Team Optimal Control of Coupled Major-Minor Subsystems with Mean-Field Sharing

Jalal Arabneydi and Aditya Mahajan

Abstract— In this paper, we investigate team optimal control of coupled major-minor subsystems with mean-field sharing. In such a model, there is one major subsystem that directly influences the dynamics of n homogeneous minor subsystems; however, the minor subsystems influence the dynamics of the major subsystem and each other only through their mean behaviour (indirectly). In this model, the major and the minor subsystems are arbitrarily coupled in the cost and have mean-field sharing information structure. We propose a two-step approach. In the first step, we describe a mean-field sharing model with multiple types of minor subsystems as a generalization of the mean-field sharing model of Arabneydi and Mahajan, CDC 2014. In the second step, we use the results obtained in the first step to construct a dynamic programming decomposition to identify optimal strategies for the major and the minor subsystems. We present an example with numerical results to illustrate the approach.

I. INTRODUCTION

A. Motivation

Team theory—which is motivated by applications in networked control systems, economics, communication networks, sensor networks, robotics, and transportation networks, etc. — investigates the optimal design of decentralized systems consisting of coupled subsystems controlled by multiple controllers that have access to different information but need to collaborate to accomplish a common task.

Research in team theory began with with the seminal work of Marschack and Radner [1] in the context of organizational behaviour. Dynamic team problems have been investigated in systems and control following the seminal work of Witsenhansuen [2], [3] and others on dynamic team problems. The general dynamic team problem belongs in the NEXP complexity class [4]. As such, most of the research has focused on specific information structures that are amenable to analysis. We refer the reader to [5], [6] for detailed surveys.

Another approach to analyzing large-scale decentralized systems is mean-field approximation which arose in statistical physics, specially in applications consisting of a large number of interacting particles [7], [8]. In such large-scale systems, a single particle has negligible effect on others and is influenced by other particles only through their mean effect. This property may be exploited to reduce many body problems into one body problems. This idea from statistical

physics has been introduced in dynamical systems by Huang, Caines, and Mahlame [9]–[11] and Lasry and Lions [12], [13]. Similar ideas have also been used in large population games [14]. We refer the reader to [15] for a survey.

In real world scenarios, there are many applications where one big (major) subsystem can directly influence many small (minor) subsystems while small subsystems can only influence the big subsystem by their average (collective) behaviour. For example, consider a service-provider as the major subsystem and its users as the minor subsystems. The service-provider may directly influence each user by the cost of service. On the other hand, the average number of happy or unhappy users (collective behaviour) may affect the service-provider's reputation (and hence, its utility). In this paper, we study team optimal control of such systems by generalizing our previous results on mean-field sharing [16].

The rest of this paper is organized as follows. In Section II, we present three models of mean-field sharing: *basic mean-field sharing*, *mean-field sharing with major-minor* subsystems, and *mean-field sharing with multiple-types*. For brevity, in the sequel, we call these models the basic MF, MF-MM, and MF-T models, respectively. The reason for considering the MF-T model is that the MF-MM model may be viewed as a special case of the MF-T model; hence the main results of the MF-MM model can be obtained from the main results of the MF-T model. In Section III, we state the main results of the basic MF model in [16] and extend the solution methodology to the MF-T and MF-MM models. In Section IV, we present an example of the MF-MM model and in Section V, we conclude this paper.

B. Notation

We use upper-case letters to denote random variables (e.g. X) and lower-case letters to denote their realizations (e.g. x). Short-hand notation $X_{a:b}$ is used for the vector $(X_a, X_{a+1}, \ldots, X_b)$ and bold letter is used to denote vectors e.g. $\mathbf{Y} = (Y^1, \ldots, Y^n)$ where n is the size of vector $\mathbf{Y}$. $\mathbb{P}(\cdot)$ is probability of an event, $\mathbb{1}(\cdot)$ is indicator function of a set, and $\mathbb{E}[\cdot]$ is expectation of a random variable. $\mathbb{N}$ and $\mathbb{R}^+$ refer to the set of natural numbers and positive real numbers, respectively.

II. MEAN FIELD MODELS

A. Basic Mean-Field Sharing Model

We start with the description of the mean-field sharing model presented in [16]. We refer to this model as the basic mean-field (MF) model.

Jalal Arabneydi and Aditya Mahajan are with Department of Electrical Engineering, McGill University, 3480 University St., Montreal, QC, Canada. `jalal.arabneydi@mail.mcgill.ca` and `aditya.mahajan@mcgill.ca`

Consider a discrete time decentralized control system with $n \in \mathbb{N}$ homogeneous subsystems. Let $X_t^i \in \mathcal{X}$ and $U_t^i \in \mathcal{U}$ denote the state and the control action of subsystem i, $i \in \{1, \ldots, n\}$, at time $t, t \in \mathbb{N}$, respectively. $\mathcal{X}$ and $\mathcal{U}$ are finite sets that do not depend on i. The *mean-field* of the system refers to the empirical distribution of the state of all subsystems. The mean-field at time t, denoted by Z_t, is given by

$$Z_t = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_t^i} \qquad (1)$$

where δ_x denotes a Dirac measure on $\mathcal{X}$ with a point mass at x. Subsystem i evolves as follows:

$$X_{t+1}^i = f_t(Z_t, X_t^i, U_t^i, W_t^i), \quad i \in \{1, \ldots, n\} \qquad (2)$$

where f_t denotes the dynamics at time t (that do not depend on i) and $\{W_t^i\}_{t=1}^T$, $i \in \{1, \ldots, n\}$, are independent processes. The *primitive random variables* $(X_1^1, \ldots, X_1^n, \{W_t^1\}_{t=1}^T, \ldots, \{W_t^n\}_{t=1}^T)$ are mutually independent and defined on a common probability space. The probability distribution on X_1^i and W_t^i do not depend on i.

