Les Cahiers du GERAD ISSN: 0711–2440

Mean Field Linear Quadratic Teams

J. Arabneydi
A. Mahajan

G–2015–121

November 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Mean Field Linear Quadratic Teams

Jalal Arabneydi
Aditya Mahajan

GERAD & Department of Electrical and Computer Engineering, McGill University, Montreal (Quebec) Canada, H3A 0E9

j.arabneydi@gmail.com
aditya.mahajan@mcgill.ca

November 2015

Les Cahiers du GERAD
G–2015–121

Abstract: In this paper, we investigate team optimal control of a population of heterogeneous LQ (Linear Quadratic) agents. The population consists of finite distinct sub-populations so that agents in each sub-population are homogeneous. For each agent, the state evolves linearly (i.e. linear dynamics) and the cost is quadratic in state and action. The agents are coupled in both dynamics and cost through the empirical mean (also called mean-field) of states and actions of agents. Each agent observes its local state and the mean-field. This information structure is called mean-field sharing information structure and it is a non-classical decentralized information structure. The objective of agents is to team up with each other to minimize the total cost. We identify the team-optimal solution and show that it is unique and linear. The optimal gains are computed by the solution of appropriate Riccati equations. One of the key salient features of our approach is that the computational complexity of our solution does not depend on the number of agents, yet it depends on the number of sub-populations. This implies that the optimal strategy can be computed without any knowledge on the number of agents. We generalize our results to tracking problem, infinite horizon, and infinite population.

Acknowledgments: Authors gratefully acknowledge the support of Groupe d'études et de recherche en analyse des décisions (GERAD) and the Natural Sciences and Engineering Research Council of Canada through Grant NSERC-RGPIN 402753-11 for funding this research, without which the present study could not have been completed.

1 Introduction

In this paper, we study a class of team-optimal control systems that we call mean-field LQ (Linear Quadratic) teams. In general, mean-field LQ teams belong to the class of models that are broadly classified as mean-field LQG systems and emerge in various applications including smart grids [1], communication [2], finance [3], emergent behaviour [4], etc.

In mean-field LQ teams, the system consists of a finite number heterogeneous agents with linear dynamics and quadratic costs that are coupled through the empirical mean of states (also called the mean-field) and actions of agents. Every agent observes the local state of itself and the mean-field. The agents need to cooperate to minimize a common (mean-field coupled) cost. Mean-field LQ Teams emerge in diverse applications. For example, in smart grids, each demand (consumer) is influenced by the aggregate consumption of all demands in the grid. In particular, the aggregate consumption of demands has a direct impact on the power generation (in the grid) or the price (in the market) or both. The higher consumption, the higher generation or higher price or both. Hence, each demand is affected by the aggregate consumption (i.e. mean-field) of all demands. This scenario may be modelled as mean-field LQ teams.

In general, finding team-optimal solution of mean-field LQ teams is challenging. In particular,

1) Mean-field LQ teams are *conceptually* difficult due to the decentralized nature of information available to agents. The agents need to cooperate with each other to minimize a common cost function while they have different information. This discrepancy in information makes it difficult to establish cooperation among agents. In particular, the information structure is non-classical which, in general, is conceptually more difficult than other classes, i.e., classical and partial nested information structures. We refer reader to [5] for more details.
2) Mean-field LQ teams are *computationally* difficult for large number of agents because the computational complexity of finding the team-optimal solution increases exponentially with respect to the number of agents.

To best of our knowledge, there does not exist any general approach in the literature that provides the team-optimal solution of mean-field LQ teams for finite number of agents. However, there are few results in some special cases. For example, such a team-optimal solution is computed in [6], where agents are only coupled in the cost (no coupling in dynamics), variables are scalar, and total cost function is in the form of discounted infinite horizon.

Due to the above difficulties, most of the results in the literature focus attention on the case of countably infinite number of agents [6–13]. The key intuition behind these results is that in the case of countably infinite number of agents, the action of a single agent has no effect on the dynamics of the mean-field. This decoupling reduces the problem to one in which one generic agent is interacting with the mass. A consistent solution of this interaction provides an approximately optimal solution to the case with asymptotically large number of agents.

In this paper, we take an alternative approach and provide the team-optimal solution in general case. We assume every agent observes its own local state and the mean-field of the system. We analyze the problem with arbitrary number (not necessarily large) of agents and show that the optimal strategy is unique, identical within sub-populations, and linear in the local state and the mean-field. To compute the optimal gains, we derive decoupled Riccati equations that do not depend on the number of agents, thus the results are valid for *any* number of agents.

This paper is organized as follows. We present the model and problem formulation in Section 2.1 and main results in Section 2.2. We then describe two different variations of the model of Section 2.1 with their main results as follows: major-minor in Section 2.4 and tracking problem in Section 2.5. In Section 3, we provide the proofs of the results given in Section 2. We extend the main results to infinite horizon setup in Section 4 and infinite population in Section 5. At last, we conclude the paper in Section 6.

1.1 Notation

Given a set $\mathcal{A}$, $|\mathcal{A}|$ denotes its size. Given vectors x, y, z (of possibly different dimensions), $\text{vec}(x, y, z)$ denotes $[x^\intercal, y^\intercal, z^\intercal]^\intercal$. Superscripts index agents (indexed by i) or types (indexed by k). Given a set $\mathcal{N}$ of agents and states $x^i, i \in \mathcal{N}$, (all of same dimension), $\langle (x^i)_{i\in\mathcal{N}} \rangle$ denotes the mean-field $\frac{1}{|\mathcal{N}|}\sum_{i=1}^{|\mathcal{N}|} x^i$ of $(x^i)_{i\in\mathcal{N}}$. Given a set $\mathcal{A}$ of states x^n (of possibly different dimensions), $n \in \mathcal{A}$, we use bold letters $\mathbf{x}$ to denote $\mathbf{x} = \text{vec}(x^1, \dots, x^{|\mathcal{A}|})$. Given a random variable x, $\mathbb{E}[x]$ denotes its mean and $\text{var}(x)$ denotes its variance. Upper case letters A, B, etc. denote matrices; lower case letters x, y, etc. denote (column) vectors. Given a matrix A, $\text{Tr}(A)$ denotes its trace. Given a square matrix A, $A \geq 0$ (respectively $A > 0$) denotes that A is positive semi-definite (respectively positive definite). For any matrix A or vector x, $A^\intercal$ and $x^\intercal$ denote their transpose, respectively. We also use the short hand notation $x_{a:b}$ for $\text{vec}(x_a, x_{a+1}, \dots, x_b)$. $\mathbb{R}$ refers to the set of real numbers.

2 Finite population models and results

2.1 Model and problem formulation

Consider a heterogeneous population of N agents where each agent belongs to one of K possible types, $\{1, \dots, K\}$. Let $\mathcal{K} = \{1, \dots, K\}$ denote the set of types (of sub-populations) and for any $k \in \mathcal{K}$, $\mathcal{N}^k$ denote the sub-population of type k and $\mathcal{N}$ denote the entire population i.e. $\mathcal{N} = \cup_{k\in\mathcal{K}}\mathcal{N}^k$.

The state of agent i, $i \in \mathcal{N}$, is denoted by x_t^i and its action by u_t^i at time t. For type $k \in \mathcal{K}$ and agent $i \in \mathcal{N}^k$, the state x_t^i belongs to $\mathbb{R}^{d_x^k}$ and action u_t^i belongs to $\mathbb{R}^{d_u^k}$. For ease of notation, we denote the joint state by $\mathbf{x}_t = (x_t^i)_{i\in\mathcal{N}}$ and joint action by $\mathbf{u}_t = (u_t^i)_{i\in\mathcal{N}}$.

The *mean-field* of states[1] $\bar{x}_t^k$ of sub-population of type k, $k \in \mathcal{K}$, is the empirical mean of the states of all agents in that sub-population, i.e.,

$$\bar{x}_t^k := \langle (x_t^i)_{i\in\mathcal{N}^k} \rangle = \frac{1}{|\mathcal{N}^k|} \sum_{i\in\mathcal{N}^k} x_t^i, \quad k \in \mathcal{K}. \tag{1}$$

The mean-field of states of the entire population is denoted by $\bar{\mathbf{x}}_t$ as follows:

$$\bar{\mathbf{x}}_t = \text{vec}(\bar{x}_t^1, \dots, \bar{x}_t^K). \tag{2}$$

Similarly, the mean-field of actions $\bar{u}_t^k$ of sub-population of type k, $k \in \mathcal{K}$, is the empirical mean of the actions of all agents in that sub-population, i.e.,

$$\bar{u}_t^k := \langle (u_t^i)_{i\in\mathcal{N}^k} \rangle = \frac{1}{|\mathcal{N}^k|} \sum_{i\in\mathcal{N}^k} u_t^i, \quad k \in \mathcal{K}. \tag{3}$$

The mean-field of actions of the entire population is denoted by $\bar{\mathbf{u}}_t$ as follows:

$$\bar{\mathbf{u}}_t = \text{vec}(\bar{u}_t^1, \dots, \bar{u}_t^K). \tag{4}$$

A summary of the above notation is presented in Table 1 for ease of reference.

