Optimal Solutions to Infinite-Player Stochastic Teams and Mean-Field Teams

Sina Sanjari, *Student Member, IEEE*, and Serdar Yüksel, *Member, IEEE*

***Abstract*—We study stochastic static teams with countably infinite number of decision makers (DMs), with the goal of obtaining (globally) optimal policies under a decentralized information structure. We present sufficient conditions to connect the concepts of team optimality and person-by-person optimality for static teams with countably infinite number of DMs. We show that under uniform integrability and uniform convergence conditions, an optimal policy for static teams with countably infinite number of DMs can be established as the limit of sequences of optimal policies for static teams with N DMs as $N \to \infty$. Under the presence of a symmetry condition, we relax the conditions and this leads to optimal results for a large class of mean-field optimal team problems where the existing results have been limited to person-by-person optimality and not global optimality (under strict decentralization). In particular, we establish the optimality of symmetric (i.e., identical) policies for such problems. As a further condition, this optimality result leads to an existence result for mean-field teams. We consider a number of illustrative examples where the theory is applied to setups with either infinitely many DMs or an infinite-horizon stochastic control problem reduced to a static team.**

***Index Terms*—Average cost optimization, decentralized control, mean-field theory, stochastic teams.**

I. Introduction

A DECENTRALIZED control system, or a team, consists of a collection of decision makers (DMs)/agents acting together to optimize a common cost function, but not necessarily sharing all the available information. Teams whose initial states, observations, cost function, or the evolution dynamics are random or are disturbed by some external noise processes are called *stochastic teams*. At each time stage, each agent only has access to some parts of the global information. If each agent's information depends only on primitive random variables, the team is *static*. If at least one agent's information is affected by an action of another agent, the team is said to be *dynamic*.

Manuscript received August 8, 2018; revised August 12, 2018, June 19, 2019, and December 12, 2019; accepted April 28, 2020. Date of publication May 15, 2020; date of current version February 26, 2021. This work was supported by the Natural Sciences and Engineering Research Council of Canada. This paper was presented in part at the 57th IEEE Conference on Decision and Control, Miami Beach, FL, USA, 2018 [43].

The authors are with the Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: 16sss3@queensu.ca; yuksel@mast.queensu.ca).

Digital Object Identifier 10.1109/TAC.2020.2994899

On teams with finitely many DMs, Marschak [36] studied optimal static teams and Radner [40] developed foundational results on optimality and established connections between person-by-person optimality, stationarity, and team-optimality. Radner's results were generalized in [30] by relaxing optimality conditions. A summary of these results is that in the context of static team problems, convexity of the cost function, subject to minor regularity conditions, may suffice for the global optimality of person-by-person-optimal solutions. In the particular case for linear quadratic Gaussian (LQG) static teams, this result leads to the optimality of linear policies [40], which also applies for dynamic LQG problems under specific information structures (ISs) (to be discussed further below) [23]. These results are applicable for static teams with finite number of DMs. In our article, the focus is on teams with infinitely many DMs.

A. Connections With the Literature on Mean-Field Games/Teams

On the case with infinitely many DMs, a related set of results involves mean-field games: mean-field games (see, e.g., [24], [25], [34]) can be viewed as limit models of symmetric nonzero-sum noncooperative N-player games with a mean-field interaction as $N \to \infty$. The uniqueness and nonuniqueness results have been established for mean-field games in both the partial differential equations (PDE) and probabilistic setting [4], [11], [34]. In [4], examples have been provided to show the existence of multiple solutions to the mean-field games when uniqueness conditions in [11] and [34] are violated. The mean-field approach designs policies for both cases of games with infinitely many players, as well as games with very large number of players where the equilibrium policies for the former are shown to be ϵ-equilibria for the latter [12], [24], [42]. These results, while very useful for establishing equilibria or in the context of team problems, person-by-person-optimal policies, does not guarantee the ϵ-global optimality among all policies. That is, ϵ-person-by-person optimality is not sufficient for ϵ-global optimality since in the limit, one typically only finds equilibrium policies without establishing their uniqueness (which would imply global optimality for team problems) [29], [37], [45]. Related to such problems, in the economic theory literature, Mas-Colell and Schmeidler [37], [45] considered *Cournot-Nash* equilibria. This Cournot-Nash equilibrium concept corresponds to a mean-field equilibrium for a static problem. However, such an equilibrium does not necessarily imply global optimality in the context of team problems, as discussed above.

Recently, mean-field team problems have also been studied: Social optima for mean-field LQG control problems under both centralized and a specific decentralized IS have been considered in [26] and [47]. In [2], a setup is considered where DMs share some information on the mean-field in the system, and through showing that the performance of a corresponding centralized system can be realized under a decentralized IS, global optimality is established. In this article, we follow an approach where optimality for every N is established and also optimality holds as $N \to \infty$ for the limit policy. The papers [27] and [28] have studied a continuous-time setup where a major agent is present; by considering the social impact for each individual player, they showed person-by-person optimal policies asymptotically minimize the social cost [26]. By approximating the mean-field term, the authors bound the induced approximation error of order $O(N^{\frac{-1}{2}} + \epsilon_N)$ where ϵ_N goes to zero as the number of players $N \to \infty$ [26]. In [9], mean-field team problems with mixed players have been considered where minor agents act together to minimize a common cost against a major player. Also, for the linear quadratic (LQ) setup, under the assumption that DMs apply identical policy in addition to some technical assumptions on the cost function and transition probabilities of Markov chains, Arabneydi and Aghdam [1] showed that the expected cost achieved by a suboptimal fully decentralized strategy is on $\epsilon(n)$ neighborhood of the optimal cost achieved when mean-field (empirical distribution of states) has been shared, where n is the number of players. Such results on mean-field teams either show global optimality through equivalence to the performance of a centralized setup (considering specific sharing patterns on the mean-field model) or typically only assume person-by-person-optimality. In this article, we will establish global optimality under a completely decentralized IS; however, certain technical conditions will be imposed.

B. Connections With the Literature on Limits of Finite Player Games/Teams

There exist contributions where games with finitely many players are studied, their equilibrium solutions are obtained, and the limit is taken. Along this direction, the connection between Nash equilibrium of symmetric N-player games and an optimal solution of mean-field games has been addressed in [3], [5], [7], [17], [18], and [31]. The goal is to find sufficient conditions such that the limit of the sequences of Nash equilibrium for the N-player games identify as a solution of the corresponding mean-field game as $N \to \infty$. Convergence of Nash equilibria of symmetric N-player games to the corresponding mean-field games for stationary continuous-time problems with ergodic costs has been investigated in [5] and [17]. Moreover, such a convergence of Nash equilibria for symmetric N-player games to the corresponding mean-field solution for a broad class of continuous time symmetric games has been established in [18] under uniform integrability and exchangeability (symmetry) conditions (see [18, Th. 5.1 and conditions (T) and (S)]) provided that the cost function and dynamics admit the structural restrictions. In [31], assumptions on equilibrium policies of the large population mean-field symmetric stochastic differential games have been relaxed to allow the convergence of asymmetric approximate Nash equilibria to a weak solution of the mean-field game [31, Th. 2.6]. In a discrete-time setup, Biswas [7] considered convergence of Nash equilibria for games with the mean-field interaction and with ergodic costs for Markov processes. The convergence result has been derived under an existence assumption on the mean-field solution and an additional convexity condition (see [7, Th. 5.1 and condition (A7)]). In contrast, in the context of stochastic teams with countably infinite number of DMs, the gap between person-by-person optimality (Nash equilibrium in the game-theoretic context) and global team optimality is significant since a perturbation of finitely many policies fails to deviate the value of the expected cost; thus, person by person optimality is a weak condition for such a setup, and hence the results presented in the aforementioned papers may be inconclusive regarding global optimality of the limit equilibrium. This observation motivates us to investigate the connection between person-by-person-optimality and global team optimality in stochastic teams with countably infinite DMs. Compared with [3], [5], [7], [17], and [18] where only the convergence of a sequence of Nash equilibria for symmetric games with the mean-field interaction has been studied, we show that, under sufficient conditions, sequences of optimal policies for teams with N number of DMs as $N \to \infty$ converge to a team optimal policy for static teams with countably infinite number of DMs.

Related to mean-field team problems, a limit theory for *mean-field type problems* (also called *Mckean–Vlasov stochastic control problems*) has been established in [10] and [32]. In [10] and [32], the connection between solutions of N-player differential control systems and solutions of Mckean–Vlasov control problems has been investigated. It has been shown that the sequence of empirical measures of pairs of states and ϵ_N-centralized optimal controls (under the classical IS since all the information available are completely shared between players) converges in distribution as $N \to \infty$ to limit points in the set of pairs of states and optimal controls of the Mckean–Vlasov problem [32] (see Remark 3). In contrast, our focus is on the ISs of DMs. Here, under convexity of the cost function and symmetry, we show the convergence of a sequence of decentralized optimal policies of N-DM teams to an optimal policy of mean-field teams as $N \to \infty$.

C. Connections With the Literature on LQG Games/Teams

There has been a number of studies focusing on the LQG setup (in addition to [26] and [47]). A close study is [35] where LQG static teams with countably infinite number of DMs have been studied and sufficient conditions for global optimality have been established. In this article, we utilize some of the results from [35]; however, compared with [35], we propose sufficient conditions for team optimality on average cost problems for a general setup: except convexity, no specific structure is presumed *a priori* on the cost function. For our analysis, we do not restrict the setup to the LQG one, where often direct methods can be applied building on [30] and [40], and operator theory involving matrix algebra; in addition, we also study the mean-field setting. In fact, for a general setup of static teams, we introduce sufficient conditions (see Theorems 5 and 6) such that the optimal cost and optimal policies of static teams with countably infinite number of DMs is obtained as a limit of the optimal cost and optimal policies for static teams with N number of DMs as $N \to \infty$. In [20], LQG team problems with infinitely many DMs have been considered for a setup where the cost function is the expected inner-product of an infinite dimensional vector (and to allow for a Hilbert theoretic formulation, finiteness of the infinite sum of the moments of individual random variables is imposed) and linearity and uniqueness of optimal policies have been established; the finiteness (of the infinite summation) restriction rules out the setup in this article. In [39], infinite horizon decentralized stochastic control problems containing a remote controller and a collection of local controllers dealing with linear models have been addressed for a setup where the cost is quadratic and the communication model satisfies a specified sharing pattern of information between a local controller and remote controller. Under the assumed sharing pattern (common information), the connections between the optimal solution and the coupled algebraic Riccati equation for Markov jump linear systems and its convergence to the coupled fixed point equations have been utilized to show the optimality of the solution [39].

As a further motivation for our article, we note that for dynamic team problems, Ho and Chu [23] have introduced a technique such that dynamic partially nested LQG team problems can be reduced to static team problems (we also note that Witsenhausen [48] showed that under an absolute continuity condition, any sequential dynamic team can be reduced to a static one). For infinite-horizon dynamic team problems, this reduction leads to a static team with countably many DMs; thus, leading to a different setup where our results in this article will be applicable. We will study a particular example as a case study. In particular, the question of whether partially nested dynamic LQG teams admit optimal policies under an expected average cost criterion, in its most general form, has not been conclusively addressed despite the presence of results, which impose linearity *a priori* for the optimal policies under such ISs [41]. We hope that our solution approach can be utilized in the future to develop a complete theory for such problems.

D. Contributions

1) For a general setup of static teams, we show that (see Theorem 6), under a uniform integrability condition (see Remark 2), if sequences of team optimal policies of DMs $i = 1, \dots, N$ of static teams with N number of DMs converge uniformly in $i = 1, \dots, N$ (see (b) in Theorem 6), then the corresponding limit policies are team optimal for the static team with countably infinite number of DMs, under the expected average cost criteria.
2) We establish global optimality results for mean-field teams under strict decentralization of the IS for both teams with large numbers of players and infinitely many players. Toward this end, we introduce a notion of symmetrically optimal teams (see Definition 6) to obtain a global optimality result under relaxed sufficient conditions (see Section IV). Under mild conditions on action spaces and observations of DMs, through concentration of measures arguments, we establish the convergence of optimal policies for symmetric mean-field teams with N DMs to the corresponding optimal policy of mean-field teams (see Section IV). In addition, we establish an existence result for optimal policies on mean-field teams under relaxed conditions on action spaces and the cost function (see Theorem 12).
3) We apply our results to a number of illustrative examples: We first consider LQG and LQ (non-Gaussian) average cost problems with state coupling (see Sections V-A and V-B). We also consider LQG average cost problems with control coupling (see Section V-C). In addition, we show that the team optimal policy of LQG teams with classical IS (see Section V-D) is obtained using the technique proposed in this article. This is important since this result, while is well-known in the stochastic control literature, has not been investigated using static reduction proposed in [23] and hence this approach can be viewed as a step to address optimal solutions for infinite-horizon partially nested dynamic LQG problems, which can be reduced to a static team with countably infinite number of DMs.

The organization of the article is as follows. Preliminaries and the problem statement are presented in Section II. Section III contains our main results including sufficient conditions for team optimality and asymptotic optimality for a general setup of static teams with countably infinite number of DMs. Section IV discusses symmetric and mean-field teams, and applications are presented in Section V. Section VI presents concluding remarks.

II. Problem Formulation

A. Preliminaries

Before presenting our main results, we introduce preliminaries following the presentation in [53], in particular, we introduce the characterizations laid out by Witsenhausen, through his *Intrinsic Model* [49]; further characterizations and classifications of ISs are introduced comprehensively in [52]. Suppose there is a predefined order in which the DMs act. Such systems are called *sequential systems*. The action and measurement spaces are standard Borel spaces, that is, Borel subsets of complete, separable, and metric spaces. The *Intrinsic Model* for sequential teams is defined as follows.

1) There exists a collection of *measurable spaces* $\{(\Omega, \mathcal{F}), (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{V}^i, \mathcal{V}^i), i \in \mathcal{N}\}$, specifying the system's distinguishable events, and control and measurement spaces, where $\mathcal{N}$ is either $\{1, \dots, N\}$ or $\mathbb{N}$ ($\mathbb{N}$ denotes the set of natural numbers). In this model (described in discrete time), any action applied at any given time $t \in \mathcal{N}$ is regarded as applied by a DM DM^i for $i \in \mathcal{N}$, who acts only once. The pair $(\Omega, \mathcal{F})$ is a measurable space (on which an underlying probability may be defined). The pair $(\mathbb{U}^i, \mathcal{U}^i)$ denotes the measurable space from which the action, u^i, of DM i is selected. The pair $(\mathbb{V}^i, \mathcal{V}^i)$ denotes the measurable observation/measurement space.
2) There is a *measurement constraint* to establish the connection between the observation variables and the system's distinguishable events. The $\mathbb{V}^i$-valued observation variables are given by $v^i = h^i(\omega, \underline{u}^{[1,i-1]})$, where $\underline{u}^{[1,i-1]} = \{u^k, k \leq i-1\}$, h^i are given measurable functions and u^k denotes the action of DM^k. Hence, v^i induces $\sigma(v^i)$ over $\Omega \times \prod_{k=1}^{i-1} \mathbb{U}^k$.
3) The set of admissible control laws $\underline{\gamma} = \{\gamma^1, \gamma^2, \dots\}$, also called *designs* or *policies*, are measurable control functions, so that $u^i = \gamma^i(v^i)$. Let Γ^i denote the set of all admissible policies for DM^i.
4) There is a *probability measure* $\mathbb{P}$ on $(\Omega, \mathcal{F})$ describing the probability space on which the system is defined.

