Linear-Quadratic-Gaussian Mean-Field-Game with Partial Observation and Common Noise

Alain Bensoussan* Xinwei Feng† Jianhui Huang‡

November 14, 2019

Keywords: Mean-field games; ϵ-Nash equilibrium; Forward-backward stochastic differential equation; Kalman filtering; Partial observation; Consistency condition.

Abstract

This paper considers a class of linear-quadratic-Gaussian (LQG) mean-field games (M-FGs) with partial observation structure for individual agents. Unlike other literature, there are some special features in our formulation. First, the individual state is driven by some common-noise due to the external factor and the state-average thus becomes a random process instead of a deterministic quantity. Second, the sensor function of individual observation depends on state-average thus the agents are coupled in triple manner: not only in their states and cost functionals, but also through their observation mechanism. The decentralized strategies for individual agents are derived by the Kalman filtering and separation principle. The consistency condition is obtained which is equivalent to the wellposedness of some forward-backward stochastic differential equation (FBSDE) driven by common noise. Finally, the related ϵ-Nash equilibrium property is verified.

1 Introduction

The starting point of our work is the recently well-studied mean-field games (MFGs) for large-population system (sometimes, it is also termed multi-agent system (MAS)). The large-population system arises naturally in various fields such as economics, engineering, social science and operational research, etc. For example, dynamic economic models involving competing agents ([11], [24], [35]); wireless power control, shared data buffer modeling and traffic engineering ([13], [18], [23], [27]); synchronization of coupled nonlinear oscillators ([37]); crowd and consensus dynamics ([8], [29]), etc. The most significant feature of large-population system is the existence of a large number of individually negligible agents (or players) which are interrelated in their dynamics and (or) cost functionals via the state-average or more generally, the generated empirical measure over the whole population. Due to this highly complicated coupling feature, it is intractable for a given agent to study the centralized optimization strategies based on the information of all

*International Center for Decision and Risk Analysis Jindal School of Management, the University of Texas at Dallas and School of Data Sciences, City University of Hong Kong. Email address: axb046100@utdallas.edu

†Corresponding author. Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan, Shandong 250100, China. Email address: xwfeng@sdu.edu.cn

‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong. Email address: majhuang@polyu.edu.hk

its peers in large-population system. In fact, this will bring considerably high computational complexity in large-scale. Alternatively, one reasonable and practical direction is to investigate the related decentralized strategies based on local information only. By local information, we mean the related strategies should be designed upon the individual state of given agent only together with some mass-effect quantities but can be computed in off-line manner.

Along this research direction, one efficient and tractable methodology to decentralized strategies is the MFGs which generally leads to a coupled system of HJB equation and Fokker-Planck (FP) equation in nonlinear case. In principle, the procedure of MFGs consists of the following four main steps (see [3], [5], [6], [19], [20], [25], etc): in Step 1, it is necessary to analyze the asymptotic behavior of state-average when the agent number N tends to infinity and introduce the related state-average limiting term. Of course, this limiting term is undetermined at this moment thus it should be treated as some exogenous "frozen" term; Step 2 turns to study the related limiting optimization problem (which is also called auxiliary or tracking problem) by replacing the state-average by its frozen limit term. The initial highly-coupled optimization problems of all agents are thus decoupled and only parameterized by this generic frozen limit. The related decentralized optimal strategy can be analyzed using standard control techniques such as dynamic programming principle (DPP) or stochastic maximum principle (SMP) (see e.g., [38]). As a result, some HJB equation due to DPP or Hamiltonian system due to SMP will be obtained to characterize this decentralized optimality; Step 3 aims to determine the frozen state-average limit by some consistency condition: when applying the optimal decentralized strategies derived in Step 2, the state-average limit should be reproduced as the agents number tends to infinity. Accordingly, some fixed-point analysis should be applied here and some FP equation will be introduced and coupled with the HJB equation in Step 2. As the necessary verification, Step 4 will show that the derived decentralized strategies should possess the ϵ-Nash equilibrium properties.

For further analysis details of MFGs, the interested readers are referred to [12] for a survey of mean-field games focusing on the partial differential equation aspect and related real applications; [3] for more recent MFG studies and the related mean-field type control; [5] for the probabilistic analysis of a large class of stochastic differential games for which the interaction between the players is of mean-field type; [7] for the mean-field game where considerable interrelated banks share the system risk and common noise; [32] for a class of risk-sensitive mean-field stochastic differential games; [21] for MFGs with nonlinear diffusion dynamics and their relations to McKean-Vlasov particle system; [16] for the dynamic optimization of large-population system with partial information and the associated MFG; [31] for nonlinear filtering theory for partially observed stochastic dynamical systems of McKean–Vlasov type stochastic differential equations. It is remarkable that there exists a substantial literature body to the study of MFGs in linear-quadratic-Gaussian (LQG) framework. Here, we mention a few of them which are more relevant to our current work: [4] for the linear-quadratic mean field games via the stochastic maximum principle and adjoint equation, [1] for the N-person linear differential mean-field games with explicit solution, [17] for the mean-field LQG games with a major player and a large number of minor players, [20] for the mean-field LQG games with nonuniform agents through the state-aggregation by empirical distribution, [28] for the mean-field LQG mixed games with continuum-parameterized minor players; [15] for linear-quadratic-Gaussian MFGs having a major agent and numerous heterogeneous minor agents in the presence of mean-field interactions.

In this paper, we discuss the mean-field games in the framework of partial observation. Specially, we consider a large population system wherein all agents are coupled in their state evolutions and cost functionals. However, due to the realistic factors such as finite datum, latent process or imperfect information, each agent can only access some noisy observation on his own state. Based on this partial observation, each agent aims to analyze the decentralized strategy with the help of Kalman filtering and separation principle but in large-population setting. On the other hand, unlike most existing MFG literature, we assume the states of all agents are governed by some underlying common-noise. This common noise can be interpreted as some exogenous and generic factors such as the macro-economic scenario, tax policy, interest rate or exchange rate. It follows these factors should influence all participants in a given large-population economy. In fact, the effect of such common noise becomes more significant when we consider a given industry sector with considerable small firms. Actually, the dynamic behaviors of all these firms should be regularized by the same external competition mechanism. For example, suppose all these firms produce the same type products hence their individual production plans will depend on the quoted price of same raw materials, or the same underlying tax regulation applied. The presence of common noise makes the state-average limit in MFG analysis become some stochastic process instead of deterministic quantity.

In our work, the random state-average limit enters both the auxiliary state and observation dynamics (refer Eq. (5)-(6) below). As a result, there arise some measurability and adaptness issues (e.g., to verify the filtration generated by uncontrolled observation process equals that of the controlled observation process) when constructing the admissible control set and analyzing the related state-observation separation principle (see [2], [10], etc.). Such issues make our analysis different from the MFG with partial information discussed in [19] where no common noise added. Thus, their state-average limit is still deterministic and the standard separation principle via Kalman filtering technique can be applied directly therein without additional adaptness issues. As a solution, we give a modified separation to state and observation by taking into account random state-average limit (but without any assumption to its Gaussian-Markov property) and then verify the related observation filtration equivalence. Based on it, we can get some separation principle and derive the related decentralized control strategies. Moreover, the consistency condition will be established by the resulting decentralized strategies through some fixed-point analysis. Here, we connect the consistency condition to the well-posedness of some forward-backward stochastic differential equation (FBSDE). Moreover, we present some decoupling results of this FBSDE via some asymmetric Riccati equation system.

As a response to above discussions, this paper investigates a class of LQG MFGs with partial observation and common noise. The reminder of this paper is structured as follows: Section 2 gives the problem formulation. The decentralized strategies are derived by Kalman filtering method and the consistency condition is also established through some FBSDE system. Section 3 verifies the ϵ-Nash equilibrium of the decentralized strategies. Section 4 gives some numerical computations to illustrate the theoretical results. Section 5 concludes our work and presents some future research directions.

2 LQG MFGs with Partial Observation

Consider a finite horizon $[0,T]$ for fixed $T>0$. $(\Omega,\mathcal{F},P)$ is a complete probability space on which a standard $(d+m\times N)$-dimensional Brownian motion $\{W(t),W_i(t),\ 1\leq i\leq N\}$ is defined. Here,

d denotes the dimension of Brownian motion of common noise, m the dimension of Brownian motion of individual noise, and N is the number of agents in large population. $\mathbb{R}^n$ ($\mathbb{R}^{n\times k}$) denotes the n ($n \times k$)-dimensional Euclidean space with its norm denoted by $|\cdot|$. We denote the set of symmetric $n\times n$ matrices with real elements by S^n. Here, n, k denote the dimensions of state and control variable respectively. If $M \in S^n$ is positive (semi)definite, we write $M > \ (\geq)\ 0$. For given filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$, let $L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ ($L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n\times k})$) denote the space of all $\mathcal{F}_t$-progressively measurable processes with values in $\mathbb{R}^n$ ($\mathbb{R}^{n\times k}$) satisfying $\mathbb{E}\int_0^T |x(t)|^2 dt < +\infty$; $L^2(0,T;\mathbb{R}^n)$ ($L^2(0,T;\mathbb{R}^{n\times k})$) the space of all deterministic functions with values in $\mathbb{R}^n$ ($\mathbb{R}^{n\times k}$) satisfying $\int_0^T |x(t)|^2 dt < +\infty$; $L^\infty(0,T;\mathbb{R}^n)$ ($L^\infty(0,T;\mathbb{R}^{n\times k})$) the space of uniformly bounded functions with values in $\mathbb{R}^n$ ($\mathbb{R}^{n\times k}$); $C([0,T];\mathbb{R}^n)$ ($C([0,T];\mathbb{R}^{n\times k})$) the space of continuous functions with values in $\mathbb{R}^n$ ($\mathbb{R}^{n\times k}$). If $M(\cdot) \in L^\infty(0,T;S^n)$ and $M(t) > \ (\geq)\ 0$ for every $t \in [0,T]$, $M(\cdot)$ is positive (semi)definite, and denoted by $M(\cdot) > \ (\geq)\ 0$. For a given vector or matrix M, M' stands for its transpose.

We consider a large-population system with N individual agents $\{\mathcal{A}_i\}_{1\leq i\leq N}$. The state x_i for i^{th} agent $\mathcal{A}_i$ satisfies the following linear stochastic system:

$$\begin{cases} dx_i(t) = [A_{\theta_i}(t)x_i(t) + B(t)u_i(t) + a_{\theta_i}(t)x^{(N)}(t) + m(t)]dt + \sigma(t)dW_i(t) + \widetilde{\sigma}(t)dW(t), \\ x_i(0) = x, \end{cases} \tag{1}$$

with $x^{(N)}(\cdot) \triangleq \frac{1}{N}\sum_{i=1}^N x_i(\cdot)$ denoting the state-average of population. Here, W_i is the individual noise while W is the common noise due to underlying common factors; A_{θ_i}, B denote the drift parameters of state and control; a_{θ_i} is the state-coupling parameter; $\sigma, \widetilde{\sigma}$ denote the diffusion coefficients. Similar setup of common noise can be found in [7]. $\mathcal{A}_i$ can access the following additive white-noise partial observation:

$$\begin{cases} dY_i(t) = [H(t)x_i(t) + \widetilde{H}_{\theta_i}(t)x^{(N)}(t) + h(t)]dt + dV_i(t), \\ Y_i(0) = 0, \end{cases} \tag{2}$$

where $\{V_i\}_{1\leq i\leq N}$ stand for l-dimensional Brownian motions. Here, $\widetilde{H}_{\theta_i}$ is introduced in sensor function of (2) to characterize the coupling effects due to interactions of agents in large population system. If $\widetilde{H} = 0$, Equation (2) becomes the additive white-noise observation which is commonly seen in (linear) filtering literature (e.g., [2], [22], [30]). Define the observable filtration $\mathcal{F}^i = \{\mathcal{F}^i_t\}_{0\leq t\leq T}$ of $\mathcal{A}_i$ with $\mathcal{F}^i_t \triangleq \sigma\{Y_i(s), W(s); 0 \leq s \leq t\}$ and the filtration of common noise $\mathcal{F}^w = \{\mathcal{F}^w_t\}_{0\leq t\leq T}$ with $\mathcal{F}^w_t \triangleq \sigma\{W(s); 0 \leq s \leq t\}$.