At time t, controller i observes the history of the mean-field $Z_{1:t}$ and its own local state X_t^i. Hence, the information set of controller i, $i \in \{1, \ldots, n\}$, at time t is

$$I_t^i = \{Z_{1:t}, X_t^i\}.$$

Controller i chooses the control action $U_t^i \in \mathcal{U}$ according to

$$U_t^i = g_t^i(Z_{1:t}, X_t^i) \qquad (3)$$

where g_t^i is the *control law* of controller i.

To simplify the solution and ensure fairness between all controllers, we assume the following.

Assumption 1 *At any time t, the control laws at all controllers are identical i.e. $g_t^i = g_t^j$ for any $i, j \in \{1, \ldots, n\}$. Therefore, we drop the superscripts and denote the control law at every controller at time t as g_t.*

As illustrated in [16], in general, Assumption 1 leads to a loss in performance. Nonetheless, it is a standard assumption in the literature on large-scale systems for reasons of simplicity, fairness, and robustness.

The system runs for a finite time horizon T, $T \in \mathbb{N}$. We refer to the collection of control laws $\mathbf{g} := (g_1, \ldots, g_T)$ over time T as the *control strategy* of the system. Let $\mathcal{G}_{BMF}$ be a set of all such control strategies.

At time t, the system incurs a cost that depends on joint state $\mathbf{X}_t := (X_t^1, \ldots, X_t^n)$ and joint action $\mathbf{U}_t := (U_t^1, \ldots, U_t^n)$ as follows

$$\ell_t(\mathbf{X}_t, \mathbf{U}_t). \qquad (4)$$

The performance of any strategy $\mathbf{g}$ is measured by the following expected total cost

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[\sum_{t=1}^{T} \ell_t(\mathbf{X}_t, \mathbf{U}_t) \right] \qquad (5)$$

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

Problem 1 *For the basic MF model described above, identify a control strategy $\mathbf{g}^*$ to minimize the expected total cost $J(\mathbf{g})$ given by (5).*

In [16], the basic MF model was analyzed using the common information approach of [17]. It was shown that the mean-field Z_t is an information state for an appropriately defined coordinated system. Using this information state, an appropriate dynamic programming decomposition to identify an optimal control strategy was derived.

B. Mean-Field Sharing with Major-Minor Subsystems (MF-MM) Model

In this paper, we investigate a setup where, in addition to n homogeneous subsystems (minor subsystems), there is one major subsystem that directly affects the evolution of all minor subsystems and the cost. We call this model, *mean-field sharing with major and minor subsystems* (MF-MM for short). This model is motivated by the scenarios described in Section I-A. In mean-field games, a similar model was introduced by [18] and further analyzed in [11].

As before, let $X_t^i \in \mathcal{X}$ and $U_t^i \in \mathcal{U}$ denote the state and the control action of minor subsystem $i, i \in \{1, \ldots, n\}$, at time t, $t \in \mathbb{N}$. In addition, let $X_t^0 \in \mathcal{X}^0$ and $U_t^0 \in \mathcal{U}^0$ denote the state and the control action of the *major subsystem* at time t. The sets $\mathcal{X}^0$, $\mathcal{X}$, $\mathcal{U}^0$, and $\mathcal{U}$ are finite. Let $Z_t = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_t^i}$ denote the mean-field of the *minor subsystems*. The dynamics of the major and minor subsystems are given by

$$X_{t+1}^0 = f_t^0(Z_t, X_t^0, U_t^0, W_t^0),$$
$$X_{t+1}^i = f_t(Z_t, X_t^0, X_t^i, U_t^i, W_t^i), \quad i \in \{1, \ldots, n\}$$

where $\{W_t^i\}_{t=1}^T$, $i \in \{1, \ldots, n\}$, are independent processes. As in the basic model, the primitive random variables $(X_1^0, X_1^1, \ldots, X_1^n, \{W_t^0\}_{t=1}^T, \{W_t^1\}_{t=1}^T, \ldots, \{W_t^n\}_{t=1}^T)$ are mutually independent. The dynamics f_t and the probability distribution of X_1^i and W_t^i do not depend on i.

Note that, in addition to the mean-field, the dynamics of the minor subsystems are also influenced by the state of the major subsystem. The information sets of the major and minor subsystems are

$$I_t^0 = \{Z_{1:t}, X_{1:t}^0\}, \quad I_t^i = \{Z_{1:t}, X_{1:t}^0, X_t^i\}. \qquad (6)$$

Thus, the history of the mean-field and the major subsystem's state are observed by both the major and minor controllers. In addition, the minor controller observes its local state as well.