2.1.1 Dynamics

Agents of the same type have identical dynamics. The dynamics of agents are coupled through the mean-field of states and actions. In particular, for type $k \in \mathcal{K}$, the state of agent $i \in \mathcal{N}^k$ evolves as follows.

$$x_{t+1}^i = A_t^k x_t^i + B_t^k u_t^i + D_t^k \bar{\mathbf{x}}_t + E_t^k \bar{\mathbf{u}}_t + w_t^i, \tag{5}$$

where A_t^k, B_t^k, D_t^k, and E_t^k are matrices of appropriate dimensions, the initial state x_1^i is a random variable, and $\{w_t^i\}_{t=1}^T$ is a noise process. Let $\mathbf{w}_t = (w_t^i)_{i\in\mathcal{N}}$. We make the following assumptions on the primitive random variables:

[1]In the rest of the paper, we refer to mean-field of states simply as mean-field.

Table 1: Summary of the notation used in this paper

Notation used for agent $i \in \mathcal{N}^k$ of type $k \in \mathcal{K}$	
$x_t^i \in \mathbb{R}^{d_x^k}$	State of agent i
$u_t^i \in \mathbb{R}^{d_u^k}$	Action of agent i
Notation used for sup-population of type $k \in \mathcal{K} = \{1, \dots, K\}$	
$\mathcal{N}^k$	Entire sub-population of type k
$\bar{x}_t^k = \langle (x_t^i)_{i \in \mathcal{N}^k} \rangle$	Mean-field of states at time t
$\bar{u}_t^k = \langle (u_t^i)_{i \in \mathcal{N}^k} \rangle$	Mean-field of actions at time t
Notation used for entire population	
$\mathcal{N} = \bigcup_{k \in \mathcal{K}} \mathcal{N}^k$	Entire population
$\mathbf{x}_t = (x_t^i)_{i \in \mathcal{N}}$	Joint state of entire population at time t
$\mathbf{u}_t = (u_t^i)_{i \in \mathcal{N}}$	Joint action of entire population at time t
$\bar{\mathbf{x}}_t = \mathrm{vec}(\bar{x}_t^1, \dots, \bar{x}_t^K)$	Mean-field of states of entire population at t
$\bar{\mathbf{u}}_t = \mathrm{vec}(\bar{u}_t^1, \dots, \bar{u}_t^K)$	Mean-field of actions of entire population at t

Assumption (A1) *The primitive random variables $\{\mathbf{x}_1, \{\mathbf{w}_t\}_{t=1}^T\}$ are mutually independent.*

Assumption (A2) *For all $i \in \mathcal{N}$ and for all t, w_t^i has zero mean and finite variance; in addition, x_1^i has finite variance.*

Note that the initial joint state $\mathbf{x}_1$ and the joint noise $\mathbf{w}_t, t \geq 1$, may be correlated across agents. For some of the results, we assume a stronger version of (A1) as follows.

Assumption (A1′) *In addition to (A1), for all $i \in \mathcal{N}$, $(x_1^i)_{i \in \mathcal{N}}$ are independent and for each t, $(w_t^i)_{i \in \mathcal{N}}$ are independent. Also, for each type $k \in \mathcal{K}$, $(x_1^i)_{i \in \mathcal{N}^k}$ are identically distributed and for each t, $(w_t^i)_{i \in \mathcal{N}^k}$ are identically distributed.*

2.1.2 Per-step cost

At time t, the system incurs a cost that depends on the local state and action of the agents and the mean-field of states and actions as follows. For $t = 1, \dots, T - 1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \bar{\mathbf{x}}_t^\intercal P_t^x \bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal P_t^u \bar{\mathbf{u}}_t + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \left[{x_t^i}^\intercal Q_t^k x_t^i + {u_t^i}^\intercal R_t^k u_t^i \right] \right] \tag{6}$$

and $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) = \bar{\mathbf{x}}_T^\intercal P_T^x \bar{\mathbf{x}}_T + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} {x_T^i}^\intercal Q_T^k x_T^i \right], \tag{7}$$

where P_t^x, P_t^u, Q_t^k, and R_t^k are symmetric matrices of appropriate dimension that satisfy the following conditions:

$$\begin{aligned} Q_t^k &\geq 0, \qquad \forall k \in \mathcal{K}, \quad \mathrm{diag}\{Q_t^1, \dots, Q_t^K\} + P_t^x \geq 0, \\ R_t^k &> 0, \qquad \forall k \in \mathcal{K}, \quad \mathrm{diag}\{R_t^1, \dots, R_t^K\} + P_t^u > 0. \end{aligned} \tag{8}$$

Note that we do not require P_t^x and P_t^u to be semi-positive definite as long as above inequities hold.

2.1.3 Observation model and information structure

Agent i perfectly observes its local state x_t^i and the global mean-field $\bar{\mathbf{x}}_t$. Agents perfectly recall all the data they observe. Thus, agent i chooses action U_t^i as follows.

$$u_t^i = g_t^i(x_{1:t}^i, u_{1:t-1}^i, \bar{\mathbf{x}}_{1:t}). \tag{9}$$

We call the above observation model *mean-field sharing* information structure. The function g_t^i is called the *control law* of agent i. The collection $\mathbf{g}^i = (g_1^i, g_2^i, \dots, g_T^i)$ is called the *control strategy* of agent i. The collection $\mathbf{g} = (\mathbf{g}^i)_{i \in \mathcal{N}}$ is called the *control strategy* of the system.

The performance of strategy $\mathbf{g}$ is given by

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\Bigg[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)\Bigg], \tag{10}$$

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

2.1.4 The optimization problem

We are interested in the following optimization problem.

Problem 1 *In the model described above, find a strategy* $\mathbf{g}^*$ *that minimizes* (10), *i.e.,*

$$J^* := J^*(\mathbf{g}^*) = \inf_{\mathbf{g}} J(\mathbf{g}), \tag{11}$$

where the infimum is taken over all strategies of form (9).

We presented the model in its simplest form. The results presented below also apply to the following variations of the basic model.

1) The per-step cost has cross-terms of $(x_t^i, \bar{\mathbf{x}}_t)$ and $(u_t^i, \bar{\mathbf{u}}_t)$ as follows:

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \bar{\mathbf{x}}_t^\intercal P_t^x \bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal P_t^u \bar{\mathbf{u}}_t + \sum_{k \in \mathcal{K}} \Bigg[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \Big[{x_t^i}^\intercal Q_t^k x_t^i + {x_t^i}^\intercal S_t^{x,k} \bar{\mathbf{x}}_t + {u_t^i}^\intercal S_t^{u,k} \bar{\mathbf{u}}_t + {u_t^i}^\intercal R_t^k u_t^i\Big]\Bigg]$$

   This cost can be re-written in the form of (6) as follows:

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \bar{\mathbf{x}}_t^\intercal (P_t^x + S_t^x) \bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal (P_t^u + S_t^u) \bar{\mathbf{u}}_t + \sum_{k \in \mathcal{K}} \Bigg[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \Big[{x_t^i}^\intercal Q_t^k x_t^i + {u_t^i}^\intercal R_t^k u_t^i\Big]\Bigg],$$

   where

$$S_t^x := \operatorname{vec}(S_t^{x,1}, \dots, S_t^{x,K}), \quad S_t^u := \operatorname{vec}(S_t^{u,1}, \dots, S_t^{u,K}).$$

2) The per-step cost has cross-terms of (x_t^i, u_t^i), $(x_t^i, \bar{\mathbf{u}}_t)$, $(u_t^i, \bar{\mathbf{x}}_t)$ and $(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$. This can be treated in the same manner as cross-terms are treated in the centralized LQR.
3) The per-step cost is to minimize a tracking error. We give the details of this case in Section 2.5.

2.2 Main result

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

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

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t \tag{12}$$

   where the above gains are obtained by the solution of $K+1$ Riccati equations: one for computing each $\breve{L}_t^k$, $k \in \mathcal{K}$, and one for $\bar{L}_t := \operatorname{vec}(\bar{L}_t^1, \dots, \bar{L}_t^K)$.

