

OPTIMAL POLICIES FOR CONVEX SYMMETRIC STOCHASTIC DYNAMIC TEAMS AND THEIR MEAN-FIELD LIMIT*

SINA SANJARI[†] AND SERDAR YÜKSEL[†]

Abstract. This paper studies convex stochastic dynamic team problems with finite and infinite time horizons under decentralized information structures. First, we introduce two notions called exchangeable teams and symmetric information structures. We show that in convex exchangeable team problems an optimal policy exhibits a symmetry structure. We give a characterization for such symmetrically optimal teams for a general class of convex dynamic team problems under a mild conditional independence condition. In addition, through concentration of measure arguments, we establish the convergence of optimal policies for teams with N decision makers to the corresponding optimal policies for symmetric mean-field teams with infinitely many decision makers. As a by-product, we present an existence result for convex mean-field teams, where the main contribution of our paper is with respect to the information structure in the system when compared with the related results in the literature that have assumed either a classical information structure or a static information structure. We also apply these results to the important special case of linear quadratic Gaussian (LQG) team problems, where while for partially nested LQG team problems with finite time horizons it is known that the optimal policies are linear, for infinite horizon problems the linearity of optimal policies has not been established in full generality. We also study average cost finite and infinite horizon dynamic team problems with a symmetric partially nested information structure and obtain globally optimal solutions where we establish linearity of optimal policies.

Key words. stochastic teams, average cost optimization, decentralized control, mean-field teams

AMS subject classifications. 93E20, 90B99, 49J55

DOI. 10.1137/19M1284294

1. Introduction and literature review. Team problems consist of a collection of decision makers (DMs) or agents acting together to optimize a common cost function, but not necessarily sharing all the available information. The term *stochastic teams* refers to the class of team problems where there exist randomness in the initial states, observations, cost realizations, or the evolution of the dynamics. At each time stage, each agent only has partial access to the global information which is defined by the *information structure* of the problem [45]. If there is a predefined order in which the DMs act, then the team is called a *sequential team*. For sequential teams, 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*. Information structures can be further categorized as *classical*, *partially nested*, or *nonclassical*. An information structure is *classical* if the information of decision maker i (DM^i) contains all of the information available to DM^k for $k < i$. An information structure is *partially nested* if whenever the action of DM^k , for some $k < i$, affects the information of DM^i , then the information of DM^i contains the information of DM^k . An information structure which is not partially nested is *nonclassical*. A detailed review is presented in [49].

Obtaining structural results in team problems is important toward establishing both existence and computational/approximation methods for optimal policies. In this paper, we define the notion of exchangeable teams and symmetric information

*Received by the editors August 28, 2019; accepted for publication (in revised form) November 1, 2020; published electronically March 1, 2021. A summary of some of the results here is included in [37].

<https://doi.org/10.1137/19M1284294>

Funding: This research was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

[†]Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (16sss3@queensu.ca, yuksel@mast.queensu.ca).

structures, and we show that, for convex exchangeable dynamic teams with finite horizons, optimal policies exhibit a symmetry structure (Theorem 2.7). For any number of DMs, this symmetry structure is more relaxed when compared with the symmetry results developed earlier, e.g., in [38, 36], which focused on problems under a static information structure, and is applicable for dynamic teams which may not admit a static reduction, as long as convexity in policies holds for the team problem.

There have been many studies involving decentralized stochastic control with infinitely many DMs. In particular, when the coupling among the DMs is only through some aggregate/average effect, such problems can be viewed within the umbrella of mean-field games [26, 21], which were introduced as a limit model for noncooperative symmetric N -player differential games with a mean-field interaction as $N \rightarrow \infty$. The solution concept in game theory is often Nash equilibrium, and often under various characterizations of it in dynamic Bayesian setups. In the context of decentralized stochastic control or teams, these would correspond to person-by-person optimal solutions, and hence not necessarily globally optimal solutions.

Nonetheless, on the existence as well as uniqueness and nonuniqueness results on equilibria, there have been several studies for mean-field games [26, 5, 14, 28, 22, 7, 17]. There have also been several studies for mean-field games where the limits of sequences of Nash equilibria have been investigated as the number of DMs $N \rightarrow \infty$ (see, e.g., [15, 25, 6, 26, 4]). We refer interested readers to [13, 11] for a literature review and a detailed summary of some recent results on mean-field games.

Some notable relevant studies from the mean-field literature are the following. In [15], through a concentration of measures argument, it has been shown that sequences of ϵ_N -local (for each player) Nash equilibria for N player games converge to a solution for the mean-field game under exchangeability of the initial states and weak convergence of normalized occupational measures to a deterministic measure [15, Theorem 5.1]. In [23], assumptions on equilibrium policies of the large population mean-field symmetric stochastic differential games have been presented to allow the convergence of asymmetric approximate Nash equilibria to a weak solution of the mean-field game [23, Theorem 2.6].

However, in these studies the information structures are restricted to the following models. In [15] the information structure is assumed to be static since strategies of each player are assumed to be adapted to the filtration generated by his/her initial states and Wiener process (also called *distributed open-loop* controllers in the mean-field games literature [23, 14, 13]); see Remark 2 for details of this discussion. Convergence of Nash equilibria induced by closed-loop controllers to a weak semi-Markov mean-field equilibrium has been established in [25] for finite horizon mean-field game problems, where the classical information structure (i.e., what would be a centralized problem in the team theoretic setup) has been considered. For infinite horizon problems, in [12], an example of ergodic differential games with mean-field coupling has been constructed such that limits of sequences of expected costs induced by symmetric Nash-equilibrium policies of N -player games capture expected costs induced by many more Nash-equilibrium policies including a mean-field equilibrium and social optima. In [25], the classical information structure (a centralized problem) has been considered, where in [12] it has been assumed that players have access to all the history of states of all players but not controls (we note that in the team problem setup through using a classical result of Blackwell [9] in the case where each DM knows all the history of states of all DMs, optimal policies can be realized as one in the centralized problem where just the global state is a sufficient statistic). Moreover, under relaxed regularity conditions on dynamics and the cost function, a limit theory has been established for controlled McKean–Vlasov dynamics [24] under the classical information structure, where through a similar analysis as in [15, 23], it has

been shown that the empirical measure of pairs of states and ϵ_N -open-loop optimal controls converges weakly as $N \rightarrow \infty$ to limit points in the set of pairs of states and optimal controls of the McKean–Vlasov problem.

The above highlights the intricacies due to the information structure aspects: different from the studies above, we consider information structures that are not necessarily static or classical. Also, in this paper, we work with global optimality and not only mean-field equilibria and we show the existence of a globally optimal policy for mean-field team problems. On the other hand, in our paper since we work under the convexity assumption, the information structure does not allow for the mean-field coupling in the dynamics. We also note that in prior work [38], we studied static teams where under convexity and more restrictive symmetry conditions, global optimality of a limit policy of a sequence of N -DM optimal policies has been established.

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. For teams and social optima control problems, the analysis has primarily focused on the linear quadratic gaussian (LQG) model or Markov chains where the centralized performance has been shown to be achieved asymptotically by decentralized controllers (see, e.g., [20, 2, 3]).

We also obtain existence results on optimal policies for the setups considered. Compared to the results on the existence of a globally optimal policy in team problems where (finite) N -DM team problems have been considered [47, 16, 50, 34], we study convex team problems with countably infinite number of DMs.

Parts of our results in this paper correspond to LQG teams. In [18], it has been shown that for teams with finite number of DMs, dynamic teams with a partially nested information structure can be reduced to a static one [18, 46] where Radner's theorem concludes global optimality of linear policies for LQG team problems [31]. However, for average cost infinite horizon, partially nested, LQG dynamic team problems so far there has been no universal result establishing that a globally optimal policy is linear, time-invariant, and stabilizing, and this has been often imposed a priori. In [33], the problem of designing a linear, time-invariant, stabilizing, state feedback optimal policy for decentralized \mathcal{H}_2 -optimization problems, which satisfy the quadratic invariance property, has been addressed by reparametrizing the problem as a convex problem (via Youla parameterization). In [32], it has been shown that for sequential team problems involving linear systems, quadratic invariance and the partially nested property are equivalent. For a class of partially ordered systems, state space techniques have been utilized to obtain optimal, linear, time-invariant, state feedback controllers for \mathcal{H}_2 -optimization problems with sparsity constraints [41]. A similar result has been established in [42] where linearity and time invariance have been imposed a priori. In [27], \mathcal{H}_2 -optimization output feedback problems with two players have been considered and optimality results have been established when the optimal policies are restricted to linear, time-invariant, stabilizing policies. However, the results in [27, 33, 41, 42] are inconclusive regarding global optimality. Our contribution here is to consider average cost infinite horizon dynamic team problems without restricting the set of policies to those that are linear, time-invariant, and stabilizing, unlike the results in [27, 33, 41, 42]. We note again that the optimality of linear policies for infinite horizon LQG problems is an open problem in its generality and we provide positive results for a class of such problems.

Contributions. In view of the discussion above, our paper makes the following contributions.

- (i) We define a notion of exchangeable teams and symmetric information structures, and we show that, for convex exchangeable dynamic teams with finite horizons, optimal policies exhibit a symmetry structure (Theorem 2.7). For any number of DMs, this symmetry structure is more relaxed when compared with the symmetry results developed in [38, 36] and is applicable for dynamic teams which may not admit a static reduction, as long as convexity in policies holds for the team problem.
- (ii) For convex mean-field teams with a symmetric information structure, through concentration of measure arguments, we establish the convergence of optimal policies for mean-field teams with N DMs to the corresponding optimal policies for mean-field teams (see Theorem 3.2).
- (iii) We establish an existence result for the class of convex mean-field teams with a symmetric information structure (see Theorem 3.4) for finite horizon problems, where, as noted in the literature review, related results assumed more restrictive information structures which are either static or classical.
- (iv) We also apply our results to LQG dynamic teams for finite horizon problems (see section 4). For LQG dynamic teams with a symmetric partially nested information pattern, we obtain an optimal policy for finite horizon problems (see section 4.1). We also apply convex mean-field results to LQG mean-field teams with a symmetric partially nested information structure (see section 4.1) and obtain a globally optimal policy. Building on the result above, we also obtain a globally optimal policy for average cost LQG team problems.

The organization of the paper is as follows. We study convex exchangeable dynamic teams with finite horizons in section 2, and we study mean-field teams in section 3. We obtain globally optimal solutions for finite horizon problems with a symmetric partially nested information structure and LQG mean-field teams in section 4.1, and we discuss average cost LQG team problems with a symmetric information structure in section 4.2, respectively.

Notation. \mathbb{R} and \mathbb{N} denote the set of real numbers and natural numbers, respectively. We denote trace of a matrix A as $\text{Tr}(A)$. We denote A^T as the transpose of a matrix A and $A^{(T)}$ to show the dependence of a matrix A to $T \in \mathbb{N}$. For any random variables $z^{1:N} := (z^1, \dots, z^N)$, we define $z^{-i} := (z^1, \dots, z^{i-1}, z^{i+1}, \dots, z^N)$, and $\mathcal{M}_{r,q}$ denotes the space of $r \times q$ matrices.

1.1. Preliminaries. In this section, we introduce Witsenhausen's *intrinsic model* for sequential teams [45] (we generalize this definition to infinite number of DMs). Consider *sequential systems* and assume 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.

- There exists a collection of *measurable spaces* $\{(\Omega, \mathcal{F}), (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{Y}^i, \mathcal{Y}^i), i \in \mathcal{N}\}$, specifying the system's distinguishable events, and control and measurement spaces. The set \mathcal{N} denotes the collection of DMs. The set \mathcal{N} can be a finite set $\{1, 2, \dots, N\}$ or a countable set \mathbb{N} . The pair (Ω, \mathcal{F}) is a measurable space (on which an underlying probability may be defined). The pair $(\mathbb{U}^i, \mathcal{U}^i)$ denotes the Borel space from which the action u^i of DM ^{i} is selected. The pair $(\mathbb{Y}^i, \mathcal{Y}^i)$ denotes the Borel observation/measurement space.
- There is a *measurement constraint* to establish the connection between the observation variables and the system's distinguishable events. The \mathbb{Y}^i -valued observation variables are given by $y^i = h^i(\omega, \underline{u}^{1:i-1})$, where $\underline{u}^{1:i-1} = \{u^k, k \leq i-1\}$ and h^i 's are measurable functions.

- The set of admissible control laws $\underline{\gamma} = \{\gamma^i\}_{i \in \mathcal{N}}$, also called *designs* or *policies*, are measurable control functions, so that $u^i = \gamma^i(y^i)$. Let Γ^i denote the set of all admissible policies for DM^i and let $\Gamma = \prod_{i \in \mathcal{N}} \Gamma^i$.
- There is a *probability measure* P on (Ω, \mathcal{F}) describing the probability space on which the system is defined.

Under the intrinsic model, every DM acts separately. However, depending on the information structure, it may be convenient to consider a collection of DMs as a single DM acting at different time instances. In fact, in the classical stochastic control, this is the standard approach.

2. Finite horizon convex dynamic team problems with a symmetric information structure. In this section, we characterize symmetry in dynamic team problems. According to the discussion above, by considering a collection of DMs as a single DM ($i = 1, \dots, N$) acting at different time instances ($t = 0, \dots, T-1$), we define a team problem with (NT) -DMs as a team with N -DMs:

(i) Let the observation and action spaces be Borel subsets of \mathbb{R}^n for a positive integer n and be identical for each DM ($i = 1, \dots, N$) with $\mathbb{Y}_i := \mathbf{Y} = \prod_{t=0}^{T-1} \mathbb{Y}^t$, $\mathbb{U}_i := \mathbf{U} = \prod_{t=0}^{T-1} \mathbb{U}^t$, respectively. The sets of all admissible policies are denoted by $\mathbf{\Gamma} = \prod_{i=1}^N \Gamma_i = \prod_{i=1}^N \prod_{t=0}^{T-1} \Gamma^t$.

(ii) For $i = 1, \dots, N$, $y_t^i := h_t^i(x_0^{1:N}, \zeta_{0:t}^{1:N}, u_{0:t-1}^{1:N})$ represents the observation of DM^i at time t (h_t^i s are Borel measurable functions).

