big], \tag{43}
$$

where

$$
\mathcal{L}_4 := \begin{pmatrix} 0 \\ L_4 \end{pmatrix}, \ \mathcal{C}_{i5} := \begin{pmatrix} C_i \\ 0 \end{pmatrix}, \ \mathcal{L}_{i5} := \begin{pmatrix} L_{i5} \\ 0 \end{pmatrix}, \ i = 0, 1, 2, 3.
$$

We have the following theorem and its proof is left in the Appendix.

Theorem 2.2 *Let* $(\mathbf{A2.1})-(\mathbf{A2.4})$ *and* $(\mathbf{A2.5})-(\mathbf{A2.8})$ *in the Appendix hold. Let* $(\mathcal{P}_1(\cdot),\mathcal{P}_2(\cdot),\mathcal{P}_3(\cdot),\mathcal{P}_4(\cdot))$ *satisfy the following system of Riccati equations*

$$\begin{cases}
0=\dot{\mathcal{P}}_1+\mathcal{P}_1\mathcal{A}_0+\mathcal{A}_0\mathcal{P}_1+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_1+\mathcal{P}_1\big(\mathcal{B}_1^\top\Sigma_1+\mathcal{B}_2^\top\Sigma_2+\mathcal{B}_3^\top\Sigma_3\big)+\mathcal{A}_1\Sigma_1+\mathcal{A}_2\Sigma_2+\mathcal{A}_3\Sigma_3+\mathcal{Q}_2,\\
0=\dot{\mathcal{P}}_2+\mathcal{P}_1\widehat{\mathcal{A}}_0+\mathcal{A}_0\mathcal{P}_2+\mathcal{H}_1\mathcal{P}_1+\mathcal{H}_1\mathcal{P}_2+\mathcal{P}_2(\mathcal{A}_0+\widehat{\mathcal{A}}_0)+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_1+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_2\\
\quad+\mathcal{P}_1\big(\mathcal{B}_1^\top\widehat{\Sigma}_1+\mathcal{B}_2^\top\widehat{\Sigma}_2+\mathcal{B}_3\widehat{\Sigma}_3\big)+\mathcal{P}_2\big[\mathcal{B}_1^\top(\Sigma_1+\widehat{\Sigma}_1)+\mathcal{B}_2^\top(\Sigma_2+\widehat{\Sigma}_2)+\mathcal{B}_3^\top(\Sigma_3+\widehat{\Sigma}_3)\big]\\
\quad+\mathcal{A}_1\widehat{\Sigma}_1+\mathcal{A}_2\widehat{\Sigma}_2+\mathcal{A}_3\widehat{\Sigma}_3+\widehat{\mathcal{A}}_2(\Sigma_2+\widehat{\Sigma}_2)+\widehat{\mathcal{A}}_3(\Sigma_3+\widehat{\Sigma}_3),\\
0=\dot{\mathcal{P}}_3+\mathcal{P}_3\mathcal{A}_0+\mathcal{A}_0\mathcal{P}_3+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_3+\mathcal{P}_1\mathcal{C}_0(\mathcal{P}_1+\mathcal{P}_2)+\mathcal{P}_3\big(\mathcal{B}_0+\mathcal{C}_0\big)(\mathcal{P}_1+\mathcal{P}_3)\\
\quad+\mathcal{P}_1\big(\mathcal{B}_1^\top\widetilde{\Sigma}_1+\mathcal{B}_2^\top\widetilde{\Sigma}_2+\mathcal{B}_3^\top\widetilde{\Sigma}_3\big)+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_1^\top+\mathcal{P}_3\mathcal{B}_1^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_1^\top\big)(\Sigma_1+\widetilde{\Sigma}_1)+\mathcal{A}_1\widetilde{\Sigma}_1+\mathcal{A}_2\widetilde{\Sigma}_2\\
\quad+\mathcal{A}_3\widetilde{\Sigma}_3+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_2^\top+\mathcal{P}_3\mathcal{B}_2^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_2^\top\big)(\Sigma_2+\widetilde{\Sigma}_2)+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_3^\top+\mathcal{P}_3\mathcal{B}_3^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_3^\top\big)(\Sigma_3+\widetilde{\Sigma}_3),\\
0=\dot{\mathcal{P}}_4+\mathcal{P}_4(\mathcal{A}_0+\widehat{\mathcal{A}}_0+\overline{\mathcal{A}}_0)+\mathcal{A}_0\mathcal{P}_4+\mathcal{H}_2\mathcal{P}_3+\mathcal{H}_2\mathcal{P}_4+\mathcal{P}_3(\widehat{\mathcal{A}}_0+\overline{\mathcal{A}}_0)+\mathcal{P}_3(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_2+\mathcal{P}_4)\\
\quad+\mathcal{P}_1\overline{\mathcal{A}}_0+\mathcal{H}_1+\big[\overline{\mathcal{A}}_0^\top+\mathcal{P}_2(\mathcal{C}_0+\widetilde{\mathcal{C}}_0)+\mathcal{P}_3\widetilde{\mathcal{C}}_0+\mathcal{P}_4(\mathcal{B}_0+\mathcal{C}_0+\widetilde{\mathcal{C}}_0)\big](\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\\
\quad+\big(\mathcal{A}_1+\mathcal{P}_1\mathcal{B}_1^\top\big)\overline{\Sigma}_1+\big(\mathcal{A}_2+\mathcal{P}_1\mathcal{B}_2^\top\big)\overline{\Sigma}_2+\big(\mathcal{A}_3+\mathcal{P}_1\mathcal{B}_3^\top\big)\overline{\Sigma}_3+\mathcal{P}_2\mathcal{B}_1^\top(\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
\quad+\big(\widehat{\mathcal{A}}_2+\mathcal{P}_2\mathcal{B}_2^\top\big)(\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\big(\widehat{\mathcal{A}}_3+\mathcal{P}_2\mathcal{B}_3^\top\big)(\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_1^\top+\mathcal{P}_3\mathcal{B}_1^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_1^\top\big)(\widehat{\Sigma}_1+\overline{\Sigma}_1)\\
\quad+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_2^\top+\mathcal{P}_3\mathcal{B}_2^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_2^\top\big)(\widehat{\Sigma}_2+\overline{\Sigma}_2)+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_3^\top+\mathcal{P}_3\mathcal{B}_3^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_3^\top\big)(\widehat{\Sigma}_3+\overline{\Sigma}_3)\\
\quad+\big[\overline{\mathcal{A}}_1^\top+\mathcal{P}_1\overline{\mathcal{C}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_1^\top+\overline{\mathcal{C}}_0\big)+\mathcal{P}_3\overline{\mathcal{C}}_0+\mathcal{P}_4\big(\mathcal{B}_1^\top+\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0\big)\big](\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
\quad+\big[\overline{\mathcal{A}}_2^\top+\mathcal{P}_1\overline{\mathcal{D}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0\big)+\mathcal{P}_3\overline{\mathcal{D}}_0+\mathcal{P}_4\big(\mathcal{B}_2^\top+\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0\big)\big](\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)\\
\quad+\big[\overline{\mathcal{A}}_3^\top+\mathcal{P}_1\overline{\mathcal{E}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{E}}_0\big)+\mathcal{P}_3\overline{\mathcal{E}}_0+\mathcal{P}_4\big(\mathcal{B}_3^\top+\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{E}}_0\big)\big](\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3),\\
\mathcal{P}_1(T)=\mathcal{G}_2,\ \mathcal{P}_2(T)=0,\ \mathcal{P}_3(T)=0,\ \mathcal{P}_4(T)=0,
\end{cases}\tag{44}$$

where $\Sigma_i,\widehat{\Sigma}_i,\widetilde{\Sigma}_i,\overline{\Sigma}_i$ *in the above depend on* P_i *for* $i=1,2,3,4$ *(The definitions of them are in the Appendix.). Let* $\check{\check{X}}(\cdot)$ *be the* $\mathcal{G}_t^1\cap\mathcal{G}_t^2$*-adapted solution to*

$$\begin{cases}
d\check{\check{X}}(t)=\big[\mathcal{A}_0+\widehat{\mathcal{A}}_0+\overline{\mathcal{A}}_0+(\mathcal{B}_0+\mathcal{C}_0+\widetilde{\mathcal{C}}_0)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)+(\mathcal{B}_1^\top+\widetilde{\mathcal{B}}_1^\top+\overline{\mathcal{C}}_0)(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
\qquad+(\mathcal{B}_2^\top+\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0)(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+(\mathcal{B}_3^\top+\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{E}}_0)(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\check{X}}(t)dt\\
\qquad+\big[\mathcal{A}_3+\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3+(\mathcal{B}_3+\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)+(\mathcal{B}_3+\widetilde{\mathcal{C}}_3+\overline{\mathcal{C}}_3)(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
\qquad+(\mathcal{D}_3+\widetilde{\mathcal{D}}_3+\overline{\mathcal{D}}_3)(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+(\mathcal{E}_3+\widetilde{\mathcal{E}}_3+\overline{\mathcal{E}}_3)(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\check{X}}(t)dW_3(t),\ t\in[0,T],\\
\check{\check{X}}(0)=X_0,
\end{cases}\tag{45}$$

$\check{X}(\cdot)$ *be the* $\mathcal{G}_t^2$*-adapted solution to*

$$\begin{cases}
d\check{X}(t)=\Big\{\big[\mathcal{A}_0+(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_3)+(\mathcal{B}_1+\widetilde{\mathcal{B}}_1)^\top(\Sigma_1+\widetilde{\Sigma}_1)+(\mathcal{B}_2+\widetilde{\mathcal{B}}_2)^\top(\Sigma_2+\widetilde{\Sigma}_2)\\
\qquad+(\mathcal{B}_3+\widetilde{\mathcal{B}}_3)^\top(\Sigma_3+\widetilde{\Sigma}_3)\big]\check{X}(t)+\big[\widehat{\mathcal{A}}_0+\overline{\mathcal{A}}_0+(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_2+\mathcal{P}_4)+(\mathcal{B}_1+\widetilde{\mathcal{B}}_1)^\top(\widehat{\Sigma}_1+\overline{\Sigma}_1)\\
\qquad+(\mathcal{B}_2+\widetilde{\mathcal{B}}_2)^\top(\widehat{\Sigma}_2+\overline{\Sigma}_2)+(\mathcal{B}_3+\widetilde{\mathcal{B}}_3)^\top(\widehat{\Sigma}_3+\overline{\Sigma}_3)+\widetilde{\mathcal{C}}_0(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\\
\qquad+\overline{\mathcal{C}}_0(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)+\overline{\mathcal{D}}_0(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\overline{\mathcal{E}}_0(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\check{X}}(t)\Big\}dt\\
\qquad+\sum_{i=2,3}\Big\{\big[\mathcal{A}_i+(\mathcal{B}_i+\widetilde{\mathcal{B}}_i)(\mathcal{P}_1+\mathcal{P}_3)+(\mathcal{B}_i+\widetilde{\mathcal{C}}_i)(\Sigma_1+\widetilde{\Sigma}_1)+(\mathcal{D}_i+\widetilde{\mathcal{D}}_i)(\Sigma_2+\widetilde{\Sigma}_2)\\
\qquad+(\mathcal{E}_i+\widetilde{\mathcal{E}}_i)(\Sigma_3+\widetilde{\Sigma}_3)\big]\check{X}(t)+\big[\widehat{\mathcal{A}}_i+\overline{\mathcal{A}}_i+(\mathcal{B}_i+\widetilde{\mathcal{B}}_i)(\mathcal{P}_2+\mathcal{P}_4)+\overline{\mathcal{B}}_i(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\\
\qquad+(\mathcal{B}_i+\widetilde{\mathcal{C}}_i)(\widehat{\Sigma}_1+\overline{\Sigma}_1)+(\mathcal{D}_i+\widetilde{\mathcal{D}}_i)(\widehat{\Sigma}_2+\overline{\Sigma}_2)+(\mathcal{E}_i+\widetilde{\mathcal{E}}_i)(\widehat{\Sigma}_3+\overline{\Sigma}_3)+\overline{\mathcal{C}}_i(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
\qquad+\overline{\mathcal{D}}_i(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\overline{\mathcal{E}}_i(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\check{X}}(t)\Big\}dW_i(t),\ t\in[0,T],\\
\check{X}(0)=X_0,
\end{cases}\tag{46}$$

$\hat{X}(\cdot)$ be the $\mathcal{G}_t^1$-adapted solution to

$$
\begin{cases}
d\hat{X}(t) = \Big\{\big[\mathcal{A}_0 + \widehat{\mathcal{A}}_0 + \mathcal{B}_0(\mathcal{P}_1+\mathcal{P}_2) + \mathcal{B}_1^\top(\Sigma_1+\widehat{\Sigma}_1) + \mathcal{B}_2^\top(\Sigma_2+\widehat{\Sigma}_2) + \mathcal{B}_3^\top(\Sigma_3+\widehat{\Sigma}_3)\big]\hat{X}(t) \\
\qquad + \big[\overline{\mathcal{A}}_0 + \mathcal{B}_0(\mathcal{P}_3+\mathcal{P}_4) + \mathcal{B}_1^\top(\widetilde{\Sigma}_1+\overline{\Sigma}_1) + \mathcal{B}_2^\top(\widetilde{\Sigma}_2+\overline{\Sigma}_2) + \mathcal{B}_3^\top(\widetilde{\Sigma}_3+\overline{\Sigma}_3) \\
\qquad + (\mathcal{C}_0+\widetilde{\mathcal{C}}_0)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + (\widetilde{\mathcal{B}}_1^\top+\widetilde{\mathcal{B}}_2^\top+\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{C}}_0)(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1) \\
\qquad + \overline{\mathcal{D}}_0(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2) + \overline{\mathcal{E}}_0(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\hat{X}}(t)\Big\}dt \\
\quad + \sum_{i=1,3}\Big\{\big[\mathcal{A}_i+\widehat{\mathcal{A}}_i+\mathcal{B}_i(\mathcal{P}_1+\mathcal{P}_2)+\mathcal{B}_i(\Sigma_1+\widehat{\Sigma}_1)+\mathcal{D}_i(\Sigma_2+\widehat{\Sigma}_2)+\mathcal{E}_i(\Sigma_3+\widehat{\Sigma}_3)\big]\hat{X}(t) \\
\qquad + \big[\overline{\mathcal{A}}_i+\mathcal{B}_i(\mathcal{P}_3+\mathcal{P}_4)+\mathcal{B}_i(\widetilde{\Sigma}_1+\overline{\Sigma}_1)+\mathcal{D}_i(\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\mathcal{E}_i(\widetilde{\Sigma}_3+\overline{\Sigma}_3) \\
\qquad + (\widetilde{\mathcal{B}}_i+\overline{\mathcal{B}}_i)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + (\widetilde{\mathcal{C}}_i+\overline{\mathcal{C}}_i)(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1) \\
\qquad + (\widetilde{\mathcal{D}}_i+\overline{\mathcal{D}}_i)(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2) + (\widetilde{\mathcal{E}}_i+\overline{\mathcal{E}}_i)(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\hat{X}}(t)\Big\}dW_i(t),\ t\in[0,T], \\
\hat{X}(0) = X_0,
\end{cases}
\tag{47}
$$

and $X(\cdot)$ be the $\mathcal{F}_t$-adapted solution to

$$
\begin{cases}
dX(t) = \Big\{\big(\mathcal{A}_0+\mathcal{B}_0\mathcal{P}_1+\mathcal{B}_1^\top\Sigma_1+\mathcal{B}_2^\top\Sigma_2+\mathcal{B}_3^\top\Sigma_3\big)X(t) + \big(\widehat{\mathcal{A}}_0+\mathcal{B}_0\mathcal{P}_2+\mathcal{B}_1^\top\widehat{\Sigma}_1+\mathcal{B}_2^\top\widehat{\Sigma}_2+\mathcal{B}_3^\top\widehat{\Sigma}_3\big)\hat{X}(t) \\
\qquad + \big[\mathcal{B}_0\mathcal{P}_3+\mathcal{C}_0(\mathcal{P}_1+\mathcal{P}_3)+\mathcal{B}_1^\top\widetilde{\Sigma}_1+\mathcal{B}_2^\top\widetilde{\Sigma}_2+\mathcal{B}_3^\top\widetilde{\Sigma}_3+\widetilde{\mathcal{B}}_1^\top(\Sigma_1+\widetilde{\Sigma}_1)+\widetilde{\mathcal{B}}_2^\top(\Sigma_2+\widetilde{\Sigma}_2) \\
\qquad + \widetilde{\mathcal{B}}_3^\top(\Sigma_3+\widetilde{\Sigma}_3)\big]\check{X}(t) + \big[\overline{\mathcal{A}}_0+\mathcal{B}_0\mathcal{P}_4+\mathcal{C}_0(\mathcal{P}_2+\mathcal{P}_4)+\widetilde{\mathcal{C}}_0(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) \\
\qquad + \mathcal{B}_1^\top\overline{\Sigma}_1+\mathcal{B}_2^\top\overline{\Sigma}_2+\mathcal{B}_3^\top\overline{\Sigma}_3+\widetilde{\mathcal{B}}_1^\top(\widehat{\Sigma}_1+\overline{\Sigma}_1)+\widetilde{\mathcal{B}}_2^\top(\widehat{\Sigma}_2+\overline{\Sigma}_2)+\widetilde{\mathcal{B}}_3^\top(\widehat{\Sigma}_3+\overline{\Sigma}_3) \\
\qquad + \overline{\mathcal{C}}_0(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1)+\overline{\mathcal{D}}_0(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\overline{\mathcal{E}}_0(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\hat{X}}(t)\Big\}dt \\
\quad + \sum_{i=1}^{3}\Big\{\big(\mathcal{A}_i+\mathcal{B}_i\mathcal{P}_1+\mathcal{B}_i\Sigma_1+\mathcal{D}_i\Sigma_2+\mathcal{E}_i\Sigma_3\big)X(t)+\big(\widehat{\mathcal{A}}_i+\mathcal{B}_i\mathcal{P}_2+\mathcal{B}_i\widehat{\Sigma}_1+\mathcal{D}_i\widehat{\Sigma}_2+\mathcal{E}_i\widehat{\Sigma}_3\big)\hat{X}(t) \\
\qquad + \big[\mathcal{B}_i\mathcal{P}_3+\widetilde{\mathcal{B}}_i(\mathcal{P}_1+\mathcal{P}_3)+\mathcal{B}_i\widetilde{\Sigma}_1+\mathcal{D}_i\widetilde{\Sigma}_2+\mathcal{E}_i\widetilde{\Sigma}_3+\widetilde{\mathcal{C}}_i(\Sigma_1+\widetilde{\Sigma}_1)+\widetilde{\mathcal{D}}_i(\Sigma_2+\widetilde{\Sigma}_2)+\widetilde{\mathcal{E}}_i(\Sigma_3+\widetilde{\Sigma}_3)\big]\check{X}(t) \\
\qquad + \big[\overline{\mathcal{A}}_i+\mathcal{B}_i\mathcal{P}_4+\widetilde{\mathcal{B}}_i(\mathcal{P}_2+\mathcal{P}_4)+\overline{\mathcal{B}}_i(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)+\mathcal{B}_i\overline{\Sigma}_1+\mathcal{D}_i\overline{\Sigma}_2+\mathcal{E}_i\overline{\Sigma}_3 \\
\qquad + \widetilde{\mathcal{C}}_i(\widehat{\Sigma}_1+\overline{\Sigma}_1)+\widetilde{\mathcal{D}}_i(\widehat{\Sigma}_2+\overline{\Sigma}_2)+\widetilde{\mathcal{E}}_i(\widehat{\Sigma}_3+\overline{\Sigma}_3)+\overline{\mathcal{C}}_i(\Sigma_1+\widehat{\Sigma}_1+\widetilde{\Sigma}_1+\overline{\Sigma}_1) \\
\qquad + \overline{\mathcal{D}}_i(\Sigma_2+\widehat{\Sigma}_2+\widetilde{\Sigma}_2+\overline{\Sigma}_2)+\overline{\mathcal{E}}_i(\Sigma_3+\widehat{\Sigma}_3+\widetilde{\Sigma}_3+\overline{\Sigma}_3)\big]\check{\hat{X}}(t)\Big\}dW_i(t),\ t\in[0,T], \\
X(0) = X_0.
\end{cases}
\tag{48}
$$

Then the feedback representation for the optimal control $u_2^*(\cdot)$ of the leader is given by

$$
\begin{aligned}
u_2^*(t) = &- N_2^{-1}(t)\Big\{\Big[\mathcal{C}_{05}^\top(\mathcal{P}_1+\mathcal{P}_3)+\sum_{i=1}^{3}\mathcal{C}_{i5}^\top(\Sigma_i+\widetilde{\Sigma}_i)\Big]\check{X}(t)+\Big[\mathcal{L}_4^\top+\mathcal{C}_{05}^\top(\mathcal{P}_2+\mathcal{P}_4) \\
&+\sum_{i=1}^{3}\mathcal{C}_{i5}^\top(\widehat{\Sigma}_i+\overline{\Sigma}_i)+\mathcal{L}_{05}^\top(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)+\sum_{i=1}^{3}\mathcal{L}_{i5}^\top(\Sigma_i+\widehat{\Sigma}_i+\widetilde{\Sigma}_i+\overline{\Sigma}_i)\Big]\check{\hat{X}}(t)\Big\}.
\end{aligned}
\tag{49}
$$

Remark 2.2 In the literature [15], the explicit optimal cost was given in (47) by the solution to some Riccati equations. However, as mentioned in Remark 2.1, in our overlapping information, the completion of squares method is invalid. Up to now, there are no explicit results about this.

However, since the problem in this paper is the LQ case, by **(A2.4)** it is easy to check that the concavity/convexity conditions in the verification theorem of Proposition 2.4 of Shi et al. [23] hold, then $u_2^*(\cdot)$ given by (49) is really optimal.

