Linear Quadratic Mean Field Teams: Optimal and Approximately Optimal Decentralized Solutions

Jalal Arabneydi, Student Member, IEEE, and Aditya Mahajan, Senior Member, IEEE

Abstract—We consider team optimal control of decentralized systems with linear dynamics, quadratic costs, and arbitrary disturbance that consist of multiple sub-populations with exchangeable agents (i.e., exchanging two agents within the same sub-population does not affect the dynamics or the cost). Such a system is equivalent to one where the dynamics and costs are coupled across agents through the mean-field (or empirical mean) of the states and actions (even when the primitive random variables are non-exchangeable). Two information structures are investigated. In the first, all agents observe their local state and the mean-field of all sub-populations; in the second, all agents observe their local state but the mean-field of only a subset of the sub-populations. Both information structures are non-classical (i.e., exchanging two agents within the same sub-population is not possible or is not observed). The approximation error is inversely proportional to the size of the sub-populations whose mean-fields are not observed. The corresponding gains are determined by the solution of $K + 1$ decoupled standard Riccati equations, where $K$ is the number of sub-populations. The dimensions of the Riccati equations do not depend on the size of the sub-populations; thus the solution complexity is independent of the number of agents. Generalizations to major-minor agents, tracking cost, weighted mean-field, and infinite horizon are provided. The results are illustrated using an example of demand response in smart grids.

Index Terms—Stochastic dynamic teams, multi-agent systems, decentralized control, non-classical information structures, linear quadratic systems, team theory, large-scale systems.

I. INTRODUCTION

A. Motivation

Team optimal control of decentralized systems has been an important research topic since the mid 1960s. Many of the initial research results were negative and showed that even simple dynamical systems with two agents can be difficult to design—even in the celebrated linear quadratic Gaussian (LQG) framework. In particular, non-linear strategies can outperform the best linear strategy [2]; even if attention is restricted to linear strategies, the best linear strategy may not have a finite dimensional representation [3]. Since then, various solution methodologies for the optimal control of decentralized systems have been proposed and there has been considerable progress in understanding the nature of system dynamics and the information structure under which these methodologies work. See [4] and references therein for an overview.

In spite of this progress, there is a big gap between the theory and applications of optimal decentralized control. On the one hand, the envisioned applications—which include networked control systems, swarm robotics, and modern power systems—often consist of multiple interconnected dynamical systems and controllers. On the other hand, explicit optimal solutions are available only for systems with a few (often two or three) controllers [5]–[7]. The model and results presented in this paper attempt to reduce the gap between theory and applications.

In particular, we study decentralized control systems in which the dynamics and cost satisfy a property that we call exchangeability. In a dynamical system, we say agents $i$ and $j$ are exchangeable if exchanging (or interchanging) agents $i$ and $j$ does not affect the dynamics or the cost (the formal definition is given below). Or, equivalently, the dynamics and the cost do not depend on the index assigned to the two agents.

In many applications of decentralized systems, the system may be partitioned into sub-populations where all agents within a sub-population are exchangeable. We call such systems as systems with partially exchangeable agents. In this paper, we develop a framework for the design of optimal decentralized control for such systems.

B. System with partially exchangeable agents

To formally define exchangeability, consider a multi-agent dynamical system where $\mathcal{N}$ denotes the set of agents. The state and action of agent $i$, $i \in \mathcal{N}$, at time $t$ are denoted by $x^i_t$ and $u^i_t$, where $x^i_t \in \mathcal{X}^i$ and $u^i_t \in \mathcal{U}^i$. Let $x_t = (x^i_t)_{i \in \mathcal{N}}$ and $u_t = (u^i_t)_{i \in \mathcal{N}}$ denote the state and action of the entire system. The dynamics are given by

$$x_{t+1} = f_t(x_t, u_t, w_t),$$

where $f_t$ is system dynamics and $\{w_t\}_{t \geq 1}$, where $w_t = (w^i_t)_{i \in \mathcal{N}}$ and $w^i_t \in \mathcal{W}^i$, is the disturbance noise process. A per-step cost $c_t(x_t, u_t)$ is incurred at each time $t$.

For now, we do not specify the information structure as we want to identify the system properties that do not depend on the information structure.

1For example, consider an aggregator that provides demand response as a service by controlling the air conditioners in multiple neighborhoods in a city. The air-conditioners could be partitioned into sub-populations based on their tonnage and type (window, split, or packages AC). To the first-level of approximation, all air conditioners with the same tonnage and type have same dynamics and cost—and, therefore, are exchangeable. Similar situations arise in swarm robotics (where the subpopulations correspond to robots with different capabilities), and other engineering applications.
For any state $x$ and agents $i,j \in N$, let $\sigma_{ij} x$ denote the state when agents $i$ and $j$ are exchanged. For example, if $x = (x^1, x^2, x^3, x^4, x^5)$, then $\sigma_{23} x = (x^1, x^4, x^3, x^2, x^5)$. Similar interpretation holds for $\sigma_{ij} u$ and $\sigma_{ij} w$.

**Definition 1 (Exchangeable agents)** A pair $(i,j)$ of agents is exchangeable if the following conditions hold:

1. $X^i = X^j$, $U^i = U^j$, and $W^i = W^j$, i.e., the states, actions, and disturbances of agents $i$ and $j$ have the same dimensions.
2. For any $t$, and any $x_t$, $u_t$, and $w_t$,
   $$f_t(\sigma_{ij} x_t, \sigma_{ij} u_t, \sigma_{ij} w_t) = f_t(x_t, u_t, w_t),$$
   i.e., exchanging agents $i$ and $j$ does not affect the system dynamics.
3. For any $t$, and any $x_t$ and $u_t$,
   $$c_t(\sigma_{ij} x_t, \sigma_{ij} u_t) = c_t(x_t, u_t),$$
   i.e., exchanging agents $i$ and $j$ does not affect the cost.

**Definition 2 (Exchangeable set of agents)** A set $S$ of agents, $S \subseteq N$, is exchangeable if every pair of agents in $S$ is exchangeable.

**Definition 3 (System with partially exchangeable agents)**

The multi-agent system described above is called a system with partially exchangeable agents if the set $N$ of agents can be partitioned into $K$ disjoint subsets $N^k$, $k \in K := \{1, \ldots, K\}$, such that for each $k \in K$, the set $N^k$ of agents is exchangeable.

In this paper, we investigate optimal decentralized control of linear quadratic system (i.e., a system where dynamics are linear and the per-step cost is quadratic) with partially exchangeable agents. In a subsequent paper, we will investigate systems with controlled Markovian dynamics.

C. Notation

For a set $N$, $|N|$ denotes its size. For a matrix $A$, $A^T$ denotes its transpose, $\text{Tr}(A)$ denotes its trace; if $A$ is square, $A \geq 0$ (respectively $A > 0$) denotes that $A$ is positive semi-definite (respectively positive definite). For matrices $A$ and $B$ of appropriate size, $A \leq B$ means $B - A \geq 0$, $\text{diag}(A,B)$ denotes a block diagonal matrix with diagonal terms $A$ and $B$, $\sqrt{A}$ denotes a matrix $C$ such that $C^T C = A$, $A \circ B$ denotes Hadamard product, and $A \otimes B$ denotes Kronecker product. For matrices $A$, $B$, and $C$ with the same number of columns, $\text{rows}(A,B,C)$ denotes the matrix $[A^T, B^T, C^T]^T$. For vectors $x$, $y$, and $z$, $\text{vec}(x, y, z)$ denotes the vector $[x^T, y^T, z^T]^T$.

Superscripts index agents (indexed by $i$) or sub-populations (indexed by $k$). Given a set $N$ of agents and states $x^i$, $i \in N$, bold $x$ denotes $\text{vec}(x^1, \ldots, x^n)$; when all states are of the same dimension, $(x^i)_{i \in N}$ denotes the mean-field $\frac{1}{|N|} \sum_{i=1}^{|N|} x^i$ of $(x^i)_{i \in N}$. For vectors and matrices, we use the short hand notation $x_{1:t}$ or $A_{1:t}$ to denote $(x_1, \ldots, x_t)$ and $(A_1, \ldots, A_t)$, respectively.

$\mathbb{R}$, $\mathbb{R}_2$, and $\mathbb{R}_{>0}$ denote the sets of real, non-negative real, and positive real numbers, respectively. $I_{n \times m}$ denotes $n \times m$ matrix of ones, $I_n$ denotes $n \times n$ identity matrix. We omit the subscripts when the dimensions are clear from the context. For a random variable $x$, $\mathbb{E}[x]$ and $\text{var}(x)$ denote its mean and variance, respectively.

Given horizon $T$ and matrices $A_{1:T}$ and $Q_{1:T}$, the notation $M_{1:T} = \text{DLE}_{T}(A_{1:T}, Q_{1:T})$ means that $M_{1:T}$ is the solution of the finite horizon discrete Lyapunov equation, i.e., $M_T = Q_T$, and for $t \in \{T - 1, \ldots, 1\}$, $M_t = A_t^T M_{t+1} A_t + Q_t$.

Similarly, given a horizon $T$ and matrices $A_{1:T}$, $B_{1:T}$, $Q_1$, and $R_1$, the notation $M_{1:T} = \text{DRE}_{T}(A_{1:T}, B_{1:T}, Q_{1:T}, R_{1:T})$ means that $M_{1:T}$ is the solution of the finite horizon discrete Riccati equation, i.e., $M_T = Q_T$, and for $t \in \{T - 1, \ldots, 1\}$, $M_t = -A_t^T M_{t+1} B_t (B_t^T M_{t+1} B_t + R_t)^{-1} B_t^T M_{t+1} A_t + A_t^T M_{t+1} A_t + Q_t$.

Given a discount factor $\beta \in [0,1]$ and matrices $A, B, Q$, and $R$, the notation $M = \text{DALE}_\beta(A, Q)$ means that $M$ is the solution of the discrete algebraic Lyapunov equation

$$M = \beta A^T M A + Q,$$

and the notation $M = \text{DARE}_\beta(A, B, Q, R)$ means that $M$ is the solution of the discrete algebraic Riccati equation

$$M = -\beta A^T M B (B^T M B + \beta^{-1} R)^{-1} B^T M A + \beta A^T M A + Q.$$

II. PROBLEM FORMULATION AND LITERATURE OVERVIEW

A. Linear quadratic system with partially exchangeable agents

1. System Model: Suppose the dynamics (1) are linear, i.e.,
   $$x_{t+1} = A_t x_t + B_t u_t + w_t,$$
   where $A_t$ and $B_t$ are matrices of appropriate dimensions and $\{x_t, \{w_t\}_{t=1}^T\}$ are random variables defined on a common probability space. The cost is quadratic, i.e., for $t \in \{1, \ldots, T-1\}$,
   $$c_t(x_t, u_t) = x_t^T Q_t x_t + u_t^T R_t u_t,$$
   and $t = T$,
   $$c_T(x_T) = x_T^T Q_T x_T,$$
   where $Q_t$ and $R_t$ are matrices of appropriate dimensions. Furthermore, assume that the above system is partially exchangeable, i.e., agents $N$ can be partitioned into $K$ disjoint sub-populations $N^k, k \in K := \{1, \ldots, K\}$, such that for each $k \in K$, the agents $N^k$ are exchangeable. Moreover, for any sub-population $k \in K$ and agent $i \in N^k$, state $x^i_k$ takes values in $\mathbb{R}^{d^k}$ and action $u^i_k$ takes values in $\mathbb{R}^{d^k}$.