2) Riccati equations: *Let*

$$\bar{A}_t := \operatorname{diag}(A_t^1, \dots, A_t^K) + \operatorname{vec}(D_t^1, \dots, D_t^K),$$
$$\bar{B}_t := \operatorname{diag}(B_t^1, \dots, B_t^K) + \operatorname{vec}(E_t^1, \dots, E_t^K),$$
$$\bar{Q}_t := \operatorname{diag}(Q_t^1, \dots, Q_t^K), \bar{R}_t := \operatorname{diag}(R_t^1, \dots, R^K).$$

For $t = 1, \dots, T-1$*:*

$$\breve{L}_t^k = -\left(B_t^{k\intercal} \breve{M}_{t+1}^k B_t^k + R_t^k\right)^{-1} B_t^{k\intercal} \breve{M}_{t+1}^k A_t^k \tag{13}$$

and

$$\bar{L}_t = -\left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t, \tag{14}$$

where $\{\breve{M}_t^k\}_{t=1}^T$ *and* $\{\bar{M}_t\}_{t=1}^T$ *are the solutions of following Riccati equations:*

$$\breve{M}_T^k = Q_T^k, \quad \bar{M}_T = \bar{Q}_T + P_T^x, \tag{15}$$

and for $t = T-1, \dots, 1$,

$$\breve{M}_t^k = -A_t^{k\intercal} \breve{M}_{t+1}^k B_t^k \left(B_t^{k\intercal} \breve{M}_{t+1}^k B_t^k + R_t^k\right)^{-1} B_t^{k\intercal} \breve{M}_{t+1}^k A_t^k + A_t^{k\intercal} \breve{M}_{t+1}^k A_t^k + Q_t^k, \tag{16}$$

and

$$\bar{M}_t = -\bar{A}_t^\intercal \bar{M}_{t+1} \bar{B}_t \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{A}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{Q}_t + P_t^x. \tag{17}$$

3) Optimal performance: *Let*

$$\breve{\Sigma}_t^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \operatorname{var}(w_t^i - \bar{w}_t^k), \quad \bar{\Sigma}_t := \operatorname{var}(\bar{\mathbf{w}}_t),$$
$$\breve{\Xi}^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \operatorname{var}(x_1^i - \bar{x}_1^k), \quad \bar{\Xi} := \operatorname{var}(\bar{\mathbf{x}}_1),$$
$$\breve{\mu}^i := \frac{1}{\sqrt{|\mathcal{N}^k|}} \mathbb{E}(x_1^i - \bar{x}_1^k), \quad \bar{\mu} := \mathbb{E}(\bar{\mathbf{x}}_1).$$

Then, the optimal cost is given by

$$J^* = \sum_{t=1}^T \left[\sum_{k \in \mathcal{K}} \operatorname{Tr}\left(\breve{\Sigma}_t^k \breve{M}_{t+1}^k\right) + \operatorname{Tr}(\bar{\Sigma}_t \bar{M}_{t+1})\right] + \sum_{k \in \mathcal{K}} \operatorname{Tr}\left(\breve{\Xi}^k \breve{M}_1^k\right) + \operatorname{Tr}(\bar{\Xi} \bar{M}_1) + \left[\sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{N}^k} \breve{\mu}^{i\intercal} \breve{M}_1^k \breve{\mu}^i\right] + \bar{\mu}^\intercal \bar{M}_1 \bar{\mu}. \tag{18}$$

To implement the optimal control strategies:

- all agents must compute $\bar{L}_{1:T-1}$ by solving the Riccati equation (15) and (17),
- agents of type k must compute $\breve{L}_{1:T-1}^k$ by solving the Riccati equation (15) and (16).

Then, an individual agent i of type k, upon observing the local state x_t^i and the global mean-field $\bar{\mathbf{x}}_t$, chooses its local control action according to (12).

Remark 1 *Under (A1′) and (A2), for each type* $k \in \mathcal{K}$*, let* Σ_t^k*,* Ξ^k*, and* μ^k *denote* $\operatorname{var}(w_t^i)$*,* $\operatorname{var}(x_1^i)$*, and* $\mathbb{E}(x_1^i)$*, respectively, for* $i \in \mathcal{N}^k$*. Then,*

$$\breve{\Sigma}_t^k = \frac{|\mathcal{N}^k| - 1}{|\mathcal{N}^k|} \Sigma_t^k, \quad \bar{\Sigma}_t = \operatorname{diag}(\Sigma_t^1, \dots, \Sigma_t^K),$$
$$\breve{\Xi}^k = \frac{|\mathcal{N}^k| - 1}{|\mathcal{N}^k|} \Xi^k, \quad \bar{\Xi} = \operatorname{diag}(\Xi^1, \dots, \Xi^K),$$
$$\breve{\mu}^k = 0, \quad \bar{\mu} = \operatorname{vec}(\mu^1, \dots, \mu^K).$$

2.3 Salient features of the result

1. The linear quadratic mean-field model presented in Section 2.1 is a decentralized system with non-classical information structure that is neither partially nested nor quadratic invariant; yet linear control laws are optimal.
2. All agents in a sub-population of a particular type have identical optimal control laws. Although the agents of the same type are exchangeable (i.e., if $i, j \in \mathcal{N}^k$, then interchanging i and j does not affect the dynamics or the cost), in general, it is not optimal to use identical control laws at exchangeable agents (see [14] for a counterexample).
3. The solution and the solution complexity depend on the number of types but not on the number of agents of each type. In particular, the Riccati equations of (15)–(17) do not depend on $|\mathcal{N}^k|, k \in \mathcal{K}$.
4. Consider the above model with a centralized information structure, i.e., at time t, all agents have access to $(\mathbf{x}_{1:t}, \mathbf{u}_{1:t-1})$. As part of the proof of Theorem 1 (see Section 3), we show that the optimal control laws under centralized information are implementable under mean-field sharing. Hence, the optimal decentralized performance, given by (18), is the same as the optimal centralized performance.
5. From an implementation point of view, the above feature has an interesting consequence. If we have the freedom to design the information structure, then there is no advantage of sharing anything beyond the mean-field. Note that the mean-field can be shared using distributed consensus algorithms.
6. When the number of agents for all sub-populations goes to infinity, the Riccati equations remain the same; however, the mean-field becomes a deterministic process that can be pre-computed (using $\bar{A}_{1:t}, \bar{B}_{1:t}, \bar{L}_{1:t}$). Therefore, the mean-field sharing information structure is informationally equivalent to the completely decentralized information structure (where agent i knows only $(x^i_{1:t}, u^i_{1:t-1})$). Thus, when every sub-population is infinite, the optimal control laws under completely centralized information (i.e. $(\mathbf{x}_{1:t}, \mathbf{u}_{1:t-1})$) are implementable under completely decentralized information structure (i.e. $(x^i_{1:t}, u^i_{1:t-1})$). We present a generalization of this model in Section 5.

2.4 Special case of a major agent and population of minor agents

Consider the model of Section 2.1 with the population of N (minor) agents and one additional agent, called the major agent. As in Section 2.1, the population of N minor agents consists of K types $\{1, \dots, K\}$.

For the minor agent i, $i \in \mathcal{N}$, the state is denoted by x^i_t and the action is denoted by u^i_t. For the major agent, the state is denoted by x^0_t and the action is denoted by u^0_t. We assume $x^0_t \in \mathbb{R}^{d^0_x}$ and $u^0_t \in \mathbb{R}^{d^0_u}$.

The mean-field of states and actions of minor agents are given by $\bar{\mathbf{x}}_t = \text{vec}(\bar{x}^1_t, \dots, \bar{x}^K_t)$ and $\bar{\mathbf{u}}_t = \text{vec}(\bar{u}^1_t, \dots, \bar{u}^K_t)$ respectively where $\bar{x}^k_t$ and $\bar{u}^k_t$ are given by (1) and (3), respectively.

2.4.1 Dynamics

The state of major agent evolves as follows.

$$x^0_{t+1} = A^0_t x^0_t + B^0_t u^0_t + D^0_t \bar{\mathbf{x}}_t + E^0_t \bar{\mathbf{u}}_t + w^0_t. \tag{19}$$

The state of minor agent i with type k, $i \in \mathcal{N}^k$, evolves as follows.

$$x^i_{t+1} = A^k_t x^i_t + B^k_t u^i_t + D^k_t \bar{\mathbf{x}}_t + E^k_t \bar{\mathbf{u}}_t + H^{x,k}_t x^0_t + H^{u,k}_t u^0_t + w^i_t. \tag{20}$$

2.4.2 Per-step cost

At time t, the system incurs a cost that depends on the local state, local action, mean-field of states and actions of minor agents and the local state and action of major agent as follows. For $t = 1, \dots, T-1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t, x_t^0, u_t^0) = {x_t^0}^\intercal Q_t^0 x_t^0 + {u_t^0}^\intercal R_t^0 u_t^0 + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \left[{x_t^i}^\intercal Q_t^k x_t^i + {u_t^i}^\intercal R_t^k u_t^i \right] \right] + \bar{\mathbf{x}}_t^\intercal P_t^x \bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal P_t^u \bar{\mathbf{u}}_t + 2{x_t^0}^\intercal P_t^{x,0} \bar{\mathbf{x}}_t + 2{u_t^0}^\intercal P_t^{u,0} \bar{\mathbf{u}}_t, \quad (21)$$

and $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T, x_T^0) = {x_T^0}^\intercal Q_T^0 x_T^0 + \sum_{k \in \mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} {x_T^i}^\intercal Q_T^k x_T^i \right] + \bar{\mathbf{x}}_T^\intercal P_T^x \bar{\mathbf{x}}_T + 2{x_T^0}^\intercal P_T^{x,0} \bar{\mathbf{x}}_T, \quad (22)$$

where $P_t^{x,0}$, P_t^x, $P_t^{u,0}$, P_t^u, Q_t^0, Q_t^k, R_t^0, and R_t^k are symmetric matrices of appropriate dimension. Let

$$\bar{Q}_t := \operatorname{diag}(Q_t^0, Q_t^1, \dots, Q_t^K), \quad \bar{R}_t := \operatorname{diag}(R_t^0, R_t^1, \dots, R_t^K),$$
$$\bar{P}_t^x := \begin{bmatrix} 0 & P_t^{x,0} \\ P_t^{x,0} & P_t^x \end{bmatrix}, \qquad \bar{P}_t^u := \begin{bmatrix} 0 & P_t^{u,0} \\ P_t^{u,0} & P_t^u \end{bmatrix}.$$