As in the basic model, we assume that the control laws of all minor controllers are identical. Thus, the control actions are generated as follows

$$U_t^0 = g_t^0(Z_{1:t}, X_{1:t}^0), \quad U_t^i = g_t(Z_{1:t}, X_{1:t}^0, X_t^i).$$

The collection $(\mathbf{g}^0, \mathbf{g})$, where $\mathbf{g}^0 = (g_1^0, \ldots, g_T^0)$ and $\mathbf{g} = (g_1, \ldots, g_T)$, is called the *control strategy* of the system. Let $\mathcal{G}_{MF-MM}$ be a set of all such control strategies.

At time t, system incurs a cost $\ell_t(\mathbf{X}_t, X_t^0, \mathbf{U}_t, U_t^0)$ that depends on $\mathbf{X}_t, X_t^0, \mathbf{U}_t, U_t^0$. The performance of any strategy $(\mathbf{g}^0, \mathbf{g})$ is quantified by the following expected total cost

$$J(\boldsymbol{g}^0, \boldsymbol{g}) = \mathbb{E}^{(\boldsymbol{g}^0, \boldsymbol{g})}\left[\sum_{t=1}^{T} \ell_t(\mathbf{X}_t, X_t^0, \mathbf{U}_t, U_t^0)\right] \qquad (7)$$

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

Problem 2 *For the MF-MM model, identify a control strategy $(\mathbf{g}^{0^*}, \mathbf{g}^*)$ to minimize the expected total cost given by* (7).

As mentioned earlier in this section, similar generalization of the basic MF model in the context of mean-field games (MFG) has been considered in [11], [18][1]. Introducing a major agent gives rise to a non-trivial conceptual difficulty in MFG. To explain this difficulty, let us describe the solution approach used in MFG. Assume that all players follow a control strategy $\boldsymbol{g} = (g_1, \ldots, g_T)$, where $g_t : \mathcal{X} \mapsto \mathcal{U}$. In the case of infinite players, this strategy gives rise to a *deterministic* trajectory of the mean-field $\{Z_t\}_{t=1}^T$. Given a deterministic mean-field, the problem of finding the best response at an agent is a centralized stochastic control problem. The corresponding optimal strategy is called the best response to $\{Z_t\}_{t=1}^T$. A strategy-mean field pair $(\boldsymbol{g}, \{Z_t\}_{t=1}^T)$ is called *consistent* if $\{Z_t\}_{t=1}^T$ is generated by $\boldsymbol{g}$ and $\boldsymbol{g}$ is the best response to $\{Z_t\}_{t=1}^T$.

In the major-minor setup described above, the mean-field generated by a strategy $\boldsymbol{g} = (g_1, \ldots, g_T)$, where $g_t : \mathcal{X}^0 \times \mathcal{X} \mapsto \mathcal{U}$, depends on the sample realization of $X_{1:T}^0$; hence, it is not deterministic. Thus, one can not directly use the solution approach described above.

In this paper, we generalize the results in [16] to the major-minor model; in particular we obtain a dynamic programming decomposition to determine an optimal strategy. In contrast to the conceptual difficulties that arise in the solution methodology of MFG, our solution is straightforward.

We first describe a mean-field model with multiple types (MF-T for short, explained below) and show that an appropriate change of variables reduces the MF-T model to the basic MF model. We then show that the MF-MM model is a special case of the MF-T model.

C. Mean Field Sharing with Multiple Types (MF-T) Model

In this section, we describe the model of *mean-field sharing with multiple types* (MF-T for short). This model is a generalization of the basic MF model. Instead of homogeneous subsystems, the subsystems are heterogeneous. In particular, subsystems are allowed to have different types such that, within each type, subsystems are homogeneous.

Let Θ denote the set of all different types of subsystems. Assume we have $k \in \mathbb{N}$ different types i.e. $\Theta = \{1, \ldots, k\}$. Let subsystem $i \in \{1, \ldots, n\}$ have a type $\theta^i \in \Theta$. Let $X_t^i \in$

$\mathcal{X}(\theta^i)$ denote the state of subsystem i and $u_t^i \in \mathcal{U}(\theta^i)$ denote its control action. Assume $\mathcal{X}(\theta)$ and $\mathcal{U}(\theta)$ are finite sets, $\forall \theta \in \Theta$. Let Z_t be as follows:

$$Z_t = \frac{1}{n} \sum_{i=1}^{n} \delta_{\theta^i, X_t^i}. \qquad (8)$$

For example, suppose there are $n = 5$ subsystems with two different types i.e. $\Theta = \{1, 2\}$. The state space of type $\theta = 1$ is $\mathcal{X}(1) = \{1, 2\}$ and the state space of type $\theta = 2$ is $\mathcal{X}(2) = \{2, 3, 4\}$. Subsystems $i = 1, 3$ are of type 1 and subsystems $i = 2, 4, 5$ are of type 2. At time t, $x_t^1 = 2, x_t^2 = 3, x_t^3 = 2, x_t^4 = 4, x_t^5 = 2$, then $z_t(\theta = 1, x = 1) = 0, z_t(\theta = 1, x = 2) = \frac{2}{5}, z_t(\theta = 2, x = 2) = \frac{1}{5}, z_t(\theta = 2, x = 3) = \frac{1}{5}, z_t(\theta = 2, x = 4) = \frac{1}{5}$. Hence, mean-field $z_t = (0, \frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$.