The mean-field of states $x^k$ of sub-population $k$, $k \in K$, is defined as the empirical mean of the states of all agents in that sub-population, i.e.,

$$\bar{x}^k_t := \frac{1}{|N^k|} \sum_{i \in N^k} x^i_t, \quad k \in K.$$

$^3$In the sequel, we refer to mean-field of the states simply as mean field.
Similarly, the mean-field of the actions $\bar{u}_t^k$ of sub-population $k$, $k \in K$, is defined as the empirical mean of the actions of all agents in that sub-population, i.e.,

$$\bar{u}_t^k := \frac{1}{|K|^k} \sum_{i \in N^K} u_i^k, \quad k \in K.$$ 

The mean-field of states and actions of the entire population are denoted by $\bar{x}_t$ and $\bar{u}_t$, respectively, i.e.,

$$\bar{x}_t = \text{vec}(\bar{x}_t^1, \ldots, \bar{x}_t^K), \quad \bar{u}_t = \text{vec}(\bar{u}_t^1, \ldots, \bar{u}_t^K).$$

For ease of reference, the notation is summarized in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Notation for agent $i \in N^K$ belonging to sub-population $k \in K$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i^t \in \mathbb{R}^{d_x}$</td>
</tr>
<tr>
<td>$u_i^t \in \mathbb{R}^{d_u}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Notation for sup-population $k \in K = {1, \ldots, K}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X^K_t = \langle {x_i^t}_{i \in N^K} \rangle$</td>
</tr>
<tr>
<td>$\bar{x}<em>t^k = \langle {x_i^t}</em>{i \in N^K} \rangle$</td>
</tr>
<tr>
<td>$\bar{u}<em>t^k = \langle {u_i^t}</em>{i \in N^K} \rangle$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Notation used for entire population</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N' = \bigcup_{k \in K} X^K_t$</td>
</tr>
<tr>
<td>$x_t = (x_t^i)_{i \in N'}$</td>
</tr>
<tr>
<td>$u_t = (u_t^i)_{i \in N'}$</td>
</tr>
<tr>
<td>$\bar{x}_t = \text{vec}(\bar{x}_t^1, \ldots, \bar{x}_t^K)$</td>
</tr>
<tr>
<td>$\bar{u}_t = \text{vec}(\bar{u}_t^1, \ldots, \bar{u}_t^K)$</td>
</tr>
</tbody>
</table>

### Proposition 1

In the linear quadratic system with partially exchangeable agents described above, there exist matrices $\{A_i^k, B_i^k, D_i^k, E_i^k, Q_i^k, R_i^k\}_{k \in K}$ and $P^x_t$ and $P^u_t$ such that the dynamics of agent $i \in N^K$ of sub-population $k, k \in K$, may be written as

$$x_{i+1}^t = A_i^k x_i^t + B_i^k u_i^t + D_i^k x_t + E_i^k \bar{u}_t + u_i^t; \quad (5)$$

the per-step cost at time $t \in \{1, \ldots, T-1\}$, may be written as

$$c_t(x_t, u_t, x_{i,t}, \bar{u}_t) = x_t^T P_t^x x_t + u_t^T P_t^u u_t$$

$$+ \sum_{k \in K} \sum_{i \in N^k} \frac{1}{|N|^k} \left[ (x_i^t)^T Q_i^k x_i^t + (u_i^t)^T R_i^k u_i^t \right]; \quad (6)$$

and the per-step cost at time $t = T$, may be written as

$$c_T(x_T, x_T) = x_T^T P_T^x x_T + \sum_{k \in K} \sum_{i \in N^k} \frac{1}{|N|^k} (x_i^t)^T Q_i^k x_i^t. \quad (7)$$

The proof is presented in Appendix A.

### Remark 1

In general, the matrices $\{A_i^k, B_i^k, D_i^k, E_i^k, Q_i^k, R_i^k\}$ and $(P_t^x, P_t^u)$ may depend on the number $\{|N^K|\}_{k \in K}$ of agents in the sub-populations, but their dimensions do not.

Thus, any linear quadratic system with partial exchangeable agents irrespective of the information structure is equivalent to a mean-field coupled system with the same information structure. In the rest of this paper, we investigate the optimal control of such systems under the following two information structures.

2) **Observation model and information structure:** We consider two information structures; in both, agents perfectly recall all data that they observe. In the first information structure, we call *mean field sharing* and denote by MFS-IS, every agent $i \in N$ perfectly observes its local state $x_i^t$ and the global mean-field $\bar{x}_t$. Thus, the data $I_t^i$ available to agent $i$ at time $t$ is given by

$$I_t^i = (x_i^{1:t}, u_i^{1:t-1}, \bar{x}_1:t). \quad (MFS-IS)$$

In the second information structure, which we call *partial mean field sharing* and denote by PMFS-IS, there exists a subset $S$ of the sub-populations $K$ such that every agent $i \in N$ perfectly observes its local state $x_i^t$ and the mean-fields of sub-populations $S$, i.e., $\{x_i^t\}_{k \in S}$. We use $S' \subset K \setminus S$. The data $I_t^i$ available to agent $i$ at time $t$ is given by

$$I_t^i = (x_i^{1:t}, u_i^{1:t-1}, (\bar{x}_i^t)_{k \in S}). \quad (PMFS-IS)$$

Under both information structures, agent $i$ chooses $u_i^t$ as follows:

$$u_i^t = g_i(I_t^i). \quad (8)$$

The function $g_i$ is called the control law of agent $i$ at time $t$. The collection $g^i = \{g_1^i, g_2^i, \ldots, g_T^i\}$ is called the control strategy of agent $i$. The collection $g = (g^i)_{i \in N}$ is called the control strategy of the system. The performance of strategy $g$ is given by

$$J(g) = \mathbb{E} \left[ \sum_{t=1}^{T-1} c_t(x_t, u_t, \bar{x}_t, \bar{u}_t) + c_T(x_T, x_T) \right], \quad (9)$$

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy $g$.

3) **The optimization problem:** We are interested in the following optimization problem.

**Problem 1** In the model described above, find a strategy $g^*$ that minimizes (9), i.e.,

$$J^* := J(g^*) = \inf_g J(g),$$

where the infimum is taken over all strategies of form (8).

### B. Conceptual difficulties

There are several conceptual difficulties in solving Problem 1 because it has a non-classical information structure. Information structure refers to the set of information known to all agents at all times. If every decision maker knows the observations and actions of all decision makers that acted before it, then the information structure is said to be classical; if every decision maker knows the observations and actions of all decision makers whose actions effect its observations, then the information structure is said to be partially nested; otherwise, the information structure is said to be non-classical [8], [9]. For linear quadratic systems with classical or partially nested information structures, when the primitive random variables are jointly Gaussian, the optimal control action is a linear (or affine) function of the observations. This is not necessarily

*In classical information structure with state feedback, the optimal control action is linear function of the state and this result holds even when the primitive random variables are not Gaussian.*
the case when the information structure is non-classical as is illustrated by the Witsenhausen counterexample [2], which presents a linear quadratic Gaussian model with non-classical information structure where non-linear strategies outperform the best linear strategy. The model presented in this paper is neither classical nor partially nested nor the primitive random variables are necessarily Gaussian, so it is not known a priori whether there is no loss of optimality in restricting attention to linear strategies.

Even when linear strategies are not optimal, sometimes attention is restricted to linear strategies because they are simple and easy to implement. For systems with non-classical information structure, the problem of finding the best linear strategy need not be convex; it is convex only for special sparsity pattern such as funnel causality [10] and quadratic invariance [11]. Furthermore, as is illustrated by the Whittle and Rudge counterexample [3], even when the problem of finding the best linear strategy is convex, the best linear strategy might not have a finite dimensional representation.

Finally, the usual curse of dimensionality is exasperated in systems with non-classical information structure. Even in systems with finite state and action spaces, the complexity of finding the optimal control strategy belongs to NEXP complexity class [12].

C. Contributions of the paper

1) We show that linear control laws are team optimal for MFS-IS (even when the noise processes are not Gaussian). As argued earlier, MFS-IS does not fall into the class of information structures for which linear strategies are known to be optimal. We show that the corresponding gains are computed by solving $K+1$ decoupled Riccati equations (where $K$ is the number of sub-populations) (Theorem 1).

2) We propose a certainty equivalence linear strategy for PMFS-IS and show that the error satisfies a Lyapunov equation. The approximation error converges to zero at a rate that is inversely proportional to the number of agents in the sub-populations whose mean-fields are not observed (Theorem 2).

3) The salient feature of our main results is that the solution complexity does not depend on the number of agents in each sub-population; rather, it only depends on the number of sub-populations. Furthermore, the optimal gains can be computed in a decentralized manner such that each agent simply needs to solve at most two (rather than all) Riccati equations.

4) We show that our results generalize to variations of the basic model that are not partially exchangeable including: systems where the objective is to optimally track reference trajectories (Sec. IV-C) and systems where agents have individual weights (Sec. IV-D).

5) When the dynamics and the per-step cost are time-homogeneous, we show that our results extend to infinite horizon setups: both for the discounted cost setup with any discount factor in $(0,1)$ and for average-cost per unit time setup. For both setups, the optimal control strategy for MFS-IS and the approximately optimal control strategy for PMFS-IS are time-homogeneous and the corresponding gains are computed by solving $K+1$ decoupled algebraic Riccati equations.

D. Literature overview

Our model and results for MFS-IS are similar in spirit to those obtained in [13] under stronger modeling assumptions. In [13], the authors consider a homogeneous population of dynamically decoupled agents which are coupled in the cost through a weighted mean-field term. Two models are investigated: (a) hard-constraint model where the weighted mean-field of actions must equal a pre-specified linear function of the weighted mean-field of states; and (b) soft-constraint model where the above hard constraint is relaxed by penalizing it in the cost. For both models, the authors show that the optimal centralized control laws are linear in the local state and the mean field; the corresponding gains are computed by two decoupled Riccati equations. In section IV-D, we generalize our results to the case when a weighted empirical mean-field is shared. In contrast to [13], we consider multiple sub-populations and allow agents to be coupled in dynamics. Note that approximation results similar to those for partial mean-field sharing were not considered in [13].

Our results have similar features to those obtained for centralized linear quadratic mean-field control [14], [15]. In these models, the dynamics and the cost depend on the statistical mean-field of the state and action. Such a model may be viewed as a special case of our model when we restrict to a single homogeneous sub-population and consider the limit of infinite number of agents (and therefore the empirical mean and the statistical mean are the same). Our proof technique, which relies on a simple change of variables, is conceptually simpler than that of [14], [15].\textsuperscript{5} It is worth highlighting that the linear quadratic mean-field control model is a centralized control problem and the results of [14], [15] do not apply to the multi-agent models that we consider.

Recently, an iterative bidding strategy was proposed in [17] for the optimal control multi-agent systems with decoupled dynamics that are coupled through a constraint. For LQG agents, the scheme operates as follows: at each time, a coordinator sets a price profile for all future times; agents submit a bid profile for all future times; the coordinator updates the prices and the process continues until the bids have converged. Agents choose the first value of their bid as their action and the above process is repeated at the next time step. In this scheme, agents do not need to know the system dynamics of other agents. In contrast, we assume that the system dynamics are common knowledge to all agents. However, in our model, agents only need to share the mean-field of their states (which can be computed using a consensus algorithm) rather than iteratively sharing the bid profile for all future times.

A decomposition-coordination approach for optimal decentralized control of deterministic linear quadratic systems was proposed in [16]. In [16], a matrix dynamical optimization method is used.

\textsuperscript{5}In [14], first coupled forward and backward stochastic differential equations are derived and then they are decoupled into two Riccati equations using the four step technique of [16]. In [15], a matrix dynamical optimization method is used.
proposed in [18], [19]. This is an iterative approach. Each iteration consists of two steps: (i) a decomposition step in which each agent assumes decoupled dynamics and costs and computes its local control trajectory by solving an optimal tracking problem from pre-specified linear offsets for the dynamics and a reference trajectory for the cost; (ii) a coordination step in which the linear offsets for the dynamics and reference trajectories for the cost are computed for all agents from the pre-specified control trajectories. It is shown that this iterative process converges to the optimal centralized solution. In contrast to such decomposition-coordination methods, our proposed solution is not iterative. The optimal gains for all agents are computed in a single step by solving Riccati equations. Furthermore, our solution methodology works for deterministic as well as stochastic systems.

A related solution approach called mean-field games (MFG) was proposed in [20]–[28] to compute approximate Nash equilibrium for large population games. The main idea is to assume an infinite large size of each sub-population and solve a set of two coupled equations: a Hamilton-Jacobi-Bellman (HJB) equation to compute the best response of a generic agent playing against a “mass trajectory” and a Fokker-Planck-Kolmogorov (FPK) equation to compute the mass trajectory from the strategy of a generic agent. It is shown that a solution to these equations exists under appropriate conditions. The resulting strategies are $\varepsilon$-Nash when the sub-populations are finite, where the approximation error is $O(1/\sqrt{n})$, where $n$ denotes the size of the smallest sub-population. For linear quadratic systems, the coupled HJB-FPK equations simplify to $K$ Riccati equations and two coupled forward and backward ODEs. In contrast, in our solution there is an additional Riccati equation instead of the coupled forward-backward equations. The coupled equations in MFG depend on the initial mean-field while the Riccati equations in our solution do not. The key difference between our results and the results in the MFG literature is that we obtain team optimal strategies of a decentralized control problem while in the MFG literature one typically obtains either Nash or Markov perfect equilibrium strategies of a large population dynamic game problem. These solution concepts are different.

The approach of mean-field games was used to obtain team optimal solution of linear quadratic systems with decoupled dynamics in [29]. It is shown that the MFG solution is $\varepsilon$-socially optimal (with $\varepsilon \in O(1/\sqrt{n})$). We obtain a similar result for dynamically coupled agents with $\varepsilon \in O(1/n)$.

It should be noted that identifying team-optimal control laws for systems with coupled dynamics is significantly more challenging than for systems with decoupled dynamics. This is because, when the agent dynamics are decoupled (and the primitive random variables are Gaussian), the information structure is partially nested, so one may restrict attention to linear strategies. Furthermore, for a finite horizon system, team-optimal strategies may be obtained by solving a set of linear equations.\(^6\) In contrast, when the system dynamics are coupled, the information structure is non-classical and there is no general solution methodology to obtain a team-optimal solution.

III. MAIN RESULTS

A. Exact solution for MFS-IS

We impose following standard assumptions on the model described in Proposition 1:

**Assumption (A1)** The primitive random variables \(\{x_1,\{w_t\}_{t=1}^T\}\) have zero mean, finite variance, and are mutually independent.

**Remark 2** Note that we do not require the primitive random variables to be Gaussian. Nor do we require the initial state \(x_1\) and the disturbance \(w_t\) to be independent or exchangeable across agents.

