On Convex Stochastic Dynamic Teams with Infinite Horizons and Decision Makers

Sina Sanjari and Serdar Yüksel

Abstract—This paper studies convex dynamic team problems with finite and infinite time horizons. 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. While for partially nested LQG team problems with finite time horizon 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 and typically not only linearity but also time-invariance and stability properties are imposed a priori in the literature. In this paper, average cost finite and infinite horizon dynamic team problems with a symmetric partially nested information structure are studied and globally optimal solutions are obtained where we establish linearity of optimal policies. In addition, for convex mean field teams with a symmetric information structure, through concentration of measure arguments, we show the convergence of optimal policies for mean field teams with $N$ decision makers to the corresponding optimal policies of mean field teams. Moreover, we discuss average cost infinite horizon state feedback problems for LQG dynamic teams with sparsity and delay constraints.

Index Terms—Stochastic teams, average cost optimization, decentralized control, mean field teams

I. INTRODUCTION

Team problems consist of a collection of decision makers 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 (IS) of the problem. If there is a pre-defined order in which the decision makers 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 categorized as classical, partially nested, and non-classical. An IS is classical if the information of decision maker $i$ ($\text{DM}_i$) contains all of the information available to $\text{DM}_k$ for $k < i$. An IS is partially nested, if whenever the action of $\text{DM}_k$, for some $k < i$, affects the information of $\text{DM}_i$, then the information of $\text{DM}_i$ contains the information of $\text{DM}_k$. An IS which is not partially nested is non-classical. Dynamic teams with partially nested information structures include those with delayed sharing patterns, have been introduced in [1], [2]. In [2], it has been shown that for teams with finite number of DMs, under an invertibility condition, dynamic teams with a partially nested information structure can be reduced to a static one (static reduction [2], [3]) where Radner’s theorem concludes the global optimality of a linear policy for LQG team problems [4]. Hence, the linearity of the optimal policy has been established for finite horizon LQG partially nested teams. However, for average cost infinite horizon, partially nested, LQG dynamic team problems so far there has been no universal result showing that a globally optimal policy is linear, time-invariant, and stabilizing.

In [5], the problem of designing a linear, time-invariant, stabilizing, state feedback optimal policy for decentralized $\mathcal{H}_2$-optimization problems, which satisfy the quadratically invariant property, has been addressed by reparametrizing the problem as a convex problem (via Youla parametrization). In [6], 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 set (POSET), causal 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 [7]. Using the spectral factorization approach, a similar result has been established in [8] where linearity and time invariance has been imposed a priori. In [9], $\mathcal{H}_2$-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 [9], [5], [7], [8] are inconclusive regarding global optimality. In this paper, we search for globally optimal solutions and we do not restrict the search for optimal policies over those that are linear, time-invariant, and stabilizing, hence, we justify global optimality for a class of problems.

Average cost, infinite horizon, team problems with the one step delay sharing pattern have not been fully addressed in the literature, despite the presence of results where finiteness of the state space and actions and additional technical assumptions on the information structure and the cost function have been imposed [10]. In addition, in [11], under the assumption that initial states and disturbances are independent, the state feedback optimal policy has been obtained for finite horizon LQG team problems with a partially nested information structure admitting delay and sparsity constraints [11]. This approach assumes fully observed state feedback with independent initial states and disturbances. The result has been extended to average cost infinite horizon LQG problems where under some technical assumptions, the convergence of the solution to the solution of the corresponding algebraic
Riccati equations [11] has been shown; however, the global optimality of the limit solution has not been established. In this paper, we establish the global optimality.

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

(i) We define the notion of exchangeable teams and symmetric information structures, and we show that, for convex exchangeable dynamic teams, optimal policies exhibit a symmetry structure. This symmetry structure is more relaxed when compared with the symmetry results developed in [12], [13] and is applicable for dynamic teams which may not admit a static reduction, as long as convexity on policies holds for the team problem. For LQG dynamic teams with a symmetric partially nested information pattern, we obtain an optimal policy for finite horizon problems.

(ii) Building on the result above, we obtain a globally optimal policy for average cost team problems with a symmetric partially nested information structure on a tree. 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.

(iii) For convex mean field teams with a symmetric information structure, through concentration of measure arguments, we show the convergence of optimal polices for mean field teams with $N$ decision makers to the corresponding optimal policies for mean field teams.