(iii) Let $(\underline{\zeta}^{1:N}) := (\underline{\zeta}^1, \dots, \underline{\zeta}^N)$ where $\underline{\zeta}^i := (x_0^i, \zeta_{0:T-1}^i)$ denotes all the uncertainty associated with DM^i including his/her initial states. We assume that $(\underline{\zeta}^i)$ takes values in Ω_ζ (where at each time instance t , it takes value in Ω_{ζ_t}). Let μ denote the law of $\underline{\zeta}^{1:N}$.

(iv) Define the expected cost function of $\underline{\gamma}^{1:N}$ as $J_T(\underline{\gamma}^{1:N}) = E^{\mathcal{I}^{1:N}}[c(\underline{\zeta}^{1:N}, \underline{u}^{1:N})]$ for some Borel measurable cost function $c : \prod_{i=1}^N (\Omega_\zeta \times \mathbf{U}) \rightarrow \mathbb{R}_+$, where $\underline{\gamma}^{1:N} = (\underline{\gamma}^1, \underline{\gamma}^2, \dots, \underline{\gamma}^N)$ and $\underline{\gamma}^i = \gamma_{0:T-1}^i$ for $i = 1, \dots, N$.

Now, we present the definition of symmetric information structures (note that symmetric information structures can be classical, partially nested, or nonclassical).

DEFINITION 2.1. *Let the information of DM^i acting at time t be described as $I_t^i := \{y_t^i\}$. The information structure of a sequential N -DM team problem is symmetric if*

(i) $y_t^i = h_t(x_0^i, x_0^{-i}, \zeta_{0:t}^i, \zeta_{0:t}^{-i}, u_{0:t-1}^i, u_{0:t-1}^{-i})$ where h_t is identical for all $i = 1, \dots, N$ (note that function's arguments depend on i).

We note that the above definition can be generalized to be applicable for teams with countably infinite DMs and infinite horizon problems.

The symmetric information structure can also be interpreted and defined as a graph, which has often been the common method to describe information structures in control theory, relating DMs and their information through directed edges. Consider $G(V, \mu)$ as a directed graph with $V = \{1, \dots, NT\}$ nodes and where $\mu \subset V \times V$ determines the directed edges between nodes; this represents the dependency notation in the information of nodes, i.e., (i, j) denotes a directed edge from i to j , $i \rightarrow j$, which means u^i affects y^j through the relation $y^j = h^j(\omega, \underline{u}^{1:j-1})$ defined in the intrinsic model (see section 1.1). We denote by $\downarrow j$ the set of nodes i such that $i \rightarrow j$ (ancestors), and $\downarrow\downarrow j := \{\downarrow j\} \cup \{j\}$. Similarly, we can define descendants by $\uparrow j$. We can define a collection of DMs as a single DM ($i = 1, \dots, N$) acting at different time instances ($t = 0, \dots, T-1$) on a graph with a symmetric information structure (two examples are shown in Figures 2.1 and 2.2). Assume as follows:

(i) There exists a node $\{i\}$ (root node), ω_0 . Each subgraph represents a single DM acting at time instances $t = 0, \dots, T-1$, and there exists a finite number of sub-

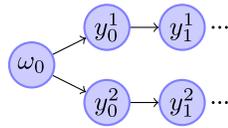


FIG. 2.1. A tree structure of a symmetric dynamic team.

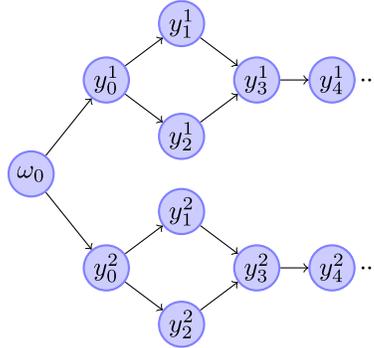


FIG. 2.2. An example of a graph structure of a symmetric dynamic team.

graphs $G_p(\hat{V}, \hat{\mu})$ such that $\cup_{p=1}^N G_p \cup \{i\} = G$, where G_p s are isomorphic (see, e.g., [43]) for all $p = 1, \dots, N$, i.e., for every node with directed edges in each subgraph there exists a unique node with identical directed edges in the corresponding subgraphs, where $\hat{V} = \{0, \dots, T - 1\}$, and G_p^k refers to a node k in G_p for all $p = 1, \dots, N$ and $k = 0, \dots, T - 1$,

(ii) Sharing of the information is symmetric across subgraphs, i.e., for $p, s = 1, \dots, N$, and $k, j = 0, \dots, T - 1$, and for every edge from a node G_p^k to a node G_s^j , there exists an edge from a node G_p^k to nodes G_{-p}^j , where G_{-p}^j denotes $(G_1^j, \dots, G_{p-1}^j, G_{p+1}^j, \dots, G_N^j)$, and also there exist edges from nodes G_{-p}^k to a node G_p^j .

Now, we present an exchangeability hypothesis on the cost function. First, we recall the definition of an *exchangeable* finite set of random variables.

DEFINITION 2.2. *Random variables (x^1, x^2, \dots, x^N) defined on a common probability space are exchangeable if for any permutation σ of the set $\{1, \dots, N\}$ (a mapping $\sigma : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$),*

$$P((x^\sigma)^1 \in A^1, (x^\sigma)^2 \in A^2, \dots, (x^\sigma)^N \in A^N) = P(x^1 \in A^1, x^2 \in A^2, \dots, x^N \in A^N)$$

for any measurable $\{A^1, \dots, A^N\}$ and $(x^\sigma)^i := x^{\sigma(i)}$ for all $i \in \{1, \dots, N\}$.

Assumption 2.1. For any permutation σ of the set $\{1, \dots, N\}$, we have for all ω_0

$$(2.1) \quad c(\omega_0, (\underline{\zeta}^\sigma)^{1:N}, (\underline{u}^\sigma)^{1:N}) = c(\omega_0, \underline{\zeta}^{1:N}, \underline{u}^{1:N}),$$

where $(\underline{\zeta}^\sigma)^{1:N} := (\underline{\zeta}^{\sigma(1)}, \dots, \underline{\zeta}^{\sigma(N)})$ and $(\underline{u}^\sigma)^{1:N} := (\underline{u}^{\sigma(1)}, \dots, \underline{u}^{\sigma(N)})$.

Here, we recall some definitions and results from [50, section 3.3] on convexity of static and dynamic team problems required to follow the result in this paper.

DEFINITION 2.3 ([50, section 3.3]). *An N -DM team problem (static or dynamic) is convex in policies if for any two team policies $\underline{\gamma}_T^{1:N}$ and $\tilde{\gamma}_T^{1:N}$ in the set $\{\underline{\gamma}_T^{1:N} \in \Gamma : J_T(\underline{\gamma}_T^{1:N}) < \infty\}$, and for any $\alpha \in (0, 1)$, we have*

$$J_T \left(\alpha \underline{\gamma}_T^{1:N} + (1 - \alpha) \tilde{\gamma}_T^{1:N} \right) \leq \alpha J_T \left(\underline{\gamma}_T^{1:N} \right) + (1 - \alpha) J_T \left(\tilde{\gamma}_T^{1:N} \right).$$

The above definition can also be applied to infinite horizon and/or teams with countably infinite number of DMs. We recall sufficient conditions for convexity of static and dynamic team problems following [50, section 3.3].

THEOREM 2.4 ([50, section 3.3]). *Consider a sequential team problem, and assume action spaces are convex, and the expected cost function is finite for all admissible policies $\underline{\gamma} \in \mathbf{\Gamma}$ (or alternatively, restrict the set to those leading to the finite cost). Then*

(i) *for static team problems convexity of the cost function in actions is sufficient for convexity of the team problem in policies;*

(ii) *for dynamic team problems with a static reduction, convexity of the team problem in policies is equivalent to the convexity of its static reduction;*

(iii) *in particular, for partially nested dynamic teams with a static reduction (more generally, for stochastically partially nested team problems [50, section 3.3]) if the cost function is convex in actions, then the reduced team problem with an equivalent information structure (via control sharing according to the partial nested information structure) is convex on $\mathbf{\Gamma}$.*

The conditions above, however, are only sufficient conditions [50, Example 1]. We note, however, that as a corollary for (ii) above, for the LQG setup, under partial nestedness, convexity in policies holds as a consequence of Radner's theorem; we will study this case in section 4. On the other hand, not all LQG problems are convex: the celebrated counterexample of Witsenhausen [44] demonstrates that under non-classical information structures, even LQG problems may not be convex and optimal policies may not be linear.

2.1. Optimality of symmetric policies for convex dynamic teams with a symmetric information structure. In the following, we define notions of exchangeable and symmetrically optimal teams analogous to [38, 36] for dynamic teams.

DEFINITION 2.5 (exchangeable teams). *An N -DM team is exchangeable if the value of the expected cost function is invariant under every permutation of policies of DMs, i.e., $J_T(\underline{\gamma}_T^1, \underline{\gamma}_T^2, \dots, \underline{\gamma}_T^N) = J_T((\underline{\gamma}_T^\sigma)^1, \dots, (\underline{\gamma}_T^\sigma)^N)$.*

DEFINITION 2.6 (symmetrically optimal teams). *A team is symmetrically optimal if for every given policy $\underline{\gamma}_T = (\underline{\gamma}_T^1, \dots, \underline{\gamma}_T^N)$, there exists an identically symmetric policy (i.e., each DM has the same policy, $\tilde{\gamma}_T = (\tilde{\gamma}_T^1, \dots, \tilde{\gamma}_T^N)$, and $\tilde{\gamma}_T^i = \tilde{\gamma}_T^j$ for all $i, j = 1, \dots, N$) which performs at least as well as the given policy.*

Remark 1. The concepts of exchangeable and symmetrically optimal dynamic teams in this paper are generalizations of those for static teams in [38, 36]. However, here, the value of the expected cost function may not be invariant under exchanging γ_t^i with γ_k^j for $k \neq t$, $k, t = 0, \dots, T - 1$, and for $i, j = 1, \dots, N$.

Here, we give a characterization for exchangeable and symmetrically optimal dynamic teams.

THEOREM 2.7. *Consider dynamic team problems with a symmetric information structure under Assumption 2.1. If*

- (a) *action spaces \mathbb{U}^t are convex for all $t = 0, \dots, T - 1$,*
- (b) *$(\underline{\zeta}^1, \dots, \underline{\zeta}^N)$ are exchangeable,*
- (c) *for all policies $\underline{\gamma} \in \mathbf{\Gamma}$, and for all $A = A^1 \times \dots \times A^N$ where $A^i \in \mathcal{Y}^i$,*

(2.2)

$$\begin{aligned} & \prod_{t=0}^{T-1} P\left(y_t^{1:N} \in A \mid x_0^{1:N}, \zeta_{0:t-1}^{1:N}, y_{0:t-1}^{1:N}, \gamma_0^1(y_0^1), \dots, \gamma_{t-1}^1(y_{t-1}^1), \gamma_0^2(y_0^2), \dots, \gamma_{t-1}^N(y_{t-1}^N)\right) \\ &= \prod_{t=0}^{T-1} \prod_{i=1}^N P\left(y_t^i \in A^i \mid x_0^i, \zeta_{0:t-1}^i, y_{\downarrow t}^{\downarrow\downarrow i}, \gamma_{\downarrow t}^{\downarrow\downarrow i}(y_{\downarrow t}^{\downarrow\downarrow i})\right), \end{aligned}$$

where $y_{\downarrow t}^{\downarrow\downarrow i} := \{y_p^j | u_p^j \text{ affects } y_t^i \text{ for all } p = 0, \dots, t-1 \text{ and for all } j = 1, \dots, N\}$ and $(\gamma_{\downarrow t}^{\downarrow\downarrow i}(y_{\downarrow t}^{\downarrow\downarrow i}))$ can be defined similarly,

- (i) then, the team problem is exchangeable.
- (ii) Furthermore, if the team problem is convex in policies (see Theorem 2.4), then the team is symmetrically optimal.

Proof. We first show that for any permutation $\sigma \in S$, $J_T((\underline{\gamma}_T^\sigma)^1, \dots, (\underline{\gamma}_T^\sigma)^N) = J_T(\underline{\gamma}_T^1, \dots, \underline{\gamma}_T^N)$, i.e., the team is exchangeable. We have

$$\begin{aligned} & J_T\left((\underline{\gamma}_T^\sigma)^1, \dots, (\underline{\gamma}_T^\sigma)^N\right) \\ (2.3) \quad &= \int c\left(\underline{\zeta}^{1:N}, (\underline{\gamma}_T^\sigma)^1(y^1), \dots, (\underline{\gamma}_T^\sigma)^N(y^N)\right) \mu(dx_0^{1:N}, d\zeta_{0:T-1}^{1:N}) \\ & \quad \times \prod_{t=0}^{T-1} \prod_{i=1}^N P\left(dy_t^i \mid x_0^i, \zeta_{0:t-1}^i, y_{\downarrow t}^{\downarrow\downarrow i}, (\gamma^\sigma)_{\downarrow t}^{\downarrow\downarrow i}(y_{\downarrow t}^{\downarrow\downarrow i})\right) \end{aligned}$$

$$\begin{aligned} (2.4) \quad &= \int c\left((\underline{\zeta}^\sigma)^{1:N}, (\underline{\gamma}_T^\sigma)^1((\underline{y}^\sigma)^1), \dots, (\underline{\gamma}_T^\sigma)^N((\underline{y}^\sigma)^N)\right) \\ & \quad \times \mu\left(d(x_0^\sigma)^{1:N}, d(\zeta_{0:T-1}^\sigma)^{1:N}\right) \\ & \quad \times \prod_{t=0}^{T-1} \prod_{i=1}^N P\left(d(y^\sigma)_t^i \mid (x_0^\sigma)^i, (\zeta_{0:t-1}^\sigma)^i, (y^\sigma)_{\downarrow t}^{\downarrow\downarrow i}, (\gamma^\sigma)_{\downarrow t}^{\downarrow\downarrow i}((y^\sigma)_{\downarrow t}^{\downarrow\downarrow i})\right) \end{aligned}$$

$$\begin{aligned} (2.5) \quad &= \int c\left(\underline{\zeta}^{1:N}, \underline{\gamma}_T^1(y^1), \dots, \underline{\gamma}_T^N(y^N)\right) \mu(dx_0^{1:N}, d\zeta_{0:T-1}^{1:N}) \\ & \quad \times \prod_{t=0}^{T-1} \prod_{i=1}^N P\left(dy_t^i \mid x_0^i, \zeta_{0:t-1}^i, y_{\downarrow t}^{\downarrow\downarrow i}, \gamma_{\downarrow t}^{\downarrow\downarrow i}(y_{\downarrow t}^{\downarrow\downarrow i})\right) \\ &= J_T\left(\underline{\gamma}_T^1, \dots, \underline{\gamma}_T^N\right), \end{aligned}$$

where (2.3) follows from condition (c). Equality (2.4) follows from exchanging $y^i, \underline{\zeta}^i$ with $(\underline{y}^\sigma)^i, (\underline{\zeta}^\sigma)^i$ by relabeling them, respectively. Since the information structure is symmetric, (2.1) and condition (b) imply (2.5). Hence, the team is exchangeable. Let $\underline{\gamma}_T^* = (\underline{\gamma}_T^{1*}, \dots, \underline{\gamma}_T^{N*})$ be a given policy. Consider $\tilde{\underline{\gamma}}_T$ as a convex combination of all possible permutations of policies by averaging them. Since action spaces are convex by condition (a), $\tilde{\underline{\gamma}}_T$ is a control policy. Following from convexity of the cost function in policies, we have

$$J_T(\tilde{\underline{\gamma}}_T) := J_T\left(\sum_{\sigma \in S} \frac{1}{|S|} \underline{\gamma}_T^{*,\sigma}\right) \leq \sum_{\sigma \in S} \frac{1}{|S|} J_T(\underline{\gamma}_T^{*,\sigma}) = J_T(\underline{\gamma}_T^*),$$

where $|S|$ denotes the cardinality of the set S and the inequality above follows from the hypothesis that the team problem is convex on Γ and the last equality follows from exchangeability of the team problem. This implies that the team is symmetrically optimal and completes the proof. \square

Examples will be given in sections 3 and 4.1, where Theorem 2.7 can be applied. Here, we present the result for a class of problems that admit a static reduction (see [49, section 3.7], [50, section 1.2], [19, 46]).