**Assumption (A2)** For every \(t\), \(P_t^x\), \(P_t^u\), \(Q_t^x\), and \(R_t^u\) are symmetric matrices that satisfy

\[
Q_t^k \geq 0, \quad \forall k \in \mathcal{K}, \quad \text{diag}(Q_t^1, \ldots, Q_t^K) + P_t^x \geq 0, \quad (10)
\]

\[
R_t^k \geq 0, \quad \forall k \in \mathcal{K}, \quad \text{diag}(R_t^1, \ldots, R_t^K) + P_t^u > 0. \quad (11)
\]

Note that matrices \(P_t^x\) and \(P_t^u\) are not required to be positive semi-definite as long as (10)–(11) hold.

**Theorem 1** Under (A1), (A2), and (MFS-IS), we have the following results for Problem 1.

1) Structure of optimal strategy: The optimal strategy of Problem 1 is unique and is linear in the local state and the mean-field of the system. In particular,

\[
u_t^i = \tilde{L}_t^k(x_t^i - \tilde{x}_t^k) + \tilde{L}_t^k\tilde{x}_t^i, \quad (12)
\]

where the gains \(\{\tilde{L}_t^k, \tilde{L}_t^k\}_{t=1}^T\) are obtained by the solution of \(K + 1\) Riccati equations given below: one for computing each \(\tilde{L}_t^k, k \in \mathcal{K}\), and one for \(\tilde{L}_t := \text{rows}(\tilde{L}_t^1, \ldots, \tilde{L}_t^K)\).

2) Riccati equations: Let

\[\begin{align*}
\tilde{A}_t := \text{diag}(A_1^k, \ldots, A^K_k) + \text{rows}(D_1^k, \ldots, D^K_k), \\
\tilde{B}_t := \text{diag}(B_1^k, \ldots, B^K_k) + \text{rows}(E_1^k, \ldots, E^K_k), \\
Q_t := \text{diag}(Q_1^k, \ldots, Q^K_k), \\
R_t := \text{diag}(R_1^k, \ldots, R^K_k).
\end{align*}\]

Then, for \(t \in \{1, \ldots, T - 1\}\), define:

\[
\tilde{L}_t^k = -\left((B_t^k)^T\tilde{M}_t^{k+1}B_t^k + R_t^k\right)^{-1}(B_t^k)^T\tilde{M}_t^{k+1}A_t^k,
\]

\[
\tilde{L}_t = -\left(\tilde{B}_t^T\tilde{M}_t\tilde{B}_t + \tilde{R}_t^u\right)^{-1}\tilde{B}_t^T\tilde{M}_t\tilde{A}_t,
\]

where \(\{\tilde{M}_t^k\}_{t=1}^T\) and \(\{\tilde{M}_t\}_{t=1}^T\) are the solutions of following Riccati equations:

\[
\tilde{M}_t^k = \text{DRE}_T(\tilde{A}_T^k, \tilde{B}_T^k, Q_T^k, \tilde{R}_T^k), \quad (13)
\]

\[
\tilde{M}_t = \text{DRE}_T(\tilde{A}_T, \tilde{B}_T, \tilde{Q}_T, \tilde{R}_T^u + P_t^x, \tilde{R}_t^u + P_t^u). \quad (14)
\]

\(^6\)It is shown in [9] that a finite horizon system with partially nested information structure may be converted to a static team by an appropriate change of variables. The optimal control laws for such a static team may be obtained by solving a set of linear equations [30]. The key conceptual challenge in such problem is to identify sufficient statistics such that the optimal control laws can be computed efficiently and the results can generalize to infinite-horizon setup.
3) Optimal performance: Let
\[ \Sigma_i^k := \frac{1}{|N|} \sum_{i \in N_k} \text{var}(u_i^t - \bar{u}_i^k), \quad \Sigma_t := \text{var}(\bar{u}_t), \]
\[ \Xi^k := \frac{1}{|N|} \sum_{i \in N_k} \text{var}(x_i^t - \bar{x}_i^t), \quad \Xi := \text{var}(\bar{x}_t). \]

Then, the optimal cost is given by
\[ J^* = \sum_{k \in K} \left( \sum_{t \in T} \text{Tr}(\Xi^k \bar{M}_t^k) \right) + \text{Tr}(\Xi \bar{M}_T) \]
\[ + \sum_{t=1}^{T-1} \left[ \sum_{k \in K} \text{Tr}(\Xi^k \bar{M}_{t+1}^k) + \text{Tr}(\Sigma_t \bar{M}_{t+1}) \right]. \quad (15) \]

Remark 4 An interesting feature of the solution is that all agents in a particular sub-population use identical control laws. This is a feature of the linear quadratic system and not of exchangeability.

Remark 5 We assumed that there are no cross-terms of the form \( x^T S u \) in the per-step cost of (3) and (4). If such cross-terms are present, there will be cross-terms involving \( (x_i^t, u_i^t) \), \( (x_i^t, \bar{u}_i) \), and \( (\bar{x}_t, u_i^t) \) in the equivalent mean-field model presented in Proposition 1. These cross-terms can be treated in the standard manner as cross-terms are treated in centralized LQR.

Remark 6 Suppose in addition to (A1), we have that \( \{x_i^t, \{u_i^t\}_{t=1}^2\}_{i \in N} \) are independent and for any \( k \in K \), \( x_i^t \) is i.i.d. with variance \( \Xi^k \) and \( \{u_i^t\}_{i \in N_k} \) is i.i.d. with variance \( \Sigma^k \). Then, we have
\[ \Sigma_i^k = \frac{|N|}{|N|} \Sigma^k, \quad \Sigma_t = \text{diag}(\Sigma^1, \ldots, \Sigma^K), \]
\[ \Xi^k = \frac{|N|}{|N|} \Xi^k, \quad \Xi = \text{diag}(\Xi^1, \ldots, \Xi^K). \]

The expression of total cost (15) can be simplified accordingly.

B. Approximate solution for PMFS-IS

In this section, we consider Problem 1 under PMFS-IS. Based on the results of Theorem 1, we propose a certainty equivalence strategy for PMFS-IS and show that the performance of this strategy is close to the optimal performance under MFS-IS. We impose the following assumptions on the model.

Assumption (A1a) In addition to (A1), for any \( k \in S \) and \( k' \in S' \), initial states \( \{x_i^1, \{u_i^t\}_{i \in N} \} \) are independent of \( \{x_i^1, \{u_i^t\}_{i \in N} \} \).

Assumption (A1b) The primitive random variables \( \{x_i^1, \{u_i^t\}_{i \in N} \} \) are independent. For any \( k, k' \in K \), there exist finite matrices \( c_k^0 \) and \( c_k^0 \) such that
\[ \sup_{i \in N_k} \text{var}(x_i^1) \leq c_k^0, \quad \sup_{t \leq T} \text{var}(u_i^t) \leq c_k^0. \]

Assumption (A3) The dynamics \( \{A_k^t, B_k^t, D_k^t, E_k^t\}_{k \in K} \), cost \( \{Q_k^t, P_k^t\}_{k \in K} \), and covariance bounds \( \{c_k^0, c_k^0\}_{k \in K} \) do not depend on the sizes \( |N^1|, \ldots, |N^{|K|}| \) of the sub-populations.

Since we are comparing the system performance under two information structures, we use different notation for the two. Under MFS-IS, the state and action of agent \( i \) are denoted by \( x_i^t \) and \( u_i^t \). Assume that \( u_i^t \) is generated as per Theorem 1. Under PMFS-IS, the state and action of agent \( i \) are denoted by \( s_i^t \) and \( v_i^t \). The dynamics are same as (5). In particular for agent \( i \) of sub-population \( k \in K \), \( s_i^1 = x_i^1 \) and
\[ s_i^{t+1} = A_k^t s_i^t + B_k^t u_i^t + D_k^t \bar{s}_t + E_k^t \bar{v}_t + w_i^t, \quad (16) \]

where
\[ s_t = \text{vec}(s_i^1, \ldots, s_i^K), \quad \bar{s}_t = \text{vec}(\{s_i^t\}_{i \in N_k}), \]
\[ v_t = \text{vec}(v_i^1, \ldots, v_i^K), \quad \bar{v}_t = \text{vec}(\{v_i^t\}_{i \in N_k}). \]

Define a (mean-field) approximation process \( \{z_t\}_{t=1}^T \) as follows: \( z_t = \text{vec}(z_i^1, \ldots, z_i^K) \), where for any \( k \in K \),
\( z_k^t \in \mathbb{R}^{d_k^*} \); the initial state \( z_1^k \) is given by \( z_1^k \) for \( k \in S \) and is 0 for \( k \notin S \). The process evolves as:

\[
\begin{cases}
    \overline{s}_{t+1}^k = \overline{s}_1^k, & k \in S,
    \\
    \overline{A}^k_{t+1} = (B^k \bar{L}^k + D^k + E^k \bar{L}^k)\overline{z}_t, & k \in S^c,
\end{cases}
\]  

where \( \bar{L}_t \) is as defined in Theorem 1. Note that the approximation process \( \{\overline{s}_k^t\}_{t=1}^T \) is adapted to the filtration \( \{\{\overline{s}_k^t\}_{k \in S}\}_{t=1}^T \) which is known at all agents. Therefore, at time \( t \), \( \overline{z}_t \) can be computed at all agents.

Now, consider the following certainty equivalence strategy for PMFS-IS: for agent \( i \) of sub-population \( k \in K \),

\[
v_i^t = \bar{L}_t^i (s_i^t - \tilde{s}_t^i) + \bar{E}_t^i z_t.
\]

The above strategy is similar to the optimal strategy for MFS-IS [given by (12) in Theorem 1] except that the mean field \( \{s_i^t\}_{k \in K}^t \) has been replaced by its approximation \( \tilde{s}_t^i \).

For ease of exposition, let \( d_c := \sum_{k \in K} d_k^c \) and matrix \( H = \text{rows}(H^1, \ldots, H^K) \) be a binary matrix such that

\[
H_k = \begin{cases}
    0_{d_k^c \times d_k^c}, & k \in S, \\
    1_{d_k^c \times d_k^c}, & k \in S^c.
\end{cases}
\]

Let \( \bar{J} \) denote the performance of strategy (18) and \( J^* \) denote the optimal performance under MFS-IS. Then, the difference in performance \( \bar{J} - J^* \) is bounded. In particular, we have

**Theorem 2** Assume (A1a), (A2), and (PMFS-IS). Then,

1. The performance loss is given by

\[
\bar{J} - J^* = \text{Tr}(\bar{X}_1 \bar{M}_1) + \sum_{t=1}^{T-1} \text{Tr}(\bar{W}_t \bar{M}_{t+1}),
\]

where \( \bar{X}_1 = \mathbb{1}_{2d_d \times 2d_d} \otimes [H \circ \text{var}(x_1)], \bar{W}_t = \mathbb{1}_{2d_d \times 2d_d} \otimes [H \circ \text{var}(\tilde{w}_t)] \), and \( \bar{M}_{1:T} \) is the solution of following Lyapunov equation:

\[
\bar{M}_{1:T} = \text{DLE}_T(\tilde{A}_{1:T}, \tilde{Q}_{1:T}),
\]

where

\[
\tilde{A} = \begin{bmatrix}
    \tilde{A}_1 - (\mathbb{1}_{d_c \times d_c} - H) \circ \tilde{A}_2^T,
    \\
    0 & H \circ \tilde{A}_2
\end{bmatrix},
\]

and \( \tilde{Q}_1 = \text{diag}(-\tilde{Q}_1, \tilde{Q}_2^T) \) where \( \tilde{A}_1 = \tilde{A}_1 + B_1 \bar{L}_t \), \( \tilde{A}_2 = \tilde{A}_2 + B_2 \bar{L}_t \), \( \tilde{Q}_1 = P_t^1 + Q_t + \bar{L}_t (P_t^1 + R_t) \bar{L}_t^T \), \( \tilde{Q}_2 = P_t^2 + Q_t + \bar{L}_t (P_t^2 + R_t) \bar{L}_t^T \), and \( \bar{L}_t = \text{diag}(\bar{L}_1^1, \ldots, \bar{L}_k^1) \).

2. Let \( n = \min_{k \in S^c}(|N^k|) \). Under (A1b) and (A3),

\[
\bar{J} - J^* \in \mathcal{O}\left(\frac{T}{n}\right).
\]

The result is proved in Section VI.

**Remark 7** As the number of agents in each sub-population \( k \in S^c \), becomes large, the approximation error \( \bar{J} - J^* \) goes to zero; therefore, PMFS-IS is as informative as MFS-IS.

Note that when the mean-field of all sub-populations are shared, then \( S = K \) and, therefore, \( H \) is zero. Consequently, the approximation error given by (19) is zero. Hence, the result of Theorem 2 is consistent with that of Theorem 1.

