Linear quadratic mean-field game-team analysis: a mixed coalition approach

Jianhui Huang [a], Zhenghong Qiu [a], Shujun Wang [b], Zhen Wu [c]

[a] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong*

[b] *School of Management, Shandong University, Jinan, Shandong 250100, China*

[c] *School of Mathematics, Shandong University, Jinan, Shandong 250100, China*

Abstract

Mean-field theory has been extensively explored in decision analysis of large-scale (LS) systems but traditionally in "pure" cooperative or competitive settings. This leads to the so-called mean-field game (MG) or mean-field team (MT). This paper introduces a new class of LS systems with cooperative inner layer and competitive outer layer, so a "mixed" mean-field analysis is proposed for distributed game-team strategy. A novel asymptotic mixed-equilibrium-optima is also proposed and verified.

Key words: Mean-field game, Mean-field team, Game-team mixed strategy, Mixed-equilibrium-optima

1 Introduction

This paper is inspired by (dynamic) decision of large-scale (LS) systems for which the most salient feature, in linear-quadratic (LQ) setting, is existence of sufficiently many negligible agents interacted among their states or objectives via empirical state-average or control-average. Weakly-coupled LS systems have been found broad applications across economics, finance, biology and engineering. Interested readers may refer [10], [12], etc. For LS systems with highly complex interactions, mean-field theory provides an effective scheme to study its decision *asymptotically* as population size $N \to \infty$.

In principle, decisions of LS systems can be classified into *non-cooperative* game or *cooperative* team, relying on coalition structure formalized by involved agents in underlying population. Accordingly, mean-field game (MG) or mean-field team (MT) arise naturally when studying LS asymptotically. Along game-line (MG), all agents, $\{\mathcal{A}_i\}_{i=1}^N$, are non-cooperative with competitive objectives $\{\mathcal{J}_i\}_{i=1}^N$ thus some Nash type equilibrium should be analyzed. There have accumulated vast literature on MG study, see e.g. [2], [3], [5], [6], [8], [9], [10], [12], [14], [15], [21]. Parallel to game, MT forms another appealing line along which $\{\mathcal{A}_i\}_{i=1}^N$ are cooperative to same team (social) objective $\mathcal{J}_{soc}^{(N)} = \sum \mathcal{J}_i$. MT has also been extensively explored from a wide range of perspectives, e.g., [11], [16], [17], [22]. All aforementioned (MG, MT) attempts, albeit well explored, have only been built on a "pure" basis whereby all LS agents are purely cooperative or competitive. However, in reality, LS system often displays some "mixed" behaviors in its agents' organization with combined game and team. To be explained later, pure basis seems over idealized to fit reality, so some expanded analysis on more realistic mixed basis is strongly suggested.

On one hand, from a "practical" viewpoint, various real environments suggest LS to display some "mix" decision patterns. Indeed, mixed LS come often from economics, engineering, or management, etc., when two or more competitive networks co-exist, e.g., duopoly market with distributed franchisees; adversarial networks with operation knots (e.g., [4], [7]). Specifically, a mixed basis is well-posed whenever classical *two-person game* or general multiple-person-game are framed with *decentralized* decisions on their distributed *sub-units*. By this, original two (multiple) persons or entities still play game, while

all sub-units within each entity formalize two (multiple) teams. So a "mixed" basis arises with tangled game and team. On the other hand, from a "mathematical" viewpoint: conceptually, any LS decision can be captured by its underlying **coalition matrix** (to be introduced in Section 2.3), an $N \times N$ matrix satisfying given block structure properties. Exactly, MG and MT are captured by its two extreme cases: *identical* and *one* matrix, respectively ($\mathbf{C}_1$ and $\mathbf{C}_2$). However, besides them, there exists other types coalition matrix which should feature other meaningful LS decision patterns. More general mixed structure can be identified using other variants of coalition matrix.

Motivated by above discussions, this paper formulates and analyzes a certain type of *mixed* game-team problem involving two levels of interactions: cooperative inner layer for weakly-coupled agents and competitive outer layer between sub-systems. This yields an adversarial LS networks with complex interactions inside and outside. Accordingly, our main contributions are as follows:

(1) A new *Mix* game-team is introduced and studied, which offers a more general LS setting including those classical ones on mean-field game and team.

(2) A coalition matrix representation is formulated to characterize competitive and cooperative structure among weakly-coupled decision makers. Specifically, classical MG and MT correspond to two extreme cases of such coalition structures.

(3) A bilateral person-by-person optimality condition and bilateral variation decomposition are proposed for the first time to design distributed strategy. Wellposedness of associated consistency condition (CC) is also discussed under mild conditions.

(4) A new notion of asymptotic mixed-equilibrium-optima is proposed to mixed LS study, and verified to the derived distributed mixed-game-team strategy asymptotically.

The remainder of this paper is organized as follows. Section 2 gives some preliminary notations and a general formulation of mixed game-team via coalition matrix. Section 3 makes some synthesis analysis on variation decomposition. Section 4 aims to some distributed mixed strategy design via a bilateral auxiliary problem and resulting CC system. Asymptotic mixed-equilibrium-optima is studied in Section 5. Section 6 concludes this work.

2 Problem formulation

2.1 Preliminary

Throughout this paper, $\prod$ denotes the Cartesian product. $\mathbb{R}^{n \times m}$ and $\mathbb{S}^n$ denote the sets of all $(n \times m)$ real matrices and all $(n \times n)$ symmetric matrices respectively. For a vector or matrix A, A^T denotes transpose of A. We denote $\|\cdot\|$ as the standard Euclidean norm and $\langle \cdot, \cdot \rangle$ as the standard Euclidean inner product. For a vector v and a symmetric matrix S, $\|v\|_S^2 := \langle Sv, v \rangle = v^T Sv$. $S > 0 \ (\geq 0)$ means that S is positive (semi-positive) definite, and $S \gg 0$ means that, $S - \varepsilon I \geq 0$, for some $\varepsilon > 0$. Moreover, we suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is an complete probability space, and $W = (W_1, \cdots, W_N)$ is a N-dimensional standard Brownian motion de-

fined on it. $\{\mathcal{F}(t)\}_{t \geq 0}$ is the natural filtration generated by $W_i(s), 0 \leq s \leq t, 1 \leq i \leq N$, augmented by all the $\mathbb{P}$−null sets of $\mathcal{F}$, and $\mathbb{F} := \{\mathcal{F}(t)\}_{t \geq 0}$. $\{\mathcal{F}_i(t)\}_{t \geq 0}$ is the natural filtration generated by $W_i(s), 0 \leq s \leq t$, augmented by all the $\mathbb{P}$−null sets of $\mathcal{F}_i$, and $\mathbb{F}_i := \{\mathcal{F}_i(t)\}_{t \geq 0}, 1 \leq i \leq N$. For a generic Euclidean space $\mathcal{E}$ and filtration $\mathbb{K}$, we introduce the following spaces: $L^\infty(0, T; \mathcal{E}) = \{x : [0, T] \to \mathcal{E} \mid x \text{ is bounded and deterministic}\}$, $C([0, T]; \mathcal{E}) = \{x : [0, T] \to \mathcal{E} \mid x \text{ is continuous}\}$, $L^2_{\mathbb{K}}(0, T; \mathcal{E}) = \{x : [0, T] \times \Omega \to \mathcal{E} \mid x \text{ is } \mathbb{K}\text{-progressively}$

measurable, $\|x\|_{L^2} := \left(\mathbb{E} \int_0^T \|x(t)\|^2 dt \right)^{\frac{1}{2}} < \infty \}$.

2.2 Large-scale system with weak-coupling

In this paper, on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ we consider an LS system with N agents, denoted by $\{\mathcal{A}_i\}_{i \in \mathcal{I}}$, and $\mathcal{I} = \{1, \cdots, N\}$ denotes the index set of the agents. The aggregation of all agents is denoted by $\mathcal{A} := \{\mathcal{A}_i\}_{i \in \mathcal{I}}$. The state process of the ith agent $\mathcal{A}_i$ is modeled by a controlled linear SDE on finite time horizon $[0, T]$:

$$\mathcal{S}_i : \quad dx_i(t) = (A_i x_i(t) + B_i u_i(t) + F_i x^{(N)}(t)) dt \\ + \sigma_i dW_i(t), \quad x_i(0) = \xi_i, \quad (1)$$

and for the sake of notation simplicity, we denote $\mathbf{u} := (u_1, \cdots, u_N)$ and $\mathbf{x} := (x_1, \cdots, x_N)$. Then $\mathcal{S} := \{\mathcal{S}_i\}_{i \in \mathcal{I}}$ form a weakly coupled LS system, since each agent $\mathcal{A}_i$ is coupled with the others via $x^{(N)}$, and $\mathcal{A}_i$ could only provide weak influence to the others at $O\left(\frac{1}{N}\right)$ order. To evaluate each control law u_i, we introduce the following individual cost functional:

$$\mathcal{J}_i(\xi_0; \mathbf{u}) = \mathcal{J}_i(\xi_0; u_i, \mathbf{u}_{-i}) \\ = \frac{1}{2} \mathbb{E} \left\{ \int_0^T \left[\|x_i(t) - \Gamma_i x^{(N)}(t)\|_{Q_i}^2 + \|u_i(t)\|_{R_i}^2 \right] dt \right\}, \quad (2)$$

where $\mathbf{u}_{-i} = (u_1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_N)$. LS system (1)-(2) is in linear quadratic (LQ) setting that is well documented in mean-field literature (e.g. [2], [10], [22], etc). Henceforward, for notation simplicity, we will drop t if no confusion. For $i \in \mathcal{I}$, the decentralized admissible strategy set for the ith agent is given by $\mathcal{U}_i = \left\{ u_i | u_i \in L^2_{\mathbb{F}_i}(0, T; \mathbb{R}^m) \right\}$; the centralized admissible strategy set for the ith agent is given by $\mathcal{U}_i^c = \left\{ u_i | u_i \in L^2_{\mathbb{F}}(0, T; \mathbb{R}^m) \right\}$.

For LS (1), centralized decision on full information $\mathbb{F}$ becomes inefficient due to coupling. Instead, we prefer decentralized one on distributed $\mathbb{F}_i$ via mean-field scheme.

Along this line, recall the well-studied MG and MT:

$$\begin{cases} \textbf{MG:} \inf_{u_i(\cdot)\in\mathcal{U}_i} \mathcal{J}_i(u_i;\mathbf{u}_{-i}); \\ \textbf{MT:} \inf_{\mathbf{u}(\cdot)\in\mathcal{U}^N} \mathcal{J}_{soc}^{(N)}(\mathbf{u}) := \sum_{i=1}^N \mathcal{J}_i(\mathbf{u}), \quad \mathcal{U}^N := \prod_{i=1}^N \mathcal{U}_i. \end{cases} \tag{3}$$

MG, MT are both *decentralized* on $L^2_{\mathbb{F}_i}(0,T;\mathbb{R}^m)$ or $\prod_{i=1}^N L^2_{\mathbb{F}_i}(0,T;\mathbb{R}^m)$. So, they differ from mean-field-type control ([1], [18], etc.), still centralized on $\mathbb{F}$. But, MT aims to unified *team cost* $\mathcal{J}_{soc}^{(N)}$ by *cooperative* team-decision $\mathbf{u} = (u_1, \cdots, u_N)$, whereas MG to *competitive* $\mathcal{J}_i$ by individual u_i. Our inspiration is to recast MG, MT from a *coalition matrix* insight.