The dynamics of subsystem i is given by

$$X_{t+1}^i = f_t(\theta^i, Z_t, X_t^i, U_t^i, W_t^i), i \in \{1, \ldots, n\}, \theta^i \in \Theta \quad (9)$$

where the primitive random variables $(X_1^1, \ldots, X_1^n, \{W_t^1\}_{t=1}^T, \ldots, \{W_t^n\}_{t=1}^T)$ are mutually independent. The dynamics f_t and the probability distribution of X_1^i and W_t^i may depend on type $\theta^i \in \Theta$ but they do not directly depend on i. In other words, if subsystems i and j have the same type i.e. $\theta^i = \theta^j$, then their dynamics and probability distribution of their respective primitive random variables are the same. In addition, type of each subsystem is fixed and every subsystem a priori knows the types of all other subsystems.

Similar to the basic model, we assume that the control laws of all controllers with the same type are identical. Thus, the control actions are generated as follows:

$$U_t^i = g_t(\theta^i, Z_{1:t}, X_t^i) \quad i \in \{1, \ldots, n\}, \theta^i \in \Theta. \qquad (10)$$

The collection $(\mathbf{g}(1), \ldots, \mathbf{g}(k))$, where $\mathbf{g}(\theta) := (g_1(\theta, \cdot), \ldots, g_T(\theta, \cdot)), \theta \in \Theta$, is called the *control strategy* of the system. Let $\mathcal{G}_{MF-T}$ be a set of all such control strategies. At time t, the system incurs a cost depending on joint state $\mathbf{X}_t$ and joint action $\mathbf{U}_t$ that is given by

$$\ell_t(\mathbf{X}_t, \mathbf{U}_t). \qquad (11)$$

The performance of any strategy $(\mathbf{g}(1), \ldots, \mathbf{g}(k))$ is measured by the following expected total cost

$$J(\mathbf{g}(1), \ldots, \mathbf{g}(k)) = \mathbb{E}^{\mathbf{g}(1), \ldots, \mathbf{g}(k)}\left[\sum_{t=1}^{T} \ell_t(\mathbf{X}_t, \mathbf{U}_t)\right] \qquad (12)$$

where the expectation is with respect to a joint measure induced on all system variables by the choice of $(\mathbf{g}(1), \ldots, \mathbf{g}(k))$. We are interested in the following optimization problem.

Problem 3 *For the MF-T model, identify a control strategy $(\mathbf{g}(1)^*, \ldots, \mathbf{g}(k)^*)$ to minimize the expected total cost given by* (12).

III. MAIN RESULTS

In this section, we present the main results of three mean-field models described in Section II.

---

[1] In MFG, the subsystems are only allowed to be weakly coupled in the cost (through the mean-field Z_t and state of the major subsystem X_t^0) while in this work subsystems are allowed to be arbitrarily coupled in the cost.

A. Basic MF Model

The basic MF model was investigated in [16]. The key idea in [16] is to construct a *coordinated system* using the common information approach of [17] and show that the mean-field Z_t is an information state for the coordinated system. Using this information state, we can obtain a dynamic program to identify an optimal control strategy. For completeness, we summarize the main steps below.

Following [17], we construct a coordinated system from original system described in Section II-A. Consider a virtual *coordinator* that observes the common information between all subsystems by time t i.e. $Z_{1:t}$. Based on $Z_{1:t}$, the coordinator decides control action $\Gamma_t : \mathcal{X} \rightarrow \mathcal{U}$ as follows:

$$\Gamma_t = \psi_t(Z_{1:t}). \tag{13}$$

We refer to ψ_t as *coordinator*'s control law at time t. In the coordinated system, each controller is a passive agent that uses its local state X_t^i and the mapping Γ_t, which is called *prescription*, to generate

$$U_t^i = \Gamma_t(X_t^i), \quad i \in \{1, \ldots, n\}. \tag{14}$$

Note that Γ_t and ψ_t do not depend on i. The dynamics of each subsystem and the cost function are the same as in the original problem.

It is shown in [17] that the coordinated system is equivalent to the original system. In particular, for any ψ_t, we can construct g_t given by

$$g_t(Z_{1:t}, X_t^i) := \psi_t(Z_{1:t})(X_t^i)$$

such that the expected total cost under strategy $\boldsymbol{g}$ in the original system is the same as that under strategy $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_T)$ in the coordinated system.

Thus, instead of searching for an optimal strategy $\boldsymbol{g} = (g_1, \ldots, g_T)$ in the original system, we may, equivalently, search for an optimal coordination strategy $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_T)$ in the coordinated system.

It is shown in [16] that Z_t is an information state for coordinated system. In particular, the cost given in (4) may be viewed as a function of (Z_t, Γ_t) (Lemma 1) and the evolution of mean-field Z_{t+1} depends only on the current information $\{Z_t, \Gamma_t\}$ (Lemma 2).