LEMMA 2.8. *Consider a dynamic team problem with a symmetric partially nested information structure (see Definition 2.1) which admits a static reduction. Under Assumption 2.1, and assumptions (a), (b), (c) of Theorem 2.7, if the cost function is jointly convex in $\underline{u}^1, \dots, \underline{u}^N$ P -almost surely, then the team is symmetrically optimal.*

We note again that here by symmetry, we mean symmetry across the DMs.

Proof. The proof follows from Theorems 2.4(iii) and 2.7 since the team is convex on Γ under the static reduction which is equivalent to the dynamic problem. \square

Hence, it follows that if a static reduction of an exchangeable, symmetrically optimal, dynamic team exists, then it is exchangeable and symmetrically optimal.

3. Convex mean-field teams with a symmetric information structure.

In the following, we establish global optimality results for convex mean-field teams with a symmetric information structure (that is not necessarily partially nested).

Define state dynamics and observations as

$$(3.1) \quad x_{t+1}^i = f_t(x_t^i, u_t^i, w_t^i),$$

$$(3.2) \quad y_t^i = h_t(x_{0:t}^i, u_{0:t-1}^i, v_{0:t}^i),$$

where functions f_t and h_t are measurable functions. The information structure of DM^{*i*} at time t is $I_t^i = \{y_t^i\}$, and $\zeta_t^i := (w_t^i, v_t^i)$ (with $\zeta_0^i := (x_0^i, w_0^i, v_0^i)$) denotes uncertainty corresponding to dynamics and observations at time t for DM^{*i*} which are exogenous random vectors in a standard Borel space. Denote $\mathbb{X} \subseteq \mathbb{R}^m$, $\mathbb{U} \subseteq \mathbb{R}^{m'}$, $\mathbb{Y} \subseteq \mathbb{R}^{m''}$, \mathbb{W} , and \mathbb{V} as the state space, action space, observation space, and space of disturbances of dynamics and observations of DMs at each time instance $t = 0, \dots, T-1$, respectively, where m, m' , and m'' are positive integers.

Problem ($\mathcal{P}_T^{N, \text{MF}}$). Consider N -DM teams with the expected cost function of $\underline{\gamma}_T^{1:N}$ as

$$(3.3) \quad J_T^N(\underline{\gamma}_T^N) = \frac{1}{N} \sum_{t=0}^{T-1} \sum_{i=1}^N E^{\underline{\gamma}_T^{1:N}} \left[c \left(\omega_0, x_t^i, u_t^i, \frac{1}{N} \sum_{p=1}^N u_t^p, \frac{1}{N} \sum_{p=1}^N x_t^p \right) \right],$$

where $\omega_0 : (\Omega, \mathcal{F}) \rightarrow (\Omega_0, \mathcal{F}_0)$ is an exogenous random vector in the standard Borel space and $\underline{\gamma}_T^{1:N} = \gamma_{0:T-1}^{1:N}$, and the cost function satisfies the following assumption.

Problem ($\mathcal{P}_T^{\infty, \text{MF}}$). Consider mean-field teams with the expected cost function of $\underline{\gamma}_T$ as

$$(3.4) \quad J_T^\infty(\underline{\gamma}_T) = \limsup_{N \rightarrow \infty} J_T^N(\underline{\gamma}_T),$$

where $J_T^N(\cdot)$ is defined in (3.3), $\underline{\gamma}_T^i = \gamma_{0:T-1}^i$ for $i \in \mathbb{N}$ and $\underline{\gamma}_T = \{\underline{\gamma}_T^i\}_{i \in \mathbb{N}}$.

Assumption 3.1. Assume

- (a) function $f_t : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \rightarrow \mathbb{X}$ is continuous in its first and second arguments for all w_t^i and for each $i \in \mathbb{N}$ and uniformly bounded,
- (b) function $h_t : \prod_{k=0}^t \mathbb{X} \times \prod_{k=0}^{t-1} \mathbb{U} \times \prod_{k=0}^t \mathbb{V} \rightarrow \mathbb{Y}$ is continuous in states and actions for all $v_{0:t}^i$ and for each $i \in \mathbb{N}$, and

- (c) the cost function in (3.3), $c : \Omega_0 \times \mathbb{X} \times \mathbb{U} \times \mathbb{U} \times \mathbb{X} \rightarrow \mathbb{R}_+$, is continuous in its second, third, fourth, and fifth arguments for all ω_0 .

3.1. Mean-field optimal policies as limits of optimal N -DM teams. In the following, we first establish global optimality results under Assumption 3.2 (see Theorem 3.2), then we establish the result under a more relaxed assumption, Assumption 3.3 (see Theorem 3.3).

Assumption 3.2. Assume

- (i) (x_0^1, x_0^2, \dots) are independent and identically distributed (i.i.d.) random vectors conditioned on ω_0 ,
(ii) for $t = 0, \dots, T-1$, $\{w_t^i\}_{i \in \mathbb{N}}$ are i.i.d. random vectors, and for $i \in \mathbb{N}$, $\{w_t^i\}_{t=0}^{T-1}$ are mutually independent, and independent of ω_0 and (x_0^1, x_0^2, \dots) . For $t = 0, \dots, T-1$, $\{v_t^i\}_{i \in \mathbb{N}}$ are i.i.d. random vectors, and for $i \in \mathbb{N}$, $\{v_t^i\}_{t=0}^{T-1}$ are mutually independent, and independent of ω_0 , (x_0^1, x_0^2, \dots) , and w_t^i 's for $i \in \mathbb{N}$ and $t = 0, \dots, T-1$.

Assumption 3.3. Assume that conditioned on ω_0 , (x_0^1, x_0^2, \dots) are exchangeable random vectors.

Later on we will establish an existence theorem under Assumption 3.2, and we note that the proof under Assumption 3.2 is more direct. This is why two separate theorems will be presented, and the proof of the latter will be built on that of the former.

LEMMA 3.1. *Consider a team defined as $(\mathcal{P}_T^{N,MF})$ with a symmetric information structure. Assume the team problem is convex in policies. Let the action space be compact and convex for each DMs. Under Assumptions 3.1 and 3.2, the team is symmetrically optimal.*

Proof. The proof follows from Theorem 2.7. \square

THEOREM 3.2. *Consider a team defined as $(\mathcal{P}_T^{\infty,MF})$ with $(\mathcal{P}_T^{N,MF})$ having a symmetric information structure for every N . Assume for every N the team problem is convex in policies. Let the action space be compact and convex for each DM. Under Assumptions 3.1 and 3.2, if there exists a sequence of optimal policies for $(\mathcal{P}_T^{N,MF})$, $\{\gamma_T^{*,N}\}_N$, which converges (for every DM due to the symmetry) pointwise to $\gamma_T^{*,\infty}$ as $N \rightarrow \infty$, then $\gamma_T^{*,\infty}$ (which is identically symmetric) is an optimal policy for $(\mathcal{P}_T^{\infty,MF})$.*

Proof. Following from Lemma 3.1, one can consider a sequence of N -DM teams which are symmetrically optimal that defines $(\mathcal{P}_T^{N,MF})$ and whose limit is identified with $(\mathcal{P}_T^{\infty,MF})$. Define

$$(3.5) \quad Q_N(B) := \frac{1}{N} \sum_{i=1}^N \delta_{\beta_N^i}(B), \quad \text{where} \quad \beta_N^i := (\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i),$$

$$\tilde{Q}_N(B) := \frac{1}{N} \sum_{i=1}^N \delta_{\beta_\infty^i}(B), \quad \text{where} \quad \beta_\infty^i := (\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i),$$

where $\delta_Y(\cdot)$ denotes the Dirac measure for any random vector Y , and $B \in \mathcal{Z} := \mathbf{U} \times \mathbf{Y} \times \mathbf{S}$, $\mathbf{U} := (\prod_{t=0}^{T-1} \mathbb{U})$, $\mathbf{Y} := (\prod_{t=0}^{T-1} \mathbb{Y})$, $\mathbf{S} := (\prod_{t=0}^{T-1} \mathbb{S}) = \mathbb{X} \times (\prod_{t=0}^{T-1} \mathbb{W} \times \mathbb{V})$, $\mathbf{X} = (\prod_{t=0}^{T-1} \mathbb{X})$, $\underline{y}^i = (y_0^i, \dots, y_{T-1}^i)$, and $\underline{\zeta}^i = (\zeta_0^i, \dots, \zeta_{T-1}^i)$.

In the following, first we show that conditioned on ω_0 , Q_N converges P -almost surely to $Q = \mathcal{L}(\beta_\infty^1 | \omega_0)$ in w -s topology (coarsest topology on $\mathcal{P}(\mathbf{U} \times \mathbf{Y} \times \mathbf{S})$ under which $\int f(u, y, \zeta) Q_N(du, dy, d\zeta) : \mathcal{P}(\mathbf{U} \times \mathbf{Y} \times \mathbf{S}) \rightarrow \mathbb{R}$ is continuous for every measurable and bounded function f which is continuous in u and y but need not to be continuous in ζ (see, e.g., [39] and [47, Theorem 5.6])). Then, we show that

$$\limsup_{N \rightarrow \infty} J_T^N(\tilde{\gamma}_T^{*,N}) = J_T^\infty(\tilde{\gamma}_T^{*,\infty}),$$

where $\tilde{\gamma}_T^{*,N} := (\gamma_T^{*,N}, \gamma_T^{*,N}, \dots, \gamma_T^{*,N})$ and $\tilde{\gamma}_T^{*,\infty} := (\gamma_T^{*,\infty}, \gamma_T^{*,\infty}, \dots)$.

Step 1. In this step, we show that conditioned on ω_0 , Q_N converges P -almost surely to Q in w -s topology. First, we show that for every continuous and bounded function g in actions and observations, for every ω_0 on a set of P -measure one,

$$(3.6) \quad P\left(\left\{\omega \in \Omega \mid \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N \left[g\left(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - g\left(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right)\right] = 0\right) \mid \omega_0\right\} = 1.\right.$$

Following from symmetry of the information structure and Lemma 3.1, every DM applies an identical optimal policy $\gamma_T^{*,N}$ and since functions f_t and h_t are identical for each DM, conditioned on ω_0 , $(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ and $(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors. For every $\epsilon > 0$ and for every function g continuous and bounded in actions and observations, we have P -almost surely

$$(3.7) \quad \begin{aligned} & \lim_{N \rightarrow \infty} P\left(\left|\int g dQ_N - \int g d\tilde{Q}_N\right| \geq \epsilon \mid \omega_0\right) \\ & \leq \epsilon^{-1} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E\left[\left|g\left(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - g\left(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right)\right| \mid \omega_0\right] \end{aligned}$$

$$(3.8) \quad = \epsilon^{-1} \lim_{N \rightarrow \infty} E\left[\left|g\left(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - g\left(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right)\right| \mid \omega_0\right]$$

$$(3.9) \quad = \epsilon^{-1} E\left[\lim_{N \rightarrow \infty} \left|g\left(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - g\left(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right)\right| \mid \omega_0\right] = 0,$$

where (3.7) follows from Markov's inequality, the triangle inequality, and the definition of the empirical measure, and (3.8) follows from the fact that $(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ and $(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors. Since g is bounded and continuous, the dominated convergence theorem implies (3.9). Hence, for every subsequence, there exists a subsubsequence such that P -almost surely $P(\{\omega \in \Omega \mid \lim_{N \rightarrow \infty} (\int g dQ_N - \int g d\tilde{Q}_N) = 0\} \mid \omega_0) = 1$.

Now, we show that conditioned on ω_0 , $\{\tilde{Q}_N\}_N$ converges weakly to Q P -almost surely. Since conditioned on ω_0 , $(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors, the strong law of large numbers implies that P -almost surely

$$(3.10) \quad P\left(\left\{\omega \in \Omega \mid \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N g\left(\gamma_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - E\left[g\left(\gamma_T^{*,\infty}(\underline{y}^1), \underline{y}^1, \underline{\zeta}^1\right) \mid \omega_0\right]\right) = 0\right\} \mid \omega_0\right) = 1,$$

hence, $P(\{\omega \in \Omega \mid \lim_{N \rightarrow \infty} (\int g d\tilde{Q}_N - \int g dQ) = 0\} \mid \omega_0) = 1$ P -almost surely.