2.3 A coalition matrix representation

First, we introduce *nominal cost vector* $\mathbf{J} := (\mathcal{J}_1, \cdots, \mathcal{J}_N)^T$ formalized by all $\{\mathcal{J}_i\}_{i=1}^N$. For fixed $\mathcal{A}_i$, we call $\mathcal{J}_i$ its *principal* cost, while $\{\mathcal{J}_k\}_{k\neq i}$ *marginal* costs. Next, set *effective cost vector* $\mathbf{K} := (\mathcal{K}_1, \cdots, \mathcal{K}_N)^T$ with $\mathcal{K}_i$ the effective cost really "targeted" by $\mathcal{A}_i$. In general, $\mathcal{K}_i \neq \mathcal{J}_i$. By matrix operation, $\mathbf{K} = \mathbf{C}\cdot\mathbf{J}$ for some *coalition matrix* $\mathbf{C} = [c_{i,j}]_{1\leq i,j\leq N}$ for which some typical forms are listed below as $\mathbf{C}_1$-$\mathbf{C}_5$:

Fig. 1. $\mathbf{C}-$form coalition matrices.

Notably, all coalition $\mathbf{C}-$forms above share a common *block* feature: the diagonal "covered" by $1 \leq n_c \leq N$ square **one** matrix $\mathbb{1}_s$ (i.e., all entries are 1) of possible varying sizes (e.g., $s_1, \cdots, s_{n_c}$) as exhibited by the general $\mathbf{C}_5$. Noting $\mathbf{C}_5$ nests $\mathbf{C}_4$ if $n_c = 2$, $s_1 = N_1$, $s_2 = N_2 = N - N_1$ and properly setting its off-diagonal blocks; $\mathbf{C}_4$ further nests $\mathbf{C}_3$ by $\alpha = \beta = -1$ and $\mathbf{C}_2$ by $\alpha = \beta = 1$ ($n_c = 1$, $s_1 = N$); also, $\mathbf{C}_5$ nests $\mathbf{C}_1$ by $n_c = N$, $s_i \equiv 1$. The meaning of "coalition" is illustrated by classifications below:

$$\begin{cases} \mathbf{C} = \mathbf{C}_1 = \mathbf{I}_N \Longrightarrow (\mathbf{K} = \mathbf{J}) \Leftrightarrow (\mathcal{K}_i = \mathcal{J}_i) \\ \quad \Longrightarrow \mathbf{MG} \Leftrightarrow \text{(non-cooperative)} \{\mathcal{A}_i\}_{i=1}^N; \\ \mathbf{C} = \mathbf{C}_2 = \mathbb{1}_N \Longrightarrow (\mathcal{K}_i \equiv \mathcal{J}_{soc}^{(N)} = \sum \mathcal{J}_k) \\ \quad \Longrightarrow \mathbf{MT} \Leftrightarrow \text{(cooperative)} \{\mathcal{A}_i\}_{i=1}^N. \end{cases} \tag{4}$$

So, different $\mathbf{C}-$forms induce different *coalitions* among $\{\mathcal{A}_i\}_{i=1}^N$. Here, we do not distinguish more subtle *exogenous* or *endogenous* coalition formation (e.g., [13], [20], [23]) as they are economics biased. We observe $\mathbf{C}_1, \mathbf{C}_2$ as exactly two extreme $\mathbf{C}-$forms which are leading to "pure" MG or MT. But various other $\mathbf{C}-$forms $1 < n_c < N$ do exist such as $\mathbf{C}_3 - \mathbf{C}_5$ which should connect to other meaningful structures beyond "pure"

game/team. In fact, by setting $\mathcal{I}_1 = \{\theta_1, \cdots, \theta_{N_1}\}, \mathcal{I}_2 = \{\vartheta_1, \cdots, \vartheta_{N_2}\}$ and $N_1 + N_2 = N$, where θ_i, ϑ_j are all positive indexes for $1 \leq i \leq N_1$, $1 \leq j \leq N_2$, we may proceed more along typical $\mathbf{C}_4$:

$$\begin{aligned} &\mathbf{C} = \mathbf{C}_4 \Longrightarrow \mathcal{K}_i = \\ &\begin{cases} \mathbf{J}_{\text{mix}}^{(1)} := \mathcal{J}_{soc}^{1,(N)} + \alpha\mathcal{J}_{soc}^{2,(N)}, \quad \mathcal{J}_{soc}^{1,(N)} := \sum_{k\in\mathcal{I}_1} \mathcal{J}_k, \\ \mathbf{J}_{\text{mix}}^{(2)} := \mathcal{J}_{soc}^{2,(N)} + \beta\mathcal{J}_{soc}^{1,(N)}, \quad \mathcal{J}_{soc}^{2,(N)} := \sum_{k\in\mathcal{I}_2} \mathcal{J}_k, \end{cases} \end{aligned} \tag{5}$$

where $\alpha,\ \beta \in \mathbb{R}$. For $k = 1, 2$, let $\pi_k^{(N)} = \frac{N_k}{N}$, then $\pi^{(N)} = (\pi_1^{(N)}, \pi_2^{(N)})$ is a probability vector representing the empirical distribution of $\mathcal{I}_1$ and $\mathcal{I}_2$. Thus, the original LS system divides into two sub-systems: LS1 := $\{\mathcal{A}_k\}_{k\in\mathcal{I}_1}$, LS2 := $\{\mathcal{A}_k\}_{k\in\mathcal{I}_2}$. All agents inside LS1/LS2 aim the same cost so formalize two cooperative *internal* MTs. LS1 and LS2 remain competitive outside if $\alpha < 0$, $\beta < 0$; while cooperative outside if $\alpha > 0$, $\beta > 0$.

2.4 Mixed game-team via coalition

Now, we are ready to formally introduce the "mixed" game-team (**Mix**) via coalition matrix representation $\mathbf{C}_4$ as

$$(\textbf{Mix}) \begin{cases} \textbf{LS1:} \inf_{\mathbf{u}_1(\cdot)\in\mathbf{U}^{(1)}} \mathbf{J}_{\text{mix}}^{(1)}(\mathbf{u}_1; \mathbf{u}_2), \\ \quad \mathbf{u}_1 := (\cdots, u_k, \cdots)|_{k\in\mathcal{I}_1}, \ \mathbf{U}^{(1)} := \prod_{k\in\mathcal{I}_1} \mathcal{U}_k; \\ \textbf{LS2:} \inf_{\mathbf{u}_2(\cdot)\in\mathbf{U}^{(2)}} \mathbf{J}_{\text{mix}}^{(2)}(\mathbf{u}_2; \mathbf{u}_1), \\ \quad \mathbf{u}_2 := (\cdots, u_k, \cdots)|_{k\in\mathcal{I}_2}, \ \mathbf{U}^{(2)} := \prod_{k\in\mathcal{I}_2} \mathcal{U}_k. \end{cases} \tag{6}$$

To better illustrate the categories of **LS1** and **LS2**, we need rewrite the dynamics $\mathcal{S}_i$ in (1) as the following homogeneous form:

$$\begin{cases} dx_i(t) = \Big(A_1 x_i(t) + B_1 u_i(t) + F_1 x^{(N)}(t)\Big) dt \\ \qquad\quad + \sigma_1(t) dW_i(t), \quad x_i(0) = \xi_i, \quad i \in \mathcal{I}_1, \\ dx_j(t) = \Big(A_2 x_j(t) + B_2 u_j(t) + F_2 x^{(N)}(t)\Big) dt \\ \qquad\quad + \sigma_2(t) dW_j(t), \quad x_j(0) = \eta_j, \quad j \in \mathcal{I}_2, \end{cases} \tag{7}$$

where $x^{(N)}(\cdot) = \frac{1}{N}\Big(\sum_{i\in\mathcal{I}_1} x_i(\cdot) + \sum_{j\in\mathcal{I}_2} x_j(\cdot)\Big)$. For the sake of notation simplicity, we denote $\mathbf{x}_1 := (x_{\theta_1}, \cdots, x_{\theta_{N_1}})$, $\mathbf{x}_2 := (x_{\vartheta_1}, \cdots, x_{\vartheta_{N_2}})$. we also rewrite the individual cost functionals (2) as

$$
\begin{cases}
\mathcal{J}_i(\mathbf{u}_1(\cdot), \mathbf{u}_2(\cdot)) \\
= \dfrac{1}{2}\mathbb{E}\int_0^T \Big[\|x_i(t) - \Gamma_1 x^{(N)}(t)\|_{Q_1}^2 + \|u_i(t)\|_{R_1}^2\Big]dt, \\
\mathcal{J}_j(\mathbf{u}_1(\cdot), \mathbf{u}_2(\cdot)) \\
= \dfrac{1}{2}\mathbb{E}\int_0^T \Big[\|x_j(t) - \Gamma_2 x^{(N)}(t)\|_{Q_2}^2 + \|u_j(t)\|_{R_2}^2\Big]dt.
\end{cases} \tag{8}
$$

The agents in $\{\mathcal{A}_i\}_{i\in\mathcal{I}_1}$ and $\{\mathcal{A}_j\}_{j\in\mathcal{I}_2}$ are cooperative respectively and their social cost functionals $\mathbf{J}_{\mathrm{mix}}^{(1)}$, $\mathbf{J}_{\mathrm{mix}}^{(2)}$ are given by (5). We impose the following general assumptions, which are commonly used in LQ models, on the coefficients:

(H1) A_k, F_k, $\Gamma_k \in \mathbb{R}^{n\times n}$, $B_k \in \mathbb{R}^{n\times m}$, $\sigma_k \in L^\infty(0, T; \mathbb{R}^n)$, $k = 1, 2$.

(H2) $Q_k \in \mathbb{S}^n$, $R_k \in \mathbb{S}^m$, $Q_k > 0$, $R_k > 0$, $k = 1, 2$.

(H3) $\{\xi_i\}_{i=1}^{N_1}$ are independent identically distributed (i.i.d) with mathematical expectation $\mathbb{E}\xi$; $\{\eta_j\}_{j=1}^{N_2}$ are i.i.d with mathematical expectation $\mathbb{E}\eta$.

(H4) There exists a probability mass vector $\pi = (\pi_1, \pi_2)$ such that $\lim_{N\to+\infty}\pi^{(N)} = \pi$, $\min_{1\le k\le 2}\pi_k > 0$.

Thus, we can propose the following homogeneous mixed game-team problem:

Problem 2.1 *Find a centralized strategy set* $(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)$, *where* $\bar{\mathbf{u}}_1 = (\bar{u}_{\theta_1}, \cdots, \bar{u}_{\theta_{N_1}})$, $\bar{\mathbf{u}}_2 = (\bar{u}_{\vartheta_1}, \cdots, \bar{u}_{\vartheta_{N_2}})$, $\bar{u}_i \in \mathcal{U}_i$, $\bar{u}_j \in \mathcal{U}_j$, $i \in \mathcal{I}_1$, $j \in \mathcal{I}_2$, *such that*

$$
\textbf{(Mix)} \begin{cases}
\textbf{LS1:} \quad \mathbf{J}_{\mathrm{mix}}^{(1)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2) = \inf_{\mathbf{u}_1\in\mathbf{U}^{(1)}}\mathbf{J}_{\mathrm{mix}}^{(1)}(\mathbf{u}_1, \bar{\mathbf{u}}_2), \\
\textbf{LS2:} \quad \mathbf{J}_{\mathrm{mix}}^{(2)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2) = \inf_{\mathbf{u}_2\in\mathbf{U}^{(2)}}\mathbf{J}_{\mathrm{mix}}^{(2)}(\bar{\mathbf{u}}_1, \mathbf{u}_2).
\end{cases} \tag{9}
$$

(Mix) is meaningful, not only in mathematics as verified by coalition arguments above, but also in practice as supported by real motivations (see [4], [7], etc.).

3 Variation synthesis analysis

Initially, we synthesize all response components for $\mathbf{J}_{\mathrm{mix}}^{(1)}$ in (9), from a bilateral viewpoint of **LS1**, **LS2**. In what follows, we only focus on the viewpoint of **LS1**, and the similar argument can be applied to **LS2**.