Finally, the optimal control $u_1^*(\cdot)$ of the follower can also be represented in a "nonanticipating" way. In fact, by (25), noticing (49), (41), (100) and (120) (in the Appendix), we obtain

$$
\begin{aligned}
u_1^*(t)=&-\Big[N_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)B_i(t)\Big]^{-1}\bigg\{\Big[B_0^\top(t)P_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)A_i(t)\Big]\hat{x}^{u_1^*,u_2^*}(t)\\
&+B_0^\top(t)P_1(t)\hat{\phi}^*(t)+B_1^\top(t)\hat{\beta}_1^*(t)+B_3^\top(t)\hat{\beta}_3^*(t)+\Big[\sum_{i=1}^3 B_i^\top(t)P_1(t)C_i(t)\Big]\hat{u}_2^*(t)\bigg\}\\
=&-\Big[N_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)B_i(t)\Big]^{-1}\bigg\{\begin{pmatrix} B_0^\top(t)P_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)A_i(t) & 0\end{pmatrix}\hat{X}(t)\\
&+\begin{pmatrix}0 & B_0^\top(t)P_1(t)\end{pmatrix}\hat{Y}(t)+\begin{pmatrix}0 & B_1^\top(t)\end{pmatrix}\hat{Z}_1(t)+\begin{pmatrix}0 & B_3^\top(t)\end{pmatrix}\hat{Z}_3(t)\\
&-\Big(\sum_{i=1}^3 B_i^\top(t)P_1(t)C_i(t)\Big)N_2^{-1}(t)\Big[\mathcal{C}_{05}^\top(\mathcal{P}_1+\mathcal{P}_3)+\sum_{i=1}^3\mathcal{C}_{i5}^\top(\Sigma_i+\widetilde{\Sigma}_i)+\mathcal{L}_4^\top+\mathcal{C}_{05}^\top(\mathcal{P}_2+\mathcal{P}_4)\\
&+\sum_{i=1}^3\mathcal{C}_{i5}^\top(\widehat{\Sigma}_i+\overline{\Sigma}_i)+\mathcal{L}_{05}^\top(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)+\sum_{i=1}^3\mathcal{L}_{i5}^\top(\Sigma_i+\widehat{\Sigma}_i+\widetilde{\Sigma}_i+\overline{\Sigma}_i)\Big]\check{\hat{X}}(t)\bigg\}\\
=&-\Big[N_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)B_i(t)\Big]^{-1}\bigg[\begin{pmatrix} B_0^\top(t)P_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)A_i(t) & 0\end{pmatrix}\\
&+\begin{pmatrix}0 & B_0^\top(t)P_1(t)\end{pmatrix}(\mathcal{P}_1+\mathcal{P}_2)+\begin{pmatrix}0 & B_1^\top(t)\end{pmatrix}(\Sigma_1+\widehat{\Sigma}_1)+\begin{pmatrix}0 & B_3^\top(t)\end{pmatrix}(\Sigma_3+\widetilde{\Sigma}_3)\bigg]\hat{X}(t)\\
&-\Big[N_1(t)+\sum_{i=1}^3 B_i^\top(t)P_1(t)B_i(t)\Big]^{-1}\bigg\{\begin{pmatrix}0 & B_0^\top(t)P_1(t)\end{pmatrix}(\mathcal{P}_3+\mathcal{P}_4)+\begin{pmatrix}0 & B_1^\top(t)\end{pmatrix}(\widetilde{\Sigma}_1+\overline{\Sigma}_1)\\
&+\begin{pmatrix}0 & B_3^\top(t)\end{pmatrix}(\widetilde{\Sigma}_3+\overline{\Sigma}_3)-N_2^{-1}(t)\Big(\sum_{i=1}^3 B_i^\top(t)P_1(t)C_i(t)\Big)\Big[\mathcal{L}_4^\top+(\mathcal{C}_{05}+\mathcal{L}_{05})^\top(\mathcal{P}_1+\mathcal{P}_2\\
&+\mathcal{P}_3+\mathcal{P}_4)+\sum_{i=1}^3(\mathcal{C}_{i5}+\mathcal{L}_{i5})^\top(\Sigma_i+\widetilde{\Sigma}_i+\widehat{\Sigma}_i+\overline{\Sigma}_i)\Big]\bigg\}\check{\hat{X}}(t),
\end{aligned}
\tag{50}
$$

which is observable for the follower.

Up to now, the Stackelberg equilibrium strategy $(u_1^*(\cdot),u_2^*(\cdot))$ is obtained, which is represented as the state estimate feedback form in (50) and (49).

3. A special case: Control independent diffusions

In this section, we consider the problem for the special $n=1$ case with control independent diffusions and constant parameters. In this case, the problem can be solved and the system of Riccati equations has a decoupling structure.

We consider the scalar state process $x^{u_1,u_2}(\cdot)$ which satisfies the linear SDE

$$
\begin{cases}
dx^{u_1,u_2}(t)=\big[A_0x^{u_1,u_2}(t)+B_0u_1(t)+C_0u_2(t)\big]dt+A_1x^{u_1,u_2}(t)dW_1(t)\\
\qquad\qquad\quad +A_2x^{u_1,u_2}(t)dW_2(t)+A_3x^{u_1,u_2}(t)dW_3(t),\ t\in[0,T],\\
x^{u_1,u_2}(0)=x_0.
\end{cases}
\tag{51}
$$

Here $u_1(\cdot)$ and $u_2(\cdot)$ are both scalar-valued and A_0,B_0,C_0,A_1,A_2,A_3 are constants. We define the admissible control sets $\mathcal{U}_1,\mathcal{U}_2$ as in Section 2.

In step 1, for any chosen $u_2(\cdot)$, the follower wishes to select a $u_1^*(\cdot)\in\mathcal{U}_1$ to minimize the cost functional

$$J_1(u_1(\cdot),u_2(\cdot))=\frac{1}{2}\mathbb{E}\bigg[\int_0^T\Big(Q_1\big|x^{u_1,u_2}(t)\big|^2+N_1u_1^2(t)\Big)dt+G_1\big|x^{u_1,u_2}(T)\big|^2\bigg]. \tag{52}$$

Here $Q_1,G_1\geq 0, N_1>0$ are constants. In step 2, after the follower's optimal control $u_1^*(\cdot)$ is announced, the leader would like to choose a $u_2^*(\cdot)\in\mathcal{U}_2$ to minimize

$$J_2(u_1^*(\cdot),u_2(\cdot))=\frac{1}{2}\mathbb{E}\bigg[\int_0^T\Big(Q_2\big|x^{u_1^*,u_2}(t)\big|^2+N_2u_2^2(t)\Big)dt+G_2\big|x^{u_1^*,u_2}(T)\big|^2\bigg], \tag{53}$$

where $Q_2,G_2\geq 0, N_2>0$ are constants. We wish to find the Stackelberg equilibrium strategy $(u_1^*(\cdot),u_2^*(\cdot))\in\mathcal{U}_1\times\mathcal{U}_2$.

3.1. **Problem of The Follower**

For given control $u_2(\cdot)$, let $u_1^*(\cdot)$ be a $\mathcal{G}_t^1$-adapted optimal control of the follower, and the corresponding optimal state be $x^{u_1^*,u_2}(\cdot)$. Now the follower's Hamiltonian function (13) writes

$$H_1\big(t,x,u_1,u_2,q,k_1,k_2,k_3\big):=q(A_0x+B_0u_1+C_0u_2)+A_1k_1x+A_2k_2x+A_3k_3x-\frac{1}{2}Q_1x^2-\frac{1}{2}N_1u_1^2. \tag{54}$$

And (14) yields that

$$0=N_1u_1^*(t)-B_0\hat{q}(t), \tag{55}$$

where the $\mathcal{F}_t$-adapted process quadruple $(q(\cdot),k_1(\cdot),k_2(\cdot),k_3(\cdot))$ satisfies the adjoint BSDE

$$\begin{cases}-dq(t)=\big[A_0q(t)+A_1k_1+A_2k_2+A_3k_3-Q_1x^{u_1^*,u_2}(t)\big]dt-k_1dW_1(t)\\ \qquad\qquad -k_2dW_2(t)-k_3dW_3(t),\ t\in[0,T],\\ q(T)=-G_1x^{u_1^*,u_2}(T),\end{cases} \tag{56}$$

which is a special case of (15). Repeating the same approach as in Section 2.1, we obtain the following theorem.

Theorem 3.1 *Let $P(\cdot)$ satisfy*

$$\begin{cases}\dot{P}(t)+\big(2A_0+A_1^2+A_2^2+A_3^2\big)P(t)-N_1^{-1}B_0^2P^2(t)+Q_1=0,\ t\in[0,T],\\ P(T)=G_1.\end{cases} \tag{57}$$

For chosen $u_2(\cdot)$ of the leader, $u_1^*(\cdot)$ defined by the following is a feedback representation for the optimal control of the follower:

$$u_1^*(t)=-N_1^{-1}B_0\big[P(t)\hat{x}^{u_1^*,u_2}(t)+\hat{\phi}(t)\big], \tag{58}$$

where $(\hat{x}^{u_1^*,u_2}(\cdot),\hat{\phi}(\cdot),\hat{\beta}_1(\cdot),\hat{\beta}_3(\cdot))$ is the unique $\mathcal{G}_t^1$-adapted solution to

$$\begin{cases}d\hat{x}^{u_1^*,u_2}(t)=\big[\big(A_0-N_1^{-1}B_0^2P(t)\big)\hat{x}^{u_1^*,u_2}(t)-N_1^{-1}B_0^2\hat{\phi}(t)+C_0\hat{u}_2(t)\big]dt\\ \qquad\qquad +A_1\hat{x}^{u_1^*,u_2}(t)dW_1(t)+A_3\hat{x}^{u_1^*,u_2}(t)dW_3(t),\\ -d\hat{\phi}(t)=\big[\big(A_0-N_1^{-1}B_0^2P(t)\big)\hat{\phi}(t)+A_1\hat{\beta}_1(t)+A_3\hat{\beta}_3(t)+P(t)C_0\hat{u}_2(t)\big]dt\\ \qquad\qquad -\hat{\beta}_1(t)dW_1(t)-\hat{\beta}_3(t)dW_3(t),\ t\in[0,T],\\ \hat{x}^{u_1,u_2}(0)=x_0,\quad \hat{\phi}(T)=0.\end{cases} \tag{59}$$

3.2. **Problem of The Leader**

The leader keeps in mind that the follower takes $u_1^*(\cdot)$ by (58), then his state equation (31) writes

$$\begin{cases} dx^{u_2}(t) = \big[A_0 x^{u_2}(t) - N_1^{-1}B_0^2 P(t)\hat{x}^{u_2}(t) - N_1^{-1}B_0^2\hat{\phi}(t) + C_0 u_2(t)\big]dt \\ \qquad\qquad + A_1 x^{u_2}(t)dW_1(t) + A_2 x^{u_2}(t)dW_2(t) + A_3 x^{u_2}(t)dW_3(t), \\ -d\hat{\phi}(t) = \Big\{\big[A_0 - N_1^{-1}B_0^2 P(t)\big]\hat{\phi}(t) + A_1\hat{\beta}_1(t) + A_3\hat{\beta}_3(t) + P(t)C_0\hat{u}_2(t)\Big\}dt \\ \qquad\qquad - \hat{\beta}_1(t)dW_1(t) - \hat{\beta}_3(t)dW_3(t),\ t\in[0,T], \\ x^{u_2}(0) = x_0, \quad \hat{\phi}(T) = 0. \end{cases} \tag{60}$$

The problem of the leader is to select a $\mathcal{G}_t^2$-adapted optimal control $u_2^*(\cdot)$ such that the cost functional

$$J_2(u_2(\cdot)) = \frac{1}{2}\mathbb{E}\Big[\int_0^T \big[Q_2|x^{u_2}(t)|^2 + N_2 u_2^2(t)\big]dt + G_2|x^{u_2}(T)|^2\Big] \tag{61}$$

is minimized.

Suppose that there exists a $\mathcal{G}_t^2$-adapted optimal control $u_2^*(\cdot)$ of the leader, and his optimal state is $(x^*(\cdot),\hat{\phi}^*(\cdot),\hat{\beta}_1^*(\cdot),\hat{\beta}_3^*(\cdot)) \equiv (x^{u_2^*}(\cdot),\hat{\phi}^*(\cdot),\hat{\beta}_1^*(\cdot),\hat{\beta}_3^*(\cdot))$. Now the leader's Hamiltonian function (34) reduces to

$$\begin{aligned} &H_2\big(t,x^{u_2},u_2,\phi,\beta_1,\beta_3;p,y,z_1,z_2,z_3\big) := y\big[A_0x^{u_2} - N_1^{-1}B_0^2P(t)\hat{x}^{u_2} - N_1^{-1}B_0^2\hat{\phi} + C_0u_2\big] + p\big\{A_1\hat{\beta}_1 \\ &\quad + A_3\hat{\beta}_3 + [A_0 - N_1^{-1}B_0^2P(t)]\hat{\phi} + P(t)C_0\hat{u}_2\big\} + z_1A_1x^{u_2} + z_2A_2x^{u_2} + z_3A_3x^{u_2} + \frac{1}{2}\big[Q_2|x^{u_2}|^2 + N_2u_2^2\big], \end{aligned} \tag{62}$$

where the $\mathcal{F}_t$-adapted process quintuple $(p(\cdot),y(\cdot),z_1(\cdot),z_2(\cdot),z_3(\cdot))$ satisfies the adjoint equation

$$\begin{cases} dp(t) = \Big\{-N_1^{-1}B_0^2y(t) + \big[A_0 - N_1^{-1}B_0^2P(t)\big]p(t)\Big\}dt + A_1p(t)dW_1(t) + A_3p(t)dW_3(t), \\ -dy(t) = \big[A_0y(t) - N_1^{-1}B_0^2P(t)\hat{y}(t) + A_1z_1(t) + A_2z_2(t) + A_3z_3(t) + Q_2x^*(t)\big]dt \\ \qquad\qquad - z_1(t)dW_1(t) - z_2(t)dW_2(t) - z_3(t)dW_3(t),\ t\in[0,T], \\ p(0) = 0, \quad y(T) = G_2x^*(T), \end{cases} \tag{63}$$

which is a special case of (35). Similarly, we have

$$u_2^*(t) = -N_2^{-1}C_0\big[\check{y}(t) + P(t)\check{p}(t)\big], \tag{64}$$

where

$$\begin{cases} d\check{x}^*(t) = \big[A_0\check{x}^*(t) - N_1^{-1}B_0^2P(t)\check{\hat{x}}^*(t) - N_1^{-1}B_0^2\check{\hat{\phi}}^*(t) - N_2^{-1}C_0^2\check{y}(t) - N_2^{-1}C_0^2P(t)\check{p}(t)\big]dt \\ \qquad\qquad + A_2\check{x}^*(t)dW_2(t) + A_3\check{x}^*(t)dW_3(t), \\ d\check{p}(t) = \Big\{-N_1^{-1}B_0^2\check{y}(t) + \big[A_0 - N_1^{-1}B_0^2P(t)\big]\check{p}(t)\Big\}dt + A_3\check{p}(t)dW_3(t), \\ -d\check{y}(t) = \big[A_0\check{y}(t) - N_1^{-1}B_0^2P(t)\check{\hat{y}}(t) + A_1\check{z}_1(t) + A_2\check{z}_2(t) + A_3\check{z}_3(t) + Q_2\check{x}^*(t)\big]dt \\ \qquad\qquad - \check{z}_2(t)dW_2(t) - \check{z}_3(t)dW_3(t), \\ -d\check{\hat{\phi}}^*(t) = \Big\{\big[A_0 - N_1^{-1}B_0^2P(t)\big]\check{\hat{\phi}}^*(t) + A_1\check{\hat{\beta}}_1^*(t) + A_3\check{\hat{\beta}}_3^*(t) - N_2^{-1}C_0^2P(t)\check{\hat{y}}(t) \\ \qquad\qquad - N_2^{-1}C_0^2P^2(t)\check{\hat{p}}(t)\Big\}dt - \check{\hat{\beta}}_3^*(t)dW_3(t),\ t\in[0,T], \\ \check{x}^*(0) = x_0, \quad \check{p}(0) = 0, \quad \check{y}(T) = G_2\check{x}^*(T), \quad \check{\hat{\phi}}^*(T) = 0, \end{cases} \tag{65}$$

and

$$
\begin{cases}
d\check{\hat{x}}^*(t) = \Big\{\big[A_0 - N_1^{-1}B_0^2P(t)\big]\check{\hat{x}}^*(t) - N_1^{-1}B_0^2\check{\hat{\phi}}^*(t) - N_2^{-1}C_0^2\check{y}(t) - N_2^{-1}C_0^2P(t)\check{p}(t)\Big\}dt \\
\qquad\qquad + A_3\check{\hat{x}}^*(t)dW_3(t), \\
d\check{p}(t) = \Big\{ - N_1^{-1}B_0^2\check{y}(t) + \big[A_0 - N_1^{-1}B_0^2P(t)\big]\check{p}(t)\Big\}dt + A_3\check{p}(t)dW_3(t), \\
-d\check{y}(t) = \Big\{\big[A_0 - N_1^{-1}B_0^2P(t)\big]\check{y}(t) + A_1\check{\hat{z}}_1(t) + A_2\check{\hat{z}}_2(t) + A_3\check{\hat{z}}_3(t) + Q_2\check{\hat{x}}^*(t)\Big\}dt \\
\qquad\qquad - \check{\hat{z}}_3(t)dW_3(t),\ t\in[0,T], \\
\check{\hat{x}}^*(0) = x_0, \quad \check{p}(0) = 0, \quad \check{y}(T) = G_2\check{\hat{x}}^*(T).
\end{cases}
\tag{66}
$$

The solvability of (63), (65) and (66) will be analyzed in the following context. As in Section 2, we will proceed to represent $u_2^*(\cdot)$ of (64) as the state estimate feedback form, via some Riccati type equations.

First, we rewrite the optimal state equation (60) as

$$
\begin{cases}
dx^*(t) = \big[A_0x^*(t) - N_1^{-1}B_0^2P(t)\hat{x}^*(t) - N_1^{-1}B_0^2\hat{\phi}^*(t) - N_2^{-1}C_0^2\check{y}(t) - N_2^{-1}C_0^2P(t)\check{p}(t)\big]dt \\
\qquad\qquad + A_1x^*(t)dW_1(t) + A_2x^*(t)dW_2(t) + A_3x^*(t)dW_3(t), \\
-d\hat{\phi}^*(t) = \Big\{\big[A_0 - N_1^{-1}B_0^2P(t)\big]\hat{\phi}^*(t) + A_1\hat{\beta}_1^*(t) + A_3\hat{\beta}_3^*(t) - N_2^{-1}C_0^2P(t)\check{y}(t) - N_2^{-1}C_0^2P^2(t)\check{p}(t)\Big\}dt \\
\qquad\qquad - \hat{\beta}_1^*(t)dW_1(t) - \hat{\beta}_3(t)^*dW_3(t),\ t\in[0,T], \\
x^*(0) = x_0, \quad \hat{\phi}^*(T) = 0.
\end{cases}
\tag{67}
$$

Define X, Y, Z_1, Z_2, Z_3 as (41) and

$$
\begin{cases}
\mathcal{A}_0 := \begin{pmatrix} A_0 & 0 \\ 0 & A_0 - N_1^{-1}B_0^2P(t) \end{pmatrix},\ \mathcal{A}_1 := \begin{pmatrix} A_1 & 0 \\ 0 & A_1 \end{pmatrix},\ \mathcal{A}_2 := \begin{pmatrix} A_2 & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{A}_3 := \begin{pmatrix} A_3 & 0 \\ 0 & A_3 \end{pmatrix},\ \mathcal{B}_0 := \begin{pmatrix} 0 & -N_1^{-1}B_0^2 \\ -N_1^{-1}B_0^2 & 0 \end{pmatrix},\ \overline{\mathcal{B}}_0 := \begin{pmatrix} -N_1^{-1}B_0^2P(t) & 0 \\ 0 & 0 \end{pmatrix}, \\
\mathcal{C}_0 := \begin{pmatrix} -N_2^{-1}C_0 & 0 \\ 0 & 0 \end{pmatrix}, \widetilde{\mathcal{C}}_0 := \begin{pmatrix} 0 & -N_2^{-1}C_0^2P(t) \\ 0 & 0 \end{pmatrix},\ \widehat{\mathcal{C}}_0 := \begin{pmatrix} 0 & 0 \\ -N_2^{-1}C_0^2P(t) & 0 \end{pmatrix}, \\
\overline{\mathcal{C}}_0 := \begin{pmatrix} 0 & 0 \\ 0 & -N_2^{-1}C_0^2P^2(t) \end{pmatrix},\ \mathcal{Q}_2 := \begin{pmatrix} Q_2 & 0 \\ 0 & 0 \end{pmatrix},\ \mathcal{G}_2 := \begin{pmatrix} G_2 & 0 \\ 0 & 0 \end{pmatrix},\ X_0 := \begin{pmatrix} x_0 \\ 0 \end{pmatrix},
\end{cases}
$$

then we have

$$
u_2^*(t) = -N_2^{-1}\Big[\begin{pmatrix} C_0 & 0 \end{pmatrix}\check{Y}(t) + \begin{pmatrix} C_0P(t) & 0 \end{pmatrix}\check{\hat{X}}(t)\Big],
\tag{68}
$$

and

$$
\begin{cases}
dX(t) = \big[\mathcal{A}_0X(t) + \overline{\mathcal{B}}_0\hat{X}(t) + \widetilde{\mathcal{C}}_0\check{\hat{X}}(t) + \mathcal{B}_0Y(t) + \mathcal{C}_0\check{Y}(t)\big]dt + \mathcal{A}_1X(t)dW_1(t) \\
\qquad\qquad + \mathcal{A}_2X(t)dW_2(t) + \mathcal{A}_3X(t)dW_3(t), \\
-dY(t) = \big[\mathcal{Q}_2X(t) + \mathcal{A}_0Y(t) + \overline{\mathcal{B}}_0\hat{Y}(t) + \widehat{\mathcal{C}}_0\check{\hat{Y}}(t) + \overline{\mathcal{C}}_0\check{\hat{X}}(t) + \mathcal{A}_1Z_1(t) \\
\qquad\qquad + \mathcal{A}_2Z_2(t) + \mathcal{A}_3Z_3(t)\big]dt - Z_1dW_1(t) - Z_2dW_2(t) - Z_3dW_3(t),\ t\in[0,T], \\
X(0) = X_0, \quad Y(T) = \mathcal{G}_2X(T),
\end{cases}
\tag{69}
$$

where the equations for $\hat{X}(\cdot), \check{X}(\cdot), \check{\hat{X}}(\cdot)$ are

$$
\begin{cases}
d\hat{X}(t) = \big[(\mathcal{A}_0 + \overline{\mathcal{B}}_0)\hat{X}(t) + \widetilde{\mathcal{C}}_0\check{\hat{X}}(t) + \mathcal{B}_0\hat{Y}(t) + \mathcal{C}_0\check{\hat{Y}}(t)\big]dt + \mathcal{A}_1\hat{X}(t)dW_1(t) + \mathcal{A}_3\hat{X}(t)dW_3(t),\ t\in[0,T], \\
\hat{X}(0) = X_0,
\end{cases}
\tag{70}
$$