**Corollary 1** When the mean-field is not shared, i.e., \( S = \emptyset \), the approximation error \( \bar{J} - J^* \) is

\[
\text{Tr}(\text{var}(x_1)(\bar{M}_1^1 - \tilde{M}_1^1)) + \sum_{t=1}^{T-1} \text{Tr}(\text{var}(\tilde{w}_t)(\bar{M}_t^1 - \tilde{M}_t^1)),
\]

where \( \tilde{M}_{1:T}^1 \) and \( \tilde{M}_{1:T}^2 \) are the solutions of following two decoupled Lyapunov equations:

\[
\tilde{M}_{1:T}^1 = \text{DLE}_T(\tilde{A}_{1:T}^1, \tilde{Q}_{1:T}^1), \quad \tilde{M}_{1:T}^2 = \text{DLE}_T(\tilde{A}_{1:T}^2, \tilde{Q}_{1:T}^2).
\]

**Proof:** When \( S = \emptyset \), \( H \) is \( 1_{d_c \times d_c} \); thus, \( \bar{A}_t \) is block diagonal. Consequently, the Lyapunov equation (20) decouples into the two smaller Lyapunov equations given above.

**IV. SPECIAL CASES AND GENERALIZATIONS**

In this section, we present two special cases and two generalizations of Problem 1. Due to space limitations, we only present the results for MFS-IS (i.e., the analogue of Theorem 1); the results for PMFS (i.e., the analogue of Theorem 2) may be derived in a similar manner.

**A. Special case 1: major and minor agents**

Suppose there exist \( M \subseteq K \) sub-populations with only 1 agent, i.e., \( |N^k| = 1, k \in M \). Then, for every \( k \in M \), \( \bar{x}_k = x_k^1 \). The rest of the dynamics and cost are the same as in Section II-A. Since the dynamics are coupled through the mean-field, the states of the agents of sub-populations \( M \) directly influence the dynamics of all other agents and the per-step cost. For this reason, such agents are called major agents. A variant of the above model with a single major agent was first introduced in [32] and other variations have been investigated in [33]–[35].

For above model, result of Theorem 1 simplifies as follows.

**Corollary 2** For any sub-population \( k \in K \setminus M \) and minor agent \( i \in N^k \), \( u_i^t \) is given by (12). For any major agent \( i \in N^k \), \( k \in M \), the control law is given by \( u_i^t = \bar{L}_i^k x_i \).

Note that for \( k \in M \), \( \bar{L}_i^k \) is not needed to compute \( u_i^t \); so we do not need a Riccati equation to compute \( \bar{M}_{1:T}^k \).

**B. Special case 2: no local controls**

Suppose that for all \( k \in K \), \( B_k^0 = 0 \) and \( R_k^0 = 0 \). Moreover, assume that there exists a vector \( \theta_k = (\theta_k^1, \ldots, \theta_k^{|N^k|}), \theta_k^{|N^k|} \in \mathbb{R}^{d_d \times d_d}, k \in K \), such that \( E_k^0 = \bar{E}_k^0 \theta_k^1 \) for all \( k \in K \) and \( P_k^0 = \theta_k^1 P_k^0 \theta_k \). In addition, let \( \theta_k^{k+} \) denote the right inverse of \( \theta_k^k \) (i.e., \( \theta_k^k \theta_k^{k+} = I_{d_d} \)), which is assumed to exist. This implies that the dynamics and cost are given as follows. Let \( \tilde{u}_t := \theta_k^1 u_t = \sum_{k \in K} \theta_k^1 \tilde{u}_k^t \).

Then, for agent \( i \in N^k \) of sub-population \( k \in K \), we have

\[
x_{i,t+1} = A_k^i x_{i,t} + D_k^i x_t + \bar{E}_k^i \tilde{u}_t + w_i^t.
\]
At time $t \in \{1, \ldots, T-1\}$, the per-step cost is given by,
\[
c_t(x_t, u_t, \bar{x}_t, \tilde{u}_t) = \bar{x}_t^T P_t^* \bar{x}_t + \bar{u}_t^T P_u^* \bar{u}_t + \sum_{k \in K, i \in \mathcal{N}^k} \frac{1}{|\mathcal{N}^k|} (x_t - r_{i, t})^T Q_i x_{i, t},
\]
and $t = T$,
\[
c_T(x_T, \bar{x}_T) = \bar{x}_T^T P_{T, T}^* \bar{x}_T + \sum_{k \in K, i \in \mathcal{N}^k} \frac{1}{|\mathcal{N}^k|} (x_T - r_{i, T})^T Q_i x_{i, t}.\]

**Corollary 3** For the model described above, the optimal control law is given as follows. For all $k \in \mathcal{K}$ and $i \in \mathcal{N}^k$,
\[
u^*_i = \theta_i^{k, \bar{L}_i \tilde{x}_i},
\]
where $\{\bar{L}_1, \ldots, \bar{L}_K\} := \bar{L}_i$ is given as in Theorem 1 but with $\bar{B}_t$ replaced by $\bar{B}_t = \text{rows}(\bar{E}_t^1, \ldots, \bar{E}_t^K)$ and $P_u^*$ replaced by $P_{T, T}^*$. The proof is presented in Appendix B.

**Remark 8** Note that for the model defined above, each agent only needs to observe the mean-field of its sub-population (rather than the mean-field of entire population). Thus, this result is similar in spirit to [36, Theorem 1].

**C. Generalization 1: tracking cost function**

Consider a tracking problem in which we are given a tracking signal $\{s^k_t\}_{t=1}^{T}$, $s^k_t \in \mathbb{R}^d_s$ for the mean-field of sub-population $k \in \mathcal{K}$ and a tracking signal $\{r^i_{t}\}_{t=1}^{T}$, $r^i_t \in \mathbb{R}^d_r$, for each agent $i \in \mathcal{N}^k$.

Define $\bar{r}^i_t := \langle r^i_{t} \rangle_{i \in \mathcal{N}^k}, k \in \mathcal{K}$, $\bar{r}_i := \text{vec}(\bar{r}^1_t, \ldots, \bar{r}^K_t)$, and $s_t = \text{vec}(s^1_t, \ldots, s^K_t)$. The tracking cost is as follows. For $t \in \{1, \ldots, T-1\}$,
\[
c_t(x_t, u_t, \bar{x}_t, \tilde{u}_t) = (x_t - s_t)^T P_t^* (x_t - s_t) + \tilde{u}_t^T P_u^* \tilde{u}_t + \sum_{k \in K, i \in \mathcal{N}^k} \frac{1}{|\mathcal{N}^k|} (x_t - r^i_t)^T Q_i (x_t - r^i_t) + (u_t^i)^T R_i^k u_t^i,
\]
and for $t = T$,
\[
c_T(x_T, \bar{x}_T) = (x_T - s_T)^T P_{T, T}^* (x_T - s_T) + \sum_{k \in K, i \in \mathcal{N}^k} \frac{1}{|\mathcal{N}^k|} (x_T - r^i_T)^T Q_i (x_T - r^i_T).
\]

We assume that, in addition to MFS-IS specified in Section II-A, agent $i$ also knows signals $\{r^i_{t}, \bar{r}_i, s_t\}_{t=1}^T$. The rest of the model is the same as in Section II-A.

**Theorem 3** Under (A1), (A2), and (MFS-IS), the optimal strategy is unique and given by
\[
u^*_i = \bar{L}_i^* x_t - \tilde{x}_i^* + \bar{F}_k u^*_i + \bar{F}_u^* \tilde{u}_i, \tag{21}
\]
where the gains $\{\bar{L}_i^*, \bar{F}_k^*, \bar{F}_u^* \}_{k=1}^{T}$ are obtained by the solution of $K + 1$ Riccati equations defined in Theorem 1 and the gains $\{\bar{L}_i^*, \bar{F}_k^*, \bar{F}_u^* \}_{k=1}^{T}$ and the correction signals $\{v^*_i, \bar{v}_i\}_{t=1}^T$ are given as follows. Let $\{\bar{M}_i t \}_{t=1}^T$ and $\{\bar{M}_i^k t \}_{t=1}^T$ be the solutions of $K + 1$ Riccati equations defined in Theorem 1. For $t \in \{1, \ldots, T-1\}$, the gains $\{\bar{F}_k^*, \bar{F}_k^* \}_{k=1}^{T-1}$ are given by
\[
\bar{F}_t^k = \left( (B_{t}^1)^T \bar{M}_{t+1}^1 B_{t}^1 + R_{t}^k \right)^{-1} B_{t}^1 \bar{M}_{t+1}^1 B_{t}^1,
\]
and rows $(\bar{F}_1, \ldots, \bar{F}_K) := \bar{F}_t$, where
\[
\bar{F}_t = (B_{t}^1 \bar{M}_{t+1}^1 B_{t} + R_{t}^k)^{-1} B_{t}^1.
\]
The correction signals $\{v^*_i, \bar{v}_i\}_{t=1}^T$ are given recursively as follows: for $t = T$,
\[
v_T^* = Q_k^* \bar{F}_t^*, \quad \bar{v}_T = \bar{Q}_T \bar{F}_T + P_{T}^* s_T, \tag{22}
\]
and for $t \in \{T - 1, \ldots, 1\}$,
\[
v^*_t = (A_i^t + B_i^t \bar{F}_t^*)^T v_{i+1}^* + Q_i^t r^i_t, \tag{23}
\]
\[
\bar{v}_t = (A_i^t + B_i^t \bar{F}_t^*)^T \bar{v}_{i+1} + Q_i^t \bar{r}_i + P_{T}^* s_t. \tag{24}
\]

The proof is presented in Appendix C. To implement the optimal control strategies:

- all agents must compute $\bar{L}_1:T-1$ and $\bar{F}_1:T-1$ by solving Riccati equation (14) and compute the global correction signal $\bar{v}_T$ by solving backward equations (22) and (24);
- agents of sub-population $k$ must compute $\bar{L}_k^*_{1:T-1}$ and $\bar{F}_k^*_{1:T-1}$ by solving Riccati equation (13),
- an individual agent $i$ of sub-population $k$ must compute a local correction signal $v^*_i$ by solving backward equations (22) and (23).

Then, an individual agent $i$ of sub-population $k$, upon observing the local state $x^i_t$ and the global mean-field $\bar{x}_t$, chooses its local control action according to (21).

**D. Generalization 2: weighted mean-field**

Suppose there are weights $(a^i, \lambda^i, b^i)$ associated with each agent $i \in \mathcal{N}$ such that $a^i, \lambda^i \in \mathbb{R}$ and $b^i \in \mathbb{R}_{>0}$. For each sub-population $k \in \mathcal{K}$ define the weighted mean-field of states and actions as follows.
\[
\bar{x}^k_{i, t} = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \lambda^i x^i_{i, t}, \quad u^k_{i, t} = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \lambda^i u^i_{i, t},
\]
\[
\bar{x}^i_{k, t} = \text{vec}(\bar{x}^{1, \lambda} \bar{x}^{2, \lambda} \ldots \bar{x}^{K, \lambda}), \quad u^i_{k, t} = \text{vec}(u^{1, \lambda} \bar{u}^{2, \lambda} \ldots \bar{u}^{K, \lambda}).
\]

Also, define $\bar{a}^k_{i, t} = \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \lambda^i a^i_{i, t}$. For sub-population $k \in \mathcal{K}$, the state of agent $i \in \mathcal{N}^k$ evolves as follows.
\[
x^i_{t+1} = A^i_{k} x^i_{k, t} + B^i_{k, t} u^i_{k, t} + a^i (D^i_{k} \bar{x}^i_{k, t} + E^i_{k} \bar{u}^i_{k, t}) + w^i_{t}.
\]

The per-step cost is given by
\[
c_t(x_t, u_t, \bar{x}_t, \tilde{u}_t) = (\bar{x}_t)^T P_{T, T}^* \bar{x}_t + (u_t)^T P_{T, T}^* u_t
\]
\[
+ \sum_{k \in K, i \in \mathcal{N}^k} \frac{b^i}{|\mathcal{N}^k|} \left( (x_t)^T Q_t x_t + (u_t)^T R_t u_t \right),
\]
and the terminal cost is given by
\[
c_T(x_T, \bar{x}_T) = (\bar{x}_T)^T P_{T, T}^* \bar{x}_T + \sum_{k \in K, i \in \mathcal{N}^k} \frac{b^i}{|\mathcal{N}^k|} \left( (x_T)^T Q_T x_T \right).
\]

Such models arise in applications where the interaction between two homogeneous agents is not symmetric but depends on their weights. For example, in wireless networks, the interference caused at the base-station depends on the distance of the agents from the base-station.

In the above model, agents are not partially exchangeable. Nonetheless, we are able to explicitly identify optimal control strategies under the following assumptions.
Assumption (A4) For each sub-population \( k \in \mathcal{K} \) and each agent \( i \in \mathcal{N}^k \), \( a^i b^i = \lambda^i a^i b^i \).

Given a sub-population \( k \in \mathcal{K} \), examples of weights that satisfy (A4) are: for all \( i \in \mathcal{N}^k \), (i) \( a^i = 0 \), (ii) \( a^i = 1 \) and \( b^i = \lambda^i \), (iii) \( a^i = \lambda^i \), \( b^i = 1 \), and \( \frac{1}{|\mathcal{K}|} \sum_{i \in \mathcal{N}^k} \lambda^i = 1 \). To simplify the exposition, define \( \mu^k := 2 - \frac{1}{|\mathcal{K}|} \sum_{i \in \mathcal{N}^k} (\lambda^i)^2 \).