Lemma 1 (Lemma 3 of [16]) *The per-step cost may be written as a function of Z_t and Γ_t. In particular, there exits a function $\hat{\ell}$ such that $\mathbb{E}[\ell_t(\mathbf{X}_t, \mathbf{U}_t)|Z_{1:t}, \Gamma_{1:t}] =: \hat{\ell}_t(Z_t, \Gamma_t)$.*

Lemma 2 (Lemma 4 of [16]) *For any choice $\gamma_{1:t}$ of $\Gamma_{1:t}$, any realization $z_{1:t}$ of $Z_{1:t}$, and any z,*

$$\mathbb{P}(Z_{t+1}=z|Z_{1:t}=z_{1:t}, \Gamma_{1:t}=\gamma_{1:t})=\mathbb{P}(Z_{t+1}=z|Z_t=z_t, \Gamma_t=\gamma_t).$$

From Lemmas 1 and 2, one can conclude that Z_t is a controlled Markov process with control action Γ_t. Thus, standard results for MDPs can be used to obtain the optimal strategy as follows.

Theorem 1 (Theorem 1 of [16]) *Within basic mean-field control strategies $\mathcal{G}_{BMF}$, there is no loss of optimality in restricting attention to Markovian strategy i.e. $U_t^i =$*

$g_t(Z_t, X_t^i)$. *Furthermore, an optimal strategy $\boldsymbol{g}^*$ is obtained by solving the following dynamic program. Define recursively value functions $V_{T+1}(z) := 0, \forall z$, and for $t = T, \ldots, 1$,*

$$V_t(z_t):=\min_{\gamma_t}(\hat{\ell}_t(z_t,\gamma_t)+\mathbb{E}[V_{t+1}(Z_{t+1})|Z_t=z_t, \Gamma_t=\gamma_t]) \tag{15}$$

where $\gamma_t : \mathcal{X} \mapsto \mathcal{U}$. Let $\psi_t^*(z_t)$ denote the argmin in the right-hand side of (15). Define $g_t^*(z, x) := \psi_t^*(z)(x), \forall x, z$. Then, $\boldsymbol{g}^* = (g_1^*, \ldots, g_T^*)$ is an optimal strategy.

In [16], it is shown that the above results extend to the case when all subsystems observe a noisy version of the mean field by defining the information state as the posterior probability of the mean-field given noisy observations.

B. MF-T Model

The MF-T model is a straightforward generalization of the basic MF model and can be solved in a similar manner as Section III-A. For succinctness, we simply present a change of variable argument that lets us view the MF-T model as a special case of the basic MF model.

Consider the MF-T model in Section II-C. Define $\tilde{X}_t^i := (\theta^i, X_t^i)$ as an augmented state that takes value in state space $\tilde{\mathcal{X}} = \cup_{\theta \in \Theta} (\{\theta\} \times \mathcal{X}(\theta))$. With this augmented state space, the MF-T model is converted to the basic MF model. In particular, mean-field of the MF-T model (8) may be written as $Z_t = \frac{1}{n} \sum_{i=1}^{n} \delta_{\tilde{X}_t^i}$ which is in the form of (1). Similarly, the dynamics given by (9) may be written as

$$\tilde{X}_{t+1}^i = \tilde{f}_t(Z_t, \tilde{X}_t^i, U_t^i, W_t^i)$$

which is of the form of (2). Since the type of subsystems are fixed and a priori known to every subsystem, the primitive random variables $(\tilde{X}_1^1, \ldots, \tilde{X}_1^n, \{W_t^1\}_{t=1}^T, \ldots, \{W_t^n\}_{t=1}^T)$ are mutually independent conditioned on the types of subsystems. Note that dynamics $\tilde{f}_t$ and probability distribution on $\tilde{X}_t^i$ and W_t^i may only depend on the types and do not directly depend on the index i.

In addition, control law of the MF-T model (10) may be formulated as

$$U_t^i = \tilde{g}_t(Z_{1:t}, \tilde{X}_t^i)$$

which is in the form of (3). Furthermore, the cost given by (11) may be written as $\tilde{\ell}_t(\tilde{\mathbf{X}}_t, \mathbf{U}_t)$ which is of the form of (4). Thus, similar to Section III-A, we define function $\tilde{\psi}_t$ (coordinator's control law) as a mapping from the history of mean-field $Z_{1:t}$ to $\tilde{\Gamma}_t$ in the same fashion as (13) such that $\tilde{\Gamma}_t = \tilde{\psi}_t(Z_{1:t})$ where $U_t^i = \tilde{\Gamma}_t(\tilde{X}_t^i), i \in \{1, \ldots, n\}$, which is of the form of (14).

Hence, an immediate consequence of the results of Section III-A are the following:

Lemma 3 *The per-step cost given in $\tilde{\ell}_t(\tilde{\mathbf{X}}_t, \mathbf{U}_t)$ may be written as a function of Z_t and $\tilde{\Gamma}_t$. In particular, there exits a function $\hat{\ell}$ such that $\mathbb{E}[\tilde{\ell}_t(\tilde{\mathbf{X}}_t, \mathbf{U}_t)|Z_{1:t}, \tilde{\Gamma}_{1:t}] =: \hat{\ell}_t(Z_t, \tilde{\Gamma}_t)$.*

Lemma 4 *For any choice $\tilde{\gamma}_{1:t}$ of $\tilde{\Gamma}_{1:t}$, any realization $z_{1:t}$ of $Z_{1:t}$, and any z,*

$$\mathbb{P}(Z_{t+1}=z|Z_{1:t}=z_{1:t}, \tilde{\Gamma}_{1:t}=\tilde{\gamma}_{1:t})=\mathbb{P}(Z_{t+1}=z|Z_t=z_t, \tilde{\Gamma}_t=\tilde{\gamma}_t).$$

Similar to Section III-A, based on Lemmas 3 and 4, one can conclude that Z_t is a controlled Markov process with control action $\tilde{\Gamma}_t$. Thus, standard results for MDPs can be used to obtain the optimal strategy as follows.