$$\begin{cases} d\check{X}(t)=\big[\mathcal{A}_0\check{X}(t)+(\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0)\check{\check{X}}(t)+(\mathcal{B}_0+\mathcal{C}_0)\check{Y}(t)\big]dt+\mathcal{A}_2\check{X}(t)dW_2(t)+\mathcal{A}_3\check{X}(t)dW_3(t),\ t\in[0,T], \\ \check{X}(0)=X_0, \end{cases} \tag{71}$$

and

$$\begin{cases} d\check{\check{X}}(t)=\big[(\mathcal{A}_0+\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0)\check{\check{X}}(t)+(\mathcal{B}_0+\mathcal{C}_0)\check{\check{Y}}(t)\big]dt+\mathcal{A}_3\check{\check{X}}(t)dW_3(t),\ t\in[0,T], \\ \check{\check{X}}(0)=X_0, \end{cases} \tag{72}$$

respectively. Defining $\mathcal{P}_1(\cdot),\mathcal{P}_2(\cdot),\mathcal{P}_3(\cdot),\mathcal{P}_4(\cdot)$ as (100) in the Appendix, by Itô's formula we obtain

$$\begin{aligned} dY(t)=&\Big\{\big(\dot{\mathcal{P}}_1+\mathcal{P}_1\mathcal{A}_0+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_1\big)X(t)+\big[\dot{\mathcal{P}}_2+\mathcal{P}_2(\mathcal{A}_0+\overline{\mathcal{B}}_0)+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_1+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_2+\mathcal{P}_1\overline{\mathcal{B}}_0\big]\hat{X}(t)\\ &+\big[\dot{\mathcal{P}}_3+\mathcal{P}_3\mathcal{A}_0+\mathcal{P}_3(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_3)+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_3+\mathcal{P}_1\mathcal{C}_0\mathcal{P}_3+\mathcal{P}_1\mathcal{C}_0\mathcal{P}_1\big]\check{X}(t)\\ &+\big[\dot{\mathcal{P}}_4+\mathcal{P}_4(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3)+(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3)(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_4+\mathcal{P}_4\big(\mathcal{A}_0+\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0\big)\\ &+\mathcal{P}_4(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_4+(\mathcal{P}_1+\mathcal{P}_2)\widetilde{\mathcal{C}}_0+\mathcal{P}_1\mathcal{C}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_1+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_3+\mathcal{P}_3\mathcal{B}_0\mathcal{P}_2\\ &+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_3+\mathcal{P}_3\big(\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0\big)+\mathcal{P}_3\mathcal{C}_0\mathcal{P}_2\big]\check{\check{X}}(t)\Big\}dt+\big[\mathcal{P}_1\mathcal{A}_1X(t)+\mathcal{P}_2\mathcal{A}_1\hat{X}(t)\big]dW_1(t)\\ &+\big[\mathcal{P}_1\mathcal{A}_2X(t)+\mathcal{P}_3\mathcal{A}_2\check{X}(t)\big]dW_2(t)+\big[\mathcal{P}_1\mathcal{A}_3X(t)+\mathcal{P}_2\mathcal{A}_3\hat{X}(t)+\mathcal{P}_3\mathcal{A}_3\check{X}(t)+\mathcal{P}_4\mathcal{A}_3\check{\check{X}}(t)\big]dW_3(t)\\ =&-\Big\{\big(\mathcal{Q}_2+\mathcal{A}_0\mathcal{P}_1\big)X(t)+\big(\mathcal{A}_0\mathcal{P}_2+\overline{\mathcal{B}}_0\mathcal{P}_1+\overline{\mathcal{B}}_0\mathcal{P}_2\big)\hat{X}(t)+\mathcal{A}_0\mathcal{P}_3\check{X}(t)+\big[\mathcal{A}_0\mathcal{P}_4+\overline{\mathcal{B}}_0\mathcal{P}_3+\overline{\mathcal{B}}_0\mathcal{P}_4+\overline{\mathcal{C}}_0\\ &+\widehat{\mathcal{C}}_0(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\big]\check{\check{X}}(t)+\mathcal{A}_1Z_1(t)+\mathcal{A}_2Z_2(t)+\mathcal{A}_3Z_3(t)\Big\}dt+Z_1(t)dW_1(t)\\ &+Z_2(t)dW_2(t)+Z_3(t)dW_3(t). \end{aligned} \tag{73}$$

Comparing the diffusion terms on both sides of (73), we directly get

$$\begin{cases} Z_1(t)=\mathcal{P}_1\mathcal{A}_1X(t)+\mathcal{P}_2\mathcal{A}_1\hat{X}(t),\quad Z_2(t)=\mathcal{P}_1\mathcal{A}_2X(t)+\mathcal{P}_3\mathcal{A}_2\check{X}(t),\\ Z_3(t)=\mathcal{P}_1\mathcal{A}_3X(t)+\mathcal{P}_2\mathcal{A}_3\hat{X}(t)+\mathcal{P}_3\mathcal{A}_3\check{X}(t)+\mathcal{P}_4\mathcal{A}_3\check{\check{X}}(t). \end{cases} \tag{74}$$

It is worth pointing out that, comparing with the four steps in the control-dependent case of Section 2, the current case is rather simple to obtain (74). Similarly, we obtain

$$\begin{cases} 0=\dot{\mathcal{P}}_1+\mathcal{P}_1\mathcal{A}_0+\mathcal{A}_0\mathcal{P}_1+\mathcal{A}_1\mathcal{P}_1\mathcal{A}_1+\mathcal{A}_2\mathcal{P}_1\mathcal{A}_2+\mathcal{A}_3\mathcal{P}_1\mathcal{A}_3+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_1+\mathcal{Q}_2,\\ 0=\dot{\mathcal{P}}_2+\mathcal{P}_2(\mathcal{A}_0+\overline{\mathcal{B}}_0)+(\mathcal{A}_0+\overline{\mathcal{B}}_0)\mathcal{P}_2+\mathcal{A}_1\mathcal{P}_2\mathcal{A}_1+\mathcal{A}_3\mathcal{P}_2\mathcal{A}_3+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_1\\ \qquad+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_2+\mathcal{P}_1\overline{\mathcal{B}}_0+\overline{\mathcal{B}}_0\mathcal{P}_1,\\ 0=\dot{\mathcal{P}}_3+\mathcal{P}_3\mathcal{A}_0+\mathcal{A}_0\mathcal{P}_3+\mathcal{A}_2\mathcal{P}_3\mathcal{A}_2+\mathcal{A}_3\mathcal{P}_3\mathcal{A}_3+\mathcal{P}_3(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_1+\mathcal{P}_1(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_3\\ \qquad+\mathcal{P}_3(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_3+\mathcal{P}_1\mathcal{C}_0\mathcal{P}_1,\\ 0=\dot{\mathcal{P}}_4+\mathcal{P}_4\big(\mathcal{A}_0+\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0\big)+\big(\mathcal{A}_0+\overline{\mathcal{B}}_0+\widehat{\mathcal{C}}_0\big)\mathcal{P}_4+\mathcal{A}_3\mathcal{P}_4\mathcal{A}_3+\mathcal{P}_4(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3)+(\mathcal{P}_1\\ \qquad+\mathcal{P}_2+\mathcal{P}_3)(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_4+\mathcal{P}_4(\mathcal{B}_0+\mathcal{C}_0)\mathcal{P}_4+\mathcal{P}_3\big(\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0\big)+(\overline{\mathcal{B}}_0+\widehat{\mathcal{C}}_0)\mathcal{P}_3+\mathcal{P}_2\mathcal{B}_0\mathcal{P}_3+\mathcal{P}_3\mathcal{B}_0\mathcal{P}_2\\ \qquad+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_3+\mathcal{P}_3\mathcal{C}_0\mathcal{P}_2+(\mathcal{P}_1+\mathcal{P}_2)\widetilde{\mathcal{C}}_0+\widehat{\mathcal{C}}_0(\mathcal{P}_1+\mathcal{P}_2)+\mathcal{P}_1\mathcal{C}_0\mathcal{P}_2+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_1+\mathcal{P}_2\mathcal{C}_0\mathcal{P}_2+\overline{\mathcal{C}}_0,\\ \mathcal{P}_1(T)=\mathcal{G}_2,\ \mathcal{P}_2(T)=0,\ \mathcal{P}_3(T)=0,\ \mathcal{P}_4(T)=0. \end{cases} \tag{75}$$

In this case, the above system of Riccati equations is decoupling! So we can solve firstly $\mathcal{P}_1(\cdot)$, then $\mathcal{P}_2(\cdot)$, thirdly $\mathcal{P}_3(\cdot)$ and finally $\mathcal{P}_4(\cdot)$. However, since the first equation for $\mathcal{P}_1(\cdot)$ is not standard, we still leave the solvability of (75) open in this paper.

We have the following theorem.

Theorem 3.2 *Let $(\mathcal{P}_1(\cdot),\mathcal{P}_2(\cdot),\mathcal{P}_3(\cdot),\mathcal{P}_4(\cdot))$ satisfy (75), $\check{\hat{X}}(\cdot)$ be the $\mathcal{G}_t^1\cap\mathcal{G}_t^2$-adapted solution to*

$$\begin{cases} d\check{\hat{X}}(t)=\big[\mathcal{A}_0+\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0+(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\big]\check{\hat{X}}(t)dt+\mathcal{A}_3\check{\hat{X}}(t)dW_3(t),\ t\in[0,T], \\ \check{\hat{X}}(0)=X_0, \end{cases} \tag{76}$$

$\check{X}(\cdot)$ be the $\mathcal{G}_t^2$-adapted solution to

$$\begin{cases} d\check{X}(t)=\Big\{\big[\mathcal{A}_0+(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_3)\big]\check{X}(t)+\big[\overline{\mathcal{B}}_0+\widetilde{\mathcal{C}}_0+(\mathcal{B}_0+\mathcal{C}_0)(\mathcal{P}_1+\mathcal{P}_3)\big]\check{\hat{X}}(t)\Big\}dt \\ \qquad\qquad +\mathcal{A}_2\check{X}(t)dW_2(t)+\mathcal{A}_3\check{X}(t)dW_3(t),\ t\in[0,T], \\ \check{X}(0)=X_0, \end{cases} \tag{77}$$

$\hat{X}(\cdot)$ be the $\mathcal{G}_t^1$-adapted solution to

$$\begin{cases} d\hat{X}(t)=\Big\{\big[\mathcal{A}_0+\overline{\mathcal{B}}_0+\mathcal{B}_0(\mathcal{P}_1+\mathcal{P}_2)\big]\hat{X}(t)+\big[\widetilde{\mathcal{C}}_0+\mathcal{B}_0(\mathcal{P}_1+\mathcal{P}_2)+\mathcal{C}_0(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\big]\check{\hat{X}}(t)\Big\}dt \\ \qquad\qquad +\mathcal{A}_1\hat{X}(t)dW_1(t)+\mathcal{A}_3\hat{X}(t)dW_3(t),\ t\in[0,T], \\ \hat{X}(0)=X_0, \end{cases} \tag{78}$$

and $X(\cdot)$ be the $\mathcal{F}_t$-adapted solution to

$$\begin{cases} dX(t)=\Big\{(\mathcal{A}_0+\mathcal{B}_0\mathcal{P}_1)X(t)+(\overline{\mathcal{B}}_0+\mathcal{B}_0\mathcal{P}_2)\hat{X}(t)+\big[\mathcal{B}_0\mathcal{P}_3+\mathcal{C}_0(\mathcal{P}_1+\mathcal{P}_3)\big]\check{X}(t)+\big[\widetilde{\mathcal{C}}_0+\mathcal{B}_0\mathcal{P}_4 \\ \qquad\qquad +\mathcal{C}_0(\mathcal{P}_2+\mathcal{P}_4)\big]\check{\hat{X}}(t)\Big\}dt+\mathcal{A}_1X(t)dW_1(t)+\mathcal{A}_2X(t)dW_2(t)+\mathcal{A}_3X(t)dW_3(t),\ t\in[0,T], \\ X(0)=X_0. \end{cases} \tag{79}$$

Then the feedback representation for the optimal control $u_2^*(\cdot)$ of the leader is given by

$$\begin{aligned} u_2^*(t)=&-N_2^{-1}\begin{pmatrix} C_0P(t) & 0 \end{pmatrix}(\mathcal{P}_1+\mathcal{P}_3)\check{X}(t) \\ &-N_2^{-1}\Big[\begin{pmatrix} C_0P(t) & 0 \end{pmatrix}(\mathcal{P}_2+\mathcal{P}_4)+\begin{pmatrix} C_0P(t) & 0 \end{pmatrix}\Big]\check{\hat{X}}(t). \end{aligned} \tag{80}$$

Proof. The conclusion is easily obtained from (100), (70), (71), (72), (69), (68). □

Finally, for the follower, by (58), noticing (41) and (100), we obtain

$$\begin{aligned} u_1^*(t)=&-N_1^{-1}B_0\big[P(t)\hat{x}^*(t)+\hat{\phi}^*(t)\big]=-N_1^{-1}\Big[\begin{pmatrix} B_0P(t) & 0 \end{pmatrix}\hat{X}(t)+\begin{pmatrix} 0 & B_0 \end{pmatrix}\hat{Y}(t)\Big] \\ =&-N_1^{-1}\Big[\begin{pmatrix} B_0P(t) & 0 \end{pmatrix}+\begin{pmatrix} 0 & B_0 \end{pmatrix}\big(\mathcal{P}_1+\mathcal{P}_2\big)\Big]\hat{X}(t)-N_1^{-1}\begin{pmatrix} 0 & B_0 \end{pmatrix}\big(\mathcal{P}_3+\mathcal{P}_4\big)\check{\hat{X}}(t). \end{aligned} \tag{81}$$

4. A continuous-time principal-agent problem

This section is devoted to studying the continuous-time principal-agent problem with overlapping information (Example 1.1 of Section 1), which naturally motivates the research for the problem in previous sections. The financial framework is a generalization of the work by Williams [31].

In order to apply the results in Section 3, we define $X:=(y,m)^\top$ and then

$$\begin{cases} dX(t)=\big[\widetilde{r}X(t)+\widetilde{B}e(t)+\alpha_1c(t)+\alpha_2s(t)+\alpha_3d(t)\big]dt+\widetilde{\sigma}_1dW_1(t)+\widetilde{\sigma}_2dW_2(t)+\widetilde{\sigma}_3dW_3(t),\ t\in[0,T], \\ X(0)=X_0\in\mathbb{R}^2, \end{cases} \tag{82}$$

and

$$
\begin{aligned}
J_1(e(\cdot),c(\cdot),s(\cdot),d(\cdot)) &= \frac{1}{2}\mathbb{E}\bigg[\int_0^T \big[c^2(t)-e^2(t)+\langle \widetilde{G}_1X(t),X(t)\rangle\big]dt+\langle \widetilde{G}_1X(T),X(T)\rangle\bigg],\\
J_2(e(\cdot),c(\cdot),s(\cdot),d(\cdot)) &= \frac{1}{2}\mathbb{E}\bigg[\int_0^T \big[d^2(t)-s^2(t)+\langle \widetilde{G}_2X(t),X(t)\rangle\big]dt+\langle \widetilde{G}_2X(T),X(T)\rangle\bigg],
\end{aligned}
\tag{83}
$$

where

$$
\begin{cases}
X_0:=\begin{pmatrix} y_0\\ m_0\end{pmatrix},\ \widetilde{r}:=\begin{pmatrix} r & 0\\ 0 & r\end{pmatrix},\ \widetilde{B}:=\begin{pmatrix} B\\ 0\end{pmatrix},\ \alpha_1:=\begin{pmatrix} 0\\ -1\end{pmatrix},\ \alpha_2:=\begin{pmatrix} -1\\ 1\end{pmatrix},\ \alpha_3:=\begin{pmatrix} -1\\ 0\end{pmatrix},\\
\widetilde{\sigma}_1:=\begin{pmatrix} \sigma_1\\ \bar{\sigma}_1\end{pmatrix},\ \widetilde{\sigma}_2:=\begin{pmatrix} \sigma_2\\ \bar{\sigma}_2\end{pmatrix},\ \widetilde{\sigma}_3:=\begin{pmatrix} \sigma_3\\ \bar{\sigma}_3\end{pmatrix},\ \widetilde{G}_1:=\begin{pmatrix} 0 & 0\\ 0 & 1\end{pmatrix},\ \widetilde{G}_2:=\begin{pmatrix} 1 & 0\\ 0 & 0\end{pmatrix}.
\end{cases}
$$

For the follower (agent)'s problem, first the leader (principal) announces his control $s(\cdot),d(\cdot)$. Following the step in Section 3.1, we have

$$
e^*(t)=\widetilde{B}^\top\big[P(t)\hat{X}(t)+\Phi(t)\big],\ c^*(t)=-\alpha_1^\top\big[P(t)\hat{X}(t)+\Phi(t)\big],
\tag{84}
$$

where 2×2-matrix-valued function $P(\cdot)$ satisfies

$$
\dot{P}(t)+P(t)\widetilde{r}+\widetilde{r}^\top P(t)+P(t)\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)P(t)^\top+\widetilde{G}_1=0,\ t\in[0,T],\ P(T)=\widetilde{G}_1,
\tag{85}
$$

and $\mathbb{R}^2$-valued, $\mathcal{G}_t^1$-adapted process quadruple $(\hat{X}(\cdot),\Phi(\cdot),\Pi_1(\cdot),\Pi_3(\cdot))$ satisfies

$$
\begin{cases}
d\hat{X}(t)=\Big\{\big[\widetilde{r}+\widetilde{B}\widetilde{B}^\top P(t)-\alpha_1\alpha_1^\top P(t)\big]\hat{X}(t)+\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)\Phi(t)+\alpha_2\hat{s}(t)+\alpha_3\hat{d}(t)\Big\}dt\\
\qquad\qquad +\widetilde{\sigma}_1dW_1(t)+\widetilde{\sigma}_3dW_3(t),\\
-d\Phi(t)=\Big\{\big[\widetilde{r}+\widetilde{B}\widetilde{B}^\top P(t)-\alpha_1\alpha_1^\top P(t)\big]\Phi(t)+P(t)\alpha_2\hat{s}(t)+P(t)\alpha_3\hat{d}(t)\Big\}dt\\
\qquad\qquad -\Pi_1(t)dW_1(t)-\Pi_3(t)dW_3(t),\ t\in[0,T],\\
\hat{X}(0)=X_0,\quad \Phi(T)=0.
\end{cases}
\tag{86}
$$

For the leader (principal)'s problem, the state now writes

$$
\begin{cases}
dX(t)=\Big\{\widetilde{r}X(t)+\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)P(t)\hat{X}(t)+\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)\Phi(t)+\alpha_2s(t)+\alpha_3d(t)\Big\}dt\\
\qquad\qquad +\widetilde{\sigma}_1dW_1(t)+\widetilde{\sigma}_2dW_2(t)+\widetilde{\sigma}_3dW_3(t),\\
-d\Phi(t)=\Big\{\big[\widetilde{r}+\widetilde{B}\widetilde{B}^\top P(t)-\alpha_1\alpha_1^\top P(t)\big]\Phi(t)+P(t)\alpha_2\hat{s}(t)+P(t)\alpha_3\hat{d}(t)\Big\}dt\\
\qquad\qquad -\Pi_1(t)dW_1(t)-\Pi_3(t)dW_3(t),\ t\in[0,T],\\
X(0)=X_0,\quad \Phi(T)=0.
\end{cases}
\tag{87}
$$

Following the step in Section 3.2, we introduce $\mathbb{R}^2$-valued, $\mathcal{F}_t$-adapted process quintuple $(\widetilde{p}(\cdot),\widetilde{y}(\cdot),z_1(\cdot),z_2(\cdot),z_3(\cdot))$ which satisfies the adjoint equation

$$
\begin{cases}
d\widetilde{p}(t)=\Big\{\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)\widetilde{y}(t)+\big[\widetilde{r}^\top+\widetilde{B}\widetilde{B}^\top P(t)+\alpha_1\alpha_1^\top P(t)\big]\widetilde{p}(t)\Big\}dt,\\
-d\widetilde{y}(t)=\big[\widetilde{r}\widetilde{y}(t)+\big(\widetilde{B}\widetilde{B}^\top-\alpha_1\alpha_1^\top\big)P(t)\hat{\widetilde{y}}(t)+\widetilde{G}_2X^*(t)\big]dt-z_1(t)dW_1(t)\\
\qquad\qquad -z_2(t)dW_2(t)-z_3(t)dW_3(t),\ t\in[0,T],\\
\widetilde{p}(0)=0,\quad \widetilde{y}(T)=\widetilde{G}_2X^*(T),
\end{cases}
\tag{88}
$$

then

$$s^*(t) = -\alpha_2\check{\widetilde{y}}(t) - \alpha_2 P(t)\check{\check{\widetilde{y}}}(t),\ d^*(t) = \alpha_3\check{\widetilde{y}}(t) + \alpha_3 P(t)\check{\check{\widetilde{y}}}(t). \tag{89}$$