Assumption (A2a) For every \( t, P^x_t, P^u_t, Q^k_t \), and \( R^k_t \) are symmetric matrices that satisfy
\[
Q^k_t \geq 0, \quad \forall k \in \mathcal{K},
\]
\[
R^k_t > 0, \quad \forall k \in \mathcal{K},
\]
where the gains \( \{L^k_t, L^k_t\}_{t=0}^{T-1} \) are obtained by the solution of \( K \times \) Riccati equations defined in Theorem 1 when \( A_t, B_t, Q_t, R_t \) are replaced by
\[
\bar{A}_t := \text{diag}(A^1_t, \ldots, A^K_t) + \text{rows}(a_1^{1, \lambda} D^1_t, \ldots, a^K_{\lambda} D^K_t),
\]
\[
\bar{B}_t := \text{diag}(B^1_t, \ldots, B^K_t) + \text{rows}(a_1^{1, \lambda} E^1_t, \ldots, a^K_{\lambda} E^K_t),
\]
\[
\bar{Q}_t := \text{diag}(\mu^1 Q^1_t, \ldots, \mu^K Q^K_t),
\]
\[
\bar{R}_t := \text{diag}(\mu^1 R^1_t, \ldots, \mu^K R^K_t).
\]

The proof is presented in Appendix D.

Remark 9 The optimal strategy depends on the weights and, even within a sub-population, the gains of the mean-field terms are different for different agents.

Remark 10 If the dynamics of the agents are decoupled, i.e., \( a^i = 0 \) for all agents, then the results of Theorem 4 are similar to the model with soft constraints discussed in [13].

Note that if \( a^i = b^i = \lambda^i = 1 \) for all agents, then the weighted mean-field model reduces to the basic model described in Proposition 1 and the result of Theorem 4 reduces to that of Theorem 1.

V. PROOF OF THEOREM 1

We start with the model presented in Proposition 1. The proof proceeds in three steps.

- **Step 1:** We use a coordinate transformation to construct a system that is isomorphic to the original system.

- **Step 2:** We construct an auxiliary system which is system of Step 1 with classical information structure (i.e., all decisions are made by a single agent).

- **Step 3:** We show that the optimal control laws of the auxiliary system can be implemented using MFS-IS. A fortiori, they are also optimal for MFS-IS.

A. Step 1: A coordinate transformation

Define \( \tilde{x}_t = x_t - \bar{x}^k \) and \( \tilde{u}_t = u_t - \bar{u}^k \) and consider the following coordinate transformation \( T \) of the state and action spaces: \( T \vec{(x_t)}_{i \in \mathcal{N}} = \vec{(x_t)}_{i \in \mathcal{N}, \bar{x}_t} \) and \( T \vec{(u_t)}_{i \in \mathcal{N}} = \vec{(u_t)}_{i \in \mathcal{N}, \bar{u}_t} \). Under this transformation, the dynamics (5) may be written as
\[
\tilde{x}_t = A^k_t \tilde{x}_t + B^k_t \tilde{u}_t + \bar{w}_t,
\]
where \( \bar{w}_t \) and \( \tilde{x}_t \) are defined as in Theorem 1.

The per-step cost \( c_t(x_t, u_t, \bar{x}_t, \bar{u}_t) \) can also be written in terms of the transformed variables. For that matter, we need the following result that is similar to the Parallel-Axis Theorem (or Huygens-Steiner Theorem) in mechanics [37]:

\[\text{Lemma 1} \quad \text{For any } x = \vec{x} \in \mathcal{X}^N \text{ and } \bar{x} = \langle x \rangle, \text{ let } \tilde{x} = x - \bar{x}, \ i \in \{1, \ldots, N\}. \text{ Then, for any matrix } Q \text{ of appropriate dimension,}
\]
\[
\frac{1}{N} \sum_{i=1}^{N} (\tilde{x}_i)^T Q \tilde{x}_i = \frac{1}{N} \sum_{i=1}^{N} (\tilde{x}_i)^T Q \tilde{x}_i + \tilde{x}^T Q \bar{x}.
\]

Proof: The result follows from elementary algebra and the observation that \( \sum_{i=1}^{N} \tilde{x}_i = 0 \).

An immediate consequence of Lemma 1 is the following:

\[\text{Corollary 4} \quad \text{For time } t, t \in \{1, \ldots, T\}, \text{ there exist functions } \{c^k_t\}_{k \in \mathcal{K}} \text{ and } c^e_t \text{ such that}
\]
\[c_t(x_t, u_t, \bar{x}_t, \bar{u}_t) = c^e_t(x_t, \bar{u}_t) + \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{N}^k} c^k_t(\tilde{x}_i, \tilde{u}_i),
\]
where
\[c^e_t(x_t, \bar{u}_t) = \bar{x}^T(Q_t + P^x_t) \bar{x}_t + \bar{u}^T(R_t + P^u_t) \bar{u}_t,
\]
\[c^k_t(\tilde{x}_i, \tilde{u}_i) = \frac{1}{|\mathcal{K}|} \left[ (\tilde{x}_i)^T Q^k_t \tilde{x}_i + (\tilde{u}_i)^T R^k_t \tilde{u}_i \right],
\]
and for \( t = T \),
\[c_T(x_T, \bar{x}_T) = c^e_T(x_T) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} c^k_T(\tilde{x}_i),
\]
where
\[c_T(x_T) = x^T(Q_T + P^x_T) x_T, \quad c^k_T(\tilde{x}_i) = \frac{1}{|\mathcal{K}|} (\tilde{x}_{i})^T Q^k_T \tilde{x}_{i}.
\]

Since the transformation \( T \) is an isomorphism, the transformed model with dynamics (25) and (26) and the per-step cost (27) and (28) is equivalent to the original model in Proposition 1, irrespective of the information structure.
B. Step 2: An auxiliary system

Consider an auxiliary system with state $\hat{x}_t = \text{vec}(\{\hat{x}_i^j\}_{j \in \mathcal{N}_i}, \hat{x}_t)$ and action $\hat{u}_t = \text{vec}(\{\hat{u}_i^j\}_{j \in \mathcal{N}_i}, \hat{u}_t$) (which is the same as the transformed model of Step 1). There is a single centralized agent that chooses $\hat{u}_t$ based on the observations. In particular, the centralized agent observes $\hat{x}_t$ and chooses $\hat{u}_t$ according to

$$\hat{u}_t = \hat{g}_t(\hat{x}_{1:t}, \hat{u}_{1:t-1}). \quad (29)$$

The performance of strategy $\hat{g} := (\hat{g}_1, \ldots, \hat{g}_T)$ is given by

$$\hat{J}(\hat{g}) = \mathbb{E}[\sum_{t=1}^{T-1} c_t(\hat{x}_t, \hat{u}_t, \hat{x}_t, \hat{u}_t) + c_T(\hat{x}_T, \hat{x}_T)], \quad (30)$$

where the expectation is with respect to the measure induced on all system variables by the choice of strategy $\hat{g}$. We are interested in the following optimization problem.

**Problem 2** In the auxiliary system, find strategy $\hat{g}^*$ that minimizes (30), i.e.,

$$\hat{J}^* := \hat{J}(\hat{g}^*) = \inf_{\hat{g}} \hat{J}(\hat{g}),$$

where the infimum is taken over all strategies of the form (29).

Let $J^*$ and $\hat{J}^*$ denote the optimal cost for Problem 1 and Problem 2, respectively. Since the per-step cost is the same in both cases, but Problem 2 is centralized, we have that

$$J^* \geq \hat{J}^*.$$

We identify the optimal control laws for the auxiliary system and show that these laws can be implemented in, and therefore are optimal for, the original decentralized system.

C. Step 3: The Optimal Solution of the Auxiliary System

The auxiliary system is a stochastic linear quadratic system. So, the optimal control laws are linear and the optimal gains are given by the solution of an appropriate Riccati equation. However, the dimension of the state $\hat{x}_t$, and therefore the dimension of the Riccati equation, increases with the number of agents. To overcome this challenge, we present an alternative approach that involves solving $K + 1$ Riccati equations that do not depend on the number of agents.

Since the auxiliary system is a stochastic linear quadratic system, the certainty equivalence principle [38, Theorem 6.1] holds. Therefore, the optimal control law is identical to the control law of the corresponding deterministic system, whose dynamics are given as follows: for $k \in \mathcal{K}$ and $i \in \mathcal{N}_k$

$$\hat{x}_{t+1}^k = A_t^k \hat{x}_t^k + B_t^k \hat{u}_t^k, \quad \hat{x}_{t+1} = \hat{A}_t \hat{x}_t + \hat{B}_t \hat{u}_t,$$

and whose per-step cost is $\hat{c}_t(\hat{x}_t, \hat{u}_t)$ given by Corollary 4. Under (A2), the deterministic centralized linear quadratic system is strictly convex; hence, the solution is unique [38, Theorem 4.1].

Note that this system consists of $(N + 1)$ components: $N$ components with state $\hat{x}_t^k$ and action $\hat{u}_t^k$, $i \in \mathcal{N}$, and one component with state $\hat{x}_t$ and action $\hat{u}_t$. The first $N$ components are split into $K$ classes of identical components— one for each sub-population. The components have decoupled dynamics and decoupled cost. Thus, the optimal control law of each class may be identified separately. In particular, from [38, Theorem 4.1], we have that

**Theorem 5** The optimal control strategy of the auxiliary system (i.e., Problem 2) is unique and given by

$$\hat{u}_t = \hat{L}_t \hat{x}_t, \quad \text{and for } k \in \mathcal{K}, i \in \mathcal{N}_k, \quad \hat{u}_t^k = \hat{L}_k \hat{x}_t^k,$$

where the gains $\{\hat{L}_t, \hat{L}_k\}_{t=1}^{T-1}$ are given as in Theorem 1.

Now, we transform the optimal centralized solution, given by Theorem 5, back to the original model (by taking the inverse of coordinate transformation used in Step 1), to get

$$\hat{u}_t^i \hat{u}_t + \hat{u}_t^k = \hat{L}_t^i (\hat{x}_t^i + \hat{x}_t^k) + \hat{L}_t^k \hat{x}_t,$$

Note that the above control laws are implementable under MFS-IS. Therefore, the solution of Problem 2 coincides with the solution of Problem 1 with MFS-IS.

VI. Proof of Theorem 2

A. Preliminary results

We use the same transformation as Step 1 in Section V-A. In particular, for any $k \in \mathcal{K}$ and $i \in \mathcal{N}_k$, define $\bar{x}_t^i := x_t^i - \bar{x}_t^i$, $\bar{u}_t^i := u_t^i - \bar{u}_t^i$, $\bar{s}_t^i := s_t^i - \bar{s}_t^i$, and $\bar{v}_t^i := v_t^i - \bar{v}_t^i$. Then, we have

**Lemma 2** For all $t$, $\bar{s}_t^i = \bar{x}_t^i$ and $\bar{v}_t^i = \bar{x}_t^i$. Consequently,

$$\hat{J} - J = \sum_{t=1}^{T} \left[ c_t(\bar{s}_t^i, \bar{v}_t^i) - \tilde{c}_t(\hat{x}_t^i, \hat{u}_t^i) \right]. \quad (31)$$

**Proof:** We prove the first part by induction. Note that $\bar{x}_t^i = \bar{s}_t^i$ and $\bar{u}_t^i = \bar{L}_t^i \bar{x}_t^i = \bar{L}_t^i \bar{s}_t^i$. This forms the basis of induction. Now assume that $\bar{s}_t^i = \bar{x}_t^i$ and $\bar{v}_t^i = \bar{u}_t^i$ and consider time $t + 1$. Then,

$$\bar{s}_{t+1}^i = A_t^i \bar{s}_t^i + B_t^i \bar{u}_t^i = A_t^i \bar{x}_t^i + B_t^i \bar{u}_t^i = \bar{x}_{t+1}^i.$$

Moreover, $\bar{v}_{t+1}^i = \bar{L}_t^i \bar{x}_{t+1}^i = \bar{L}_t^i \bar{x}_{t+1}^i = \bar{v}_{t+1}^i$. Thus, the result is true by induction. Equation (31) immediately follows from the first part and Corollary 4.

Next we simplify (31) in terms of the following relative errors: For any $k \in \mathcal{K}$, define

$$\hat{z}_t^k = \bar{x}_t^k - z_t^k \text{ and } \hat{z}_t^k = \bar{s}_t^k - z_t^k.$$  

Let $\zeta_t = \text{vec}(\zeta_1^t, \ldots, \zeta_k^t)$ and $\xi_t = \text{vec}(\xi_1^t, \ldots, \xi_k^t)$. For ease of exposition, let vector $h = \text{vec}(h^1, \ldots, h^K)$ be binary such that $h^k = 0$ on $x_t^i$ if $k \in S$ and $h^k = 1$ on $x_t^i$ if $k \in S^c$.