3.1 Variation decomposition

Let $\bar{\mathbf{u}}_1$, $\bar{\mathbf{u}}_2$ be centralized optimal strategies of the agents in $\mathcal{I}_1$, $\mathcal{I}_2$. We now perturb u_i and keep $\bar{u}_{-i} = (\bar{u}_{\theta_1}, \cdots, \bar{u}_{i-1}, \bar{u}_{i+1}, \cdots, \bar{u}_{\theta_{N_1}})$, $\bar{\mathbf{u}}_2$ fixed. For $k \neq i$, denote the perturbation $\delta u_i = u_i - \bar{u}_i$, $\delta x_i = x_i - \bar{x}_i$, $\delta x_k = x_k - \bar{x}_k$, $\delta x^{(N)} = \frac{1}{N}\sum_{k\in\mathcal{I}_1\cup\mathcal{I}_2}\delta x_k$, and $\delta\mathcal{J}_{\mathrm{soc}}^{1,(N)}$, $\delta\mathcal{J}_{\mathrm{soc}}^{2,(N)}$ are the first variations (Fréchet

differentials) of $\mathcal{J}_{\mathrm{soc}}^{1,(N)}$, $\mathcal{J}_{\mathrm{soc}}^{2,(N)}$ w.r.t. δu_i. Therefore,

$$
\begin{cases}
d\delta x_i(\delta u_i) = \Big[A_1\delta x_i(\delta u_i) + B_1\delta u_i + F_1\delta x^{(N)}(\delta u_i)\Big]dt, \\
\quad \delta x_i(0) = 0, \\
d\delta x_k(\delta u_i) = \Big[A_1\delta x_k(\delta u_i) + F_1\delta x^{(N)}(\delta u_i)\Big]dt, \\
\quad \delta x_k(0) = 0, \ k \in \mathcal{I}_1, \ k \neq i, \\
d\delta x_k(\delta u_i) = \Big[A_2\delta x_k(\delta u_i) + F_2\delta x^{(N)}(\delta u_i)\Big]dt, \\
\quad \delta x_k(0) = 0, \ k \in \mathcal{I}_2.
\end{cases}
$$

Then we derive the following lemma.

Lemma 3.1 $\delta\mathbf{J}_{mix}^{(1)}(\delta u_i)$ *can be represented as*

$$
\delta\mathbf{J}_{mix}^{(1)}(\delta u_i) = \mathbb{E}\int_0^T\Big[\big\langle\bar{\Theta}_1, \delta x_i\big\rangle + \big\langle\Theta_2, \delta u_i\big\rangle + \Big\langle\bar{\Theta}_3, \sum_{k\in\mathcal{I}_1, k\neq i}\delta x_k\Big\rangle
$$
$$
+ \Big\langle\bar{\Theta}_4, \sum_{k\in\mathcal{I}_2}\delta x_k\Big\rangle + \sum_{k\in\mathcal{I}_1, k\neq i}\big\langle\Theta_5^k, \delta x_k\big\rangle + \sum_{k\in\mathcal{I}_2}\big\langle\Theta_6^k, \delta x_k\big\rangle\Big]dt + \varepsilon_1^i, \tag{10}
$$

where

$$
\varepsilon_1^i = \mathbb{E}\int_0^T\Big[\big\langle\Theta_1 - \bar{\Theta}_1, \delta x_i\big\rangle + \Big\langle\Theta_3 - \bar{\Theta}_3, \sum_{k\in\mathcal{I}_1, k\neq i}\delta x_k\Big\rangle
$$
$$
+ \Big\langle\Theta_4 - \bar{\Theta}_4, \sum_{k\in\mathcal{I}_2}\delta x_k\Big\rangle\Big]dt,
$$

$$
\begin{cases}
\Theta_1 = Q_1\bar{x}_i - \Big(Q_1\Gamma_1\bar{x}^{(N)} + \Gamma_1^T Q_1\pi_1^{(N)}\frac{\sum_{k\in\mathcal{I}_1}\bar{x}_k}{N_1} - \pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1\bar{x}^{(N)}\Big) \\
\qquad - \alpha\Big(\Gamma_2^T Q_2\pi_2^{(N)}\frac{\sum_{k\in\mathcal{I}_2}\bar{x}_k}{N_2} - \pi_2^{(N)}\Gamma_2^T Q_2\Gamma_2\bar{x}^{(N)}\Big), \\
\Theta_3 = -\Big(Q_1\Gamma_1\bar{x}^{(N)} + \Gamma_1^T Q_1\pi_1^{(N)}\frac{\sum_{k\in\mathcal{I}_1}\bar{x}_k}{N_1} - \pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1\bar{x}^{(N)}\Big) \\
\qquad - \alpha\Big(\Gamma_2^T Q_2\pi_2^{(N)}\frac{\sum_{k\in\mathcal{I}_2}\bar{x}_k}{N_2} - \pi_2^{(N)}\Gamma_2^T Q_2\Gamma_2\bar{x}^{(N)}\Big), \\
\Theta_4 = -\alpha\Big(Q_2\Gamma_2\bar{x}^{(N)} + \Gamma_2^T Q_2\pi_2^{(N)}\frac{\sum_{k\in\mathcal{I}_2}\bar{x}_k}{N_2} - \pi_2^{(N)}\Gamma_2^T Q_2\Gamma_2\bar{x}^{(N)}\Big) \\
\qquad - \Big(\Gamma_1^T Q_1\pi_1^{(N)}\frac{\sum_{k\in\mathcal{I}_1}\bar{x}_k}{N_1} - \pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1\bar{x}^{(N)}\Big), \\
\Theta_2 = R_1, \quad \Theta_5^k = Q_1\bar{x}_k, \quad \Theta_6^k = \alpha Q_2\bar{x}_k, \\
\bar{\Theta}_1 = Q_1\bar{x}_i - \big(Q_1\Gamma_1\bar{\mathbf{m}} + \Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1 - \pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\big) \\
\qquad - \alpha\big(\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2 - \pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\big), \\
\bar{\Theta}_3 = -\big(Q_1\Gamma_1\bar{\mathbf{m}} + \Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1 - \pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\big) \\
\qquad - \alpha\big(\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2 - \pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\big), \\
\bar{\Theta}_4 = -\alpha\big(Q_2\Gamma_2\bar{\mathbf{m}} + \Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2 - \pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\big) \\
\qquad - \big(\Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1 - \pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\big).
\end{cases}
$$

Here, $\bar{\mathbf{m}}_1$, $\bar{\mathbf{m}}_2$ *and* $\bar{\mathbf{m}} = \pi_1\bar{\mathbf{m}}_1 + \pi_2\bar{\mathbf{m}}_2$ *are mean-field (MF) approximations of* $\frac{1}{N_1}\sum_{k\in\mathcal{I}_1}\bar{x}_k$, $\frac{1}{N_2}\sum_{k\in\mathcal{I}_2}\bar{x}_k$ *and* $\bar{x}^{(N)}$, *respectively.*

3.2 Bilateral duality

Lemma 3.2 $\delta\mathbf{J}_{mix}^{(1)}(\delta u_i)$ *can further be represented as*

$$\delta\mathbf{J}_{mix}^{(1)}(\delta u_i) = \mathbb{E}\int_0^T\Big[\Big\langle\bar{\Theta}_1+\pi_1 F_1^T\mathbb{E}p_k^{(1)*}+\pi_2 F_2^T\mathbb{E}p_k^{(2)*}$$
$$+\pi_1 F_1^T p^{(1)*}+\pi_2 F_2^T p^{(2)*},\delta x_i\Big\rangle+\langle\Theta_2,\delta u_i\rangle\Big]dt \qquad (11)$$
$$+\varepsilon_1^i+\varepsilon_2^i+\varepsilon_3^i,$$

where

$$\varepsilon_2^i=\mathbb{E}\int_0^T\Big[\Big\langle\bar{\Theta}_3,\sum_{k\in\mathcal{I}_1,k\neq i}\delta x_k-x^{(1),**}\Big\rangle+\Big\langle\bar{\Theta}_4,\sum_{k\in\mathcal{I}_2}\delta x_k-x^{(2),**}\Big\rangle$$
$$+\frac{1}{N_1}\sum_{k\in\mathcal{I}_1,k\neq i}\Big\langle\Theta_5^k,N_1\delta x_k-x_k^{(1),*}\Big\rangle$$
$$+\frac{1}{N_2}\sum_{k\in\mathcal{I}_2}\Big\langle\Theta_6^k,N_2\delta x_k-x_k^{(2),*}\Big\rangle\Big]dt,$$

$$\varepsilon_3^i=\mathbb{E}\int_0^T\Big\langle\pi_1 F_1^T\Big(\frac{\sum_{k\in\mathcal{I}_1,k\neq i}p_k^{(1)}}{N_1}-\mathbb{E}p_k^{(1)*}\Big)$$
$$+\pi_2 F_2^T\Big(\frac{\sum_{k\in\mathcal{I}_2}p_k^{(2)}}{N_2}-\mathbb{E}p_k^{(2)*}\Big)+\pi_1 F_1^T\Big(p^{(1)}-p^{(1)*}\Big)$$
$$+\pi_2 F_2^T\Big(p^{(2)}-p^{(2)*}\Big),\delta x_i\Big\rangle dt,$$

and

$$\begin{cases}dx_k^{(1),*}=\Big[A_1 x_k^{(1),*}+\pi_1 F_1(\delta x_i+x^{(1),**}+x^{(2),**})\Big]dt,\\dx_k^{(2),*}=\Big[A_2 x_k^{(2),*}+\pi_2 F_2(\delta x_i+x^{(1),**}+x^{(2),**})\Big]dt,\\dx^{(1),**}=\Big[A_1 x^{(1),**}+\pi_1 F_1(x^{(1),**}+x^{(2),**}+\delta x_i)\Big]dt,\\dx^{(2),**}=\Big[A_2 x^{(2),**}+\pi_2 F_2(x^{(1),**}+x^{(2),**}+\delta x_i)\Big]dt,\\x_k^{(1),*}(0)=0,\ x_k^{(2),*}(0)=0,\ x^{(1),**}(0)=0,\ x^{(2),**}(0)=0,\end{cases}$$

$$\begin{cases}dp_k^{(1)}=-\Big(\Theta_5^k+A_1^T p_k^{(1)}\Big)dt+q_k^{(1)}dW_k+\sum_{k'\neq k}q_{kk'}^{(1)}dW_{k'},\\dp_k^{(2)}=-\Big(\Theta_6^k+A_2^T p_k^{(2)}\Big)dt+q_k^{(2)}dW_k+\sum_{k'\neq k}q_{kk'}^{(2)}dW_{k'},\\dp^{(1)}=-\Big(\bar{\Theta}_3+\pi_1\frac{\sum_{k\in\mathcal{I}_1,k\neq i}F_1^T p_k^{(1)}}{N_1}+\pi_2\frac{\sum_{k\in\mathcal{I}_2}F_2^T p_k^{(2)}}{N_2}\\\qquad+A_1^T p^{(1)}+\pi_1 F_1^T p^{(1)}+\pi_2 F_2^T p^{(2)}\Big)dt+\sum q_{k'}^{(1)}dW_{k'},\\dp^{(2)}=-\Big(\bar{\Theta}_4+\pi_1\frac{\sum_{k\in\mathcal{I}_1,k\neq i}F_1^T p_k^{(1)}}{N_1}+\pi_2\frac{\sum_{k\in\mathcal{I}_2}F_2^T p_k^{(2)}}{N_2}\\\qquad+\pi_1 F_1^T p^{(1)}+\pi_2 F_2^T p^{(2)}+A_2^T p^{(2)}\Big)dt+\sum q_{k'}^{(2)}dW_{k'},\\p_k^{(1)}(T)=0,\ p_k^{(2)}(T)=0,\ p^{(1)}(T)=0,\ p^{(2)}(T)=0,\end{cases}$$

$$\begin{cases}\widehat{\mathcal{P}}_k^{(2)*}:dp_k^{(1)*}=-\Big(\Theta_5^k+A_1^T p_k^{(1)*}\Big)dt+q_k^{(1)*}dW_k,\\\widehat{\mathcal{P}}_k^{(1)*}:dp_k^{(2)*}=-\Big(\Theta_6^k+A_2^T p_k^{(2)*}\Big)dt+q_k^{(2)*}dW_k,\\\widehat{\mathcal{P}}^{(2)*}:dp^{(1)*}=-\Big[\bar{\Theta}_3+\pi_1 F_1^T\mathbb{E}p_k^{(1)*}+\pi_2 F_2^T\mathbb{E}p_k^{(2)*}\\\qquad+\Big(A_1^T+\pi_1 F_1^T\Big)p^{(1)*}+\pi_2 F_2^T p^{(2)*}\Big]dt,\\\widehat{\mathcal{P}}^{(1)*}:dp^{(2)*}=-\Big[\bar{\Theta}_4+\pi_1 F_1^T\mathbb{E}p_k^{(1)*}+\pi_2 F_2^T\mathbb{E}p_k^{(2)*}\\\qquad+\pi_1 F_1^T p^{(1)*}+\Big(\pi_2 F_2^T+A_2^T\Big)p^{(2)*}\Big]dt,\\p_k^{(1)*}(T)=0,\ p_k^{(2)*}(T)=0,\ p^{(1)*}(T)=0,\ p^{(2)*}(T)=0.\end{cases}$$