In (1), (2), θ_i is a dynamic parameter for agent $\mathcal{A}_i$ in the heterogeneous population. For sake of brief notations, we only set the coefficients $(A, a, \widetilde{H})$ to be dependent on θ_i. In case other coefficients for $\mathcal{A}_i$ also depend on θ_i, the analysis is similar and we will not present its full details here. For θ_i, we assume it takes values from a finite set $\Theta = \{1, 2, \cdots, K\}$, i.e., there are K different types of heterogeneous agents (see [17] for similar setup). For example, if $\theta_i = k$, then $\mathcal{A}_i$ is called a $k-$type agent. In this paper, we are interested in the asymptotic behavior as N tends to infinity. For $1 \leq k \leq K$, introduce

$$\mathcal{I}_k = \{i|\theta_i = k, 1 \leq i \leq N\}, \qquad N_k = |\mathcal{I}_k|,$$

where N_k is the cardinality of index set $\mathcal{I}_k$. For $1 \leq k \leq K$, let $\chi^{(N)}_k = \frac{N_k}{N}$, then $\chi^{(N)} = (\chi^{(N)}_1, \cdots, \chi^{(N)}_K)$ is a probability vector representing the empirical distribution of $\theta_1, \cdots, \theta_N$. We introduce the following assumption:

(A1) There exists a probability mass vector $\chi = (\chi_1, \cdots, \chi_K)$ such that $\lim_{N \longrightarrow +\infty} \chi^{(N)} = \chi$ and $\min_{1 \leq k \leq K} \chi_k > 0$.

The implication of (A1) is that if the population size $N \longrightarrow +\infty$, the proportion of $k-$type agents becomes stable for each k and the number of each type agents tends to infinity. Otherwise, the agents in given type with bounded size should be excluded from consideration when analyzing asymptotic behavior as $N \longrightarrow +\infty$.

Remark 2.1 *Hereafter, the time variable t will often be suppressed to simplify the notations and presentations.*

For $1 \leq i \leq N$, the admissible control set $\mathcal{U}_i$ of agent i is defined as

$$\mathcal{U}_i := \{u_i(\cdot) | u_i(\cdot) \in L^2_{\mathcal{F}^i}(0, T; \mathbb{R}^k)\}.$$

Let $u = (u_1, \cdots, u_i, \cdots, u_N)$ denote the strategy set of all N agents; $u_{-i} = (u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_N)$ the strategy set except $\mathcal{A}_i$. The cost functional of $\mathcal{A}_i$ is assumned to be:

$$\mathcal{J}_i(u_i(\cdot), u_{-i}(\cdot)) = \mathbb{E}\Big[\int_0^T \Big((x_i - x^{(N)})' Q (x_i - x^{(N)}) + u_i' R u_i\Big)\, dt + x_i'(T) G x_i(T)\Big]. \tag{3}$$

Here, Q, R are state and control weight matrix in running cost, while G the terminal weight of state. We set the following assumptions on the coefficients:

(A2) $\{A_k\}_{k=1}^K \in L^\infty(0, T; \mathbb{R}^{n\times n}), B \in L^\infty(0, T; \mathbb{R}^{n\times k}), \{a_k\}_{k=1}^K \in L^\infty(0, T; \mathbb{R}^{n\times n}),$ $m \in L^2(0, T; \mathbb{R}^n), \sigma \in L^2(0, T; \mathbb{R}^{n\times m}), \widetilde{\sigma} \in L^2(0, T; \mathbb{R}^{n\times d})$;

(A3) $H, \{\widetilde{H}_k\}_{k=1}^K \in L^\infty(0, T; \mathbb{R}^{l\times n}), h \in L^2(0, T; \mathbb{R}^l)$;

(A4) $Q \in L^\infty(0, T; S^n), Q \geq 0, R(\cdot) \in L^\infty(0, T; S^k), R \geq \delta I$, for some $\delta > 0$, $G \in S^n, G \geq 0$.

Under (A2), for any $u_i \in \mathcal{U}_i$, the state equation (1) admits a unique strong solution (e.g., [38]). Under (A4), the cost functional (3) is well-defined.

Now, we formulate the problem to find a Nash equilibrium of mean-field game with partial observation (**PO**).

Problem (PO). Find the strategies set $\bar{u} = (\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_N)$ such that for $i = 1, 2, \cdots, N$,

$$\mathcal{J}_i(\bar{u}_i(\cdot), \bar{u}_{-i}(\cdot)) = \inf_{u_i(\cdot) \in \mathcal{U}_i} \mathcal{J}_i(u_i(\cdot), \bar{u}_{-i}(\cdot)).$$

To study (**PO**), one efficient methodology is the mean-field LQG games which relates the "centralized" LQG problems via the limiting state-average, as the agent number tends to infinity. Define the state-average of all agents

$$x^{(N)} \triangleq \frac{1}{N}\sum_{i=1}^N x_i = \frac{1}{N}\sum_{k=1}^K \sum_{i\in\mathcal{I}_k} x_i = \sum_{k=1}^K \chi_k^{(N)} x_k^{(N)}, \tag{4}$$

where $x_k^{(N)} \triangleq \frac{1}{N_k}\sum_{i\in\mathcal{I}_k} x_i$ denotes the state-average of all $k-$type agents.

As explained in the introduction, the centralized strategies for **Problem (PO)** are rather complicate and infeasible to be applied when the number of the agents tends to infinity. Alternatively, we investigate the decentralized strategies via the limiting problem with the help of the frozen limiting state-average. To this end, first we figure out the representation of the limiting process by the heuristic arguments. Based on this, we can find the decentralized strategies by the consistency condition and verify the asymptotic Nash equilibrium of the derived decentralized strategies. Since $\lim_{N\longrightarrow\infty}\chi^{(N)}=\chi$, by (4), we may approximate $x^{(N)}$, $\{x_k^{(N)}\}_{k=1}^N$ by x^0, $\{x_k^0\}_{k=1}^K$, respectively, where x^0, $\{x_k^0\}_{k=1}^K$ should have the following relation

$$x^0=\sum_{k=1}^K\chi_k x_k^0.$$

1

Define the state filter for $\mathcal{F}_t^i$ as

$$\hat{x}_i(t)\triangleq\mathbb{E}[x_i(t)|\mathcal{F}_t^i].$$

Then $\hat{x}^{(N)}(\cdot)\triangleq\frac{1}{N}\sum_{i=1}^N\hat{x}_i(\cdot)$ denotes the average of state filters. Similarly, $\hat{x}^{(N)}(\cdot)$ can be approximated by $\hat{x}^0(\cdot)=\sum_{k=1}^K\chi_k\hat{x}_k^0(\cdot)$ where $\hat{x}_k^0(\cdot)\triangleq\lim_{N\longrightarrow+\infty}\frac{1}{N_k}\sum_{i\in\mathcal{I}_k}\hat{x}_i(\cdot)$. Moreover, due to the common noise, $x^0,x_k^0,\hat{x}^0,\hat{x}_k^0$ should be adapted to filtration $\{\mathcal{F}_t^w\}$ and this can be verified in our late analysis. Now, we introduce the limiting state dynamics

$$\begin{cases}dy_i=[A_{\theta_i}y_i+Bu_i+a_{\theta_i}x^0+m]dt+\sigma dW_i+\widetilde{\sigma}dW,\\ y_i(0)=x,\end{cases}\tag{5}$$

and limiting observation process

$$\begin{cases}d\bar{Y}_i=[Hy_i+\widetilde{H}_{\theta_i}x^0+h]dt+dV_i,\\ \bar{Y}_i(0)=0.\end{cases}\tag{6}$$

The limiting cost functional is given by

$$J_i(u_i(\cdot))=\ \mathbb{E}\Big[\int_0^T\Big((y_i-x^0)'Q(y_i-x^0)+u_i'Ru_i\Big)\,dt+y_i'(T)Gy_i(T)\Big].\tag{7}$$

Note that (5)-(7) are limiting versions of (1)-(3) when the mean field term, $x^{(N)}$, is replaced by x^0, which will be determined later in the paper. Before formulating the limiting LQG MFG, we should first analyze the control-observation information structure as the observation process depends on the admissible control applied, and vice versa, the admissible control should be adapted to observation process. To this end, we will use the separation method which is originally obtained by Wonham [36] and is systematically introduced in the book Bensoussan [2]. See also Wang and Wu [33], Wang Wu and Xiong [34] for the backward separation method. Introduce the processes $\alpha_i(\cdot),\beta_i(\cdot)$ by

$$\begin{cases}d\alpha_i=[A_{\theta_i}\alpha_i+m]dt+\sigma dW_i,\\ \alpha_i(0)=x,\end{cases}\tag{8}$$

and

$$\begin{cases} d\beta_i = [H\alpha_i + h]dt + dV_i, \\ \beta_i(0) = 0. \end{cases} \tag{9}$$

Note that the processes $\alpha_i(\cdot), \beta_i(\cdot)$ correspond to the state and observation processes when there is neither control nor x^0 (more precisely the control and x^0 are 0). Further introduce

$$\begin{cases} dx_i^1 = [A_{\theta_i} x_i^1 + Bu_i + a_{\theta_i} x^0]dt + \widetilde{\sigma} dW, \\ x_i^1(0) = 0, \end{cases} \tag{10}$$

and

$$\begin{cases} dz_i^1 = [Hx_i^1 + \widetilde{H}_{\theta_i} x^0]dt, \\ z_i^1(0) = 0. \end{cases} \tag{11}$$

It follows that for any control $u_i(\cdot)$,

$$y_i(t) = \alpha_i(t) + x_i^1(t), \quad \bar{Y}_i(t) = \beta_i(t) + z_i^1(t). \tag{12}$$

Define $\mathcal{F}_{u,t}^{\bar{Y}_i,W} \triangleq \sigma\{\bar{Y}_i(s), W(s); 0 \leq s \leq t\}$, $\mathcal{F}_t^{\beta_i,W} \triangleq \sigma\{\beta_i(s), W(s); 0 \leq s \leq t\}$, $\mathcal{F}_t^{\beta_i} \triangleq \sigma\{\beta_i(s); 0 \leq s \leq t\}$. Here, the subscript u in $\mathcal{F}_{u,t}^{\bar{Y}_i,W}$ emphasizes its dependence on control. We define the following (restricted) admissible control set $\bar{\mathcal{U}}_i$ for limiting partial observation:

$$\bar{\mathcal{U}}_i := \Big\{ u_i(\cdot) | u_i(\cdot) \in L^2(0,T;\mathbb{R}^k),\ u_i(\cdot) \text{ is adapted to } \mathcal{F}_{u,t}^{\bar{Y}_i,W} \text{ and } \mathcal{F}_t^{\beta_i,W} \Big\},\ 1 \leq i \leq N.$$

Now formulate the following limiting partial observation (**LPO**) LQG game.

Problem (LPO). For the i^{th} agent, $i = 1, 2, \cdots, N$, find $\bar{u}_i(\cdot) \in \bar{\mathcal{U}}_i$ satisfying

$$J_i(\bar{u}_i(\cdot)) = \inf_{u_i(\cdot) \in \bar{\mathcal{U}}_i} J_i(u_i(\cdot)).$$

Then $\bar{u}_i(\cdot)$ is called an optimal control for Problem (**LPO**).