Letting

$$\mathcal{X} = \begin{pmatrix} X^* \\ \widetilde{p} \end{pmatrix},\ Y = \begin{pmatrix} \widetilde{y} \\ \Phi^* \end{pmatrix},\ Z_1 = \begin{pmatrix} z_1 \\ \Pi_1^* \end{pmatrix},\ Z_2 = \begin{pmatrix} z_2 \\ 0 \end{pmatrix},\ Z_3 = \begin{pmatrix} z_3 \\ \Pi_3^* \end{pmatrix}, \tag{90}$$

and

$$\begin{cases} \mathcal{A}_0 := \begin{pmatrix} \widetilde{r} & 0 \\ 0 & \widetilde{r} + \widetilde{B}\widetilde{B}^\top P(t) - \alpha_1\alpha_1^\top P(t) \end{pmatrix},\ \overline{\mathcal{A}}_0 := \begin{pmatrix} \widetilde{B}\widetilde{B}^\top P(t) - \alpha_1\alpha_1^\top P(t) & 0 \\ 0 & 0 \end{pmatrix}, \\ \mathcal{B}_0 := \begin{pmatrix} 0 & \widetilde{B}\widetilde{B}^\top - \alpha_1\alpha_1^\top \\ \widetilde{B}\widetilde{B}^\top - \alpha_1\alpha_1^\top & 0 \end{pmatrix},\ \widetilde{\alpha}_2 := \begin{pmatrix} \alpha_2 \\ 0 \end{pmatrix},\ \widetilde{\alpha}_3 := \begin{pmatrix} \alpha_3 \\ 0 \end{pmatrix}, \\ \Sigma_1 := \begin{pmatrix} \widetilde{\sigma}_1 \\ 0 \end{pmatrix},\ \Sigma_2 := \begin{pmatrix} \widetilde{\sigma}_2 \\ 0 \end{pmatrix},\ \Sigma_3 := \begin{pmatrix} \widetilde{\sigma}_3 \\ 0 \end{pmatrix},\ \mathcal{G}_2 := \begin{pmatrix} \widetilde{G}_2 & 0 \\ 0 & 0 \end{pmatrix},\ \mathcal{X}_0 := \begin{pmatrix} X_0 \\ 0 \end{pmatrix}, \\ \overline{\alpha}_2 := \begin{pmatrix} 0 \\ P(t)\alpha_2 \end{pmatrix},\ \overline{\alpha}_3 := \begin{pmatrix} 0 \\ P(t)\alpha_3 \end{pmatrix},\ \Lambda_1 := \begin{pmatrix} Z_1 \\ \beta_1 \end{pmatrix},\ \Lambda_2 := \begin{pmatrix} Z_2 \\ 0 \end{pmatrix},\ \Lambda_3 := \begin{pmatrix} Z_3 \\ \beta_3 \end{pmatrix}, \end{cases}$$

then we have

$$\begin{cases} d\mathcal{X}(t) = \big[\mathcal{A}_0\mathcal{X}(t) + \overline{\mathcal{A}}_0\hat{\mathcal{X}}(t) + \mathcal{B}_0 Y(t) + \widetilde{\alpha}_2 s^*(t) + \widetilde{\alpha}_3 d^*(t)\big]dt \\ \qquad\qquad + \Sigma_1 dW_1(t) + \Sigma_2 dW_2(t) + \Sigma_3 dW_3(t), \\ -dY(t) = \Big\{\big[\mathcal{G}_2\mathcal{X}(t) + \mathcal{A}_0 Y(t) + \overline{\mathcal{A}}_0\hat{Y}(t) + \overline{\alpha}_2\hat{s}^*(t) + \overline{\alpha}_3\hat{d}^*(t)\Big\}dt \\ \qquad\qquad - \Pi_1(t)dW_1(t) - \Pi_3(t)dW_3(t),\ t \in [0,T], \\ \mathcal{X}(0) = \mathcal{X}_0, \quad Y(T) = \mathcal{G}_2\mathcal{X}(T). \end{cases} \tag{91}$$

As Section 3.2, letting $Y(t) = \mathcal{P}_1(t)\mathcal{X}(t) + \mathcal{P}_2(t)\hat{\mathcal{X}}(t) + \mathcal{P}_3(t)\check{\mathcal{X}}(t) + \mathcal{P}_4(t)\check{\hat{\mathcal{X}}}(t)$, then we get

$$\begin{cases} s^*(t) = -\big(\widetilde{\alpha}_2^\top\mathcal{P}_1 + \widetilde{\alpha}_2^\top\mathcal{P}_3\big)\check{\mathcal{X}}(t) - \big(\widetilde{\alpha}_2^\top\mathcal{P}_2 + \widetilde{\alpha}_2^\top\mathcal{P}_4 + \overline{\alpha}_2^\top\big)\check{\hat{\mathcal{X}}}(t), \\ d^*(t) = \big(\widetilde{\alpha}_3^\top\mathcal{P}_1 + \widetilde{\alpha}_3^\top\mathcal{P}_3\big)\check{\mathcal{X}}(t) + \big(\widetilde{\alpha}_3^\top\mathcal{P}_2 + \widetilde{\alpha}_3^\top\mathcal{P}_4 + \overline{\alpha}_3^\top\big)\check{\hat{\mathcal{X}}}(t), \end{cases} \tag{92}$$

where 4×4-matrix-valued functions $\mathcal{P}_1(\cdot),\mathcal{P}_2(\cdot),\mathcal{P}_3(\cdot),\mathcal{P}_4(\cdot)$ satisfy

$$\begin{cases} 0 = \dot{\mathcal{P}}_1 + \mathcal{P}_1\mathcal{A}_0 + \mathcal{A}_0^\top\mathcal{P}_1 + \mathcal{P}_1\mathcal{B}_0\mathcal{P}_1 + \mathcal{G}_2,\ \mathcal{P}_1(T) = \mathcal{G}_2, \\ 0 = \dot{\mathcal{P}}_2 + \mathcal{P}_2(\mathcal{A}_0 + \overline{\mathcal{A}}_0) + (\mathcal{A}_0 + \overline{\mathcal{A}}_0)^\top\mathcal{P}_2 + \mathcal{P}_1\mathcal{B}_0\mathcal{P}_2 + \mathcal{P}_2\mathcal{B}_0\mathcal{P}_1 + \mathcal{P}_2\mathcal{B}_0\mathcal{P}_2 + \mathcal{P}_1\overline{\mathcal{A}}_0 + \overline{\mathcal{A}}_0^\top\mathcal{P}_1,\ \mathcal{P}_2(T) = 0, \\ 0 = \dot{\mathcal{P}}_3 + \mathcal{A}_0^\top\mathcal{P}_3 + \mathcal{P}_3\mathcal{A}_0 + \mathcal{P}_1\big(\mathcal{B}_0 + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\big)\mathcal{P}_3 + \mathcal{P}_3\big(\mathcal{B}_0 + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\big)\mathcal{P}_1 \\ \qquad + \mathcal{P}_3\big(\mathcal{B}_0 + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\big)\mathcal{P}_3 + \mathcal{P}_1\big(\widetilde{\alpha}_3\widetilde{\alpha}_3^\top - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\big)\mathcal{P}_1,\ \mathcal{P}_3(T) = 0, \\ 0 = \dot{\mathcal{P}}_4 + \mathcal{P}_4(\mathcal{A}_0 + \overline{\mathcal{A}}_0) + (\mathcal{A}_0 + \overline{\mathcal{A}}_0)^\top\mathcal{P}_4 + \mathcal{P}_2\big(\widetilde{\alpha}_2\overline{\alpha}_2^\top - \widetilde{\alpha}_3\overline{\alpha}_3^\top\big) + \mathcal{P}_4\big(\mathcal{P}_1\mathcal{B}_0 - \mathcal{P}_1\widetilde{\alpha}_2\widetilde{\alpha}_2^\top + \mathcal{P}_1\widetilde{\alpha}_3\widetilde{\alpha}_3^\top \\ \qquad + \mathcal{P}_2\mathcal{B}_0 - \mathcal{P}_2\widetilde{\alpha}_2\widetilde{\alpha}_2^\top + \mathcal{P}_2\widetilde{\alpha}_3\widetilde{\alpha}_3^\top - \overline{\alpha}_2\widetilde{\alpha}_2^\top + \overline{\alpha}_3\widetilde{\alpha}_3^\top\big) + \big(\mathcal{B}_0\mathcal{P}_1 - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_1 + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_1 + \mathcal{B}_0\mathcal{P}_2 - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_2 \\ \qquad + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_2 - \widetilde{\alpha}_2\overline{\alpha}_2^\top + \widetilde{\alpha}_3\overline{\alpha}_3^\top\big)\mathcal{P}_4 + \mathcal{P}_4\big(\mathcal{B}_0 - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top + \widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_4 - \mathcal{P}_1\widetilde{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_2 + \mathcal{P}_1\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_2 \\ \qquad - \mathcal{P}_2\widetilde{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_1 + \mathcal{P}_2\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_1 + \mathcal{P}_2\big(\mathcal{B}_0 - \widetilde{\alpha}_2\widetilde{\alpha}_2^\top - \widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_3 - \mathcal{P}_2\widetilde{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_2 + \mathcal{P}_2\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_2 - \mathcal{P}_1\widetilde{\alpha}_2\overline{\alpha}_2^\top \\ \qquad + \mathcal{P}_1\widetilde{\alpha}_3\overline{\alpha}_3^\top - \mathcal{P}_2\widetilde{\alpha}_2\overline{\alpha}_2^\top + \mathcal{P}_2\widetilde{\alpha}_3\overline{\alpha}_3^\top - \overline{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_1 + \overline{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_1 - \overline{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_2 + \overline{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_2 - \overline{\alpha}_2\widetilde{\alpha}_2^\top\mathcal{P}_3 \\ \qquad + \overline{\alpha}_3\widetilde{\alpha}_3^\top\mathcal{P}_3 + \overline{\mathcal{A}}_0^\top\mathcal{P}_3 - \overline{\alpha}_2\overline{\alpha}_2^\top + \overline{\alpha}_3\overline{\alpha}_3^\top,\ \mathcal{P}_4(T) = 0, \end{cases} \tag{93}$$

$\mathbb{R}^4$-valued, $\mathcal{G}_t^2$-adapted processes $\check{\mathcal{X}}(\cdot)$ satisfies

$$\begin{cases} d\check{\mathcal{X}}(t)=\Big\{\big[\mathcal{A}_0+\big(\mathcal{B}_0-\widetilde{\alpha}_2\widetilde{\alpha}_2^\top+\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)(\mathcal{P}_1+\mathcal{P}_3)\big]\check{\mathcal{X}}(t)+\big[\overline{\mathcal{A}}_0+\big(\mathcal{B}_0-\widetilde{\alpha}_2\widetilde{\alpha}_2^\top+\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)(\mathcal{P}_2+\mathcal{P}_4) \\ \qquad -\widetilde{\alpha}_2\overline{\alpha}_2^\top+\widetilde{\alpha}_3\overline{\alpha}_3^\top\big]\check{\check{\mathcal{X}}}(t)\Big\}dt+\Sigma_2dW_2(t)+\Sigma_3dW_3(t),\ t\in[0,T], \\ \check{\mathcal{X}}(0)=\mathcal{X}_0, \end{cases} \tag{94}$$

and $\mathbb{R}^4$-valued, $\mathcal{G}_t^2\cap\mathcal{G}_t^2$-adapted processes $\check{\check{\mathcal{X}}}(\cdot)$ satisfies

$$\begin{cases} d\check{\check{\mathcal{X}}}(t)=\big[\mathcal{A}_0+\overline{\mathcal{A}}_0+\big(\mathcal{B}_0-\widetilde{\alpha}_2\widetilde{\alpha}_2^\top+\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) \\ \qquad -\widetilde{\alpha}_2\overline{\alpha}_2^\top+\widetilde{\alpha}_3\overline{\alpha}_3^\top\big]\check{\check{\mathcal{X}}}(t)dt+\Sigma_3dW_3(t),\ t\in[0,T], \\ \check{\check{\mathcal{X}}}(0)=\mathcal{X}_0. \end{cases} \tag{95}$$

For the follower, by (84), noticing (90), we obtain

$$\begin{cases} e^*(t)=\widetilde{B}^\top\big[P(t)\hat{X}^*(t)+\Phi^*(t)\big]=\Big[\begin{pmatrix}\widetilde{B}^\top P(t) & 0\end{pmatrix}\hat{\mathcal{X}}(t)+\begin{pmatrix}0 & \widetilde{B}^\top\end{pmatrix}\hat{Y}(t)\Big] \\ \qquad =\Big[\begin{pmatrix}\widetilde{B}^\top P(t) & 0\end{pmatrix}+\begin{pmatrix}0 & \widetilde{B}^\top\end{pmatrix}(\mathcal{P}_1+\mathcal{P}_2)\Big]\hat{\mathcal{X}}(t)+\begin{pmatrix}0 & \widetilde{B}^\top\end{pmatrix}(\mathcal{P}_3+\mathcal{P}_4)\check{\check{\mathcal{X}}}(t), \\ c^*(t)=-\alpha_1^\top\big[P(t)\hat{X}^*(t)+\Phi^*(t)\big]=\Big[\begin{pmatrix}-\alpha_1^\top P(t) & 0\end{pmatrix}\hat{\mathcal{X}}(t)+\begin{pmatrix}0 & -\alpha_1^\top\end{pmatrix}\hat{Y}(t)\Big] \\ \qquad =\Big[\begin{pmatrix}-\alpha_1^\top P(t) & 0\end{pmatrix}+\begin{pmatrix}0 & -\alpha_1^\top\end{pmatrix}(\mathcal{P}_1+\mathcal{P}_2)\Big]\hat{\mathcal{X}}(t)+\begin{pmatrix}0 & -\alpha_1^\top\end{pmatrix}(\mathcal{P}_3+\mathcal{P}_4)\check{\check{\mathcal{X}}}(t), \end{cases} \tag{96}$$

where $\mathbb{R}^4$-valued, $\mathcal{G}_t^1$-adapted process $\hat{\mathcal{X}}(\cdot)$ satisfies

$$\begin{cases} d\hat{\mathcal{X}}(t)=\Big\{\big[\mathcal{A}_0+\overline{\mathcal{A}}_0+\mathcal{B}_0(\mathcal{P}_1+\mathcal{P}_2)\big]\hat{\mathcal{X}}(t)+\big[\mathcal{B}_0(\mathcal{P}_3+\mathcal{P}_4)-\big(\widetilde{\alpha}_2\widetilde{\alpha}_2^\top-\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)(\mathcal{P}_1+\mathcal{P}_2 \\ \qquad +\mathcal{P}_3+\mathcal{P}_4)\big]\check{\check{\mathcal{X}}}(t)-\widetilde{\alpha}_2\overline{\alpha}_2^\top+\widetilde{\alpha}_3\overline{\alpha}_3^\top\Big\}dt+\Sigma_1dW_1(t)+\Sigma_3dW_3(t),\ t\in[0,T], \\ \hat{\mathcal{X}}(0)=\mathcal{X}_0. \end{cases} \tag{97}$$

And the optimal state equation of the leader is

$$\begin{cases} d\mathcal{X}(t)=\Big\{\big(\mathcal{A}_0+\mathcal{B}_0\mathcal{P}_1\big)\mathcal{X}(t)+\big(\overline{\mathcal{A}}_0+\mathcal{B}_0\mathcal{P}_2\big)\hat{\mathcal{X}}(t)+\big[\big(\mathcal{B}_0-\widetilde{\alpha}_2\widetilde{\alpha}_2^\top+\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_3 \\ \qquad -\big(\widetilde{\alpha}_2\widetilde{\alpha}_2^\top-\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_1\big]\check{\mathcal{X}}(t)+\big[\big(\mathcal{B}_0+\widetilde{\alpha}_2\widetilde{\alpha}_2^\top-\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_4-\big(\widetilde{\alpha}_2\widetilde{\alpha}_2^\top-\widetilde{\alpha}_3\widetilde{\alpha}_3^\top\big)\mathcal{P}_2 \\ \qquad -\widetilde{\alpha}_2\overline{\alpha}_2^\top+\widetilde{\alpha}_3\overline{\alpha}_3^\top\big]\check{\check{\mathcal{X}}}(t)\Big\}dt+\Sigma_1dW_1(t)+\Sigma_2dW_2(t)+\Sigma_3dW_3(t),\ t\in[0,T], \\ \mathcal{X}(0)=\mathcal{X}_0. \end{cases} \tag{98}$$

Finally, we rewrite the above Stackelberg equilibrium strategy $(e^*(\cdot),c^*(\cdot),s^*(\cdot),d^*(\cdot))$, with respect to the asset of the principal $y(\cdot)$ and that of the agent $m(\cdot)$. In fact, let

$$\mathcal{P}_k\equiv\big(\mathcal{P}_k^{i,j}\big)_{4\times4},k=1,2,3,4,\ \text{and}\ P\equiv\big(P^{i,j}\big)_{2\times2},$$

where $\mathcal{P}^{i,j}$, $P^{i,j}$ denotes the elements of the matrices. Then by (92) and (96) we have

$$
\begin{aligned}
s^*(t) &= \big(-\mathcal{P}_1^{1,1}+\mathcal{P}_1^{2,1}-\mathcal{P}_3^{1,1}+\mathcal{P}_3^{2,1}\big)\check{y}(t)+\big(-\mathcal{P}_2^{1,1}+\mathcal{P}_2^{2,1}-\mathcal{P}_4^{1,1}+\mathcal{P}_4^{2,1}\big)\check{\hat{y}}(t)\\
&\quad+\big(-\mathcal{P}_1^{1,2}+\mathcal{P}_1^{2,2}-\mathcal{P}_3^{1,2}+\mathcal{P}_3^{2,2}\big)\check{m}(t)+\big(-\mathcal{P}_2^{1,2}+\mathcal{P}_2^{2,2}-\mathcal{P}_4^{1,2}+\mathcal{P}_4^{2,2}\big)\check{\hat{m}}(t)\\
&\quad+\begin{pmatrix} -\mathcal{P}_1^{1,3}+\mathcal{P}_1^{2,3}-\mathcal{P}_3^{1,3}+\mathcal{P}_3^{2,3}\\ -\mathcal{P}_1^{1,4}+\mathcal{P}_1^{2,4}-\mathcal{P}_3^{1,4}+\mathcal{P}_3^{2,4}\\ -\mathcal{P}_2^{1,3}+\mathcal{P}_2^{2,3}-\mathcal{P}_4^{1,3}+\mathcal{P}_4^{2,3}-P^{1,1}+P^{1,2}\\ -\mathcal{P}_2^{1,4}+\mathcal{P}_2^{2,4}-\mathcal{P}_4^{1,4}+\mathcal{P}_4^{2,4}-P^{2,1}+P^{2,2}\end{pmatrix}^{\top}\begin{pmatrix}\check{\tilde{p}}(t)\\ \check{\hat{\tilde{p}}}(t)\end{pmatrix},\\
d^*(t) &= -\big(\mathcal{P}_1^{1,1}+\mathcal{P}_3^{1,1}\big)\check{y}(t)-\big(\mathcal{P}_2^{1,1}+\mathcal{P}_4^{1,1}\big)\check{\hat{y}}(t)-\big(\mathcal{P}_1^{1,2}+\mathcal{P}_3^{1,2}\big)\check{m}(t)\\
&\quad-\big(\mathcal{P}_2^{1,2}+\mathcal{P}_4^{1,2}\big)\check{\hat{m}}(t)-\begin{pmatrix}\mathcal{P}_1^{1,3}+\mathcal{P}_3^{1,3}\\ \mathcal{P}_1^{1,4}+\mathcal{P}_3^{1,4}\\ \mathcal{P}_2^{1,3}+\mathcal{P}_4^{1,3}+P^{1,1}\\ \mathcal{P}_2^{1,4}+\mathcal{P}_4^{1,4}+P^{2,1}\end{pmatrix}^{\top}\begin{pmatrix}\check{\tilde{p}}(t)\\ \check{\hat{\tilde{p}}}(t)\end{pmatrix}, \qquad (99)\\
e^*(t) &= B\big(P^{1,1}+\mathcal{P}_1^{3,1}+\mathcal{P}_2^{3,1}\big)\hat{y}(t)+B\big(\mathcal{P}_3^{3,1}+\mathcal{P}_4^{3,1}\big)\check{\hat{y}}(t)+B\big(P^{1,2}+\mathcal{P}_1^{3,2}+\mathcal{P}_2^{3,2}\big)\hat{m}(t)\\
&\quad+B\big(\mathcal{P}_3^{3,2}+\mathcal{P}_4^{3,2}\big)\check{\hat{m}}(t)+B\begin{pmatrix}\mathcal{P}_1^{3,3}+\mathcal{P}_2^{3,3} & \mathcal{P}_1^{3,4}+\mathcal{P}_2^{3,4} & \mathcal{P}_3^{3,3}+\mathcal{P}_4^{3,3} & \mathcal{P}_3^{3,4}+\mathcal{P}_4^{3,4}\end{pmatrix}\begin{pmatrix}\hat{\tilde{p}}(t)\\ \check{\hat{\tilde{p}}}(t)\end{pmatrix},\\
c^*(t) &= \big(P^{2,1}+\mathcal{P}_1^{4,1}+\mathcal{P}_2^{4,1}\big)\hat{y}(t)+\big(\mathcal{P}_3^{4,1}+\mathcal{P}_4^{4,1}\big)\check{\hat{y}}(t)+\big(P^{2,2}+\mathcal{P}_1^{4,2}+\mathcal{P}_2^{4,2}\big)\hat{m}(t)\\
&\quad+\big(\mathcal{P}_3^{4,2}+\mathcal{P}_4^{4,2}\big)\check{\hat{m}}(t)+\begin{pmatrix}\mathcal{P}_1^{4,3}+\mathcal{P}_2^{4,3} & \mathcal{P}_1^{4,4}+\mathcal{P}_2^{4,4} & \mathcal{P}_3^{4,3}+\mathcal{P}_4^{4,3} & \mathcal{P}_3^{4,4}+\mathcal{P}_4^{4,4}\end{pmatrix}\begin{pmatrix}\hat{\tilde{p}}(t)\\ \check{\hat{\tilde{p}}}(t)\end{pmatrix},
\end{aligned}
$$

where $(\check{\hat{y}}(t),\check{\hat{m}}(t),\check{\hat{\tilde{p}}}(t))\equiv\check{\hat{\mathcal{X}}}(t)$ satisfies (95), $(\hat{y}(t),\hat{m}(t),\hat{\tilde{p}}(t))\equiv\hat{\mathcal{X}}(t)$ satisfies (97) and $(\check{y}(t),\check{m}(t),\check{\tilde{p}}(t))\equiv\check{\mathcal{X}}(t)$ satisfies (94).

5. Concluding remarks

In this paper, we have discussed the stochastic LQ Stackelberg differential game with overlapping information. This kind of game problem possesses two attractive features worthy of being highlighted. First, the game problem has the asymmetric and overlapping information between the two players, which was not considered in Yong [35], Øksendal et al. [19] and Bensoussan et al. [2]. Stochastic filtering technique is introduced to compute the optimal filtering estimates for the corresponding adjoint processes, which perform as the solution to some nonstandard stochastic filtering equations. Second, the Stackelberg equilibrium is represented in its state estimate feedback form, under some appropriate assumptions on the coefficient matrices in the state equation and the cost functional. Some new system of Riccati equations are first introduced in this paper, to deal with the leader's problem.

Many interesting and important problems remain open. The solvability of the system of Riccati equations (44) (even for its special case (75)) is a challenging problem. It is worthy to study the numerical approximation of its solution. The problem with random coefficients is interesting. In fact, when $A(t)$ is a deterministic function for any t, it can be obtained that $\mathbb{E}[A(t)X(t)|\mathcal{F}_t]=A(t)\mathbb{E}[X(t)|\mathcal{F}_t]$. Throughout our paper, this fact has played an essential role. However, it will be no longer valid if $A(t)$ is a random matrix for any t. Therefore, for the case of random coefficient case, one might have to find a different approach. The problem is still under investigation. Problems with time delay (Xu and Zhang [33], Xu et al. [34]), with partial observation (Huang et al. [11], Wang et al. [28], Shi et al. [23]) and of mean-field type (Wang et al. [29], Moon and Başar [17], Lin et al. [15]) which are important and reasonable for applications and more technically demanding in its filtering procedure, are highly desirable for further research. These topics will be considered in our future work.