An asymptotic "Fréchet response" holds for $\delta J_i^{(1)}=$

$\lim\delta\mathbf{J}_{mix}^{(1)}(\delta u_i)$:

LS1: $\delta J_i^{(1)}=\langle\Theta^\dagger(\bar{x}_i,\mathbf{m},\mathbf{P}^*),\delta x_i\rangle+\langle\Theta^{\dagger\dagger}(\bar{u}_i),\delta u_i\rangle$
$$(12)$$
with $\Theta^{\dagger\dagger}=R_1\bar{u}_i$ and $\Theta^\dagger$ is revised on $\big(\bar{\Theta}_1,\bar{\Theta}_3,\bar{\Theta}_4,\Theta_5^k,\Theta_6^k\big)$.
Here, $\mathbf{m}:=(\bar{\mathbf{m}},\bar{\mathbf{m}}_1,\bar{\mathbf{m}}_2)$, $\widehat{\mathbf{P}}^*:=(\widehat{\mathcal{P}}_k^{(1)*},\widehat{\mathcal{P}}^{(1)*};\widehat{\mathcal{P}}_k^{(2)*},\widehat{\mathcal{P}}^{(2)*})$.

4 Distributed design

4.1 Auxiliary control

Motivated by (11), we introduce the following auxiliary problem for $i\in\mathcal{I}_1$:

Problem 4.1 *Minimize $J_i(u_i)$ over $u_i\in\mathcal{U}_i$ where*

$$\begin{cases}dx_i=(A_1 x_i+B_1 u_i+F_1\bar{\mathbf{m}})dt+\sigma_1 dW_i,\ x_i(0)=\xi_i,\ i\in\mathcal{I}_1,\\J_i^{(1)}(u_i)=\frac{1}{2}\mathbb{E}\int_0^T[\langle Q_1 x_i,x_i\rangle+2\langle S_1,x_i\rangle+\langle R_1 u_i,u_i\rangle]dt,\\S_1=-\Big(Q_1\Gamma_1\bar{\mathbf{m}}+\Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1-\pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\Big)\\\qquad-\alpha\Big(\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2-\pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\Big)+\pi_1 F_1^T\mathbb{E}p_k^{(1)*}\\\qquad+\pi_2 F_2^T\mathbb{E}p_k^{(2)*}+\pi_1 F_1^T p^{(1)*}+\pi_2 F_2^T p^{(2)*}.\end{cases}$$

The mean-field terms $\bar{\mathbf{m}}$, $\bar{\mathbf{m}}_1$, $\bar{\mathbf{m}}_2$, $p_k^{(1)*}$, $p_k^{(2)*}$ will be determined by the CC system later. Similarly, for $j\in\mathcal{I}_2$, by following the procedure in Section 3, we can also introduce the following auxiliary problem

Problem 4.2 *Minimize $J_j(u_j)$ over $u_j\in\mathcal{U}_j$ where*

$$\begin{cases}dx_j=(A_2 x_j+B_2 u_j+F_2\bar{\mathbf{m}})dt+\sigma_2 dW_j,\ x_j(0)=\eta_j,\ j\in\mathcal{I}_2,\\J_j^{(2)}(u_j)=\frac{1}{2}\mathbb{E}\int_0^T[\langle Q_2 x_j,x_j\rangle+2\langle S_2,x_j\rangle+\langle R_2 u_j,u_j\rangle]dt,\\S_2=-\Big(Q_2\Gamma_2\bar{\mathbf{m}}+\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2-\pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\Big)\\\qquad-\alpha\Big(\Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1-\pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\Big)+\pi_2 F_2^T\mathbb{E}\hat{p}_k^{(2)*}\\\qquad+\pi_1 F_1^T\mathbb{E}\hat{p}_k^{(1)*}+\pi_2 F_2^T\hat{p}^{(2)*}+\pi_1 F_1^T\hat{p}^{(1)*},\end{cases}$$

where the limiting duality on δu_j yields parallel $\widehat{\mathbf{P}}^*:=(\widehat{\mathcal{P}}_k^{(1)*},\widehat{\mathcal{P}}^{(1)*};\widehat{\mathcal{P}}_k^{(2)*},\widehat{\mathcal{P}}^{(2)*})$. Here, it holds that

$$\begin{cases}\widehat{\mathcal{P}}_k^{(2)*}:d\hat{p}_k^{(2)*}=-\Big(\hat{\Theta}_5^k+A_2^T\hat{p}_k^{(2)*}\Big)dt+\hat{q}_k^{(2)*}dW_k,\\\widehat{\mathcal{P}}_k^{(1)*}:d\hat{p}_k^{(1)*}=-\Big(\hat{\Theta}_6^k+A_1^T\hat{p}_k^{(1)*}\Big)dt+\hat{q}_k^{(1)*}dW_k,\\\widehat{\mathcal{P}}^{(2)*}:d\hat{p}^{(2)*}=-\Big[\hat{\bar{\Theta}}_4+\pi_2 F_2^T\mathbb{E}\hat{p}_k^{(2)*}+\pi_1 F_1^T\mathbb{E}\hat{p}_k^{(1)*}\\\qquad+\Big(A_2^T+\pi_2 F_2^T\Big)\hat{p}^{(2)*}+\pi_1 F_1^T\hat{p}^{(1)*}\Big]dt,\\\widehat{\mathcal{P}}^{(1)*}:d\hat{p}^{(1)*}=-\Big[\hat{\bar{\Theta}}_3+\pi_2 F_2^T\mathbb{E}\hat{p}_k^{(2)*}+\pi_1 F_1^T\mathbb{E}\hat{p}_k^{(1)*}\\\qquad+\pi_2 F_2^T\hat{p}^{(2)*}+\Big(\pi_1 F_1^T+A_1^T\Big)\hat{p}^{(1)*}\Big]dt,\\\hat{p}_k^{(2)*}(T)=\hat{p}_k^{(1)*}(T)=\hat{p}^{(2)*}(T)=\hat{p}^{(1)*}(T)=0,\end{cases}$$

and

$$\begin{cases} \hat{\bar{\Theta}}_4 = -\left(Q_2\Gamma_2\bar{\mathbf{m}}+\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2-\pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\right) \\ \qquad -\beta\left(\Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1-\pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\right), \\ \hat{\bar{\Theta}}_3 = -\beta\left(Q_1\Gamma_1\bar{\mathbf{m}}+\Gamma_1^T Q_1\pi_1\bar{\mathbf{m}}_1-\pi_1\Gamma_1^T Q_1\Gamma_1\bar{\mathbf{m}}\right) \\ \qquad -\left(\Gamma_2^T Q_2\pi_2\bar{\mathbf{m}}_2-\pi_2\Gamma_2^T Q_2\Gamma_2\bar{\mathbf{m}}\right), \\ \hat{\Theta}_5^k = Q_2\bar{x}_k, \quad \hat{\Theta}_6^k = \beta Q_1\bar{x}_k. \end{cases}$$

The above analysis constructs a bilateral auxiliary control problem

(Bilateral auxiliary problem)

$$\begin{cases} \mathcal{A}_i : \inf_{u_i(\cdot)\in\mathcal{U}_i} J_i^{(1)}(u_i;\mathbf{m},\mathbf{P}^*), \\ \mathcal{A}_j : \inf_{u_j(\cdot)\in\mathcal{U}_j} J_j^{(2)}(u_j;\mathbf{m},\widehat{\mathbf{P}}^*). \end{cases} \tag{13}$$

Then, upon distributed $\mathcal{U}_i$ or $\mathcal{U}_j$, two generic $\mathcal{A}_i, \mathcal{A}_j$ aim to optimize $J_i^{(1)}$ or $J_j^{(2)}$. Applying standard LQ method, under some mild (say, convexity) condition, above bilateral auxiliary problem can be solved with optimal pair $(\check{x}_i(\check{u}_i), \check{u}_i)$ for $\mathcal{A}_i$, $(\check{x}_j(\check{u}_j), \check{u}_j)$ for $\mathcal{A}_j$, which depend on $(\mathbf{m},\mathbf{P}^*)$ or $(\mathbf{m},\widehat{\mathbf{P}}^*)$ by (13). Note that $\mathbf{m}$ is not specified yet. By stochastic maximum principle, we have the following result for the bilateral auxiliary problem:

Proposition 4.1 *Under (H1)-(H4), Problem 4.1-4.2 are both uniquely solvable, and the optimal auxiliary controls $\bar{u}_i, \bar{u}_j$ are determined by the following Hamiltonian systems*

$$\begin{cases} d\check{x}_i = (A_1\check{x}_i+B_1\check{u}_i+F_1\bar{\mathbf{m}})dt+\sigma_1 dW_i, \quad \check{x}_i(0)=\xi_i, \\ d\check{x}_j = (A_2\check{x}_j+B_2\check{u}_j+F_2\bar{\mathbf{m}})dt+\sigma_2 dW_j, \quad \check{x}_j(0)=\eta_j, \\ dy_i = -\left(A_1^T y_i+Q_1\check{x}_i+S_1\right)+z_i dW_i, \quad y_i(T)=0, \\ dy_j = -\left(A_2^T y_j+Q_2\check{x}_j+S_2\right)+z_j dW_j, \quad y_j(T)=0, \\ R_1\check{u}_i+B_1^T y_i = 0, \quad R_2\check{u}_j+B_2^T y_j = 0. \end{cases}$$

4.2 Consistency condition

This sub-step aims to synthesize *all* generic behaviors $\{\check{x}_i(\check{u}_i(\mathbf{m},\mathbf{P}^*))\}_{i\in\mathcal{I}_1}$ in LS1, $\{\check{x}_j(\check{u}_j(\mathbf{m},\widehat{\mathbf{P}}^*))\}_{j\in\mathcal{I}_2}$ in LS2, to match aggregated $\mathbf{m}$ across LS. Noting LS1 itself is homogenous (although LS is not), optimal auxiliary controls $\{\check{u}_i\}_{i\in\mathcal{I}_1}$ are thus symmetric and so is $\{\check{x}_i(\check{u}_i)\}_{i\in\mathcal{I}_1}$. Parallel holds for $\{\check{x}_j(\check{u}_j)\}_{j\in\mathcal{I}_2}$. Then, applying de Finetti's theorem to realized $\{\check{x}_i(\check{u}_i(\mathbf{m},\mathbf{P}^*))\}_{i\in\mathcal{I}_1}$, $\{\check{x}_j(\check{u}_j(\mathbf{m},\widehat{\mathbf{P}}^*))\}_{j\in\mathcal{I}_2}$, we yield a bilateral fixed-point CC equivalence (noticing that $\mathbf{m} = (\bar{\mathbf{m}}, \bar{\mathbf{m}}_1, \bar{\mathbf{m}}_2)$):

$$\textbf{(CC)}\begin{cases} \text{LS1:} \quad \bar{\mathbf{m}}_1 = \mathbb{E}(\bar{x}_i(\bar{u}_i(\bar{\mathbf{m}}_1,\bar{\mathbf{m}}_2,\mathbf{P}^*))), \\ \qquad \mathbf{P}^* = \mathbf{P}^*(\bar{x}_i(\bar{u}_i(\bar{\mathbf{m}}_1,\bar{\mathbf{m}}_2,\mathbf{P}))); \\ \text{LS2:} \quad \bar{\mathbf{m}}_2 = \mathbb{E}(\bar{x}_j(\bar{u}_j(\bar{\mathbf{m}}_1,\bar{\mathbf{m}}_2,\widehat{\mathbf{P}}^*))), \\ \qquad \widehat{\mathbf{P}}^* = \widehat{\mathbf{P}}^*(\bar{x}_j(\bar{u}_j(\bar{\mathbf{m}}_1,\bar{\mathbf{m}}_2,\widehat{\mathbf{P}}^*))). \end{cases} \tag{14}$$