Then, above matrices satisfy the following conditions:

$$\begin{aligned} Q_t^k &\geq 0, \quad \forall k \in \mathcal{K}, \quad \bar{Q}_t + \bar{P}_t^x \geq 0, \\ R_t^k &> 0, \quad \forall k \in \mathcal{K}, \quad \bar{R}_t + \bar{P}_t^u > 0. \end{aligned} \quad (23)$$

2.4.3 Information structure

The major agent observes its local state and the mean-field of states of minor agents and chooses action according to

$$u_t^0 = g_t^0(x_{1:t}^0, u_{1:t-1}^0, \bar{\mathbf{x}}_{1:t}). \quad (24)$$

In addition to its local state, the minor agent i perfectly observes the mean-field of states of minor agents and the local state of major agent and chooses action according to

$$u_t^i = g_t^i(x_{1:t}^i, u_{1:t-1}^i, \bar{\mathbf{x}}_{1:t}, x_{1:t}^0). \quad (25)$$

The performance of joint strategy $(\mathbf{g}, \mathbf{g}^0)$, where $\mathbf{g}^0 := \{g_1^0, \dots, g_T^0\}$, is given by

$$J_{MM}(\mathbf{g}, \mathbf{g}^0) = \mathbb{E}^{(\mathbf{g}, \mathbf{g}^0)} \left[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t, x_t^0, u_t^0) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T, x_T^0) \right] \quad (26)$$

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy $(\mathbf{g}, \mathbf{g}^0)$. We are interested in the following optimization problem.

Problem 2 *In the model described above, find a joint strategy* $(\mathbf{g}^*, {\mathbf{g}^0}^*)$ *that minimizes* (26), *i.e.,*

$$J_{MM}^* := J_{MM}^*(\mathbf{g}^*, {\mathbf{g}^0}^*) = \inf_{\mathbf{g}, \mathbf{g}^0} J_{MM}(\mathbf{g}, \mathbf{g}^0). \quad (27)$$

where the infimum is taken over all strategies of the form (24) *and* (25).

A variation of the above model was first introduced in [8] and other variations have been investigated in [15–18].

Corollary 1 *Under (A1) and (A2), we have the following results for Problem 2.*

1) Structure of optimal strategy: *The optimal strategy for Problem 2 is unique and is linear in local state, the mean-field of the minor agents, and the state of major agent. In particular,*

$$u_t^0 = \hat{L}_t^0 x_t^0 + \bar{L}_t^0 \bar{\mathbf{x}}_t, \tag{28}$$

and

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t + \hat{L}_t^k x_t^0, \tag{29}$$

where the above gains are computed by the solution of $K+1$ Riccati equations: one for computing each $\breve{L}_t^k$, $k \in \mathcal{K}$ (which are the same as in Theorem 1), and one for $L_t := \begin{bmatrix} \hat{L}_t^0 & \bar{L}_t^0 \\ \hat{L}_t^1 & \bar{L}_t^1 \\ \vdots & \vdots \\ \hat{L}_t^K & \bar{L}_t^K \end{bmatrix}$.

2) Riccati equations: *The Riccati equations for $\breve{L}_t^k$ are the same as in Theorem 1. Let*

$$\bar{A}_t := \text{diag}\{A_t^0, A_t^1, \dots, A_t^K\} + \begin{bmatrix} 0 & D_t^0 \\ H_t^{x,1} & D_t^1 \\ \vdots & \vdots \\ H_t^{x,K} & D_t^K \end{bmatrix},$$

$$\bar{B}_t := \text{diag}\{B_t^0, B_t^1, \dots, B_t^K\} + \begin{bmatrix} 0 & E_t^0 \\ H_t^{u,1} & E_t^1 \\ \vdots & \vdots \\ H_t^{u,K} & E_t^K \end{bmatrix}.$$

For $t = 1, \dots, T-1$:

$$L_t = -\left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + \bar{P}_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t, \tag{30}$$

where $\{\bar{M}_t\}_{t=1}^T$ is the solution of the following Riccati equation:

$$\bar{M}_T = \bar{Q}_T + \bar{P}_T^x. \tag{31}$$

and for $t = T-1, \dots, 1$,

$$\bar{M}_t = -\bar{A}_t^\intercal \bar{M}_{t+1} \bar{B}_t \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + \bar{P}_t^u\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{A}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{Q}_t + \bar{P}_t^x. \tag{32}$$

3) Optimal performance: *Let $\breve{\Sigma}_t^k$, $\breve{\Xi}^k$, $\mu^i, i \in \mathcal{N}^k, k \in \mathcal{K}$, be defined as in Theorem 1. Let*

$$\bar{\Sigma}_t := \text{var}(\text{vec}(w_t^0, \bar{\mathbf{w}}_t)), \quad \bar{\Xi} := \text{var}(\text{vec}(x_1^0, \bar{\mathbf{x}}_1)),$$
$$\bar{\mu} := \mathbb{E}(\text{vec}(x_1^0, \bar{\mathbf{x}}_1)).$$

Then, the optimal cost is given by

$$J_{MM}^* = \sum_{t=1}^T \left[\sum_{k\in\mathcal{K}} \text{Tr}\left(\breve{\Sigma}_t^k \breve{M}_{t+1}^k\right) + \text{Tr}(\bar{\Sigma}_t \bar{M}_{t+1})\right] + \sum_{k\in\mathcal{K}} \text{Tr}\left(\breve{\Xi}^k \breve{M}_1^k\right)$$
$$+ \text{Tr}(\bar{\Xi}\bar{M}_1) + \left[\sum_{k\in\mathcal{K}} \sum_{i\in\mathcal{N}^k} \breve{\mu}^{i\intercal} \breve{M}_1^k \breve{\mu}^i\right] + \bar{\mu}^\intercal \bar{M}_1 \bar{\mu}. \tag{33}$$

The proof is presented in Section 3.3.

2.5 Generalization to tracking cost function

Consider a tracking problem in which we are given a tracking signal $\{s_t^k\}_{t=1}^T$, $s_t^k \in \mathbb{R}^{d_x^k}$ for the mean-field of type $k, k \in \mathcal{K}$, and a tracking signal $\{r_t^i\}_{t=1}^T$, $r_t^i \in \mathbb{R}^{d_x^k}$, for each agent $i \in \mathcal{N}^k$.

Define $\bar{r}_t^k := \langle (r_t^i)_{i\in\mathcal{N}^k}\rangle, k \in \mathcal{K}$, $\bar{\mathbf{r}}_t := \text{vec}(\bar{r}_t^1 \ldots, \bar{r}_t^K)$, and $\mathbf{s}_t = \text{vec}(s_t^1, \ldots, s_t^K)$. The tracking cost is as follows. For $t = 1, \ldots, T-1$,

$$c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = (\bar{\mathbf{x}}_t - \mathbf{s}_t)^\intercal P_t^x (\bar{\mathbf{x}}_t - \mathbf{s}_t) + \bar{\mathbf{u}}_t^\intercal P_t^u \bar{\mathbf{u}}_t + \sum_{k\in\mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i\in\mathcal{N}^k} \left[(x_t^i - r_t^i)^\intercal Q_t^k (x_t^i - r_t^i) + {u_t^i}^\intercal R_t^k u_t^i \right] \right],$$

and for $t = T$,

$$c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) = (\bar{\mathbf{x}}_T - \mathbf{s}_T)^\intercal P_T^x (\bar{\mathbf{x}}_T - \mathbf{s}_T) + \sum_{k\in\mathcal{K}} \left[\frac{1}{|\mathcal{N}^k|} \sum_{i\in\mathcal{N}^k} (x_T^i - r_T^i)^\intercal Q_T^k (x_T^i - r_T^i) \right].$$

We assume that, in addition to the observation specified in Section 2.1.3, agent i also knows $\{r_t^i, \bar{\mathbf{r}}_t, \mathbf{s}_t\}_{t=1}^T$. The rest of the model is the same as in Section 2.1. The performance of strategy $\mathbf{g}$ is given by

$$J_T(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T) \right], \tag{34}$$

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

Problem 3 *In the model described above, find a strategy $\mathbf{g}^*$ that minimizes (34), i.e.,*

$$J_T^* := J_T^*(\mathbf{g}^*) = \inf_{\mathbf{g}} J_T(\mathbf{g}), \tag{35}$$

where the infimum is taken over all strategies of form (9).

Theorem 2 *Under (A1) and (A2), we have the following results for Problem 3.*

1) Structure of optimal strategy: *The optimal strategy for Problem 3 is unique and is linear in local state and the mean-field of the system. In particular,*

$$u_t^i = \breve{L}_t^k (x_t^i - \bar{x}_t^k) + \bar{L}_t^k \bar{\mathbf{x}}_t + \breve{F}_t^k v_t^i + \bar{F}_t^k \bar{v}_t, \tag{36}$$

where the above gains are obtained by the solution of $K+1$ Riccati equations defined in Theorem 1. In particular, gains $\{\breve{L}_t^k, \bar{L}_t^k\}_{t=1}^{T-1}$ are the same as in Theorem 1.