Acknowledgements

The authors would like to thank the editor and the two anonymous referees for their constructive and insightful comments for improving the quality of this work. The main content of this paper is presented by the first author in The 40th Conference on Stochastic Processes and their Applications, Gothenberg, Sweden, June 2018, and The 12th AIMS Conference on Dynamic Systems, Differential Equations and Applications, Taibei, Taiwan, July 2018. Many thanks for discussions and suggestions with Professor Shuenn-Jyi Sheu and Professor Jiongmin Yong. The first and second authors would like to thank Department of Mathematics and SUSTech International Center for Mathematics, Southern University of Science and Technology for their hospitality during their visit to Shenzhen.

Appendix: Proof of Theorem 2.2

Proof of Theorem 2.2. We wish to decouple FBSDE (42). For this target, let

$$Y(t) = \mathcal{P}_1(t)X(t) + \mathcal{P}_2(t)\hat{X}(t) + \mathcal{P}_3(t)\check{X}(t) + \mathcal{P}_4(t)\check{\hat{X}}(t), \tag{100}$$

where $\mathcal{P}_1(\cdot), \mathcal{P}_2(\cdot), \mathcal{P}_3(\cdot), \mathcal{P}_4(\cdot)$ are all differentiable, deterministic $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ matrix-valued functions with $\mathcal{P}_1(T) = \mathcal{G}_2$, $\mathcal{P}_2(T) = 0$, $\mathcal{P}_3(T) = 0$, $\mathcal{P}_4(T) = 0$. From (42), the equations for $\hat{X}(\cdot)$, $\check{X}(\cdot)$, $\check{\hat{X}}(\cdot)$ are

$$\begin{cases} d\hat{X}(t) = \Big[(\mathcal{A}_0 + \widehat{\mathcal{A}}_0)\hat{X} + \overline{\mathcal{A}}_0\check{\hat{X}} + \mathcal{B}_0\hat{Y} + (\mathcal{C}_0 + \widetilde{\mathcal{C}}_0)\check{\hat{Y}} + \mathcal{B}_1^\top \hat{Z}_1 + (\widetilde{\mathcal{B}}_1^\top + \overline{\mathcal{C}}_0)\check{\hat{Z}}_1 + \mathcal{B}_2^\top \hat{Z}_2 + (\widetilde{\mathcal{B}}_2^\top + \overline{\mathcal{D}}_0)\check{\hat{Z}}_2 \\ \qquad + \mathcal{B}_3^\top \hat{Z}_3 + (\widetilde{\mathcal{B}}_3^\top + \overline{\mathcal{E}}_0)\check{\hat{Z}}_3\Big]dt + \sum_{i=1,3}\Big[(\mathcal{A}_i + \widehat{\mathcal{A}}_i)\hat{X} + (\widetilde{\mathcal{A}}_i + \overline{\mathcal{A}}_i)\check{\hat{X}} + \mathcal{B}_i\hat{Y} + (\widetilde{\mathcal{B}}_i + \overline{\mathcal{B}}_i)\check{\hat{Y}} \\ \qquad + \mathcal{C}_i\hat{Z}_1 + (\widetilde{\mathcal{C}}_i + \overline{\mathcal{C}}_i)\check{\hat{Z}}_1 + \mathcal{D}_i\hat{Z}_2 + (\widetilde{\mathcal{D}}_i + \overline{\mathcal{D}}_i)\check{\hat{Z}}_2 + \mathcal{E}_i\hat{Z}_3 + (\widetilde{\mathcal{E}}_i + \overline{\mathcal{E}}_i)\check{\hat{Z}}_3\Big]dW_i(t),\ t \in [0,T], \\ \hat{X}(0) = X_0, \end{cases} \tag{101}$$

$$\begin{cases} d\check{X}(t) = \Big[\mathcal{A}_0\check{X} + (\widehat{\mathcal{A}}_0 + \overline{\mathcal{A}}_0)\check{\hat{X}} + (\mathcal{B}_0 + \mathcal{C}_0)\check{Y} + \widetilde{\mathcal{C}}_0\check{\hat{Y}} + (\mathcal{B}_1 + \widetilde{\mathcal{B}}_1)^\top \check{Z}_1 + \overline{\mathcal{C}}_0\check{\hat{Z}}_1 + (\mathcal{B}_2 + \widetilde{\mathcal{B}}_2)^\top \check{Z}_2 + \overline{\mathcal{D}}_0\check{\hat{Z}}_2 \\ \qquad + (\mathcal{B}_3 + \widetilde{\mathcal{B}}_3)^\top \check{Z}_3 + \overline{\mathcal{E}}_0\check{\hat{Z}}_3\Big]dt + \sum_{i=2,3}\Big[(\mathcal{A}_i + \widetilde{\mathcal{A}}_i)\check{X} + (\widehat{\mathcal{A}}_i + \overline{\mathcal{A}}_i)\check{\hat{X}} + (\mathcal{B}_i + \widetilde{\mathcal{B}}_i)\check{Y} + \overline{\mathcal{B}}_i\check{\hat{Y}} \\ \qquad + (\mathcal{C}_i + \widetilde{\mathcal{C}}_i)\check{Z}_1 + \overline{\mathcal{C}}_i\check{\hat{Z}}_1 + (\mathcal{D}_i + \widetilde{\mathcal{D}}_i)\check{Z}_2 + \overline{\mathcal{D}}_i\check{\hat{Z}}_2 + (\mathcal{E}_i + \widetilde{\mathcal{E}}_i)\check{Z}_3 + \overline{\mathcal{E}}_i\check{\hat{Z}}_3\Big]dW_i(t),\ t \in [0,T], \\ \check{X}(0) = X_0, \end{cases} \tag{102}$$

and

$$\begin{cases} d\check{\hat{X}}(t) = \Big[(\mathcal{A}_0 + \widehat{\mathcal{A}}_0 + \overline{\mathcal{A}}_0)\check{\hat{X}} + (\mathcal{B}_0 + \mathcal{C}_0 + \widetilde{\mathcal{C}}_0)\check{\hat{Y}} + (\mathcal{B}_1^\top + \widetilde{\mathcal{B}}_1^\top + \overline{\mathcal{C}}_0)\check{\hat{Z}}_1 + (\mathcal{B}_2^\top + \widetilde{\mathcal{B}}_2^\top + \overline{\mathcal{D}}_0)\check{\hat{Z}}_2 \\ \qquad + (\mathcal{B}_3^\top + \widetilde{\mathcal{B}}_3 + \overline{\mathcal{E}}_0)\check{\hat{Z}}_3\Big]dt + \Big[(\mathcal{A}_3 + \widetilde{\mathcal{A}}_3 + \widehat{\mathcal{A}}_3 + \overline{\mathcal{A}}_3)\check{\hat{X}} + (\mathcal{B}_3 + \widetilde{\mathcal{B}}_3 + \overline{\mathcal{B}}_3)\check{\hat{Y}} \\ \qquad + (\mathcal{C}_3 + \widetilde{\mathcal{C}}_3 + \overline{\mathcal{C}}_3)\check{\hat{Z}}_1 + (\mathcal{D}_3 + \widetilde{\mathcal{D}}_3 + \overline{\mathcal{D}}_3)\check{\hat{Z}}_2 + (\mathcal{E}_3 + \widetilde{\mathcal{E}}_3 + \overline{\mathcal{E}}_3)\check{\hat{Z}}_3\Big]dW_3(t),\ t \in [0,T], \\ \check{\hat{X}}(0) = X_0. \end{cases} \tag{103}$$

Applying Itô's formula to (100), we obtain

$$\begin{aligned} dY(t) = \Big\{&\big(\dot{\mathcal{P}}_1 + \mathcal{P}_1\mathcal{A}_0 + \mathcal{P}_1\mathcal{B}_0\mathcal{P}_1\big)X + \big[\dot{\mathcal{P}}_2 + \mathcal{P}_1\widehat{\mathcal{A}}_0 + \mathcal{P}_2\big(\mathcal{A}_0 + \widehat{\mathcal{A}}_0\big) + \mathcal{P}_1\mathcal{B}_0\mathcal{P}_2 + \mathcal{P}_2\mathcal{B}_0\mathcal{P}_1 + \mathcal{P}_2\mathcal{B}_0\mathcal{P}_2\big]\hat{X} \\ &+ \big[\dot{\mathcal{P}}_3 + \mathcal{P}_3\mathcal{A}_0 + \mathcal{P}_1\mathcal{B}_0\mathcal{P}_3 + \mathcal{P}_1\mathcal{C}_0(\mathcal{P}_1 + \mathcal{P}_2) + \mathcal{P}_3\big(\mathcal{B}_0 + \mathcal{C}_0\big)\big(\mathcal{P}_1 + \mathcal{P}_3\big)\big]\check{X} + \big[\dot{\mathcal{P}}_4 + \mathcal{P}_4\big(\mathcal{A}_0 + \widehat{\mathcal{A}}_0 + \overline{\mathcal{A}}_0\big) \end{aligned}$$

$$
\begin{aligned}
&+\mathcal{P}_1\overline{\mathcal{A}}_0+\mathcal{P}_2\overline{\mathcal{A}}_0+\mathcal{P}_4\big(\mathcal{B}_0+\mathcal{C}_0+\widetilde{\mathcal{C}}_0\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_1\mathcal{B}_0\mathcal{P}_4+\mathcal{P}_1\mathcal{C}_0\big(\mathcal{P}_3+\mathcal{P}_4\big)\\
&+\mathcal{P}_1\widetilde{\mathcal{C}}_0\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\mathcal{C}_0+\widetilde{\mathcal{C}}_0\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\mathcal{B}_0\big(\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_3\big(\widehat{\mathcal{A}}_0+\overline{\mathcal{A}}_0\big)\\
&+\mathcal{P}_3\big(\mathcal{B}_0+\mathcal{C}_0\big)\big(\mathcal{P}_2+\mathcal{P}_4\big)+\mathcal{P}_3\widetilde{\mathcal{C}}_0\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_1^\top Z_1+\mathcal{P}_2\mathcal{B}_1^\top\hat{Z}_1+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_1^\top+\mathcal{P}_3\mathcal{B}_1^\top\\
&+\mathcal{P}_3\widetilde{\mathcal{B}}_1^\top\big)\check{Z}_1+\mathcal{P}_1\mathcal{B}_2^\top Z_2+\mathcal{P}_2\mathcal{B}_2^\top\hat{Z}_2+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_2^\top+\mathcal{P}_3\mathcal{B}_2^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_2^\top\big)\check{Z}_2+\mathcal{P}_1\mathcal{B}_3^\top Z_3+\mathcal{P}_2\mathcal{B}_3^\top\hat{Z}_3\\
&+\big(\mathcal{P}_1\widetilde{\mathcal{B}}_3^\top+\mathcal{P}_3\mathcal{B}_3^\top+\mathcal{P}_3\widetilde{\mathcal{B}}_3^\top\big)\check{Z}_3+\big[\mathcal{P}_1\overline{\mathcal{C}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_1^\top+\overline{\mathcal{C}}_0\big)+\mathcal{P}_3\overline{\mathcal{C}}_0+\mathcal{P}_4\big(\mathcal{B}_1^\top+\widetilde{\mathcal{B}}_1^\top+\overline{\mathcal{C}}_0\big)\big]\check{\hat{Z}}_1\\
&+\big[\mathcal{P}_1\overline{\mathcal{D}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0\big)+\mathcal{P}_3\overline{\mathcal{D}}_0+\mathcal{P}_4\big(\mathcal{B}_2^\top+\widetilde{\mathcal{B}}_2^\top+\overline{\mathcal{D}}_0\big)\big]\check{\hat{Z}}_2\\
&+\big[\mathcal{P}_1\overline{\mathcal{E}}_0+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{E}}_0\big)+\mathcal{P}_3\overline{\mathcal{E}}_0+\mathcal{P}_4\big(\mathcal{B}_3^\top+\widetilde{\mathcal{B}}_3^\top+\overline{\mathcal{E}}_0\big)\big]\check{\hat{Z}}_3\Big\}dt\\
+&\Big\{\big(\mathcal{P}_1\mathcal{A}_1+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_1\big)X+\big[\mathcal{P}_1\widehat{\mathcal{A}}_1+\mathcal{P}_2\big(\mathcal{A}_1+\widehat{\mathcal{A}}_1\big)+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_1+\mathcal{P}_2\big)\big]\hat{X}\\
&+\big[\mathcal{P}_1\mathcal{B}_1\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}+\big[\mathcal{P}_1\overline{\mathcal{A}}_1+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_1\overline{\mathcal{B}}_1\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{A}}_1+\overline{\mathcal{A}}_1\big)+\mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_3+\mathcal{P}_4\big)\\
&+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_1+\overline{\mathcal{B}}_1\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_1Z_1+\mathcal{P}_2\mathcal{C}_1\hat{Z}_1+\mathcal{P}_1\widetilde{\mathcal{C}}_1\check{Z}_1+\mathcal{P}_1\mathcal{D}_1Z_2\\
&+\mathcal{P}_2\mathcal{D}_1\hat{Z}_2+\mathcal{P}_1\widetilde{\mathcal{D}}_1\check{Z}_2+\mathcal{P}_1\mathcal{E}_1Z_3+\mathcal{P}_2\mathcal{E}_1\hat{Z}_3+\mathcal{P}_1\widetilde{\mathcal{E}}_1\check{Z}_3+\big(\mathcal{P}_1\overline{\mathcal{C}}_1+\mathcal{P}_2\widetilde{\mathcal{C}}_1+\mathcal{P}_2\overline{\mathcal{C}}_1\big)\check{\hat{Z}}_1\\
&+\big(\mathcal{P}_1\overline{\mathcal{D}}_1+\mathcal{P}_2\widetilde{\mathcal{D}}_1+\mathcal{P}_2\overline{\mathcal{D}}_1\big)\check{\hat{Z}}_2+\big(\mathcal{P}_1\overline{\mathcal{E}}_1+\mathcal{P}_2\widetilde{\mathcal{E}}_1+\mathcal{P}_2\overline{\mathcal{E}}_1\big)\check{\hat{Z}}_3\Big\}dW_1(t)\\
+&\Big\{\big(\mathcal{P}_1\mathcal{A}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_1\big)X+\big(\mathcal{P}_1\widehat{\mathcal{A}}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_2\big)\hat{X}+\big[\mathcal{P}_1\mathcal{B}_2\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_3\big)\\
&+\mathcal{P}_3\big(\mathcal{A}_2+\widetilde{\mathcal{A}}_2\big)+\mathcal{P}_3\big(\mathcal{B}_2+\widetilde{\mathcal{B}}_2\big)\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}+\big[\mathcal{P}_1\overline{\mathcal{A}}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_2\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_1\overline{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_3\big(\widehat{\mathcal{A}}_2+\overline{\mathcal{A}}_2\big)+\mathcal{P}_3\big(\mathcal{B}_2+\widetilde{\mathcal{B}}_2\big)\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_3\widetilde{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_2Z_1+\big(\mathcal{P}_1\widetilde{\mathcal{C}}_2+\mathcal{P}_3\mathcal{C}_2+\mathcal{P}_3\widetilde{\mathcal{C}}_2\big)\check{Z}_1+\mathcal{P}_1\mathcal{D}_2Z_2\\
&+\big(\mathcal{P}_1\widetilde{\mathcal{D}}_2+\mathcal{P}_3\mathcal{D}_2+\mathcal{P}_3\widetilde{\mathcal{D}}_2\big)\check{Z}_2+\mathcal{P}_1\mathcal{E}_2Z_3+\big(\mathcal{P}_1\widetilde{\mathcal{E}}_2+\mathcal{P}_3\mathcal{E}_2+\mathcal{P}_3\widetilde{\mathcal{E}}_2\big)\check{Z}_3\\
&+\big(\mathcal{P}_1\overline{\mathcal{C}}_2+\mathcal{P}_3\overline{\mathcal{C}}_2\big)\check{\hat{Z}}_1+\big(\mathcal{P}_1\overline{\mathcal{D}}_2+\mathcal{P}_3\overline{\mathcal{D}}_2\big)\check{\hat{Z}}_2+\big(\mathcal{P}_1\overline{\mathcal{E}}_2+\mathcal{P}_3\overline{\mathcal{E}}_2\big)\check{\hat{Z}}_3\Big\}dW_2(t)\\
+&\Big\{\big(\mathcal{P}_1\mathcal{A}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_1\big)X+\big[\mathcal{P}_1\widehat{\mathcal{A}}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_2+\mathcal{P}_2(\mathcal{A}_3+\widehat{\mathcal{A}}_3)+\mathcal{P}_2\mathcal{B}_3\big(\mathcal{P}_1+\mathcal{P}_2\big)\big]\hat{X}\\
&+\big[\mathcal{P}_1\mathcal{B}_3\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_3\big)+\mathcal{P}_3\big(\mathcal{A}_3+\widetilde{\mathcal{A}}_3\big)+\mathcal{P}_3\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}\\
&+\Big[\mathcal{P}_1\overline{\mathcal{A}}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_3\big(\mathcal{P}_2+\mathcal{P}_4\big)+\mathcal{P}_1\overline{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_2\mathcal{B}_3\big(\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_3\big(\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_3\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3\big)\big(\mathcal{P}_2+\mathcal{P}_4\big)+\mathcal{P}_3\overline{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_4\big(\mathcal{A}_3+\widetilde{\mathcal{A}}_3+\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_4\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\Big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_3Z_1+\mathcal{P}_2\mathcal{C}_3\hat{Z}_1+\big(\mathcal{P}_1\widetilde{\mathcal{C}}_3+\mathcal{P}_3\mathcal{C}_3\\
&+\mathcal{P}_3\widetilde{\mathcal{C}}_3\big)\check{Z}_1+\mathcal{P}_1\mathcal{D}_3Z_2+\mathcal{P}_2\mathcal{D}_3\hat{Z}_2+\big(\mathcal{P}_1\widetilde{\mathcal{D}}_3+\mathcal{P}_3\mathcal{D}_3+\mathcal{P}_3\widetilde{\mathcal{D}}_3\big)\check{Z}_2+\mathcal{P}_1\mathcal{E}_3Z_3+\mathcal{P}_2\mathcal{E}_3\hat{Z}_3\\
&+\big(\mathcal{P}_1\widetilde{\mathcal{E}}_3+\mathcal{P}_3\mathcal{E}_3+\mathcal{P}_3\widetilde{\mathcal{E}}_3\big)\check{Z}_3+\big[\mathcal{P}_1\overline{\mathcal{C}}_3+\mathcal{P}_2\widetilde{\mathcal{C}}_3+\mathcal{P}_2\overline{\mathcal{C}}_3+\mathcal{P}_3\overline{\mathcal{C}}_3+\mathcal{P}_4\big(\mathcal{C}_3+\widetilde{\mathcal{C}}_3+\overline{\mathcal{C}}_3\big)\big]\check{\hat{Z}}_1\\
&+\big[\mathcal{P}_1\overline{\mathcal{D}}_3+\mathcal{P}_2\widetilde{\mathcal{D}}_3+\mathcal{P}_2\overline{\mathcal{D}}_3+\mathcal{P}_3\overline{\mathcal{D}}_3+\mathcal{P}_4\big(\mathcal{D}_3+\widetilde{\mathcal{D}}_3+\overline{\mathcal{D}}_3\big)\big]\check{\hat{Z}}_2\\
&+\big[\mathcal{P}_1\overline{\mathcal{E}}_3+\mathcal{P}_2\widetilde{\mathcal{E}}_3+\mathcal{P}_2\overline{\mathcal{E}}_3+\mathcal{P}_3\overline{\mathcal{E}}_3+\mathcal{P}_4\big(\mathcal{E}_3+\widetilde{\mathcal{E}}_3+\overline{\mathcal{E}}_3\big)\big]\check{\hat{Z}}_3\Big\}dW_3(t)\\
=-&\Big\{(\mathcal{Q}_2+\mathcal{A}_0\mathcal{P}_1)X+(\mathcal{A}_0\mathcal{P}_2+\mathcal{H}_1\mathcal{P}_1+\mathcal{H}_1\mathcal{P}_2)\hat{X}+\mathcal{A}_0\mathcal{P}_3\check{X}+\big[\mathcal{H}_1+\mathcal{A}_0\mathcal{P}_4+\mathcal{H}_2\mathcal{P}_3\\
&+\mathcal{H}_2\mathcal{P}_4+\overline{\mathcal{A}}_0^\top\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{A}_1Z_1+\overline{\mathcal{A}}_1^\top\check{\hat{Z}}_1+\mathcal{A}_2Z_2+\widehat{\mathcal{A}}_2\hat{Z}_2+\overline{\mathcal{A}}_2^\top\check{\hat{Z}}_2\\
&+\mathcal{A}_3Z_3+\widehat{\mathcal{A}}_3\hat{Z}_3+\overline{\mathcal{A}}_3^\top\check{\hat{Z}}_3\Big\}dt+Z_1dW_1(t)+Z_2dW_2(t)+Z_3dW_3(t).
\end{aligned}
\tag{104}
$$