So $\mathbf{m}$ can be specified. The CC system is given by

$$\begin{cases} d\mathbf{x} = (\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{u}+\mathbf{F}\mathbb{E}\mathbf{x})dt+\boldsymbol{\sigma}_1 dW_i+\boldsymbol{\sigma}_2 dW_j, \\ d\mathbf{y} = \left(\bar{\mathbf{A}}\mathbf{x}+\bar{\mathbf{B}}\mathbf{y}+\widetilde{\mathbf{F}}\mathbb{E}\mathbf{x}+\widetilde{\mathbf{H}}\mathbb{E}\mathbf{y}\right)dt+\mathbf{z}_1 dW_i+\mathbf{z}_2 dW_j, \\ \mathbf{R}\mathbf{u}+\widetilde{\mathbf{B}}^T\mathbf{y} = 0, \quad \mathbf{x}(0)=\mathbf{x}_0, \quad \mathbf{y}(T)=\mathbf{0}, \end{cases} \tag{15}$$

where

$$\mathbf{x} = (\check{x}_i, \check{x}_j), \quad \mathbf{y} = (y_i, p^{(1)*}, p^{(2)*}, p_i^{(1)*}, p_j^{(2)*}, y_j, \hat{p}^{(2)*}, \hat{p}^{(1)*}, \hat{p}_j^{(2)*}, \hat{p}_i^{(1)*}),$$
$$\mathbf{x}_0 = (\xi_i, \eta_j), \quad \mathbf{z}_1 = (z_i, 0, 0, q_i^{(1)*}, 0, 0, 0, 0, 0, \hat{q}_i^{(1)*}),$$
$$\mathbf{z}_2 = (0, 0, 0, 0, q_j^{(2)*}, z_j, 0, 0, \hat{q}_j^{(2)*}, 0), \quad \mathbf{u} = (\check{u}_i, \check{u}_j),$$
$$\mathbf{A} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \mathbf{F} = \begin{pmatrix} \pi_1 F_1 & \pi_2 F_1 \\ \pi_1 F_2 & \pi_2 F_2 \end{pmatrix}, \mathbf{R} = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix},$$

$$\bar{\mathbf{A}} = \begin{pmatrix} -Q_1 & 0 \\ 0 & 0 \\ 0 & 0 \\ -Q_1 & 0 \\ 0 & -\alpha Q_2 \\ 0 & -Q_2 \\ 0 & 0 \\ 0 & 0 \\ 0 & -Q_2 \\ -\beta Q_1 & 0 \end{pmatrix}, \tilde{\mathbf{B}} = \begin{pmatrix} B_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ B_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{\sigma}_1 = \begin{pmatrix} \sigma_1 \\ 0 \end{pmatrix}, \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 \\ \sigma_2 \end{pmatrix},$$

$$\bar{\mathbf{B}} = \begin{pmatrix} -A_1^T & -\pi_1 F_1^T & -\pi_2 F_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -A_1^T-\pi_1 F_1^T & -\pi_2 F_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\pi_1 F_1^T & -A_2^T-\pi_2 F_2^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -A_1^T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -A_2^T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -A_2^T & -\pi_2 F_2^T & -\pi_1 F_1^T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -A_2^T-\pi_2 F_2^T & -\pi_1 F_1^T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\pi_2 F_2^T & -A_1^T-\pi_1 F_1^T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_2^T & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_1^T \end{pmatrix},$$

$$\widetilde{\mathbf{F}} = \begin{pmatrix} \pi_1(Q_1\Gamma_1+\Gamma_1^T Q_1-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(Q_1\Gamma_1+\alpha\Gamma_2^T Q_2-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) \\ \pi_1(Q_1\Gamma_1+\Gamma_1^T Q_1-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(Q_1\Gamma_1+\alpha\Gamma_2^T Q_2-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) \\ \pi_1(\Gamma_1^T Q_1+\alpha Q_2\Gamma_2-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(\alpha Q_2\Gamma_2+\alpha\Gamma_2^T Q_2-\pi_1\Gamma_1^T Q_1\Gamma_1-\alpha\pi_2\Gamma_2^T Q_2\Gamma_2) \\ 0 & 0 \\ 0 & 0 \\ \pi_1(Q_2\Gamma_2+\beta\Gamma_1^T Q_1-\beta\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(Q_2\Gamma_2+\Gamma_2^T Q_2-\beta\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) \\ \pi_1(Q_2\Gamma_2+\beta\Gamma_1^T Q_1-\beta\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(Q_2\Gamma_2+\Gamma_2^T Q_2-\beta\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) \\ \pi_1(\beta Q_1\Gamma_1+\beta\Gamma_1^T Q_1-\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) & \pi_2(\beta Q_1\Gamma_1+\Gamma_1^T Q_2-\pi_1\Gamma_1^T Q_1\Gamma_1-\pi_2\Gamma_2^T Q_2\Gamma_2) \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\widetilde{\mathbf{H}} = \begin{pmatrix} 0 & -\pi_1 F_1^T & -\pi_2 F_2^T & -\pi_1 F_1^T & -\pi_2 F_2^T & 0 & 0 & 0 & 0 & 0 \\ 0 & -\pi_1 F_1^T & -\pi_2 F_2^T & -A_1^T-\pi_1 F_1^T & -\pi_2 F_2^T & 0 & 0 & 0 & 0 & 0 \\ 0 & -\pi_1 F_1^T & -A_2^T-\pi_2 F_2^T & -\pi_1 F_1^T & -\pi_2 F_2^T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\pi_2 F_2^T & -\pi_1 F_1^T & -\pi_2 F_2^T & -\pi_1 F_1^T \\ 0 & 0 & 0 & 0 & 0 & -A_2^T-\pi_2 F_2^T & -\pi_1 F_1^T & -\pi_2 F_2^T & -\pi_1 F_1^T \\ 0 & 0 & 0 & 0 & 0 & -\pi_2 F_2^T & -A_1^T-\pi_1 F_1^T & -\pi_2 F_2^T & -\pi_1 F_1^T \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

and the mean field terms are determined by $\bar{\mathbf{m}}_1 = \mathbb{E}\check{x}_i$, $\bar{\mathbf{m}}_2 = \mathbb{E}\check{x}_j$.

Proposition 4.2 *If Riccati equations*

$$\begin{cases} \dot{\mathbf{P}}+\mathbf{P}(\mathbf{A}+\mathbf{F})-\left(\widetilde{\mathbf{H}}+\bar{\mathbf{B}}\right)\mathbf{P}-\mathbf{P}\mathbf{B}\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\mathbf{P} \\ \qquad -\left(\bar{\mathbf{A}}+\widetilde{\mathbf{F}}\right) = 0, \\ \dot{\mathbf{K}}+\mathbf{K}\left(\mathbf{A}-\mathbf{B}\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\mathbf{P}\right)-\left(\bar{\mathbf{B}}+\mathbf{P}\mathbf{B}\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\right)\mathbf{K} \\ \qquad +\mathbf{K}\mathbf{B}\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\mathbf{K}+\mathbf{P}\mathbf{F}-\widetilde{\mathbf{F}}-\widetilde{\mathbf{H}}\mathbf{P} = 0, \\ \mathbf{P}(T) = 0, \quad \mathbf{K}(T) = 0 \end{cases} \tag{16}$$

admit solutions $\mathbf{P}\in C([0,T];\mathbb{R}^{10n\times 2n}), \mathbf{K}\in C([0,T];\mathbb{R}^{10n\times 2n})$, then Problem 2.1 admits a feedback form decentralized control $\tilde{\mathbf{u}} = \Theta_1\tilde{\mathbf{x}}+\Theta_2(\mathbb{E}\tilde{\mathbf{x}}-\tilde{\mathbf{x}})$, where $\tilde{\mathbf{u}} = (\tilde{u}_i, \tilde{u}_j)$, $\tilde{\mathbf{x}} = (\tilde{x}_i, \tilde{x}_j)$, and

$$\Theta_1 = -\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\mathbf{P}, \quad \Theta_2 = -\mathbf{R}^{-1}\widetilde{\mathbf{B}}^T\mathbf{K}.$$

The realized state $\tilde{\mathbf{x}} = (\tilde{x}_i, \tilde{x}_j)$ satisfies the following dy-

namic

$$\begin{cases} d\tilde{x}_i = \left(A_1\tilde{x}_i + B_1\tilde{u}_i + F_1\tilde{x}^{(N)}\right)dt + \sigma_1 dW_i, & \tilde{x}_i(0) = \xi_i, \\ d\tilde{x}_j = \left(A_2\tilde{x}_j + B_2\tilde{u}_j + F_2\tilde{x}^{(N)}\right)dt + \sigma_2 dW_j, & \tilde{x}_j(0) = \eta_j. \end{cases}$$

For the solvability of (16), we have the following result

Proposition 4.3 *Let* Ψ, Φ *be fundamental matrices* $\Psi = \begin{pmatrix} \mathbf{A+F} & -\mathbf{BR}^{-1}\tilde{\mathbf{B}}^{\scriptscriptstyle T} \\ \tilde{\mathbf{A}}+\tilde{\mathbf{F}} & \tilde{\mathbf{H}}+\tilde{\mathbf{B}} \end{pmatrix}$ *and* $\Phi = \begin{pmatrix} \mathbf{A}-\mathbf{BR}^{-1}\tilde{\mathbf{B}}^{\scriptscriptstyle T}\mathbf{P} & \mathbf{BR}^{-1}\tilde{\mathbf{B}}^{\scriptscriptstyle T} \\ -\left(\mathbf{PF}-\tilde{\mathbf{F}}-\tilde{\mathbf{H}}\mathbf{P}\right) & \tilde{\mathbf{B}}+\mathbf{PBR}^{-1}\tilde{\mathbf{B}}^{\scriptscriptstyle T} \end{pmatrix}$. *If*

$$\begin{cases} \left[(0,I)e^{\Psi(T-t)}\begin{pmatrix}0\\I\end{pmatrix}\right]^{-1} \in L^1(0,T;\mathbb{R}^{10n\times10n}), \\ \left[(0,I)e^{\Phi(T-t)}\begin{pmatrix}0\\I\end{pmatrix}\right]^{-1} \in L^1(0,T;\mathbb{R}^{10n\times10n}), \end{cases}$$

then (16) *admits unique solutions* $\mathbf{P}$, $\mathbf{K}$ *as*

$$\begin{cases} \mathbf{P}(t) = -\left[(0,I)e^{\Psi(T-t)}\begin{pmatrix}0\\I\end{pmatrix}\right]^{-1}(0,I)e^{\Psi(T-t)}\begin{pmatrix}I\\0\end{pmatrix}, \\ \mathbf{K}(t) = -\left[(0,I)e^{\Phi(T-t)}\begin{pmatrix}0\\I\end{pmatrix}\right]^{-1}(0,I)e^{\Phi(T-t)}\begin{pmatrix}I\\0\end{pmatrix}. \end{cases}$$
(17)