(iv) For LQG dynamic teams with the one step delayed information structure, we show that under observability and controllability conditions, a globally optimal policy for the average cost infinite horizon dynamic problems can be obtained as the pointwise limit of the sequence of optimal policies for finite horizon problems as $T \rightarrow \infty$.

Hence, we establish linearity of an optimal policy.

The organization of the paper is as follows: we obtain globally optimal solutions for finite horizon problems with a symmetric partially nested information structure in Section III. We study mean field teams in Section IV and we discuss average cost LQG team problems with a symmetric information structure and with the one step delayed sharing pattern and sparsity in Section V-A and Section V-B, respectively.

**II. Preliminaries**

We first introduce preliminaries following the presentation in [14], in particular, we introduce the characterizations laid out by Witsenhausen, through his Intrinsic Model [15]. 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 $\{(\Omega, \mathcal{F}), (U^i, U^i), (Y^i, Y^i), i \in \mathcal{N}\}$, specifying the system’s distinguishable events, and control and measurement spaces. In this model (described in discrete time), any action applied at any given time is regarded as applied by an individual decision maker (DM), who acts only once. The pair $(\Omega, \mathcal{F})$ is a measurable space (on which an underlying probability may be defined). The pair $(U^i, U^i)$ denotes the measurable space from which the action, $u^i$, of DM $i$ is selected. The pair $(Y^i, Y^i)$ denotes the measurable observation/measurement space.
- There is a measurement constraint to establish the connection between the observation variables and the system’s distinguishable events. The $Y^i$-valued observation variables are given by $y^i = h^i(\omega, y^{[i, i-1]})$, where $y^{[i, i-1]} = \{u^k, k \leq i \}$, $h^i$’s are measurable functions and $u^k$ denotes the action of $D M^k$. Hence, $y^i$ induces $\sigma(y^i)$ over $\Omega \times \prod_{k=1}^{i-1} U^k$.
- The set of admissible control laws $\gamma_i = \{\gamma^i\}_{i \in \mathcal{N}}$, also called designs or policies, are measurable control functions, so that $u^i = \gamma^i(y^i)$. Let $\Gamma^i$ 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 $(\Omega, \mathcal{F})$ describing the probability space on which the system is defined.

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

Notations: $\mathbb{R}$ and $\mathbb{N}$ denote the set of real numbers and natural numbers, respectively. We denote trace of a matrix $A$ as $Tr(A)$. We denote that a random vector $X$ is independent of a random vector $Y$ by $X \perp Y$. We denote $A^T$ as the transpose of a matrix $A$ and $A(T)$ to show the dependence of a matrix $A$ to $T \in \mathbb{N}$.

**III. Finite horizon Convex Dynamic team problems with a Symmetric Information structure**

In this section, we obtain explicit recursions for an optimal policy for finite horizon convex dynamic teams satisfying a symmetric information structure. To this end, we first define the notion of a symmetric information structure on a graph. Consider $G(V, \mu)$ as a directed information graph with $V = \{1, \ldots , n\}$ 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 $j$. $i \rightarrow j$ means $u^i$ affects $y^j$ through the relation $y^j = h^j(\omega, y^{[i, i-1]})$ defined in the intrinsic model (see Section II). We denote by $\downarrow j$ as the set of nodes $i$ such that $i \rightarrow j$ (ancestors), and $\uparrow j = \{ \downarrow j \} \cup \{ j\}$. Similarly, we can define descendents by $\uparrow j$. In the following, we define concepts required for our main theorem.

**Definition 1.** The information structure of a sequential team problem is symmetric if

(i) there exists a node $\{i\}$ and finite number of sub-graphs $G_p(\tilde{V}, \tilde{\mu})$ such that $\bigcup_{p=1}^N G_p \cup \{ i\} = G$ and $G_p$ is isomorphic (see e.g., [16]) for all $p = 1, \ldots , N$, i.e., for every node with directed edges in each sub-graph there exists a unique node with identical directed edges in the corresponding sub-graphs, where $\tilde{V} = \{0, \ldots , T - 1\}$, and $G_p^k$ refers to a node $k$ in $G_p$ for all $p = 1, \ldots , N$ and $k = 0, \ldots , T - 1$, and

(ii) sharing of the information between sub-graphs is symmetric, i.e., for $p, s = 1, \ldots , N$, and $k, j = 0, \ldots , T - 1$, and
for every edge from a node \( G_k^+ \) to a node \( G_k^- \), there exists an edge from a node \( G_k^- \) to nodes \( G^j_{p-1} \), where \( G^j_{p-1} \) denotes \( \{G^1_{p-1}, \ldots, G^j_{p-1}, \ldots, G^N_{p-1}\} \), and also there exist edges from nodes \( G_{-k}^- \) to a node \( G_{-k}^+ \).