Under this intrinsic model, a sequential team problem is *dynamic* if the information available to at least one DM is affected by the action of at least one other DM. A team problem is *static*, if for every DM, the information available is only affected by exogenous disturbances; that is, no other DM can affect the information of any given DM.

ISs can also be categorized as *classical*, *quasi-classical*, or *nonclassical*. An IS $\{v^i, i \in \mathcal{N}\}$ is *classical* if v^i contains all of the information available to DM^k for $k < i$. An IS is *quasi-classical* or *partially nested*, if whenever u^k, for some $k < i$, affects v^i through the measurement function h^i, v^i contains v^k (that is $\sigma(v^k) \subset \sigma(v^i)$). An IS, which is not partially nested, is *nonclassical*.

($\mathcal{P}'_N$) Let $N = |\mathcal{N}|$ be the number of control actions taken, and each of these actions is taken by a different DM, where $\mathcal{N} := \{1, \ldots, N\}$. Let $\underline{\gamma}_N = \{\gamma^1, \ldots, \gamma^N\}$ and let $\mathbf{\Gamma}_N = \prod_i^N \Gamma^i$ be the space of admissible policies for the team with N-DMs. Assume an expected cost function is defined as

$$J_N(\underline{\gamma}_N) = E^{\underline{\gamma}_N}[c(\omega_0, \underline{u}_N)] \tag{1}$$

for some Borel measurable cost function $c : \Omega_0 \times \prod_{k=1}^N \mathbb{U}^k \to \mathbb{R}$ where $E^{\underline{\gamma}_N}[c(\omega_0, \underline{u}_N)] := E[c(\omega_0, \gamma^1(v^1), \ldots, \gamma^N(v^N))]$ and we define ω_0 as the cost function relevant exogenous random variable as $\omega_0 : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega_0, \mathcal{B}(\Omega_0))$. Here, we have the notation $\underline{u}_N := \{u^i, i \in \mathcal{N}\}$ and $\mathcal{B}(\cdot)$ denotes the Borel σ-field.

Definition 1: Team optimal solution for ($\mathcal{P}'_N$) [52].

For a given stochastic team problem with a given IS, a policy (strategy) N-tuple $\underline{\gamma}_N^* := (\gamma^{1^*}, \ldots, \gamma^{N^*}) \in \mathbf{\Gamma}_N$ is *optimal* (*team-optimal solution*) for ($\mathcal{P}'_N$) if

$$J_N(\underline{\gamma}_N^*) = \inf_{\underline{\gamma}_N \in \mathbf{\Gamma}_N} J(\underline{\gamma}_N) =: J_N^*.$$

Definition 2: Person-by-person optimal solution [52].

For a given N-DM stochastic team with a fixed IS, an N-tuple of strategies $\underline{\gamma}_N^* := (\gamma^{1^*}, \ldots, \gamma^{N^*})$ constitutes a *person-by-person optimal* solution for ($\mathcal{P}'_N$) if, for all $\beta \in \Gamma^i$ and all $i \in \mathcal{N}$, the following inequalities hold:

$$J_N^* := J_N(\underline{\gamma}_N^*) \leq J_N(\underline{\gamma}_N^{-i*}, \beta)$$

where $(\underline{\gamma}_N^{-i*}, \beta) := (\gamma^{1*}, \ldots, \gamma^{(i-1)*}, \beta, \gamma^{(i+1)*}, \ldots, \gamma^{N*})$.

To simplify notations, let for any $1 \leq k \leq N$, $\underline{\gamma}_N^{-k} := \{\gamma^i, i \in \{1, \cdots, N\} \setminus \{k\}\}$.

Definition 3 (Stationary solution [40]): A policy $\underline{\gamma}_N(.)$ is stationary if $J(\underline{\gamma}_N) < \infty$, and for all $i = 1, \ldots, N$, $\mathbb{P}$-almost surely

$$\nabla_{u^i} \mathbb{E}\left[c(\omega_0, (\underline{\gamma}_N^{-i}, u^i)) \bigg| v^i\right]\bigg|_{u^i = \gamma^i(v^i)} = 0$$

where ∇_{u^i} denotes the gradient with respect to u^i.

In this section, without abuse of notations, we sometimes used γ^i as $\gamma^i(v^i)$. In the following, we present some related existing results for static teams with N DMs. The following is known as Radner's theorem [40]. Radner proposed the first result to connect the stationarity concept and global team optimality.

Theorem 1 (see [40]): If

a) $c(\omega_0, \underline{u}_N)$ is convex and differentiable in $\underline{u}_N$ for $\mathbb{P}$- almost surely;
b) $\inf_{\underline{\gamma}_N \in \mathbf{\Gamma}_N} J_N(\underline{\gamma}_N) > -\infty$;
c) $J_N(.)$ is locally finite at $\underline{\gamma}_N^*$ [40];
d) $\underline{\gamma}_N^*$ is stationary;

then $\underline{\gamma}_N^*$ is globally optimal for ($\mathcal{P}'_N$).

Radner's theorem fails in some applications because of the restrictive local finiteness assumption. Krainak *et al.* [30] relaxed assumptions and presented sufficient conditions for team optimality on static teams.

Theorem 2 (see [30]): Assume that, for every fixed ω_0, $c(\omega_0, \underline{u}_N)$ is convex differentiable in $\underline{u}_N$. Suppose (b) in Theorem 1 holds. Let $\underline{\gamma}_N^* \in \mathbf{\Gamma}_N$, and assume that $\mathbb{E}[c(\omega_0, \underline{\gamma}_N^*(\underline{v}_N))] < \infty$. If, for all $\underline{\gamma}_N \in \mathbf{\Gamma}_N$ with $\mathbb{E}[c(\omega_0, \underline{\gamma}_N(\underline{v}_N))] < \infty$

$$\mathbb{E}\left[\sum_{i=1}^N c_{u^i}(\omega_0, \underline{\gamma}_N^*)(\gamma^i - \gamma^{i*})\right] \geq 0 \tag{2}$$

where $c_{u^i}(\omega_0, \underline{\gamma}_N^*)$ is the partial derivative of $c(\omega_0, \underline{u}_N)$ with respect to u^i valued in $\underline{u}_N = \underline{\gamma}_N^*$, then $\underline{\gamma}_N^*$ is an optimal team policy for ($\mathcal{P}'_N$). Moreover, if $c(\omega_0, \underline{u}_N)$ is strictly convex in $\underline{u}_N$ $\mathbb{P}$-almost surely, then $\underline{\gamma}_N^*$ is $\mathbb{P}$-a.s. unique.

Since the set of admissible policies is generally uncountable, checking (2) is difficult. Krainak *et al.* [30] further developed relaxed conditions under which stationarity of a policy implies its optimality.

Theorem 3 (see [30]): Assume that, for every fixed $\omega_0 \in \Omega_0$, $c(\omega_0, \underline{u}_N)$ is a convex differentiable function of $\underline{u}_N$ and suppose (b) in Theorem 1 holds. Assume that $\underline{\gamma}_N^* \in \mathbf{\Gamma}_N$ is a stationary policy. Let, for all $\underline{\gamma}_N \in \mathbf{\Gamma}_N$ with $\mathbb{E}[c(\omega_0, \underline{\gamma}_N(\underline{v}_N))] < \infty$

$$\mathbb{E}\left[c_{u^i}(\omega_0, \underline{\gamma}_N^*)(\gamma^i - \gamma^{i*})\right] < \infty \text{ for } i = 1, \ldots, N. \tag{3}$$

Then, $\underline{\gamma}_N^*$ is a team optimal policy for ($\mathcal{P}'_N$). If $c(\omega_0, \underline{u}_N)$ is strictly convex in $\underline{u}_N$, $\mathbb{P}$-a.s., then $\underline{\gamma}_N^*$ is unique.

Furthermore, (3) can be replaced by the following more checkable conditions [52]: Let Γ^i be Hilbert space for each $i = 1, \ldots, N$ and $\mathbb{E}[c(\omega_0, \underline{\gamma}_N(\underline{v}_N))] < \infty$ for all $\underline{\gamma}_N \in \mathbf{\Gamma}_N$. Moreover, let

$$\mathbb{E}\left[c_{u^i}(\omega_0, \underline{\gamma}_N^*) \bigg| v^i\right] \in \Gamma^i, i = 1, \ldots, N. \tag{4}$$

The above conditions follows directly from (3) when Γ^i is a Hilbert space for all $i = 1, 2, \ldots, N$. This condition can be checked for some applications; for example, LQ teams [52].

B. Problem Statement

($\mathcal{P}_\infty$) Consider a team with countably infinitely many DMs. Let $\mathbf{\Gamma} = \prod_{i \in \mathbb{N}} \Gamma^i$ be a countable but an infinite product policy space. We assume $\mathbb{U}^i = \mathbb{R}^n$, and $\mathbb{V}^i = \mathbb{R}^m$ for all $i \in \mathbb{N}$, where n and m are positive integers. Let $c : \Omega_0 \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, and the expected cost be

$$J(\underline{\gamma}) = \limsup_{N \to \infty} \frac{1}{N} \mathbb{E}^{\underline{\gamma}}\left[\sum_{i=1}^N c\left(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p\right)\right] \tag{5}$$

where we denote $\mathbb{E}^{\underline{\gamma}}[\sum_{i=1}^N c(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p)] := \mathbb{E}[\sum_{i=1}^N c(\omega_0, \gamma^i(v^i), \frac{1}{N}\sum_{p=1}^N \gamma^p(v^p))]$.

Definition 4: Team optimal solution for ($\mathcal{P}_\infty$).

For a given stochastic team problem with a given IS, a policy $\underline{\gamma}^* := (\gamma^{1*}, \gamma^{2*}, \ldots) \in \mathbf{\Gamma}$ is *optimal* for ($\mathcal{P}_\infty$) if

$$J(\underline{\gamma}^*) = \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}) =: J^*.$$

Our goal in this article is to establish conditions for a team policy to be optimal, and also connect the optimal cost and policies for $(\mathcal{P}_\infty)$ and $(\mathcal{P}_N)$. To this end, we redefine $(\mathcal{P}_N)$ for our problem statement as follows:

$(\mathcal{P}_N)$ Let $N = |\mathcal{N}|$ be the number of control actions taken and $\underline{\gamma}_N = \{\gamma^1, \ldots, \gamma^N\}$ and let $\mathbf{\Gamma}_N = \prod_i^N \Gamma^i$ space of admissible policies for the team with N-DMs. Assume an expected cost function is defined as

$$J_N(\underline{\gamma}_N) = \frac{1}{N}\mathbb{E}^{\underline{\gamma}_N}\left[\sum_{i=1}^N c(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p)\right]. \quad (6)$$

We will investigate the relations between the sequence of solutions to (6) and the solution to (5). We note that our main result is on the connection between $(\mathcal{P}_\infty)$ and $(\mathcal{P}_N)$.

III. Optimal Policies for Teams With Infinitely Many DMs

A. Sufficient Conditions of Optimality

In the following, we propose sufficient conditions of team optimality for $(\mathcal{P}_\infty)$. We often follow [30], and the result is an extension of [30] to a general setup of static teams with countably infinite number of DMs. We also note a related analysis in [35]. We will use the following theorem for LQ static teams with countably infinite number of DMs (see Section V-B).

Assumption 1: Let

A1) $c(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p)$ be a $\mathbb{R}_+$-valued jointly convex function of second and third arguments and differentiable in u^i with continuous partial derivatives, for every $\omega_0 \in \Omega_0$.

A2) for some $\underline{\gamma}^* \in \mathbf{\Gamma}$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbb{E}^{\underline{\gamma}^*}\left[c\left(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p\right)\right] < \infty. \quad (7)$$

We note that the cost function is differentiable in u^i which means that the cost is totally differentiable in u^i, i.e., $\frac{d}{du^i}c(\omega_0, u^i, \frac{1}{N}\sum_{p=1}^N u^p) = \frac{\partial}{\partial u^i}c(\omega_0, u^i, \mu_N) + \frac{1}{N}\frac{\partial}{\partial \mu_N}c(\omega_0, u^i, \mu_N)$.

Theorem 4: Assume (A1) holds and (A2) holds for $\underline{\gamma}^* \in \mathbf{\Gamma}$. If for all $\underline{\gamma} \in \mathbf{\Gamma}$ with $J(\underline{\gamma}) < \infty$

$$\limsup_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\sum_{i=1}^N\sum_{k=1}^N c_{u^k}(\omega_0, \gamma^{i*}, \mu_N^*)(\gamma^k - \gamma^{k*})\right] \geq 0 \quad (8)$$

where $\mu_N^* = \frac{1}{N}\sum_{p=1}^N \gamma^{p*}(v^p)$, then $\underline{\gamma}^*$ is a globally optimal team policy for $(\mathcal{P}_\infty)$.

Proof: Under (A1), the required derivatives in (8) in the direction of u^i exist and the chain rule of derivatives can be applied since this implies that the cost function is Fréchet differentiable in u^i [19]. Now, we use the convexity property to justify interchanging the expectation and the derivation similar to [30, Th. 2], then we use (7) and (8) to establish the global optimality of $\underline{\gamma}^*$ for $(\mathcal{P}_\infty)$. Under (A1), we have for every $\alpha \in (0, 1]$

$$\sum_{i=1}^N c\left(\omega_0, \gamma^{i*} + \alpha\delta^i, \mu_N^* + \frac{\alpha}{N}\sum_{p=1}^N \delta^p\right) - c(\omega_0, \gamma^{i*}, \mu_N^*)$$

$$\leq \alpha\sum_{i=1}^N\left(c(\omega_0, \gamma^i, \mu_N) - c(\omega_0, \gamma^{i*}, \mu_N^*)\right)$$

where $\mu_N = \frac{1}{N}\sum_{p=1}^N \gamma^p(v^p)$ and $\delta^i = \gamma^i - \gamma^{i*}$. Let

$$h_N^{\omega_0}(\alpha) := \frac{1}{\alpha}\left[\frac{1}{N}\sum_{i=1}^N c\left(\omega_0, \gamma^{i*} + \alpha\delta^i, \mu_N^* + \frac{\alpha}{N}\sum_{p=1}^N \delta^p\right) - c(\omega_0, \gamma^{i*}, \mu_N^*)\right].$$

Hence, [14, Proposition 6.3.2] implies that $h_N^{\omega_0}(\alpha)$ is a monotone nonincreasing function as $\alpha \to 0$ in $\alpha \in [0, 1]$ and bounded from above by $h_N^{\omega_0}(1)$. Thus, by [14, Corollary 6.3.3], $h'_{+,N}(\omega_0, 0) := \lim_{\alpha\to 0} h_N^{\omega_0}(\alpha)$ exists. Since $h_N^{\omega_0}(\alpha)$ is a monotonic nonincreasing function as $\alpha \to 0$ in $\alpha \in [0, 1]$ and bounded above by $h_N^{\omega_0}(1)$, and since $J(\underline{\gamma}^*)$ and $J(\underline{\gamma})$ are finite, we can choose N large enough such that $\mathbb{E}(h_N^{\omega_0}(1)) < \infty$. Now, we can use the monotone convergence theorem (see [22, page. 170]) to interchange the limit and the expectation

$$\lim_{\alpha\to 0}\mathbb{E}(h_N^{\omega_0}(\alpha)) = \mathbb{E}\left(\lim_{\alpha\to 0} h_N^{\omega_0}(\alpha)\right) = \mathbb{E}(h'_{+,N}(\omega_0, 0)). \quad (9)$$