Hence, through choosing a suitable subsequence, for every $\omega_0 \in \Omega_0$ on a set of P -measure one, for every function g continuous and bounded in actions and observations and measurable and bounded in uncertainty and initial states

$$\lim_{N \rightarrow \infty} \left| \int gdQ_N - \int gdQ \right| \leq \lim_{N \rightarrow \infty} \left(\left| \int gdQ_N - \int gd\tilde{Q}_N \right| + \left| \int gd\tilde{Q}_N - \int gdQ \right| \right) = 0,$$

hence, conditioned on ω_0 , Q_N converges weakly to Q P -almost surely. We note that the convergence is in the weak convergence topology, but since ζ^i 's are exogenous with a fixed probability measure, the convergence is also in the w - s topology.

Step 2. Following from (3.1) and (3.2), we have

$$(3.11) \quad x_t^i = f_{t-1}(f_{t-2}(\dots f_0(x_0^i, u_0^i, w_0^i))), u_{t-1}^i, w_{t-1}^i = \tilde{f}_{t-1}(\zeta^i, u_{0:t-1}^i),$$

$$(3.12) \quad y_t^i = h_t(x_0^i, \tilde{f}_0(\zeta^i, u_0^i), \dots, \tilde{f}_{t-1}(\zeta^i, u_{0:t-1}^i), u_{0:t-1}^i, v_{0:t}^i) = \tilde{h}_t(\zeta^i, u_{0:t-1}^i),$$

where following from Assumption 3.1, \tilde{f}_{t-1} and \tilde{h}_t are continuous in actions. Hence, under Assumption 3.1(c), we have

$$(3.13) \quad \begin{aligned} & \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{T-1} E^{\gamma_T^{*,1:N}} \left[c \left(\omega_0, x_t^i, u_t^i, \frac{1}{N} \sum_{p=1}^N u_t^p, \frac{1}{N} \sum_{p=1}^N x_t^p \right) \right] \\ &= \frac{1}{N} \sum_{i=1}^N E \left[\tilde{c} \left(\omega_0, \zeta^i, \gamma_T^{*,N}(\underline{y}^i), \frac{1}{N} \sum_{p=1}^N \gamma_T^{*,N}(y^p), \frac{1}{N} \sum_{p=1}^N \Lambda \left(\gamma_T^{*,N}(y^p), \zeta^p \right) \right) \right], \end{aligned}$$

where (3.13) is true following from (3.11) for some functions $\tilde{c} : \Omega_0 \times \mathbf{S} \times \mathbf{U} \times \mathbf{U} \times \mathbf{X} \rightarrow \mathbb{R}_+$ which are continuous in its last three arguments and a function $\Lambda : \mathbf{U} \times \mathbf{S} \rightarrow \mathbf{X}$ which is continuous in actions. Hence, by induction and rewriting observations as a functions of policies of the past DMs ($\gamma_{\downarrow t}^{*,N}(y_{\downarrow t}^i)$) since $\gamma_T^{*,N}$ converges to $\gamma_T^{*,\infty}$, the induced cost by $\gamma_T^{*,N}$ also converges to the cost induced by $\gamma_T^{*,\infty}$ P -almost surely.

Step 3. We have

$$(3.14) \quad \begin{aligned} & \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{T-1} E^{\gamma_T^{*,1:N}} \left[c \left(\omega_0, x_t^i, u_t^i, \frac{1}{N} \sum_{p=1}^N u_t^p, \frac{1}{N} \sum_{p=1}^N x_t^p \right) \right] \\ &= \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E \left[\tilde{c} \left(\omega_0, \zeta^i, \gamma_T^{*,N}(\underline{y}^i), \frac{1}{N} \sum_{p=1}^N \gamma_T^{*,N}(y^p), \frac{1}{N} \sum_{p=1}^N \Lambda \left(\gamma_T^{*,N}(y^p), \zeta^p \right) \right) \right] \end{aligned}$$

$$(3.15) \quad \begin{aligned} & \geq \liminf_{N \rightarrow \infty} E \left[E \left[\int_{\mathcal{Z}} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right. \right. \\ & \quad \left. \left. \times Q_N(du, dy, d\zeta) \middle| \omega_0 \right] \right] \end{aligned}$$

$$(3.16) \quad \begin{aligned} & \geq E \left[E \left[\liminf_{N \rightarrow \infty} \int_{\mathcal{Z}} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right. \right. \\ & \quad \left. \left. \times Q_N(du, dy, d\zeta) \middle| \omega_0 \right] \right] \end{aligned}$$

$$(3.17) \quad \begin{aligned} & \geq E \left[E \left[\int_{\mathcal{Z}} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) \right. \right. \\ & \quad \left. \left. \times Q(du, dy, d\zeta) \middle| \omega_0 \right] \right], \end{aligned}$$

where (3.14) follows from (3.13). Inequality (3.15) follows from (3.5) and replacing limsup with liminf, and (3.16) follows from Fatou's lemma. In the following, we justify (3.17). Since conditioned on ω_0 , Q_N converges weakly to Q P -almost surely, we have $Q_N(du \times \mathbf{Y} \times \mathbf{S})$ converges weakly to $Q(du \times \mathbf{Y} \times \mathbf{S})$ P -almost surely conditioned on ω_0 ; hence, the compactness of \mathbf{U} implies that conditioned on ω_0 , P -almost surely

$$(3.18) \quad \frac{1}{N} \sum_{i=1}^N \gamma_T^{*,N}(\underline{y}^i) = \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}) \xrightarrow{N \rightarrow \infty} \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}) = E[\gamma_T^{*,\infty}(\underline{y}^1) | \omega_0].$$

Since conditioned on ω_0 , Q_N converges weakly to Q P -almost surely, we have $Q_N(du \times \mathbf{Y} \times d\zeta)$ converges P -almost surely to $Q(du \times \mathbf{Y} \times d\zeta)$ in w - s topology conditioned on ω_0 . Following from (3.11), since f_t s are bounded and continuous in actions, Λ is bounded and continuous in actions; hence, this implies that conditioned on ω_0 , P -almost surely

$$(3.19) \quad \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \xrightarrow{N \rightarrow \infty} \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta).$$

Since the cost function \tilde{c} is continuous in its last three arguments, P -almost surely

$$\begin{aligned} & \lim_{N \rightarrow \infty} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \\ &= \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right). \end{aligned}$$

Define a nonnegative bounded function

$$\begin{aligned} G_N^M &:= \min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right\}, \\ G^M &:= \min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) \right\}, \end{aligned}$$

where the sequence $\{G^M\}_M$ converges P -almost surely as $M \rightarrow \infty$ to

$$G := \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right).$$

We have P -almost surely

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \int_{\mathcal{Z}} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) Q_N(du, dy, d\zeta) \\ (3.20) & \geq \lim_{M \rightarrow \infty} \liminf_{N \rightarrow \infty} \int_{\mathcal{Z}} G_N^M Q_N(du, dy, d\zeta) \end{aligned}$$

$$(3.21) = \lim_{M \rightarrow \infty} \int_{\mathcal{Z}} G^M Q(du, dy, d\zeta)$$

$$(3.22) = \int_{\mathcal{Z}} \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) Q(du, dy, d\zeta),$$

where (3.20) is true since

$$\tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \geq G_N^M.$$

Equality (3.21) follows from the generalized convergence theorem in [40, Theorem 3.5] since G_N^M is bounded and continuously converges to G^M , i.e., P -almost surely

$$(3.23) \quad \lim_{N \rightarrow \infty} \min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u_N, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ = \min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) \right\},$$

when $u_N \rightarrow u$ as $N \rightarrow \infty$. The monotone convergence theorem implies (3.22). Hence, (3.17) holds, which implies $\limsup_{N \rightarrow \infty} J_T^N(\tilde{\gamma}_T^{*,N}) = J_T^\infty(\tilde{\gamma}_T^{*,\infty})$, and this completes the proof following from [38, Theorem 5]. Here, for completeness we present the proof which is similar to the analysis of the proof [38, Theorem 5] for dynamic teams,

$$(3.24) \quad \inf_{\underline{\gamma}_T} J_T^\infty(\underline{\gamma}_T) \leq \limsup_{N \rightarrow \infty} J_T^N(\tilde{\gamma}_T^{*,\infty}) \\ = \limsup_{N \rightarrow \infty} J_T^N(\underline{\gamma}_T^{*,N}) = \limsup_{N \rightarrow \infty} \inf_{\underline{\gamma}_T^{1:N}} J_T^N(\underline{\gamma}_T^{1:N}) \\ = \limsup_{N \rightarrow \infty} \inf_{\underline{\gamma}_T} J_T^N(\underline{\gamma}_T) \\ \leq \inf_{\underline{\gamma}_T} \limsup_{N \rightarrow \infty} J_T^N(\underline{\gamma}_T) = \inf_{\underline{\gamma}_T} J^\infty(\underline{\gamma}_T),$$

where (3.24) is true since the restriction $\underline{\gamma}_T$ to the first N components is $\underline{\gamma}_T^{1:N}$. This implies that $\tilde{\gamma}_T^{*,\infty}$ is globally optimal. \square

Remark 2. On the connection between finitely many DMs and infinitely many DMs, we note a closely related work on mean-field games by Fischer [15], where the information structure is assumed to be static since the policy of each player is assumed to be adapted to the filtration generated by his/her initial states and Wiener process (also called in the mean-field games literature, somewhat nonstandard in the control literature, *open-loop distributed* controllers [23], [14, pp. 72–76]). This means that the information of each DM is not affected by any of the actions of the other DMs. For dynamic teams, there are two difficulties: (1) obtaining variational equations is challenging since fixing policies of DMs and perturbing only DM's policies perturbs the observation of other DMs and hence the controls $u^{-i*} = (\gamma^{1*}(y^1), \dots, \gamma^{i-1*}(y^{i-1}), \gamma^{i+1*}(y^{i+1}), \dots, \gamma^{N*}(y^N))$; (2) solutions of variational equations which give person-by-person optimal policies are inconclusive for global optimality due to the lack of convexity in general.

Remark 3. We also note additional related works by Lacker [24, 25], where either convergence of open-loop controllers or convergence of Nash equilibria induced by closed-loop controllers (where controls are measurable path-dependent functions of states, $u_t^i = \phi(t, \mathbf{x}_{0:t})$, where $\mathbf{x}_{0:t} = (x_{0:t}^1, \dots, x_{0:t}^N)$ and ϕ is a measurable function) or Markovian controllers ($u_t^i = \phi(t, \mathbf{x}_t)$, where $\mathbf{x}_t = (x_t^1, \dots, x_t^N)$) have been considered. In [25], the information structure is classical (a centralized problem) since players have access to all the information available to previous DMs.

Remark 4. In Lemma 3.1 and Theorem 3.2, we considered a nonclassical information structure for teams defined as $(\mathcal{P}_T^{\infty, \text{MF}})$ with a convex expected cost in policies. For teams defined as $(\mathcal{P}_T^{\infty, \text{MF}})$ with a symmetric partially nested information structure which admit static reduction, the above result holds, and similar to the proof of Theorem 3.2, it can be proven under the assumption that the cost functions is convex in actions (since convexity of the cost function in actions is a sufficient condition

for convexity of the expected cost function in policies for this class of problems [50, Theorem 3.7]).

Remark 5. Assumptions that action spaces are compact and f_t s are bounded can be relaxed by assuming that

$$(A1) \quad \sup_{N \geq 1} E[|\underline{\gamma}_T^{*,N}(\underline{y}^1)|^{1+\delta}] < \infty \text{ for some } \delta > 0,$$

$$(A2) \quad \sup_{N \geq 1} E[|\Lambda(\underline{\gamma}_T^{*,N}(\underline{y}^1), \underline{\zeta}^1)|^{1+\tilde{\delta}}] < \infty \text{ for some } \tilde{\delta} > 0.$$

That is because, following from the pointwise convergence of $\underline{\gamma}_T^{*,N}$ and continuity of Λ in actions, the above uniform integrability assumption justifies exchanging the limit and the expectation required to establish the convergence in (3.18) and (3.19) using a similar analysis of (3.9) and an argument of (3.10) based on the strong law of large numbers. This result is particularly important for LQG models (we use this remark in section 4).

THEOREM 3.3. *Consider a team defined as $(\mathcal{P}_T^{\infty, MF})$ with $(\mathcal{P}_T^{N, MF})$ having a symmetric information structure for every N . Assume for every N the team problem is convex in policies. Let the action space be compact and convex for each DM, and assume Assumptions 3.1, 3.2(ii), and 3.3 hold. If there exists a sequence of optimal policies for $(\mathcal{P}_T^{N, MF})$, $\{\underline{\gamma}_T^{*,N}\}_N$, which converges (for every DM due to the symmetry) pointwise to $\underline{\gamma}_T^{*,\infty}$ as $N \rightarrow \infty$, then $\underline{\gamma}_T^{*,\infty}$ (which is identically symmetric) is an optimal policy for $(\mathcal{P}_T^{\infty, MF})$.*

Proof. Under Assumptions 3.2(ii) and 3.3, for every $A^i \in \mathcal{B}(\mathbf{S})$, and $A^i = B^i \times \prod_{t=0}^{T-1} (D_t^i \times E_t^i)$ (where $B^i \in \mathcal{B}(\mathbb{X})$, $D_t^i \in \mathcal{B}(\mathbb{W})$, and $E_t^i \in \mathcal{B}(\mathbb{V})$), for all $N \in \mathbb{N}$, and permutations σ , we have P -almost surely

$$(3.25) \quad \begin{aligned} & P(\underline{\zeta}^1 \in A^1, \dots, \underline{\zeta}^N \in A^N | \omega_0) \\ &= P(x_0^1 \in B^1, \dots, x_0^N \in B^N | \omega_0) \prod_{i=1}^N \prod_{t=0}^{T-1} P(w_t^i \in D_t^i) P(v_t^i \in E_t^i) \end{aligned}$$

$$(3.26) \quad \begin{aligned} &= P((x_0^\sigma)^1 \in B^1, \dots, (x_0^\sigma)^N \in B^N | \omega_0) \\ &\quad \times \prod_{i=1}^N \prod_{t=0}^{T-1} P((w_t^\sigma)^i \in D_t^i) P((v_t^\sigma)^i \in E_t^i) \\ &= P((\underline{\zeta}^\sigma)^1 \in A^1, \dots, (\underline{\zeta}^\sigma)^N \in A^N | \omega_0), \end{aligned}$$

where (3.25) follows from Assumption 3.2(ii), and (3.26) follows from Assumption 3.3. Hence, $(\underline{\zeta}^1, \underline{\zeta}^2, \dots)$ are exchangeable conditioned on ω_0 .