Two examples are shown in Fig. 1, and Fig. 2. First, we recall the definition of exchangeable finite set of random variables.

**Definition 2.** Random variables (vectors) \( x^1, x^2, \ldots, x^N \) are exchangeable if any permutation, \( \sigma \), of the set \( \{1, \ldots, N\} \) does not change the joint probability measures of random variables (vectors), i.e.,

\[
P(d(x^1), d(x^2), \ldots, d(x^N)) = P(dx^1, dx^2, \ldots, dx^N).
\]

**Assumption 1.** Assume

(i) the information of each decision maker is described as \( I_i = \{y^i_0, u^i_0, y^i_{1:t}, y^i_{1:t}\} \), where \( u^i_{1:t} \) corresponds to the actions of DM\(i\)'s (includes DM\(i\) itself) at time instances \( p \) where the action of DM\(i\)s at time \( p \) affects the observation of DM\(i\) at time \( t \). For \( i = 1, \ldots, N \), \( y^i_t = h_i(\zeta^i, u^i_{0:t-1}, y^i_{0:t-1}) \) represents the observation of DM\(i\) at time \( t \), and \( u^i_{0:t-1} = (u^i_{0:t-1}, \ldots, u^i_{0:t-1}, y^i_{0:t-1}, \ldots, y^i_{0:t-1}) \), \( (ht) \) is a measurable function independent of \( i \), and we denote \( \zeta^i \) to represent all the uncertainty of the sub-graph \( i \),

(ii) observation spaces, action spaces are standard Borel spaces and are identical for each DM \( Y = \prod_{k=0}^{T-1} \mathcal{Y}^k, \mathcal{U} = \prod_{k=0}^{T-1} \mathcal{U}^k \), respectively. The sets of all admissible policies are denoted by \( \Pi = \prod_{i=1}^N \prod_{t=0}^{T-1} \Pi^t \),

(iii) node \( i \) in Definition 1 represents initial states of decision makers \( \omega_i = (x^i_0, \ldots, x^i_0) \) and each sub-graph represents a single DM acting at time instances \( t = 0, \ldots, T - 1 \),

(iv) the expected cost function is defined as

\[
J_T(\gamma^1_T) = \mathbb{E} \sum_{t=0}^{T-1} [c(\xi^{1:N}, \gamma^{1:N})],
\]

**Fig. 1:** A tree structure of a symmetric dynamic team.

**Fig. 2:** An example of the graph structure of a symmetric dynamic team.

for some Borel measurable cost function \( c : \Omega \times \prod_{i=1}^N \mathcal{X} \to \mathbb{R} \), where \( \gamma^1_N = (\gamma^1_T, \gamma^2_T, \ldots, \gamma^N_T) \) and \( \gamma^1_T = \gamma^1_{0:T-1} \) for \( i = 1, \ldots, N \). For any permutation \( \sigma \in \Sigma \) of the set \( \{1, \ldots, N\} \), where \( \Sigma \) is the set of all possible permutations, we have \( P \)-almost surely,

\[
c(\xi^{1:N}, (\omega^i)^{1:N}) = c(\xi^{1:N}, (\omega^i)^{1:N}). \quad (1)
\]

In the following, we define notions of exchangeable and symmetrically optimal teams analogous to [12], [13] for dynamic teams.

**Definition 3.** (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., \( J_T(\gamma^1_T, \gamma^2_T, \ldots, \gamma^N_T) = J_T(\gamma^1_T, \gamma^2_T, \ldots, \gamma^N_T) \).

**Definition 4.** (Symmetrically optimal teams)

A team is symmetrically optimal if there exists a team optimal policy where each DM has the identical policy, i.e., \( \gamma^1_T = (\gamma^2_T, \ldots, \gamma^N_T) \), and \( \gamma^1_T = \gamma^2_T \) for all \( i, j = 1, \ldots, N \), where \( \gamma^1_T \) is an team optimal policy.