2) Riccati equations: *Let $\{\breve{M}_t^k\}_{t=1}^T$ and $\{\bar{M}_t\}_{t=1}^T$ be the solution of (K+1) Riccati equations defined in Theorem 1. For $t = 1, \ldots, T-1$:*

$$\breve{F}_t^k = \left({B_t^k}^\intercal \breve{M}_{t+1}^k B_t^k + R_t^k \right)^{-1} {B_t^k}^\intercal, \tag{37}$$

and

$$\bar{F}_t = \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P_t^u \right)^{-1} \bar{B}_t^\intercal, \tag{38}$$

where $\bar{F}_t =: \text{vec}(\bar{F}_t^1, \ldots, \bar{F}_t^K)$. For $t = T$,

$$v_T^i = Q_T^k r_T^i, \quad \bar{v}_T = \bar{Q}_T \bar{\mathbf{r}}_T + P_T^x \mathbf{s}_T \tag{39}$$

and for $t = T-1, \ldots, 1$,

$$v_t^i = (A_t^k - B_t^k \breve{L}_t^k)^\intercal v_{t+1}^i + Q_t^k r_t^i \tag{40}$$

and

$$\bar{v}_t = (\bar{A}_t - \bar{B}_t \bar{L}_t)^\intercal \bar{v}_{t+1} + \bar{Q}_t \bar{\mathbf{r}}_t + P_t^x \mathbf{s}_t. \tag{41}$$

3) Optimal performance: *For* $t = T$*,*

$$\alpha^i_T = {r^i_T}^\intercal Q^k_T r^i_T, \quad \bar{\alpha}_T = \bar{\mathbf{r}}_T^\intercal \bar{Q}_T \bar{\mathbf{r}}_T + {\mathbf{s}_T}^\intercal P^x_T \mathbf{s}_T, \tag{42}$$

and for $t = T-1, \ldots, 1$*,*

$$\alpha^i_t = -2{v^i}^\intercal_{t+1} B^k_t ({B^k_t}^\intercal \breve{M}^k_{t+1} B^k_t + R^k_t)^{-1} {B^k_t}^\intercal v^i_{t+1} + {r^i_t}^\intercal Q^k_t r^i_t + \alpha^i_{t+1} \tag{43}$$

and

$$\bar{\alpha}_t = -2\bar{v}^\intercal_{t+1} \bar{B}_t (\bar{B}^\intercal_t \bar{M}_{t+1} \bar{B}_t + \bar{R}_t + P^u_t)^{-1} \bar{B}^\intercal_t \bar{v}_{t+1} + \bar{\mathbf{r}}^\intercal_t \bar{Q}_t \bar{\mathbf{r}}_t + {\mathbf{s}_t}^\intercal P^x_t \mathbf{s}_t + \bar{\alpha}_{t+1}. \tag{44}$$

Then,

$$J^*_T = J^* + \bar{\alpha}_1 + \sum_{k \in \mathcal{K}} \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \alpha^i_1 - \sum_{k \in \mathcal{K}} {\bar{r}^k_t}^\intercal Q^k_t \bar{r}^k_t. \tag{45}$$

The proof is presented in Section 3.4. To implement the optimal control strategies:

- all agents must compute $\bar{L}_{1:T-1}$ and $barF_{1:T-1}$ by solving the Riccati equation (15) and (17) and compute the global reference trajectory $\bar{v}_{1:T}$ by solving the backward equation (39) and (41),
- agents of type k must compute $\breve{L}^k_{1:T-1}$ and $\breve{F}^k_{1:T-1}$ by solving the Riccati equation (15) and (16),
- an individual agent i of type k computes a local reference trajectory $v^i_{1:T}$ by solving the backward equation (39) and (40).

Then, an individual agent i of type k, upon observing the local state x^i_t and the global mean-field $\bar{\mathbf{x}}_t$, chooses its local control action according to (36).

3 Proof of the result for finite population

The main idea of the proof is as follows. We construct an auxiliary system whose state, control actions, and per-step cost are equivalent to $\mathbf{x}_t$, $\mathbf{u}_t$, and $c_t(\cdot)$, respectively (modulo a change of variables that we describe later). However, this auxiliary system is centrally controlled by a single agent that has access to all the information available to the N decentralized agents in the original system. We show that the optimal centralized solution of this auxiliary system can be implemented in the original decentralized system, and is therefore also optimal for the decentralized system.

3.1 The auxiliary system

Define $\breve{x}^i_t = x^i_t - \bar{x}^k_t$ and $\breve{u}^i_t = u^i_t - \bar{u}^k_t$. The auxiliary system is a centralized system with state $\mathring{\mathbf{x}}_t = \text{vec}((\breve{x}^i_t)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$ and action $\mathring{\mathbf{u}}_t = \text{vec}((\breve{u}^i_t)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$. Note that $\mathring{\mathbf{x}}_t$ is equivalent to $\mathbf{x}_t$ and $\mathring{\mathbf{u}}_t$ is equivalent to $\mathbf{u}_t$.

The dynamics are the same as the model in Section 2. In particular,

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t + \breve{w}^i_t, \tag{46}$$

where $\breve{w}^i_t := w^i_t - \bar{w}^k_t$ and $\bar{w}^k_t := \langle (w^i_t)_{i \in \mathcal{N}^k} \rangle$ and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_t + \bar{\mathbf{w}}_t \tag{47}$$

where $\bar{\mathbf{w}}_t := \text{vec}(\bar{w}^1_t, \ldots, \bar{w}^K_t)$. The per-step cost of the auxiliary model is given by $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ at $t \le T-1$ and terminal cost $c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)$ at $t = T$. In the auxiliary system, there is a *single centralized agent* that chooses $\mathring{\mathbf{u}}_t$ based on the observations. In particular, the centralized agent observes $\mathring{\mathbf{x}}_t$ and chooses $\mathring{\mathbf{u}}_t$ according to

$$\mathring{\mathbf{u}}_t = \mathring{g}_t(\mathring{\mathbf{x}}_{1:t}, \mathring{\mathbf{u}}_{1:t-1}). \tag{48}$$

The performance of strategy $\mathring{\mathbf{g}} := (\mathring{g}_1, \dots, \mathring{g}_T)$ is given by

$$\mathring{J}(\mathring{\mathbf{g}}) = \mathbb{E}^{\mathring{\mathbf{g}}}\Big[\sum_{t=1}^{T-1} c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)\Big], \tag{49}$$

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

Problem 4 *In the auxiliary model, find strategy $\mathring{\mathbf{g}}^*$ that minimizes (49), i.e.,*

$$\mathring{J}^* := \mathring{J}^*(\mathring{\mathbf{g}}^*) = \inf_{\mathring{\mathbf{g}}} \mathring{J}(\mathring{\mathbf{g}}), \tag{50}$$

where the infimum is taken over all strategies of the form (48).

Let J^* and $\mathring{J}^*$ denote the optimal cost for Problem 1 and Problem 4, respectively. Since the per-step cost is the same in both cases, but Problem 4 is centralized, we have that $J^* \geq \mathring{J}^*$. 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.

A critical step in the proof is to rewrite the per-step cost $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ and terminal cost $c_T(\mathbf{x}_T, \bar{\mathbf{x}}_T)$ in terms of $\mathring{\mathbf{x}}_t$ and $\mathring{\mathbf{u}}_t$. For that matter, we need the following key result.

Lemma 1 *For any $\mathbf{x} = \operatorname{vec}(x^1, \dots, x^N)$ and $\bar{x} = \langle \mathbf{x} \rangle$, let $\breve{x}^i = x^i - \bar{x}$, $i \in \{1, \dots, N\}$. Then, for any matrix Q of appropriate dimension,*

$$\frac{1}{N}\sum_{i=1}^{N} {x^i}^\intercal Q x^i = \frac{1}{N}\sum_{i=1}^{N} {\breve{x}^i}^\intercal Q \breve{x}^i + \bar{x}^\intercal Q \bar{x}. \tag{51}$$

Proof. The result follows from elementary algebra and the observation that $\sum_{i=1}^{N} \breve{x}^i = 0$. □

Note that Lemma 1 is similar to Huygens–Steiner theorem in physics-mechanics [19].