From [30, Lemma 1], we have $\mathbb{E}(h'_{+,N}(\omega_0, 0)) = \frac{1}{N}\mathbb{E}(\sum_{i=1}^N\sum_{k=1}^N c_{u^k}(\omega_0, \gamma^{i*}, \mu_N^*)\delta^k)$. Define

$$F_{\underline{\gamma}_N}^N(\alpha) := \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^N c\left(\omega_0, \gamma^{i*} + \alpha\delta^i, \mu_N^* + \frac{\alpha}{N}\sum_{p=1}^N \delta^p\right)\right).$$

Note that $F_{\underline{\gamma}_N}^N(\alpha)$ exists for $\alpha \in [0, 1]$ since $\mathbb{E}(h_N^{\omega_0}(\alpha)) \leq \mathbb{E}(h_N^{\omega_0}(1)) < \infty$, and $\mathbb{E}(\frac{1}{N}\sum_{i=1}^N c(\omega_0, \gamma^{i*}, \mu_N^*)) < \infty$. Therefore, one can write $F'^N_{\underline{\gamma}_N^+}(0) = \lim_{\alpha\to 0}\mathbb{E}(h^{\omega_0}(\alpha))$, and

$$F'^N_{\underline{\gamma}_N^+}(0) = \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^N\sum_{k=1}^N c_{u^k}(\omega_0, \gamma^{i*}, \mu_N^*)(\gamma^k - \gamma^{k*})\right).$$

Thus, we can write

$$J(\underline{\gamma}) - J(\underline{\gamma}^*) = \limsup_{N\to\infty} F_{\underline{\gamma}_N}^N(1) - \limsup_{N\to\infty} F_{\underline{\gamma}_N}^N(0) \quad (10)$$

$$= \limsup_{N\to\infty} F_{\underline{\gamma}_N}^N(1) - \liminf_{N\to\infty} F_{\underline{\gamma}_N}^N(0) \quad (11)$$

$$\geq \limsup_{N\to\infty}\frac{F_{\underline{\gamma}_N}^N(1) - F_{\underline{\gamma}_N}^N(0)}{1} \quad (12)$$

$$\geq \limsup_{N\to\infty} F'^N_{\underline{\gamma}_N^+}(0) \geq 0 \quad (13)$$

where (11) follows from (A2) and (7), and $-\liminf_{N\to\infty} a_N = \limsup_{N\to\infty} -a_N$, $\limsup_{N\to\infty} a_N + \limsup_{N\to\infty} b_N \geq \limsup_{N\to\infty}(a_N + b_N)$ imply (12), and (13) holds since $F_{\underline{\gamma}_N}^N(.)$ is a convex function using [14, Corollary 6.3.3], and since $a_N \geq b_N$, then $\limsup_{N\to\infty} a_N \geq \limsup_{N\to\infty} b_N$. Finally, the last inequality follows from (8); hence, $J(\underline{\gamma}) - J(\underline{\gamma}^*) \geq 0$, and the proof is completed.

In some applications, (8) can be difficult to check since it must be satisfied for all $\underline{\gamma} \in \mathbf{\Gamma}$ with $J(\underline{\gamma}) < \infty$. In the next section, we address this issue by introducing a constructive approach for static teams with countably infinite number of DMs as a limit of a sequence of team optimal policies of the corresponding static teams with finite number of DMs. In the following, we propose sufficient conditions to approximate the optimal cost and a team optimal policy for static teams with countably infinite number of DMs using the optimal cost and an optimal policy for static teams with N DMs. We note that our first result here is based on [35, Th. 1], which considered an equality. We denote $\underline{\gamma}|_N \in \mathbf{\Gamma_N}$ as a restriction of $\underline{\gamma} \in \mathbf{\Gamma}$ to the first N components.

Theorem 5: Let $\underline{\gamma}_N^* \in \mathbf{\Gamma}_N$ be an optimal policy for $(\mathcal{P}_N)$ as (6) (see [21], [30], [53] for sufficient conditions). If there exists $\underline{\gamma}^* \in \mathbf{\Gamma}$, with $J(\underline{\gamma}^*) < \infty$, satisfying

$$\limsup_{N\to\infty} J_N(\underline{\gamma}_N^*) \geq J(\underline{\gamma}^*) \tag{14}$$

then, $\underline{\gamma}^*$ is a globally team optimal policy for $(\mathcal{P}_\infty)$.

Proof: We have

$$J(\underline{\gamma}^*) \leq \limsup_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}^{\underline{\gamma}_N^*}\left(c(\omega_0, u^i, \mu_N)\right) \tag{15}$$

$$= \limsup_{N\to\infty} \inf_{\underline{\gamma}_N \in \mathbf{\Gamma}_N} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}^{\underline{\gamma}_N}\left(c(\omega_0, u^i, \mu_N)\right) \tag{16}$$

$$= \limsup_{N\to\infty} \inf_{\underline{\gamma} \in \mathbf{\Gamma}} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}^{\underline{\gamma}}\left(c(\omega_0, u^i, \mu_N)\right) \tag{17}$$

$$\leq \inf_{\underline{\gamma} \in \mathbf{\Gamma}} \limsup_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}^{\underline{\gamma}}\left(c(\omega_0, u^i, \mu_N)\right)$$

$$= \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}) \tag{18}$$

where $\mu_N := \frac{1}{N}\sum_{p=1}^{N} u^p$ and (15) follows from (14), and (16) is true since γ_N^* is a team optimal policy for $(\mathcal{P}_N)$ [see (6)]. Furthermore, (17) follows from the fact that $[\underline{\gamma}|_N : \underline{\gamma} \in \mathbf{\Gamma}] = \mathbf{\Gamma_N}$, where $\underline{\gamma}|_N$ is $\underline{\gamma}$ restricted to the first N components. ■

Remark 1: Under (A2), one can replace (14) with

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \left[\mathbb{E}^{\underline{\gamma}_N^*}\left(c(\omega_0, u^i, \mu_N)\right) - \mathbb{E}^{\underline{\gamma}^*}\left(c(\omega_0, u^i, \mu_N)\right)\right] \geq 0. \tag{19}$$

The above theorem and remark will be useful for some applications (see for example Section V-D).

B. Asymptotically Optimal Policies as a Limit of Finite Team Optimal Policies

In the following, we present a sufficient condition for (14). The following result also presents a constructive method to obtain optimal policies using asymptotic analysis.

Theorem 6: Assume

a) for every N, there exist $\underline{\gamma}_N^* \in \mathbf{\Gamma}_N$ for $(\mathcal{P}_N)$ [see (6)],

b) let $\omega \in B$ for some $B \in \mathcal{F}$ event of $\mathbb{P}$ measure one, for every fixed $v^i(\omega)$, $\gamma_N^{i*}(v^i)$ converges to $\gamma_\infty^{i*}(v^i)$ uniformly in $i = 1, 2, \ldots, N$, i.e.,

$$\lim_{N\to\infty} \sup_{1\leq i\leq N} |\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| = 0 \mathbb{P} - a.s.$$

c) there exists a $\mathbb{P}$-integrable function $g(\omega_0, \underline{v})$ such that, for every N

$$\frac{1}{N}\sum_{i=1}^{N} c\left(\omega_0, \gamma_\infty^{i*}(v^i), \frac{1}{N}\sum_{p=1}^{N}\gamma_\infty^{p*}(v^p)\right) \leq g(\omega_0, \underline{v})$$

where $\underline{v} = (v^1, v^2, \ldots)$, then $\underline{\gamma}^*$, a team optimal policy for $(\mathcal{P}_\infty)$, is a pointwise limit of $\underline{\gamma}_N^*$, an optimal policy for $(\mathcal{P}_N)$, i.e., $\gamma^{i*}(v^i) = \lim_{N\to\infty} \gamma_N^{i*}(v^i) = \gamma_\infty^{i*}(v^i)$ $\mathbb{P}$-almost surely.

Proof: According to Theorem 5, we only need to show that

$$\limsup_{N\to\infty} J_N(\underline{\gamma}_N^*) \geq \liminf_{N\to\infty} J_N(\underline{\gamma}_N^*)$$

$$\geq E\left(\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} c(\omega_0, \gamma_N^{i*}(v^i), \mu_N^*)\right)$$

$$= \lim_{N\to\infty} J_N(\underline{\gamma}_\infty^*)$$

where $\mu_N^* = \frac{1}{N}\sum_{p=1}^{N} \gamma_N^{p*}(v^p)$ and the second inequality follows from Fatou's lemma (since the cost function is nonnegative). In the following, we justify the equality above. On a set of $\mathbb{P}$ measure one, $\omega \in B$ where $B \in \mathcal{F}$, for every fixed $v^i(\omega)$ in this set, define $\underline{v}(\omega) = (v^1(\omega), v^2(\omega), \ldots)$ and $\underline{v}_N(\omega) = (v^1(\omega), \ldots, v^N(\omega))$. We follow three steps to prove the theorem.

Step 1: We show that on a set of $\mathbb{P}$ measure one, $\omega \in B$ where $B \in \mathcal{F}$, for every fixed $v^i(\omega)$ in this set $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}(\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)) = 0$. For a fixed $\underline{v}$, following from (b) for a given $\delta_{\omega,\underline{v}_N} := \sup_{1\leq i\leq N} |\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| > 0$, there exists $\hat{N}(\delta_{\omega,\underline{v}_N}) \in \mathbb{N}$ such that for $N > \hat{N}(\delta_{\omega,\underline{v}_N})$, $|\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| \leq \delta_{\omega,\underline{v}_N}$ for every $i = 1, \ldots, N$, where $\lim_{N\to\infty} \delta_{\omega,\underline{v}_N} = 0$ $\mathbb{P}$-almost surely. We have $\mathbb{P}$-almost surely

$$\left|\frac{1}{N}\sum_{i=1}^{N}\left(\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)\right)\right| < \frac{1}{N}\sum_{i=1}^{N} \delta_{\omega,\underline{v}_N} = \delta_{\omega,\underline{v}_N}$$

and since $\lim_{N\to\infty} \sup_{1\leq i\leq N} |\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| = 0$, we have $\lim_{N\to\infty} \delta_{\omega,\underline{v}_N} = 0$. Hence, we can show that $\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \gamma_N^{i*}(v^i) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \gamma_\infty^{i*}(v^i)$. Following from continuity, $c(\omega_0, \gamma_N^{i*}(v^i), \mu_N^*)$ converges to $c(\omega_0, \gamma_\infty^{i*}(v^i), \lim_{N\to\infty} \mu_\infty^*)$ $\mathbb{P}$-a.s. for every $i = 1. \ldots, N$.

Step 2: We show that $c(\omega_0, \gamma_N^{i*}(v^i), \mu_N^*)$ converges to $c(\omega_0, \gamma_\infty^{i*}(v^i), \lim_{N\to\infty} \mu_\infty^*)$ uniformly in $i = 1, \ldots, N$ $\mathbb{P}$-almost surely, where $\mu_\infty^* = \frac{1}{N}\sum_{p=1}^{N} \gamma_\infty^{p*}(v^p)$. By continuity of the cost function, we have for a given $\epsilon_{\omega,\underline{v}_N} > 0$, there exists $\delta_{\omega,\underline{v}_N} > 0$ such that $|\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| < \delta_{\omega,\underline{v}_N}$, and $|\frac{1}{N}\sum_{i=1}^{N}(\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i))| < \delta_{\omega,\underline{v}_N}$ implies $|c(\omega_0, \gamma_N^{i*}(v^i), \mu_N^*) - c(\omega_0, \gamma_\infty^{i*}(v^i), \mu_\infty^*)| < \epsilon_{\omega,\underline{v}_N}$ $\mathbb{P}$-almost surely for every $i = 1, \ldots, N$. Following from (Step 1), we have for $N > \hat{N}(\delta_{\omega,\underline{v}_N}(\epsilon_{\omega,\underline{v}_N}))$, $|\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| < \delta_{\omega,\underline{v}_N}$, and $|\frac{1}{N}\sum_{i=1}^{N}(\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i))| < \delta_{\omega,\underline{v}_N}$. Hence, $\mathbb{P}$-a.s.

$$|c(\omega_0, \gamma_N^{i*}(v^i), \mu_N^*) - c(\omega_0, \gamma_\infty^{i*}(v^i), \mu_\infty^*)| < \epsilon_{\omega,\underline{v}_N}$$

where $\lim_{N\to\infty} \epsilon_{\omega,\underline{v}_N} = 0$.

Step 3: In this step, we show that $\mathbb{P}$-a.s.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(c(\omega_0,\gamma_N^{i*}(v^i),\mu_N^*)-c(\omega_0,\gamma_\infty^{i*}(v^i),\mu_\infty^*)\right)=0.$$

According to (Step 2), for $N>\hat{N}(\delta_{\omega,\underline{v}_N}(\epsilon_{\omega,\underline{v}_N}))$, we have $\mathbb{P}$-a.s.

$$\left|\frac{1}{N}\sum_{i=1}^{N}c(\omega_0,\gamma_N^{i*}(v^i),\mu_N^*)-c(\omega_0,\gamma_\infty^{i*}(v^i),\mu_\infty^*)\right|<\epsilon_{\omega,\underline{v}_N}.$$

Following from (c), we can interchange the limit and the integral using the dominated convergence theorem, and the proof is completed.

Remark 2: One can relax conditions in Theorem 6 as follows.

i) relax (a) by considering a sequence of ϵ_N-optimal policy, where ϵ_N are non-negative and converges to zero as $N\to\infty$,

ii) relax (c) with a uniform integrability condition, which is satisfied if the following expression is finite (see [6, Th. 3.5])

$$\sup_{N\geq 1}\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N}c\left(\omega_0,\gamma_\infty^{i*}(v^i),\frac{1}{N}\sum_{i=1}^{N}\gamma_\infty^{i*}(v^i)\right)\right|^{1+\epsilon}\right]$$

for some $\epsilon>0$. This new condition can be checked in some applications (see Section V). The result follows from [6, Th. 3.5],

iii) relax the convergence $\mathbb{P}$-almost surely in (b) by considering convergence in probability, i.e.,

$$\lim_{N\to\infty}\mathbb{P}\left(\sup_{1\leq i\leq N}\left|\gamma_N^{i*}(v^i)-\gamma_\infty^{i*}(v^i)\right|\geq\epsilon\right)=0$$

hence similar to the proof of Theorem 6 (Step 1), using continuous mapping theorem (see for example, [6, page 20]), we can show that $c(\omega_0,\gamma_N^{i*}(v^i),\mu_N^*)$ converges to $c(\omega_0,\gamma_\infty^{i*}(v^i),\lim_{N\to\infty}\mu_\infty^*)$ in probability. Similarly, the result of (Step 2) holds in probability. Using [6, Th. 3.5], under the uniform integrability of $X_N:=\frac{1}{N}\sum_{i=1}^{N}c(\omega_0,\gamma_\infty^{i*}(v^i),\frac{1}{N}\sum_{i=1}^{N}\gamma_\infty^{i*}(v^i))$ and under the convergence in probability of X_N to $X:=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}c(\omega_0,\gamma_\infty^{i*}(v^i),\frac{1}{N}\sum_{i=1}^{N}\gamma_\infty^{i*}(v^i))$, we can conclude that $\mathbb{E}(X_N)\to\mathbb{E}(X)$. This relaxation can be useful when the weak law of large numbers can be invoked to check (c), but the strong law of large numbers (SLLNs) fails to apply.

We apply the results of this section to two examples in Sections V-A and V-B.