**Remark 1.** The concepts of exchangeable and symmetrically optimal dynamic teams in this paper are generalizations of those for static teams in [12], [13]. Here, the value of the cost function may not be invariant under exchanging \( \gamma^i_k \) with \( \gamma^k_j \) for \( k \neq j, k, t = 0, \ldots, T - 1 \), and for \( i, j = 1, \ldots, N \). Hence, the concept of symmetrically optimal teams in this paper does not necessarily imply that an optimal policy is identical through time, i.e., \( \gamma^i_T \) may not be identical to \( \gamma^j_T \) for \( k \neq j, k, t = 0, \ldots, T - 1 \), and for \( i = 1, \ldots, N \).

Here, we give a characterization for exchangeable and symmetrically optimal dynamic teams. In the following theorem, the cost function may not be quadratic and the system may not be linear Gaussian.

**Theorem 1.** Consider a team problem with a symmetric partially nested information structure (see Definition 1) under Assumption (1) if

(a) action spaces are convex for each DM,

(b) \( \xi = (\xi^1, \ldots, \xi^N) \) are exchangeable,

(c) for all policies \( \gamma \in \Pi \),

\[
\prod_{t=1}^{T-1} \mathbb{P}(dy^{1:N} | y^{1:j|1:N}, \gamma^{1:j|1:N}) (y^{1:j|1:N}) \mathbb{P}(dy^{j|0} | y^{j|0}) = \prod_{t=1}^{T-1} \mathbb{P}(dy^{1:N} | y^{1:j|1:N}, \gamma^{1:j|1:N}) (y^{1:j|1:N}) \mathbb{P}(dy^{j|0} | y^{j|0}), \quad (2)
\]

where \( \downarrow \downarrow (1:N) \) denotes \( (\downarrow \downarrow 1, \ldots, \downarrow \downarrow N) \) (see Corollary (2) for sufficient conditions of (3)), then the team is exchangeable. Moreover, if the team problem is convex in policies (see [14, Section 3.3]), i.e., for all \( \sum_{i=1}^N \alpha_i = 1 \) and \( \alpha_i \in (0, 1) \),

\[
J_T \left( \sum_{i=1}^N \alpha_i \gamma^i_T \right) \leq \sum_{i=1}^N \alpha_i J_T \left( \gamma^i_T \right),
\]

then the team is symmetrically optimal.

**Proof.** We first show that for any permutation \( \sigma \in \Sigma \) of policies \( J_T((\sigma^1_T)^{1:N}, \ldots, (\sigma^N_T)^{1:N}) = J_T((1_T)^{1:N}, \ldots, (N_T)^{1:N}), \) i.e., a
team is exchangeable. We have,

\[
J_T \left((\gamma_1^{\sigma_1}, \ldots, \gamma_T^{\sigma_T})^N\right) \\
= \int c \left(\xi_{1:N}, (\gamma_1^{\sigma_1})^{(1)}(\gamma_T^{\sigma_T})^{N} \right) \mathbb{P}(d\xi) \\
\times \prod_{i=1}^{T-1} \prod_{N} \mathbb{P}\left(d\gamma_{i+1}^{\sigma_{i+1}, \sigma_{i+1}} \mid (\gamma_{i+1}^{\sigma_{i+1}})^{i}, (\gamma_T^{\sigma_T})^{N}\right) \mathbb{P}\left(d\gamma_{0}^0 \mid x_0^0\right) \\
= \int c \left(\xi_{1:N}, (\gamma_1^{\sigma_1})^{(1)}(\gamma_T^{\sigma_T})^{N}\right) \mathbb{P}(d\xi) \\
\times \prod_{i=1}^{T-1} \prod_{N} \mathbb{P}\left(d\gamma_{i+1}^{\sigma_{i+1}} \mid (\gamma_{i+1}^{\sigma_{i+1}})^{i}, \gamma_T^{\sigma_T}\right) \mathbb{P}\left(d\gamma_{0}^0 \mid \omega_0\right) \\
= J_T(\gamma_1^{\sigma_1}, \ldots, \gamma_T^{\sigma_T}),
\]