Comparing the diffusion terms $dW_1(t), dW_2(t), dW_3(t)$ on both sides of (104) respectively, we have

$$
\begin{aligned}
Z_1(t) =& \big(\mathcal{P}_1\mathcal{A}_1+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_1\big)X+\big[\mathcal{P}_1\widehat{\mathcal{A}}_1+\mathcal{P}_2\big(\mathcal{A}_1+\widehat{\mathcal{A}}_1\big)+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_2+\mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_1+\mathcal{P}_2\big)\big]\hat{X}\\
&+\big[\mathcal{P}_1\mathcal{B}_1\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}+\big[\mathcal{P}_1\overline{\mathcal{A}}_1+\mathcal{P}_1\mathcal{B}_1\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_1\overline{\mathcal{B}}_1\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{A}}_1+\overline{\mathcal{A}}_1\big)+\mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_3+\mathcal{P}_4\big)\\
&+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_1+\overline{\mathcal{B}}_1\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_1Z_1+\mathcal{P}_2\mathcal{C}_1\hat{Z}_1+\mathcal{P}_1\widetilde{\mathcal{C}}_1\check{Z}_1+\mathcal{P}_1\mathcal{D}_1Z_2\\
&+\mathcal{P}_2\mathcal{D}_1\hat{Z}_2+\mathcal{P}_1\widetilde{\mathcal{D}}_1\check{Z}_2+\mathcal{P}_1\mathcal{E}_1Z_3+\mathcal{P}_2\mathcal{E}_1\hat{Z}_3+\mathcal{P}_1\widetilde{\mathcal{E}}_1\check{Z}_3+\big(\mathcal{P}_1\overline{\mathcal{C}}_1+\mathcal{P}_2\widetilde{\mathcal{C}}_1+\mathcal{P}_2\overline{\mathcal{C}}_1\big)\check{\hat{Z}}_1\\
&+\big(\mathcal{P}_1\overline{\mathcal{D}}_1+\mathcal{P}_2\widetilde{\mathcal{D}}_1+\mathcal{P}_2\overline{\mathcal{D}}_1\big)\check{\hat{Z}}_2+\big(\mathcal{P}_1\overline{\mathcal{E}}_1+\mathcal{P}_2\widetilde{\mathcal{E}}_1+\mathcal{P}_2\overline{\mathcal{E}}_1\big)\check{\hat{Z}}_3,\\
Z_2(t) =& \big(\mathcal{P}_1\mathcal{A}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_1\big)X+\big(\mathcal{P}_1\widehat{\mathcal{A}}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_2\big)\hat{X}+\big[\mathcal{P}_1\mathcal{B}_2\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_3\big)\\
&+\mathcal{P}_3\big(\mathcal{A}_2+\widetilde{\mathcal{A}}_2\big)+\mathcal{P}_3\big(\mathcal{B}_2+\widetilde{\mathcal{B}}_2\big)\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}+\big[\mathcal{P}_1\overline{\mathcal{A}}_2+\mathcal{P}_1\mathcal{B}_2\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_2\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_1\overline{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_3\big(\widehat{\mathcal{A}}_2+\overline{\mathcal{A}}_2\big)+\mathcal{P}_3\big(\mathcal{B}_2+\widetilde{\mathcal{B}}_2\big)\big(\mathcal{P}_2+\mathcal{P}_4\big)\\
&+\mathcal{P}_3\overline{\mathcal{B}}_2\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_2Z_1+\big(\mathcal{P}_1\widetilde{\mathcal{C}}_2+\mathcal{P}_3\mathcal{C}_2+\mathcal{P}_3\widetilde{\mathcal{C}}_2\big)\check{Z}_1+\mathcal{P}_1\mathcal{D}_2Z_2\\
&+\big(\mathcal{P}_1\widetilde{\mathcal{D}}_2+\mathcal{P}_3\mathcal{D}_2+\mathcal{P}_3\widetilde{\mathcal{D}}_2\big)\check{Z}_2+\mathcal{P}_1\mathcal{E}_2Z_3+\big(\mathcal{P}_1\widetilde{\mathcal{E}}_2+\mathcal{P}_3\mathcal{E}_2+\mathcal{P}_3\widetilde{\mathcal{E}}_2\big)\check{Z}_3\\
&+\big(\mathcal{P}_1\overline{\mathcal{C}}_2+\mathcal{P}_3\overline{\mathcal{C}}_2\big)\check{\hat{Z}}_1+\big(\mathcal{P}_1\overline{\mathcal{D}}_2+\mathcal{P}_3\overline{\mathcal{D}}_2\big)\check{\hat{Z}}_2+\big(\mathcal{P}_1\overline{\mathcal{E}}_2+\mathcal{P}_3\overline{\mathcal{E}}_2\big)\check{\hat{Z}}_3,\\
Z_3(t) =& \big(\mathcal{P}_1\mathcal{A}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_1\big)X+\big[\mathcal{P}_1\widehat{\mathcal{A}}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_2+\mathcal{P}_2(\mathcal{A}_3+\widehat{\mathcal{A}}_3)+\mathcal{P}_2\mathcal{B}_3\big(\mathcal{P}_1+\mathcal{P}_2\big)\big]\hat{X}\\
&+\big[\mathcal{P}_1\mathcal{B}_3\mathcal{P}_3+\mathcal{P}_1\widetilde{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_3\big)+\mathcal{P}_3\big(\mathcal{A}_3+\widetilde{\mathcal{A}}_3\big)+\mathcal{P}_3\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_3\big)\big]\check{X}\\
&+\Big[\mathcal{P}_1\overline{\mathcal{A}}_3+\mathcal{P}_1\mathcal{B}_3\mathcal{P}_4+\mathcal{P}_1\widetilde{\mathcal{B}}_3\big(\mathcal{P}_2+\mathcal{P}_4\big)+\mathcal{P}_1\overline{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_2\mathcal{B}_3\big(\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_2\big(\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_3\big(\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_3\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3\big)\big(\mathcal{P}_2+\mathcal{P}_4\big)+\mathcal{P}_3\overline{\mathcal{B}}_3\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)+\mathcal{P}_4\big(\mathcal{A}_3+\widetilde{\mathcal{A}}_3+\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3\big)\\
&+\mathcal{P}_4\big(\mathcal{B}_3+\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3\big)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big)\Big]\check{\hat{X}}+\mathcal{P}_1\mathcal{B}_3Z_1+\mathcal{P}_2\mathcal{C}_3\hat{Z}_1+\big(\mathcal{P}_1\widetilde{\mathcal{C}}_3+\mathcal{P}_3\mathcal{C}_3\\
&+\mathcal{P}_3\widetilde{\mathcal{C}}_3\big)\check{Z}_1+\mathcal{P}_1\mathcal{D}_3Z_2+\mathcal{P}_2\mathcal{D}_3\hat{Z}_2+\big(\mathcal{P}_1\widetilde{\mathcal{D}}_3+\mathcal{P}_3\mathcal{D}_3+\mathcal{P}_3\widetilde{\mathcal{D}}_3\big)\check{Z}_2+\mathcal{P}_1\mathcal{E}_3Z_3+\mathcal{P}_2\mathcal{E}_3\hat{Z}_3\\
&+\big(\mathcal{P}_1\widetilde{\mathcal{E}}_3+\mathcal{P}_3\mathcal{E}_3+\mathcal{P}_3\widetilde{\mathcal{E}}_3\big)\check{Z}_3+\big[\mathcal{P}_1\overline{\mathcal{C}}_3+\mathcal{P}_2\widetilde{\mathcal{C}}_3+\mathcal{P}_2\overline{\mathcal{C}}_3+\mathcal{P}_3\overline{\mathcal{C}}_3+\mathcal{P}_4\big(\mathcal{C}_3+\widetilde{\mathcal{C}}_3+\overline{\mathcal{C}}_3\big)\big]\check{\hat{Z}}_1\\
&+\big[\mathcal{P}_1\overline{\mathcal{D}}_3+\mathcal{P}_2\widetilde{\mathcal{D}}_3+\mathcal{P}_2\overline{\mathcal{D}}_3+\mathcal{P}_3\overline{\mathcal{D}}_3+\mathcal{P}_4\big(\mathcal{D}_3+\widetilde{\mathcal{D}}_3+\overline{\mathcal{D}}_3\big)\big]\check{\hat{Z}}_2\\
&+\big[\mathcal{P}_1\overline{\mathcal{E}}_3+\mathcal{P}_2\widetilde{\mathcal{E}}_3+\mathcal{P}_2\overline{\mathcal{E}}_3+\mathcal{P}_3\overline{\mathcal{E}}_3+\mathcal{P}_4\big(\mathcal{E}_3+\widetilde{\mathcal{E}}_3+\overline{\mathcal{E}}_3\big)\big]\check{\hat{Z}}_3.
\end{aligned}
\tag{105}
$$

Next, we wish to represent each $Z_i(\cdot)$ and its filtering estimates as functionals of the "state" $X(\cdot)$ and its filtering estimates, from (105). For this target, we need the following four steps.

Step 1. Taking $\mathbb{E}\big[\mathbb{E}[\cdot|\mathcal{G}_t^2]\big|\mathcal{G}_t^1\big]$ on both sides of (105), we derive

$$
\check{\hat{Z}}_i(t)=\mathcal{M}_{i0}\check{\hat{X}}(t)+\mathcal{M}_{i1}\check{\hat{Z}}_1(t)+\mathcal{M}_{i2}\check{\hat{Z}}_2(t)+\mathcal{M}_{i3}\check{\hat{Z}}_3(t),\ i=1,2,3,
\tag{106}
$$

where

$$
\begin{cases}
\mathcal{M}_{10}:=\mathcal{P}_1(\mathcal{A}_1+\widehat{\mathcal{A}}_1+\overline{\mathcal{A}}_1)+\mathcal{P}_2(\mathcal{A}_1+\widehat{\mathcal{A}}_1+\widetilde{\mathcal{A}}_1+\overline{\mathcal{A}}_1)+(\mathcal{P}_1+\mathcal{P}_2)(\mathcal{B}_1+\widetilde{\mathcal{B}}_1+\overline{\mathcal{B}}_1)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big),\\
\mathcal{M}_{20}:=\mathcal{P}_1(\mathcal{A}_2+\widehat{\mathcal{A}}_2+\overline{\mathcal{A}}_2)+\mathcal{P}_3(\mathcal{A}_2+\widehat{\mathcal{A}}_2+\widetilde{\mathcal{A}}_2+\overline{\mathcal{A}}_2)+(\mathcal{P}_1+\mathcal{P}_3)(\mathcal{B}_2+\widetilde{\mathcal{B}}_2+\overline{\mathcal{B}}_2)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big),\\
\mathcal{M}_{30}:=\mathcal{P}_1(\mathcal{A}_3+\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3)+(\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)(\mathcal{A}_3+\widehat{\mathcal{A}}_3+\widetilde{\mathcal{A}}_3+\overline{\mathcal{A}}_3)\\
\qquad+(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)(\mathcal{B}_3+\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3)\big(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4\big),
\end{cases}
$$

$$\begin{cases} \mathcal{M}_{11} := \mathcal{P}_1\mathcal{B}_1 + \mathcal{P}_2\mathcal{C}_1 + \mathcal{P}_1\widetilde{\mathcal{C}}_1 + \mathcal{P}_1\overline{\mathcal{C}}_1 + \mathcal{P}_2\widetilde{\mathcal{C}}_1 + \mathcal{P}_2\overline{\mathcal{C}}_1,\ \mathcal{M}_{12} := \mathcal{P}_1\mathcal{D}_1 + \mathcal{P}_2\mathcal{D}_1 + \mathcal{P}_1\widetilde{\mathcal{D}}_1 + \mathcal{P}_1\overline{\mathcal{D}}_1 + \mathcal{P}_2\widetilde{\mathcal{D}}_1 + \mathcal{P}_2\overline{\mathcal{D}}_1, \\ \mathcal{M}_{13} := \mathcal{P}_1\mathcal{E}_1 + \mathcal{P}_2\mathcal{E}_1 + \mathcal{P}_1\widetilde{\mathcal{E}}_1 + \mathcal{P}_1\overline{\mathcal{E}}_1 + \mathcal{P}_2\widetilde{\mathcal{E}}_1 + \mathcal{P}_2\overline{\mathcal{E}}_1,\ \mathcal{M}_{21} := \mathcal{P}_1\mathcal{B}_2 + \mathcal{P}_1\widetilde{\mathcal{C}}_2 + \mathcal{P}_3\mathcal{C}_2 + \mathcal{P}_3\widetilde{\mathcal{C}}_2 + \mathcal{P}_1\overline{\mathcal{C}}_2 + \mathcal{P}_3\overline{\mathcal{C}}_2, \\ \mathcal{M}_{22} := \mathcal{P}_1\mathcal{D}_2 + \mathcal{P}_1\widetilde{\mathcal{D}}_2 + \mathcal{P}_3\mathcal{D}_2 + \mathcal{P}_3\widetilde{\mathcal{D}}_2 + \mathcal{P}_1\overline{\mathcal{D}}_2 + \mathcal{P}_3\overline{\mathcal{D}}_2,\ \mathcal{M}_{23} := \mathcal{P}_1\mathcal{E}_2 + \mathcal{P}_1\widetilde{\mathcal{E}}_2 + \mathcal{P}_3\mathcal{E}_2 + \mathcal{P}_3\widetilde{\mathcal{E}}_2 + \mathcal{P}_1\overline{\mathcal{E}}_2 + \mathcal{P}_3\overline{\mathcal{E}}_2, \\ \mathcal{M}_{31} := \mathcal{P}_1\mathcal{B}_3 + \mathcal{P}_1\widetilde{\mathcal{C}}_3 + \mathcal{P}_1\overline{\mathcal{C}}_3 + (\mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4)(\mathcal{C}_3 + \widetilde{\mathcal{C}}_3 + \overline{\mathcal{C}}_3),\ \mathcal{M}_{32} := \mathcal{P}_1\mathcal{D}_3 + \mathcal{P}_1\widetilde{\mathcal{D}}_3 + \mathcal{P}_1\overline{\mathcal{D}}_3 \\ \qquad + (\mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4)(\mathcal{D}_3 + \widetilde{\mathcal{D}}_3 + \overline{\mathcal{D}}_3),\ \mathcal{M}_{33} := \mathcal{P}_1\mathcal{E}_3 + \mathcal{P}_1\widetilde{\mathcal{E}}_3 + \mathcal{P}_1\overline{\mathcal{E}}_3 + (\mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4)(\mathcal{E}_3 + \widetilde{\mathcal{E}}_3 + \overline{\mathcal{E}}_3). \end{cases}$$

We rewrite (106) as

$$\begin{pmatrix} I_n - \mathcal{M}_{11} & -\mathcal{M}_{12} & -\mathcal{M}_{13} \\ -\mathcal{M}_{21} & I_n - \mathcal{M}_{22} & -\mathcal{M}_{23} \\ -\mathcal{M}_{31} & -\mathcal{M}_{32} & I_n - \mathcal{M}_{33} \end{pmatrix} \begin{pmatrix} \check{\check{Z}}_1(t) \\ \check{\check{Z}}_2(t) \\ \check{\check{Z}}_3(t) \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{10} \\ \mathcal{M}_{20} \\ \mathcal{M}_{30} \end{pmatrix} \check{\check{X}}(t). \tag{107}$$

If we assume that

(A2.5) the coefficient matrix of (107) is invertible, for any $t \in [0,T]$,

then by Cramer's rule, we have

$$\check{\check{Z}}_i(t) = (-1)^{i-1}\big(\mathbb{N}_1\big)^{-1}\big[\mathcal{M}_{10}\mathbf{M}_{1i} - \mathcal{M}_{20}\mathbf{M}_{2i} + \mathcal{M}_{30}\mathbf{M}_{3i}\big]\check{\check{X}}(t) := \mathcal{N}^i(t)\check{\check{X}}(t),\ i = 1,2,3, \tag{108}$$

where $\mathbb{N}_1$ is the determinant of the coefficient of (107), and $\mathbf{M}^{ji}(t)$ is the adjoint matrix of the (j,i) element in (107), for $j,i = 1,2,3$.

Step 2. Taking $\mathbb{E}[\cdot|\mathcal{G}_t^2]$ on both sides of (105), we get

$$\begin{aligned} \check{Z}_i(t) = {} & \widetilde{\mathcal{M}}_{i0}\check{X}(t) + \overline{\mathcal{M}}_{i0}\check{\check{X}}(t) + \widetilde{\mathcal{M}}_{i1}\check{Z}_1(t) + \overline{\mathcal{M}}_{i1}\check{\check{Z}}_1(t) \\ & + \widetilde{\mathcal{M}}_{i2}\check{Z}_2(t) + \overline{\mathcal{M}}_{i2}\check{\check{Z}}_2(t) + \widetilde{\mathcal{M}}_{i3}\check{Z}_3(t) + \overline{\mathcal{M}}_{i3}\check{\check{Z}}_3(t),\ i = 1,2,3, \end{aligned} \tag{109}$$

where

$$\begin{cases} \widetilde{\mathcal{M}}_{10} := \mathcal{P}_1\mathcal{A}_1 + \mathcal{P}_1(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1)(\mathcal{P}_1 + \mathcal{P}_3),\ \overline{\mathcal{M}}_{10} := \mathcal{P}_1(\widehat{\mathcal{A}}_1 + \overline{\mathcal{A}}_1) + \mathcal{P}_2(\mathcal{A}_1 + \widehat{\mathcal{A}}_1 + \widetilde{\mathcal{A}}_1 + \overline{\mathcal{A}}_1) + \mathcal{P}_1(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1)\big(\mathcal{P}_2 + \mathcal{P}_4\big) \\ \qquad + \big(\mathcal{P}_1\overline{\mathcal{B}}_1 + \mathcal{P}_2\mathcal{B}_1 + \mathcal{P}_2\widetilde{\mathcal{B}}_1 + \mathcal{P}_2\overline{\mathcal{B}}_1\big)\big(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4\big), \\ \widetilde{\mathcal{M}}_{20} := \mathcal{P}_1\mathcal{A}_2 + \mathcal{P}_3(\mathcal{A}_2 + \widetilde{\mathcal{A}}_2) + (\mathcal{P}_1 + \mathcal{P}_3)(\mathcal{B}_2 + \widetilde{\mathcal{B}}_2)(\mathcal{P}_1 + \mathcal{P}_3), \\ \overline{\mathcal{M}}_{20} := (\mathcal{P}_1 + \mathcal{P}_3)(\widehat{\mathcal{A}}_2 + \overline{\mathcal{A}}_2) + (\mathcal{P}_1 + \mathcal{P}_3)(\mathcal{B}_2 + \widetilde{\mathcal{B}}_2)\big(\mathcal{P}_2 + \mathcal{P}_4\big) + (\mathcal{P}_1 + \mathcal{P}_3)\overline{\mathcal{B}}_2\big(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4\big), \\ \widetilde{\mathcal{M}}_{30} := \mathcal{P}_1\mathcal{A}_3 + \mathcal{P}_3(\mathcal{A}_3 + \widetilde{\mathcal{A}}_3) + (\mathcal{P}_1 + \mathcal{P}_3)(\mathcal{B}_3 + \widetilde{\mathcal{B}}_3)(\mathcal{P}_1 + \mathcal{P}_3), \\ \overline{\mathcal{M}}_{30} := (\mathcal{P}_1 + \mathcal{P}_3)(\widehat{\mathcal{A}}_3 + \overline{\mathcal{A}}_3) + (\mathcal{P}_2 + \mathcal{P}_4)(\mathcal{A}_3 + \widehat{\mathcal{A}}_3 + \widetilde{\mathcal{A}}_3 + \overline{\mathcal{A}}_3) + (\mathcal{P}_1 + \mathcal{P}_3)(\mathcal{B}_3 + \widetilde{\mathcal{B}}_3)\big(\mathcal{P}_2 + \mathcal{P}_4\big) \\ \qquad + \big[(\mathcal{P}_2 + \mathcal{P}_2)(\mathcal{B}_3 + \widetilde{\mathcal{B}}_3 + \overline{\mathcal{B}}_3) + \mathcal{P}_1\overline{\mathcal{B}}_3 + \mathcal{P}_3\widetilde{\mathcal{B}}_3\big]\big(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4\big), \\ \widetilde{\mathcal{M}}_{11} := \mathcal{P}_1\mathcal{B}_1 + \mathcal{P}_1\widetilde{\mathcal{C}}_1, \quad \overline{\mathcal{M}}_{11} := \mathcal{P}_2\mathcal{C}_1 + \mathcal{P}_1\overline{\mathcal{C}}_1 + \mathcal{P}_2\widetilde{\mathcal{C}}_1 + \mathcal{P}_2\overline{\mathcal{C}}_1, \quad \widetilde{\mathcal{M}}_{12} := \mathcal{P}_1\mathcal{D}_1 + \mathcal{P}_1\widetilde{\mathcal{D}}_1, \\ \overline{\mathcal{M}}_{12} := \mathcal{P}_2\mathcal{D}_1 + \mathcal{P}_1\overline{\mathcal{D}}_1 + \mathcal{P}_2\widetilde{\mathcal{D}}_1 + \mathcal{P}_2\overline{\mathcal{D}}_1, \quad \widetilde{\mathcal{M}}_{13} := \mathcal{P}_1\mathcal{E}_1 + \mathcal{P}_1\widetilde{\mathcal{E}}_1, \quad \overline{\mathcal{M}}_{13} := \mathcal{P}_2\mathcal{E}_1 + \mathcal{P}_1\overline{\mathcal{E}}_1 + \mathcal{P}_2\widetilde{\mathcal{E}}_1 + \mathcal{P}_2\overline{\mathcal{E}}_1, \\ \widetilde{\mathcal{M}}_{21} := \mathcal{P}_1\mathcal{B}_2 + \mathcal{P}_1\widetilde{\mathcal{C}}_2 + \mathcal{P}_3\mathcal{C}_2 + \mathcal{P}_3\widetilde{\mathcal{C}}_2, \quad \overline{\mathcal{M}}_{21} := \mathcal{P}_1\overline{\mathcal{C}}_2 + \mathcal{P}_3\overline{\mathcal{C}}_2, \quad \widetilde{\mathcal{M}}_{22} := \mathcal{P}_1\mathcal{D}_2 + \mathcal{P}_1\widetilde{\mathcal{D}}_2 + \mathcal{P}_3\mathcal{D}_2 + \mathcal{P}_3\widetilde{\mathcal{D}}_2, \\ \overline{\mathcal{M}}_{22} := \mathcal{P}_1\overline{\mathcal{D}}_2 + \mathcal{P}_3\overline{\mathcal{D}}_2, \quad \widetilde{\mathcal{M}}_{23} := \mathcal{P}_1\mathcal{E}_2 + \mathcal{P}_1\widetilde{\mathcal{E}}_2 + \mathcal{P}_3\mathcal{E}_2 + \mathcal{P}_3\widetilde{\mathcal{E}}_2, \quad \overline{\mathcal{M}}_{23} := \mathcal{P}_1\overline{\mathcal{E}}_2 + \mathcal{P}_3\overline{\mathcal{E}}_2, \\ \widetilde{\mathcal{M}}_{31} := \mathcal{P}_1\mathcal{B}_3 + \mathcal{P}_1\widetilde{\mathcal{C}}_3 + \mathcal{P}_3\widetilde{\mathcal{C}}_3 + \mathcal{P}_3\widetilde{\mathcal{C}}_3, \quad \overline{\mathcal{M}}_{31} := (\mathcal{P}_1 + \mathcal{P}_3)\overline{\mathcal{C}}_3 + (\mathcal{P}_2 + \mathcal{P}_4)(\mathcal{C}_3 + \widetilde{\mathcal{C}}_3 + \overline{\mathcal{C}}_3), \\ \widetilde{\mathcal{M}}_{32} := \mathcal{P}_1\mathcal{D}_3 + \mathcal{P}_1\widetilde{\mathcal{D}}_3 + \mathcal{P}_3\widetilde{\mathcal{D}}_3 + \mathcal{P}_3\widetilde{\mathcal{D}}_3, \quad \overline{\mathcal{M}}_{32} := (\mathcal{P}_1 + \mathcal{P}_3)\overline{\mathcal{D}}_3 + (\mathcal{P}_2 + \mathcal{P}_4)(\mathcal{D}_3 + \widetilde{\mathcal{D}}_3 + \overline{\mathcal{D}}_3), \\ \widetilde{\mathcal{M}}_{33} := \mathcal{P}_1\mathcal{E}_3 + \mathcal{P}_1\widetilde{\mathcal{E}}_3 + \mathcal{P}_3\widetilde{\mathcal{E}}_3 + \mathcal{P}_3\widetilde{\mathcal{E}}_3, \quad \overline{\mathcal{M}}_{33} := (\mathcal{P}_1 + \mathcal{P}_3)\overline{\mathcal{E}}_3 + (\mathcal{P}_2 + \mathcal{P}_4)(\mathcal{E}_3 + \widetilde{\mathcal{E}}_3 + \overline{\mathcal{E}}_3). \end{cases}$$