Hence, following from a similar argument as in the proof of Theorem 2.7 (by considering ω_0 in the cost function and the law of total expectation (by first conditioning on ω_0), under Assumptions 3.2(ii) and 3.3, one can consider a sequence of N -DM teams which are symmetrically optimal that defines $(\mathcal{P}_T^{N, MF})$ and whose limit is identified with $(\mathcal{P}_T^{\infty, MF})$. Since initial states are not necessarily independent conditioned on ω_0 , we cannot establish that $(\underline{\gamma}_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors conditioned on ω_0 which has been used in (3.10) in Step 1 of the proof of Theorem 3.2 to show that Q_N converges weakly to Q P -almost surely.

However, we note that since $(\underline{\zeta}^1, \underline{\zeta}^2, \dots)$ are exchangeable conditioned on ω_0 , for every $A^i \in \mathcal{B}(\mathbf{S})$ and $C \in \mathcal{B}(\Omega_0)$, and for all $N \in \mathbb{N}$, and permutations σ , we have

$$(3.27) \quad P(\underline{\zeta}^1 \in A^1, \dots, \underline{\zeta}^N \in A^N, \omega_0 \in C) = P((\underline{\zeta}^\sigma)^1 \in A^1, \dots, (\underline{\zeta}^\sigma)^N \in A^N, \omega_0 \in C).$$

Let $\alpha^i := (\omega_0, \underline{\zeta}^i)$. Hence, (3.27) implies that $(\alpha^1, \alpha^2, \dots)$ is exchangeable. Following from [1, Proposition 3.8(a)], there exists a random vector $z \in [0, 1]$ such that $(\underline{\zeta}^1, \underline{\zeta}^2, \dots)$ are i.i.d. random vectors conditioned on (ω_0, z) .

Let $\tilde{\omega}_0 := (\omega_0, z)$. Hence, under Assumptions 3.2(ii) and 3.3, conditioned on $\tilde{\omega}_0$, $(\underline{\zeta}^1, \underline{\zeta}^2, \dots)$ are i.i.d. random vectors. Following from standard stochastic realization results [10, Lemma 3.1], we can represent any stochastic kernel in a functional form, with almost sure equivalence, $\underline{\zeta}^i = g(\tilde{\omega}_0, \theta^i)$ for some independent θ^i and measurable g (note that following from exchangeability, g is identical for all $i \in \mathbb{N}$ and $(\theta^1, \theta^2, \dots)$ are i.i.d. random vectors).

Since conditioned on $\tilde{\omega}_0$, $(\underline{\zeta}^1, \underline{\zeta}^2, \dots)$ are i.i.d. random vectors, $(\underline{\gamma}_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors conditioned on $\tilde{\omega}_0$; hence for every $\tilde{\omega}_0$ on a set of P -measure one, we have for every continuous and bounded function g in actions and observations, by the strong law of large numbers,

$$P\left(\left\{\omega \in \Omega \mid \lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{i=1}^N g\left(\underline{\gamma}_T^{*,\infty}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i\right) - E\left[g\left(\underline{\gamma}_T^{*,\infty}(\underline{y}^1), \underline{y}^1, \underline{\zeta}^1\right) \mid \tilde{\omega}_0\right] \right| = 0\right\} \mid \tilde{\omega}_0\right) = 1.$$

Hence, following from an identical analysis as that of Step 1 of the proof of Theorem 3.2, conditioned on $\tilde{\omega}_0$, Q_N converges weakly to Q , for every $\tilde{\omega}_0$ on a set of P -measure one.

Following from the representation $\underline{\zeta}^i = g(\tilde{\omega}_0, \theta^i)$, we have

$$\begin{aligned} & \limsup_{N \rightarrow \infty} E\left[E\left[\int_{\mathcal{Z}} \tilde{c}\left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta)\right) \right. \right. \\ & \quad \left. \left. \times Q_N(du, dy, d\zeta) \mid \omega_0\right]\right] \\ (3.28) \quad & = \limsup_{N \rightarrow \infty} \int_{\Omega_0 \times [0,1]} E\left[\int_{\mathcal{Z}} \tilde{c}\left(\tilde{\omega}_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta)\right) \right. \\ & \quad \left. \times Q_N(du, dy, d\zeta) \mid \tilde{\omega}_0\right] P(d\tilde{\omega}_0), \end{aligned}$$

where (3.28) follows from the fact that for all $N \in \mathbb{N}$, and for every $A^i \in \mathcal{B}(\mathbf{S})$

$$\int_{\Omega_0} P(\underline{\zeta}^1 \in A^1, \dots, \underline{\zeta}^N \in A^N \mid \omega_0) P(d\omega_0) = \int_{\Omega_0 \times [0,1]} \prod_{i=1}^N \eta(\underline{\zeta}^i \in A^i \mid z, \omega_0) P(dz, d\omega_0),$$

and with slight abuse of notation we use the same notation, \tilde{c} , for the cost function after transformation in (3.28). The rest of the proof is identical to that of Theorem 3.2. \square

3.2. An existence theorem on globally optimal policies for dynamic mean-field team problems with a symmetric information structure. An implication of Theorem 3.2 is the following existence result on globally optimal policies for mean-field team problems. In particular, we will establish the existence of a converging subsequence, in an appropriate sense, for a sequence of optimal policies for N -DM teams with an increasing number of DMs. For the following theorem, we do not establish the pointwise convergence, but by Theorem 3.2, if a sequence of optimal policies for $(\mathcal{P}_T^{N, MF}, \{\underline{\gamma}_T^{*,N}\}_N)$, converges pointwise, a global optimal policy exists. To this end, we allow DMs to apply randomized policies. For each DM (DM ^{i} for

$i \in \mathbb{N}$), a probability measure $P \in \mathcal{P}(\Omega_0 \times \mathbb{X} \times \prod_{t=0}^{T-1} (\mathbb{W} \times \mathbb{V}) \times \prod_{t=0}^{T-1} (\mathbb{U} \times \mathbb{Y}))$ is a policy induced by a randomized policy if and only if for every $t = 0, \dots, T-1$ and for all continuous and bounded function g

$$(3.29) \quad \begin{aligned} & \int g(\omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t}^i, u_{0:t}^i) P(d\omega_0, dx_0^i, d\zeta_{0:t-1}^i, dy_{0:t}^i, du_{0:t}^i) \\ &= \int g(\omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t}^i, u_{0:t}^i) \mu^i(dx_0^i, d\zeta_{0:t-1}^i | \omega_0) P(d\omega_0) \\ & \quad \times \prod_{k=0}^t \Pi_k^i(dy_k^i | y_k^i) p_k^i(dy_k^i | \omega_0, x_0^i, \zeta_{0:k-1}^i, y_{0:k-1}^i, u_{0:k-1}^i), \end{aligned}$$

for a stochastic kernel Π_k^i on \mathbb{U} given \mathbb{Y} , where p_k^i is the transition kernel characterizing the observations of DM^i at time t ,

$$\begin{aligned} & p_k^i(y_k^i \in \cdot | \omega_0, x_0^i, \zeta_{0:k-1}^i, y_{0:k-1}^i, u_{0:k-1}^i) \\ & := P(h_k(x_{0:k}^i, u_{0:k-1}^i, v_{0:k}^i) \in \cdot | \omega_0, x_0^i, \zeta_{0:k-1}^i, y_{0:k-1}^i, u_{0:k-1}^i), \end{aligned}$$

and μ^i is a fixed probability measure on initial states and disturbances of DM^i conditioned on ω_0 . This equivalency follows from the fact that continuous and bounded functions form a separating class [8, p. 12] and [48, Theorem 2.2].

First, we present an absolute continuity assumption on observations of DMs.

Assumption 3.4. For every DM^i and $t = 0, \dots, T-1$, there exists a function $\psi_t^i : \mathbb{Y} \times \Omega_0 \times \mathbb{X} \times \prod_{k=0}^{t-1} (\mathbb{W} \times \mathbb{V}) \times \prod_{k=0}^{t-1} (\mathbb{Y} \times \mathbb{U}) \rightarrow \mathbb{R}_+$ continuous in actions, and a probability measure ν_t^i on \mathbb{Y} such that for all Borel sets $A = A^1 \times \dots \times A^N$,

$$\begin{aligned} & P(y_t^{1:N} \in A | \omega_0, x_0^{1:N}, \zeta_{0:t-1}^{1:N}, y_{0:t-1}^{1:N}, u_{0:t-1}^{1:N}) \\ &= \prod_{i=1}^N \int_{A^i} \psi_t^i(y_t^i, \omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t-1}^i, u_{0:t-1}^i) \nu_t^i(dy_t^i). \end{aligned}$$

This assumption allows us to obtain an independent measurements reduction (see [47, section 2.2]). For example, if v_t^i for all $i \in \mathbb{N}$ and $t = 0, \dots, T-1$ are i.i.d with a probability measure admitting a density function so that the observation of each DM^i at time t is $y_t^i = \tilde{h}_t(x_t^i, u_{\downarrow t}^i) + v_t^i$, where \tilde{h}_t is continuous, then Assumption 3.4 holds [16, Lemma 5.1].

THEOREM 3.4. Consider a team defined as $(\mathcal{P}_T^{\infty, MF})$ with $(\mathcal{P}_T^{N, MF})$ having a symmetric information structure for every N . Assume for every N the team problem is convex in policies and the action space is convex. Assume further that without any loss, the optimal policies can be restricted to those with $E(\phi_i(u^i)) \leq K$ for some finite K , where $\phi_i : \mathbb{U} \rightarrow \mathbb{R}_+$ is lower semicontinuous (moment condition). Under Assumptions 3.1 and 3.2 if either

- (i) Assumption 3.4 holds (with no further assumptions on the information structure of each DM^i for $i \in \mathbb{N}$ through time $t = 0, \dots, T-1$), or
- (ii) for each DM^i for $i \in \mathbb{N}$ through time $t = 0, \dots, T-1$, there exists a static reduction with the classical information structure (i.e., under a static reduction, the information structure is expanding such that $\sigma(y_t^i) \subset \sigma(y_{t+1}^i)$ where σ denotes the σ -field),

then there exists an optimal policy for $(\mathcal{P}_T^{\infty, MF})$.

Since the space of policies that are deterministic (where Π_k^i in (3.29) are indicator functions) is not closed under the weak convergence topology (e.g., as an implication

of [50, Theorem 2.7]), we allow for randomization in the policies and therefore the limit policy is not necessarily deterministic according to the above result; however, it is identical for each DM.

Proof. We use individually randomized policies and we show that for every sequence of N -DM optimal policies, there exists a subsequence which converges to an optimal independently randomized policy for the mean-field limit under an appropriate topology defined by the product topology where each coordinate is endowed with the weak convergence topology. In Step 1, we show that for each finite N -DM team problem, optimal policies are deterministic and symmetric and we consider the independently randomized policies induced by such policies $\{P_N\}_N$ (where $P_N \in \mathcal{P}(\mathbf{Y} \times \mathbf{U})$) for each DM satisfying (3.29) as our sequence to be studied. We also define the sequence of empirical measures induced by these policies, Q_N , as (3.5).

In Step 2, we show that for every sequence of policies satisfying a moment condition there exists a subsequence such that policies $\{P_n\}_n$ for each DM and a subsequence of empirical measures $\{Q_n\}_n$ induced by these policies (where $n \in \mathbb{I}$ is the index set of a convergent subsequence) converge weakly to a limit P -almost surely, that is, for a set of P -measure one, for every bounded function g which is continuous in actions and observations and measurable in uncertainties,

$$P\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} \left(\int g dQ_n - \int g dQ\right) = 0\right\} \mid \omega_0\right) = 1.$$

To this end, we first show that for each DM a sequence $\{P_n\}_n$ is tight; then we show that the sequence of empirical measures $\{Q_n\}_n$ induced by these policies converges weakly to a limit P -almost surely.

In Step 3, we show that the set of policies for each DM is closed under the weak convergence topology; hence, the limit policy satisfies the required measurability/conditional independence constraints (that is, the limit policy satisfies (3.29)). In Step 4, we use the lower semicontinuity argument to show that the expected cost function under the induced limit policy is less than or equal to the expected cost achieved by the sequence of N -DM optimal policies.

Step 1. Under Assumptions 3.1, 3.1(c), and 3.2, and by condition (i) using [47, Theorem 5.2], or condition (ii) using [47, Theorem 5.6], there exists a deterministic optimal policy for each finite N -DM team problem. Action spaces are convex and the team problem is convex in policies, hence, using Lemma 3.1, one can consider a sequence of N -DM teams which are symmetrically optimal that defines $(\mathcal{P}_T^{N,MF})$ and whose limit is identified with $(\mathcal{P}_T^{\infty,MF})$. Hence, for each N -DM team problem, we consider symmetric randomized optimal policies.

Step 2. In the following, we first show that the set of policies $P_N \in \mathcal{P}(\mathbf{Y} \times \mathbf{U})$ for each DM satisfying (3.29) and the moment condition is tight; then, by symmetry, we show that $\{Q_N\}_N$ induced by this set of policies is tight. We use the fact that conditioned on ω_0 , $(\gamma_T^{*,N}(\underline{y}^i), \underline{y}^i, \underline{\zeta}^i)$ are i.i.d. random vectors (this follows from symmetry of the information structure and Lemma 3.1 since every DM applies the identical policy $\gamma_T^{*,N}$) and also since the space of control policies is tight under the weak convergence for each DM (see, e.g., [47, proof of Theorem 4.7]).

Since actions of DMs do not affect the observations of others, the policy spaces are decoupled from the actions of other DMs. Since we can restrict the search for optimal policies over those satisfying the moment condition, the fact that $\nu \rightarrow \int \nu(dx)g(x)$ is lower semicontinuous for a continuous function g [47, proof of Theorem 4.7] implies that the marginals on \mathbf{U} satisfying the moment condition are tight under the weak convergence topology. Hence, the collection of all probability measures with these

tight marginals is also tight (see, e.g., [48, proof of Theorem 2.4]). This implies that the sequence of randomized policies satisfying the moment condition is tight.