where the equality (3) follows from Assumption (ii) and Assumption (c). Equality (4) follows from Assumption (ii) by exchanging \(y_i, \xi_i\) with \((\gamma_{i+1}^{\sigma_{i+1}}), (\gamma_T^{\sigma_T})^{N}\), respectively. Equality (i) and Assumption (b) imply (5). Hence, the team is exchangeable. Let \(\hat{\gamma}_T^* = (\gamma_1^{\sigma_1}, \ldots, \gamma_T^{\sigma_T})\) be a team optimal policy. Consider \(\hat{\gamma}_T^*\) as a convex combination of all possible permutations of policies by averaging them. Since action spaces are convex, \(\hat{\gamma}_T^*\) is a control policy. Following from convexity of the team in policies, we have for \(\sum_{\sigma \in \Sigma} \alpha_{\sigma} = 1\), and \(\alpha_{\sigma} \in (0, 1)\),

\[
J_T(\gamma_1^{\sigma_1}, \ldots, \gamma_T^{\sigma_T}) := J_T\left(\sum_{\sigma \in \Sigma} \alpha_{\sigma} \gamma_1^{\sigma_1} \ldots \gamma_T^{\sigma_T}\right) \\
\leq \sum_{\sigma \in \Sigma} \alpha_{\sigma} J_T(\gamma_T^*) \\
= J_T(\gamma_T^*),
\]

where the inequality follows from the hypothesis that the team problem is convex on \(\Gamma\) and the last equality follows from exchangeability of the team problem. This implies that \(\gamma_T^*\) is team optimal and the dynamic team is symmetrically optimal.

**Remark 2.** The proof of Theorem [7] shows that the above result also holds for non-classical dynamic team problems which are convex on \(\Gamma\) (see [14, Section 3.3] for a characterization of such tasks).

Here, we present the result for the class of problems that admit a static reduction (see [17, Section 3.7], [14, Section 1.2], [18], [3]).

**Lemma 1.** Consider a dynamic team problem with a symmetric partially nested information structure (see Definition [2]) which admits a static reduction. Under Assumption (i) and Assumptions (a), (b), (c) of Theorem [7], if the cost function is jointly convex on \(u_1^0, \ldots, u_N^T\) \(\mathbb{P}\)-almost surely, then the team is symmetrically optimal (through decision makers and not necessarily through time, i.e., the permutation, \(\sigma\) (see Definition [2]), is of the set \(\{1, \ldots, N\}\) and not \(\{0, \ldots, T-1\}\) (see Remark [7]).

**Proof.** The result follows from [14, Theorem 7] and Theorem [1] since the team problem is convex on \(\Gamma\).

**Lemma 2.** If a static reduction of an exchangeable, symmetrically optimal, dynamic team exists, then it is exchangeable and symmetrically optimal (through decision makers and not necessarily through time, i.e., the permutation, \(\sigma\) (see Definition [2]), is of the set \(\{1, \ldots, N\}\) and not \(\{0, \ldots, T-1\}\) (see Remark [7]).

**Proof.** The result follows from Lemma [1] and [12, Lemma 1].

**Remark 3.** Following from Remark [7] the result of Lemma [7] can not be obtained directly from [12, Lemma 1] without invoking Theorem [7] and using the generalization of the notions of exchangeability and symmetrically optimal teams introduced in this paper (see Definition [2] and Definition [3]) (through decision makers and not necessarily through time, i.e., the permutation, \(\sigma\) (see Definition [2]), is of the set \(\{1, \ldots, N\}\) and not \(\{0, \ldots, T-1\}\) (see Remark [7]). We note that substitutable teams, which have been defined in [19], are different from exchangeable teams defined here.

In the next subsection, we consider the LQG setup where the above theorem can be utilized.

**A. Symmetric partially nested LQG dynamic teams on a graph**

In the following, we consider decentralized problems where Theorem [1] can be utilized and the optimal policy can be obtained. First, we formulate LQG problems with a symmetric partially nested information pattern. Consider the following dynamics. Let \(i = 1, 2, \) and

\[
x_{i+1} = Ax_i + Bu_i + w_i,
\]

where \(\mathbb{E}(x_0^0(x_0^0)^T) \neq 0\).