Remark 2.2 *In [19], a class of LQG MFG with noisy observation is also discussed but without the introduction of common noise. Thus, the limit state-average in [19] is deterministic and it does not bring any measurability issue to the observation-control analysis.*

With the definition of admissibility, it is immediate from (12) that

$$\text{if } u_i(\cdot) \text{ is admissible, then } \mathcal{F}_{u,t}^{\bar{Y}_i,W} = \mathcal{F}_t^{\beta_i,W}.$$

Thus we have

$$\hat{y}_i(t) = \mathbb{E}\Big[y_i(t) | \mathcal{F}_{u,t}^{\bar{Y}_i,W} \Big] = \mathbb{E}\Big[y_i(t) | \mathcal{F}_t^{\beta_i,W} \Big].$$

Noting that $W(\cdot)$ is independent of $W_i(\cdot), V_i(\cdot)$, we get $W(\cdot)$ is independent of $\alpha_i(\cdot), \beta_i(\cdot)$. Then it follows $\mathbb{E}\big(\alpha_i(t) | \mathcal{F}_t^{\beta_i,W}\big) = \mathbb{E}\big(\alpha_i(t) | \mathcal{F}_t^{\beta_i}\big) = \hat{\alpha}_i(t)$, where $\hat{\alpha}_i$ satisfies the Kalman filtering equation (e.g. [2], Section 1.2)

$$\begin{cases} d\hat{\alpha}_i = \big[A_{\theta_i} \hat{\alpha}_i + m \big] dt + P_{\theta_i} H' \big[d\beta_i - (H\hat{\alpha}_i + h)dt \big], \\ \hat{\alpha}_i(0) = x, \end{cases} \tag{13}$$

and P_{θ_i} is the unique solution of the Riccati equation

$$\begin{cases} \dot{P}_{\theta_i} = A_{\theta_i} P_{\theta_i} + P_{\theta_i} A'_{\theta_i} - P_{\theta_i} H' H P_{\theta_i} + \sigma\sigma', \\ P_{\theta_i}(0) = 0. \end{cases} \tag{14}$$

Noting $x_i^1(\cdot) \in \mathcal{F}_t^{\beta_i, W}$, we have $\hat{y}_i = \hat{\alpha}_i + x_i^1$. Besides,

$$d\beta_i - (H\hat{\alpha}_i + h)dt = d\bar{Y}_i - (H\hat{y}_i + \widetilde{H}_{\theta_i} x^0 + h)dt.$$

Therefore,

$$\begin{cases} d\hat{y}_i = \big[A_{\theta_i}\hat{y}_i + Bu_i + a_{\theta_i} x^0 + m\big]dt + P_{\theta_i} H'\big[d\bar{Y}_i - (H\hat{y}_i + \widetilde{H}_{\theta_i} x^0 + h)dt\big] + \widetilde{\sigma} dW, \\ \hat{y}_i(0) = x. \end{cases} \tag{15}$$

Introduce the innovation process

$$I_i(t) = \beta_i(t) - \int_0^t [H(s)\hat{\alpha}_i(s) + h(s)]ds,$$

which is adapted to $\mathcal{F}_t^{\beta_i, W}$. Let $\Lambda_{\theta_i} \in L^\infty(0, T; \mathbb{R}^{n\times n})$, $\lambda_{\theta_i} \in L^2_{\mathcal{F}^w}(0, T; \mathbb{R}^n)$ be the parameters of a feedback $\Lambda_{\theta_i} x_i + \lambda_{\theta_i}$. Consider

$$\begin{cases} d\eta_i = [(A_{\theta_i} + B\Lambda_{\theta_i})\eta_i + a_{\theta_i} x^0 + m + B\lambda_{\theta_i}]dt + P_{\theta_i} H' dI_i + \widetilde{\sigma} dW, \\ \eta_i(0) = x. \end{cases} \tag{16}$$

It is clearly that $\eta_i(\cdot) \in \mathcal{F}_t^{\beta_i, W}$. Define $u_i(t) = \Lambda_{\theta_i}(t)\eta_i(t) + \lambda_{\theta_i}(t)$, then $u_i(\cdot)$ is square integrable and adapted to $\mathcal{F}_t^{\beta_i, W}$. Further we have

$$dI_i = d\beta_i - (H\hat{\alpha}_i + h)dt = d\bar{Y}_i - (H\eta_i + \widetilde{H}_{\theta_i} x^0 + h)dt.$$

Plugging this into (16), we have

$$d\eta_i = [(A_{\theta_i} + B\Lambda_{\theta_i} - P_{\theta_i} H' H)\eta_i + a_{\theta_i} x^0 + m + B\lambda_{\theta_i} - P_{\theta_i} H'(\widetilde{H}_{\theta_i} x^0 + h)]dt + P_{\theta_i} H' d\bar{Y}_i + \widetilde{\sigma} dW.$$

Therefore,

$$\begin{aligned} \eta_i(t) =& \Phi(t)x + \Phi(t)\int_0^t \Phi^{-1}(s)\Big[a_{\theta_i} x^0 + m + B\lambda_{\theta_i} - P_{\theta_i} H'(\widetilde{H}_{\theta_i} x^0 + h)\Big]ds \\ &+ \Phi(t)\int_0^t \Phi^{-1}(s) P_{\theta_i} H' d\bar{Y}_i + \Phi(t)\int_0^t \Phi^{-1}(s)\widetilde{\sigma} dW, \end{aligned}$$

where

$$\begin{cases} d\Phi(t) = (A_{\theta_i} + B\Lambda_{\theta_i} - P_{\theta_i} H' H)\Phi(t)dt, \\ \Phi(0) = I. \end{cases}$$

Then $\eta_i(\cdot)$, and consequently $u_i(\cdot)$, are adapted to $\mathcal{F}_{u,t}^{\bar{Y}_i, W}$. It follows that $u_i(\cdot)$ is an admissible control. Naturally $\eta_i(\cdot)$ is the corresponding Kalman filter, and $u_i(t) = \Lambda_{\theta_i}(t)\eta_i(t) + \lambda_{\theta_i}(t)$ is a feedback on it.

Introduce the following two equations of π_{θ_i} and γ_{θ_i} respectively:

$$\begin{cases} \dot{\pi}_{\theta_i} + \pi_{\theta_i} A_{\theta_i} + A'_{\theta_i}\pi_{\theta_i} - \pi_{\theta_i} BR^{-1}B'\pi_{\theta_i} + Q = 0, \\ \pi_{\theta_i}(T) = G, \end{cases} \tag{17}$$

and

$$\begin{cases} d\gamma_{\theta_i} + \big[(A'_{\theta_i} - \pi_{\theta_i} BR^{-1}B')\gamma_{\theta_i} + \pi_{\theta_i}(a_{\theta_i}x^0 + m) - Qx^0\big]dt + \xi_{\theta_i} dW = 0, \\ \gamma_{\theta_i}(T) = 0. \end{cases} \tag{18}$$

Under (A2)-(A4), (14), (17) are standard Riccati equations which admit a unique solution $P_{\theta_i}, \pi_{\theta_i} \in C([0,T];\mathbb{R}^{n\times n})$. Moreover, under (A2)-(A4), the linear backward stochastic differential equation (LBSDE) (18) admits a unique adaptive solution pair $(\gamma_{\theta_i}, \xi_{\theta_i}) \in L^2_{\mathcal{F}^w_t}(0,T;\mathbb{R}^n) \times L^2_{\mathcal{F}^w_t}(0,T;\mathbb{R}^{n\times d})$. Note that $\xi_{\theta_i}(\cdot)$ is introduced in solution pair to ensure $\gamma_{\theta_i}(\cdot)$ to be adapted to $\mathcal{F}^w_t$. Now we present the following result.

Lemma 2.1 *Let* (A2)-(A4) *hold and* $P_{\theta_i}, \pi_{\theta_i} \in C([0,T];\mathbb{R}^{n\times n})$ *are solution of* (14), (17) *respectively,* $(\gamma_{\theta_i}, \xi_{\theta_i}) \in L^2_{\mathcal{F}^w_t}(0,T;\mathbb{R}^n) \times L^2_{\mathcal{F}^w_t}(0,T;\mathbb{R}^{n\times d})$ *is the solution pair of* (18). *Then the optimal control of* (**LPO**) *is*

$$\bar{u}_i(t) = -R^{-1}(t)B'(t)\pi_{\theta_i}(t)\hat{y}_i(t) - R^{-1}(t)B'(t)\gamma_{\theta_i}(t), \tag{19}$$

where $\hat{y}_i(t)$ *satisfies the following filtering equation*

$$\begin{cases} d\hat{y}_i = \big[A_{\theta_i}\hat{y}_i - BR^{-1}B'(\pi_{\theta_i}\hat{y}_i + \gamma_{\theta_i}) + a_{\theta_i}x^0 + m\big]dt + P_{\theta_i}H'\big[d\bar{Y}_i - (H\hat{y}_i + \widetilde{H}_{\theta_i}x^0 + h)dt\big] \\ \qquad\quad + \widetilde{\sigma}dW, \\ \hat{y}_i(0) = x. \end{cases} \tag{20}$$

Proof. Suppose the optimal control $\bar{u}_i(\cdot)$ can be written by a linear feedback: $\bar{u}_i = \Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i}$ for $\Lambda_{\theta_i}, \lambda_{\theta_i}$ to be determined (this can be verified in our later analysis). Here, $\hat{y}_i(\cdot)$ is the Kalman filter corresponding to $\bar{u}_i$, and $y_i(\cdot), \bar{Y}_i(\cdot)$ are the corresponding state and observation to $\bar{u}_i$ respectively. Then the following relations hold:

$$\begin{cases} d\hat{y}_i = \big[(A_{\theta_i} + B\Lambda_{\theta_i})\hat{y}_i + a_{\theta_i}x^0 + m + B\lambda_{\theta_i}\big]dt + P_{\theta_i}H'\big[d\bar{Y}_i - (H\hat{y}_i + \widetilde{H}_{\theta_i}x^0 + h)dt\big] + \widetilde{\sigma}dW, \\ \hat{y}_i(0) = x, \\ \bar{u}_i = \Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i}, \\ dy_i = [A_{\theta_i}y_i + B\bar{u}_i + a_{\theta_i}x^0 + m]dt + \sigma dW_i + \widetilde{\sigma}dW, \\ y_i(0) = x, \\ d\bar{Y}_i = (Hy_i + \widetilde{H}_{\theta_i}x^0 + h)dt + dV_i, \ \bar{Y}_i(0) = 0. \end{cases} \tag{21}$$

Let $\mu(\cdot)$ be adapted to $\mathcal{F}^{\beta_i,W}_t$ and $\mathcal{F}^{\bar{Y}_i,W}_{u,t}$. Consider the state $y^\mu_i(\cdot)$ and the observation $\bar{Y}^\mu_i(\cdot)$ corresponding to $u_i(\cdot)$, where $u_i(t) = \Lambda_{\theta_i}(t)\hat{y}^\mu_i(t) + \lambda_{\theta_i}(t) + \mu(t) \in \mathcal{F}^{\beta_i,W}_t$ and $\mathcal{F}^{\bar{Y}_i,W}_{u,t}$, $\hat{y}^\mu_i(t)$ is the related Kalman filter. Then we can write for any $\mu(\cdot) \in \mathcal{F}_t^{\beta_i,W}$ and $\mathcal{F}_{u,t}^{\bar{Y}_i,W}$

$$\begin{cases} d\hat{y}_i^\mu = \big[(A_{\theta_i} + B\Lambda_{\theta_i})\hat{y}_i^\mu + a_{\theta_i}x^0 + m + B\lambda_{\theta_i} + B\mu\big]dt + P_{\theta_i}H'\big[d\bar{Y}_i^\mu - (H\hat{y}_i^\mu + \widetilde{H}_{\theta_i}x^0 + h)dt\big] \\ \qquad\quad + \widetilde{\sigma}dW, \\ \hat{y}_i^\mu(0) = x, \\ u_i = \Lambda_{\theta_i}\hat{y}_i^\mu + \lambda_{\theta_i} + \mu, \\ dy_i^\mu = [A_{\theta_i}y_i^\mu + Bu_i + a_{\theta_i}x^0 + m]dt + \sigma dW_i + \widetilde{\sigma}dW, \\ y_i^\mu(0) = x, \\ d\bar{Y}_i^\mu = (Hy_i^\mu + \widetilde{H}_{\theta_i}x^0 + h)dt + dV_i(t),\ \bar{Y}_i^\mu(0) = 0. \end{cases} \tag{22}$$