Remark 4.1 *Actually, for different* α *and* β, *we may reach different kinds of mean field problems and find some interesting phenomena. Details can be referred to Appendix A.*

4.3 Distributed design

Thus, we can conclude the following procedure of deriving the mean-field strategy:

Step 1 By (17), the solution $(\mathbf{P},\mathbf{K})$ of Riccati equations (16) can be obtained.
Step 2 For any agent $\mathcal{A}_i$ (or $\mathcal{A}_j$) in $\mathcal{I}_1$ (or $\mathcal{I}_2$), besides its own information, it still need the information of another generic agent denoted by $\mathcal{A}_j$ (or $\mathcal{A}_i$) in its opposite team $\mathcal{I}_2$ (or $\mathcal{I}_1$). Then it can obtain its feedback form mean-field strategy by

$$\begin{pmatrix}\tilde{u}_i\\\tilde{u}_j\end{pmatrix} = \Theta_1\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix} + \Theta_2\left(\mathbb{E}\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix} - \begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix}\right),$$

where $\Theta_1 = -\mathbf{R}^{-1}\tilde{\mathbf{B}}^T\mathbf{P}$ and $\Theta_2 = -\mathbf{R}^{-1}\tilde{\mathbf{B}}^T\mathbf{K}$.
Step 3 The realized states $\tilde{x}_i$, $\tilde{x}_j$ satisfy the following bilateral closed-loop system:

$$\begin{aligned} d\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix} &= \left\{\left[\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix} + \begin{pmatrix}B_1&0\\0&B_2\end{pmatrix}(\Theta_1-\Theta_2)\right]\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix}\right. \\ &\quad + \begin{pmatrix}B_1&0\\0&B_2\end{pmatrix}\Theta_2\mathbb{E}\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix} + \begin{pmatrix}F_1\\F_2\end{pmatrix}\tilde{x}^{(N)}\right\}dt \\ &\quad + \begin{pmatrix}\sigma_1\\0\end{pmatrix}dW_i + \begin{pmatrix}0\\\sigma_2\end{pmatrix}dW_j, \quad \begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix}(0) = \begin{pmatrix}\xi_i\\\eta_j\end{pmatrix}. \end{aligned}$$

Remark 4.2 *It should be noticed that there are MF terms in* **Step 2-3**. *To break up this couple feather, we should take the expectation to the system, which becomes an ordinary differential equation (ODE). By solving the ODE, we derive* $\mathbb{E}\begin{pmatrix}\tilde{x}_i\\\tilde{x}_j\end{pmatrix}$ *which implies the MF term, and then we focus on the SDE without MF terms.*

5 Performance analysis

We present **(Mix)** analysis in a three-step procedure by highlighting its novelty via pairwise comparisons to "pure" MG and MT. All limits below are in $N \to +\infty$ sense.

We continue to analyze the performance of **(Mix)** strategy $(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)$ derived by Section 4, in an asymptotic mixed-equilibrium-optima (**AMEO**) sense (it combines *equilibrium* due to game, and (social) *optima* due to team):

$$\textbf{(AMEO)}\begin{cases} \exists\, \rho,\ \lambda > 0, \text{s.t. } \text{LS1}: \\ \left|\dfrac{1}{N_1}\mathbf{J}_{\text{mix}}^{(1)}(\widehat{\mathbf{u}}_1, \bar{\mathbf{u}}_2) - \dfrac{1}{N_1}\mathbf{J}_{\text{mix}}^{(1)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)\right| \\ = O(N_1^{-\rho} + \epsilon_N^\lambda), \\ \exists\, \rho,\ \lambda > 0, \text{s.t. } \text{LS2}: \\ \left|\dfrac{1}{N_2}\mathbf{J}_{\text{mix}}^{(1)}(\bar{\mathbf{u}}_1, \widehat{\mathbf{u}}_2) - \dfrac{1}{N_2}\mathbf{J}_{\text{mix}}^{(1)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)\right| \\ = O(N_2^{-\rho} + \epsilon_N^\lambda), \end{cases}$$
(18)

where $\epsilon_N = \sup_{1 \le k \le 2}\left|\pi_k^{(N)} - \pi_k\right|$.

AMEO poses an *inside-outside-mixed* concept: social (team) optima inside, and Nash (game) equilibrium outside (LS1,LS2), along perturbed $\widehat{\mathbf{u}}_1 := (\cdots \hat{u}_k \cdots)|_{k \in \mathcal{I}_1}$, $\widehat{\mathbf{u}}_2 := (\cdots \hat{u}_k \cdots)|_{k \in \mathcal{I}_2}$. AMEO mainly includes three sub-steps as below.

(1) Study convergence behavior of realized empirical LS1, LS2 averages to off-lined CC system (cf.(14)) in an L^2-norm. A key point here is a refined bilateral forward-backward SDE (FBSDE) estimates on $(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)$.
(2) Estimate two upper bound(s) for candidate perturbation(s) $(\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2)$ respectively subject to $\mathbf{J}_{\text{mix}}^{(1)}(\widehat{\mathbf{u}}_1, \bar{\mathbf{u}}_2) \le \mathbf{J}_{\text{mix}}^{(1)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2), \mathbf{J}_{\text{mix}}^{(1)}(\bar{\mathbf{u}}_1, \widehat{\mathbf{u}}_2) \le \mathbf{J}_{\text{mix}}^{(2)}(\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2)$. Two keys: some tailor-made FBSDE stability estimates; $\mathbf{J}_{\text{mix}}^{(1)}$ convexity by underlying (Q,R) positiveness in (2).

(3) Formulate $\mathbf{J}_{\mathrm{mix}}^{(1)}$ in (18) as a *quadratic* functional on $(\mathbf{u}_1, \mathbf{u}_2)$ by specifying its second-order-operator(s) for convexity, and first-order ones for gradient. Its Fréchet derivatives to team-wise $(\mathbf{u}_1, \mathbf{u}_2)$, and componentwise $\{u_k\}_{k \in \mathcal{I}}$ can thus be structured. Then, AMEO can be verified by combining $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)$ bounds in step (ii), a near-stationarity (asymptotic zero Fréchet derivative) by step (i), and (uniform) *convexity* of $\mathbf{J}_{\mathrm{mix}}^{(1)}$.

Now, we give the main result of this work, whose proof is based on some lemmas. Please refer to Appendix B-E for details.

Theorem 5.1 *Under (H1)-(H4), the mean-field strategy $(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$ satisfies the following asymptotic optimality*

$$
\begin{cases}
\dfrac{1}{N_1}\left(\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) - \inf_{\mathbf{u}_1}\mathbf{J}_{\mathrm{mix}}^{(1)}(\mathbf{u}_1, \tilde{\mathbf{u}}_2)\right) = O(N_1^{-\frac{1}{2}} + \epsilon_N), \\[2ex]
\dfrac{1}{N_2}\left(\mathbf{J}_{\mathrm{mix}}^{(2)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) - \inf_{\mathbf{u}_2}\mathbf{J}_{\mathrm{mix}}^{(2)}(\tilde{\mathbf{u}}_1, \mathbf{u}_2)\right) = O(N_2^{-\frac{1}{2}} + \epsilon_N).
\end{cases}
\tag{19}
$$

Hence, $(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$ is an asymptotic Nash equilibrium for the game between $\mathcal{I}_1$ and $\mathcal{I}_2$.

Proof For L^2 bounded candidate $\hat{\mathbf{u}}_1$, by letting $\delta\mathbf{u}_1 = \tilde{\mathbf{u}}_1 - \hat{\mathbf{u}}_1$ we have

$$
\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) - \mathbf{J}_{\mathrm{mix}}^{(1)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) = 2\langle \mathbf{M}_2\tilde{\mathbf{u}}_1 + \mathbf{M}_1, \delta\mathbf{u}_1\rangle + o(\delta\mathbf{u}_1).
$$

Here, $\langle \mathbf{M}_2\tilde{\mathbf{u}}_1 + \mathbf{M}_1, \cdot\rangle$ is the Fréchet derivative of $J_{soc}^{(N)}$ on $\tilde{\mathbf{u}}_i$. Due to the linearity, we also get

$$
\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) - \mathbf{J}_{\mathrm{mix}}^{(1)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)
$$
$$
= 2\sum_{i=1}^{N_1}\langle \mathbf{M}_2\tilde{\mathbf{u}} + \mathbf{M}_1, \delta\mathbf{u}_1^i\rangle + o(\delta\mathbf{u}),
$$

where $\delta\mathbf{u}_1^i := (0, \cdots, 0, \tilde{u}_i - \hat{u}_i, 0, \cdots, 0)$. Based on the synthesis analysis in Section 3, we derive

$$
\langle \mathbf{M}_2\tilde{\mathbf{u}} + \mathbf{M}_1, \delta\mathbf{u}_1^i\rangle = \delta\mathbf{J}_{\mathrm{mix}}^{(1)}(\delta u_i)
$$
$$
= \mathbb{E}\int_0^T \Big[\Big\langle \bar{\Theta}_1 + \pi_1 F_1^T \mathbb{E}p_k^{(1)*} + \pi_2 F_2^T \mathbb{E}p_k^{(2)*} + \pi_1 F_1^T p^{(1)*}
$$
$$
+ \pi_2 F_2^T p^{(2)*}, \delta x_i\Big\rangle + \langle \Theta_2, \delta u_i\rangle\Big]dt + \varepsilon_1^i + \varepsilon_2^i + \varepsilon_3^i
$$
$$
= \varepsilon_1^i + \varepsilon_2^i + \varepsilon_3^i.
$$

By virtue of some FBSDE estimations, we obtain

$$
0 \leq \mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) - \mathbf{J}_{\mathrm{mix}}^{(1)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)
$$
$$
\leq \sum_{i=1}^{N_1}\left|\varepsilon_1^i + \varepsilon_2^i + \varepsilon_3^i\right| = N_1 O(N_1^{-\frac{1}{2}} + \epsilon_N).
$$

The first equation of (19) is established. Applying similar argument, we obtain the second one. $\square$

6 Conclusion

This paper investigates a new class of "mixed" mean-field analysis. The $\mathbf{C}_4$−type *coalition matrix* analysis is related to game-team problem. A novel bilateral person-by-person optimality is introduced and wellposedness of related CC system is investigated. Game-team strategies are designed and a novel asymptotic mixed-equilibrium-optima is proposed. An interesting work for further study is to consider the $\mathbf{C}_5$−type game-team problem referred in this work.

References

[1] D. Andersson and B. Djehiche (2010). A maximum principle for SDEs of mean-field type. Applied Mathematics and Optimization, 63(3), 341-356.

[2] A. Bensoussan, K. Sung, S. Yam, and S. Yung (2016). Linear-quadratic mean field games. Journal of Optimization Theory and Applications, 169, 496-529.

[3] R. Buckdhan, J. Li, and S. Peng (2014). Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents. SIAM Journal on Control and Optimization, 52, 451-492.

[4] P. J. Buckley and M. C. Casson (1998). Analyzing foreign market entry strategies: extending the internalization approach. Journal of International Business Studies, 29(3), 539-561.

[5] P. Cardaliaguet (2010). Notes on Mean Field Games, Technical report.

[6] R. Carmona and F. Delarue (2013). Probabilistic analysis of mean-field games. SIAM Journal on Control and Optimization, 51, 2705-2734.

[7] X. Du, P. He and J. R. Martins(2020). A B-spline-based generative adversarial network model for fast interactive airfoil aerodynamic optimization. In AIAA Scitech 2020 Forum (p. 2128).

[8] D. Gomes and J. Saude (2014). Mean field games models–a brief survey. Dyn. Games Appl., 4, 110-154.