Theorem 2 *Within MF-T control strategies $\mathcal{G}_{MF-T}$, there is no loss of optimality in restricting attention to Markovian strategy i.e. $U_t^i = g_t(\theta^i, Z_t, X_t^i)$. Furthermore, an optimal strategy $(g(1)^*, \ldots, g(k)^*)$ is obtained by solving the following dynamic program. Define recursively value functions $V_{T+1}(z) := 0, \forall z$, and for $t = T, \ldots, 1$,*

$$V_t(z_t):=\min_{\tilde{\gamma}_t}(\hat{\ell}_t(z_t,\tilde{\gamma}_t)+\mathbb{E}[V_{t+1}(Z_{t+1})|Z_t=z_t,\tilde{\Gamma}_t=\tilde{\gamma}_t]) \quad (16)$$

where $\tilde{\gamma}_t = (\gamma_t(1), \ldots, \gamma_t(k))$ and $\gamma_t(\theta) : \mathcal{X}(\theta) \mapsto \mathcal{U}(\theta), \forall \theta \in \Theta$. Let $\psi_t^*(z_t)$ denote the argmin in the right-hand side of (16). Define $g_t^*(\theta, z, x) := \psi_t^*(z)(\theta, x), \forall \theta, x, z$. Then, the collection of strategies $\mathbf{g}(\theta)^* = (g_1(\theta, \cdot)^*, \ldots, g_T(\theta, \cdot)^*), \forall \theta \in \Theta$, is optimal.

C. MF-MM Model

In this section, we show that the MF-MM model is a special case of the MF-T model described in Section II-C. In particular, the system has two types: major and minor. There is only one major subsystem; hence the mean-field is equivalent to (Z_t, X_t^0). Let $\Gamma_t : \mathcal{X} \mapsto \mathcal{U}$ and $\Gamma_t^0 : \mathcal{X}^0 \mapsto \mathcal{U}^0$. Then, by virtue of Lemma 3 and Lemma 4, we have that (Z_t, X_t^0) is a controlled Markov process with control action (Γ_t, Γ_t^0). Note that this result is simplified further because X_t^0 is part of the local information at the major subsystem as well as the common information. Therefore, it can be shown that, the function $\Gamma_t^0 : \mathcal{X}^0 \mapsto \mathcal{U}^0$ may be replaced by the action U_t^0. Thus, from Lemma 3 and Lemma 4, we get the following:

Lemma 5 *The per-step cost $\ell_t(\mathbf{X}_t, X_t^0, \mathbf{U}_t, U_t^0)$ may be written as a function of (Z_t, X_t^0) and (Γ_t, U_t^0). In particular, there exits a function $\hat{\ell}$ such that*

$$\mathbb{E}[\ell_t(\mathbf{X}_t, X_t^0, \mathbf{U}_t, U_t^0)|Z_{1:t}, X_{1:t}^0, \Gamma_{1:t}, U_{1:t}^0]=:\hat{\ell}_t(Z_t, X_t^0, \Gamma_t, U_t^0).$$

Lemma 6 *For any choice of $(\gamma_{1:t}, u_{1:t}^0)$ of $(\Gamma_{1:t}, U_{1:t}^0)$, any realization $(z_{1:t}, x_{1:t}^0)$ of $(Z_{1:t}, X_{1:t}^0)$, and any (z, x^0),*

$$\mathbb{P}(Z_{t+1}=z,X_{t+1}^0=x^0|Z_{1:t}=z_{1:t},X_{1:t}^0=x_{1:t}^0,\Gamma_{1:t}=\gamma_{1:t},U_{1:t}^0=u_{1:t}^0)$$
$$= \mathbb{P}(Z_{t+1}=z, X_{t+1}^0=x^0|Z_t=z_t, X_t^0=x_t^0, \Gamma_t=\gamma_t, U_t^0=u_t^0)$$

From Lemmas 5 and 6, we can conclude that (Z_t, X_t^0) is a controlled Markov process with control actions (Γ_t, U_t^0). Hence, we can use the standard results for MDPs to obtain the optimal strategies.

Theorem 3 *Within MF-MM control strategies $\mathcal{G}_{MF-MM}$, there is no loss of optimality in restricting attention to Markovian strategy i.e. $U_t^0 = g_t^0(Z_t, X_t^0)$ and $U_t^i = g_t(Z_t, X_t^0, X_t^i)$. Furthermore, an optimal strategy $(\mathbf{g}^{0*}, \mathbf{g}^*)$ is obtained by solving the following dynamic program. Define recursively value functions $V_{T+1}(z, x^0) := 0, \forall z, x^0$, and for $t = T, \ldots, 1$,*

$$V_t(z_t, x_t^0) := \min_{\gamma_t, u_t^0}(\hat{\ell}_t(z_t, x_t^0, \gamma_t, u_t^0)$$
$$+\mathbb{E}[V_{t+1}(Z_{t+1},X_{t+1}^0)|Z_t=z_t,X_t^0=x_t^0,\Gamma_t=\gamma_t,U_t^0 = u_t^0]) \quad (17)$$