Comparing (21) and (22), we have

$$d\bar{Y}_i^\mu - (H\hat{y}_i^\mu + \widetilde{H}_{\theta_i}x^0 + h)dt = d\beta_i - (H\hat{\alpha}_i + h)dt = d\bar{Y}_i - (H\hat{y}_i + \widetilde{H}_{\theta_i}x^0 + h)dt. \tag{23}$$

Set $\widetilde{X}_i(t) \triangleq \hat{y}_i^\mu(t) - \hat{y}_i(t)$, and introduce $y_i^{1,\mu}(\cdot), y_i^1(\cdot)$ such that $\hat{y}_i^\mu(t) = \hat{\alpha}_i(t) + y_i^{1,\mu}(t)$ and $\hat{y}_i(t) = \hat{\alpha}_i(t) + y_i^1(t)$. It follows that

$$\hat{y}_i^\mu - \hat{y}_i = y_i^{1,\mu} - y_i^1 = y_i^\mu - y_i = \widetilde{X}_i,$$

and

$$d\widetilde{X}_i = (A_{\theta_i} + B\Lambda_{\theta_i})\widetilde{X}_i dt + B\mu dt, \quad \widetilde{X}_i(0) = 0.$$

Compute the value of the cost functional as follows

$$\begin{aligned} J_i(u_i) =& \mathbb{E}\Big\{\int_0^T \Big[(y_i - x^0)'Q(y_i - x^0) + 2(y_i - x^0)'Q\widetilde{X}_i + \widetilde{X}_i'Q\widetilde{X}_i \\ &+ (\Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i})' \cdot R(\Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i}) + 2(\Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i})'R(\Lambda_{\theta_i}\widetilde{X}_i + \mu) + (\Lambda_{\theta_i}\widetilde{X}_i + \mu)'(\Lambda_{\theta_i}\widetilde{X}_i + \mu)\Big]dt \\ &+ y_i(T)'Gy_i(T) + 2y_i(T)'G\widetilde{X}_i(T) + \widetilde{X}_i(T)'G\widetilde{X}_i(T)\Big\}. \end{aligned}$$

Hence

$$J_i(u_i) = J_i(\bar{u}_i) + \mathbb{E}\Big\{\int_0^T \Big[\widetilde{X}_i'Q\widetilde{X}_i + (\Lambda_{\theta_i}\widetilde{X}_i + \mu)'R(\Lambda_{\theta_i}\widetilde{X}_i + \mu)\Big]dt + \widetilde{X}_i'(T)G\widetilde{X}_i(T)\Big\} + 2\mathbb{X}_i,$$

where

$$\mathbb{X}_i = \mathbb{E}\Big\{\int_0^T \Big[\widetilde{X}_i'Qy_i - \widetilde{X}_i'Qx^0 + (\Lambda_{\theta_i}\widetilde{X}_i + \mu)'R(\Lambda_{\theta_i}\hat{y}_i + \lambda_{\theta_i})\Big]dt + \widetilde{X}_i'(T)Gy_i(T)\Big\}.$$

Notice that

$$\mathbb{E}\Big[\widetilde{X}_i'(t)R(t)y_i(t)\Big] = \mathbb{E}\Big[\widetilde{X}_i'(t)R(t)\mathbb{E}(y_i(t)|\mathcal{F}_{u,t}^{\bar{Y}_i,W})\Big] = \mathbb{E}\Big[\widetilde{X}_i'(t)R(t)\hat{y}_i(t)\Big].$$

Then we have

$$\mathbb{X}_i = \mathbb{E}\Big\{\int_0^T \Big[\widetilde{X}_i'Qy_i - \widetilde{X}_i'Qx^0 + (\Lambda_{\theta_i}\widetilde{X}_i + \mu)'R(\Lambda_{\theta_i}y_i + \lambda_{\theta_i})\Big]dt + \widetilde{X}_i'(T)Gy_i(T)\Big\}.$$

Define

$$p_i(t) = \pi_{\theta_i}(t)y_i(t) + \gamma_{\theta_i}(t),$$

where $\pi_{\theta_i}(\cdot), \gamma_{\theta_i}(\cdot)$ are given by (17) and (18). Applying Itô's formula to $\widetilde{X}_i'(t)p_i(t)$, integrating between 0 and T, and taking the expectation, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\widetilde{X}_i'(T)Gy_i(T)\Big]\\
=&\mathbb{E}\Big\{\int_0^T \Big[\widetilde{X}_i'\Big(A_{\theta_i} + B\Lambda_{\theta_i}\Big)'p_i + \mu'B'p_i + \widetilde{X}_i'\dot{\pi}_{\theta_i}y_i\\
&\quad + \widetilde{X}_i'\pi_{\theta_i}\Big(A_{\theta_i}y_i + B\Lambda_{\theta_i}\hat{y}_i + B\lambda_{\theta_i} + a_{\theta_i}x^0 + m\Big)\Big]dt + \int_0^T \widetilde{X}_i' d\gamma_{\theta_i}\Big\}.
\end{aligned} \tag{24}$$

Substituting (24) into $\mathbb{X}_i$, it follows that $\mathbb{X}_i = 0$ and

$$J_i(u_i) = J_i(\bar{u}_i) + \mathbb{E}\Big\{\int_0^T \Big[\widetilde{X}_i'Q\widetilde{X}_i + (\Lambda_{\theta_i}\widetilde{X}_i + \mu)'R(\Lambda_{\theta_i}\widetilde{X}_i + \mu)\Big]dt + \widetilde{X}_i'(T)G\widetilde{X}_i(T)\Big\}.$$

with $\Lambda_{\theta_i} = -R^{-1}B'\pi_{\theta_i}, \lambda_{\theta_i} = -R^{-1}B'\gamma_{\theta_i}$. The optimal μ is $\mu = 0$ as in this case, $\widetilde{X}_i \equiv 0$, which implies the optimality of $\bar{u}_i$. □

Now, we aim to derive the consistency condition satisfied by the decentralized strategies. In below, for two matrices A, B, $A \otimes B$ denotes their Kronecker product.

Lemma 2.2 *Let* (A1)-(A4) *hold, then state-average limit* $x^0 = \sum_{j=1}^K \chi_j x_j^0$ *where the set of aggregate quantities* $\bar{z} = [(x_1^0)', \cdots, (x_K^0)']'$ *and* $\bar{\gamma} = [(\gamma_1)', \cdots, (\gamma_K)']'$ *satisfies the following consistency condition:*

$$\begin{cases}
d\bar{z} = \Big[\bar{\mathbf{A}}\bar{z} + \bar{\mathbf{B}}\bar{\gamma} + \bar{\mathbf{m}}\Big]dt + \bar{\sigma}dW,\\
d\bar{\gamma} = -\Big[\check{\mathbf{A}}\bar{z} + \bar{\mathbf{G}}'\bar{\gamma} + \bar{\mathbf{s}}\Big]dt - \bar{\xi}dW,\\
\bar{z}(0) = \bar{\mathbf{x}}, \ \ \bar{\gamma}(T) = 0,
\end{cases} \tag{25}$$

with

$$\begin{cases}
\bar{\mathbf{A}} = \bar{\mathbf{G}} + \bar{\mathbf{a}} \otimes \chi, \quad \chi = [\chi_1, \cdots, \chi_K], \quad \bar{\mathbf{a}} = [a_1', \cdots, a_K']',\\
\check{\mathbf{A}} = \bar{\mathbf{q}} \otimes \chi, \quad \bar{\mathbf{q}} = [(\pi_1 a_1 - Q)', \cdots, (\pi_K a_K - Q)']',\\
\bar{\mathbf{m}} = [m', \cdots, m']', \quad \bar{\sigma} = [\widetilde{\sigma}', \cdots, \widetilde{\sigma}']', \quad \bar{\xi} = [\xi_1', \cdots, \xi_K']',\\
\bar{\mathbf{s}} = [(\pi_1 a_1)', \cdots, (\pi_K a_k)']' \cdot m, \quad \bar{\mathbf{x}} = [x', \cdots, x']',
\end{cases} \tag{26}$$

and

$$\bar{\mathbf{G}} = \begin{pmatrix} A_1 - BR^{-1}B'\pi_1 & & \\ & \ddots & \\ & & A_K - BR^{-1}B'\pi_K \end{pmatrix},$$

and

$$\bar{\mathbf{B}} = \begin{pmatrix} -BR^{-1}B' & & \\ & \ddots & \\ & & -BR^{-1}B' \end{pmatrix}.$$

Proof. It follows from Lemma 2.1 that the (decentralized) strategy $\widetilde{u}_i(t)$ of Problem (**PO**) is given by

$$\widetilde{u}_i = -R^{-1}B'\pi_{\theta_i}\hat{x}_i - R^{-1}B'\gamma_{\theta_i}, \tag{27}$$

with

$$\begin{aligned} d\hat{x}_i =& \Big[A_{\theta_i}\hat{x}_i - BR^{-1}B'(\pi_{\theta_i}\hat{x}_i + \gamma_{\theta_i}) + a_{\theta_i}x^0 + m\Big]dt \\ &+ P_{\theta_i}H'\Big[\Big(H\big(x_i - \hat{x}_i\big) + \widetilde{H}_{\theta_i}\big(x^{(N)} - x^0\big)\Big)dt + dV_i\Big] + \widetilde{\sigma}dW. \end{aligned}$$

Taking summation for $i \in \mathcal{I}_k$ and let $N \longrightarrow +\infty$,

$$d\hat{x}_k^0 = \Big[\Big(A_k - BR^{-1}B'\pi_k - P_kH'H\Big)\hat{x}_k^0 - BR^{-1}B'\gamma_k + m + a_kx^0 + P_kH'Hx_k^0\Big]dt + \widetilde{\sigma}dW.$$

Substituting (27) into (1), we have

$$dx_i = [A_{\theta_i}x_i - BR^{-1}B'\pi_{\theta_i}\hat{x}_i - BR^{-1}B'\gamma_{\theta_i} + a_{\theta_i}x^{(N)} + m]dt + \sigma dW_i + \widetilde{\sigma}dW.$$

Taking summation for $i \in \mathcal{I}_k$, and let $N \longrightarrow +\infty$,

$$dx_k^0 = \Big[A_kx_k^0 + a_kx^0 - BR^{-1}B'\pi_k\hat{x}_k^0 - BR^{-1}B'\gamma_k + m\Big]dt + \widetilde{\sigma}dW.$$

It follows that

$$x_k^0(t) = \hat{x}_k^0(t), \qquad a.s., a.e. \tag{28}$$

for any $t \in [0, T]$. With (18), we have for $k = 1, 2, \cdots, K$,

$$\begin{cases} dx_k^0 = \big[A_kx_k^0 - BR^{-1}B'\pi_kx_k^0 + a_kx^0 - BR^{-1}B'\gamma_k + m\big]\,dt + \widetilde{\sigma}dW, \quad x_k^0 = x, \\ d\gamma_k + \big[(A_k' - \pi_kBR^{-1}B')\gamma_k + \pi_k(a_kx^0 + m) - Qx^0\big]\,dt + \xi_kdW = 0, \quad \gamma_k(T) = 0. \end{cases} \tag{29}$$

Write the above systems in compact form for $k = 1, 2, \cdots, K$, we formulate (25). □

Similar to [4], suppose $\bar{\gamma} = K\bar{z} + \Phi$, thus we have the following matrix-valued equations for K and Φ:

$$\begin{cases} \dot{K} + K\bar{\mathbf{A}} + \bar{\mathbf{G}}'K + K\bar{\mathbf{B}}K + \check{\mathbf{A}} = 0, \\ \dot{\Phi} + \big(\bar{\mathbf{G}}' + K\bar{\mathbf{B}}\big)\Phi + (\bar{\mathbf{s}} + K\bar{\mathbf{m}}) = 0, \\ K(T) = 0, \qquad \Phi(T) = 0. \end{cases} \tag{30}$$

K in (30) is nonsymmetric Riccati equation. We first state the following result based on [4] (Proposition 3.2) which is a version of Radon's lemma for nonsymmetric Riccati equation. Suppose two-point boundary problem

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} \xi^1 \\ -\eta^1 \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{A}} & \bar{\mathbf{B}} \\ \check{\mathbf{A}} & \bar{\mathbf{G}}' \end{pmatrix}\begin{pmatrix} \xi^1 \\ \eta^1 \end{pmatrix}, \\ \xi^1(t_0) = 0, \eta^1(T) = 0, \end{cases}$$

admits a unique solution for any $t_0 \in [0, T]$, respectively. Then there is a unique solution $K(\cdot)$ to the nonsymmetric Riccati equation (30). Then, applying the Banach fixed point theorem for two-point boundary problem, we have the following general existence result to nonsymmetric Riccati equation (see [26] for more details):

Propsition 2.1 *Let* (A1)-(A4) *hold, there exists a unique solution of* (30) *if* $L<1$ *where*

$$L = T\|\check{\mathbf{A}}\|_T\|\bar{\mathbf{B}}\|_T \cdot \exp((2\|\bar{\mathbf{A}}\|_T + 2\|\bar{\mathbf{G}}\|_T + \|\bar{\mathbf{B}}\|_T + \|\check{\mathbf{A}}\|_T)T)$$

and $\|\cdot\|_T$ *denotes the super-norm of matrix-valued function on* $[0,T]$.