Putting (108) into (109), we get

$$\begin{aligned}\check{Z}^i(t) &= \widetilde{\mathcal{M}}_{i0}\check{X}(t) + \Big[\overline{\mathcal{M}}_{i0} + \overline{\mathcal{M}}_{i1}\mathcal{N}_1 + \overline{\mathcal{M}}_{i2}\mathcal{N}_2 + \overline{\mathcal{M}}_{i3}\mathcal{N}_3\Big]\check{\check{X}}(t) + \widetilde{\mathcal{M}}_{i1}\check{Z}_1(t) + \widetilde{\mathcal{M}}_{i2}\check{Z}_2(t) + \widetilde{\mathcal{M}}_{i3}\check{Z}_3(t)\\ &:= \widetilde{\mathcal{M}}^{i0}(t)\check{X}(t) + \overline{\mathcal{N}}^{i0}(t)\check{\check{X}}(t) + \widetilde{\mathcal{M}}_{i1}\check{Z}_1(t) + \widetilde{\mathcal{M}}_{i2}\check{Z}_2(t) + \widetilde{\mathcal{M}}_{i3}\check{Z}_3(t),\ i = 1, 2, 3.\end{aligned} \tag{110}$$

We rewrite (110) as

$$\begin{pmatrix} I_n - \widetilde{\mathcal{M}}_{11} & -\widetilde{\mathcal{M}}_{12} & -\widetilde{\mathcal{M}}_{13} \\ -\widetilde{\mathcal{M}}_{21} & I_n - \widetilde{\mathcal{M}}_{22} & -\widetilde{\mathcal{M}}_{23} \\ -\widetilde{\mathcal{M}}_{31} & -\widetilde{\mathcal{M}}_{32} & I_n - \widetilde{\mathcal{M}}_{33} \end{pmatrix} \begin{pmatrix} \check{Z}_1(t) \\ \check{Z}_2(t) \\ \check{Z}_3(t) \end{pmatrix} = \begin{pmatrix} \widetilde{\mathcal{M}}_{10}\check{X}(t) + \overline{\mathcal{N}}_{10}\check{\check{X}}(t) \\ \widetilde{\mathcal{M}}_{20}\check{X}(t) + \overline{\mathcal{N}}_{20}\check{\check{X}}(t) \\ \widetilde{\mathcal{M}}_{30}\check{X}(t) + \overline{\mathcal{N}}_{30}\check{\check{X}}(t) \end{pmatrix}. \tag{111}$$

Similarly, if we assume that

(A2.6) the coefficient matrix of (111) is invertible, for any $t \in [0, T]$,

then we have

$$\begin{aligned}\check{Z}_i(t) &= (-1)^{i-1}\big(\mathbb{N}_2\big)^{-1}\Big[\big(\widetilde{\mathcal{M}}_{10}\check{X}(t) + \overline{\mathcal{N}}_{10}\check{\check{X}}(t)\big)\widetilde{\mathbf{M}}_{1i} - \big(\widetilde{\mathcal{M}}_{20}\check{X}(t) + \overline{\mathcal{N}}_{20}\check{\check{X}}(t)\big)\widetilde{\mathbf{M}}_{2i}\\ &\qquad + \big(\widetilde{\mathcal{M}}_{30}\check{X}(t) + \overline{\mathcal{N}}_{30}(t)\check{\check{X}}(t)\big)\widetilde{\mathbf{M}}_{3i}\Big]\\ &= (-1)^{i-1}\big(\mathbb{N}_2\big)^{-1}\Big[\widetilde{\mathcal{M}}_{10}\widetilde{\mathbf{M}}_{1i} - \widetilde{\mathcal{M}}_{20}\widetilde{\mathbf{M}}_{2i} + \widetilde{\mathcal{M}}_{30}\widetilde{\mathbf{M}}_{3i}\Big]\check{X}(t)\\ &\qquad + \big(\mathbb{N}_2\big)^{-1}\Big[\overline{\mathcal{N}}_{10}\widetilde{\mathbf{M}}_{1i} - \overline{\mathcal{N}}_{20}\widetilde{\mathbf{M}}_{2i} + \overline{\mathcal{N}}_{30}\widetilde{\mathbf{M}}_{3i}\Big]\check{\check{X}}(t) := \widetilde{\mathcal{N}}_i\check{X}(t) + \overline{\mathcal{N}}_i\check{\check{X}}(t),\ i = 1, 2, 3,\end{aligned} \tag{112}$$

where $\mathbb{N}_2$ is the determinant of the coefficient of (111), and $\widetilde{\mathbf{M}}_{ji}$ is the adjoint matrix of the (j, i) element in (111), for $j, i = 1, 2, 3$.

Step 3. Taking $\mathbb{E}[\cdot|\mathcal{G}_t^1]$ on both sides of (105), we obtain

$$\hat{Z}_i(t) = \widehat{\mathcal{M}}_{i0}\hat{X}(t) + \overline{\overline{\mathcal{M}}}_{i0}\check{\hat{X}}(t) + \widehat{\mathcal{M}}_{i1}\hat{Z}_1(t) + \overline{\overline{\mathcal{M}}}_{i1}\check{\hat{Z}}_1(t) + \widehat{\mathcal{M}}_{i2}\hat{Z}_2(t) + \overline{\overline{\mathcal{M}}}_{i2}\check{\hat{Z}}_2(t) + \widehat{\mathcal{M}}_{i3}\hat{Z}_3(t) + \overline{\overline{\mathcal{M}}}_{i3}\check{\hat{Z}}_3(t), \tag{113}$$

$i = 1, 2, 3$, where

$$\begin{cases} \widehat{\mathcal{M}}_{10} := (\mathcal{P}_1 + \mathcal{P}_2)(\mathcal{A}_1 + \widehat{\mathcal{A}}_1) + (\mathcal{P}_1 + \mathcal{P}_2)\mathcal{B}_1(\mathcal{P}_1 + \mathcal{P}_2), \quad \widehat{\mathcal{M}}_{20} := \mathcal{P}_1(\mathcal{A}_2 + \widehat{\mathcal{A}}_2) + \mathcal{P}_1\mathcal{B}_2(\mathcal{P}_1 + \mathcal{P}_2),\\ \overline{\overline{\mathcal{M}}}_{10} := (\mathcal{P}_1 + \mathcal{P}_2)\overline{\mathcal{A}}_1 + \mathcal{P}_2\widetilde{\mathcal{A}}_1 + (\mathcal{P}_1 + \mathcal{P}_2)\mathcal{B}_1(\mathcal{P}_3 + \mathcal{P}_4) + (\mathcal{P}_1 + \mathcal{P}_2)(\widetilde{\mathcal{B}}_1 + \overline{\mathcal{B}}_1)(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4),\\ \overline{\overline{\mathcal{M}}}_{20} := \mathcal{P}_1\overline{\mathcal{A}}_2 + \mathcal{P}_3(\mathcal{A}_2 + \widehat{\mathcal{A}}_2 + \widetilde{\mathcal{A}}_2 + \overline{\mathcal{A}}_2) + \mathcal{P}_1\mathcal{B}_2(\mathcal{P}_3 + \mathcal{P}_4) + (\mathcal{P}_1\widetilde{\mathcal{B}}_2 + \mathcal{P}_1\overline{\mathcal{B}}_2 + \mathcal{P}_3\mathcal{B}_2 + \mathcal{P}_3\widetilde{\mathcal{B}}_2\\ \qquad + \mathcal{P}_3\overline{\mathcal{B}}_3)(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4), \quad \widehat{\mathcal{M}}_{30} := (\mathcal{P}_1 + \mathcal{P}_2)(\mathcal{A}_3 + \widehat{\mathcal{A}}_3) + (\mathcal{P}_1 + \mathcal{P}_2)\mathcal{B}_3(\mathcal{P}_1 + \mathcal{P}_2),\\ \overline{\overline{\mathcal{M}}}_{30} := \mathcal{P}_1\overline{\mathcal{A}}_3 + \mathcal{P}_2(\widetilde{\mathcal{A}}_3 + \widehat{\mathcal{A}}_3) + (\mathcal{P}_3 + \mathcal{P}_4)(\mathcal{A}_3 + \widehat{\mathcal{A}}_3 + \widetilde{\mathcal{A}}_3 + \overline{\mathcal{A}}_3) + (\mathcal{P}_1 + \mathcal{P}_2)\mathcal{B}_3(\mathcal{P}_3 + \mathcal{P}_4)\\ \qquad + \big[(\mathcal{P}_1 + \mathcal{P}_2)(\widetilde{\mathcal{B}}_3 + \overline{\mathcal{B}}_3) + (\mathcal{P}_3 + \mathcal{P}_4)(\mathcal{B}_3 + \widetilde{\mathcal{B}}_2 + \overline{\mathcal{B}}_3)\big](\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4),\\ \widehat{\mathcal{M}}_{11} := \mathcal{P}_1\mathcal{B}_1 + \mathcal{P}_2\mathcal{C}_1, \quad \overline{\overline{\mathcal{M}}}_{11} := \mathcal{P}_1\widetilde{\mathcal{C}}_1 + \mathcal{P}_1\overline{\mathcal{C}}_1 + \mathcal{P}_2\widetilde{\mathcal{C}}_1 + \mathcal{P}_2\overline{\mathcal{C}}_1, \quad \widehat{\mathcal{M}}_{12} := \mathcal{P}_1\mathcal{D}_1 + \mathcal{P}_2\mathcal{D}_1,\\ \overline{\overline{\mathcal{M}}}_{12} := \mathcal{P}_1\widetilde{\mathcal{D}}_1 + \mathcal{P}_1\overline{\mathcal{D}}_1 + \mathcal{P}_2\widetilde{\mathcal{D}}_1 + \mathcal{P}_2\overline{\mathcal{D}}_1, \quad \widehat{\mathcal{M}}_{13} := \mathcal{P}_1\mathcal{E}_1 + \mathcal{P}_2\mathcal{E}_1, \quad \overline{\overline{\mathcal{M}}}_{13} := \mathcal{P}_1\widetilde{\mathcal{E}}_1 + \mathcal{P}_1\overline{\mathcal{E}}_1 + \mathcal{P}_2\widetilde{\mathcal{E}}_1 + \mathcal{P}_2\overline{\mathcal{E}}_1,\\ \widehat{\mathcal{M}}_{21} := \mathcal{P}_1\mathcal{B}_2, \quad \overline{\overline{\mathcal{M}}}_{21} := \mathcal{P}_1\widetilde{\mathcal{C}}_2 + \mathcal{P}_1\overline{\mathcal{C}}_2 + \mathcal{P}_3\mathcal{C}_2 + \mathcal{P}_3\widetilde{\mathcal{C}}_2 + \mathcal{P}_3\overline{\mathcal{C}}_2, \quad \widehat{\mathcal{M}}_{22} := \mathcal{P}_1\mathcal{D}_2,\\ \overline{\overline{\mathcal{M}}}_{22} := \mathcal{P}_1\widetilde{\mathcal{D}}_2 + \mathcal{P}_1\overline{\mathcal{D}}_2 + \mathcal{P}_3\mathcal{D}_2 + \mathcal{P}_3\widetilde{\mathcal{D}}_2 + \mathcal{P}_3\overline{\mathcal{D}}_2, \quad \widehat{\mathcal{M}}_{23} := \mathcal{P}_1\mathcal{E}_2, \quad \overline{\overline{\mathcal{M}}}_{23} := \mathcal{P}_1\widetilde{\mathcal{E}}_2 + \mathcal{P}_1\overline{\mathcal{E}}_2 + \mathcal{P}_3\mathcal{E}_2 + \mathcal{P}_3\widetilde{\mathcal{E}}_2 + \mathcal{P}_3\overline{\mathcal{E}}_2,\\ \widehat{\mathcal{M}}_{31} := \mathcal{P}_1\mathcal{B}_3 + \mathcal{P}_2\mathcal{C}_3, \quad \overline{\overline{\mathcal{M}}}_{31} := (\mathcal{P}_1 + \mathcal{P}_2)(\widetilde{\mathcal{C}}_3 + \overline{\mathcal{C}}_3) + (\mathcal{P}_3 + \mathcal{P}_4)(\mathcal{C}_3 + \widetilde{\mathcal{C}}_3 + \overline{\mathcal{C}}_3),\ \widehat{\mathcal{M}}_{32} := \mathcal{P}_1\mathcal{D}_3 + \mathcal{P}_2\mathcal{D}_3,\\ \overline{\overline{\mathcal{M}}}_{32} := (\mathcal{P}_1 + \mathcal{P}_2)(\widetilde{\mathcal{D}}_3 + \overline{\mathcal{D}}_3) + (\mathcal{P}_3 + \mathcal{P}_4)(\mathcal{D}_3 + \widetilde{\mathcal{D}}_3 + \overline{\mathcal{D}}_3),\ \widehat{\mathcal{M}}_{33} := \mathcal{P}_1\mathcal{E}_3 + \mathcal{P}_2\mathcal{E}_3,\\ \overline{\overline{\mathcal{M}}}_{33} := (\mathcal{P}_1 + \mathcal{P}_2)(\widetilde{\mathcal{E}}_3 + \overline{\mathcal{E}}_3) + (\mathcal{P}_3 + \mathcal{P}_4)(\mathcal{E}_3 + \widetilde{\mathcal{E}}_3 + \overline{\mathcal{E}}_3). \end{cases}$$

Putting (108) into (113), we get

$$\begin{aligned}\hat{Z}_i(t) &= \widehat{\mathcal{M}}_{i0}\hat{X}(t) + \Big[\overline{\overline{\mathcal{M}}}_{i0} + \overline{\overline{\mathcal{M}}}_{i1}\mathcal{N}_1 + \overline{\overline{\mathcal{M}}}_{i2}\mathcal{N}_2 + \overline{\overline{\mathcal{M}}}_{i3}\mathcal{N}_3\Big]\check{\check{X}}(t) + \widehat{\mathcal{M}}_{i1}\hat{Z}_1(t) + \widehat{\mathcal{M}}_{i2}\hat{Z}_2(t) + \widehat{\mathcal{M}}_{i3}\hat{Z}_3(t) \\ &:= \widehat{\mathcal{M}}_{i0}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{i0}\check{\check{X}}(t) + \widehat{\mathcal{M}}_{i1}\hat{Z}_1(t) + \widehat{\mathcal{M}}_{i2}\hat{Z}_2(t) + \widehat{\mathcal{M}}_{i3}\hat{Z}_3(t), \ i = 1,2,3.\end{aligned} \tag{114}$$

We rewrite (114) as

$$\begin{pmatrix} I_n - \widehat{\mathcal{M}}_{11} & -\widehat{\mathcal{M}}_{12} & -\widehat{\mathcal{M}}_{13} \\ -\widehat{\mathcal{M}}_{21} & I_n - \widehat{\mathcal{M}}_{22} & -\widehat{\mathcal{M}}_{23} \\ -\widehat{\mathcal{M}}_{31} & -\widehat{\mathcal{M}}_{32} & I_n - \widehat{\mathcal{M}}_{33} \end{pmatrix} \begin{pmatrix} \hat{Z}_1(t) \\ \hat{Z}_2(t) \\ \hat{Z}_3(t) \end{pmatrix} = \begin{pmatrix} \widehat{\mathcal{M}}_{10}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{10}\check{\check{X}}(t) \\ \widehat{\mathcal{M}}_{20}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{20}\check{\check{X}}(t) \\ \widehat{\mathcal{M}}_{30}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{30}\check{\check{X}}(t) \end{pmatrix}. \tag{115}$$

Similarly, if we assume that

(A2.7) the coefficient matrix of (115) is invertible, for any $t \in [0,T]$,

then we have

$$\begin{aligned}\hat{Z}_i(t) &= (-1)^{i-1}\big(\mathbb{N}_3\big)^{-1}\Big[\big(\widehat{\mathcal{M}}_{10}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{10}\check{\check{X}}(t)\big)\widehat{\mathbf{M}}_{1i} - \big(\widehat{\mathcal{M}}_{20}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{20}\check{\check{X}}(t)\big)\widehat{\mathbf{M}}_{2i} \\ &\qquad + \big(\widehat{\mathcal{M}}_{30}\hat{X}(t) + \overline{\overline{\mathcal{N}}}_{30}\check{\check{X}}(t)\big)\widehat{\mathbf{M}}_{3i}\Big] \\ &= (-1)^{i-1}\big(\mathbb{N}_3\big)^{-1}\Big[\widehat{\mathcal{M}}_{10}\widehat{\mathbf{M}}_{1i} - \widehat{\mathcal{M}}_{20}\widehat{\mathbf{M}}_{2i} + \widehat{\mathcal{M}}_{30}\widehat{\mathbf{M}}_{3i}\Big]\hat{X}(t) \\ &\qquad + \big(\mathbb{N}_3\big)^{-1}\Big[\overline{\overline{\mathcal{N}}}_{10}\widehat{\mathbf{M}}_{1i} - \overline{\overline{\mathcal{N}}}_{20}\widehat{\mathbf{M}}_{2i} + \overline{\overline{\mathcal{N}}}_{30}\widehat{\mathbf{M}}_{3i}\Big]\check{\check{X}}(t) := \widehat{\mathcal{N}}_i\hat{X}(t) + \overline{\overline{\mathcal{N}}}_i\check{\check{X}}(t), \ i = 1,2,3,\end{aligned} \tag{116}$$

where $\mathbb{N}_3$ is the determinant of the coefficient of (115), and $\widehat{\mathbf{M}}_{ji}$ is the adjoint matrix of the (j,i) element in (115), for $j,i = 1,2,3$.

Step 4. Putting (108), (112) and (116) into (105), we have

$$Z_i(t) = \Gamma_{i0}X(t) + \widehat{\Gamma}_{i0}\hat{X}(t) + \widetilde{\Gamma}_{i0}\check{X}(t) + \overline{\Gamma}_{i0}\check{\check{X}}(t) + \mathcal{P}_1\mathcal{B}_iZ_1(t) + \mathcal{P}_1\mathcal{D}_iZ_2(t) + \mathcal{P}_1\mathcal{E}_iZ_3(t), \ i = 1,2,3, \tag{117}$$

where

$$\begin{cases} \Gamma_{10} := \mathcal{P}_1\mathcal{A}_1 + \mathcal{P}_1\mathcal{B}_1\mathcal{P}_1, \quad \Gamma_{20} := \mathcal{P}_1\mathcal{A}_2 + \mathcal{P}_1\mathcal{B}_2\mathcal{P}_1, \quad \Gamma_{30} := \mathcal{P}_1\mathcal{A}_3 + \mathcal{P}_1\mathcal{B}_3\mathcal{P}_1, \\ \widehat{\Gamma}_{10} := \mathcal{P}_1\widehat{\mathcal{A}}_1 + \mathcal{P}_2\big(\mathcal{A}_1 + \widehat{\mathcal{A}}_1\big) + \mathcal{P}_1\mathcal{B}_1\mathcal{P}_2 + \mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_1 + \mathcal{P}_2\big) + \mathcal{P}_2\mathcal{C}_1\widehat{\mathcal{N}}_1 + \mathcal{P}_2\mathcal{D}_1\widehat{\mathcal{N}}_2 + \mathcal{P}_2\mathcal{E}_1\widehat{\mathcal{N}}_3, \\ \widetilde{\Gamma}_{10} := \mathcal{P}_1\mathcal{B}_1\mathcal{P}_3 + \mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_1 + \mathcal{P}_3\big) + \mathcal{P}_1\widetilde{\mathcal{C}}_1\widetilde{\mathcal{N}}_1 + \mathcal{P}_1\widetilde{\mathcal{D}}_1\widetilde{\mathcal{N}}_2 + \mathcal{P}_1\mathcal{E}_1\widetilde{\mathcal{N}}_3, \quad \widehat{\Gamma}_{20} := \mathcal{P}_1\widehat{\mathcal{A}}_2 + \mathcal{P}_1\mathcal{B}_2\mathcal{P}_2, \\ \overline{\Gamma}_{10} := \mathcal{P}_1\overline{\mathcal{A}}_1 + \mathcal{P}_1\mathcal{B}_1\mathcal{P}_4 + \mathcal{P}_1\widetilde{\mathcal{B}}_1\big(\mathcal{P}_2 + \mathcal{P}_4\big) + \mathcal{P}_1\overline{\mathcal{B}}_1\big(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4\big) + \mathcal{P}_2\big(\widetilde{\mathcal{A}}_1 + \overline{\mathcal{A}}_1\big) \\ \qquad + \mathcal{P}_2\mathcal{B}_1\big(\mathcal{P}_3 + \mathcal{P}_4\big) + \mathcal{P}_2\big(\widetilde{\mathcal{B}}_1 + \overline{\mathcal{B}}_1\big)\big(\mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4\big) + \mathcal{P}_2\mathcal{C}_1\overline{\overline{\mathcal{N}}}_1 + \mathcal{P}_1\widetilde{\mathcal{C}}_1\overline{\mathcal{N}}_1 \\ \qquad + \mathcal{P}_2\mathcal{D}_1\overline{\overline{\mathcal{N}}}_2 + \mathcal{P}_1\widetilde{\mathcal{D}}_1\overline{\mathcal{N}}_2 + \mathcal{P}_2\mathcal{E}_1\overline{\overline{\mathcal{N}}}_3 + \mathcal{P}_1\widetilde{\mathcal{E}}_1\overline{\mathcal{N}}_3 + \big(\mathcal{P}_1\overline{\mathcal{C}}_1 + \mathcal{P}_2\widetilde{\mathcal{C}}_1 + \mathcal{P}_2\overline{\mathcal{C}}_1\big)\mathcal{N}_1, \\ \widetilde{\Gamma}_{20} := \mathcal{P}_1\mathcal{B}_2\mathcal{P}_3 + \mathcal{P}_1\widetilde{\mathcal{B}}_2\big(\mathcal{P}_1 + \mathcal{P}_3\big) + \mathcal{P}_3\big(\mathcal{A}_2 + \widetilde{\mathcal{A}}_2\big) + \mathcal{P}_3\big(\mathcal{B}_2 + \widetilde{\mathcal{B}}_2\big)\big(\mathcal{P}_1 + \mathcal{P}_3\big) \\ \qquad + \big(\mathcal{P}_1\widetilde{\mathcal{C}}_2 + \mathcal{P}_3\mathcal{C}_2 + \mathcal{P}_3\widetilde{\mathcal{C}}_2\big)\widetilde{\mathcal{N}}_1 + \big(\mathcal{P}_1\widetilde{\mathcal{D}}_2 + \mathcal{P}_3\mathcal{D}_2 + \mathcal{P}_3\widetilde{\mathcal{D}}_2\big)\widetilde{\mathcal{N}}_2 + \big(\mathcal{P}_1\widetilde{\mathcal{E}}_2 + \mathcal{P}_3\mathcal{E}_2 + \mathcal{P}_3\widetilde{\mathcal{E}}_2\big)\widetilde{\mathcal{N}}_3, \end{cases}$$