where $\gamma_t : \mathcal{X} \mapsto \mathcal{U}$. Let $\psi_t^*(z_t, x_t^0)$ denote the argmin in the right-hand side of (17) such that $\gamma_t^* = \psi_t^*(1)(z_t, x_t^0)$ and $u_t^{0*} = \psi_t^*(2)(z_t, x_t^0)$. Define $g_t^*(z, x^0, x) := \psi_t^*(1)(z, x^0)(x)$ and $g_t^{0*}(z, x^0) := \psi_t^*(2)(z, x^0), \forall x, z, x^0$. Then, the collection of strategies $\mathbf{g}^* = (g_1^*, \ldots, g_T^*)$ and $\mathbf{g}^{0*} = (g_1^{0*}, \ldots, g_T^{0*})$ is optimal.

Remark 1 *At time t, the optimal action of major subsystem U_t^{0*} depends on the mean field of minor subsystems Z_t and the state of major subsystem X_t^0 i.e. $U_t^{0*} = g_t^{0*}(Z_t, X_t^0)$. Moreover, the optimal action of minor subsystem i, U_t^{i*}, depends on its own local state X_t^i in addition to Z_t and X_t^0 i.e. $U_t^{i*} = g_t^*(Z_t, X_t^0, X_t^i)$.*

IV. EXAMPLE

In this section, we present an example of the MF-MM model where a service provider (major subsystem) interacts with its users (minor subsystems). Roughly speaking, we answer the question of how a service provider should design strategies for itself and the users such that the service is not only profitable but also customer-satisfactory.

Consider a service provider (internet provider, utility provider, etc.) who serves $n \in \mathbb{N}$ users (customers). Let $X_t^i \in \mathcal{X} = \{0, 1\}$, $i \in \{1, \ldots, n\}$, denote whether user i is active ($X_t^i = 1$) or passive ($X_t^i = 0$) at time t and $X_t^0 \in \mathcal{X}^0$ denote the available capacity of the system. At time t, the service provider chooses U_t^0, the capacity of the system at time $t + 1$. Thus, the dynamics of the state of the service provider is

$$X_{t+1}^0 = U_t^0$$

where $\mathcal{X}^0 = \mathcal{U}^0$ is a finite set consisting of all feasible capacities.

Dynamics of users are represented by $P(u)$ where $P(u)$ denotes the controlled transition matrix under action $u \in \mathcal{U} = \{0, 1, 2\}$, i.e.

$$[P(u)]_{xy} = \mathbb{P}(X_{t+1}^i = y \mid X_t^i = x, U_t^i = u), \quad x, y \in \mathcal{X}.$$

The service provider can affect the evolution of user i by choosing action $u^i \in \{0, 1, 2\}$. Action $u^i = 0$ is a *free action* under which user i evolves in an uncontrolled manner, i.e. $P(0) = Q$, where Q represents the *natural* dynamics of the users. Action $u^i \neq 0$ is a *forcing action* under which service provider forces user i to switch to state $(u^i - 1)$. However, there is $\epsilon_{u^i} \in [0, 1]$ probability that user i does *not* switch to state $(u^i - 1)$ and evolves according to its natural dynamics. Thus,

$$P(u) = (1 - \epsilon_u)\mathbf{K}_u + \epsilon_u Q$$

where $\mathbf{K}_u$ is a $k \times k$ matrix where column $u \in \{1, 2\}$ is all ones, and other columns are all zeros. The goal is that the service provider efficiently manages the users and the capacity of the system in order to minimize costs while satisfying the users. Thus, the loss function is given by

$$\ell_t^0(Z_t, X_t^0, U_t^0) + \sum_{i=1}^n \ell_t(U_t^i)$$

where ℓ_t^0 is as follows:

$$\ell_t^0(Z_t, X_t^0, U_t^0) = S(U_t^0) + a|U_t^0 - X_t^0| - G(Z_t, X_t^0), \ a \in \mathbb{R}^+$$

and $\ell_t(U_t^i) = H(U_t^i)$. In ℓ_t^0, the first term is associated with the cost of capacity and the second term refers to patching and dispatching capacity. The third term corresponds to the benefit (proportional to the number of active users) and the penalty of unavailable service (proportional to the number of active users that do not receive service) i.e.

$$G(Z_t, X_t^0) = \begin{cases} bnZ_t(1) & nZ_t(1) \le X_t^0 \\ bX_t^0 - c(nZ_t(1) - X_t^0) & nZ_t(1) > X_t^0 \end{cases}$$

where $b \in \mathbb{R}^+$ indicates the rate of benefit and $c \in \mathbb{R}^+$ determines the rate of penalty of unavailable service. Note that $Z_t(1) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_t^i = 1)$ is the average number of active users at time t. In addition, ℓ_t is the cost associated with forcing users.