Given the special structure on $\check{\mathbf{A}}$, a relaxed condition is given below which is obtained in [4]:

Propsition 2.2 *Let* (A1)-(A4) *hold. Suppose* $\check{\mathbf{A}}$ *is invertible, let* $\phi(t,s)$ *is the fundamental solution to* $\bar{\mathbf{G}}$ *and* $\|\phi\|_T = \sup_{0\le t\le T}\sqrt{\int_t^T \|\phi'(s,t)\check{\mathbf{A}}_s^{\frac{1}{2}}\|^2 ds}$, $\|\bar{\mathbf{A}}-\bar{\mathbf{G}}\|_T = \sup_{0\le t\le T}\|(\bar{\mathbf{A}}-\bar{\mathbf{G}})_t\check{\mathbf{A}}_t^{-\frac{1}{2}}\|$. *Then there exists a unique solution of* (30) *if*

$$\sqrt{T}\|\phi\|_T\|\bar{\mathbf{A}}-\bar{\mathbf{G}}\|_T < 1.$$

Proof. Applying the similar procedures in Theorem III.6 in [4], we can obtain the condition above for forward-backward SDE. Details are omitted. □

Unlike the condition in terms of two-point boundary problems, the condition of Proposition 2.2 is given by matrix norm which is more checkable. For its illustration, we present some numerical example in Section 4.

Finally, we obtain the estimation of the solution of (25) which will be used in the following section.

Lemma 2.3 *There exists a constant* c *such that*

$$\sup_{1\le k\le K}\mathbb{E}\sup_{0\le t\le T}|x_k^0(t)|^2 + \sup_{1\le k\le K}\mathbb{E}\sup_{0\le t\le T}|\gamma_k(t)|^2 \le c. \tag{31}$$

Proof. By (29), it follows from the standard estimations for SDE and BSDE that, there exists a constant c that

$$\mathbb{E}\sup_{0\le t\le T}|\gamma_k(t)|^2 \le c\mathbb{E}\int_0^T (|x^0(t)|^2 + |m(t)|^2)dt,$$

and

$$\mathbb{E}\sup_{0\le t\le T}|x_k^0(t)|^2 \le c\mathbb{E}\int_0^T (|x^0(t)|^2 + |\gamma_k(t)|^2 + |m(t)|^2 + |\widetilde{\sigma}(t)|^2)dt.$$

Therefore,

$$\sum_{k=1}^K \mathbb{E}\sup_{0\le t\le T}|x_k^0(t)|^2 \le c\sum_{k=1}^K \mathbb{E}\int_0^T |x_k^0(t)|^2 dt + c\mathbb{E}\int_0^T (|m(t)|^2 + |\widetilde{\sigma}(t)|^2)dt.$$

Hence there exists a constant c such that

$$\sup_{1\le k\le K}\mathbb{E}\sup_{0\le t\le T}|x_k^0(t)|^2 \le c,$$

and

$$\sup_{1\le k\le K}\mathbb{E}\sup_{0\le t\le T}|\gamma_k(t)|^2 \le c.$$

□

3 ϵ-Nash Equilibrium for Problem (PO)

Now we show $(\widetilde{u}_1, \widetilde{u}_2, \cdots, \widetilde{u}_N)$ satisfies the ϵ-Nash equilibrium for (**PO**). Here, for $1 \le i \le N$, $\widetilde{u}_i$ is given by (27) and γ_{θ_i} satisfies the consistent condition (25). We first give the definition of ϵ-Nash equilibrium.

Definition 3.1 *A set of controls $u_i(\cdot) \in \mathcal{U}_i$, $1 \le i \le N$, for N agents is called an ϵ-Nash equilibrium with respect to the costs $\mathcal{J}_i$, $1 \le i \le N$, if there exists $\epsilon = \epsilon_N \ge 0$ and $\epsilon_N \to 0$ as $N \to \infty$ such that for any fixed $1 \le i \le N$, we have*

$$\mathcal{J}_i(u_i, u_{-i}) \le \mathcal{J}_i(u_i', u_{-i}) + \epsilon_N, \tag{32}$$

when any alternative control $u_i'(\cdot) \in \mathcal{U}_i$ is applied by $\mathcal{A}_i$.

Our main result in this section is as follows.

Theorem 3.1 *Let* (A1)-(A4) *hold, then $(\widetilde{u}_1, \widetilde{u}_2, \cdots, \widetilde{u}_N)$ is an ϵ-Nash equilibrium of Problem* (**PO**) *with $\epsilon = O(\frac{1}{\sqrt{N}} + \varepsilon_N)$, where $\varepsilon_N := \sup_{1 \le k \le K} |\chi_k^{(N)} - \chi_k| \to 0$ as $N \to \infty$.*

For $\bar{u}_i(\cdot)$ defined in (19) and any $u_i(\cdot) \in \mathcal{U}_i$, we have

$$\mathcal{J}_i(\widetilde{u}_i, \widetilde{u}_{-i}) - \mathcal{J}_i(u_i, \widetilde{u}_{-i}) \le \mathcal{J}_i(\widetilde{u}_i, \widetilde{u}_{-i}) - J_i(\bar{u}_i(\cdot)) + J_i(u_i(\cdot)) - \mathcal{J}_i(u_i, \widetilde{u}_{-i}).$$

Therefore, in order to show that $(\widetilde{u}_1, \widetilde{u}_2, \cdots, \widetilde{u}_N)$ satisfies the ϵ-Nash equilibrium, we will study $\mathcal{J}_i(\widetilde{u}_i, \widetilde{u}_{-i}) - J_i(\bar{u}_i(\cdot))$ and $J_i(u_i(\cdot)) - \mathcal{J}_i(u_i, \widetilde{u}_{-i})$ in the following subsections, respectively.

3.1 Estimation of $|\mathcal{J}_i(\widetilde{u}_i, \widetilde{u}_{-i}) - J_i(\bar{u}_i(\cdot))|$

In order to estimate $|\mathcal{J}_i(\widetilde{u}_i, \widetilde{u}_{-i}) - J_i(\bar{u}_i(\cdot))|$, first we need to obtain the estimations of the difference between the optimal state-average and the frozen term(see Lemma 3.2) and the difference between the decentralized and centralized states and filters(see Lemma 3.3). For $k \in \Theta, i \in \mathcal{I}_k$, applying $\widetilde{u}_i(\cdot)$ for $\mathcal{A}_i$, we have the following close-loop state

$$\begin{cases} dx_i = [A_k x_i - BR^{-1}B'\big(\pi_k \hat{x}_i + \mathbf{e}_k(K\bar{z} + \Phi)\big) + a_k x^{(N)} + m]dt + \sigma dW_i + \widetilde{\sigma} dW, \\ d\hat{x}_i = \Big[A_k \hat{x}_i - BR^{-1}B'\big(\pi_k \hat{x}_i + \mathbf{e}_k\left(K\bar{z} + \Phi\right)\big) + a_k x^0 + m\Big]dt \\ \qquad + P_k H'\Big[dY_i - \big(H\hat{x}_i + \widetilde{H}_k x^0 + h\big)dt\Big] + \widetilde{\sigma} dW, \\ dY_i = [Hx_i + \widetilde{H}_k x^{(N)} + h]dt + dV_i, \\ x_i(0) = x,\ \hat{x}_i(0) = x,\ Y_i(0) = 0, \end{cases} \tag{33}$$

where for $1 \le k \le K$, $\mathbf{e}_k$ is the $n \times (nK)$ matrix with the $n \times n$ identity matrix I_n located in its $k-th$ block and other blocks are null matrix, that is $\mathbf{e}_k = [0_{n\times n}, \cdots, 0_{n\times n}, I_n, 0_{n\times n}, \cdots, 0_{n\times n}]$. The auxiliary system (of limiting problem) is given by

$$\begin{cases} dy_i = [A_k y_i - BR^{-1}B'\big(\pi_k \hat{y}_i + \mathbf{e}_k\left(K\bar{z} + \Phi\right)\big) + a_k x^0 + m]dt + \sigma dW_i + \widetilde{\sigma} dW, \\ d\hat{y}_i = \Big[A_k \hat{y}_i - BR^{-1}B'\big(\pi_k \hat{y}_i + \mathbf{e}_k\left(K\bar{z} + \Phi\right)\big) + a_k x^0 + m\Big]dt \\ \qquad + P_k H'\Big[d\bar{Y}_i - \big(H\hat{y}_i + \widetilde{H}_k x^0 + h\big)dt\Big] + \widetilde{\sigma} dW, \\ d\bar{Y}_i = [Hy_i + \widetilde{H}_k x^0 + h]dt + dV_i, \\ y_i(0) = x,\ \hat{y}_i(0) = x,\ \bar{Y}_i(0) = 0. \end{cases} \tag{34}$$

Based on (33), we derive that

$$\begin{cases} dx_k^{(N)} =[(A_k x_k^{(N)} - BR^{-1}B'\big(\pi_k \hat{x}_k^{(N)} + \mathbf{e}_k(K\bar{z}+\Phi)\big) + a_k x^{(N)} + m]dt \\ \qquad\qquad + \dfrac{1}{N_k}\displaystyle\sum_{i\in\mathcal{I}_k} \sigma dW_i + \widetilde{\sigma} dW, \\ d\hat{x}_k^{(N)} =\Big[A_k \hat{x}_k^{(N)} - BR^{-1}B'\big(\pi_k \hat{x}_k^{(N)} + \mathbf{e}_k(K\bar{z}+\Phi)\big) + a_k x^0 + m\Big]dt \\ \qquad\qquad + \widetilde{\sigma} dW + P_k H'\Big[dY_k^{(N)} - \big(H\hat{x}_k^{(N)} + \widetilde{H}_k x^0 + h\big)dt\Big], \\ dY_k^{(N)} =[Hx_k^{(N)} + \widetilde{H}_k x^{(N)} + h]dt + \dfrac{1}{N_k}\displaystyle\sum_{i\in\mathcal{I}_k} dV_i, \\ x_k^{(N)}(0) =x,\ \hat{x}_k^{(N)}(0) = x,\ Y_k^{(N)}(0) = 0, \end{cases} \tag{35}$$

where $Y_k^{(N)} = \frac{1}{N_k}\sum_{i\in\mathcal{I}_k} Y_i$.

For (34) and (35), applying the same method as in Lemma 2.3 and using (31), we have the following result.