Since every DM applies an identical policy and since observations are conditionally i.i.d., a countably infinite product of space of policies of DMs is tight (where each coordinate is tight in the weak convergence topology). Hence, there exists a subsequence of policies $\tilde{P}_n \in \mathcal{P}(\prod_i(\mathbf{Y} \times \mathbf{U}))$ (as a product of policies of DMs) converging weakly to a limit \tilde{P} (each coordinate converges weakly) P -almost surely. Furthermore, since every DM applies an identical policy, conditioned on ω_0 , actions induced by an identical randomized policies, observations, and disturbances are i.i.d. through DMs. Hence, following from a similar argument as in Step 1 of the proof of Theorem 3.2, a subsequence of empirical measures $\{Q_n\}_{n \in \mathbb{I}}$ converges P -almost surely to Q in w -s topology. We note that the convergence is under the weak convergence topology, but since ζ^i 's are exogenous with a fixed marginal, the convergence is also in the w -s topology.

Step 3. In this step, we show that each coordinate of the space of policies (space of policies for each DM) is closed under the weak convergence topology. This in particular implies that the space of policies is closed under the product topology and using (Step 1), we can conclude that the space of control policies is compact under the product topology where each coordinate is weakly compact.

Assume P_n is a policy for DM^{*i*} induced by a randomized policy converging weakly to P_∞ . In fact, condition (i) or (ii) leads to the closedness of the set of policies (see (3.29)) induced by P_n . If Assumption 3.4 holds, then by the discussion in the proof of [47, Theorem 5.2], each coordinate of policy spaces corresponding to DM^{*i*} acting through time is closed under the weak convergence topology. Also, if condition (ii) holds, then [47, Theorem 5.6] leads to the same conclusion. Hence, each coordinate of space of policies (corresponding to DM^{*i*}) is closed under the weak convergence topology (since each coordinate of the space of policies is a finite product of space of policies for each DM at time instances $t = 0, \dots, T-1$). Hence, following from Step 2, there exists a subsequence $\{Q_n\}_{n \in \mathbb{I}}$ converging weakly to Q P -almost surely where Q is induced by a randomized policy in the set of policies satisfying (3.29) and the limit policy is admissible and satisfies the required measurability/conditional independence constraints.

For every $t = 0, \dots, T-1$, let P_n^{*,ω_0} be a probability measure on actions, observations, and uncertainties induced by optimal randomized policies for each DM (which is identical because of symmetry) for N -DM teams conditioned on ω_0 , i.e., a probability measure that satisfies

$$\begin{aligned} & \int g(\omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t}^i, u_{n,0:t}^{i,*}) P_n^{*,\omega_0}(dx_0^i, d\zeta_{0:t-1}^i, dy_{0:t}^i, du_{n,0:t}^{i,*}) \\ (3.30) \quad &= \int g(\omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t}^i, u_{n,0:t}^{i,*}) \mu^i(dx_0^i, d\zeta_{0:t-1}^i | \omega_0) \\ & \quad \times \prod_{k=0}^t \Pi_k^{*,n}(du_{n,k}^{*,i} | y_k^i) p_k(dy_k^i | \omega_0, x_0^i, \zeta_{0:k-1}^i, y_{0:k-1}^i, u_{n,0:k-1}^{i,*}), \end{aligned}$$

for all bounded functions g which is continuous in actions and observations and measurable in other arguments. We denote $\underline{u}_n^{i,*} := (u_{n,0}^{i,*}, \dots, u_{n,T-1}^{i,*})$ as the action of DM^{*i*} through time induced by $\Pi_k^{*,n}$. Similarly, we denote P^{*,ω_0} as a probability measure induced by the limit policy, i.e., a probability measure satisfying (3.30) induced by $\Pi_k^{*,\infty}$ where $\underline{u}_\infty^{i,*} := (u_{\infty,0}^{i,*}, \dots, u_{\infty,T-1}^{i,*})$ is the action of DM^{*i*} through time induced by $\Pi_k^{*,\infty}$.

Step 4. Now, we show that the expected cost function under the limit randomized policy is less than or equal to the expected cost achieved by $\limsup_{n \rightarrow \infty} J_T^n(\tilde{\underline{\tau}}_T^{*,n})$.

Since the cost function is continuous in states and actions, under the reduction (conditions (i) or (ii)), we have P -almost surely

$$(3.31) \quad \begin{aligned} & \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{T-1} c \left(\omega_0, x_t^i, u_t^i, \frac{1}{N} \sum_{p=1}^N u_t^p, \frac{1}{N} \sum_{p=1}^N x_t^p \right) \prod_{i=1}^N \prod_{t=0}^{T-1} \psi_t \left(y_t^i, \omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t-1}^i, u_{0:t-1}^i \right) \\ &= \frac{1}{N} \sum_{i=1}^N \bar{c} \left(\omega_0, \underline{\zeta}^i, \underline{u}^i, \frac{1}{N} \sum_{p=1}^N u^p, \frac{1}{N} \sum_{p=1}^N \Lambda(\underline{u}^p, \underline{\zeta}^p) \right) \prod_{i=1}^N \underline{\psi} \left(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}^i \right), \end{aligned}$$

where (3.31) is true following from (3.2) and Assumption 3.1 for some functions $\bar{c} : \Omega_0 \times \mathbf{S} \times \mathbf{U} \times \mathbf{U} \times \mathbf{X} \rightarrow \mathbb{R}_+$ continuous in states and actions and a function $\Lambda : \mathbf{U} \times \mathbf{S} \rightarrow \mathbf{X}$ continuous in actions and

$$\prod_{i=1}^N \underline{\psi} \left(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}^i \right) := \prod_{i=1}^N \prod_{t=0}^{T-1} \psi_t \left(y_t^i, \omega_0, x_0^i, \zeta_{0:t-1}^i, y_{0:t-1}^i, u_{0:t-1}^i \right).$$

We have

$$(3.32) \quad \begin{aligned} & \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{T-1} E_{\mathcal{L}_T^{*,1:N}} \left[c \left(\omega_0, x_t^i, u_t^i, \frac{1}{N} \sum_{p=1}^N u_t^p, \frac{1}{N} \sum_{p=1}^N x_t^p \right) \right] \\ & \geq \lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \int \int_{\mathcal{Z}} \min \left\{ M, \bar{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ & \quad \times Q_N(du, dy, d\zeta) \prod_{i=1}^{\infty} P_N^{*,\omega_0}(d\underline{u}_N^{i,*}, d\underline{y}^i, d\underline{\zeta}^i) \prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}_N^{i,*}) P(d\omega_0) \end{aligned}$$

$$(3.33) \quad \begin{aligned} & \geq \lim_{M \rightarrow \infty} \lim_{n \rightarrow \infty} \int \int_{\mathcal{Z}} \min \left\{ M, \bar{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_n(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_n(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ & \quad \times Q_n(du, dy, d\zeta) \prod_{i=1}^{\infty} P_n^{*,\omega_0}(d\underline{u}_n^{i,*}, d\underline{y}^i, d\underline{\zeta}^i) \prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}_n^{i,*}) P(d\omega_0) \end{aligned}$$

$$(3.34) \quad \begin{aligned} & = \lim_{M \rightarrow \infty} \int \lim_{n \rightarrow \infty} \int_{\mathcal{Z}} \min \left\{ M, \bar{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_n(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_n(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ & \quad \times Q_n(du, dy, d\zeta) \prod_{i=1}^{\infty} P_n^{*,\omega_0}(d\underline{u}_n^{i,*}, d\underline{y}^i, d\underline{\zeta}^i) \prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}_n^{i,*}) P(d\omega_0) \end{aligned}$$

$$(3.35) \quad \begin{aligned} & = \lim_{M \rightarrow \infty} \int \int_{\mathcal{Z}} \min \left\{ M, \bar{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ & \quad \times Q(du, dy, d\zeta) \prod_{i=1}^{\infty} P^{*,\omega_0}(d\underline{u}_\infty^{i,*}, d\underline{y}^i, d\underline{\zeta}^i) \prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}_\infty^{i,*}) P(d\omega_0) \end{aligned}$$

$$(3.36) \quad \begin{aligned} & = \int \int_{\mathcal{Z}} \bar{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta) \right) \\ & \quad \times Q(du, dy, d\zeta) \prod_{i=1}^{\infty} P^{*,\omega_0}(d\underline{u}_\infty^{i,*}, d\underline{y}^i, d\underline{\zeta}^i) \prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \underline{\zeta}^i, \underline{u}_\infty^{i,*}) P(d\omega_0), \end{aligned}$$

where (3.32) follows from the definition of empirical measures and by integrating over the set $(\prod_{i=n_t+1}^{\infty} \mathbf{Y} \times \mathbf{S})$ and since P -almost surely

$$\begin{aligned} & \min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right) \right\} \\ & \leq \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_N(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_N(du \times \mathbf{Y} \times d\zeta) \right). \end{aligned}$$

Inequality (3.33) is true since \limsup is the greatest convergent subsequential limit for a bounded sequence and (3.34) follows from the dominated convergence theorem. We note that $\{Q_n\}_n$ is induced by $\underline{u}_n^{i,*}$ for each DM. Since $\{Q_n\}_{n \in I}$ converges weakly to Q P -almost surely, by the moment condition and Remark 5, a similar argument as Step 3 of the proof of Theorem 3.2 implies that P -almost surely

$$\begin{aligned} \int_{\mathbf{U}} u Q_n(du \times \mathbf{Y} \times \mathbf{S}) & \xrightarrow{n \rightarrow \infty} \int_{\mathbf{U}} u Q(du \times \mathbf{Y} \times \mathbf{S}), \\ \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_n(du \times \mathbf{Y} \times d\zeta) & \xrightarrow{n \rightarrow \infty} \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q(du \times \mathbf{Y} \times d\zeta). \end{aligned}$$

Hence, (3.35) follows from [40, Theorem 3.5] since

$$\min \left\{ M, \tilde{c} \left(\omega_0, \zeta, u, \int_{\mathbf{U}} u Q_n(du \times \mathbf{Y} \times \mathbf{S}), \int_{\mathbf{U} \times \mathbf{S}} \Lambda Q_n(du \times \mathbf{Y} \times d\zeta) \right) \right\}$$

is bounded and nonnegative and continuously converges in u P -almost surely (see (3.23)). That is because, conditioned on $\omega_0, \underline{y}^i$ are i.i.d. random vectors (thanks to the symmetry), the space of policies is compact under the product topology (with the weak convergence topology for each coordinate (for each DM)), $\prod_{i=1}^{\infty} \underline{\psi}(\underline{y}^i, \omega_0, \zeta^i, \underline{u}_n^{i,*})$ converges in the product topology, and the cost function and $\underline{\psi}$ are continuous. Finally, (3.36) follows from the monotone convergence theorem. Hence, the proof is completed. \square

Remark 6. For the existence result, to show that the set of policies induced by independently randomized policies for each DM through time $t = 0, \dots, T - 1$ (see (3.29)) is closed under the weak convergence topology, we utilized the results in [47, section 5.2] which are more general than those in [16, 50]. We note that the extension of the existence results in [47, section 5.2] to our setup is not immediate since the conclusion of Step 3 cannot be established rigorously without considering the technical steps involving infinite dimensions and limit arguments.

4. Symmetric LQG dynamic teams. In the section, we consider the LQG setup where the results of sections 2 and 3 can be applied. We first consider N -DM LQG problems where we use Theorem 2.7 to show that the globally optimal policies are symmetric. Then, based on symmetry, we calculate N -DM optimal policies for such problems. Next, using Theorems 3.2 and 3.3, we show the convergence of N -DM optimal policies to optimal policies of LQG mean-field teams with countably infinite number of DMs. Finally, we consider infinite horizon problems where we use symmetry and convergence results to obtain global optimal policies for such problems.

4.1. Symmetric partially nested LQG dynamic teams on a graph. In the following, we consider decentralized problems where Theorem 2.7 can be utilized and the optimal policy can be obtained. First, we formulate LQG problems with a symmetric partially nested information structure. Consider the following dynamics. Let $i = 1, 2$, and

$$(4.1) \quad x_{t+1}^i = Ax_t^i + Bu_t^i + w_t^i.$$

Problem (\mathcal{P}_T). Consider the expected cost function of $(\underline{\gamma}_T^1, \underline{\gamma}_T^2)$ as

$$(4.2) \quad J_T(\underline{\gamma}_T^1, \underline{\gamma}_T^2) = E^{(\underline{\gamma}_T^1, \underline{\gamma}_T^2)} \left[\frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^2 (x_t^i)^T Q x_t^i + (u_t^i)^T R u_t^i + (u_t^1)^T \tilde{R} u_t^2 + (u_t^2)^T \tilde{R} u_t^1 \right],$$

where $\underline{\gamma}_T^i = (\gamma_{0:T-1}^i)$, and $R, \tilde{R} > 0$ and $Q \geq 0$. Let

$$(4.3) \quad y_t^i = H_t \zeta_{0:t-1}^i + \sum_{j=0}^{t-1} D_{tj} u_j^i,$$

where $\zeta_t^i = (w_t^i, v_t^i)$ with $\zeta_0^i = (x_0^i, w_0^i, v_0^i)$. Let $n, m, s \in \mathbb{N}$ and $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{R}^s$, $\mathbb{U} = \mathbb{R}^m$, $w_t^i \in \mathbb{R}^n$, $v_t^i \in \mathbb{R}^n$, $A \in \mathcal{M}_{n,n}$, $B \in \mathcal{M}_{n,m}$, $R \in \mathcal{M}_{m,m}$, $Q \in \mathcal{M}_{n,n}$, $\tilde{R} \in \mathcal{M}_{m,m}$, $H_t \in \mathcal{M}_{s,n(2t+1)}$, and $D_{tj} \in \mathcal{M}_{s,m}$. Let the information structure of each DM be $I_t^i = \{y_t^i, y_{\downarrow t}^i\}$.

In the following, we show that the above dynamic teams are symmetrically optimal under sufficient conditions on the observations and initial states.

COROLLARY 4.1. *For a fixed T , consider a finite horizon team problem defined above as (\mathcal{P}_T). If x_0^1 and x_0^2 are exchangeable zero mean Gaussian random vectors and w_t^i s and v_t^i s are i.i.d. Gaussian random vectors for $i = 1, 2$ and independent for all $t = 0, \dots, T-1$ and also independent of initial states, then the dynamic team is symmetrically optimal.*

Proof. Since the dynamic team is LQG with a partially nested information structure, a static reduction exists and the expected cost is convex in policies under static reductions (see [18] and Theorem 2.4(iii)). Assumption 2.1 is satisfied following from (4.2). We need to show assumptions of Theorem 2.7 hold. Following from the hypothesis on disturbances and initial states, assumption (b) holds. Assumption (c) holds following from Assumption 2.7 and since given (x_0^1, x_0^2) , $(y_{0:T-1}^1, y_{0:T-1}^2)$ are independent. Hence, Theorem 2.7 completes the proof. \square

Here, we consider a class of LQG dynamic teams with a tree information structure where we utilize Corollary 4.1 and we obtain an explicit recursion for the optimal policy.