An immediate consequence of Lemma 1 is the following:

Corollary 2 *For any time t, $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t)$ such that for $t = 1, \dots, T-1$,*

$$\mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t) = \bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}_t^k(\breve{x}_t^i, \breve{u}_t^i),$$

and $t = T$,

$$\mathring{c}_T(\mathring{\mathbf{x}}_T) = \bar{c}_T(\bar{\mathbf{x}}_T) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}_T^k(\breve{x}_T^i),$$

where for $t = 1, \dots, T-1$,

$$\begin{aligned}
\bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) &= \bar{\mathbf{x}}_t^\intercal(\bar{Q}_t + P_t^x)\bar{\mathbf{x}}_t + \bar{\mathbf{u}}_t^\intercal(\bar{R}_t + P_t^u)\bar{\mathbf{u}}_t, \\
\breve{c}_t^k(\breve{x}_t^i, \breve{u}_t^i) &= \frac{1}{|\mathcal{N}^k|}\left[{\breve{x}_t^i}^\intercal Q_t^k \breve{x}_t^i + {\breve{u}_t^i}^\intercal R_t^k \breve{u}_t^i\right],
\end{aligned}$$

and $t = T$,

$$\begin{aligned}
\bar{c}_T(\bar{\mathbf{x}}_T) &= \bar{\mathbf{x}}_T^\intercal(\bar{Q}_T + P_T^x)\bar{\mathbf{x}}_T, \\
\breve{c}_T^k(\breve{x}_T^i) &= \frac{1}{|\mathcal{N}^k|}\left[{\breve{x}_T^i}^\intercal Q_T^k \breve{x}_T^i\right].
\end{aligned}$$

Note that the auxiliary model has linear dynamics and in Corollary 2 we have shown that the cost is quadratic in the state and the control actions. Thus, the optimal control actions are linear in the state and the corresponding optimal gains can be obtained by solving an appropriate Riccati equation. However, the size of state $\mathring{\mathbf{x}}_t$ of the auxiliary system increases with the number of agents (e.g. N), thus, a naive attempt to obtain an optimal solution will involve solving for $\mathcal{O}(N^2)$ dimensional Riccati equations. We present an alternative approach in the next section that involves solving $K+1$ Riccati equations whose dimensions are independent of N.

3.2 The optimal solution of the auxiliary system

The auxiliary system is a stochastic linear quadratic system. From the certainty equivalence principle [20], we know that the optimal control law is unique and identical to the control law in the corresponding deterministic problem, whose dynamics are given by

$$\breve{x}^i_{t+1} = A^k_t \breve{x}^i_t + B^k_t \breve{u}^i_t, \tag{52}$$

and

$$\bar{\mathbf{x}}_{t+1} = \bar{A}_t \bar{\mathbf{x}}_t + \bar{B}_t \bar{\mathbf{u}}_{\mathbf{t}}, \tag{53}$$

and the per-step cost is $\mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t)$ given by Corollary 2.

Note that this system consists on $(N+1)$ components: N components with state $\breve{x}^i_t$ and action $\breve{u}^i_t$, $i \in \mathcal{N}$, and one component with state $\bar{\mathbf{x}}_t$ and action $\bar{\mathbf{u}}_t$. The first N components are split into K classes of identical components – one for each type. The components have decoupled dynamics and decoupled cost. Thus, the optimal control law of each class may be identified separately. Therefore, we have the following:

Theorem 3 *The optimal control strategy of auxiliary model is unique and given by*

$$\breve{u}^i_t = \breve{L}^k_t \breve{x}^i_t, \quad \bar{\mathbf{u}}_t = \bar{L}_t \bar{\mathbf{x}}_t, \quad i \in \mathcal{N}^k, k \in \mathcal{K}, \tag{54}$$

where the gains $\{\breve{L}^k_t\}_{t=1}^{T-1}$ *and* $\{\bar{L}_t\}_{t=1}^{T-1}$ *are given as in Theorem 1.*

To complete the proof of Theorem 1, note that

$$u^i_t = \breve{u}^i_t + \bar{u}^k_t = \breve{L}^k_t (x^i_t - \bar{x}^k_t) + \bar{L}^k_t \bar{\mathbf{x}}_t.$$

Thus, the control laws specified in Theorem 1 are the optimal *centralized* control laws, and, a fortrori, the optimal decentralized control laws.

3.3 Proof of Corollary 1

The major-minor model may be viewed as a special case of the model of Section 2.1. If we consider the major agent as a sub-population of a different type, say type 0, then the mean-field $\bar{x}^0_t$ of type 0 is x^0_t because $|\mathcal{N}^0| = 1$. Thus, the mean-field of the entire population is $\text{vec}(x^0_t, \bar{\mathbf{x}}_t)$. Consequently, the dynamics (19) and (20) are of the form (5); the cost (21) and (22) are of the form (6) and (7), respectively, and the information structure (24) and (25) is same as (9). Thus, we can directly use Theorem 1 to solve Problem 2. The direct use of Theorem 1 will give $K+2$ Riccati equations. However, one of these is redundant because one of the types (types 0) has a sub-population of size 1. In particular, when constructing the auxiliary system in the proof of Theorem 1 in Section 3, $\breve{x}^0_t := x^0_t - \bar{x}^0_t$ becomes zero by definition. Therefore, the solution is given by $K+1$ Riccati equations as shown in Corollary 1.

3.4 Proof of Theorem 2

As in the proof of Theorem 1 described in Section 3, define $\breve{x}^i_t = x^i_t - \bar{x}^k_t$, $\breve{u}^i_t = u^i_t - \bar{u}^k_t$, $\mathring{\mathbf{x}}_t = \text{vec}((\breve{x}^i_t)_{i \in \mathcal{N}}, \bar{\mathbf{x}}_t)$, and $\mathring{\mathbf{u}}_t = \text{vec}((\breve{u}^i_t)_{i \in \mathcal{N}}, \bar{\mathbf{u}}_t)$. We identify a cost function $\mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t)$ as in Corollary 2.

Lemma 2 *For any time t, $c_t(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t)$ such that for $t = 1, \dots, T-1$,*

$$\mathring{c}_t(\mathring{\mathbf{x}}_t, \mathring{\mathbf{u}}_t) = \bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}^k_t(\breve{x}^i_t, \breve{u}^i_t) - \sum_{k \in \mathcal{K}} \bar{r}^{k\intercal}_t Q^k_t \bar{r}^k_t,$$

and $t = T$,

$$\mathring{c}_T(\mathring{\mathbf{x}}_T) = \bar{c}_T(\bar{\mathbf{x}}_T) + \sum_{i \in \mathcal{N}^k, k \in \mathcal{K}} \breve{c}^k_T(\breve{x}^i_T) - \sum_{k \in \mathcal{K}} \bar{r}^{k\intercal}_T Q^k_T \bar{r}^k_T,$$

where for $t = 1, \dots, T-1$,

$$\bar{c}_t(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \bar{\mathbf{x}}_t - \begin{bmatrix} \bar{\mathbf{r}}_t \\ \mathbf{s}_t \end{bmatrix} \right)^\intercal \begin{bmatrix} \bar{Q}_t & 0 \\ 0 & P^x_t \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \bar{\mathbf{x}}_t - \begin{bmatrix} \bar{\mathbf{r}}_t \\ \mathbf{s}_t \end{bmatrix} \right) + \bar{\mathbf{u}}^\intercal_t (\bar{R}_t + P^u_t) \bar{\mathbf{u}}_t,$$
$$\breve{c}^k_t(\breve{x}^i_t, \breve{u}^i_t) = \frac{1}{|\mathcal{N}^k|} \left[(\breve{x}^i_t - r^i_t)^\intercal Q^k_t (\breve{x}^i_t - r^i_t) + \breve{u}^{i\intercal}_t R^k_t \breve{u}^i_t \right].$$

and $t = T$,

$$\bar{c}_T(\bar{\mathbf{x}}_T) = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \bar{\mathbf{x}}_T - \begin{bmatrix} \bar{\mathbf{r}}_T \\ \mathbf{s}_T \end{bmatrix} \right)^\intercal \begin{bmatrix} \bar{Q}_T & 0 \\ 0 & P^x_T \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \bar{\mathbf{x}}_T - \begin{bmatrix} \bar{\mathbf{r}}_T \\ \mathbf{s}_T \end{bmatrix} \right),$$
$$\breve{c}^k_T(\breve{x}^i_T) = \frac{1}{|\mathcal{N}^k|} \left[(\breve{x}^i_T - r^i_T)^\intercal Q^k_T (\breve{x}^i_T - r^i_T) \right].$$

Note that per-step cost is decomposed into terms that depend only on $(\bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)$ and terms that depend only on $(\breve{x}^i_t, \breve{u}^i_t)$ (and terms that do not depend on the control strategy). The rest of the proof follows along the same lines as the proof of Theorem 1. In particular, the auxiliary system consists of $N+1$ components; N components with state $\breve{x}^i_t$ and action $\breve{u}^i_t$, $i \in \mathcal{N}$, and one component with state $\bar{\mathbf{x}}_t$ and action $\bar{\mathbf{u}}_t$. The first N components are split into K classes. All agents in a class have identical dynamics and similar tracking cost but have different reference trajectory. Therefore, from standard results in LQR tracking problem, the optimal control law of agent $i \in \mathcal{N}^k$ of type $k \in \mathcal{K}$ is given by

$$u^i_t = \breve{u}^i_t + \bar{u}^k_t = \left[\breve{L}^k_t (x^i_t - \bar{x}^k_t) + \breve{F}^k_t v^i_t \right] + \left[\bar{L}^k_t \bar{\mathbf{x}}_t + \bar{F}^k_t \bar{v}_t \right],$$

where gains $\{\breve{L}^k_t, \bar{L}^k_t, \breve{F}^k_t, \bar{F}^k_t\}_{t=1}^{T-1}$ are identical for all agents of type k, $\bar{v}_t$ is identical for all agents of all types, and v^i_t is separate for each agent.