Lemma 3.1 *There exists a constant c such that*

$$\mathbb{E}\sup_{0\le t\le T}|y_i(t)|^2 + \mathbb{E}\sup_{0\le t\le T}|\hat{y}_i(t)|^2 + \mathbb{E}\sup_{0\le t\le T}|\bar{Y}_i(t)|^2 \le c,$$

and

$$\sup_{k\in\Theta}\mathbb{E}\sup_{0\le t\le T}|x_k^{(N)}(t)|^2 + \sup_{k\in\Theta}\mathbb{E}\sup_{0\le t\le T}|\hat{x}_k^{(N)}(t)|^2 \le c,$$

The following lemma establishes the estimations of the difference between the optimal state-average and the frozen term

Lemma 3.2

$$\sup_{k\in\Theta}\sup_{0\le t\le T}\mathbb{E}\Big|x_k^{(N)}(t) - x_k^0(t)\Big|^2 = O\Big(\frac{1}{N} + \varepsilon_N^2\Big), \tag{36}$$

$$\sup_{k\in\Theta}\sup_{0\le t\le T}\mathbb{E}\Big|x_k^{(N)}(t) - x_k^0(t)\Big|^2 = O\Big(\frac{1}{N} + \varepsilon_N^2\Big). \tag{37}$$

Proof. By (29) and (35), we get

$$\begin{cases} d\Big(x_k^{(N)} - x_k^0\Big) =\Big[A_k\big(x_k^{(N)} - x_k^0\big) - BR^{-1}B'\pi_k(\hat{x}_k^{(N)} - x_k^0)) + a_k(x^{(N)} - x^0)\Big]dt+ \\ \qquad\qquad + \dfrac{1}{N_k}\displaystyle\sum_{i\in\mathcal{I}_k}\sigma dW_i, \\ x_k^{(N)}(0) - x_k^0(0) =0, \end{cases} \tag{38}$$

and

$$\begin{cases} d\Big(\hat{x}_k^{(N)} - x_k^0\Big) =\Big[\big(A_k - BR^{-1}B'\pi_k\big)\big(\hat{x}_k^{(N)} - x_k^0\big)\Big]dt \\ \qquad\qquad + P_k H'\Big[H\big(x_k^{(N)} - \hat{x}_k^{(N)}\big) + \widetilde{H}_k\big(x^{(N)} - x^0\big)\Big]dt + P_k H'\dfrac{1}{N_k}\displaystyle\sum_{i\in\mathcal{I}_k}^{N} dV_i, \\ \hat{x}_k^{(N)}(0) - x_k^0(0) =0. \end{cases} \tag{39}$$

It follows from (38), (39) and (28) that

$$
\begin{aligned}
&\mathbb{E}\Big|x_k^{(N)}(t)-x_k^0(t)\Big|^2\\
\leq& C\mathbb{E}\int_0^t \left(\Big|x_k^{(N)}(s)-x_k^0(s)\Big|^2+\Big|\hat{x}_k^{(N)}(s)-x_k^0(s)\Big|^2+\Big|x^{(N)}(s)-x^0(s)\Big|^2\right)ds+C\mathbb{E}\Big|\int_0^t \frac{1}{N_k}\sum_{i\in\mathcal{I}_k}\sigma dW_i\Big|^2,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\Big|\hat{x}_k^{(N)}(t)-x_k^0(t)\Big|^2\\
\leq& C\mathbb{E}\int_0^t \left(\Big|\hat{x}_k^{(N)}(s)-x_k^0(s)\Big|^2+\Big|x_k^{(N)}(s)-x_k^0(s)\Big|^2+\Big|x^{(N)}(s)-x^0(s)\Big|^2\right)ds+C\mathbb{E}\Big|\int_0^t \frac{1}{N_k}\sum_{i\in\mathcal{I}_k}dV_i\Big|^2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\Big|x^{(N)}(s)-x^0(s)\Big|^2\\
=&\Big|\sum_{k=1}^K(\chi_k^{(N)}x_k^{(N)}(s)-\chi_k x_k^0(s))\Big|^2\\
=&\Big|\sum_{k=1}^K(\chi_k^{(N)}x_k^{(N)}(s)-\chi_k x_k^{(N)}(s))+\sum_{k=1}^K(\chi_k x_k^{(N)}(s)-\chi_k x_k^0(s))\Big|^2\\
\leq& C\sup_{k\in\Theta}|\chi_k^{(N)}-\chi_k|^2\sum_{k=1}^K|x_k^{(N)}(s)|^2+C\sum_{k=1}^K|x_k^{(N)}(s)-x_k^0(s)|^2.
\end{aligned}
\tag{40}
$$

and

$$
\mathbb{E}\Big|\int_0^t \frac{1}{N_k}\sum_{i\in\mathcal{I}_k}\sigma dW_i\Big|^2\sim\mathbb{E}\Big|\int_0^t \frac{1}{N_k}\sum_{i\in\mathcal{I}_k}dV_i\Big|^2=O\Big(\frac{1}{N}\Big).
$$

Then (36) and (37) follow by Gronwall's inequality. □

Considering the difference between the decentralized and centralized states and filters, we have the following estimates:

Lemma 3.3

$$
\sup_{1\leq i\leq N}\left[\sup_{0\leq t\leq T}\mathbb{E}\Big|x_i(t)-y_i(t)\Big|^2\right]=O\Big(\frac{1}{N}+\varepsilon_N^2\Big),\tag{41}
$$

$$
\sup_{1\leq i\leq N}\left[\sup_{0\leq t\leq T}\mathbb{E}\Big|\hat{x}_i(t)-\hat{y}_i(t)\Big|^2\right]=O\Big(\frac{1}{N}+\varepsilon_N^2\Big),\tag{42}
$$

$$
\sup_{1\leq i\leq N}\left[\sup_{0\leq t\leq T}\mathbb{E}\Big|Y_i(t)-\bar{Y}_i(t)\Big|^2\right]=O\Big(\frac{1}{N}+\varepsilon_N^2\Big).\tag{43}
$$

Proof. By (33) and (34), we get

$$
\begin{aligned}
&\sup_{0\leq s\leq t}\mathbb{E}\Big|x_i(s)-y_i(s)\Big|^2\\
\leq& C\int_0^t\mathbb{E}\Big|x_i(s)-y_i(s)\Big|^2ds+C\mathbb{E}\int_0^t\left[\Big|\hat{x}_i(s)-\hat{y}_i(s)\Big|^2+\Big|x^{(N)}(s)-x^0(s)\Big|^2\right]ds,
\end{aligned}
$$

and

$$\sup_{0\le s\le t}\mathbb{E}\Big|\hat{x}_i(s)-\hat{y}_i(s)\Big|^2\le C\int_0^t\mathbb{E}\Big|\hat{x}_i(s)-\hat{y}_i(s)\Big|^2ds+C\mathbb{E}\Big|Y_i(t)-\bar{Y}_i(t)\Big|^2,$$

and

$$\mathbb{E}\Big|Y_i(t)-\bar{Y}_i(t)\Big|^2\le C\int_0^t\mathbb{E}\Big|x_i(s)-y_i(s)\Big|^2ds+C\int_0^t\mathbb{E}\Big|x^{(N)}(s)-x^0(s)\Big|^2ds.$$

Recalling (40), by virtue of Lemma 3.2 and Gronwall's inequality, we obtain (41)-(43). □

The following is the main result of this subsection.

Propsition 3.1 *For* $\forall\ 1\le i\le N$,

$$\Big|\mathcal{J}_i(\widetilde{u}_i,\widetilde{u}_{-i})-J_i(\bar{u}_i)\Big|=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big).$$

Proof. Applying Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\sup_{0\le t\le T}\mathbb{E}\Big|\big|x_i(t)-x^{(N)}(t)\big|^2-\big|y_i(t)-x^0(t)\big|^2\Big|\\
\le&\sup_{0\le t\le T}\mathbb{E}\big|x_i(t)-x^{(N)}(t)-y_i(t)+x^0(t)\big|^2\\
&+2\sup_{0\le t\le T}\mathbb{E}\Big[\big|y_i(t)-x^0(t)\big|\cdot\big|x_i(t)-x^{(N)}(t)-y_i(t)+x^0(t)\big|\Big]\\
\le&\sup_{0\le t\le T}\mathbb{E}\big|x_i(t)-y_i(t)-(x^{(N)}(t)-x^0(t))\big|^2\\
&+2\Big(\sup_{0\le t\le T}\mathbb{E}\big|y_i(t)-x^0(t)\big|^2\Big)^{\frac12}\cdot\Big(\sup_{0\le t\le T}\mathbb{E}\big|x_i(t)-y_i(t)-(x^{(N)}(t)-x^0(t))\big|^2\Big)^{\frac12}\\
=&O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big),
\end{aligned}$$

where the last equality is obtained by using the results of Lemmas 3.1, 3.2 and 3.3. Similarly, by (19), (27) and (42), applying the same technique we get

$$\sup_{0\le t\le T}\mathbb{E}\Big|\big|\widetilde{u}_i(t)\big|^2-\big|\bar{u}_i(t)\big|^2\Big|=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big).$$

In addition,

$$\mathbb{E}\Big|\big|x_i(T)\big|^2-\big|y_i(T)\big|^2\Big|=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big).$$

Then

$$\Big|\mathcal{J}_i(\widetilde{u}_i,\widetilde{u}_{-i})-J_i(\bar{u}_i)\Big|=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big),$$

which completes the proof. □

1 3.2 Estimation of $|\mathcal{J}_i(u_i,\widetilde{u}_{-i})-J_i(u_i(\cdot))|$

The proof of $|\mathcal{J}_i(u_i,\widetilde{u}_{-i})-J_i(u_i(\cdot))|$ is similar to the proof in Subsection 3.1. We will consider the state and cost under perturbation. Thus, we give some new notations first. For $i\in\mathcal{I}_k$, consider a perturbed control $u_i\in\mathcal{U}_i$ for $\mathcal{A}_i$ and introduce

$$\begin{cases} dl_i=[A_kl_i+Bu_i+a_kl^{(N)}+m]dt+\sigma dW_i+\widetilde{\sigma}dW,\\ dY_i^l=[Hl_i+\widetilde{H}_kl^{(N)}+h]dt+dV_i,\\ l_i(0)=x,\ Y_i^l(0)=0,\end{cases}\tag{44}$$

whereas other agents of same type still keep the control $\widetilde{u}_j, j\neq i$, i.e.,

$$\begin{cases} dl_j=&[A_{\theta_j}l_j-BR^{-1}B'\big(\pi_{\theta_j}\hat{l}_j+\mathbf{e}_{\theta_j}(K\bar{z}+\Phi)\big)+a_{\theta_j}l^{(N)}+m]dt\\ &+\sigma dW_j+\widetilde{\sigma}dW,\\ d\hat{l}_j=&\Big[A_{\theta_j}\hat{l}_j-BR^{-1}B'\big(\pi_{\theta_j}\hat{l}_j+\mathbf{e}_{\theta_j}\big(K\bar{z}+\Phi\big)+a_{\theta_j}x^0+m\Big]dt\\ &+\widetilde{\sigma}dW+PH'\Big[dY_j^l-\big(H\hat{l}_j+\widetilde{H}_{\theta_j}x^0+h\big)dt\Big],\\ dY_j^l=&[Hl_j+\widetilde{H}_{\theta_j}l^{(N)}+h]dt+dV_j,\\ l_j(0)=&x,\ \hat{l}_j(0)=x,\ Y_j^l(0)=0,\end{cases}\tag{45}$$

where $l^{(N)}(t)=\frac{1}{N}\sum_{k=1}^N l_k(t)$. By the definition of ϵ-Nash equilibrium, we need only consider the perturbation control $u_i\in\mathcal{U}_i$ such that $\mathcal{J}_i(u_i,\widetilde{u}_{-i})\leq\mathcal{J}_i(\widetilde{u}_i,\widetilde{u}_{-i})$, which implies

$$\frac{1}{2}\mathbb{E}\int_0^T u_i(t)'R(t)u_i(t)dt\leq\mathcal{J}_i(u_i,\widetilde{u}_{-i})\leq\mathcal{J}_i(\widetilde{u}_i,\widetilde{u}_{-i})=J_i(\bar{u}_i)+O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big),$$

i.e.,

$$\mathbb{E}\int_0^T|u_i(t)|^2dt\leq c_1,\tag{46}$$

2 where c_1 is a positive constant which is independent of N. Then we have the following result.