In the following section, we show that under symmetry of optimal policies, sufficient conditions of optimality can be satisfied quite effortlessly.

IV. Globally Optimal Policies for Mean-Field Teams

A. Symmetric Teams

In the following, we present sufficient conditions for team optimality in symmetric and mean-field teams. The concept of symmetry has been studied in a variety of contexts (see, e.g., [13] and [38] and many others).

Definition 5 (Exchangeable teams): An N-DM team is *exchangeable* if the value of the expected cost function [see (1)] is invariant under every permutation of policies.

We note that it is also called *totally symmetric* in a game theoretic context (see, for example, [13]).

Definition 6 (Symmetrically optimal teams): A team is *symmetrically optimal*, if for every given policy, there exists an identically symmetric policy (i.e., each DM has the same policy), which performs at least as good as the given policy.

In the following, we characterize the symmetry of the general setup for $(\mathcal{P}_N')$ [see (1)] defined in Section II-A. Clearly, the result will also hold for the $(\mathcal{P}_N)$ [see (6)] defined in Section II-B. First, we recall the definition of an *exchangeable* finite set of random variables.

Definition 7: Random variables $x^1,x^2,\ldots,x^N$ are *exchangeable* if any permutation, σ, of the set of indexes $\{1,\ldots,N\}$ fails to change the joint probability measures of random variables, i.e., $\mathbb{P}(dx^{\sigma(1)},dx^{\sigma(2)},\ldots,dx^{\sigma(N)})=\mathbb{P}(dx^1,dx^2,\ldots,dx^N)$.

Lemma 1: For a fixed N, consider an N-DM team defined as $(\mathcal{P}_N')$ [see (1)] and let the cost function be a convex function of $\underline{u}_N$ $\mathbb{P}$-almost surely. Assume the cost function is exchangeable $\mathbb{P}$-almost surely with respect to the actions, i.e., for any permutation of indexes, σ, $\mathbb{P}$-almost surely $c(\omega_0,u^1,\ldots,u^N)=c(\omega_0,u^{\sigma(1)},\ldots,u^{\sigma(N)})$. If $\mathbb{U}$ is convex, and observations of DMs are exchangeable conditioned on ω_0, then the team is symmetrically optimal.

Proof: Any permutation of policies does not deviate the value of $J_N(\underline{\gamma}_N)$ since

$$\begin{aligned}
&J_N(\underline{\gamma}_N^\sigma)\\
&=\int c(\omega_0,u^1,\ldots,u^N)\mathbb{P}_N(dv^1,\ldots,dv^N|\omega_0)\\
&\quad\times 1_{\{(\gamma^{\sigma(1)}(v^1),\ldots,\gamma^{\sigma(N)}(v^N))\}}(du^1,\ldots,du^N)\mathbb{P}(d\omega_0)\\
&=\int c(\omega_0,u^{\sigma(1)},\ldots,u^{\sigma(N)})\\
&\quad\times 1_{\{(\gamma^{\sigma(1)}(v^{\sigma(1)}),\ldots,\gamma^{\sigma(N)}(v^{\sigma(N)}))\}}(du^{\sigma(1)},\ldots,du^{\sigma(N)})\\
&\quad\times\mathbb{P}_N(dv^{\sigma(1)},\ldots,dv^{\sigma(N)}|\omega_0)\mathbb{P}(d\omega_0)\\
&=\int c(\omega_0,u^1,\ldots,u^N)1_{\{(\gamma^1(v^1),\ldots,\gamma^N(v^N))\}}(du^1,\ldots,du^N)\\
&\quad\times\mathbb{P}_N(dv^1,\ldots,dv^N|\omega_0)\mathbb{P}(d\omega_0)\\
&=J_N(\underline{\gamma}_N)
\end{aligned}\tag{20}$$

where (20) follows from the assumption that the cost function is exchangeable with respect to the actions, and the hypothesis that observations of DMs are $\mathbb{P}$-almost surely exchangeable conditioned on the random variable ω_0. Let $\underline{\gamma}_N^*=(\gamma^{1*},\gamma^{2*},\ldots,\gamma^{N*})$ be a given team policy for $(\mathcal{P}_N')$ [see (1)]. Consider $\underline{\tilde{\gamma}}_N$ as a convex combination of all possible permutations of policies by averaging them, $\sigma\in\Sigma$, where Σ is the set of all possible permutation. Since $\mathbb{U}$ is convex, $\underline{\tilde{\gamma}}_N$ is a control policy. Following from convexity of the cost function $\mathbb{P}$-almost surely, we have for $\alpha_\sigma=\frac{1}{|\Sigma|}$ (where $|\Sigma|$ denotes the cardinality of Σ)

$$J_N(\underline{\tilde{\gamma}}_N):=J_N(\sum_{\sigma\in\Sigma}\alpha_\sigma\underline{\gamma}_N^{*,\sigma})\leq\sum_{\sigma\in\Sigma}\alpha_\sigma J_N(\underline{\gamma}_N^{*,\sigma})$$

$$= \sum_{\sigma\in\Sigma} \alpha_\sigma J_N(\underline{\gamma}_N^*) \ = J_N(\underline{\gamma}_N^*)$$

where the inequality follows from convexity of the cost function $\mathbb{P}$-almost surely for every fixed realization of observations since we have

$$\mathbb{E}\left[c\left(\omega_0, \sum_{\sigma\in\Sigma}\alpha_\sigma(\underline{\gamma}_N^{*,\sigma})^1(v^1),\ldots,\sum_{\sigma\in\Sigma}\alpha_\sigma(\underline{\gamma}_N^{*,\sigma})^N(v^N)\right)\right]$$
$$\leq \mathbb{E}\left[\sum_{\sigma\in\Sigma}\alpha_\sigma c\left(\omega_0,(\underline{\gamma}_N^{*,\sigma})^1(v^1),\ldots,(\underline{\gamma}_N^{*,\sigma})^N(v^N)\right)\right]$$
$$= \sum_{\sigma\in\Sigma}\alpha_\sigma \mathbb{E}\left[c\left(\omega_0,(\underline{\gamma}_N^{*,\sigma})^1(v^1),\ldots,(\underline{\gamma}_N^{*,\sigma})^N(v^N)\right)\right]$$

where $(\underline{\gamma}_N^{*,\sigma})^j$ denotes the jth component of $\underline{\gamma}_N^{*,\sigma}$, and the inequality above follows from Jensen's inequality since the cost function is convex $\mathbb{P}$-almost surely. Hence, the team is symmetrically optimal. ■

In the following, we present another characterization of symmetrically optimal teams; this looks to be a standard result; however, a proof is included for completeness since we could not find an explicit reference.

Lemma 2: For a fixed N, consider an N-DM team defined as $(\mathcal{P}'_N)$ [see (1)] and let the cost function be a convex function of $\underline{u}_N$ $\mathbb{P}$-almost surely. Assume the set of action space for each DM is convex. If the expected cost function [see (1)] is exchangeable with respect to the policies, then the team is symmetrically optimal.

Proof: Let $\underline{\gamma}_N^* = (\gamma^{1*},\gamma^{2*},\ldots,\gamma^{N*})$ be a given team policy for $(\mathcal{P}'_N)$ [see (1)]. According to the definition of exchangeable teams, any permutation of policies, say $\hat{\underline{\gamma}}_N^* = (\gamma^{i_1*},\gamma^{i_2*},\ldots,\gamma^{i_N*})$, fails to change the value of the expected cost function, and hence achieve the same expected cost as the one induced by $\underline{\gamma}_N^*$. Consider $\tilde{\underline{\gamma}}_N$ as a uniform randomization among all possible permutations of optimal policies, since $\mathbb{U}$ is convex, then $\tilde{\underline{\gamma}}_N$ is a control policy. By convexity of the cost function, through Jensen's inequality, and the fact that any permutation of optimal policies preserves the value of the cost function, we have $J_N(\tilde{\underline{\gamma}}_N) \leq J_N(\underline{\gamma}_N^*)$. Since $\tilde{\underline{\gamma}}_N$ is also identically symmetric, the proof is completed. ■

Now, we characterize symmetrically optimal teams for $(\mathcal{P}_N)$ [see (6)].

Theorem 7: Consider an N-DM team defined as $(\mathcal{P}_N)$ [see (6)] in Section II-B. Let action spaces be convex and the cost function be convex in the second and third arguments $\mathbb{P}$-almost surely. If observations are exchangeable conditioned on ω_0, then the team is symmetrically optimal.

Proof: The cost function defined in $(\mathcal{P}_N)$ [see (6)] is exchangeable in actions; hence, under convexity of the action spaces and the cost function and following from the hypothesis that observations are exchangeable condition on ω_0, the proof is completed using Lemma 1. ■

Theorem 7 will be utilized in our analysis to follow.

B. Optimal Solutions for Mean-Field Teams as Limits of Optimal Policies for Finite Symmetric Teams

In the following, we present results for symmetrically optimal static teams. First, we focus on the case that the observations of DMs are identical and independent, then we deal with nonidentical and dependent observations under additional assumptions. As we noted earlier, mean-field games studied in [18] belong to this class in a game theoretic context; in [18], concentration of measures arguments and independence of measurements have been utilized to justify the convergence of equilibria (person-by-person-optimality in the team setup). We also note that Jovanovic *et al.* [29] and [37] have considered symmetry conditions for mean-field games. In the context of LQ mean-field teams, Arabneydi and Mahajan [2] has considered a setup where DMs share the mean field in the system either completely or partially (through showing that a centralized performance can be attained under the restricted IS). Also, for the LQ setup under the assumption that DMs apply an identical policy in addition to some technical assumptions, Arabneydi and Aghdam [1] showed that the expected cost achieved by a suboptimal fully decentralized strategy is on $\epsilon(n)$ neighborhood of the optimal expected cost achieved when mean field (empirical distribution of states) has been shared, where n is the number of players. In [28], a continuous-time setup with a major agent has been studied.

Remark 3: We note that, in [10, Ch. 6 Vol. I] and [32, Sec. 2.4], the connection between solutions of N-player differential control systems and solutions of Mckean–Vlasov control problems has been investigated under either the assumption that the IS is classical (i.e., the problem is centralized) since the controls, u_t^i, for each player are assumed to be progressively measurable with respect to the filtration generated by all initial states, $(X_0^1,\ldots,X_0^N)$ and Wiener processes of all DMs $(\{(W_s^1,\ldots,W_s^N), s\leq t\})$, or by imposing structural assumptions on the controllers where controllers assumed to belong to the open-loop class (with their definition being, somewhat nonstandard, that u_t^i are progressively measurable with respect to the filtration generated by initial states and Wiener processes instead of the path of states X_s^i for $s\leq t$) or to belong to Markovian controllers (i.e., $u_t^i = \phi^i(t, X_t^i)$ where ϕ^i are measurable functions) [10, pages 72-76],[32]. Also, in [32, Th. 2.11], it has been shown that a sequence of relaxed (measure-valued) open-loop ϵ_N-optimal policies for N-player differential control systems (with only coupling on states) converges to a relaxed open-loop Mckean–Vlasov control optimal solution. Under additional assumptions, the existence of a strong solution and a Markovian optimal solution of McKean–Vlasov solution has been established [32, Th. 2.12 and Corollary 2.13]. In the mean-field team setup, under the decentralized IS, it is not clear *a priori* whether the limsup of the expected cost function and states of dynamics for N-DM teams converge to the limit. In fact, the IS of the team problem can break the symmetry and also can prevent establishing a limit theory (for example, by considering a partial sharing of observations between DMs). Here, by focusing on the decentralized setup and by considering mean-field coupling of controls, using a convexity argument and symmetry, we show that a sequence of optimal policies for $(\mathcal{P}_N)$ converges pointwise to an optimal policy for $(\mathcal{P}_\infty)$.

Our next theorem, under the assumption that observations are independent and identically distributed (i.i.d.), utilizes a measure concentration argument to establish a convergence result.

Theorem 8: Consider a team defined as $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action space be compact and convex for each DM, and v^is be i.i.d. random variables. If there exists a sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)], $\{\underline{\gamma}_N^*\}_N$, which converges (for every DM due to the symmetry) pointwise to $\underline{\gamma}_\infty^*$ as $N\to\infty$, then $\underline{\gamma}_\infty^*$ (which is identically symmetric) is an optimal policy for $(\mathcal{P}_\infty)$.

Proof: Action spaces and the cost function are convex and following from the hypothesis that v^is are i.i.d. random variables (hence they are exchangeable conditioned on ω_0) and the result of Theorem 7, one can consider a sequence of N-DM teams, which are symmetrically optimal that defines $(\mathcal{P}_N)$ [see (6)] and whose limit is identified with $(\mathcal{P}_\infty)$. Define empirical measures on actions and observation of $Q_N(B) := \frac{1}{N}\sum_{i=1}^N \delta_{\zeta_N^i}(B)$, and $\tilde{Q}_N(B) := \frac{1}{N}\sum_{i=1}^N \delta_{\zeta_\infty^i}(B)$, where $B \in \mathcal{Z} := \mathbb{U}\times\mathbb{V}$, $\zeta_N^i := (\gamma_N^*(v^i), v^i)$, $\zeta_\infty^i := (\gamma_\infty^*(v^i), v^i)$, and $\delta_Y(\cdot)$ is the Dirac measure for any random variable Y. In the following, we first show that $\tilde{Q}_N$ converges weakly to $Q = Law(\zeta_\infty^i)$ $\mathbb{P}$-almost surely, then we show (14) holds, and we invoke Theorem 5.