**Lemma 3** Let $\hat{A}_t$ be defined as in Theorem 2. Then, $\zeta_t = h \circ \xi_t$ and $\xi_t = h \circ \zeta_t$ and

$$\begin{bmatrix} \zeta_{t+1}^k \\ \xi_{t+1}^k \end{bmatrix} = \hat{A}_t^k \begin{bmatrix} \zeta_t^k \\ \xi_t^k \end{bmatrix} + \begin{bmatrix} h \circ \bar{w}_t^k \\ h \circ \bar{w}_t^k \end{bmatrix}.$$  

**Proof:** From (16) and (18), we get

$$\bar{s}_{t+1}^k = A_t^k \bar{s}_t^k + B_t^k \bar{u}_t^k + B_t^k \bar{v}_t^k + A_t^k \bar{v}_t^k + \bar{w}_t^k,$$

$$\bar{v}_t^k = \bar{L}_t^k (\bar{s}_t^k - \bar{x}_t^k) + \bar{L}_t^k \bar{z}_t^k,$$

where $\bar{w}_t^k := \text{vec}(\bar{w}_t^i)_{i \in \mathcal{N}_k}$. Write (32) in a vectorized form,

$$\bar{s}_{t+1} = \hat{A}_t \bar{s}_t + \hat{B}_t \bar{v}_t + \bar{w}_t, \quad \bar{v}_t = \bar{L}_t \bar{z}_t + \bar{L}_t \bar{z}_t,$$
where $\tilde{w}_t = \text{vec}(\tilde{w}_t^1, \ldots, \tilde{w}_t^K)$. From Theorem 1, we can write the dynamics under the optimal strategy as follows
\[
\bar{x}_{t+1}^k = A_k^t x_t^k + (B_k^t L_t^k + D_k^t + E_k^t L_t) \bar{x}_t + \tilde{w}_t^k;
\]
and in a vectorized form,
\[
\bar{x}_{t+1} = (\bar{A}_t + B_t \bar{L}_t) \bar{x}_t + \tilde{w}_t, \quad \bar{u}_t = \bar{L}_t \bar{x}_t.
\]
Thus, the dynamics of the relative errors can be written as follows. If $k \in S$,
\[
\zeta_t^k = \frac{A_k^t \zeta_t^k + (B_k^t L_t^k + D_k^t + E_k^t L_t) \zeta_t - (D_k^t + E_k^t L_t) \bar{x}_t}{\bar{L}_t} + \tilde{w}_t^k,
\]
and if $k \in S^c$,
\[
\zeta_t^k = \frac{A_k^t \zeta_t^k + (B_k^t L_t^k + D_k^t + E_k^t L_t) \zeta_t}{\bar{L}_t} + \bar{x}_t + \tilde{w}_t^k.
\]
Combining these, gives the result of the Lemma.

**Lemma 4** For all $t$, $E[\zeta_t | F_t] = E[\zeta_t | F_t] = 0$.

**Proof:** If $k \not\in S$, $\zeta_t^k = \zeta_t^k = 0$ and if $k \in S$, $\zeta_t^k = \zeta_t^k = \bar{x}_t^k$, and from (A1a), $E[\bar{x}_t^k | F_t] = E[\bar{x}_t^k] = 0$. Therefore, $E[\zeta_t | F_t] = E[\zeta_t | F_t] = 0$. Thus, from Lemma 3 and $E[\tilde{w}_t | F_t] = 0$, we get that $E[\zeta_t | F_t] = E[\zeta_t | F_t] = 0$.

**Lemma 5** $z_t$ is measurable with respect to $F_t$, therefore, $E[z_t | F_t] = z_t$.

**Proposition 2** The relative loss is given
\[
\bar{J} - J^* = E \left[ \sum_{i=1}^{T} \zeta_t^i \zeta_t^i ^T Q_t \zeta_t^i \zeta_t^i \right].
\]

**Proof:** Recall that $\hat{c}_t(\bar{x}_t, \bar{u}_t) = \bar{x}_t^T (\bar{Q}_t + P_t^u) \bar{x}_t + \bar{u}_t^T (\bar{R}_t + P_t^u) \bar{u}_t$. The proof follows immediately from (31) and the following observation:

**Lemma 6** Let $\hat{Q}_t := \bar{Q}_t + P_t^u$ and $\hat{R}_t := \bar{R}_t + P_t^u$. Then,
\[
E[\bar{x}_t^T \hat{Q}_t \bar{x}_t - \bar{x}_t^T \hat{Q}_t \bar{x}_t | F_t] = E[\zeta_t^T \hat{Q}_t \zeta_t - \zeta_t^T \hat{Q}_t \zeta_t | F_t],
\]
and
\[
E[\bar{u}_t^T \hat{R}_t \bar{u}_t - \bar{u}_t^T \hat{R}_t \bar{u}_t | F_t] = E[\zeta_t^T \hat{R}_t \zeta_t - \zeta_t^T \hat{R}_t \zeta_t | F_t].
\]
Therefore, the proof of Proposition 2 is complete.

**Proof of Lemma 6:**
1) Substituting $\bar{s}_t = \xi_t + z_t$ and $\bar{x}_t = \zeta_t + z_t$, we get
\[
E[\bar{s}_t^T \hat{Q}_t \bar{s}_t - \bar{x}_t^T \hat{Q}_t \bar{x}_t | F_t] = E[\zeta_t^T \hat{Q}_t \zeta_t - \zeta_t^T \hat{Q}_t \zeta_t | F_t] + 2E[\zeta_t^T \hat{Q}_t z_t | F_t] - 2E[\zeta_t^T \hat{Q}_t z_t | F_t] = E[\zeta_t^T \hat{Q}_t \zeta_t - \zeta_t^T \hat{Q}_t \zeta_t | F_t],
\]
where the last two terms in (a) are zero by Lemmas 4 and 5.
2) Substituting $\bar{v}_t = \bar{L}_t \xi_t + \bar{L}_t z_t$ and $\bar{u}_t = \bar{L}_t \bar{x}_t = \bar{L}_t (\xi_t + z_t)$, we get
\[
E[\bar{v}_t^T \hat{R}_t \bar{v}_t - \bar{u}_t^T \hat{R}_t \bar{u}_t | F_t] = E[\zeta_t^T \hat{L}_t \xi_t - \zeta_t^T \hat{L}_t \xi_t | F_t] + 2E[\zeta_t^T \hat{L}_t \bar{x}_t | F_t] - 2E[\zeta_t^T \hat{L}_t \bar{x}_t | F_t] = E[\zeta_t^T \hat{L}_t \xi_t - \zeta_t^T \hat{L}_t \xi_t | F_t],
\]
where the last two terms in (b) are zero by Lemmas 4 and 5.

**B. Proof of Theorem 2**

To prove part 1, note that $\bar{J} - J^*$ is the expected total quadratic cost (given by Proposition 2) of a linear (uncontrolled) system (given by Lemma 3). Thus, $\bar{J} - J^*$ is given by (19) where $\hat{M}_{1:T}$ is the solution of the Lyapunov equation (20). Note that the variance of the initial state and noises in Lemma 3 are given as follows:
\[
\var(h \circ \bar{x}_1, h \circ \bar{x}_1) = 1_{2d_x \times 2d_x} \otimes |H \circ \var(\bar{x}_1)| = \hat{X}_1,
\]
\[
\var(h \circ \bar{w}_1, h \circ \bar{w}_1) = 1_{2d_w \times 2d_w} \otimes |H \circ \var(\bar{w}_1)| = \hat{W}_1.
\]

To prove part 2 of Theorem 2, first observe that due to (A2), matrices $\bar{A}_t$ and $\bar{Q}_t$ do not depend on $([N_1], \ldots, [N_K])$; therefore, neither does $\hat{M}_{1:T}$. Thus the only dependence on the size of the sub-population is due to $\hat{X}_1$ and $\hat{W}_1$. Under (A1b) and (A3), for any sub-population $k \in K$,
\[
\var(\bar{x}_t^k) = \frac{1}{|N_k|^2} \sum_{i \in N_k} \var(x_t^i) \leq \frac{\epsilon_k}{n},
\]
\[
\var(\bar{w}_t^k) = \frac{1}{|N_k|^2} \sum_{i \in N_k} \var(w_t^i) \leq \frac{\epsilon_k}{n}.
\]
From (A1b), $\var(\bar{x}_1) = \text{diag}(\var(\bar{x}^1_i), \ldots, \var(\bar{x}^K_i))$ and $\var(\bar{w}_1) = \text{diag}(\var(\bar{w}^1_i), \ldots, \var(\bar{w}^K_i))$. Thus,
\[
\hat{X}_1 \leq \frac{1}{n} 1_{2d_x \times 2d_x} \otimes \text{diag}(c_x^1, \ldots, c_x^K),
\]
\[
\hat{W}_1 \leq \frac{1}{n} 1_{2d_w \times 2d_w} \otimes \text{diag}(c_w^1, \ldots, c_w^K).
\]
Thus, $\hat{X}_1$ and $\hat{W}_1$ are $O(\frac{1}{n})$. From (19), we have
\[
|\bar{J} - J^*| \leq \frac{1}{n} \text{Tr} \left( \hat{X}_1 \hat{M}_1 \right) + \sum_{i=1}^{T-1} \text{Tr} \left( \hat{W}_i \hat{M}_{i+1} \right),
\]
where each of above absolute values is $O(\frac{1}{n})$. In particular, since $\hat{X}_1$ and $\hat{W}_1$ are $O(\frac{1}{n})$ and $\hat{M}_{1:T}$ do not depend on $n$, $|\text{Tr}(\hat{X}_1 \hat{M}_1)|$ and $|\text{Tr}(\hat{W}_i \hat{M}_{i+1})|$ are $O(\frac{1}{n})$.

**VII. INFINITE HORIZON**

The results presented in Sections III and IV generalize to infinite horizon setup in a natural manner. Assume that the model is time-invariant, i.e., the matrices $\{A_k^t, B_k^t, E_k^t, Q_t^k, R_t^k, P_t^k, P_t^u\}$ and covariances $\{\Sigma_t^k, \Sigma_t^*, \Xi_t^k, \Xi_t^*\}$ (defined in Theorem 1) do not depend on time; hence, we remove the subscript $t$. The rest of the model is as same as that in Section II-A.
Consider the infinite horizon discounted cost and the infinite horizon long-term average setups as follows:

**Problem 3** Given discount factor \( \beta \in (0, 1) \), find a strategy \( \mathbf{g} \) that minimizes the following cost:

\[
J_\beta(\mathbf{g}) = (1 - \beta) \mathbb{E}[\sum_{t=1}^{\infty} \beta^{t-1} \mathbf{c}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_t, \bar{\mathbf{u}}_t)],
\]

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy \( \mathbf{g} \).

**Problem 4** Find a strategy \( \mathbf{g} \) that minimizes the following cost:

\[
J_1(\mathbf{g}) = \lim_{T \to \infty} \mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} \mathbf{c}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_t, \bar{\mathbf{u}}_t)\right],
\]

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy \( \mathbf{g} \).

**Assumption (A5)** For each sub-population \( k \in \mathcal{K} \), \( (\sqrt{\beta}A^k, \sqrt{\beta}B^k) \) are stabilizable and \( (\sqrt{\beta}A^k, \sqrt{\beta}Q^k) \) are detectable. In addition, for \( A_t \) and \( B_t \) defined in Theorem 1, \( (\sqrt{\beta}A, \sqrt{\beta}B) \) are stabilizable and \( (\sqrt{\beta}A, \sqrt{\beta}Q + P^k) \) are detectable.

A. Exact solution for MFS-IS

The optimal strategy under MFS-IS is as follows.

**Theorem 6** Under (A1), (A2), (A5), and (MFS-IS), the optimal strategy for Problems 3 and 4 are linear and time homogeneous and are given by

\[
\mathbf{u}_t = \bar{L}^k(x_t - \bar{x}_t^k) + \bar{L}^k \bar{x}_t,
\]

where the gains \( \{\bar{L}^k, \bar{L}^k\} \) are obtained by the solution of \( K+1 \) algebraic Riccati equations given below: for computing each \( \bar{L}^k, k \in \mathcal{K} \), and one for \( \bar{L} := \text{rows}(\bar{L}^1, \ldots, \bar{L}^K) \). Let matrices \( A, B, \bar{Q}, \) and \( \bar{R} \) be defined as in Theorem 1; then, given \( \beta \in (0, 1] \),

\[
\bar{L}^k = -\left(B^k \bar{M}^k B^k + \beta^{-1} R^k\right)^{-1} B^k \bar{M}^k A^k,
\]

\[
\bar{L} = -\left(B^T \bar{M} B + \beta^{-1}(\bar{R} + \bar{P}^u)\right)^{-1} B^T \bar{M} A,
\]

where \( \bar{M}^k \) and \( \bar{M} \) are the solutions of the following algebraic Riccati equations:

\[
\bar{M}^k = \text{DARE}_\beta(A^k, B^k, Q^k, R^k),
\]

\[
\bar{M} = \text{DARE}_\beta(A, B, \bar{Q} + P^u, \bar{R} + P^u).
\]

In addition, the optimal performance is given by

\[
J^*_\beta = (1 - \beta) \left[\sum_{k \in \mathcal{K}} \text{Tr}\left(\bar{\Sigma}^k \bar{M}^k\right) + \text{Tr}(\bar{\Xi} \bar{M})\right] + \left[\sum_{k \in \mathcal{K}} \text{Tr}\left(\bar{\Sigma}^k \bar{M}^k\right) + \text{Tr}(\Sigma \bar{M})\right],
\]

where \( \bar{\Sigma}^k, \Sigma, \bar{\Xi}, \bar{\Xi}, \) and \( \Xi \) are defined as in Theorem 1.