4 Infinite horizon

The results presented in Section 2 generalize to infinite horizon setup in a natural manner. Assume that the model and the cost are time-invariant, i.e., $\{A^k_t, B^k_t, D^k_t, E^k_t, Q^k_t, R^k_t, P^x_t, P^u_t\}$ do not depend on time; hence, we remove the subscript t. The rest of the model is as same as that in Section 2.1. Consider the infinite horizon long-term average and the infinite horizon discounted cost setups as follows:

Problem 5 *Find a strategy $\mathbf{g}$ that minimizes the following cost:*

$$\tilde{J}(\mathbf{g}) = \lim_{T \to \infty} \mathbb{E}^{\mathbf{g}} \left[\frac{1}{T} \sum_{t=1}^{T} c(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t) \right], \tag{55}$$

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

Problem 6 *Given discount factor $\beta \in (0,1)$, find a strategy $\mathbf{g}$ that minimizes the following cost:*

$$\tilde{J}_\beta(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\left[\sum_{t=1}^{\infty} \beta^{t-1} c(\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{x}}_t, \bar{\mathbf{u}}_t)\right], \tag{56}$$

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

Assumption (A-AC) *For each type k, (A^k, B^k) are stabilizable. In addition, $(\bar{A}, \bar{B})$ defined in Theorem 1 are stabilizable.*

Theorem 4 *Under (A1), (A2), and (A-AC), we have the following results for Problem 5.*

1) Structure of optimal strategy: *The optimal strategy for Problem 5 is unique and is linear in local state and the mean-field of the system. In particular,*

$$u_t^i = \breve{L}^k(x_t^i - \bar{x}_t^k) + \bar{L}^k \bar{\mathbf{x}}_t \tag{57}$$

where the above gains are obtained by the solution of $K+1$ algebraic Riccati equations. one for computing each $\breve{L}^k, k \in \mathcal{K}$, and one for $\bar{L} := \operatorname{vec}(\bar{L}^1, \dots, \bar{L}^K)$.

2) Algebraic Riccati equations: *Let $\bar{A}, \bar{B}, \bar{Q}$, and $\bar{R}$ be defined as in Theorem 1. For $t = 1, \dots, T-1$:*

$$\breve{L}^k = -\left(B^{k\intercal} \breve{M}^k B^k + R^k\right)^{-1} B^{k\intercal} \breve{M}^k A^k \tag{58}$$

and

$$\bar{L} = -\left(\bar{B}^\intercal \bar{M} \bar{B} + \bar{R} + P^u\right)^{-1} \bar{B}^\intercal \bar{M} \bar{A}, \tag{59}$$

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

$$\breve{M}^k = -A^{k\intercal} \breve{M}^k B^k \left(B^{k\intercal} \breve{M}^k B^k + R^k\right)^{-1} B^{k\intercal} \breve{M}^k A^k + A^{k\intercal} \breve{M}^k A^k + Q^k, \tag{60}$$

and,

$$\bar{M} = -\bar{A}^\intercal \bar{M} \bar{B} \left(\bar{B}^\intercal \bar{M} \bar{B} + \bar{R} + P^u\right)^{-1} \bar{B}^\intercal \bar{M} \bar{A} + \bar{A}^\intercal \bar{M} \bar{A} + \bar{Q} + P^x. \tag{61}$$

3) Optimal performance: *Let*

$$\breve{\Sigma}^k := \frac{1}{|\mathcal{N}^k|} \sum_{i \in \mathcal{N}^k} \operatorname{var}(w_t^i - \bar{w}_t^k), \quad \bar{\Sigma} := \operatorname{var}(\bar{\mathbf{w}}_t). \tag{62}$$

Then, the optimal cost is given by

$$\tilde{J}^* = \sum_{k \in \mathcal{K}} \operatorname{Tr}\left(\Sigma^k \breve{M}^k\right) + \operatorname{Tr}(\bar{\Sigma} \bar{M}). \tag{63}$$

Proof. The proof follows along the same lines as the proof of Theorem 1. We construct an auxiliary system as in Section 3, which consists of $N+1$ components with decoupled dynamics and cost. Since the cost is infinite-horizon long run average, the optimal solution is given by appropriate algebraic Riccati equations. □

Assumption (A-Dis) *For each type k, $(\sqrt{\beta} A^k, \sqrt{\beta} B^k)$ are stabilizable. In addition, $(\sqrt{\beta}\bar{A}, \sqrt{\beta}\bar{B})$ are stabilizable, where $(\bar{A}, \bar{B})$ are defined in Theorem 1.*

Theorem 5 *Under (A1), (A2), and (A-Dis), we have the following results for Problem 6.*

1) Structure of optimal strategy: *The optimal strategy for Problem 6 is unique and is linear in local state and the mean-field of the system. In particular,*

$$u_t^i = \breve{L}^k(x_t^i - \bar{x}_t^k) + \bar{L}^k \bar{\mathbf{x}}_t \tag{64}$$

where the above gains are obtained by the solution of $K+1$ *algebraic Riccati equations: one for computing each* $\breve{L}^k, k \in \mathcal{K}$*, and one for* $\bar{L} := \text{vec}(\bar{L}^1, \ldots, \bar{L}^K)$.

2) Algebraic Riccati equations: *Let* $\bar{A}, \bar{B}, \bar{Q}$*, and* $\bar{R}$ *be defined as in Theorem 1. Then, optimal gains* $\breve{L}^k$ *and* $\bar{L}$ *are computed as in Theorem 4 where* $\breve{M}^k$ *and* $\bar{M}$ *are the solutions of the following algebraic Riccati equations:*

$$\breve{M}^k = -\beta A^{k\intercal} \breve{M}^k B^k \left(B^{k\intercal} \breve{M}^k B^k + \beta^{-1} R^k\right)^{-1} B^{k\intercal} \breve{M}^k A^k + \beta A^{k\intercal} \breve{M}^k A^k + Q^k, \tag{65}$$

and,

$$\bar{M} = -\beta \bar{A}^\intercal \bar{M} \bar{B} \left(\bar{B}^\intercal \bar{M} \bar{B} + \beta^{-1}\left(\bar{R} + P^u\right)\right)^{-1} \bar{B}^\intercal \bar{M} \bar{A} + \beta \bar{A}^\intercal \bar{M} \bar{A} + \bar{Q} + P^x. \tag{66}$$

3) Optimal performance: *Let* $\breve{\Sigma}^k$ *and* $\bar{\Sigma}$ *be as defined in (62) and* $\breve{\Xi}^k, \bar{\Xi}, \breve{\mu}^i$*, and* $\bar{\mu}$ *de defined as in Theorem 1. Then, the optimal cost is given by*

$$\tilde{J}_\beta^* = \frac{1}{1-\beta}\left[\sum_{k\in\mathcal{K}} \text{Tr}\left(\breve{\Sigma}^k \breve{M}^k\right) + \text{Tr}(\bar{\Sigma}\bar{M})\right] + \sum_{k\in\mathcal{K}} \text{Tr}\left(\breve{\Xi}^k \breve{M}_1^k\right) + \text{Tr}(\bar{\Xi}\bar{M}_1) + \left[\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{N}^k} \breve{\mu}^{i\intercal} \breve{M}_1^k \breve{\mu}^i\right] + \bar{\mu}^\intercal \bar{M}_1 \bar{\mu}. \tag{67}$$

Proof. The proof is as same as that of Theorem 4. □

5 Infinite population

Consider the scenario when the population is asymptotically large, yet the number of sub-populations (types) is finite. Let $\hat{\mathcal{K}} \subseteq \mathcal{K}$ denote a set of sub-populations (types) that are asymptotically large, i.e., $|\mathcal{N}^k| = \infty, k \in \hat{\mathcal{K}}$. We model this scenario by making the following assumptions.

Assumption (A-Indep) *At time* t*, for each type* $k \in \hat{\mathcal{K}}$*, the noises* $\{w_t^i\}_{i\in\mathcal{N}^k}$ *are i.i.d. random variables.*

Assumption (A-Inf) *Every agent knowns the initial joint mean-field* $\bar{\mathbf{x}}_1$.

We state some of the results under the following stronger assumption.

Assumption (A′-Inf) *In addition to (A-Inf), all sub-populations are asymptotically large, i.e.,* $\hat{\mathcal{K}} = \mathcal{K}$.