Given the information structure (6), the objective is to choose a control strategy that minimizes the infinite horizon discounted cost

$$J(\boldsymbol{g}) = \mathbb{E}\left[\sum_{t=1}^{\infty} \beta^t \left(\ell_t^0(Z_t, X_t^0, U_t^0) + \sum_{i=1}^n \ell_t(U_t^i) \right) \right]$$

where $\beta \in (0, 1)$ is a discount factor[2]. The optimal time-homogeneous strategies for

$$n = 100, \ \ \mathcal{X}^0 = \{50, 100\}, \ \ S(50) = 100, \ \ S(100) = 300,$$
$$H(0) = 0, \ \ H(1) = 4, \ \ H(2) = 1, \ \ \epsilon_1 = 0.1, \ \ \epsilon_2 = 0.1,$$
$$a = 2, \ \ \ b = 5, \ \ \ c = 50, \ \ \ \beta = 0.6, \ \ \ Q = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}$$

are presented below. Since the state space of minor subsystems is binary, $z(1)$ is sufficient to characterise the empirical distribution $z = [z(0), z(1)]$. Hence, for ease of presentation, we represent the optimal control law as a function of the second component $z(1)$ of $z = [z(0), z(1)]$.

$$g^*(z, x^0, x) = \begin{cases} 0 & 0 \le z(1) \le 0.53, & x^0 = 50, x = 0 \\ 1 & 0.53 < z(1) \le 0.76, & x^0 = 50, x = 0 \\ 2 & 0.76 < z(1) \le 1, & x^0 = 50, x = 0 \\ 0 & 0 \le z(1) \le 1, & x^0 = 50, x = 1 \\ 0 & 0 \le z(1) \le 0.29, & x^0 = 100, x = 0 \\ 2 & 0.29 < z(1) \le 1, & x^0 = 100, x = 0 \\ 0 & 0 \le z(1) \le 1, & x^0 = 100, x = 1 \end{cases}$$

and

$$g^{0*}(z, x^0) = \begin{cases} 50 & 0 \le z(1) \le 0.76, & x^0 = 50 \\ 100 & 0.76 < z(1) \le 1, & x^0 = 50 \\ 50 & 0 \le z(1) \le 0.29, & x^0 = 100 \\ 100 & 0.29 < z(1) \le 1, & x^0 = 100. \end{cases}$$

[2]Although we have only presented the details for finite horizon setup in this paper, the results generalize naturally to infinite horizon setup under standard assumption.

V. Conclusion

In this paper, we presented two variations of the basic mean-field sharing model: mean-field sharing with major-minor subsystems (MF-MM) and mean-field sharing with multiple types (MF-T). We demonstrated that, by an appropriate change of variables, the MF-T model may be viewed as a special case of the basic mean-field (MF) model presented in [16]. Based on the main results of the basic MF model in [16], we identified a dynamic program to identify optimal strategies for the MF-T model. In addition, we showed that the MF-MM model may be viewed as a special case of the MF-T model. Due to the special structure of the MF-MM model, we could simplify the dynamic program associated with the MF-T model further to obtain the optimal strategies for the major and minor subsystems. Note that we provided a straightforward solution for the MF-MM model in contrast to mean-field games where such a solution is non-trivial and difficult.

References

[1] J. Marschack and R. Radner, "Economic theory of teams," *New Haven: Yale University Press*, Feb. 1972.

[2] H. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. of IEEE*, vol. 59, pp. 1557–1566, Nov. 1971.

[3] H. S. Witsenhausen, "On information structures, feedback and causality," *SIAM Journal on Control*, vol. 9, no. 2, pp. 149–160, Jan. 1971.

[4] D. S. Bernstein, R. Givan, N. Immerman, and S. Zilberstein, "The complexity of decentralized control of Markov decision processes," *Math. of operations research*, vol. 27, no. 4, pp. 819–840, Nov. 2002.

[5] S. Yüksel and T. Başar, *Stochastic Networked Control Systems*. Birkhauser, 2013.

[6] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yuksel, "Information structures in optimal decentralized control," *Proc. of Conf. on Decision and Control (CDC)*, pp. 1291–1306, Dec. 2012.

[7] R. Balescue, *Equilibrium and Nonequilibrium Statistical Mechanics*. John Wiley, New York, 1975.

[8] Y. L. Klimontovich, "S-theorem," *Zeitschrift für Physik B Condensed Matter*, vol. 66, no. 1, pp. 125–127, Aug. 1987.

[9] M. Huang, P. E. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions," *IEEE Conf. on Dec. and Cont. (CDC)*, pp. 98–103, Dec. 2003.

[10] M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[11] M. Nourian and P. E. Caines, "ε-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents," *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, Aug. 2013.

[12] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, Mar. 2007.

[13] O. Guéant, J.-M. Lasry, and P.-L. Lions, "Mean field games and applications," *Paris-Princeton Lectures on Mathematical Finance*, pp. 205–266, 2011.

[14] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," *NIPS*, 2005.

[15] D. A. Gomes and J. Saude, "Mean field games models: A brief survey," *Springer, Dynamic Games and Appl.*, vol. 4, pp. 1–45, Jun. 2014.

[16] J. Arabneydi and A. Mahajan, "Team optimal control of coupled subsystems with mean field sharing," *IEEE Con. on Dec. and Cont. (CDC)*, Dec. 2014.

[17] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Tran. on Automatic Control*, vol. 58, no. 7, Jul. 2013.

[18] M. Huang, "Large-population LQG games involving a major player: the Nash certainty equivalence principle," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, Jan. 2010.