Problem ($\mathcal{P}_T^{\text{tree}}$). Consider a finite horizon expected cost (4.2) with $I_t^i = \{x_{[0:t]}^i, u_{[0:t-1]}^i\}$.

We note that this problem is a special case of (\mathcal{P}_T) since we assumed that $y_t^i = x_t^i$ for all DM^{*i*} and $t = 0, \dots, T-1$ and also all DMs have a total recall ($I_t^i = \{x_{[0:t]}^i, u_{[0:t-1]}^i\}$). For this problem, we calculate an explicit recursion for optimal policies using the symmetry established in Corollary 4.1.

THEOREM 4.2. *For a fixed T , consider a finite horizon team problem defined as ($\mathcal{P}_T^{\text{tree}}$). If (x_0^1, x_0^2) are exchangeable with an identical zero mean Gaussian distribution and w_t^i s are i.i.d. zero mean Gaussian random vectors for $i = 1, 2$ and independent for all $t = 0, \dots, T-1$ and independent of initial states, then*

$$(4.4) \quad u_t^{1*,(T)} = K_t^{(T)} x_t^1 + L_t^{(T)} E[x_0^2 | x_0^1],$$

$$(4.5) \quad u_t^{2*,(T)} = K_t^{(T)} x_t^2 + L_t^{(T)} E[x_0^1 | x_0^2],$$

where

$$(4.6) \quad K_t^{(T)} = - \left(R + B^T P_{t+1}^{(T)} B \right)^{-1} B^T P_{t+1}^{(T)} A,$$

$$(4.7) \quad P_t^{(T)} = -A^T P_{t+1}^{(T)} B^T \left(R + B^T P_{t+1}^{(T)} B \right)^{-1} B^T P_{t+1}^{(T)} A + Q + A^T P_{t+1}^{(T)} A,$$

$$(4.8) \quad L_t^{(T)} = - \left(R + B^T P_{t+1}^{(T)} B \right)^{-1} \left[\tilde{R} K_t^{(T)} G_t^{(T)} + \tilde{R} L_t^{(T)} \Sigma \right. \\ \left. + \sum_{s=t+1}^{T-1} B^T (A^T)^{s-t} P_{s+1}^{(T)} B L_s^{(T)} \right], \\ G_t^{(T)} = \prod_{s=0}^{t-1} \left(A + B K_s^{(T)} \right) + \sum_{s=1}^t \prod_{j=s}^{t-1} \left(A + B K_j^{(T)} \right) B L_{s-1}^{(T)} \Sigma,$$

where $\Sigma = E[x_0^1(x_0^2)^T](E[x_0^2(x_0^2)^T])^{-1}$, $P_T^{(T)} = 0$, $G_0^{(T)} = I$. Moreover, the optimal cost is

$$(4.9) \quad J_T(\underline{\gamma}_T^*) = \frac{2}{T} \left\{ E \left[(x_0^1)^T P_0^{(T)} x_0^1 \right] + \sum_{t=0}^{T-1} E \left[(w_t^1)^T P_t^{(T)} w_t^1 \right] \right. \\ \left. + \sum_{t=0}^{T-1} E \left[\left(E[x_0^2 | x_0^1] \right)^T \left(\left(L_t^{(T)} \right)^T B^T P_{t+1}^{(T)} B L_t^{(T)} \right) E[x_0^2 | x_0^1] \right] \right. \\ \left. + \sum_{t=1}^{T-1} E \left[(x_0^1)^T (A^T)^t P_{t+1}^{(T)} B L_t^{(T)} E[x_0^2 | x_0^1] \right] \right\}.$$

Proof. Following from [18] and Radner's theorem [31], person-by-person optimality implies global optimality due to the uniqueness of the person-by-person optimal policy. That is because the information structure is partially nested, and LQG dynamic teams can be reduced to a static one using Ho and Chu's static reduction [18]. Hence, we only need to show that the policy satisfying (4.4) and (4.5) is person-by-person optimal. We show that for DM^1 , $J(\underline{\gamma}_T^*, \underline{\gamma}_T^*) \leq J((\underline{\gamma}_T^{-t*}, \beta), \underline{\gamma}_T^*)$ for all $\beta \in \Gamma^t$ where $(\underline{\gamma}_T^{-t*}, \beta) = (\gamma_{0:t-1}^*, \beta, \gamma_{t+1:T-1}^*)$. This implies that $(\underline{\gamma}_T^*, \underline{\gamma}_T^*)$ is person-by-person optimal thanks to Corollary 4.1 since the dynamic team is symmetrically optimal (by exchanging policies $(\underline{\gamma}_T^{-t*}, \beta)$ with $\underline{\gamma}_T^*$ which implies $J(\underline{\gamma}_T^*, \underline{\gamma}_T^*) \leq J(\underline{\gamma}_T^*, (\underline{\gamma}_T^{-t*}, \beta))$ for all $\beta \in \Gamma^t$ and this implies that $(\underline{\gamma}_T^*, \underline{\gamma}_T^*)$ is the fixed point of the equation). The proof is completed by induction. Due to space constraints, we have removed the calculation. \square

Remark 7. The optimal policies (4.4) and (4.5) contain two parts which can be interpreted as follows. The first part, $k_t^{(T)} x_t^i$, is equivalent to the optimal policy of the branch (DM) by ignoring the other branch in the optimization problem (in this case, this is equivalent to the centralized policies since the information structure of each branch (DM) is centralized). The second part corresponds to the correlation term between branches (DMs).

In the following, we generalize the result of Theorem 4.2 to N -DM LQG dynamic teams. Assume that the dynamics for $i = 1, 2, \dots, N$ are defined as in (4.1).

Problem ($\mathcal{P}_T^{N, \text{tree}}$). Consider the expected cost function of $\underline{\gamma}_T^{1:N}$ as

$$(4.10) \quad J_T^N(\underline{\gamma}_T^{1:N}) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N E^{\underline{\gamma}_T^{1:N}} \left[(x_t^i)^T Q x_t^i + (u_t^i)^T R u_t^i + \sum_{j=1, j \neq i}^N (u_t^i)^T \tilde{R} u_t^j + (u_t^j)^T \tilde{R} u_t^i \right],$$

where $\underline{\gamma}_T^i = \gamma_{0:T-1}^i$ for $i = 1, \dots, N$ and $R, \tilde{R} > 0$ and $Q \geq 0$. Let $I_t^i = \{x_{[0:t]}^i, u_{[0:t-1]}^i\}$.

COROLLARY 4.3. For a fixed T and N , consider a finite horizon team problem defined as $(\mathcal{P}_T^{N, \text{tree}})$. If $(x_0^{1:N})$ are exchangeable with an identical zero mean Gaussian

distribution, and w_t^i for $i = 1, \dots, N$ are i.i.d. zero mean Gaussian random vectors, independent for $t = 0, \dots, T - 1$, and independent of initial states, then

$$u_t^{i^*,(T),(N)} = K_t^{(T)} x_t^i + L_t^{(N),(T)} \sum_{j=1, j \neq i}^N E[x_0^j | x_0^i],$$

where $K_t^{(T)}$ and $P_t^{(T)}$ satisfy (4.6) and (4.7), and $L_t^{(N),(T)}$ is a function of $K_{0:t}^{(T)}$.

Proof. The proof is similar to the one of Theorem 4.2. □

Now we consider a more general setup where using Corollary 4.1, we establish a structural result for the case where the information structure of each DM over time satisfies a structure which is identical for all DMs and is partially nested. An example of such a graph structure is depicted in Figure 2.2.

Problem (\mathcal{P}_T^N). Consider a finite horizon expected cost of $\underline{\gamma}_T^{1:N}$ as (4.10) with the information structure $I_t^i = \{y_t^i, y_{\downarrow t}^i\}$, where y_t^i is defined in (4.3) and dynamics is defined in (4.1).

THEOREM 4.4. For a fixed T and N , consider a finite horizon team problem defined as (\mathcal{P}_T^N). If (x_0^1, \dots, x_0^N) are exchangeable with an identical zero mean Gaussian distribution and w_t^i for $i = 1, \dots, N$ are i.i.d. zero mean Gaussian random vectors, independent for all $t = 0, \dots, T - 1$ and independent of initial states, then

$$u_t^{i^*,(T),(N)} = K_t^{(T)} y_t^i + L_t^{(N),(T)} \sum_{j=1, j \neq i}^N E[x_0^j | y_t^i],$$

where $K_t^{(T)}$ are obtained by considering only one DM and ignoring other DMs.

Proof. The proof is similar to that of Theorem 4.2 by [18] and Corollary 4.1. □

Remark 8. A related work is [30], where structural results for optimal policy have been obtained for finite horizon LQG problems on graphs. In our analysis above, the structural result for the optimal policy is obtained without assuming that DMs who have no common ancestors and no common descendants either have uncorrelated noise or are decoupled through the cost function. Instead, exchangeable partially nested LQG teams with correlated initial states and disturbances are considered. Moreover, here, the graph structures may not be trees in general, as opposed to [30], where a multitree structure has been imposed on a graph.

Now, we present results for LQG teams with a mean-field coupling through the cost function. First, using Corollary 4.3, we obtain globally optimal policies for N -DM teams with a mean-field coupling and correlated initial states and disturbances. Next, as an implication of Theorem 3.2, we show the convergence of optimal policies for LQG N -DM mean-field teams on a tree to the corresponding optimal policy of mean-field teams. Let $I_t^i = \{x_{[0:t]}^i, u_{[0:t-1]}^i\}$ for $i \in \mathbb{N}$, and dynamics be as (4.1).

Problem ($\mathcal{P}_{T,\text{LQG}}^{N,\text{MF}}$). Consider the expected cost function of $\underline{\gamma}_T^N$ as

$$\begin{aligned} (4.11) \quad J_T^N(\underline{\gamma}_T^N) &= \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^N E^{\underline{\gamma}_T^{1:N}} \left[(x_t^i)^T Q x_t^i + (u_t^i)^T R u_t^i \right. \\ &\quad + \frac{1}{N-1} \sum_{j=1, j \neq i}^N (u_t^i)^T \tilde{R} u_t^j + (u_t^j)^T \tilde{R} u_t^i \\ &\quad \left. + \frac{1}{N-1} \sum_{j=1, j \neq i}^N (x_t^i)^T \tilde{Q} x_t^j + (x_t^j)^T \tilde{Q} x_t^i \right], \end{aligned}$$

where $R, \tilde{R} > 0$ and $Q, \tilde{Q} \geq 0$.

Problem ($\mathcal{P}_{T,LQG}^{\infty, MF}$). Consider the expected cost function of $\underline{\gamma}_T$ as

$$(4.12) \quad J_T^\infty(\underline{\gamma}_T) = \limsup_{N \rightarrow \infty} J_T^N(\underline{\gamma}_T).$$

COROLLARY 4.5. For a fixed T and N , consider a finite horizon team problem defined as ($\mathcal{P}_{T,LQG}^{N, MF}$). If $(x_0^{1:N})$ are exchangeable zero mean Gaussian random vectors, and w_t^i 's are i.i.d. zero mean Gaussian random vectors for $i = 1, \dots, N$, independent for $t = 0, \dots, T-1$, and independent of initial states, then

$$(4.13) \quad u_t^{i*,(T),(N)} = K_t^{(T)} x_t^i + \frac{L_t^{(N),(T)}}{N-1} \sum_{j=1, j \neq i}^N E[x_0^j | x_0^i],$$

where $K_t^{(T)}$ and $P_t^{(T)}$ satisfy (4.6) and (4.7), and $L_t^{(N),(T)}$ is a function of $K_{0:t}^{(T)}$.

Proof. The proof is similar to the one in Theorem 4.2. \square

COROLLARY 4.6. For a fixed T , consider a finite horizon team problem defined as ($\mathcal{P}_{T,LQG}^{\infty, MF}$). Assume $\{x_0^i\}_{i \in \mathbb{N}}$ are exchangeable random vectors with zero mean Gaussian distribution, and w_t^i 's are i.i.d. zero mean Gaussian random vectors for $i \in \mathbb{N}$, independent for $t = 0, \dots, T-1$, and independent of initial states. If $L_t^{(N),(T)}$ in (4.13) converges pointwise as $N \rightarrow \infty$ to $L_t^{(\infty),(T)}$, then

$$u_t^{i*,(T),(\infty)} = K_t^{(T)} x_t^i + L_t^{(\infty),(T)} \Sigma x_0^i,$$

where $K_t^{(T)}$ and $P_t^{(T)}$ satisfy (4.6) and (4.7), and $\Sigma = E[x_0^1(x_0^2)^T](E[x_0^2(x_0^2)^T])^{-1}$.

Proof. Following from [1, p. 9], since $\{x_0^i\}_{i \in \mathbb{N}}$ are exchangeable Gaussian random vectors, we can describe them explicitly as $x_0^i = \omega_0 + \theta^i$, where $(\theta^1, \theta^2, \dots)$ are i.i.d. mean zero Gaussian and independent of mean zero Gaussian random vector ω_0 . Now, we invoke Theorem 3.2 (or Theorem 3.3) and Corollary 4.5 and we use Remark 5 to complete the proof. \square

4.2. Average cost infinite horizons problems for partially nested dynamic teams. In the following, we consider average cost problems with a symmetric partially nested information structure. We note that the optimality of linear policies for infinite horizon LQG problems is an open problem in its generality. In this subsection, we provide a positive result for a class of such problems.

Now, we consider an infinite horizon team problem and we use the result in section 4.1.

Problem ($\mathcal{P}_\infty^{\text{tree}}$). Consider the expected cost function of $(\underline{\gamma}^1, \underline{\gamma}^2)$ as

$$(4.14) \quad J(\underline{\gamma}^{1:2}) = \limsup_{T \rightarrow \infty} E^{(\underline{\gamma}^1, \underline{\gamma}^2)} [c(x_{0:T-1}^{1:2}, u_{0:T-1}^{1:2})],$$

where the cost function is defined as (4.2) and $I_t^i = \{x_{[0:t]}^i, u_{[0:t-1]}^i\}$.

First, we introduce a lemma essential for Theorem 4.8.