**Proof:** The proof follows along the same lines of the proof of Theorem 1. We construct an auxiliary system as in Section V, which consists of \( |\mathcal{N}| + 1 \) components with decoupled cost and dynamics coupled only through the noise. Since the costs are infinite-horizon discounted and infinite-horizon long run average, the optimal solution is given by appropriate algebraic Riccati equations. Under (A2) and (A5), these Riccati equations have a unique solution [38, Theorem 9.2].

B. Approximate solution for PMFS-IS

In this section, we propose an approximately optimal strategy for Problems 3 and 4 under PMFS-IS. Let \( \bar{L} = \text{diag}(\bar{L}^1, \ldots, \bar{L}^K) \) denote a diagonal matrix with diagonal terms of \( \bar{L}^k \) defined as in Theorem 6. We impose the following assumption.

**Assumption (A6)** \( \sqrt{\beta}(A + \bar{B}L) \) is Hurwitz matrix.

Let \( \bar{J}_\beta \) denote the performance of strategy (33) where \( \bar{x}_t \) is replaced by \( z_t \) in (17) and \( J^*_\beta \) denote the optimal performance under MFS-IS. Then, the difference in performance \( \bar{J}_\beta - J^*_\beta \) is bounded. In particular, we have the following

**Theorem 7** Assume (A1a), (A2), (A5), (A6) and (PMFS-IS). Then, for \( \beta \in (0, 1] \), we have

1) The performance loss is given by

\[
\bar{J}_\beta - J^*_\beta = (1 - \beta) \text{Tr}\left(\bar{X}_1 \bar{M}\right) + \text{Tr}(\bar{W} \bar{M}),
\]

where \( \bar{X}_1 \) and \( \bar{W} \) are time-homogeneous and defined as in Theorem 2 and \( \bar{M} \) is the solution of following algebraic Lyapunov equation:

\[
\bar{M} = \text{DARE}_\beta(\bar{A}, \bar{Q}),
\]

where \( \bar{A} \) and \( \bar{Q} \) are defined as in Theorem 2 and \( \bar{L} = \text{diag}(\bar{L}^1, \ldots, \bar{L}^K) \) and \( \bar{L} \) are computed as in Theorem 6.

2) Let \( n = \min_{k \in \mathcal{N}}(|\mathcal{N}^k|) \). Under (A1b) and (A3),

\[
\bar{J}_\beta - J^*_\beta \in \mathcal{O}\left(\frac{1}{n}\right).
\]

**Proof:** The proof follows along the same lines of the proof of Theorem 2. In particular, under (A5) and (A6), \( \sqrt{\beta}\bar{A} \) of Proposition 2 is Hurwitz; hence, the performance loss may be computed by the associated algebraic Lyapunov equation given by (35). Note that even though \( \bar{Q} \) is not positive semi-definite, the algebraic Lyapunov equation has a solution [39]. The proof of part 2 of Theorem 7 follows from (34) and observation that (i) \( \bar{M} \) given by (35) does not depend on \( n \) due to (A3); (ii) \( (\bar{X}_1, \bar{W}) \) are \( \mathcal{O}(1/n) \) due to (A1b).

**Remark 11** Assumption (A6) is always satisfied if \( D^k_i = 0 \) and \( E^k_i = 0 \) for all \( k \in \mathcal{K} \). In this case, \( \sqrt{\beta}(A + \bar{B}L) \) is \( \text{diag}(\sqrt{\beta}(A^1 + B^1 \bar{L}^1), \ldots, \sqrt{\beta}(A^K + B^K \bar{L}^K)) \), where each of the diagonal terms are Hurwitz by definition of \( \bar{L}^k \) given in Theorem 6.

Note that an infinite-horizon discounted problem with 4-tuple \( (A, B, Q, R) \) and discount factor \( \beta \) is equivalent to an undiscounted problem with 4-tuple \( (\sqrt{\beta}A, \sqrt{\beta}B, Q, R) \).
VIII. NUMERICAL EXAMPLE

To illustrate our results, we consider an example that is motivated by demand response in power systems. In demand response, the volatility in renewable generation is compensated by making small changes in the demand of a large number of loads. We model the load dynamics according to a model proposed in [40], but consider a different per-step cost.

Consider a population $N$ of space heaters that can be partitioned into $K$ disjoint sub-populations $N^k$, $k \in K := \{1, \ldots, K\}$. Each sub-population corresponds to a particular type of space heater that have similar physical characteristics such as time response and nominal temperature. For space heater $i, i \in N$, the state $x_i^k$ denotes the room temperature at time $t$. Consider a nominal temperature $x_{\text{nom}}^k$ for sub-population $k$, $k \in K$, and let $u_{\text{nom}}^k$ be the control input needed to maintain the room temperature at $x_{\text{nom}}^k$. Following [40], we linearize the dynamics of sub-population $k$ around $x_{\text{nom}}^k$, i.e.,

$$x_{t+1}^i - x_{\text{nom}}^k = a^k(x_t^i - x_{\text{nom}}^k) + b^k u_t^i + w_t^i,$$

where $u_t^i$ is the control input in addition to $u_{\text{nom}}^k$ and $w_t^i$ is a random disturbance. We assume $u_{\text{nom}}^k$ is large enough such that $(u_t^i + u_{\text{nom}}^k)$ is positive.

Let $x_{\text{des}}^i$ denote the desired temperature of user $i$. It is assumed that the mean desired temperature $\bar{x}_{\text{des}} = \text{vec}(\bar{x}_{\text{des}}^1, \ldots, \bar{x}_{\text{des}}^K)$ is known to everyone (e.g., independent system operator (ISO) could compute it and broadcast the mean value to everyone or it could be computed in a distributed manner using a consensus algorithm). For the purpose of demand response, time is divided into epochs of length $T$. At the beginning of each epoch, a central authority such as an ISO generates a reference mean temperature $m_{\text{ref}}$ and broadcasts it to all users.

During an epoch, all users collectively minimize the total expected cost $\mathbb{E}[\sum_{t=1}^T c_t]$, where the per-step cost $c_t$ is given by

$$\frac{1}{|N|} \sum_{i \in N} \left[ q(x_t^i - x_{\text{des}}^i)^2 + ru_t^i \right] + \frac{t}{T} p(m_t - m_{\text{ref}})^2,$$

where $m_t = (\sum_{i \in N} x_t^i)/|N|$. The rationale for the per-step cost is that we penalize deviations from the desired temperature (which corresponds to the user’s comfort level), the control effort, and deviation of the mean temperature from the reference prescribed by the ISO. The weight $\frac{t}{T}$ is so that we linearly add more weight to meeting global preference.

The above problem is an optimal tracking problem and the optimal strategy is given by Theorem 3. As an example, we consider the following values of the parameters: $K = 2$, $p = 30$, $q = 2$, $r = 50$, $x_{\text{des}}^i = x_1^i \sim \text{Normal}(20, 3)$, $w_t^i \sim \text{Normal}(0, 0.01)$, and

$$|N|^1 = 40, \quad a_1 = 0.5, \quad b_1 = 1.5, \quad x_{\text{nom}}^1 = 20,$$

$$|N|^2 = 100, \quad a_2 = 0.8, \quad b_2 = 1.0, \quad x_{\text{nom}}^2 = 20.$$

and consider three epochs. In the first epoch, $1 \leq t \leq 50$, there is no reference signal and the space heaters are operating around their local set temperatures; in the second epoch, $50 < t \leq 150$, $m_{\text{ref}} = 21$; in the third epoch, $150 < t \leq 250$, $m_{\text{ref}} = 19$. The resultant trajectories of a subset of the users are shown in Fig. 1.

IX. CONCLUSION

We presented team optimal control of a decentralized system with partially exchangeable agents. Partial exchangeability implies that such a system is equivalent to one where the dynamics and the cost are coupled only through the mean-field. Our two main results are as follows. First, when the mean field is observed by all agents (the MFS information structure), the linear control laws are optimal and the corresponding gains are computed by solving $K + 1$ Riccati equations, where $K$ is the number of sub-populations. The dimensions of these Riccati equations are independent of the size of sub-populations; consequently, the solution complexity depends only on the number $K$ of sub-populations (rather than the size of the entire population). Second, when the mean-field of a (possibly empty) subset of sub-populations is observed by all agents (the PMFS information structure), a linear control law based on certainty equivalence is approximately optimal.

An important practical implication of these results is that they do not suffer from the curse of dimensionality. In fact, under assumption (A3), the solution does not even depend on the number of agents and the optimal gains can be computed without being aware of the size of each sub-population. Consequently, the solution methodology generalizes to the setup where the agents in a sub-population arrive and depart according to an exogenous process (e.g., number of electric vehicles plugged in for charging in smart grids).

The raison d’être for investigating decentralized systems is that it is not possible—either physically or economically—to send all the state observation to a centralized controller. We show that when agents are partially exchangeable, we may circumvent the conceptual difficulties of decentralized control and achieve the centralized performance by sharing only the mean-field. Moreover, in view of the results of PMFS-IS, one may even decide not to share the mean-field of large sub-populations because there is only a small loss in performance in using the approximate value of the mean-field instead.

Throughout this paper, we assumed that when the mean-
field is observed, it is observed without noise. In practice (especially if the mean-field is computed using a consensus algorithm), the mean-field will be observed with noise (and the noise will be different across agents). Our results show that if all sub-populations are large, such an observation noise will not matter. (In fact, the agents may completely ignore the mean-field observations and use the approximate values instead). However, if some of the sub-populations are small, the solution approach is not obvious. In particular, in the special case when all sub-populations have one agent, the problem reduces to the general decentralized control problem with non-classical information structure. Identifying a solution methodology for this general case remains a challenging research direction.

REFERENCES


A. Proofof Proposition 1

Let \( A_{i,j} \) denote the \((i, j)\)-th block of matrix \( A_1 \). We use a similar notation for other matrices as well. Fix a sub-population \( k \in K \). If we exchange agents \( i, j \in N^k \), then property 2 of exchangeability implies that \( A_{i,j} = A_{j,i} \) and for any other agent \( n \in N \), \( A_{i,n} = A_{j,n} \) and \( A_{n,j} = A_{n,i} \). (Similar relationships hold for \( B_1 \) as well). Property 3 implies that \( Q_{i,n} = Q_{j,n} \) and \( Q_{n,j} = Q_{n,i} \). (Similar relationships hold for \( R_i \) as well). Define these by \( a_k^{i,j} \) and \( b_k^{i,j} \), respectively.

For \( i, j \in N^k \), \( A_{i,j}^{1} = A_{j,i}^{1} \) and \( B_{i,j}^{1} = B_{j,i}^{1} \). Denote these by \( a_{i,j}^{1} \) and \( b_{i,j}^{1} \), respectively.

For \( i, j \in N^k \) and \( n, m \in A_i^k \), \( A_{i,n}^{1} = A_{j,m}^{1} \) and \( B_{i,m}^{1} = B_{j,n}^{1} \). Denote these by \( a_{i,n}^{1} \) and \( b_{i,n}^{1} \), respectively.

For \( i, j \in N^k \), \( Q_{i,n}^{1} = Q_{j,m}^{1} \) and \( R_{i,n}^{1} = R_{j,m}^{1} \). Denote these by \( q_{i,n}^{1} \) and \( r_{i,n}^{1} \), respectively.