[9] Y. Hu, J. Huang and T. Nie (2018). Linear-quadratic-Gaussian mixed mean-field games with heterogeneous input constraints. SIAM Journal on Control and Optimization, 56, 2835-2877.

[10] M. Huang, P. E. Caines, and R. P. Malhamé (2007). Large-population cost-coupled LQG problems with non-uniform agents: individual-mass behavior and decentralized ε-Nash equilibria. IEEE Transactions on Automatic Control, 52, 1560-1571.

[11] M. Huang, P. E. Caines, and R. P. Malhamé (2012). Social optima in mean-field LQG control: centralized and decentralized strategies. IEEE Transactions on Automatic Control, 57, 1736-1751.

[12] J. Lasry and P. Lions (2007). Mean field games. Japanese Journal of Mathematics, 2, 229-260.

[13] T. Michalak, J. Tyrowicz, P. McBurney and M. Wooldridge (2009). Exogenous coalition formation in the e-marketplace based on geographical proximity. Electronic Commerce Research and Applications, 8(4), 203-223.

[14] J. Moon and T. Başar (2018). Linear quadratic

mean field stackelberg differential games. Automatica, 97, 200-213.

[15] S. Nguyen, D. Nguyen and G. Yin (2020). A stochastic maximum principle for switching diffusions using conditional mean-fields with applications to control problems. ESAIM: Control, Optimisation and Calculus of Variations, 26, 1-26.

[16] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang (2013). Nash, social and centralized solutions to consensus problems via mean-field control theory. IEEE Transactions on Automatic Control, 58, 639-653.

[17] G. Nuno and B. Moll (2018). Social optima in economics with heterogeneous agents. Review of Economic Dynamics, 28, 150-180.

[18] H. Pham (2016). Linear quadratic optimal control of conditional McKean-Vlasov equation with random coefficients and applications. Probab. Uncertain. Quantita. Risk, 1, 1-26.

[19] J. Sun (2021). Two-Person Zero-Sum Stochastic Linear-Quadratic Differential Games. SIAM Journal on Control and Optimization, 59(3), 1804-1829.

[20] N. Sun, W. Trockel and Z. Yang (2008). Competitive outcomes and endogenous coalition formation in an n-person game. Journal of Mathematical Economics, 44, 853-860.

[21] H. Tembine, Q. Zhu, and T. Başar (2014). Risk-sensitive mean-field games. IEEE Transactions on Automatic Control, 59, 835-850.

[22] B. Wang, J. Zhang (2017). Social optima in mean-field linear-quadratic-Gaussian models with Markov jump parameters. SIAM Journal on Control and Optimization, 55, 429-456.

[23] L. Zhou (1994). A new bargaining set of an n-person game and endogenous coalition formation. Games and Economic Behavior, 6(3), 512-526.

Appendix

Appendix A: Some cases for different α and β.

(i) Assume $\alpha = \beta = 1$, then the cost functionals of Problem 2.1 reduce to

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{1,(N)} + \mathcal{J}_{\text{soc}}^{2,(N)},$$

which leads to a combination of traditional social optima problems.

If we further assume $A_1 = A_2 := A$, $B_1 = B_2 := B$, $F_1 = F_2 := F$, $Q_1 = Q_2 := Q$, $R_1 = R_2 := R$, $\Gamma_1 = \Gamma_2 := \Gamma$, ξ_i and η_j ($1 \leq i \leq N_1$, $1 \leq j \leq N_2$) are i.i.d and denoted by ζ_k ($1 \leq k \leq N$), in this case we derive $p^{(1)*} = p^{(2)*} = \hat{p}^{(1)*} = \hat{p}^{(2)*} := \hat{p}^*$. We also obtain that $(\bar{x}_i, \bar{u}_i, y_i, p_i^{(1)*}, p_j^{(2)*})$, $(\bar{x}_j, \bar{u}_j, y_j, \hat{p}_i^{(1)*}, \hat{p}_j^{(2)*})$ are homogeneous.

(ii) Assume $\alpha = \beta = -1$, then the cost functionals of Problem 2.1 reduce to

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathcal{J}_{\text{soc}}^{1,(N)} - \mathcal{J}_{\text{soc}}^{2,(N)}, \quad \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{2,(N)} - \mathcal{J}_{\text{soc}}^{1,(N)},$$

which leads to a combination of two social optima problems (**inside**) and a zero-sum problem (**outside**). This mix problem can be viewed as a development of two-person zero-sum game problem where two-person becomes two-team (e.g. [19], etc). If we further apply conditions on the coefficients, one may simplify the CC system and details are omitted.

(iii) Assume $\alpha = 1, \beta = 0$, or $\alpha = 0, \beta = 1$, then the cost functionals reduce to

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathcal{J}_{\text{soc}}^{1,(N)} + \mathcal{J}_{\text{soc}}^{2,(N)}, \quad \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{2,(N)},$$

or

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathcal{J}_{\text{soc}}^{1,(N)}, \quad \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{2,(N)} + \mathcal{J}_{\text{soc}}^{1,(N)},$$

which leads to two classes of social optima. It means one group focuses on a social optima in itself manner; while the other group would like to consider both intra-group and inter-group cooperations.

(iv) Assume $\alpha = -1, \beta = 0$, or $\alpha = 0, \beta = -1$, then the cost functionals reduce to

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathcal{J}_{\text{soc}}^{1,(N)} - \mathcal{J}_{\text{soc}}^{2,(N)}, \quad \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{2,(N)},$$

---

or

$$\mathbf{J}_{\text{mix}}^{(1)} = \mathcal{J}_{\text{soc}}^{1,(N)}, \quad \mathbf{J}_{\text{mix}}^{(2)} = \mathcal{J}_{\text{soc}}^{2,(N)} - \mathcal{J}_{\text{soc}}^{1,(N)},$$

which means one group focuses on a social optima in itself manner; while the other group would like to collaborate within the group itself and compete with the former group.

It should be noticed that, as a special case if $N = 2$, $\mathcal{J}_{\text{soc}}^{1,(N)} = \mathcal{J}_1$, $\mathcal{J}_{\text{soc}}^{2,(N)} = \mathcal{J}_2$, and a group referred above becomes an agent. Thus above problems turn out to be optimal control problem, two-person social optima, etc.

Appendix B: Estimation of mean field coupling.

Lemma 6.1 *Under (H1)-(H4), it holds that*

$$\left\| \frac{1}{N_1} \sum_{k \in \mathcal{I}_1} \bar{x}_k - \bar{\mathbf{m}}_1 \right\|_{L^2}^2 = O(N_1^{-1} + \epsilon_N^2),$$

$$\left\| \frac{1}{N_2} \sum_{k \in \mathcal{I}_2} \bar{x}_k - \bar{\mathbf{m}}_2 \right\|_{L^2}^2 = O(N_2^{-1} + \epsilon_N^2).$$

Proof The dynamic of the realized state average is

$$\begin{cases} d\tilde{x}^{(N_1)} = \left(A_1 \tilde{x}^{(N_1)} + B_1 \tilde{u}^{(N_1)} + F_1 \pi_1^{(N)} \tilde{x}^{(N_1)} \right. \\ \qquad \left. + F_1 \pi_2^{(N)} \tilde{x}^{(N_2)} \right) dt + \frac{1}{N_1} \sum_{i \in \mathcal{I}_1} \sigma_1 dW_i, \\ d\tilde{x}^{(N_2)} = \left(A_2 \tilde{x}^{(N_2)} + B_2 \tilde{u}^{(N_2)} + F_2 \pi_1^{(N)} \tilde{x}^{(N_1)} \right. \qquad (20) \\ \qquad \left. + F_2 \pi_2^{(N)} \tilde{x}^{(N_2)} \right) dt + \frac{1}{N_2} \sum_{j \in \mathcal{I}_2} \sigma_2 dW_j, \\ \tilde{x}^{(N_1)}(0) = \xi_i, \quad \tilde{x}^{(N_2)}(0) = \eta_j, \end{cases}$$

where $\tilde{x}^{(N_1)} := \frac{1}{N_1} \sum_{k \in \mathcal{I}_1} \tilde{x}_k$, $\tilde{x}^{(N_2)} := \frac{1}{N_2} \sum_{k \in \mathcal{I}_2} \tilde{x}_k$, $\tilde{u}^{(N_1)} := \frac{1}{N_1} \sum_{k \in \mathcal{I}_1} \tilde{u}_k$, $\tilde{u}^{(N_2)} := \frac{1}{N_2} \sum_{k \in \mathcal{I}_2} \tilde{u}_k$. The dynamic of the mean field is

$$\begin{cases} d\bar{\mathbf{m}}_1 = (A_1 \bar{\mathbf{m}}_1 + B_1 \mathbb{E}\breve{u}_i + F_1 \pi_1 \bar{\mathbf{m}}_1 + F_1 \pi_2 \bar{\mathbf{m}}_2) dt, \\ d\bar{\mathbf{m}}_2 = (A_2 \bar{\mathbf{m}}_2 + B_2 \mathbb{E}\breve{u}_j + F_2 \pi_1 \bar{\mathbf{m}}_1 + F_2 \pi_2 \bar{\mathbf{m}}_2) dt, \quad (21) \\ \bar{\mathbf{m}}_1(0) = \mathbb{E}\xi, \quad \bar{\mathbf{m}}_2(0) = \mathbb{E}\eta. \end{cases}$$

By applying Cauchy inequality, BDG inequality to the difference of (20)-(21), for some positive constant K which is independent of N_1, N_2, it holds that

$$\begin{cases} \mathbb{E}\sup_{0\leq s\leq t}\|\tilde{x}^{(N_1)}(s)-\bar{\mathbf{m}}_1(s)\|^2 \\ \leq K\int_0^t \mathbb{E}\sup_{0\leq r\leq s}\|\tilde{x}^{(N_1)}(r)-\bar{\mathbf{m}}_1(r)\|^2 ds + O(N_1^{-1}+\epsilon_N^2), \\ \mathbb{E}\sup_{0\leq s\leq t}\|\tilde{x}^{(N_2)}(s)-\bar{\mathbf{m}}_2(s)\|^2 \\ \leq K\int_0^t \mathbb{E}\sup_{0\leq r\leq s}\|\tilde{x}^{(N_2)}(r)-\bar{\mathbf{m}}_2(r)\|^2 ds + O(N_2^{-1}+\epsilon_N^2). \end{cases}$$

By applying Grönwall's inequality, we have

$$\begin{cases} \mathbb{E}\sup_{0\leq t\leq T}\|\tilde{x}^{(N_1)}(t)-\bar{\mathbf{m}}_1(t)\|^2 = O(N_1^{-1}+\epsilon_N^2), \\ \mathbb{E}\sup_{0\leq t\leq T}\|\tilde{x}^{(N_2)}(t)-\bar{\mathbf{m}}_2(t)\|^2 = O(N_2^{-1}+\epsilon_N^2). \end{cases}$$

$\square$

Appendix C: Perturbations estimation.