Step 1: For every $g \in C_b(\mathcal{Z})$, where we denote $C_b(X)$ as the space of continuous and bounded functions in X, we have

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\int g dQ_N - \int g d\tilde{Q}_N\right| \ge \epsilon\right)$$

$$\le \epsilon^{-1}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\left|g(\gamma_N^*(v^i), v^i) - g(\gamma_\infty^*(v^i), v^i)\right|\right] \quad (21)$$

$$= \epsilon^{-1}\lim_{N\to\infty}\mathbb{E}\left[\left|g(\gamma_N^*(v^i), v^i) - g(\gamma_\infty^*(v^i), v^i)\right|\right] \quad (22)$$

$$= \epsilon^{-1}\mathbb{E}\left[\lim_{N\to\infty}\left|g(\gamma_N^*(v^i), v^i) - g(\gamma_\infty^*(v^i), v^i)\right|\right] = 0 \quad (23)$$

where (21) follows from Markov's inequality, the triangle inequality, and the definition of the empirical measure, and (22) follows from the hypothesis that v^is are identical random variables. Since g is bounded and continuous, the dominated convergence theorem implies (23). Hence, for every subsequence, there exists a subsequence such that $|\int g dQ_{N_{k_l}} - \int g d\tilde{Q}_{N_{k_l}}|$ converges to zero $\mathbb{P}$-almost surely as $l \to \infty$. On the other hand, since v^is are i.i.d. random variables, the SLLNs imply that $\tilde{Q}_N$ converges weakly to Q $\mathbb{P}$-almost surely, that is, $|\int g d\tilde{Q}_N - \int g dQ|$ converges to zero $\mathbb{P}$-almost surely for every $g \in C_b(\mathcal{Z})$. Hence, through choosing a suitable subsequence, Q_{N_k} converges $\mathbb{P}$-almost sure weakly to Q since for every continuous and bounded function g, we have $\mathbb{P}$-a.s.

$$\begin{aligned}&\lim_{N\to\infty}\left|\int g dQ_N - \int g dQ\right|\\ &\le \lim_{N\to\infty}\left(\left|\int g dQ_N - \int g d\tilde{Q}_N\right| + \left|\int g d\tilde{Q}_N - \int g dQ\right|\right)\\ &= 0.\end{aligned} \quad (24)$$

Step 2: Following from [16, Lemma 1.5] and [46, Th. 3.5], or [33, Th. 3.1] using the fact that the cost function is non-negative and continuous, we have

$$\begin{aligned}&\limsup_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[c\left(\omega_0, \gamma_N^*(v^i), \frac{1}{N}\sum_{i=1}^N \gamma_N^*(v^i)\right)\right]\\ &\ge \liminf_{N\to\infty}\mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V})\right)\right.\right.\\ &\qquad\left.\left.\times Q_N(du, dv)\middle|\omega_0\right]\right]\\ &\ge \mathbb{E}\left[\mathbb{E}\left[\liminf_{N\to\infty}\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V})\right)\right.\right.\\ &\qquad\left.\left.\times Q_N(du, dv)\middle|\omega_0\right]\right]\\ &\ge \mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ(du\times\mathbb{V})\right) Q(du, dv)\middle|\omega_0\right]\right]\end{aligned} \quad (25)$$

where the first inequality follows from the definition of Q_N and replacing limsup by liminf. The second inequality follows from Fatou's lemma. In the following, we justify (25). Since Q_N converges weakly to Q $\mathbb{P}$-almost surely, using continuous mapping theorem [6, page 20], we have $Q_N(du\times\mathbb{V})$ converges weakly to $Q(du\times\mathbb{V})$ $\mathbb{P}$-almost surely, hence the compactness of $\mathbb{U}$ implies $\int_{\mathbb{U}} uQ_N(du\times\mathbb{V}) \to \int_{\mathbb{U}} uQ(du\times\mathbb{V})$ $\mathbb{P}$-almost surely, and continuity of the cost function $\mathbb{P}$-almost surely implies $c(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V}))$ converges to $c(\omega_0, u, \int_{\mathbb{U}} uQ(du\times\mathbb{V}))$ $\mathbb{P}$-almost surely. Define a non-negative bounded sequence $G_N^M := \min\{M, c(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V}))\}$, where $G_M^N \uparrow G^N := c(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V}))$ as $M\to\infty$, then we have $\mathbb{P}$-almost surely

$$\begin{aligned}&\liminf_{N\to\infty}\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V})\right) Q_N(du, dv)\\ &= \lim_{M\to\infty}\liminf_{N\to\infty}\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du\times\mathbb{V})\right) Q_N(du, dv)\\ &\ge \lim_{M\to\infty}\liminf_{N\to\infty}\int_{\mathcal{Z}} G_N^M Q_N(du, dv)\\ &= \lim_{M\to\infty}\int_{\mathcal{Z}} G^M Q(du, dv)\\ &= \int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ(du\times\mathbb{V})\right) Q(du, dv)\end{aligned}$$

where the first inequality follows from the definition of G_N^M and the second equality is true using [46, Th. 3.5] since G_N^M is bounded (hence it is uniformly Q_N-integrable) and continuously converges to G^M, and the monotone convergence theorem implies the last equality. Hence, (25) holds, which implies (14), and the proof is completed using Theorem 5.

Remark 4: The proof above reveals that if $\mathbb{P}$-almost surely the sequence $\{Q_N\}_N$ converges weakly to Q, then Theorem 8 can be generalized to a class of team problems defined as $(\mathcal{P}_\infty)$ [see (5)], which may include ones with a nonconvex cost function and/or the ones with conditionally nonexchangeable observations: This relaxation contains a class of problems (see, e.g., Example 4 in Section V-C1) where one can consider a sequence of N-DM teams, which admits asymmetric optimal policies that define $(\mathcal{P}_N)$ [see (6)], but whose limit is identified with $(\mathcal{P}_\infty)$ under an optimal sequence of policies.

In the following, we relax the hypothesis that observations of DMs are independent.

Proposition 1: Consider a team defined as $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action space be compact and convex for each DM, and $v^i = h(x, z^i)$, where z^is are i.i.d. random variables. If there exists a sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)], $\{\gamma_N^*\}_N$, which converges pointwise to γ_∞^* as $N\to\infty$, then γ_∞^* (which is identically symmetric) is an optimal policy for $(\mathcal{P}_\infty)$.

Proof: Since z^is are i.i.d. random variables, observations, $v^i = h(x, z^i)$, have identical distributions (but are not independent), and similar to the proof of Theorem 8, using symmetry, one can show (23) holds. In the following, we show (24) and (25) hold

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\int g d\tilde{Q}_N - \int g dQ\right| \geq \epsilon\right)$$
$$\leq \frac{1}{\epsilon^2}\lim_{N\to\infty} \mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N} g(\gamma_\infty^*(v^i), v^i) - \mathbb{E}(g(\gamma_\infty^*(v^1), v^1))\right|^2\right] \tag{26}$$
$$= \lim_{N\to\infty}(\epsilon)^{-2}\mathbb{E}\left[\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N} L(\gamma_\infty^*(v^i), v^i)\right|^2 \middle| x\right]\right] \tag{27}$$
$$= 0 \tag{28}$$

where $L(\gamma_\infty^*(v^i), v^i) := g(\gamma_\infty^*(v^i), v^i) - \mathbb{E}(g(\gamma_\infty^*(v^1), v^1)|x)$, and (26) follows from Chebyshev's inequality, and (27) follows from the law of iterated expectations. The structure $v^i = h(x, z^i)$ implies conditional independence of v^is given x, hence, using the law of large numbers and since $g \in C_b(\mathcal{Z})$, we have (28), and this implies $\tilde{Q}_{N_k}$ converges weakly to $Law(\zeta_\infty^i|x)$ $\mathbb{P}$-almost surely as $k \to \infty$, hence through choosing a suitable subsequence, $Q_{N_{kl}}$ converges $\mathbb{P}$-almost sure weakly to $Q = Law(\zeta_\infty^i|x)$ as $l \to \infty$ and the rest of the proof to justify (25) is the same as that of Theorem 8. ∎

Remark 5: Existence of optimal policies for $(\mathcal{P}_N)$ and dynamic teams satisfying static reduction have been studied in [21] and [51]. In [51, Th. 4.8], the existence of optimal policies achieved under σ-compactness of each DM's action space and under mild conditions on the control law and the cost function. Hence, the existence of identically symmetric optimal policies for $(\mathcal{P}_N)$ [see (6)] follows from symmetry and [51, Th. 4.8]; thus, the existence result for $(\mathcal{P}_\infty)$ is obtained under assumptions of Theorem 8.

In the following, action spaces need not be compact; this is particularly important for LQG models as we will see in the next section.

Theorem 9: Consider a team defined as $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action spaces be convex for each DM. Let v^is be i.i.d. random variables. If there exists a sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)], $\{\gamma_N^*\}_N$, which converges pointwise to γ_∞^* as $N \to \infty$, and A3) for some $\delta > 0$, $\sup_{N\geq 1}\mathbb{E}(|\gamma_N^*(v^1)|^{1+\delta}) < \infty$ then γ_∞^* (which is identically symmetric) is an optimal policy for $(\mathcal{P}_\infty)$.

Proof: In the following, we just show $\int_{\mathbb{U}} uQ_N(du \times \mathbb{V}) \to \int_{\mathbb{U}} uQ(du \times \mathbb{V})$ $\mathbb{P}$-almost surely, and the rest of the proof follows from that of Theorem 8. We have

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\int_{\mathbb{U}} uQ_N(du \times \mathbb{V}) - \int_{\mathbb{U}} ud\tilde{Q}_N(du \times \mathbb{V})\right| \geq \epsilon\right)$$
$$\leq \epsilon^{-1}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\left|\gamma_N^*(v^i) - \gamma_\infty^*(v^i)\right|\right] \tag{29}$$
$$= \epsilon^{-1}\lim_{N\to\infty}\mathbb{E}\left[\left|\gamma_N^*(v^1) - \gamma_\infty^*(v^1)\right|\right] \tag{30}$$
$$= \epsilon^{-1}\mathbb{E}\left[\lim_{N\to\infty}\left|\gamma_N^*(v^1) - \gamma_\infty^*(v^1)\right|\right] = 0 \tag{31}$$

where (29) follows from Markov's inequality and the triangle inequality, and (30) is true since observations have identical distributions, and (31) follows from the uniform integrability assumption (A3) and the pointwise convergence of γ_N^* using [6, Th. 3.5]. On the other hand, SLLN implies $\mathbb{P}$-almost surely that $\int_{\mathbb{U}} u\tilde{Q}_N(du \times \mathbb{V}) = \frac{1}{N}\sum_{i=1}^{N}\gamma_\infty^*(v^i) \to \int_{\mathbb{U}} uQ(du \times \mathbb{V})$, and this completes the proof.

In the following, we present a result for monotone mean-field coupled teams.

Theorem 10: Consider a team defined as $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action spaces be convex for each DM. Let the cost function be increasing in the last argument, and v^is be i.i.d. random variables. If there exists a sequence of optimal policies for $(\mathcal{P}_N)$, $\{\gamma_N^*\}_N$ [see (6)], which converges pointwise to γ_∞^*, then γ_∞^* as $N \to \infty$ (which is identically symmetric) is an optimal policy for $(\mathcal{P}_\infty)$.

Proof: We show (14) holds, then we invoke Theorem 5. We use the same definitions in Theorem 8 for measures Q_N and Q. We have

$$\mathbb{E}\left[\mathbb{E}\left[\liminf_{N\to\infty}\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du \times \mathbb{V})\right) Q_N(du, dv)\middle|\omega_0\right]\right]$$
$$\geq \mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}}\liminf_{N\to\infty} c\left(\omega_0, u, \int_{\mathbb{U}} uQ_N(du \times \mathbb{V})\right) \times Q(du, dv)\middle|\omega_0\right]\right] \tag{32}$$
$$\geq \mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}} c\left(\omega_0, u, \int_{\mathbb{U}} uQ(du \times \mathbb{V})\right) Q(du, dv)\middle|\omega_0\right]\right] \tag{33}$$

where (32) follows from a version of Fatou's lemma in [15, Th. 1.1], and (33) is true since from the lower semicontinuity of $\int_{\mathbb{U}} uQ_N(du \times \mathbb{V})$, we have $\liminf_{N\to\infty}\int_{\mathbb{U}} uQ_N(du \times \mathbb{V}) \geq \int_{\mathbb{U}} uQ(du \times \mathbb{V})$, and continuity and the hypothesis that the cost function is increasing in the last argument imply for all $u \in \mathbb{U}$, $\liminf_{N\to\infty} c(\omega_0, u, \int_{\mathbb{U}} uQ_N(du \times \mathbb{V})) \geq c(\omega_0, u, \int_{\mathbb{U}} uQ(du \times \mathbb{V}))$ $\mathbb{P}$-almost surely, and this completes the proof. ∎

In the following, observations need not be identical or independent.

Theorem 11: Consider a team defined as $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action spaces be convex for each DM. Let (a) and (c) in Theorem 6 hold, and let observations be exchangeable conditioned on ω_0. Assume there exists a sequence $\{\gamma_N^*\}_N$ converges pointwise to γ_∞^* as $N \to \infty$, and let $\mathbb{P}$-a.s.

$$\left|\gamma_N^*(v^i) - \gamma_\infty^*(v^i)\right|^2 \leq \frac{f(v^i)h(N)}{N} \tag{34}$$

where $\lim_{N\to\infty} N^{-1}\sum_{i=1}^{N} f(v^i) < \infty$ and $\lim_{N\to\infty} h(N) = 0$. Then, a team optimal policy for $(\mathcal{P}_\infty)$ is symmetrically optimal and an optimal policy is identified as a limit of a sequence of team optimal policies for $(\mathcal{P}_N)$ [see (6)] as $N \to \infty$.

Proof: Following from the result of Theorem 7, one can consider a sequence of N-DM teams, which are symmetrically optimal that defines $(\mathcal{P}_N)$ [see (6)] and whose limit is identified with $(\mathcal{P}_\infty)$. Equivalent to (b) in Theorem 6, we can show that $\lim_{N\to\infty}\sup_{1\leq i\leq N}||\gamma_N^*(v^i) - \gamma_\infty^*(v^i)||^2 = 0$ $\mathbb{P}$-almost surely.

We have

$$\lim_{N\to\infty}\sup_{1\le i\le N}|\gamma_N^*(v^i)-\gamma_\infty^*(v^i)|^2$$
$$\le \lim_{N\to\infty}\sum_{i=1}^{N}|\gamma_N^*(v^i)-\gamma_\infty^*(v^i)|^2$$
$$\le \lim_{N\to\infty} h(N)\frac{1}{N}\sum_{i=1}^{N} f(v^i)=0$$

where the last inequality follows from (34). Hence, thanks to Theorem 6, a team optimal policy for $(\mathcal{P}_\infty)$ is the limit of a sequence of team optimal policies for $(\mathcal{P}_N)$ [see (6)] as $N\to\infty$, and hence a team optimal policy for $(\mathcal{P}_\infty)$ is symmetrically optimal and the proof is completed. ■

C. Existence Theorem on Globally Optimal Policies for Mean-Field Team Problems

An implication of our analysis is the following existence result on globally optimal policies for mean-field problems. In Theorem 8, we showed that if a pointwise limit as $N\to\infty$ of a sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)] exists, this limit is a globally optimal policy for $(\mathcal{P}_\infty)$, but under the conditions stated in the following theorem, an existence result also can be established. In the following, we relax the assumption that there exists a pointwise convergence sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)]. For the following theorem, we do not establish the pointwise convergence; but clearly if a sequence of optimal policies for $(\mathcal{P}_N)$ [see (6)] converges pointwise, a global optimal policy exists. Let $Q_N(B):=\frac{1}{N}\sum_{i=1}^{N}\delta_{\zeta_N^i}(B)$, where $B\in\mathcal{Z}:=\mathbb{U}\times\mathbb{V}$, and $\zeta_N^i:=(\gamma_N^*(v^i),v^i)$.

Theorem 12: Consider $(\mathcal{P}_\infty)$ [see (5)] with the convex cost function in the second and third arguments $\mathbb{P}$-almost surely. Let the action spaces be convex for each DM. Assume further that, without any loss, the optimal control laws can be restricted to those with $\mathbb{E}(\phi_i(u^i))\le K$ for some finite K, where $\phi_i:\mathbb{U}^i\to\mathbb{R}_+$ is lower semicontinuous. If v^is are i.i.d. random variables, then there exists an optimal policy for $(\mathcal{P}_\infty)$.

We note that the limit policy is not necessarily deterministic according to the above result; this interesting discussion is left open for further study.

Proof: We first show that $\{Q_N\}_N$ is precompact in the product space $(\mathbb{V}\times\mathbb{U})$ equipped with the weak convergence topology for each component. Then, we show that an induced policy by the limit Q achieves lower expected cost than $\limsup_{N\to\infty}J_N(\underline{\gamma}_N^*)$, and we invoke Theorem 5 to complete the proof. Action spaces and the cost function are convex and following from the hypothesis that v^is are i.i.d. random variables (hence they are exchangeable conditioned on ω_0) and the result of Theorem 7, one can consider a sequence of N-DM teams which are symmetrically optimal that defines $(\mathcal{P}_N)$ [see (6)] and whose limit is identified with $(\mathcal{P}_\infty)$.