$$\begin{cases}\overline{\Gamma}_{20} := \mathcal{P}_1\overline{\mathcal{A}}_2 + \mathcal{P}_1\mathcal{B}_2\mathcal{P}_4 + \mathcal{P}_1\widetilde{\mathcal{B}}_2(\mathcal{P}_2+\mathcal{P}_4) + \mathcal{P}_3(\widehat{\mathcal{A}}_2+\overline{\mathcal{A}}_2) + \mathcal{P}_1\overline{\mathcal{B}}_2(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\\ \qquad + \mathcal{P}_3(\mathcal{B}_2+\widetilde{\mathcal{B}}_2)(\mathcal{P}_2+\mathcal{P}_4) + \mathcal{P}_3\overline{\mathcal{B}}_2(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4)\\ \qquad + (\mathcal{P}_1\widetilde{\mathcal{C}}_2+\mathcal{P}_3\mathcal{C}_2+\mathcal{P}_3\widetilde{\mathcal{C}}_2)\overline{\mathcal{N}}_1 + (\mathcal{P}_1\widetilde{\mathcal{D}}_2+\mathcal{P}_3\mathcal{D}_2+\mathcal{P}_3\widetilde{\mathcal{D}}_2)\overline{\mathcal{N}}_2 + (\mathcal{P}_1\widetilde{\mathcal{E}}_2+\mathcal{P}_3\mathcal{E}_2+\mathcal{P}_3\widetilde{\mathcal{E}}_2)\overline{\mathcal{N}}_3\\ \qquad + (\mathcal{P}_1\overline{\mathcal{C}}_2+\mathcal{P}_3\overline{\mathcal{C}}_2)\mathcal{N}_1 + (\mathcal{P}_1\overline{\mathcal{D}}_2+\mathcal{P}_3\overline{\mathcal{D}}_2)\mathcal{N}_2 + (\mathcal{P}_1\overline{\mathcal{E}}_2+\mathcal{P}_3\overline{\mathcal{E}}_2)\mathcal{N}_3,\\ \widehat{\Gamma}_{30} := \mathcal{P}_1\widehat{\mathcal{A}}_3 + \mathcal{P}_1\mathcal{B}_3\mathcal{P}_2 + \mathcal{P}_2(\mathcal{A}_3+\widehat{\mathcal{A}}_3) + \mathcal{P}_2\mathcal{B}_3(\mathcal{P}_1+\mathcal{P}_2) + \mathcal{P}_2\mathcal{C}_3\widehat{\mathcal{N}}_1 + \mathcal{P}_2\mathcal{D}_3\widehat{\mathcal{N}}_2 + \mathcal{P}_2\mathcal{E}_3\widehat{\mathcal{N}}_3,\\ \widetilde{\Gamma}_{30} := \mathcal{P}_1\mathcal{B}_3\mathcal{P}_3 + \mathcal{P}_1\widetilde{\mathcal{B}}_3(\mathcal{P}_1+\mathcal{P}_3) + \mathcal{P}_3(\mathcal{A}_3+\widetilde{\mathcal{A}}_3) + \mathcal{P}_3(\mathcal{B}_3+\widetilde{\mathcal{B}}_3)(\mathcal{P}_1+\mathcal{P}_3)\\ \qquad + (\mathcal{P}_1\widetilde{\mathcal{C}}_3+\mathcal{P}_3\mathcal{C}_3+\mathcal{P}_3\widetilde{\mathcal{C}}_3)\widetilde{\mathcal{N}}_1 + (\mathcal{P}_1\widetilde{\mathcal{D}}_3+\mathcal{P}_3\mathcal{D}_3+\mathcal{P}_3\widetilde{\mathcal{D}}_3)\widetilde{\mathcal{N}}_2 + (\mathcal{P}_1\widetilde{\mathcal{E}}_3+\mathcal{P}_3\mathcal{E}_3+\mathcal{P}_3\widetilde{\mathcal{E}}_3)\widetilde{\mathcal{N}}_3,\\ \overline{\Gamma}_{30} := \mathcal{P}_1\overline{\mathcal{A}}_3 + \mathcal{P}_1\mathcal{B}_3\mathcal{P}_4 + \mathcal{P}_1\widetilde{\mathcal{B}}_3(\mathcal{P}_2+\mathcal{P}_4) + \mathcal{P}_1\overline{\mathcal{B}}_3(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + \mathcal{P}_2(\widetilde{\mathcal{A}}_3+\overline{\mathcal{A}}_3)\\ \qquad + \mathcal{P}_2\mathcal{B}_3(\mathcal{P}_3+\mathcal{P}_4) + \mathcal{P}_2(\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + \mathcal{P}_3(\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3)\\ \qquad + \mathcal{P}_3(\mathcal{B}_3+\widetilde{\mathcal{B}}_3)(\mathcal{P}_2+\mathcal{P}_4) + \mathcal{P}_3\overline{\mathcal{B}}_3(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + \mathcal{P}_4(\mathcal{A}_3+\widetilde{\mathcal{A}}_3+\widehat{\mathcal{A}}_3+\overline{\mathcal{A}}_3)\\ \qquad + \mathcal{P}_4(\mathcal{B}_3+\widetilde{\mathcal{B}}_3+\overline{\mathcal{B}}_3)(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + \mathcal{P}_2\mathcal{C}_3\overline{\overline{\mathcal{N}}}_1 + \mathcal{P}_2\mathcal{D}_3\overline{\overline{\mathcal{N}}}_2 + \mathcal{P}_2\mathcal{E}_3\overline{\overline{\mathcal{N}}}_3\\ \qquad + (\mathcal{P}_1\widetilde{\mathcal{C}}_3+\mathcal{P}_3\mathcal{C}_3+\mathcal{P}_3\widetilde{\mathcal{C}}_3)\overline{\mathcal{N}}_1 + (\mathcal{P}_1\widetilde{\mathcal{D}}_3+\mathcal{P}_3\mathcal{D}_3+\mathcal{P}_3\widetilde{\mathcal{D}}_3)\overline{\mathcal{N}}_2 + (\mathcal{P}_1\widetilde{\mathcal{E}}_3+\mathcal{P}_3\mathcal{E}_3+\mathcal{P}_3\widetilde{\mathcal{E}}_3)\overline{\mathcal{N}}_3\\ \qquad + [\mathcal{P}_1\overline{\mathcal{C}}_3+\mathcal{P}_2\widetilde{\mathcal{C}}_3+\mathcal{P}_2\overline{\mathcal{C}}_3+\mathcal{P}_3\overline{\mathcal{C}}_3+\mathcal{P}_4(\mathcal{C}_3+\widetilde{\mathcal{C}}_3+\overline{\mathcal{C}}_3)]\mathcal{N}_1\\ \qquad + [\mathcal{P}_1\overline{\mathcal{D}}_3+\mathcal{P}_2\widetilde{\mathcal{D}}_3+\mathcal{P}_2\overline{\mathcal{D}}_3+\mathcal{P}_3\overline{\mathcal{D}}_3+\mathcal{P}_4(\mathcal{D}_3+\widetilde{\mathcal{D}}_3+\overline{\mathcal{D}}_3)]\mathcal{N}_2\\ \qquad + [\mathcal{P}_1\overline{\mathcal{E}}_3+\mathcal{P}_2\widetilde{\mathcal{E}}_3+\mathcal{P}_2\overline{\mathcal{E}}_3+\mathcal{P}_3\overline{\mathcal{E}}_3+\mathcal{P}_4(\mathcal{E}_3+\widetilde{\mathcal{E}}_3+\overline{\mathcal{E}}_3)]\mathcal{N}_3.\end{cases}$$

We rewrite (117) as

$$\begin{pmatrix} I_n-\mathcal{P}_1\mathcal{B}_1 & -\mathcal{P}_1\mathcal{B}_2 & -\mathcal{P}_1\mathcal{B}_3\\ -\mathcal{P}_1\mathcal{D}_1 & I_n-\mathcal{P}_1\mathcal{D}_2 & -\mathcal{P}_1\mathcal{D}_3\\ -\mathcal{P}_1\mathcal{E}_1 & -\mathcal{P}_1\mathcal{E}_2 & I_n-\mathcal{P}_1\mathcal{E}_3\end{pmatrix}\begin{pmatrix} Z_1\\ Z_2\\ Z_3\end{pmatrix} = \begin{pmatrix}\Gamma_{10}X(t)+\widehat{\Gamma}_{10}\hat{X}(t)+\widetilde{\Gamma}_{10}\check{X}(t)+\overline{\Gamma}_{i0}\check{\hat{X}}(t)\\ \Gamma_{20}X(t)+\widehat{\Gamma}_{20}\hat{X}(t)+\widetilde{\Gamma}_{20}\check{X}(t)+\overline{\Gamma}_{20}\check{\hat{X}}(t)\\ \Gamma_{30}X(t)+\widehat{\Gamma}_{30}\hat{X}(t)+\widetilde{\Gamma}_{30}\check{X}(t)+\overline{\Gamma}_{30}\check{\hat{X}}(t)\end{pmatrix}. \tag{118}$$

Similarly, if we assume that

(A2.8) the coefficient matrix of (118) is invertible, for any $t \in [0,T]$,

then we have

$$\begin{aligned} Z_i(t) &= (-1)^{i-1}(\mathbb{N}_4)^{-1}\Big[\big(\Gamma_{10}X(t)+\widehat{\Gamma}_{10}\hat{X}(t)+\widetilde{\Gamma}_{10}\check{X}(t)+\overline{\Gamma}_{10}\check{\hat{X}}(t)\big)\overline{\mathbf{M}}_{1i}\\ &\qquad\qquad - \big(\Gamma_{20}X(t)+\widehat{\Gamma}_{20}\hat{X}(t)+\widetilde{\Gamma}_{20}\check{X}(t)+\overline{\Gamma}_{20}\check{\hat{X}}(t)\big)\overline{\mathbf{M}}_{2i}\\ &\qquad\qquad + \big(\Gamma_{30}X(t)+\widehat{\Gamma}_{30}\hat{X}(t)+\widetilde{\Gamma}_{30}\check{X}(t)+\overline{\Gamma}_{30}\check{\hat{X}}(t)\big)\overline{\mathbf{M}}_{3i}(t)\Big]\\ &= (-1)^{i-1}(\mathbb{N}_4)^{-1}\Big\{\big[\Gamma_{10}\overline{\mathbf{M}}_{1i}-\Gamma_{20}\overline{\mathbf{M}}_{2i}+\Gamma_{30}\overline{\mathbf{M}}_{3i}\big]X(t)+\big[\widehat{\Gamma}_{10}\overline{\mathbf{M}}_{1i}-\widehat{\Gamma}_{20}\overline{\mathbf{M}}_{2i}+\widehat{\Gamma}_{30}\overline{\mathbf{M}}_{3i}\big]\hat{X}(t)\\ &\qquad + \big[\widetilde{\Gamma}_{10}\overline{\mathbf{M}}_{1i}-\widetilde{\Gamma}_{20}\overline{\mathbf{M}}_{2i}+\widetilde{\Gamma}_{30}\overline{\mathbf{M}}_{3i}\big]\check{X}(t)+\big[\overline{\Gamma}_{10}\overline{\mathbf{M}}_{1i}-\overline{\Gamma}_{20}\overline{\mathbf{M}}_{2i}+\overline{\Gamma}_{30}\overline{\mathbf{M}}_{3i}\big]\check{\hat{X}}(t)\Big\}\\ &:= \Sigma_i X(t)+\widehat{\Sigma}_i\hat{X}(t)+\widetilde{\Sigma}_i\check{X}(t)+\overline{\Sigma}_i\check{\hat{X}}(t),\ i=1,2,3, \end{aligned} \tag{119}$$

where $\mathbb{N}_4$ is the determinant of the coefficient of (115), and $\overline{\mathbf{M}}_{ji}$ is the adjoint matrix of the (j,i) element in (118), for $j,i = 1,2,3$.

After these four steps, we have obtained that

$$Z_i(t) = \Sigma_i X(t)+\widehat{\Sigma}_i\hat{X}(t)+\widetilde{\Sigma}_i\check{X}(t)+\overline{\Sigma}_i\check{\hat{X}}(t),\ i=1,2,3. \tag{120}$$

Now, comparing the dt term in (104) and substituting (120) into it, we obtain (44). Notice that $\Sigma_i, \widehat{\Sigma}_i, \widetilde{\Sigma}_i, \overline{\Sigma}_i$ in the above depends on $P_i, i = 1, 2, 3, 4$, so the solvability of the above complicated and coupled system of Riccati's type equations is very difficult to obtain. We will not discuss this problem at the present paper for some technical reason and leave it open.

Finally, by (43), (100) and (120), we have

$$
\begin{aligned}
u_2^*(t) =& - N_2^{-1}(t)\Big[\mathcal{L}_4^\top \check{\hat{X}}(t) + \mathcal{C}_{05}^\top \check{Y}(t) + \mathcal{L}_{05}^\top \check{\hat{Y}}(t) + \sum_{i=1}^{3} \mathcal{C}_{i5}^\top \check{Z}_i(t) + \sum_{i=1}^{3} \mathcal{L}_{i5}^\top \check{\hat{Z}}_i(t)\Big] \\
=& - N_2^{-1}(t)\Big\{\Big[\mathcal{C}_{05}^\top(\mathcal{P}_1+\mathcal{P}_3) + \sum_{i=1}^{3}\mathcal{C}_{i5}^\top(\Sigma_i + \widetilde{\Sigma}_i)\Big]\check{X}(t) + \Big[\mathcal{L}_4^\top + \mathcal{C}_{05}^\top(\mathcal{P}_2+\mathcal{P}_4) \\
&+ \sum_{i=1}^{3}\mathcal{C}_{i5}^\top(\widehat{\Sigma}_i + \overline{\Sigma}_i) + \mathcal{L}_{05}^\top(\mathcal{P}_1+\mathcal{P}_2+\mathcal{P}_3+\mathcal{P}_4) + \sum_{i=1}^{3}\mathcal{L}_{i5}^\top(\Sigma_i + \widehat{\Sigma}_i + \widetilde{\Sigma}_i + \overline{\Sigma}_i)\Big]\check{\hat{X}}(t)\Big\}.
\end{aligned}
$$

Thus (49) holds. And the optimal "state" $X = \big(x^*, p\big)^\top$ of the leader admits (48), where $\hat{X}$ is determined by (47), $\check{X}$ is governed by (46), and $\check{\hat{X}}$ is given by (45). The proof is complete. $\square$

References

[1] A. Bagchi, T. Başar, Stackelberg strategies in linear-quadratic stochastic differential games. *J. Optim. Theory Appl.*, **35**(3), 443-464, 1981.

[2] A. Bensoussan, S. K. Chen, and S. P. Sethi, The maximum principle for global solutions of stochastic Stackelberg differential games. *SIAM J. Control Optim.*, **53**(4), 1956-1981, 2015.

[3] P. Cardaliaguet, C. Rainer, Stochastic differential games with asymmetric information. *Appl. Math. Optim.*, **59**(1), 1-36, 2009.

[4] D. J. Chang, H. Xiao, Linear quadratic nonzero sum differential games with asymmetric information. *Math. Prob. Engin.*, **2014**, Article ID 262314, 11 pages.

[5] L. Chen, Y. Shen, On a new paradigm of optimal reinsurance: A stochastic Stackelberg differential game between an insurer and a reinsurer. *ASTIN Bulletin*, **48**(2), 905-960, 2018.

[6] J. Cvitanić, D. Possamai, and N. Touzi, Dynamic programming approach to principal-agent problems. *Finan. & Stoch.* **22**(1), 1-37, 2018.

[7] J. Cvitanić, X. H. Wan, and J. Zhang, Optimal contracts in contituous-time models. *J. Appl. Math. Stoch.*, **2006**, Artical ID 95203, 2006.

[8] J. Cvitanić, J. F. Zhang, *Contract Theory in Continuous-Time Models*, Springer-Verlag, Berlin, 2013.

[9] N. El Karoui, S. G. Peng, and M. Quenez, Backward stochastic differential equations in finance. *Math. Finance*, **7**(1), 1-71, 1997.

[10] X. L. He, A. Prasad, and S. P. Sethi, Cooperative advertising and pricing in a dynamic stochastic supply chain: feedback Stackelberg strategies. *Prod. Oper. Manag.*, **18**(1), 78-94, 2009.

[11] J. H. Huang, G. C. Wang, and J. Xiong, A maximum principle for partial information backward stochastic control problems with applications. *SIAM J. Control Optim.*, **48**(4), 2106-2117, 2009.

[12] J. Lempa, P. Matomäki, A Dynkin game with asymmetric information. *Stochastics: An Inter. J. Proba. Stoch. Proc.*, **85**(5), 763-788, 2013.

[13] N. Li, Z. Y. Yu, Forward-backward stochastic differential equations and linear-quadratic generalized Stackelberg games. *SIAM J. Control Optim.*, **56**(6), 4148-4180, 2018.

[14] T. Li, S. P. Sethi, A review of dynamic Stackelberg game models. *Dis. Cont. Dynam. Syst., Ser. B*, **22**(1), 125-159, 2017.

[15] Y. N. Lin, X. S. Jiang, and W. H. Zhang, An open-loop Stackelberg strategy for the linear quadratic mean-field stochastic differential game. *IEEE Trans. Autom. Control*, **64**(1), 97-110, 2019.

[16] B. Liu, Z. Yin, and C. Lai, A solvable dynamic principal-agent model with linear marginal productivity. *Discrete Dyn. Nat. Soc.*, **2018**, Article ID 5282359, 17 pages.

[17] J. Moon, T. Başar, Linear quadratic mean field Stackelberg differential games. *Automatica*, **97**, 200-213, 2018.

[18] B. Øksendal, A universal optimal consumption rate for an insider. *Math. Finance*, **16**(1), 119-129, 2006.

[19] B. Øksendal, L. Sandal, and J. Ubøe, Stochastic Stackelberg equilibria with applications to time dependent newsvendor models. *J. Econ. Dyna. & Control*, **37**(7), 1284-1299, 2013.

[20] Y. Sannikov, A continuous-time version of the principal-agent problem. *Rev. Econ. Stud.*, **53**, 599-617, 2008.

[21] H. Schattler, J. Sung, The first-order approach to continuous-time principal-agent problem with exponential utility. *J. Econ. Theory*, **61**, 331-371, 1993.

[22] J. T. Shi, G. C. Wang, A new kind of linear-quadratic leader-follower stochastic differential game. In *Proc. 10th IFAC Nonlinear Control System Symposium*, 322-326, August 23-25, Monterey, USA, 2016.

[23] J. T. Shi, G. C. Wang, and J. Xiong, Leader-follower stochastic differential game with asymmetric information and applications. *Automatica*, **63**, 60-73, 2016.

[24] J. T. Shi, G. C. Wang, and J. Xiong, Linear-quadratic stochastic Stackelberg differential game with asymmetric information. *Sci. China Infor. Sci.*, **60**, 092202:1-15, 2017.

[25] M. Simaan, J. B. Cruz Jr., On the Stackelberg game strategy in non-zero games. *J. Optim. Theory Appl.*, **11**(5), 533-555, 1973.

[26] H. von Stackelberg, *The Theory of the Market Economy*, Oxford University Press, London, 1952.

[27] G. C. Wang, H. Xiao, and J. Xiong, A kind of LQ non-zero sum differential game of backward stochastic differential equations with asymmetric information. *Automatica*, **97**, 346-352, 2018.

[28] G. C. Wang, Z. Wu, and J. Xiong, Maximum principles for forward-backward stochastic control systems with correlated state and observation noises. *SIAM J. Control Optim.*, **51**(1), 491-524, 2013.

[29] G. C. Wang, C. H. Zhang, and W. H. Zhang, Stochastic maximum principle for mean-field type optimal control under partial information. *IEEE Trans. Autom. Control*, **59**(2), 522-528, 2014.

[30] N. Williams, On dynamic principal-agent problems in continuous time. *Working paper*, University of Wosconsin-Madison, 2009.

[31] N. Williams, A solvable continuous time principal agent model. *J. Econ. Theory*, **159**, 989-1015, 2015.

[32] J. Xiong, *An Introduction to Stochastic Filtering Theory*, Oxford University Press, London, 2008.

[33] J. J. Xu, H. S. Zhang, Sufficient and necessary open-loop Stackelberg strategy for two-player game with time delay. *IEEE Trans. Cyber.*, **46**(2), 438-449, 2016.

[34] J. J. Xu, J. T. Shi, and H. S. Zhang, A leader-follower stochastic linear quadratic differential game with time delay. *Sci. China Infor. Sci.*, **61**, 112202:1-13, 2018.

[35] J. M. Yong, A leader-follower stochastic linear quadratic differential games. *SIAM J. Control Optim.*, **41**(4), 1015-1041, 2002.

[36] J. M. Yong, X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.