Lemma 3.4 *There exists a constant c independent of N and j such that*

$$\sup_{1\leq j\leq N}\sup_{0\leq t\leq T}\mathbb{E}\left[|l_j(t)|^2\right]+\sup_{1\leq j\leq N}\sup_{0\leq t\leq T}\mathbb{E}\left[|\hat{l}_j(t)|^2\right]+\sup_{1\leq j\leq N}\sup_{0\leq t\leq T}\mathbb{E}\left[|Y_j(t)|^2\right]\leq c_3.$$

Proof. By (44) and (45), it holds that

$$\begin{aligned}|l_i(t)|^2\leq c_1\Big\{&|x|^2+\int_0^t\Big[|l_i(s)|^2+|u_i(s)|^2+\frac{1}{N}\sum_{j=1}^N|l_j(s)|^2+|m(s)|^2\Big]ds\\ &+\Big|\int_0^t\sigma(s)dW_i(s)\Big|^2+\Big|\int_0^t\widetilde{\sigma}(s)dW(s)\Big|^2\Big\},\end{aligned}\tag{47}$$

and for $j\neq i$,

$$\begin{aligned}|l_j(t)|^2\leq c_1\Big\{&|x|^2+\int_0^t\Big[|l_j(s)|^2+|\hat{l}_j(s)|^2+|x^0(s)|^2+\frac{1}{N}\sum_{j=1}^N|l_j(s)|^2+|m(s)|^2\Big]ds\\ &+\Big|\int_0^t\sigma(s)dW_j(s)\Big|^2+\Big|\int_0^t\widetilde{\sigma}(s)dW(s)\Big|^2\Big\},\end{aligned}\tag{48}$$

$$
\begin{aligned}
|\hat{l}_j(t)|^2 \leq c_1\Big\{ |x|^2 + \int_0^t \Big[|\hat{l}_j(s)|^2 + |x^0(s)|^2 + |m(s)|^2 + |Y_j^l(s)|^2 + |h(s)|^2 \Big] ds \\
+ \Big| \int_0^t \sigma(s) dW_j(s) \Big|^2 + \Big| \int_0^t \widetilde{\sigma}(s) dW(s) \Big|^2 \Big\},
\end{aligned} \tag{49}
$$

$$
|Y_j^l(t)|^2 \leq c_1 \Big\{ \int_0^t \Big[|l_j(s)|^2 + \frac{1}{N} \sum_{j=1}^N |l_j(s)|^2 + |h(s)|^2 \Big] ds + \Big| \int_0^t dV_j(s) \Big|^2 \Big\}, \tag{50}
$$

where c_1 is a positive constant independent of N. Thus,

$$
\begin{aligned}
& \sum_{j=1}^N \mathbb{E}\Big[|l_j(t)|^2\Big] + \sum_{j=1}^N \mathbb{E}\Big[|\hat{l}_j(t)|^2\Big] + \sum_{j=1}^N \mathbb{E}\Big[|Y_j(t)|^2\Big] \\
\leq & c_1 \Big\{ N|x|^2 + \mathbb{E} \int_0^t \Big[\sum_{j=1}^N |l_j(s)|^2 + |u_i(s)|^2 + \sum_{j=1}^N |\hat{l}_j(s)|^2 + N|x^0(s)|^2 + N|m(s)|^2 + \sum_{j=1}^N |Y_j^l(s)|^2 \\
& + |h(s)|^2 \Big] ds + \sum_{j=1}^N \mathbb{E} \Big| \int_0^t \sigma(s) dW_j(s) \Big|^2 + N\mathbb{E} \Big| \int_0^t \widetilde{\sigma}(s) dW(s) \Big|^2 + N\mathbb{E} \Big| \int_0^t dV_j(s) \Big|^2 \Big\} \\
\leq & c_1 \Big\{ N|x|^2 + \int_0^t \Big[\sum_{j=1}^N \mathbb{E}|l_j(s)|^2 + \sum_{j=1}^N \mathbb{E}|\hat{l}_j(s)|^2 + \sum_{j=1}^N \mathbb{E}|Y_j(s)|^2 \Big] ds \\
& + N\mathbb{E} \int_0^t \Big(|u_i(s)|^2 + |x^0(s)|^2 + |m(s)|^2 + |h(s)|^2 + |\sigma(s)|^2 + |\widetilde{\sigma}(s)|^2 + 1 \Big) ds \Big\}.
\end{aligned}
$$

By (46), we can see that $u_i(\cdot)$ is L^2-bounded. Then by Gronwall's inequality, it follows that there exists a constant c_2 independent of N such that

$$
\frac{1}{N} \sum_{j=1}^N \sup_{0 \leq t \leq T} \mathbb{E}\left[|l_j(t)|^2\right] + \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq t \leq T} \mathbb{E}\left[|\hat{l}_j(t)|^2\right] + \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq t \leq T} \mathbb{E}\left[|Y_j(t)|^2\right] \leq c_2. \tag{51}
$$

Plugging (51) into (47), (48), (49) and (50), it follows from Gronwall inequality that there exists a constant c_2 independent of N and j such that

$$
\sup_{0 \leq t \leq T} \mathbb{E}\left[|l_j(t)|^2\right] + \sup_{0 \leq t \leq T} \mathbb{E}\left[|\hat{l}_j(t)|^2\right] + \sup_{0 \leq t \leq T} \mathbb{E}\left[|Y_j(t)|^2\right] \leq c_3.
$$

1 □

Correspondingly, the system for agent $\mathcal{A}_i$ under control u_i in **(LPO)** is as follows

$$
\begin{cases}
dl_i^0 = [A_k l_i^0 + Bu_i + a_k x^0 + m]dt + \sigma dW_i + \widetilde{\sigma} dW, \\
dY_i^{l,0} = [H l_i^0 + \widetilde{H}_k x^0 + h]dt + dV_i, \\
l_i^0(0) = x, \ Y_i^{l,0}(0) = 0,
\end{cases} \tag{52}
$$

and for agent $\mathcal{A}_j,\ j \neq i$,

$$
\begin{cases}
dl_j^0 = [A_{\theta_j} l_j^0 - BR^{-1}B'\big(\pi_{\theta_j}\hat{l}_j^0 + \mathbf{e}_{\theta_j}(K\bar{z} + \Phi)\big) + a_{\theta_j}x^0 + m]dt + \sigma dW_j + \widetilde{\sigma} dW, \\
d\hat{l}_j^0 = \Big[A_{\theta_j}\hat{l}_j^0 - BR^{-1}B'\big(\pi_{\theta_j}\hat{l}_j^0 + \mathbf{e}_{\theta_j}(K\bar{z} + \Phi)\big) + a_{\theta_j}x^0 + m\Big]dt \\
\qquad\quad + \widetilde{\sigma} dW + PH'\Big[dY_j^{l,0} - \big(H\hat{l}_j^0 + \widetilde{H}_{\theta_j}x^0 + h\big)dt\Big], \\
dY_j^{l,0} = [Hl_j^0 + \widetilde{H}_{\theta_j}x^0 + h]dt + dV_j, \\
l_j^0(0) = x,\ \hat{l}_j^0(0) = x,\ Y_j^{l,0}(0) = 0.
\end{cases}
$$

In order to give necessary estimates in (**PO**) and (**LPO**), we introduce the intermediate state for $\mathcal{A}_i$ as

$$
\begin{cases}
dn_i = [A_k n_i + Bu_i + a_k \frac{N-1}{N} n^{(N-1)} + m]dt + \sigma dW_i + \widetilde{\sigma} dW, \\
dY_i^n = [Hn_i + \frac{N-1}{N}\widetilde{H}_k n^{(N-1)} + h]dt + dV_i, \\
n_i(0) = x,\ Y_i^n(0) = 0,
\end{cases}
$$

1 and for $j \neq i$,

$$
\begin{cases}
dn_j = [A_{\theta_j} n_j - BR^{-1}B'\big(\pi_{\theta_j}\hat{n}_j + \mathbf{e}_{\theta_j}(K\bar{z} + \Phi)\big) + a_{\theta_j}\frac{N-1}{N} n^{(N-1)} + m]dt + \sigma dW_j + \widetilde{\sigma} dW, \\
d\hat{n}_j = \Big[A_{\theta_j}\hat{n}_j - BR^{-1}B'\big(\pi_{\theta_j}\hat{n}_j + \mathbf{e}_{\theta_j}(K\bar{z} + \Phi)\big) + a_{\theta_j}x^0 + m\Big]dt \\
\qquad\quad + \widetilde{\sigma} dW + PH'\Big[dY_j^n - \big(H\hat{n}_j + \widetilde{H}_{\theta_j}x^0 + h\big)dt\Big], \\
dY_j^n = [Hn_j + \frac{N-1}{N}\widetilde{H}_{\theta_j} n^{(N-1)} + h]dt + dV_j, \\
n_j(0) = x,\ \hat{n}_j(0) = x,\ Y_j^n(0) = 0,
\end{cases}
\tag{53}
$$

where $n^{(N-1)} \triangleq \frac{1}{N-1}\sum_{j=1,j\neq i}^{N} n_j$. Define

$$
l^{(N-1)} \triangleq \frac{1}{N-1}\sum_{j=1,j\neq i}^{N} l_j, \qquad \hat{l}^{(N-1)} \triangleq \frac{1}{N-1}\sum_{j=1,j\neq i}^{N} \hat{l}_j,
$$

$$
Y_l^{(N-1)} \triangleq \frac{1}{N-1}\sum_{j=1,j\neq i}^{N} Y_j^l, \qquad Y_n^{(N-1)} \triangleq \frac{1}{N-1}\sum_{j=1,j\neq i}^{N} Y_j^n.
$$

2 By (45) and (53), we have the following estimates on these states.

Lemma 3.5

$$
\sup_{0\leq t\leq T}\mathbb{E}\Big|\hat{l}^{(N-1)} - \hat{n}^{(N-1)}\Big|^2 = O\Big(\frac{1}{N} + \varepsilon_N^2\Big),
$$

$$
\sup_{0\leq t\leq T}\mathbb{E}\Big|l^{(N-1)} - n^{(N-1)}\Big|^2 = O\Big(\frac{1}{N} + \varepsilon_N^2\Big),
$$

$$\sup_{0\le t\le T}\mathbb{E}\Big|Y_l^{(N-1)}-Y_n^{(N-1)}\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big),$$
$$\sup_{0\le t\le T}\mathbb{E}\Big|l^{(N)}-l^{(N-1)}\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big),$$
$$\sup_{0\le t\le T}\mathbb{E}\Big|\hat{n}^{(N-1)}-\hat{x}^0\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big),$$
$$\sup_{0\le t\le T}\mathbb{E}\Big|n^{(N-1)}-x^0\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big).$$

Proof. The proof is similar to that of Lemma 3.3 and omitted. □

In addition, based on Lemma 3.5, we have

Lemma 3.6

$$\sup_{0\le t\le T}\mathbb{E}\Big|l^{(N)}-x^0\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big), \tag{54}$$
$$\sup_{0\le t\le T}\mathbb{E}\Big|l_i-l_i^0\Big|^2=O\Big(\frac{1}{N}+\varepsilon_N^2\Big). \tag{55}$$

Proof. (54) follows from Lemma 3.5 directly. By (44) and (52), and using (54), we get (55). □

Finally, applying the same technique as the proof of Proposition 3.1, we obtain the following proposition.