Step 1: In the following, we show that for some subsequence $\{Q_n\}_{n\in\mathbb{I}}$ converges weakly to Q $\mathbb{P}$-almost surely, that is, $\mathbb{P}$-a.s., for every continuous and bounded function g

$$\lim_{n\to\infty}\left|\int g dQ_n-\int g dQ\right|=0$$

where $n\in\mathbb{I}$ is the index set of a converging subsequence. We use the fact that observations are i.i.d. and the space of control policies is weakly compact (see, e.g., [51, proof of Th. 4.7]). That is because we can represent the control policy spaces with the space of all joint measures on $(\mathbb{V}^i\times\mathbb{U}^i)$ for each DM with a fixed marginal on v^i [8], [53]. Since the team is static, this decouples the policy spaces from the policies of the previous DMs, and following from the hypothesis on ϕ_i and the fact that $\nu\to\int\nu(dx)g(x)$ is lower semicontinuous for a continuous function g [51, Proof of Th. 4.7], the marginals on $\mathbb{U}^i$ will be weakly compact. If the marginals are weakly compact, then the collection of all measures with these weakly compact marginals are also weakly compact (see, e.g., [50, Proof of Th. 2.4]) and hence the control policy space is weakly compact. Using Tychonoff's theorem, the countably infinite product space is also compact under the product topology, which implies compactness of the space of control policies under the product topology. Hence, there exists a subsequence $\{Q_n\}_{n\in\mathbb{I}}$ converges weakly to Q $\mathbb{P}$-almost surely.

Step 2: Now, we show that (14) holds. We have

$$\mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}}c\left(\omega_0,u,\int_{\mathbb{U}}uQ(du\times\mathbb{V})\right)Q(du,dv)\middle|\omega_0\right]\right]$$
$$=\lim_{M\to\infty}\mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}}\min\left\{M,c\left(\omega_0,u,\int_{\mathbb{U}}uQ(du\times\mathbb{V})\right)\right\}\times Q(du,dv)\middle|\omega_0\right]\right] \tag{35}$$
$$=\lim_{M\to\infty}\mathbb{E}\left[\mathbb{E}\left[\lim_{n\to\infty}\int_{\mathcal{Z}}\min\left\{M,c\left(\omega_0,u,\int_{\mathbb{U}}uQ_n(du\times\mathbb{V})\right)\right\}\times Q_n(du,dv)\middle|\omega_0\right]\right] \tag{36}$$
$$=\lim_{M\to\infty}\lim_{n\to\infty}\mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}}\min\left\{M,c\left(\omega_0,u,\int_{\mathbb{U}}uQ_n(du\times\mathbb{V})\right)\right\}\times Q_n(du,dv)\middle|\omega_0\right]\right] \tag{37}$$
$$\le\lim_{M\to\infty}\limsup_{N\to\infty}\mathbb{E}\left[\mathbb{E}\left[\int_{\mathcal{Z}}\min\left\{M,c\left(\omega_0,u,\int_{\mathbb{U}}uQ_N(du\times\mathbb{V})\right)\right\}\times Q_N(du,dv)\middle|\omega_0\right]\right] \tag{38}$$
$$\le\limsup_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[c\left(\omega_0,\gamma_N^*(v^i),\frac{1}{N}\sum_{i=1}^{N}\gamma_N^*(v^i)\right)\right] \tag{39}$$

where (35) follows from the monotone convergence theorem. Since $\{Q_n\}_{n\in I}$ converges weakly to Q $\mathbb{P}$-almost surely, we have by continuous mapping theorem (by considering a projection to the first component) $\int_{\mathbb{U}}uQ_n(du\times\mathbb{V})\to\int_{\mathbb{U}}uQ(du\times\mathbb{V})$ $\mathbb{P}$-almost surely. Following from (Step 1), (36) follows from [46, Th. 3.5]. That is because the cost function is continuous in actions, and $\min\{M,c(\omega_0,u,\int_{\mathbb{U}}uQ_n(du\times\mathbb{V}))\}$ continuously converges in u, $\min\{M,c(\omega_0,u_n,\int_{\mathbb{U}}uQ_n(du\times\mathbb{V}))\}\to\min\{M,c(\omega_0,u,\int_{\mathbb{U}}uQ(du\times\mathbb{V}))\}$ where $u_n\to u$ as $n\to\infty$. Equality (37) follows from the dominated convergence theorem since $\min\{M,c(\omega_0,u,\int_{\mathbb{U}}uQ_N(du\times\mathbb{V}))\}$ is bounded, and (38) is true since limsup is the greatest convergent subsequence limit for a bounded sequence. Finally, (39) follows from the definition of empirical measures and since for every M, $\min\{M,c(\omega_0,u,\int_{\mathbb{U}}uQ_N(du\times\mathbb{V}))\}\le c(\omega_0,u,\int_{\mathbb{U}}uQ_N(du\times\mathbb{V}))$; hence, following from Theorem 5, the randomized limit policy through subsequence is a globally optimal for $(\mathcal{P}_\infty)$. ■

We apply the results of this section in Section V-C.

V. Examples

In the following, we present a number of examples to demonstrate results in previous sections. First, we consider LQG and LQ static teams with coupling between states, then we consider LQG symmetric static teams with coupling between control actions. Moreover, we investigate dynamic infinite-horizon average cost LQG teams with the classical information structure.

A. Example 1, Static Quadratic Gaussian Teams With Coupling Between States

Consider the following observation scheme:

$$v^i = x^i + z^i \tag{40}$$

where $\{z^i\}_{i\in\mathbb{N}}$ and $\{x^i\}_{i\in\mathbb{N}}$ are i.i.d. zero mean Gaussian random variables. Let $\{z^i\}_{i\in\mathbb{N}}$ be independent of $\{x^i\}_{i\in\mathbb{N}}$. The expected cost function is defined as

$$J(\underline{\gamma}) = \limsup_{N\to\infty} \frac{1}{N} \mathbb{E}^{\underline{\gamma}} \left[\sum_{i=1}^{N} R(u^i)^2 + Q(u^i - x^i - \mu_N)^2 \right] \tag{41}$$

where $\mu_N := \frac{1}{N}\sum_{k=1}^{N} x^k$. Let R be a positive number and Q be a non-negative number.

Theorem 13: For LQG static teams as formulated above, under the measurement scheme (40), $\gamma_\infty^{i*}(v^i)$ is globally optimal for $(\mathcal{P}_\infty)$ achieved as the limit $N\to\infty$ of $\gamma_N^{i*}(v^i)$, an optimal solution for $(\mathcal{P}_N)$.

Proof: We invoke Theorem 6 to prove the theorem. The stationary policy (see Definition 3) is obtained as

$$\gamma_N^{i*} = (R+Q)^{-1} Q \left(1 + \frac{1}{N}\right) \mathbb{E}(x^i|v^i)$$

where the equality follows from the assumption that x^is are independent of z^is and x^ks, $k\neq i$ for every $i=1,2,\ldots,N$ and the assumption that random variables are mean zero. Following from [30], stationary policies are team optimal for $(\mathcal{P}_N)$ in this formulation. We have $\gamma_\infty^{i*}(v^i) = (R+Q)^{-1}Q\mathbb{E}(x^i|v^i)$. Since v^is are zero mean Gaussian random variables, we have $\mathbb{E}(x^i|v^i) = \Sigma_{x^iv^i}\Sigma^{-1}_{v^iv^i}v^i := Kv^i$, where Σ_{XY} is defined as a covariance of two random variables X and Y. We have $\mathbb{P}$-a.s.

$$\begin{aligned}\sup_{1\le i\le N} |\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)| &= \frac{Q}{R+Q} \sup_{1\le i\le N} \left|\frac{1}{N}\mathbb{E}(x^i|v^i)\right| \\ &= \frac{KQ}{R+Q} \sup_{1\le i\le N} \left|\frac{1}{N} v^i\right| \xrightarrow[N\to\infty]{} 0\end{aligned} \tag{42}$$

where (42) follows from

$$\lim_{N\to\infty} \sup_{1\le i\le N} \frac{1}{N^2}(v^i)^2 \le \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} (v^i)^2 = 0 \mathbb{P} - a.s.$$

where the first inequality is true since $(v^i)^2$s are non-negative, and equality follows from the SLLN since v^is are i.i.d. and have a finite variance, hence, (b) holds. One can show that the condition in Remark 2(ii) holds since v^is and x^is are i.i.d. random variables, hence Theorem 6 completes the proof. ■

B. Example 2, Static Non-Gaussian Teams With Coupling Between States

Let the observation scheme be (40), where $\{z^i\}_{i\in\mathbb{N}}$ and $\{x^i\}_{i\in\mathbb{N}}$ are i.i.d. zero mean random variables with finite variance. Let $\{z^i\}_{i\in\mathbb{N}}$ be independent of $\{x^i\}_{i\in\mathbb{N}}$. The expected cost function is defined as (41). Let R be a positive number and Q be a non-negative number.

Theorem 14: For LQ static teams as formulated above, under the measurement scheme (40), $\gamma_\infty^{k*}(v^k) = (R+Q)^{-1}Q\mathbb{E}(x^k|v^k)$ is globally optimal for $(\mathcal{P}_\infty)$ and is obtained as the limit of $\gamma_N^{k*}(v^k)$ as $N\to\infty$.

Proof: In the following, we use both Theorem 4 and Theorem 6. Clearly, (A1) holds, we show that (A2) holds

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N} (\gamma_\infty^{i*}(v^i))^2 R + Q(\gamma_\infty^{i*}(v^i) - x^i - \mu_N)^2\right]$$

$$= \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N} \frac{-Q^2}{Q+R}\mathbb{E}^2(x^i|v^i)(1+\frac{2}{N})(x^i+\mu_N)^2 Q\right] \tag{43}$$

$$\le \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N} \frac{-Q^2}{Q+R}\mathbb{E}^2(x^i|v^i)\right] + \lim_{N\to\infty} \frac{Q(N+3)\sigma^2}{N} \tag{44}$$

$$= \frac{-Q^2}{Q+R}\mathbb{E}\left[\mathbb{E}^2(x^1|v^1)\right] + Q\sigma^2 \tag{45}$$

where (43) follows from $\mathbb{E}(\mathbb{E}(x^i|v^i)(x^i+\mu_N)) = \mathbb{E}(\mathbb{E}(\mathbb{E}(x^i|v^i)(x^i+\mu_N)|v^i)) = (1+\frac{1}{N})\mathbb{E}(\mathbb{E}^2(x^i|v^i))$, and (44) is true since x^i and z^i are i.i.d. random variables and $\limsup_{N\to\infty} a_N + \limsup_{N\to\infty} b_N \ge \limsup_{N\to\infty}(a_N+b_N)$. We can justify (45) by defining $Y^i := (\mathbb{E}(x^i|v^i))^2$, and since Y^is are measurable functions of $\{v^i\}_{i\ge1}$, and v^is and x^is are i.i.d., Y^is are i.i.d. random variables. Similarly, one can show the other side direction for liminf. Hence, (A2) is satisfied. Now, we check (8), for every γ_∞^k with $J(\underline{\gamma}_\infty) < \infty$

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\left(\sum_{i=1}^{N}\sum_{k=1}^{N} c_{u^k}(\omega_0, \gamma^{i*}, \mu^*)(m_k)\right)$$

$$= \limsup_{N\to\infty} \frac{2Q}{N}\sum_{k=1}^{N}\mathbb{E}\left(\mathbb{E}(x^k(m_k)|v^k)\right) - \mathbb{E}\left((x^k+\mu_N)(m_k)\right) \tag{46}$$

$$= \limsup_{N\to\infty} \frac{-2Q}{N}\sum_{k=1}^{N}\mathbb{E}\left(\mu_N\left(m_k\right)\right) \tag{47}$$

$$= 2Q \limsup_{N\to\infty} \frac{1}{N^2}\sum_{k=1}^{N}\mathbb{E}\left[x^k\gamma_\infty^{k*}(v^k)\right] - \mathbb{E}\left[x^k\gamma_\infty^{k}(v^k)\right] \tag{48}$$

$$= -2Q \liminf_{N\to\infty} \frac{1}{N^2}\sum_{k=1}^{N}\mathbb{E}\left(x^k\gamma_\infty^{k}(v^k)\right) \tag{49}$$

$$\ge -2Q\sigma \liminf_{N\to\infty} \frac{1}{N^2}\sum_{k=1}^{N}\sqrt{\mathbb{E}\left[(\gamma_\infty^{k}(v^k))^2\right]} \tag{50}$$

$$\geq -2Q\sigma \liminf_{N\to\infty} \sup_{1\leq k\leq N} \sqrt{\frac{\mathbb{E}\left[(\gamma_\infty^k(v^k))^2\right]}{N^2}} = 0 \tag{51}$$

where measurability of $m_k := \gamma_\infty^k(v^k) - \gamma_\infty^{k*}(v^k)$ with respect to the σ-field generated by v^k implies (46), and (47) follows from the iterated expectations property. Since x^ps are mean zero and independent of v^k for $k \neq p$, we have (48), and (49) follows from the fact that γ_∞^{k*} is independent of k, and since v^k and x^k are i.i.d. random variables. Moreover, $J(\underline{\gamma}_\infty) < \infty$, so $\mathbb{E}(\gamma_\infty^k(v^k)) \leq \mathbb{E}((\gamma_\infty^k(v^k))^2) < \infty$, and Cauchy–Schwarz inequality implies (50), and (51) follows from

$$\liminf_{N\to\infty} \sup_{1\leq k\leq N} \frac{\mathbb{E}\left[(\gamma_\infty^k(v^k))^2\right]}{N^2}$$

$$\leq \liminf_{N\to\infty} \frac{1}{N^2}\sum_{k=1}^{N} \mathbb{E}\left[(\gamma_\infty^k(v^k))^2\right] = 0 \tag{52}$$

where (52) is true since $\mathbb{E}[(\gamma_\infty^k(v^k))^2] \geq 0$ and $\limsup_{N\to\infty} \frac{1}{N}\mathbb{E}(\sum_{k=1}^{N}(\gamma_\infty^k(v^k))^2 R) \leq J(\underline{\gamma}_\infty) < \infty$. Thus, (8) is satisfied and Theorem 4 completes the proof.