Consider an information structure, that we call *partial mean-field sharing* that is smaller than mean-field sharing information structure (but equivalently informative) as follows. In partial mean-field sharing, agent i observes the local state x_t^i and the mean-field of finite sub-populations i.e. $(\bar{x}_t^k)_{k\in\mathcal{K}\setminus\hat{\mathcal{K}}}$. Thus, agent i chooses action according to

$$u_t^i = g_t^i(x_{1:t}^i, u_{1:t-1}^i, \{(\bar{x}_\tau^k)_{k\in\mathcal{K}\setminus\hat{\mathcal{K}}}\}_{\tau=1}^t). \tag{68}$$

Theorem 6 *Under (A1), (A2), (A-Indep), and (A-Inf), the optimal control laws of Problem 1 given by Theorem 1 (under mean-field sharing information structure) are implementable under partial mean-field sharing information structure. In particular, define a process* $\{\mathbf{z}_t\}_{t=1}^T, \mathbf{z}_t := \text{vec}(z_t^1, \ldots, z_t^K)$*, that is adapted to* $\{(\bar{x}_t^k)_{k\in\mathcal{K}\setminus\hat{\mathcal{K}}}\}_{t=1}^T$ *as follows. For every* $k \in \mathcal{K}$,

$$z_t^k = \begin{cases} \bar{x}_t^k, & k \in \mathcal{K}\setminus\hat{\mathcal{K}} \\ A_{t-1}^k z_{t-1}^k + (B_{t-1}^k \bar{L}_{t-1}^k + D_{t-1}^k)\bar{\mathbf{z}}_{t-1}, & k \in \hat{\mathcal{K}}. \end{cases} \tag{69}$$

Then, the optimal control laws for Problem 1 are given by

$$u_t^i = \breve{L}_t^k(x_t^i - \bar{z}_t^k) + \bar{L}_t^k \bar{\mathbf{z}}_t, \tag{70}$$

where $\{\breve{L}_t^k, \bar{L}_t^k\}_{t=1}^{T-1}$ *are same as in Theorem 1.*

Proof. Due to the law of large numbers, an immediate consequence of (A-Indep) is that $\bar{w}_t^k = \langle (w_t^i)_{i\in\mathcal{N}^k}\rangle, k \in \hat{\mathcal{K}}$, converges to its mean (i.e. zero by virtue of (A2)) as $|\mathcal{N}^k|, k \in \hat{\mathcal{K}}$, goes to infinity. Therefore, given the optimal control laws of Problem 1, mean-field $\bar{x}_t^k, k \in \hat{\mathcal{K}}$, evolves as follows:

$$\bar{x}_{t+1}^k = A_t^k \bar{x}_t^k + B_t^k \bar{L}_t^k \bar{\mathbf{x}}_t + D_t^k \bar{\mathbf{x}}_t. \tag{71}$$

At $t = 1$, every agent knowns $\bar{\mathbf{x}}_1$ according to (A-Inf). At $t > 1$, every agent observes $(\bar{x}_t^k)_{k\in\mathcal{K}\setminus\hat{\mathcal{K}}}$ and computes $\bar{x}_t^k, k \in \hat{\mathcal{K}}$, by using (71) and one-step delayed $\bar{\mathbf{x}}_{t-1} = \text{vec}(\bar{x}_{t-1}^1, \ldots, \bar{x}_{t-1}^K)$ that is known to every agent by time $t > 1$. Hence, mean-field sharing information structure is constructable by partial mean-field sharing; consequently, optimal control laws of Problem 1 given by Theorem 1 (under mean-field sharing information structure) are implementable under partial mean-field sharing information structure as well. □

Corollary 3 *Under (A1), (A2), (A-Indep), and (A′-Inf), for Problem 1 the optimal control law is unique and given by* (12). *Moreover,* $\bar{\mathbf{x}}_t = \text{vec}(\bar{x}_t^1, \ldots, \bar{x}_t^K)$ *evolves deterministically as follows:*

$$\bar{x}_{t+1}^k = A_t^k \bar{x}_t^k + (B_t^k \bar{L}_t^k + D_t^k)\bar{\mathbf{x}}_t. \tag{72}$$

Corollary 3 implies that under (A1), (A2), (A-Indep), and (A′-Inf), the optimal solution may be interpreted as follows.

$$u_t^i = \breve{L}_t^k x_t^i + \alpha_t^k, \tag{73}$$

where α_t^k is a deterministic process (that depends on the optimal gains $\{(\breve{L}_\tau^k, \bar{L}_\tau^k)_{k\in\mathcal{K}}\}_{\tau=1}^t$ and the initial joint mean-field $\bar{\mathbf{x}}_1$). Thus, the optimal solution can be implemented under a completely decentralized information structure i.e. one in which agent i only observes x_t^i (and does not observe the mean-field $\bar{\mathbf{x}}_t$). Under this information structure, the optimal control law has the interpretation that each agent is implementing the solution of a tracking problem.

6 Conclusion

In this paper, we presented a class of decentralized control systems that we call mean-field LQ teams. Finding a team-optimal solution for mean-field LQ teams is conceptually and computationally difficult because the information structure (i.e. mean-field sharing) is a non-classical decentralized information structure. To overcome these difficulties, we took the following steps. At the first step, we constructed an auxiliary system that has access to the complete centralized information. However, a naive attempt to solve the auxiliary system involves solving Riccati equations of the size of population; hence, when population is large, the solution will be computationally very expensive. For that matter, at the second step, we used an alternative approach to solve the auxiliary system. The obtained Riccati equations do not depend on the size of population and only depend on the number of sub-populations. At the last step, we showed that the optimal strategy of centralized auxiliary system is implementable in the original decentralized system (i.e. under mean-field sharing); hence, the obtained centralized optimal solution is also optimal for the decentralized system. First implication of these results is that the decentralized performance and centralized performance are equal. Second implication is that the optimal control strategy can be computed without any knowledge on the size of population.

In particular, we identified the team-optimal solution and proved that it is unique and linear in local state and (global) mean-field. We extended our results to tracking problem, infinite horizon, and infinite population. When every population is asymptotically large, we showed that the obtained optimal strategy can be implemented under completely decentralized information structure (where each agent only observes its local state and action and does not observe mean-field).

Note that all the results of this paper are also applicable to the continuous-time. Under natural assumptions, the obtained results may be modified to continuous time by replacing the discrete-time Riccati equations with their continuous-time counterparts.

References

[1] A.C. Kizilkale and R.P. Malhame, Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads, 19th World Congress, The International Federation of Automatic Control, 1867–1872, 2014.

[2] J. Moon and T. Basar, Discrete-time LQG mean field games with unreliable communication, 2014 IEEE 53rd Annual Conference on Decision and Control (CDC), 2697–2702, 2014.

[3] A.C. Kizilkale, S. Mannor, and P.E. Caines, Large scale real-time bidding in the smart grid: A mean field framework, 2012 IEEE 51st Annual Conference on Decision and Control (CDC), 3680–3687, 2012.

[4] M. Nourian, P.E. Caines, R.P. Malhamé, and M. Huang, Leader-follower Cucker-Smale type flocking synthesized via mean field stochastic control theory, in Brain, Body and Machine, Volume 83 of the series Advances in Intelligent and Soft Computing, 283–298, 2010.

[5] A. Mahajan, N.C. Martins, M.C. Rotkowitz, and S. Yuksel, Information structures in optimal decentralized control, 2012 IEEE 51st Annual Conference on Decision and Control (CDC), 1291–1306, 2012.

[6] M. Huang, P.E. Caines, and R.P. Malhamé, Social optima in mean field LQG control: Centralized and decentralized strategies, Automatic Control, IEEE Transactions on, 57(7), 1736–1751, 2012.

[7] M. Huang, P.E. Caines, and R.P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria, Automatic Control, IEEE Transactions on, 52(9), 1560–1571, 2007.

[8] M. Huang, Large-population LQG games involving a major player: The Nash certainty equivalence principle, SIAM Journal on Control and Optimization, 48(5), 3318–3353, 2010.

[9] R. Elliott, X. Li, and Y.-H. Ni, Discrete time mean-field stochastic linear-quadratic optimal control problems, Automatica, 49(11), 3222–3233, 2013.

[10] T. Li and J.-F. Zhang, Asymptotically optimal decentralized control for large population stochastic multiagent systems, Automatic Control, IEEE Transactions on, 53(7), 1643–1660, 2008.

[11] A. Bensoussan, K. Sung, and S. Yam, Linear–quadratic time-inconsistent mean field games, Dynamic Games and Applications, 3(4), 537–552, 2013.

[12] O. Guéant, J.-M. Lasry, and P.-L. Lions, Mean field games and applications, Paris-Princeton Lectures on Mathematical Finance 2010, Springer-Verlag Berlin Heidelberg 2011, 205–266.

[13] A. Bensoussan, J. Sung, P. Yam, and S.P. Yung, Linear-quadratic mean field games, arXiv preprint arXiv:1404.5741, 2014.

[14] J. Arabneydi and A. Mahajan, Team optimal control of coupled subsystems with mean-field sharing, 2014 IEEE 53rd Annual Conference on Decision and Control (CDC), 1669–1674, 2014.

[15] P.E. Caines and A.C. Kizilkale, Mean field estimation for partially observed lqg systems with major and minor agents, Proceedings of the 19th IFAC Word Congress, 8705–8709, 2014.

[16] S.L. Nguyen and M. Huang, Mean field lqg games with a major player: Continuum parameters for minor players, 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), 1012–1017, 2011.

[17] S.L. Nguyen and M. Huang, Mean field lqg games with mass behavior responsive to a major player, 2012 IEEE 51st Annual Conference on Decision and Control (CDC), 5792–5797, 2012.

[18] J. Huang, S. Wang, and Z. Wu, Mean field linear-quadratic-gaussian (LQG) games: Major and minor players, arXiv preprint arXiv:1403.3999, 2014.

[19] T.R. Kane and D.A. Levinson, Dynamics, Theory and Applications, McGraw Hill Series in Mechanical Engineering, 1985.

[20] P.E. Caines, Linear Stochastic Systems, John Wiley & Sons, Inc., 1987.