LEMMA 4.7. Consider the sequence $\{a_T^i\}_{i=1}^T$. Assume $\lim_{T \rightarrow \infty} a_T^i = a$ for $i = 0, \dots, T-1$. If for every fixed $T \in \mathbb{N}$, $a_T^l = a_{T+1}^{l+1}$ for all $l = 0, \dots, T-1$, then $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T a_T^i = a$.

Proof. We have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T a_T^i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=0}^{T-1} a_{T-i}^1 = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=1}^T a_k^1 = a,$$

where the second equality follows from $a_T^l = a_{T+1}^{l+1}$ and the last equality follows from the Cesàro mean argument. \square

THEOREM 4.8. *Consider average cost infinite horizon team problems defined as $(\mathcal{P}_\infty^{tree})$. Assume (A, B) are stabilizable and $(A, Q^{\frac{1}{2}})$ are detectable. Assume x_0^1 and x_0^2 are exchangeable with an identical zero mean Gaussian distribution and w_t^i s are i.i.d. zero mean Gaussian random variables for $i = 1, 2$ and for all $t = 0, \dots, T - 1$ and independent of initial states. If $L_t^{(T)}$ in (4.8) converges pointwise to $L_t^{(\infty)}$ as $T \rightarrow \infty$, then the pointwise limit of the sequence of optimal policies for (\mathcal{P}_T^{tree}) is team optimal for $(\mathcal{P}_\infty^{tree})$ and stabilizes the closed-loop system,*

$$\begin{aligned} u_t^{1*,(\infty)} &= Kx_t^1 + L_t^{(\infty)} E[x_0^2|x_0^1], \\ u_t^{2*,(\infty)} &= Kx_t^2 + L_t^{(\infty)} E[x_0^1|x_0^2], \end{aligned}$$

where $K, P, L_t^{(\infty)}$ and $G_t^{(\infty)}$ are the pointwise limit of the ones for (\mathcal{P}_T^{tree}) as $T \rightarrow \infty$.

Proof. We show $\limsup_{T \rightarrow \infty} J_T(\underline{\gamma}_T^*) = J(\underline{\gamma}_\infty^*)$ and invoke [38, Theorem 5] or [29, Theorem 1] to complete the proof. From (4.9), we have

$$\lim_{T \rightarrow \infty} |J_T(\underline{\gamma}_\infty^*) - J_T(\underline{\gamma}_T^*)| \tag{4.15}$$

$$= \lim_{T \rightarrow \infty} \frac{2}{T} \left| E \left[(x_0^1)^T (P_0^{(T)} - P) x_0^1 \right] + \frac{2}{T} \sum_{t=0}^{T-1} E \left[(w_t^1)^T (P_t^{(T)} - P) w_t^1 \right] \right| \tag{4.16}$$

$$+ \frac{2}{T} \sum_{t=0}^{T-1} E \left[\left(E[x_0^2|x_0^1] \right)^T \left((L_t^{(T)})^T B^T P_{t+1}^{(T)} B L_t^{(T)} - (L_t^{(\infty)})^T B^T P B L_t^{(\infty)} \right) E[x_0^2|x_0^1] \right] \tag{4.17}$$

$$+ \frac{2}{T} \sum_{t=0}^{T-1} E \left[(x_0^1)^T (A^T)^{t-1} \left(P_{t+1}^{(T)} B L_t^{(T)} - P B L_t^{(\infty)} \right) E[x_0^2|x_0^1] \right] \tag{4.18}$$

$= 0,$

where (4.15) is zero since $P_0^{(T)}$ converges to P using Lemma 4.7 since $P_{t+1}^{(T+1)} = P_t^{(T)}$. Expression (4.17) converges to zero since $L_t^{(T)}$ in (4.8) converges pointwise to $L_t^{(\infty)}$ as $T \rightarrow \infty$, we have $\sum_{s=t+1}^\infty B^T (A^T)^{s-t} P B L_s^{(\infty)} < \infty$, and this implies that $\lim_{s \rightarrow \infty} L_s^{(\infty)} = 0$. Hence, we have that for every $\epsilon > 0$, there exists $\hat{N} > T$ such that for every $t > \hat{N}$, $|\text{Tr}[L_t^{(\infty)}(L_t^{(\infty)})^T]| < \epsilon$. We define $L_t^{(T)} = 0$ for $t > T$. Expression (4.16) is equal to zero following from Lemma 4.7 and the fact that $|\text{Tr}[L_t^{(\infty)}(L_t^{(\infty)})^T]| < \epsilon$ for every $t > \hat{N}$. Hence, equality (4.18) is true and global optimality follows from [38, Theorem 5]. The closed loop system is stable since $\limsup_{t \rightarrow \infty} E(\|x_t^1\|^2) < \infty$ following from $\|A + BK\| < 1$ (all the eigenvalues of $A + BK$ are inside of the unit circle) and since $\|L_t^{(\infty)}\|$ is uniformly bounded. \square

5. Conclusion. In this paper, we studied dynamic teams with symmetric information structures. We presented a characterization for symmetrically optimal teams for convex exchangeable team problems. For mean-field teams with symmetric information structures, we showed the convergence of optimal policies for mean-field teams with N DMs to the corresponding optimal policy of mean-field teams. We obtained

globally optimal solutions for LQG dynamic team problems with symmetric partially nested information structures. Moreover, we obtained globally optimal policies for average cost infinite horizon problems of LQG dynamic teams.

In this paper since we worked under the convexity assumption, the information structure does not allow for the mean-field coupling in the dynamics. In our recent work [35], we relaxed the convexity assumption to arrive at complementary existence and structural results.

REFERENCES

- [1] D. J. ALDOUS, I. A. IBRAGIMOV, AND J. JACOD, *Ecole d'Ete de Probabilites de Saint-Flour XIII*, 1983, Lectures Notes in Math. 1117, Springer, Berlin, 1985.
- [2] J. ARABNEYDI AND A. AGHDAM, *A certainty equivalence result in team-optimal control of mean-field coupled markov chains*, in Proceedings of the IEEE 56th Annual Conference on Decision and Control, 2017, pp. 3125–3130.
- [3] J. ARABNEYDI AND A. MAHAJAN, *Team-optimal solution of finite number of mean-field coupled LQG subsystems*, in Proceedings of the IEEE 54th Annual Conference on Decision and Control, 2015, pp. 5308–5313.
- [4] A. ARAPOSTATHIS, A. BISWAS, AND J. CARROLL, *On solutions of mean field games with ergodic cost*, J. Math. Pures Appl., 107 (2017), pp. 205–251.
- [5] M. BARDI AND M. FISCHER, *On non-uniqueness and uniqueness of solutions in finite-horizon mean field games*, ESAIM Control Optim. Calc. Var., 25 (2019).
- [6] M. BARDI AND F. S. PRIULI, *Linear-quadratic N-person and mean-field games with ergodic cost*, SIAM J. Control Optim., 52 (2014), pp. 3022–3052.
- [7] E. BAYRAKTAR AND X. ZHANG, *On non-uniqueness in mean field games*, Proc. Amer. Math. Soc., 148 (2020), pp. 4091–4106.
- [8] P. BILLINGSLEY, *Convergence of Probability Measures*, Wiley, New York, 1968.
- [9] D. BLACKWELL, *Memoryless strategies in finite-stage dynamic programming*, Ann. Math. Statist., 35 (1964), pp. 863–865.
- [10] V. S. BORKAR, *White-noise representations in stochastic realization theory*, SIAM J. Control Optim., 31 (1993), pp. 1093–1102.
- [11] P. CAINES, M. HUANG, AND R. MALHAMÉ, *Mean field games*, in Handbook of Dynamic Game Theory, Springer, New York, 2017, pp. 345–372.
- [12] P. CARDALIAGUET AND C. RAINER, *An example of multiple mean field limits in ergodic differential games*, NoDEA Nonlinear Differential Equations Appl., 27 (2020), pp. 1–19.
- [13] R. CARMONA AND F. DELARUE, *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer, New York, 2018.
- [14] R. CARMONA, F. DELARUE, AND D. LACKER, *Mean field games with common noise*, Ann. Probab., 44 (2016), pp. 3740–3803.
- [15] M. FISCHER, *On the connection between symmetric N-player games and mean field games*, Ann. Appl. Probab., 27 (2017), pp. 757–810.
- [16] 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., 53 (2015), pp. 1681–1712.
- [17] B. HAJEK AND M. LIVESAY, *On non-unique solutions in mean field games*, in Proceedings of the IEEE 58th Annual Conference on Decision and Control, 2019, pp. 1219–1224.
- [18] Y. C. HO AND K. C. CHU, *Team decision theory and information structures in optimal control problems—part I*, IEEE Trans. Automat. Control, 17 (1972), pp. 15–22.
- [19] Y. C. HO AND K. C. CHU, *On the equivalence of information structures in static and dynamic teams*, IEEE Trans. Automat. Control, 18 (1973), pp. 187–188.
- [20] M. HUANG, P. E. CAINES, AND R. P. MALHAMÉ, *Social optima in mean field LQG control: Centralized and decentralized strategies*, IEEE Trans. Automat. Control, 57 (2012), pp. 1736–1751.
- [21] M. HUANG, R. P. MALHAMÉ, AND P. E. CAINES, *Large population stochastic dynamic games: Closed-loop McLean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–252.
- [22] D. LACKER, *Mean field games via controlled martingale problems: Existence of Markovian equilibria*, Stochastic Process. Appl., 125 (2015), pp. 2856–2894.
- [23] D. LACKER, *A general characterization of the mean field limit for stochastic differential games*, Probab. Theory Related Fields, 165 (2016), pp. 581–648.
- [24] D. LACKER, *Limit theory for controlled McKean-Vlasov dynamics*, SIAM J. Control Optim., 55 (2017), pp. 1641–1672.
- [25] D. LACKER, *On the convergence of closed-loop Nash equilibria to the mean field game limit*, Ann. Appl. Probab., 30 (2020), pp. 1693–1761.

- [26] J. M. LASRY AND P. L. LIONS, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.
- [27] L. LESSARD AND S. LALL, *Optimal control of two-player systems with output feedback.*, IEEE Trans. Automat. Control, 60 (2015), pp. 2129–2144.
- [28] B. LIGHT AND G. Y. WEINTRAUB, *Mean Field Equilibrium: Uniqueness, Existence, and Comparative Statics*, Stanford University, 2018.
- [29] A. MAHAJAN, N. C. MARTINS, AND S. YÜKSEL, *Static LQG teams with countably infinite players*, in Proceedings of the IEEE 52nd Annual Conference on Decision and Control, 2013, pp. 6765–6770.
- [30] A. NAYYAR AND L. LESSARD, *Optimal Control for LQG Systems on Graphs—Part I: Structural Results*, preprint, arXiv:1408.2551, 2014.
- [31] R. RADNER, *Team decision problems*, Ann. Math. Statist., 33 (1962), pp. 857–881.
- [32] M. ROTKOWITZ, *On information structures, convexity, and linear optimality*, in Proceedings of the IEEE 47th Annual Conference on Decision and Control, 2008, pp. 1642–1647.
- [33] M. ROTKOWITZ AND S. LALL, *A characterization of convex problems in decentralized control*, IEEE Trans. Automat. Control, 51 (2006), pp. 274–286.
- [34] N. SALDI, *A topology for team policies and existence of optimal team policies in stochastic team theory*, IEEE Trans. Automat. Control, 65 (2019), pp. 310–317.
- [35] S. SANJARI, N. SALDI, AND S. YÜKSEL, *Optimality of Independently Randomized Symmetric Policies for Exchangeable Stochastic Teams with Infinitely Many Decision Makers*, preprint, arXiv:2008.11570, (2020).
- [36] S. SANJARI AND S. YÜKSEL, *Optimal stochastic teams with infinitely many decision makers and mean-field teams*, in Proceedings of the IEEE 57th Annual Conference on Decision and Control, 2018, pp. 4123–4128.
- [37] S. SANJARI AND S. YÜKSEL, *Convex symmetric stochastic dynamic teams and their mean-field limit*, in Proceedings of the IEEE 58th Annual Conference on Decision and Control, 2019, pp. 4662–4667.
- [38] S. SANJARI AND S. YÜKSEL, *Optimal solutions to infinite-player stochastic teams and mean-field teams*, IEEE Trans. Automat. Control, <https://doi.org/10.1109/TAC.2020.2994899>.
- [39] M. SCHÄL, *Conditions for optimality in dynamic programming and for the limit of n -stage optimal policies to be optimal*, Z. Wahrscheinlichkeitsth, 32 (1975), pp. 179–296.
- [40] R. SERFOZO, *Convergence of Lebesgue integrals with varying measures*, Sankhyā A, 44 (1982), pp. 380–402.
- [41] P. SHAH AND P. PARRILO, *H2 optimal decentralized control over posets: A state-space solution for state-feedback*, IEEE Trans. Automat. Control, 58 (2013), pp. 3084–3096.
- [42] J. SWIGART AND S. LALL, *Optimal controller synthesis for decentralized systems over graphs via spectral factorization*, IEEE Trans. Automat. Control, 59 (2014), pp. 2311–2323.
- [43] D. B. WEST, *Introduction to Graph Theory*, Vol. 2, Prentice, Hall, Upper Saddle River, NJ, 1996.
- [44] H. S. WITSENHAUSEN, *A counterexample in stochastic optimal control*, SIAM J. Control Optim., 6 (1968), pp. 131–147.
- [45] H. S. WITSENHAUSEN, *The intrinsic model for discrete stochastic control: Some open problems*, Lecture Notes in Econom. and Math. Systems 107, Springer, New York, 1975, pp. 322–335.
- [46] H. S. WITSENHAUSEN, *Equivalent stochastic control problems*, Math. Control Signals Systems, 1 (1988), pp. 3–11.
- [47] S. YÜKSEL, *A universal dynamic program and refined existence results for decentralized stochastic control*, SIAM J. Control Optim., 58 (2020), pp. 2711–2739.
- [48] S. YÜKSEL, *On stochastic stability of a class of non-Markovian processes and applications in quantization*, SIAM J. Control Optim., 55 (2017), pp. 1241–1260.
- [49] S. YÜKSEL AND T. BAŞAR, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*, Springer, New York, 2013.
- [50] S. YÜKSEL AND N. SALDI, *Convex analysis in decentralized stochastic control, strategic measures and optimal solutions*, SIAM J. Control Optim., 55 (2017), pp. 1–28.