One can also invoke Theorem 6 to complete the proof. One can show that the condition in Remark 2(ii) holds since v^is and x^is are i.i.d. random variables. We only justify (b). Stationary policy is team optimal for $(\mathcal{P}_N)$ in this formulation [30], hence $\gamma_N^{i*}(v^i) = (R+Q)^{-1}Q(1+\frac{1}{N})\mathbb{E}(x^i|v^i)$, so we need to show that

$$\lim_{N\to\infty} \sup_{1\leq i\leq N} \left|\gamma_N^{i*}(v^i) - \gamma_\infty^{i*}(v^i)\right| = 0 \quad \mathbb{P}-a.s.$$

Equivalently, we can show that $\mathbb{P}$-a.s.

$$\lim_{N\to\infty} \sup_{1\leq i\leq N} \frac{1}{N^2}\left(\mathbb{E}(x^i|v^i)\right)^2 \leq \lim_{N\to\infty} \frac{1}{N^2}\sum_{i=1}^{N}\left(\mathbb{E}(x^i|v^i)\right)^2 = 0$$

where the first inequality is true since $(\mathbb{E}(x^i|v^i))^2$s are non-negative, and equality follows from SLLN since

$$\mathbb{E}\left((\mathbb{E}(x^i|v^i))^2\right) = \mathbb{E}\left((x^i)^2\right) - \mathbb{E}\left((x^i - \mathbb{E}(x^i|v^i))^2\right) < \infty$$

and $(\mathbb{E}(x^i|v^i))^2$ are i.i.d. sequence of random variables since v^is are i.i.d. random variables and the proof is completed. ■

C. Example 3, LQG Symmetric Teams With Coupling Between Control Actions

Let

$$\underline{v}^i = H^i \underline{x} + \underline{z}^i \tag{53}$$

where $\{\underline{z}^i\}_{i\in\mathbb{N}}$ is independent zero mean Gaussian random vectors also independent of $\underline{x}$, with covariance $\Sigma_{jj} = N^0 > 0$. Define $\omega = (\underline{x}, \underline{z}^1, \underline{z}^2, \ldots)$, and $\omega_0 := \underline{x}$ where $\underline{x}$ is a Gaussian random vector with covariance $\mathbb{E}(\underline{x}\underline{x}^T) = \Sigma_{00}$. Let

$$J(\underline{\gamma}) = \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}^{\underline{\gamma}}\Bigg[\sum_{i=1}^{N}(u^i)^T R u^i - 2\sum_{i=1}^{N}(u^i)^T D$$

$$\times\left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right) + \left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right)^T Q\left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right)\Bigg] \tag{54}$$

where R is an appropriate dimension positive definite matrix and D, and Q are appropriate dimension positive semidefinite matrices, and $R > 2D$. In the following, we follow steps in [52, Th. 2.6.8] to obtain optimal policies for $(\mathcal{P}_N)$.

Lemma 3: Consider an N-DM LQG team formulated above, under the measurement scheme (53), the global optimal policy for $(\mathcal{P}_N)$ is linear, i.e., $\gamma_N^{k*}(v^k) = \pi_N^k v^k$. Here, $\pi_N^k \in \mathcal{M}_{n,m}(\mathbb{R})$, $n\times m$ real-valued matrix, is obtained by solving the following parallel update scheme:

$$\pi_{N,(i)}^k = -L_N\left[S^k + \frac{1}{N}\sum_{p=1,p\neq k}^{N}\pi_{N,(i)}^p H^p S^k\right] \tag{55}$$

where $L_N := (R + \frac{Q}{N^2} - \frac{2D}{N})^{-1}(\frac{Q}{N} - D)$eq, $S^k := \Sigma_{00}(H^k)^T(H^k\Sigma_{00}(H^k)^T + \Sigma_{kk})^{-1}$ and the initial points of the iterations are considered as zero functions.

Proof: By Definition 3, stationary policies satisfy the following equality for $k = 1,\ldots,N$

$$M\gamma_N^{k*}(v^k) + \left(\frac{Q}{N} - D\right)$$

$$\times\left[\mathbb{E}(\underline{x}|v^k) + \frac{1}{N}\sum_{p=1,p\neq k}^{N}\mathbb{E}\left(\gamma_N^{p*}(v^p)|v^k\right)\right] = 0 \tag{56}$$

where $M := R + \frac{Q}{N^2} - \frac{2D}{N}$, and (56) can be rewritten as $P\underline{\hat{R}}\underline{\gamma}_N^*(\underline{v}) + Pr(\omega) = 0$, where P is a block diagonal matrix with iith block $P_{ii}\beta^i(\omega) := \mathbb{E}(\beta^i(\omega)|v^i)$, $\underline{\hat{R}}$ is a matrix where $\hat{R}_{ii} := M$ and $\hat{R}_{ij} := \dfrac{1}{N}(\dfrac{Q}{N} - D)$ for every $i,j = 1,\ldots,N$, $j\neq i$, and $r(\omega) = \underline{x}$. Note that P is a projection operator defined on a Hilbert space whose operator norm is one. Now, we use the successive approximation method [52, Th. A.6.4]. According to (56), we can write for $k = 1,2,\ldots,N$

$$M\gamma_{N,(i)}^{k*}(v^k) + \epsilon\gamma_{N,(i)}^{k*}(v^k) - \epsilon\gamma_{N,(i)}^{k*}(v^k) + \left(\frac{Q}{N} - D\right)$$

$$\times\left[\mathbb{E}(\underline{x}|v^k) + \frac{1}{N}\sum_{p=1,p\neq k}^{N}\mathbb{E}\left(\gamma_{N,(i)}^{p*}(v^p)|v^k\right)\right] = 0.$$

Thus, by dividing the expression over ϵ and rearranging it, we have

$$\gamma_{N,(i)}^{k*}(v^k) = \left(1 - \frac{\hat{R}_{ii}}{\epsilon}\right)\gamma_{N,(i)}^{k*}(v^k) - \frac{1}{\epsilon}\left(\frac{Q}{N} - D\right)$$

$$\times\left[\mathbb{E}(x|v^k) + \frac{1}{N}\sum_{p=1,p\neq k}^{N}\mathbb{E}\left(\gamma_{N,(i)}^{p*}(v^p)\middle|v^k\right)\right]$$

where the initial points of the iterations are zero functions. We can write $\underline{\gamma}_N^*(\underline{v}) = P(I - \frac{1}{\epsilon}\underline{\hat{R}})\underline{\gamma}_N^*(\underline{v}) - \frac{1}{\epsilon}Pr(\omega)$. Similar to [52, Th. 2.6.5], the above sequence converges to the unique fixed point if and only if the spectral radius satisfies the following constraint $\rho(P(I - \frac{\hat{R}}{\epsilon})) = \rho(I - \frac{\hat{R}}{\epsilon}) := \lim_{k\to\infty}\sup[||A||^k]^{\frac{1}{k}} < 1$, where $A := I - \dfrac{\hat{R}}{\epsilon}$, $||A|| := \sup_{||x||<1}||Ax||$ and ρ denotes spectral radius. The first equality is true since both P and A maps $\mathbf{\Gamma}_N$ into itself and P has operator norm equal to one. The above constraint can be always satisfied by choosing $\epsilon = \frac{1}{2}(\lambda_{max}(\underline{\hat{R}}) + \lambda_{min}(\underline{\hat{R}}))$. On the other hand, since $(\underline{x}, \underline{z}^1, \ldots, \underline{z}^N)$ are jointly Gaussian, then $\gamma_N^{k*}(v^k) = \pi_N^k v^k$ for $k = 1, \dots, N$. Hence, $\gamma_{N,(i)}^{k*}(v^k) = \pi_{N,(i)}^k v^k$, and by linearity of the conditional expectation, we have $\mathbb{E}(\underline{x}|v^k) = S^k v^k$, and $\mathbb{E}(\gamma_N^{p*}(v^p)|v^k) = \pi_N^p H^p S^k v^k$. Hence, (55) holds. Following from [52], the stationary policy is globally optimal for $(\mathcal{P}_N)$, and this completes the proof. ∎

Theorem 15: Consider $(\mathcal{P}_\infty)$ with the expected cost (54). Under the following measurement scheme:

$$\underline{v}^i = H\underline{x} + \underline{z}^i \tag{57}$$

where $\underline{z}^i$s are i.i.d. Gaussian random vectors, $\gamma_\infty^{i*}(v^i) = \pi_\infty^* v^i$ is an optimal policy for $(\mathcal{P}_\infty)$ and is the pointwise limit of $\gamma_N^{i*}(v^i) = \pi_N^* v^i$, an optimal policy for $(\mathcal{P}_N)$.

Proof: In the following, we invoke Proposition 1 and Theorem 9 to prove the theorem. Under (57), the static team is symmetrically optimal and hence from (55), we have $\pi_N^* = L_N[S + N^{-1}(N-1)\pi_N^* HS]$, $\pi_\infty^* = R^{-1}D[S + \pi_\infty^* HS]$, where $L_N := (N^2R - 2DN + Q)^{-1}(N^2D - NQ)$, $S := \Sigma_{00}(H)^T(H\Sigma_{00}(H)^T + \Sigma_{kk})^{-1}$. Since for every N, we have $J_N(\underline{\gamma}_N^*) < \infty$, and since $R > 0$, we have $\sup_{N\geq 1}\mathbb{E}(||\gamma_N^*(v^1)||_2^2) < \infty$, which implies (A3). The proof is completed using the results of Proposition 1 and Theorem 9. One can also invoke Theorem 11 to justify the result. ∎

1) Example 4, Asymmetric LQG Team Problems: Here, we consider simple variation of Example 3 considered above to illustrate Remark 4. Consider the observation scheme (57), and let the expected cost function be defined as

$$\begin{aligned} J(\underline{\gamma}) = \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}^{\underline{\gamma}}\Bigg[&\sum_{i=1}^{N}(u^i)^T R u^i \\ &- 2\sum_{i=1}^{N}(u^i)^T D\left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right) \\ &+ \left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right)^T Q\left(\underline{x} + \frac{1}{N}\sum_{k=1}^{N}u^k\right) \\ &+ \frac{1}{N}\sum_{k=1}^{M}(u^k)^T\alpha_k u^k\Bigg] \end{aligned}$$

where $M \in \mathbb{Z}_+$ is independent of N. Clearly, the N-DM team admits asymmetric optimal policies for $(\mathcal{P}_N)$ with the expected cost J_N for every N. However, one can observe that the last term above goes to zero as $N \to \infty$ under a sequence of optimal policies, and hence asymptotically the expected cost would essentially be (54) and Theorem 13 implies γ_∞^* is an optimal policy since $\mathbb{P}$-almost surely the sequence Q_N converges weakly (the asymmetric term vanishes when $N \to \infty$). That is, the optimal policy designed for the symmetric problem is also a solution for the asymmetric problem since under this policy, the additional term (which is a non-negative contribution) vanishes, certifying its optimality.

D. Example 5, Multivariable Classical Linear Quadratic Gaussian Problems: Average Cost Optimality Through Static Reduction

Here, we revisit a well-known problem and a well-known solution, using the technique presented in this article. Let

$$X_{t+1} = AX_t + Bu^t + w^t$$

where $A \in \mathcal{M}_{n,n}(\mathbb{R})$, $B \in \mathcal{M}_{n,m}(\mathbb{R})$, and w^ts and X_0 are i.i.d. Gaussian random vectors with mean zero and positive variance taking values in $\mathbb{R}^n$. Let (A, B) be controllable and let

$$\begin{aligned} J(\underline{\gamma}) &= \limsup_{T\to\infty} J_T(\underline{\gamma}) \\ &:= \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}^{\underline{\gamma}}\left[\sum_{t=0}^{T-1}X_t^TQX_t + (u^t)^TRu^t\right] \end{aligned}$$

where $Q \geq 0$ and $R > 0$ are appropriate dimensions real matrices. We can write

$$\begin{aligned} J(\underline{\gamma}) = \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}^{\underline{\gamma}}\Bigg[&\sum_{t=0}^{T-1}\left(\sum_{k=1}^{t}A^{t-k}Bu^{k-1} + \sum_{k=0}^{t}A^{t-k}\zeta^k\right)^T \\ &\times Q\left(\sum_{k=1}^{t}A^{t-k}Bu^{k-1} + \sum_{k=0}^{t}A^{t-k}\zeta^k\right) + (u^t)^TRu^t\Bigg] \end{aligned}$$

where $\zeta = (X_0^T, (w^0)^T, (w^1)^T, \dots)^T$. In the following, we consider fully observed classical IS, i.e., $Y^t = X_t$, and we can write $Y^t = H^t\zeta + \sum_{j=0}^{t-1}D_{tj}u^j$, where H^t and D_{tj} are appropriate dimensional matrices. Using [23, Th. 1], we can reduce IS to the static one as $V^t = \tilde{H}_t\zeta$. According to [22, Sec. 3.5], we have $u_T^{t*} = G_T^tX_t$ for $t = 0, 1, \dots, T-1$, where $k_T^T = 0$, and

$$G_T^t = -(R + B^Tk_T^{t+1}B)^{-1}B^Tk_T^{t+1}A \tag{58}$$

$$k_T^t = Q + A^Tk_T^{t+1}A - A^Tk_T^{t+1}B(R + B^Tk_T^{t+1}B)^{-1}B^Tk_T^{t+1}A. \tag{59}$$

Theorem 16: For LQG teams with the classical IS, $u^{t*} = \lim_{T\to\infty}\gamma_T^{t*}(v^t) = \gamma_\infty^{t*}(v^t)$ is the optimal policy for $J(\underline{\gamma})$, where $\{\gamma_T^{t*}\}_T$ is a sequence of optimal policies for $\{J_T(\underline{\gamma}_T)\}_T$ with the pointwise limit γ_∞^{t*} as $T \to \infty$.

Although this result is a classical one in the literature, here, we present a new approach using the static reduction.

Proof: Since, $k_{T+1}^t = k_T^{t-1}$ for $t = 1, 2, \dots, T$, one can write (59) as

$$k_T^t = Q + A^Tk_{T-1}^tA - A^Tk_{T-1}^tB(R + B^Tk_{T-1}^tB)^{-1}B^Tk_{T-1}^tA.$$

We use Theorem 5 and Remark 1, to show that $u_\infty^{t*} = G_\infty X_t$ is team optimal, where $G_\infty = -(R + B^TC^*B)^{-1}B^TC^*A$, and following from controllability of (A, B), $C^* = \lim_{\beta\to 1}C_\beta$, a fixed point of the following recursion exists, $C_\beta(n) = Q + A^T\beta C_\beta(n-1)A - A^T\beta C_\beta(n-1)B(R + B^T\beta C_\beta(n-1)B)^{-1}B^T\beta C_\beta(n-1)A$. By comparing $C^*(n) = \lim_{\beta\to 1}C_\beta^*(n)$ and (59), we have $\lim_{T\to\infty}k_T^t = K = C^* = \lim_{n\to\infty}C^*(n)$. Hence, for $t = 0, 1, \dots, T-1$, $\lim_{T\to\infty}G_T^t = -(R + B^TKB)^{-1}B^TKA = -(R + B^TC^*B)^{-1}B^TC^*A = G_\infty$. Now, we use Remark 1 to show (14) holds

$$\begin{aligned} &\limsup_{T\to\infty}\left|J_T(\underline{\gamma}_T^*) - J_T(\underline{\gamma}_\infty^*)\right| \\ &\leq \limsup_{T\to\infty}\sup_{0\leq t\leq T-1}\Bigg|\mathbb{E}\left[\sum_{k=0}^{t}Tr\left(\zeta_k^T\left(L_T^{t,k}\right)^T(H_T^t)L_T^{t,k}\zeta_k\right)\right] \\ &- \mathbb{E}\left[\sum_{k=0}^{t}Tr\left(\zeta_k^T\left(L_\infty^{t,k}\right)^T(H_\infty^t)L_\infty^{t,k}\zeta_k\right)\right]\Bigg| \end{aligned} \tag{60}$$

$$\leq \limsup_{T\to\infty} \sup_{0\leq t\leq T-1} \left|\mathbb{E}\left[\sum_{k=0}^{t} Tr\left(\zeta_k\zeta_k^T\left(\left(L_T^{t,k}\right)^T (H_T^t)L_T^{t,k} - \left(L_\infty^{t,k}\right)^T (H_\infty^t)L_\infty^{t,k}\right)\right)\right]\right| \tag{61}$$

$$\leq \Sigma^2 \limsup_{T\to\infty} \sup_{0\leq t\leq T-1} \left|Tr\left((H_T^t)C_T^t - (H_\infty^t)C_\infty^t\right)\right| \tag{62}$$