For \( i, j \in N^k \) and \( n, m \in N^k \), \( Q_{i,n}^{1} = Q_{j,m}^{1} \) and \( R_{i,n}^{1} = R_{j,m}^{1} \). Denote these by \( p_{i,n}^{1,k,j} \) and \( p_{i,n}^{1,k,j} \), respectively.
Now, consider the dynamics according to (2), the dynamics of agent $i$ of sub-population $k$ can be written as

$$ x^i_{t+1} = A^{i,*} x_t + B^{i,*} u_t + w^i_t, \quad (36) $$

where $A^{i,*}$ and $B^{i,*}$ denote the rows corresponding to the $i$th block of $A_t$ and $B_t$. Note that

$$ A^{i,*} x_t = A^{i,*} x^i_t + \sum_{j \in \mathbb{N} \setminus i} A^{i,j,x^i_t} + \sum_{l \in K, l \neq k} \sum_{n \in \mathbb{N}^l} A^{i,n,x^i_t} = a^{i,*}_t x^i_t + d^{k,k}_t \sum_{j \in \mathbb{N} \setminus i} x^j_t + \sum_{l \in K, l \neq k} d^{k,l}_t \sum_{n \in \mathbb{N}^l} x^n_t = a^{i,*}_t x^i_t + d^{k,k}_t (|N^k| x^k_t - x^i_t) + \sum_{l \in K, l \neq k} d^{k,l}_t |N^l| x^l_t =: A^{k,*}_t x^i_t + \sum_{l \in K} D^{k,l}_t x^l_t, \quad (37) $$

where $A^k_t = a^k_t - d^{k,k}_t$ and $D^{k,l}_t = |N^l| d^{k,l}_t$. By a similar algebra, we can define $B^k_t$ and $E^{k,l}_t$ such that

$$ B^{k,*}_t u_t = B^k_t u^k_t + \sum_{l \in K} E^{k,l}_t u^l_t, \quad (38) $$

where $B^k_t = b^k_t - e^{k,k}_t$, and $E^{k,l}_t = |N^l| e^{k,l}_t$. Substituting (37) and (38) in (36), we get (5). Now consider the per-step cost given by (3). Note that

$$ x^i_t^T Q_i x^i_t = \sum_{k \in K} \sum_{l \in K} \sum_{i \in \mathbb{N}^k} \sum_{j \in \mathbb{N}^l} (x^i_t)^T Q^k_{i,l} x^j_t $$

$$ = \sum_{k \in K} \sum_{l \in K} \sum_{i \in \mathbb{N}^k} \sum_{j \in \mathbb{N}^l} (x^i_t)^T P^k_{i,l} x^j_t $$

$$ + \sum_{k \in K} \sum_{l \in K} \sum_{i \in \mathbb{N}^k} \sum_{j \in \mathbb{N}^l} (x^i_t)^T q^k_{i,l} x^j_t $$

$$ + \sum_{k \in K} \sum_{i \in \mathbb{N}^k} (x^i_t)^T q^k_{i,i}, \quad (39) $$

where $P^k_{i,l} = |N^k||N^l| p^{k,l}$ and $Q^k_{i,l} = |N^k| q^k_{i,l}$. By similar algebraic manipulation, we can show that

$$ u_t^T R_t u_t = u^k_t^T P^k_{u,l} u^k_t + \sum_{l \in K} \sum_{i \in \mathbb{N}^k} \frac{1}{|N^k|} (u^k_t)^T R^k_{u,l} u^k_t, \quad (40) $$

where $P^k_{u,l} = |N^k||N^l| p^{u,k,l}$ and $R^k_{u,l} = |N^k| r^k_{u,l}$. Substituting (39) and (40) in (3), we get (6).

### B. Proof of Corollary 3

Under the assumptions on the model, the dynamics, given by (25) and (26), simplify to

$$ \tilde{x}^i_{t+1} = A^k \tilde{x}^i_t + \tilde{u}^i_t, \quad \tilde{x}_{t+1} = \tilde{A}_t \tilde{x}_t + \tilde{B}_t \tilde{u}_t + \tilde{w}_t, \quad (41) $$

and $\tilde{c}_t(\tilde{x}_t, \tilde{u}_t)$ of Corollary 4 simplifies to

$$ \tilde{c}_t(\tilde{x}_t, \tilde{u}_t) = \tilde{x}_t^T (\tilde{Q}_t + \tilde{P}_t) \tilde{x}_t + \tilde{u}_t^T (\tilde{P}_u) \tilde{u}_t. \quad (41) $$

Thus, the $N$ subsystems corresponding to $\tilde{x}^i_t$ are uncontrolled and we need to identify $\tilde{u}_t$ to optimally control the dynamics of mean-field $\tilde{x}_t$ with per-step cost given by (41). Hence, the optimal solution is given by

$$ \tilde{u}_t = \tilde{L}_t \tilde{x}_t = \sum_{k \in K} \tilde{L}_k^k \tilde{x}^k_t, $$

where $\tilde{L}_t$ is computed as explained in Corollary 3. To complete the proof, note that if agent $i \in N^k$ of sub-population $k \in K$ chooses action $\tilde{u}^k_t = \theta^k_t \tilde{L}_k^k \tilde{x}^k_t$, then we get $\theta^k_t \tilde{u}^k_t = \tilde{L}_k^k \tilde{x}^k_t$; consequently, $\tilde{u}_t = \sum_{k \in K} \theta^k_t \tilde{u}^k_t = \sum_{k \in K} \tilde{L}_k^k \tilde{x}^k_t$.

### C. Proof of Theorem 3

As in the proof of Theorem 1 described in Section V, define $\tilde{x}_t = x_t - \tilde{x}_t^i, \tilde{u}_t = u_t - \tilde{u}_t^i, \tilde{x}_t = \text{vec}((\tilde{x}^i_t)_{i \in N^k}), \text{and} \ \tilde{u}_t = \text{vec}((\tilde{u}_t^i)_{i \in N^k})$. We identify a cost function $\{\tilde{c}_t^k\}_{k \in K}$ and $\tilde{c}_t$ as in Corollary 4.

**Corollary 5** For time $t$, $t \in \{1, \ldots, T\}$, there exist functions $\{\tilde{c}_t^k\}_{k \in K}$ and $\tilde{c}_t$ such that

$$ c_t(x_t, u_t, \bar{x}_t, \bar{u}_t) = \tilde{c}_t(x_t, \bar{u}_t) + \sum_{i \in \mathbb{N}^k, k \in K} \tilde{c}_t^k(\tilde{x}^i_t, \tilde{u}^i_t) $$

$$ - \sum_{k \in K} \tilde{c}_t^k(\tilde{x}^i_t, \tilde{u}^i_t) $$

and for $t = T$,

$$ c_T(x_t, \bar{x}_t) = \tilde{c}_T(\bar{x}_T) + \sum_{i \in \mathbb{N}^k, k \in K} \tilde{c}_T^k(\bar{x}^i_T, \bar{u}^i_T) - \sum_{k \in K} \tilde{c}_T^k(\bar{x}^i_T, \bar{u}^i_T). $$

To describe $\tilde{c}_t(\cdot)$, define $\bar{y}_t := \left[ \tilde{y}_t^T \tilde{y}_t^T \tilde{y}_t^T \right]$. Then,

$$ \tilde{c}_t(\bar{x}_t, \bar{u}_t) = \bar{y}_t^T \left[ \begin{array}{cc} \bar{Q}_T & 0 \\ 0 & \bar{P}_T \end{array} \right] \tilde{y}_T.$$

Moreover,

$$ \tilde{c}_T^k(\bar{x}^i_T, \bar{u}^i_T) = \frac{1}{|N^k|} \left[ (\bar{x}^i_T - \bar{r}^i_T)^T \bar{Q}_T (\bar{x}^i_T - \bar{r}^i_T) + (\bar{u}^i_T)^T \bar{R}_T \bar{u}^i_T \right], $$

$$ \tilde{c}_T^T(\bar{x}_T^T, \bar{r}_T^T) = \frac{1}{|N^k|} \left[ (\bar{x}_T^T - \bar{r}_T)^T \bar{Q}_T (\bar{x}_T^T - \bar{r}_T) \right]. $$

Then, define a centralized auxiliary system where the state is $\hat{x}_t = \text{vec}((\hat{x}^i_t)_{i \in N^k})$, action is $\hat{u}_t = \text{vec}((\hat{u}_t^i)_{i \in N^k})$, and the per-step cost is given by Corollary 5. Note that the per-step cost is decomposed into terms that depend only on $(\bar{x}_t, \bar{u}_t)$ and terms that depend only on $(\tilde{x}^i_t, \tilde{u}^i_t)$ (and terms that do not depend on the control strategy). The rest of the proof follows along the same lines of the proof of Theorem 1. In particular,
we consider a deterministic dynamical system and split it into $K + 1$ classes. The agents in class $k, k \in \mathcal{K},$ are solving a tracking problem whose solution is given by

$$\tilde{u}_i^k = \tilde{L}_k \tilde{x}_i^k + \tilde{F}_k v_i^k.$$  

The mean-field component is also solving a tracking problem whose solution is given by

$$\tilde{u}_i = \tilde{L}_i \tilde{x}_i + \tilde{F}_i \tilde{v}_i.$$  

The result of the Theorem follows from combining the above equations. Therefore, from standard results in LQR tracking problem, the optimal control law of agent $i \in \mathcal{N}_k$ of sub-population $k \in \mathcal{K}$ is given by

$$u_i = \bar{u}_i + \bar{u}_i^k = \left[ \tilde{L}_i \tilde{x}_i^k + \tilde{F}_i v_i^k \right] + \left[ \tilde{L}_i \tilde{x}_i + \tilde{F}_i \tilde{v}_i \right],$$

where gains $\{\tilde{L}_i, \tilde{L}_i^k, \tilde{F}_i, \tilde{F}_i^k\}_{i=1}^{T-1}$ are identical for all agents of sub-population $k$, $\tilde{u}_i$ is identical for all agents of all sub-populations, and $v_i^k$ may be different for each agent.

\[D. \ Proof \ of \ Theorem \ 4\]  

The proof follows the same lines as the proof of Theorem 1 with the following differences. The mean-field is defined as $\bar{x}_i^{k,\lambda} = \frac{1}{|N|} \sum_{t \in \mathcal{N}_k} \lambda x_i^k$ (similar interpretations hold for $\bar{u}_i^{k,\lambda}$ and $\bar{u}_i^{k,\lambda}$) and the breakeven variables are defined as $\bar{x}_i^k = x_i^k - \frac{\lambda_i^k}{\lambda_i} \bar{x}_i$ (similar interpretations hold for $\bar{u}_i^k$ and $\bar{u}_i^k$). Then, due to (44), the dynamics of $\bar{x}_i^k$ and $\bar{x}_i^{\lambda}$ are still given by (25) and (26), respectively, where $\bar{A}_i$ and $\bar{B}_i$ are defined as in Theorem 4.

The equivalent of Lemma 1 is the following:

\textbf{Lemma 7} Let $(\lambda^1, \ldots, \lambda^N) \in \mathbb{R}^N$ and $(b^1, \ldots, b^N) \in \mathbb{R}_+^N$. In addition, for any $x = \text{vec}(x^1, \ldots, x^N)$ and $\bar{x} = (\langle \lambda^k x^k \rangle)^{N}$, let $\bar{x} = x^1 - \frac{\lambda^1}{b^1} \bar{x}^1, \ldots, x^N - \frac{\lambda^N}{b^N} \bar{x}^N, \ i \in \{1, \ldots, N\}$. Then, for any matrix $Q$ of appropriate dimension,

$$\frac{1}{N} \sum_{i=1}^{N} b^i (x^i)^T Q x^i = \frac{1}{N} \sum_{i=1}^{N} b^i (\bar{x}^i)^T Q \bar{x}^i + (\bar{x}^{\lambda})^T \mu Q \bar{x}^{\lambda},$$

where $\mu := 2 - \frac{1}{N} \sum_{i=1}^{N} \frac{(\lambda^i)^2}{b^i}$.

Consequently, the equivalent of Corollary 4 is the following

\textbf{Corollary 6} \ For time $t$, $t \in \{1, \ldots, T\}$, there exist functions $\{c_k\}_{k \in \mathcal{K}}$ and $\bar{c}_t$ such that

$$c_t(x_t, u_t, x_\lambda^T, \bar{u}_t^T) = \bar{c}_t(\bar{x}_t^\lambda, \bar{u}_t^T) + \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{N}_k} \bar{c}_t(\bar{x}_t^i, \bar{u}_t^i),$$

where

$$\bar{c}_t(\bar{x}_t^\lambda, \bar{u}_t^T) = (\bar{x}_t^\lambda)^T (\bar{Q}_t + \bar{P}_t^\lambda) \bar{x}_t^\lambda + (\bar{u}_t^T)^T (\bar{R}_t + \bar{P}_t^u) \bar{u}_t^T,$n

$$\bar{c}_t(\bar{x}_t^i, \bar{u}_t^i) = b^i \left[ \frac{1}{|N|} \right] \left[ (\bar{x}_t^i)^T Q_i \bar{x}_t^i + (\bar{u}_t^i)^T R_i \bar{u}_t^i \right],$$

and for $t = T$,

$$c_T(x_T, x_T^\lambda) = \bar{c}_T(\bar{x}_T^\lambda) + \sum_{i \in \mathcal{N}_k, k \in \mathcal{K}} \bar{c}_T(\bar{x}_T^i),$$

where $\bar{Q}_t$ and $\bar{R}_t$ are defined as in Theorem 4.

The rest of the proof is the same as in Section V-C. We can show that the optimal control strategy of auxiliary model is given by

$$\bar{u}_i^\lambda = \tilde{L}_i \tilde{x}_i^\lambda \quad \text{and for} \quad k \in \mathcal{K}, \quad \bar{u}_i^k = \tilde{L}_i \tilde{x}_i^k,$$

where the gains $\{\tilde{L}_i^k, \tilde{L}_i\}_{i=1}^{T-1}$ are given as in Theorem 4. To complete the proof of Theorem 4, note that

$$u_i = \tilde{u}_i^k + \lambda_i^k \bar{u}_i^k = \tilde{L}_i \left( x_i^k - \frac{\lambda_i^k}{b^k} \bar{x}_i^k \right) + \frac{\lambda_i^k}{b^k} \tilde{L}_i \tilde{x}_i^k.$$

Thus, the control laws specified in Theorem 4 are the optimal centralized control laws, and, a fortiori, the optimal decentralized control laws.