Propsition 3.2 *For any* $1\le i\le N$*,*

$$\Big|\mathcal{J}_i(u_i,\widetilde{u}_{-i})-J_i(u_i)\Big|=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big).$$

3.3 Proof of Theorem 3.1

Combining Propositions 3.1 and 3.2, we have

$$\mathcal{J}_i(\widetilde{u}_i,\widetilde{u}_{-i})=J_i(\bar{u}_i)+O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big)\le J_i(u_i)+O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big)=\mathcal{J}_i(u_i,\widetilde{u}_{-i})+O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big).$$

Thus, Theorem 3.1 follows by taking $\epsilon=O\Big(\frac{1}{\sqrt{N}}+\varepsilon_N\Big)$. □

4 Numerical Results

Consider the case: $n=2, K=2$, thus there have two different types of agents: type-1 and type-2 respectively. The state of each agent has two components. Consider the following parameters of state, observation and cost:

$$
\begin{cases}
A_1 = \begin{pmatrix} 0.12 & 0.2 \\ 0.23 & 0.17 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.18 & 0.1 \\ 0.15 & 0.23 \end{pmatrix}, \\
a_1 = \begin{pmatrix} 0.18 & 0.2 \\ 0.1 & 0.13 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0.21 & 0.1 \\ 0.17 & 0.12 \end{pmatrix}, \\
B = (0.31, 0.22)', \chi_1 = 0.55, \chi_2 = 0.45, m = (2.7, 0.45)', \\
H = (0.23, 0.45), H_1 = (0.03, 0.08), H_2 = (-0.04, 0.05), \\
Q = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.13 \end{pmatrix}, G = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.6 \end{pmatrix}, h = 0.02, R = 2, \\
\sigma = (0.75, 0.65)', \quad \widetilde{\sigma} = (0.35, 0.85)', \quad T = 3, \quad x(0) = \mathbf{0}.
\end{cases}
$$

Corresponding to the above parameters, the matrix Riccati equations $P_1, P_2; \pi_1, \pi_2$ are all of sizes 2×2, and their solutions can be computed using Radon matrix representation (see e.g., Ch2, Theorem 4.3, [26]) after a transform on their initial or terminal conditions (recall P is forward equation with initial condition, while π is backward equation with terminal condition). Given the solution of π_1, π_2, the matrix and their norms in Proposition 2.1, 2.2 can be evaluated. In our example here, $\check{A}$ is invertible and $\sqrt{T}\|\phi\|_T\|\bar{\mathbf{A}} - \bar{\mathbf{G}}\|_T \approx 0.1197 < 1$ thus (30) admits a unique solution (K, Φ). Note that the matrix Riccati equation K is of size 4×4, and its solution can be computed using the Runge-Kutta method [14]. Given K, Φ, the state $\bar{z}$ and observation equation can be simulated using the Euler approximation scheme of [9]. The MFG strategies can be computed and we simulate the individual agent states with $N = 500$. The realized state-average for agents is also computed. The simulation results are reported by the following figures.

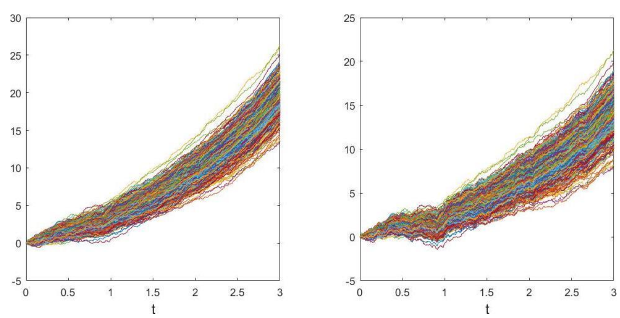

Figure 1: *Trajectories of the type-1 agents' states when N=500*

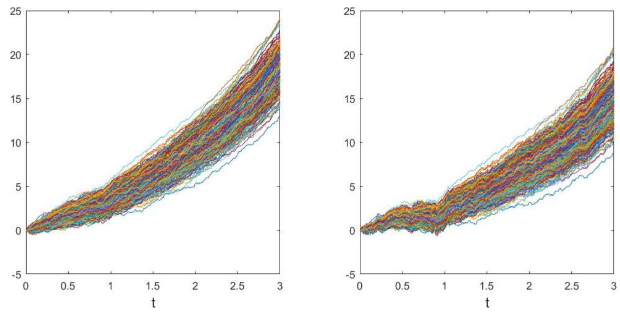

Figure 2: *Trajectories of the type-2 agents' states when N=500*

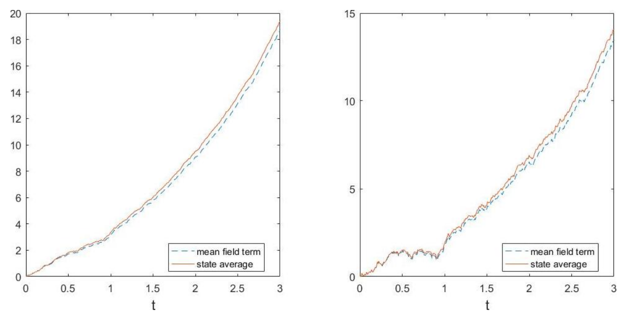

Figure 3: *Trajectories of the type-1 agents state average and the mean field term*

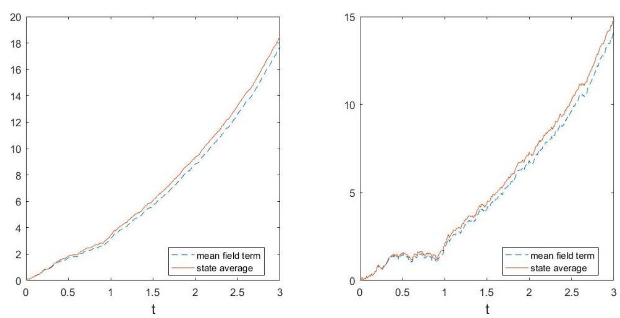

Figure 4: *Trajectories of the type-2 agents state average and the mean field term*

5 Conclusion and Future Work

We discuss mean-field games (MFGs) where each individual agent can only access partial observation on his own state. Moreover, the states of all agents are driven by a underlying common noise. The decentralized strategies are derived with the help of Kalman filtering together with consistency condition. It is notable the consistency condition is connected to the wellposedness of a FBSDE driven by the common noise. Our work suggests some future research topics. For example, the related MFGs for classical mean-variance problem but within the partial observation framework; the related MFGs where common-noise process is not observable to our agents.

Acknowledgement

A. Bensoussan acknowledges the support from grant DMS-1612880 of National Science Foundation and Research Grants Council of the Hong Kong Special Administrative Region (City U 11303316). J. Huang acknowledges the financial support partly by RGC Grant 502412, 15300514, G-YL04. The authors are grateful to Dr. Shujun Wang for helpful discussions and comments on the first version of the manuscript.

References

[1] Bardi, M. (2012). Explicit solutions of some linear-quadratic mean field games. *Networks and Heterogeneous Media*, **7**, 243-261.

[2] Bensoussan, A. (1992). Stochastic control of partially observable systems. Cambridge University Press, Cambridge.

[3] Bensoussan, A., Frehse, J., & Yam, S. (2013). Mean field games and mean field type control theory. Springer, New York.

[4] Bensoussan, A., Sung, K., Yam, S., & Yung, S. (2016). Linear-quadratic mean-field games. *J. Optimiz. Theory App.*, **169**, 496-529.

[5] Carmona, R., & Delarue, F. (2013). Probabilistic analysis of mean-field games. *SIAM J. Control Optim.*, **51**, 2705-2734.

[6] Carmona, R., & Delarue, F. (2018). Probabilistic Theory of Mean Field Games with Applications I-II. Springer Nature.

[7] Carmona, R., Fouque, J., & Sun L. (2015). Mean field games and systemic risk: a toy model. *Commun. Math. Sci.*, **13**, 911-933.

[8] Dogbé, C. (2010). Modeling crowd dynamics by the mean-field limit approach. *Math. Comput. Modelling.*, **52**, 1506-1520.

[9] E. Kloeden, P., & Platen. (1992). Numerical solution of stochastic differential equations. Springer-Verlag.

[10] Fleming, W., & Rishel, W. (1992). Deterministic and stochastic control of partially observable systems. Cambridge University Press, Cambridge.

[11] Erickson G. (1995). Differential game methods of advertising competition. *Europ. J. Oper. Res.*, **83**, 431-438.

[12] Gueant, O., Lasry, J. M., & Lions, P. L. (2011). Mean field games and applications. *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Mathematics*, 2003, 205-266.

[13] Haurie, A., & Marcotte, P. (1985). On the relationship between Nash-Cournot and Wardrop equilibria. *Networks*, **15**, 295-308.

[14] Hu, G.-D. (2007). Symplectic Runge-Kutta methods for the linear quadratic regulator problem. *J. Math. Anal.,* **1**, 293-304.

[15] Huang, J., Hu, Y., & Nie, T. (2018). Linear-quadratic-Gaussian mixed mean-field games with heterogeneous input constraints. *SIAM J. Control Optim.*, **56**, 2835-2877.

[16] Huang, J., & Wang, S. (2016). Dynamic optimization of large-population systems with partial information. *J. Optimiz. Theory App.*, **168**, 231-245.

[17] Huang, M. (2010). Large population LQG games involving a major player: The Nash certainty equivalence principle. *SIAM J. Control Optim.*, **48**, 3318-3353.

[18] Huang, M., Caines, P. E., & Malhamé, R. P. (2004). Uplink power adjustment in wireless communication systems: A stochastic control analysis. *IEEE Trans. Automat. Control*, **49**, 1693-1708.

[19] Huang, M., Caines, P. E., & Malhamé, R. P. (2006). Distributed multi-agent decision-making with partial observations: asymptotic Nash equilibria. *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan.

[20] Huang, M., Caines, P. E., & Malhamé, R. P. (2007). Large-population cost-coupled LQG problems with non-uniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automat. Control*, **52**, 1560-1571.

[21] Huang, M., Malhamé, R. P., & Caines, P. E. (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, **6**, 221-251.

[22] Kallianpur, G. (1980). Stochastic filtering theory. Springer-Verlag.

[23] Kizilkale, A. C., & Malhamé, R. P. (2014). Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads. In Proceedings of the 19th IFAC World Congress, pages 1867-1872, Cape Town, South Africa.

[24] Lambson V. E. (1984). Self-enforcing collusion in large dynamic markets. *J. Econ. Theory*, **34**, 282-291.

[25] Lasry, J. M., & Lions, P. L. (2007). Mean field games. *Japan J. Math.*, **2**, 229-260.

[26] Ma, J., & Yong, J. (1999). Forward-backward stochastic differential equations and their applications. Springer-Verlag, Berlin Heidelberg.

[27] Ma, Z., Callaway, D., & Hiskens, I. (2013). Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Trans. Automat. Control*, **21**, 67-78.

[28] Nguyen, S. L., & Huang, M. (2012). Linear-quadratic-Gaussian mixed games with continuum-parameterized minor players. *SIAM J. Control Optim.*, **50**, 2907-2937.

[29] Nourian, M., Caines, P. E., Malhamé, R. P., & Huang, M. (2013). Nash, social and centralized solutions to consensus problems via mean field control theory. *IEEE Trans. Automat. Control*, **58**, 639-653.

[30] Øksendal, B. (2003). Stochastic differential equations. Springer.

[31] Şen, N., & Caines, P. E. (2016). Nonlinear filtering theory for McKean–Vlasov type stochastic differential equations. *SIAM J. Control Optim.*, **54**, 153-174.

[32] Tembine, H., Zhu, Q., & Basar, T. (2014). Risk-sensitive mean-field stochastic differential games. *IEEE Trans. Automat. Control*, **59**, 835-850.

[33] Wang G., & Wu Z. (2008) KalmanCBucy filtering equations of forward and backward stochastic systems and applications to recursive optimal control problems. *J. Math. Anal. Appl.*, **342**, 1280-1296.

[34] Wang G., Wu Z., & Xiong J. (2015) A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information. *IEEE Trans. Automat. Control*, **60**, 2904-2916.

[35] Weintraub, Y., Benkard, L., & Van Roy, D. (2008). Markov perfect industry dynamics with many firms. *Econometrica*, **76**, 1375-1411.

[36] Wonham, W. M. (1968). On the separation theorem of stochastic control. *SIAM J. Control Optim.*, **6**, 312-326.

[37] Yin, H., Mehta, P., Meyn, S., & Shanbhag, U. (2012). Synchronization of coupled oscillators is a game. *IEEE Trans. Automat. Control*, **57**, 920-935.

[38] Yong, J., & Zhou, X. Y. (1999). Stochastic controls: Hamiltonian systems and HJB equations. Springer-Verlag, New York.