Lemma 6.2 *Under (H1)-(H4), for some positive constant K which is independent on N_1, N_2, it holds that*

$$\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \leq NK, \quad \mathbf{J}_{\mathrm{mix}}^{(2)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \leq NK.$$

Proof For $\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$, we have the following decompositions

$$\begin{cases} \mathcal{J}_i(\mathbf{u}_1,\mathbf{u}_2) = \frac{1}{2}\mathbb{E}\int_0^T \Big[\|x_i - \Gamma_1 x^{(N)}\|_{Q_1}^2 + \|u_i\|_{R_1}^2\Big]dt, \\ \mathcal{J}_j(\mathbf{u}_1,\mathbf{u}_2) = \frac{1}{2}\mathbb{E}\int_0^T \Big[\|x_j - \Gamma_2 x^{(N)}\|_{Q_2}^2 + \|u_j\|_{R_2}^2\Big]dt, \end{cases}$$

$$\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) = \mathcal{J}_{\mathrm{soc}}^{1,(N)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) + \alpha\mathcal{J}_{\mathrm{soc}}^{2,(N)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$$
$$= \frac{1}{2}\sum_{i=1}^{N_1}\mathbb{E}\int_0^T \Big[\|\tilde{x}_i - \Gamma_1 \tilde{x}^{(N)}\|_{Q_1}^2 + \|\tilde{u}_i\|_{R_1}^2\Big]dt$$
$$+ \frac{\alpha}{2}\sum_{j=1}^{N_2}\mathbb{E}\int_0^T \Big[\|\tilde{x}_j - \Gamma_2 \tilde{x}^{(N)}\|_{Q_2}^2 + \|\tilde{u}_j\|_{R_2}^2\Big]dt$$
$$\leq \frac{K}{2}\left(\sum_{i=1}^{N_1}\mathbb{E}\int_0^T \Big[\|\tilde{x}_i - \bar{\mathbf{m}}\|^2 + \|\bar{\mathbf{m}} - \tilde{x}^{(N)}\|^2 + \|\tilde{u}_i\|^2\Big]dt\right.$$
$$\left. + \alpha\sum_{j=1}^{N_2}\mathbb{E}\int_0^T \Big[\|\tilde{x}_j - \bar{\mathbf{m}}\|^2 + \|\bar{\mathbf{m}} - \tilde{x}^{(N)}\|^2 + \|\tilde{u}_j\|^2\Big]dt\right).$$

By applying Cauchy inequality, BDG inequality and Grönwall's inequality, for some positive constant K independent on N_1, N_2 we have

$$\begin{cases} \sup_{i\in\mathcal{I}_1}\mathbb{E}\sup_{0\leq t\leq T}\|\tilde{x}_i(t)-\bar{\mathbf{m}}(t)\|^2 \leq K, \\ \sup_{j\in\mathcal{I}_2}\mathbb{E}\sup_{0\leq t\leq T}\|\tilde{x}_j(t)-\bar{\mathbf{m}}(t)\|^2 \leq K. \end{cases}$$

Combined with Lemma 6.1, it holds that

$$\mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \leq NK.$$

Similar argument can be applied to $\mathbf{J}_{\mathrm{mix}}^{(2)}$ and we complete the proof. $\square$

Appendix D: Perturbations estimation.

Lemma 6.3 *Under (H1)-(H4), since we are studying the asymptotic optimality of $(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$, it is sufficient only to consider those admissible controls $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)$ performing better than $(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$. Specifically, $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)$ satisfies*

$$\sum_{i=1}^{N_1}\mathbb{E}\int_0^T \|\hat{u}_i\|^2 dt \leq NK, \quad \sum_{j=1}^{N_2}\mathbb{E}\int_0^T \|\hat{u}_j\|^2 dt \leq NK. \tag{22}$$

Proof Since $\hat{\mathbf{u}}_1$ performs better then $\tilde{\mathbf{u}}_1$, we have

$$\sum_{i=1}^{N_1}\mathbb{E}\int_0^T \|\hat{u}_i\|^2 dt \leq \mathcal{J}_{\mathrm{soc}}^{1,(N)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) + \alpha\mathcal{J}_{\mathrm{soc}}^{2,(N)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2)$$
$$= \mathbf{J}_{\mathrm{mix}}^{(1)}(\hat{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \leq \mathbf{J}_{\mathrm{mix}}^{(1)}(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \leq NK.$$

Similar argument can be applied to $\hat{\mathbf{u}}_2$ and the proof is complete. $\square$

Appendix E: Quadratic representation and Fréchet derivatives.

For the sake of notation simplicity, for a matrix M and positive number n_1 and n_2, we introduce the following notations

$$\begin{cases} \mathrm{diag}_{n_1}(M) := \begin{matrix}1\\\vdots\\n_1\end{matrix}\begin{pmatrix} M & & 0 \\ & \ddots & \\ 0 & & M \end{pmatrix} \text{ denotes an } n_1 \times n_1 \text{ size block} \\ \hspace{4cm}\text{diagonal matrix.} \\ \mathbb{1}_{n_1\times n_2}(M) := \begin{matrix}1\\\vdots\\n_1\end{matrix}\begin{pmatrix} M & \dots & M \\ \vdots & \ddots & \vdots \\ M & \dots & M \end{pmatrix} \text{ denotes an } n_1 \times n_2 \text{ size block} \\ \hspace{2.3cm}{\scriptstyle 1 \ \cdots \ n_2} \\ \hspace{4cm}\text{matrix generated by } M. \end{cases}$$

Rewrite the problem as the following high-dimensional type

$$\begin{cases} d\mathbb{X} = (\mathbb{A}\mathbb{X} + \mathbb{B}_1\mathbf{u}_1 + \mathbb{B}_2\mathbf{u}_2)dt + \sum_{i=1}^{N_1}\boldsymbol{\sigma}_i^1 dW_i + \sum_{j=1}^{N_2}\boldsymbol{\sigma}_j^2 dW_j, \\ \mathbf{J}_{\mathrm{mix}}^{(1)}(\mathbf{u}_1,\mathbf{u}_2) = \mathbb{E}\int_0^T \Big[\langle\mathbb{Q}_1\mathbf{x},\mathbf{x}\rangle + \langle\mathbb{R}_1\mathbf{u}_1,\mathbf{u}_1\rangle + \alpha\langle\mathbb{R}_2\mathbf{u}_2,\mathbf{u}_2\rangle\Big]dt, \\ \mathbf{J}_{\mathrm{mix}}^{(2)}(\mathbf{u}_1,\mathbf{u}_2) = \mathbb{E}\int_0^T \Big[\langle\mathbb{Q}_2\mathbf{x},\mathbf{x}\rangle + \langle\mathbb{R}_2\mathbf{u}_2,\mathbf{u}_2\rangle + \beta\langle\mathbb{R}_1\mathbf{u}_1,\mathbf{u}_1\rangle\Big]dt, \end{cases}$$

where

$$
\begin{cases}
\mathbb{X} = (\mathbf{x}_1, \mathbf{x}_2),\, \mathbf{u}_1 = (u_{\theta_1}, \cdots, u_{\theta_{N_1}}),\, \mathbf{u}_2 = (u_{\theta_1}, \cdots, u_{\theta_{N_2}}),\\[4pt]
\mathbb{A} = \begin{pmatrix} \operatorname{diag}_{N_1}(A_1)+\mathbb{1}_{N_1\times N_1}(\frac{F_1}{N}) & \mathbb{1}_{N_1\times N_2}(\frac{F_1}{N}) \\ \mathbb{1}_{N_2\times N_1}(\frac{F_2}{N}) & \operatorname{diag}_{N_2\times N_2}(A_2) \end{pmatrix},\, \boldsymbol{\sigma}_i^1 = \begin{pmatrix} \vdots \\ \sigma_1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix},\, \boldsymbol{\sigma}_j^2 = \begin{pmatrix} \vdots \\ 0 \\ \vdots \\ \sigma_2 \\ \vdots \\ 0 \end{pmatrix},\\[8pt]
\mathbb{B}_1 = \begin{pmatrix} \operatorname{diag}_{N_1}(B_1) \\ 0 \end{pmatrix},\, \mathbb{B}_2 = \begin{pmatrix} 0 \\ \operatorname{diag}_{N_2}(B_2) \end{pmatrix},\, \mathbb{R}_1 = \operatorname{diag}_{N_1}(R_1),\, \mathbb{R}_2 = \operatorname{diag}_{N_2}(R_2),\\[6pt]
\mathbb{Q}_1 = \begin{pmatrix} \mathbb{Q}_1^{(1,1)} & \mathbb{Q}_1^{(1,2)} \\ \mathbb{Q}_1^{(2,1)} & \mathbb{Q}_1^{(2,2)} \end{pmatrix},\, \mathbb{Q}_2 = \begin{pmatrix} \mathbb{Q}_2^{(1,1)} & \mathbb{Q}_2^{(1,2)} \\ \mathbb{Q}_2^{(2,1)} & \mathbb{Q}_2^{(2,2)} \end{pmatrix},\\[6pt]
\mathbb{Q}_1^{(1,1)} = \operatorname{diag}_{N_1}(Q_1)+\frac{1}{N}\mathbb{1}_{N_1\times N_1}\left(\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\Gamma_1^T Q_1+Q\Gamma_1+\alpha\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_1^{(1,2)} = \frac{1}{N}\mathbb{1}_{N_1\times N_2}\left(\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\alpha\Gamma_2^T Q_2+Q\Gamma_1+\alpha\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_1^{(2,1)} = \frac{1}{N}\mathbb{1}_{N_2\times N_1}\left(\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\alpha Q_2\Gamma_2+\Gamma_1^T Q+\alpha\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_1^{(2,2)} = \alpha\operatorname{diag}_{N_2}(Q_2)+\frac{1}{N}\mathbb{1}_{N_2\times N_2}\left(\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\alpha\Gamma_1^T Q_2+\alpha Q\Gamma_1+\alpha\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_2^{(1,1)} = \beta\operatorname{diag}_{N_1}(Q_1)+\frac{1}{N}\mathbb{1}_{N_1\times N_1}\left(\beta\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\beta\Gamma_1^T Q_1+\beta Q\Gamma_1+\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_2^{(1,2)} = \frac{1}{N}\mathbb{1}_{N_1\times N_2}\left(\beta\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\Gamma_2^T Q_2+\beta Q\Gamma_1+\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_2^{(2,1)} = \frac{1}{N}\mathbb{1}_{N_2\times N_1}\frac{1}{N}\left(\beta\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+Q_2\Gamma_2+\beta\Gamma_1^T Q+\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right),\\[6pt]
\mathbb{Q}_2^{(2,2)} = \operatorname{diag}_{N_2}(Q_2)+\frac{1}{N}\mathbb{1}_{N_2\times N_2}\left(\beta\pi_1^{(N)}\Gamma_1^T Q_1\Gamma_1+\Gamma_1^T Q_2+Q\Gamma_2+\frac{\pi_2^{(N)}}{N}\Gamma_2^T Q_2\Gamma_2\right).
\end{cases}
$$

Because of the game structure between $\mathcal{I}_1$ and $\mathcal{I}_2$, for agents in $\mathcal{I}_1$, they are facing with the following social optima problem

$$
\begin{cases}
d\mathbb{X} = (\mathbb{A}\mathbb{X}+\mathbb{B}_1\mathbf{u}_1+\mathbb{B}_2\tilde{\mathbf{u}}_2)\, dt+\sum_{i=1}^{N_1}\boldsymbol{\sigma}_i^1\, dW_i+\sum_{j=1}^{N_2}\boldsymbol{\sigma}_j^2\, dW_j,\\[10pt]
\min_{\mathbf{u}_1}\, \mathbf{J}_{\mathrm{mix}}^{(1)}(\mathbf{u}_1,\tilde{\mathbf{u}}_2)=\mathbb{E}\int_0^T \Big[\langle \mathbb{Q}_1\mathbb{X}, \mathbb{X}\rangle+\langle \mathbb{R}_1\mathbf{u}_1, \mathbf{u}_1\rangle \\[6pt]
\hspace{8em}+\langle \alpha\mathbb{R}_2\widetilde{\mathbf{u}}_2, \widetilde{\mathbf{u}}_2\rangle \Big]\, dt.
\end{cases}
$$

Since some bounded linear operators $(\mathbf{M}_2, \mathbf{M}_1, \mathbf{M}_0)$ are only dependent on the coefficients and $\tilde{\mathbf{u}}_2$, we can rewrite $\mathbf{J}_{\mathrm{mix}}^{(1)}$ as the following quadratic form

$$
\mathbf{J}_{\mathrm{mix}}^{(1)}(\mathbf{u}_1, \tilde{\mathbf{u}}_2) = \langle \mathbf{M}_2\mathbf{u}_1, \mathbf{u}_1\rangle + 2\langle \mathbf{M}_1, \mathbf{u}_1\rangle + \mathbf{M}_0.
$$