$$\leq \Sigma^2 \limsup_{T\to\infty} \left[\sup_{0\leq t\leq T-1} \left|Tr\left((G_T^t)^T RG_T^t C_T^t - (G_\infty)^T RG_\infty C_\infty^t\right)\right| + \sup_{0\leq t\leq T-1}\left|Tr(Qe_T^t)\right|\right] \tag{63}$$

$$\leq \Sigma^2 \limsup_{T\to\infty} \left[\sup_{0\leq t\leq T-1} \left|Tr\left((G_T^t)^T RG_T^t - (G_\infty)^T RG_\infty\right)\right| \times \sup_{0\leq t\leq T-1}\left|Tr(C_T^t)\right| + \sup_{0\leq t\leq T-1}\left|Tr\left(G_\infty^T RG_\infty e_T^t\right)\right| + \sup_{0\leq t\leq T-1}\left|Tr(Qe_T^t)\right|\right]$$

$$\leq \Sigma^2 \limsup_{T\to\infty} \left[\sup_{0\leq t\leq T-1} \left|Tr\left((G_T^t(G_T^t)^T - G_\infty G_\infty^T)R\right)\right| \times \left(\sup_{0\leq t\leq T-1}\left|Tr(e_T^t)\right| + \sup_{0\leq t\leq T-1}\left|Tr(C_\infty^t)\right|\right) + \sup_{0\leq t\leq T-1}\left|Tr\left(G_\infty^T RG_\infty e_T^t\right)\right| + \sup_{0\leq t\leq T-1}\left|Tr(Qe_T^t)\right|\right] = 0 \tag{64}$$

where $L_T^{t,k} := \prod_{p=k}^{t-1}(A+BG_T^p)$, $L_\infty^{t,k} := \prod_{p=k}^{t-1}(A+BG_\infty)$, $H_T^t = (Q+(G_T^t)^T RG_T^t)$, $H_\infty^t = (Q+(G_\infty)^T RG_\infty)$, $e_T^t := C_T^t - C_\infty^t$, and $C_T^t := [\sum_{k=0}^t L_T^{t,k}(L_T^{t,k})^T]$, $C_\infty^t := [\sum_{k=0}^t L_\infty^{t,k}(L_\infty^{t,k})^T]$ and $\Sigma^2 := \max(\sigma_{X_0}^2, \sigma_w^2)$, where $\sigma_{X_0}^2$ and σ_w^2 are the variance of each component of X_0 and w^k, respectively. Equality (60) follows from the fact that $\{w^k\}_k$ are i.i.d. and independent from X_0. Inequality (61) follows from the trace property that $Tr(ABC) = Tr(CAB)$, and (62) follows from the hypothesis that ζ_ks are i.i.d. random vectors and $Tr(ABC) = Tr(BCA)$ and (63) follows from linearity of the trace and $\sup f + g \leq \sup f + \sup g$. Inequality (64) follows from adding and subtracting $G_\infty^T RG_\infty C_T^t$ in the first term and using $Tr(AB) \leq Tr(A)Tr(B)$ for A and B positive semidefinite matrices since (59) implies that for a fixed T, $\{k_T^t\}_{t=0}^{T-1}$ is a decreasing sequence, i.e., $K > k_T^0 > k_T^1 > \ldots > k_T^{T-1}$, and hence $\{G_T^t(G_T^t)^T\}_{t=0}^{T-1}$ is a decreasing sequence. Also, from (58), we have for a fixed T, $\{(A+BG_T^t)(A+BG_T^t)^T\}_{t=0}^{T-1}$ is an increasing sequence, hence, $(G_T^t)^T RG_T^t - G_\infty^T RG_\infty$ is positive semidefinite. Finally, the last inequality follows from the definition of e_T^t and the following calculation. First note that for a fixed T, $\{Tr(e_T^t)\}_{t=0}^{T-1}$ is an increasing sequence. Hence

$$\lim_{T\to\infty} \sup_{0\leq t\leq T-1} |Tr(e_T^t)| = \lim_{T\to\infty} |Tr(e_T^{T-1})| = 0.$$

Similarly, $\lim_{T\to\infty}\sup_{0\leq t\leq T-1}|Tr(Qe_T^t)| = 0$. We have

$$\lim_{T\to\infty} \sup_{0\leq t\leq T-1} |Tr(C_\infty^t)| = \left|Tr\left[(I-(A+BG_\infty))^{-1}\right]\right| = 0$$

where $Y^{(T)}$ denotes the T power of the matrix Y and the result follows from $||A+BG_\infty|| < 1$ (following from the controllability assumption). Finally, we have

$$\lim_{T\to\infty} \sup_{0\leq t\leq T-1} \left|Tr\left[\left(G_T^t(G_T^t)^T - G_\infty G_\infty^T\right)R\right]\right| = 0$$

where the second equality follows from the aforementioned observations and since R is positive definite. Therefore, $\limsup_{T\to\infty}|J_T(\underline{\gamma}_T^*) - J(\underline{\gamma}_\infty^*)| = 0$, and the proof is completed.

Remark 6: Similarly, one can show the result for (i) $Y^t = CX^t$, (A,C) is observable and $Q = C^TC$, (ii) the discounted LQG team problems with the classical IS.

VI. Conclusion

In this article, we studied static teams with countably infinite number of DMs. We presented sufficient conditions for team optimality concerning average cost problems. Additionally, constructive results have been established to obtain the team optimal solution for static teams with countably infinite number of DMs as limits of the optimal solutions for static teams with finite number of DMs as the number of DMs goes to infinity. We also studied sufficient conditions for team optimality of symmetric static teams and mean-field teams under relaxed conditions. We recently studied convex dynamic teams with countably infinite DMs and mean-field teams [44].

Acknowledgment

The authors are grateful to Prof. N. Saldi for the detailed suggestions he has provided, and to Dr. A. Gattami, Prof. N. Martins, and Prof. A. Mahajan for related discussions on this topic.

References

[1] J. Arabneydi and A. Aghdam, "A certainty equivalence result in team-optimal control of mean-field coupled Markov chains," in *Proc. IEEE 56th Annu. Conf. Decision Control*, 2017, pp. 3125–3130.

[2] J. Arabneydi and A. Mahajan, "Team-optimal solution of finite number of mean-field coupled LQG subsystems," in *Proc. IEEE 54th Annu. Conf. Decision Control*, 2015, pp. 5308–5313.

[3] A. Arapostathis, A. Biswas, and J. Carroll, "On solutions of mean field games with ergodic cost," *J. de Mathématiques Pures et Appliquées*, vol.107, no. 2, pp. 205–251, 2017.

[4] M. Bardi and M. Fischer, "On non-uniqueness and uniqueness of solutions in finite-horizon mean field games," *ESAIM: Control, Optimis. Calculus Variations*, vol. 25, 2019, Art. no. 44.

[5] M. Bardi and F. S. Priuli, "Linear-quadratic N-person and mean-field games with ergodic cost," *SIAM J. Control Optim.*, vol. 52, no. 5, pp. 3022–3052, 2014.

[6] P. Billingsley, *Convergence of Probability Measures*. New York, NY, USA: Wiley, 1968.

[7] A. Biswas, "Mean field games with ergodic cost for discrete time Markov processes," 2015, *arXiv:1510.08968*.

[8] V. S. Borkar, "White-noise representations in stochastic realization theory," *SIAM J. Control Optim.*, vol. 31, pp. 1093–1102, 1993.

[9] R. Buckdahn, J. Li, and S. Peng, "Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents," *SIAM J. Control Optim.*, vol. 52, no. 1, pp. 451–492, 2014.

[10] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I-II*. New York, NY, USA: Springer, 2018.
[11] R. Carmona, F. Delarue, and D. Lacker, "Mean field games with common noise," *Ann. Probability*, vol. 44, no. 6, pp. 3740–3803, 2016.
[12] A. Cecchin and M. Fischer, "Probabilistic approach to finite state mean field games," *Appl. Math. Optim.*, vol. 81, pp. 253–300, 2018.
[13] P. Dasgupta and E. Maskin, "The existence of equilibrium in discontinuous economic games, I: Theory," *Rev. Econ. Stud.*, vol. 53, no. 1, pp. 1–26, 1986.
[14] R. M. Dudley, *Real Analysis and Probability*. 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.
[15] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, "Fatou's lemma for weakly converging probabilities," *Theory Probability Appl.* vol. 58, no. 4, pp. 683–689, 2014.
[16] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, "Partially observable total-cost Markov decision processes with weakly continuous transition probabilities," *Math. Oper. Res.*, vol. 41, no. 2, pp. 656–681, 2016.
[17] E. Feleqi, "The derivation of ergodic mean field game equations for several populations of players," *Dyn. Games Appl.*, vol. 3, no. 4, pp. 523–536, 2013.
[18] M. Fischer, "On the connection between symmetric N-player games and mean field games," *Ann. Appl. Probability*, vol. 27, no. 2, pp. 757–810, 2017.
[19] W. H. Fleming.,*Functions of Several Variables*. Reading, MA, USA: Addison-Wesley, 1965.
[20] A. Gattami, "Team decision problems with convex quadratic constraints," *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 3, pp. 555–563, Sep. 2017.
[21] A. Gupta, S. Yüksel, T. Başar, and C. Langbort, "On the existence of optimal policies for a class of static and sequential dynamic teams," *SIAM J. Control Optim.*, vol. 53, pp. 1681–1712, 2015.
[22] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. New York, NY, USA: Springer, 1996.
[23] Y. C. Ho and K. C. Chu, "Team decision theory and information structures in optimal control problems - Part I," *IEEE Trans. Autom. Control*, vol. 17, no. 1, pp. 15–22, Feb. 1972.
[24] M. Huang, P. E. Caines, and R. P. Malhamé, "Large population stochastic dynamic games: closed-loop Mckean–Vlasov systems and the Nash certainty equivalence principle," *Commun. Inf. Syst.*, vol. 6, pp. 221–251, 2006.
[25] 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," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.
[26] M. Huang, P. E. Caines, and R. P. Malhamé, "Social optima in mean field LQG control: Centralized and decentralized strategies," *IEEE Trans. Autom. Control*, vol. 57, no. 7, pp. 1736–1751, Jul. 2012.
[27] M. Huang and S. L. Nguyen, "Stochastic control of mean field models with mixed players," in *Proc. 18th IFAC World Congr.*, Milan, Italy, 2011, pp. 4465–4470.
[28] M. Huang and S. L. Nguyen, "Linear-quadratic mean field teams with a major agent," in *IEEE 55th Conf. Decision Control*, 2016, pp. 6958–6963.
[29] B. Jovanovic and R. W. Rosenthal, "Anonymous sequential games," *J. Math. Econ.* vol. 17, no. 1, pp. 77–87, 1988.
[30] J. C. Krainak, J. L. Speyer, and S. I. Marcus, "Static team problems – part I: Sufficient conditions and the exponential cost criterion," *IEEE Trans. Autom. Control*, vol. AC-27, no. 4, pp. 839–848, Apr. 1982.
[31] D. Lacker, "A general characterization of the mean field limit for stochastic differential games," *Probability Theory Related Fields*, vol. 165, nos. 3/4, pp. 581–648, 2016.
[32] D. Lacker, "Limit theory for controlled Mckean–Vlasov dynamics," *SIAM J. Control Optim.*, vol. 55, no. 3, pp. 1641–1672, 2017.
[33] H. J. Langen, "Convergence of dynamic programming models," *Math. Oper. Res.*, vol. 6, no. 4, pp. 493–512, 1981.
[34] J. M. Lasry and P. L. Lions, "Mean field games," *Jpn. J. Math.*, vol. 2, pp. 229–260, 2007.
[35] A. Mahajan, N. C. Martins, and S. Yüksel, "Static LQG teams with countably infinite players," in *Proc. IEEE 52nd Annu. Conf. Decision Control*, 2013, pp. 6765–6770.
[36] J. Marschak, "Elements for a theory of teams," *Manag. Sci.*, vol. 1, pp. 127–137, 1955.
[37] A. Mas-Colell, "On a theorem of Schmeidler," *J. Math. Econ.*, vol. 13, no. 3, pp. 201–206, 1984.
[38] J. Nash, "Non-cooperative games," *Ann. Math.*, vol. 54, pp. 286–295, 1951.
[39] Y. Ouyang, S. M. Asghari, and A. Nayyar, "Optimal infinite horizon decentralized networked controllers with unreliable communication," 2018, *arXiv:1806.06497*.
[40] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, pp. 857–881, 1962.
[41] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Trans. Autom. Control.*, vol. 51, no. 2, pp. 274–286, Feb. 2006.
[42] N. Saldi, T. Başar, and M. Raginsky, "Approximate Nash equilibria in partially observed stochastic games with mean-field interactions," *Math. Oper. Res.*, vol. 44, no. 3, pp. 1006–1033, 2019.
[43] S. Sanjari and S. Yüksel, "Optimal stochastic teams with infinitely many decision makers and Mean-field teams," in *Proc. IEEE 57th Annu. Conf. Decision Control*, 2018, pp. 4123–4128.
[44] S. Sanjari and S. Yüksel, "Optimal policies for convex symmetric stochastic dynamic teams and their mean-field limit," 2020, *arXiv preprint arXiv:1903.11476v2*.
[45] D. Schmeidler, "Equilibrium points of nonatomic games," *J. Stat. Physics*, vol. 7, no. 4, pp. 295–300, 1973.
[46] R. Serfozo, "Convergence of Lebesgue integrals with varying measures," *Sankhyā: The Indian J. Statist., Series A*, vol. 44, pp. 380–402, 1982.
[47] B. C. Wang and J. F. Zhang, "Social optima in mean field linear-quadratic-Gaussian models with Markov jump parameters," *SIAM J. Control Optim.*, vol. 55, no. 1, pp. 429–456, 2017.
[48] H. S. Witsenhausen, "A standard form for sequential stochastic control," *Math. Syst. Theory*, vol. 7, pp. 5–11, 1973.
[49] H. S. Witsenhausen, "The intrinsic model for discrete stochastic control: Some open problems," in *Control Theory, Numerical Methods and Computer Systems Modelling*, vol. 107. NEW YORK, NY, USA: Springer-Verlag, 1975, pp. 322–335.
[50] S. Yüksel, "On stochastic stability of a class of non-Markovian processes and applications in quantization," *SIAM J. Control Optim.*, vol. 55, pp. 1241–1260, 2017.
[51] S. Yüksel, "A universal dynamic program and refined existence results for decentralized stochastic control," 2019, *arXiv:1803.05811v3*.
[52] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. New York, NY, USA: Springer, 2013.
[53] S. Yüksel and N. Saldi, "Convex analysis in decentralized stochastic control, strategic measures and optimal solutions," *SIAM J. Control Optim.*, vol. 55, pp. 1–28, 2017.

Sina Sanjari (Student Member, IEEE) received the M.Sc. degree in electrical and computer engineering from Tarbiat Modares University, Tehran, Iran, in 2015, and the M.Sc. degree in mathematics and statistics in 2017 from Queen's University, Kingston, ON, Canada, where he is currently working toward the Ph.D. degree with the Department of Mathematics and Statistics.

His research interests include stochastic control, decentralized control, mean-field games, and applications.

Serdar Yüksel (Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001, the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana—Champaign, Champaign, IL, USA, in 2003 and 2006, respectively.

He was a Postdoctoral Researcher with Yale University before joining Queen's University as an Assistant Professor with the Department of Mathematics and Statistics, where he is currently an Associate Professor. His research interests include stochastic control, decentralized control, information theory, and probability.

Dr. Yüksel received the 2013 CAIMS/PIMS Early Career Award in applied mathematics. He is an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and *Automatica*.