**Problem (P_T):** Consider the expected cost function

\[
J_T(\gamma_1^{\sigma_1}, \gamma_T^{\sigma_T}) = \mathbb{E}\left[\sum_{t=0}^{T-1}c(x_{i+1}^t, u_{i+1}^t)\right],
\]

where \(\gamma_i^t = (\gamma_0^{0:T-1})\) for \(i = 1, 2\), and we assume that the cost function is as follows:

\[
c(x_{0:T-1}^t, u_{0:T-1}^t) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^2 (x_i^t)^T Q x_i^t + (u_i^t)^T R u_i^t \\
+ (u_1^t)^T R u_2^t + (u_2^t)^T R u_1^t,
\]

where \(R, \tilde{R} > 0\) and \(Q \geq 0\). Assume the information for each decision maker is \(\mathcal{H}_i = \{y_i, u_i^t\}, \) where \(y_i^t = H_i \xi_i + \sum_{j=0}^{T-1} D_{ij} u_j^t, \) \(\xi_i = (x_0^0, w_{0:T-1})\).

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

**Corollary 1.** If a fixed \(T\), consider a finite horizon team problem defined above as (P_T) (see (7)). If \(x_0^0\) and \(x_0^0\) are exchangeable zero mean Gaussian random vectors and \(w_i^t\) are i.i.d. Gaussian random vectors for \(i = 1, 2\), and independent for all \(t = 0, \ldots, T - 1\) and also independent
of initial states, then the dynamic team is symmetrically optimal.

**Proof.** Since the dynamic team is LQG with partially nested information structure, static reduction exists and the cost is convex in policies under a static reduction [2] and [14, Theorem 3.7]. Assumption $\mathcal{II}$ is satisfied following from the information structure. In the following, we show assumptions of Theorem $\mathcal{II}$ hold. We have, for every measurable set $A_i = \prod_{t=0}^{T-1} A_{i_t}$, $i = 1, 2$

\[
\mathbb{P}\{(\xi^1, \xi^2) \in (A_1 \times A_2)\} = \mathbb{P}\{(x_0^1, x_0^2) \in (A_0^1, A_0^2)\} \prod_{t=0}^{T-1} \mathbb{P}\{w_t^1 \in A_{t+1}^1\} \mathbb{P}\{w_t^2 \in A_{t+1}^2\}
\]

\[
= \mathbb{P}\{(x_0^1, x_0^2) \in (A_0^1, A_0^2)\} \prod_{t=0}^{T-1} \mathbb{P}\{w_t^1 \in A_{t+1}^1\} \mathbb{P}\{w_t^2 \in A_{t+1}^2\}
\]

where the first equality follows from the hypothesis that disturbances are independent random vectors and also independent of initial states, and (9) follows from the assumption that $x_0^1$ and $x_0^2$ are exchangeable and the hypothesis that $w_t$ for $i = 1, 2$ are i.i.d.; hence, this implies (b). Assumption (c) holds following from Assumption $\mathcal{II}$ and since given $\omega_0$, the set of random vectors $(y_{0:T-1})$, is mutually independent of $y_{0:T-1}$. Hence, Theorem $\mathcal{II}$ completes the proof. \qed

1 Symmetric partially nested LQG dynamic teams on a tree: Here, we consider a class of LQG dynamic teams with symmetric partially nested information structure. The information structure satisfying a tree structure where we utilize Corollary $\mathcal{III}$ and we obtain an explicit recursion for the optimal policy. Define the problem as follows:

**Problem ($P_{T}^{\mathcal{III}}$):** Consider a finite horizon expected cost $\mathcal{II}$ with the information of each DM is defined as $I^1_i = \{x^1_{i0:T-1}, w^1_{i0:T-1}\}$.

**Theorem 2.** For a fixed $T$, consider a finite horizon team problem defined as ($P_{T}^{\mathcal{III}}$) (see (7)). If $(x^1_{t0:T-1}, x^2_{t0:T-1})$ are exchangeable with an identical zero mean Gaussian distribution and $w^1_{t0:T-1}$ are i.i.d. zero mean Gaussian random for $i = 1, 2$ and independent for all $t = 0, \ldots, T - 1$ and independent of initial states, then

\[
\begin{align*}
\upsilon^*_t & = K^T_t x^1_t + L^T_t (x^1_0 | x^2_t) \\
\upsilon^*_t & = K^T_t x^2_t + L^T_t (x^1_0 | x^2_t) \end{align*}
\]

where

\[
\begin{align*}
K^T_t = & - (R + B^T P^T_{t+1} B)^{-1} B^T P^T_{t+1} A \\
P^T_t = & - A^T P^{T}_{t+1} B^T (R + B^T P^T_{t+1} B)^{-1} B^T P^T_{t+1} A \\
+ & Q + A^T P^T_{t+1} A \\
L^T_t = & - (R + B^T P^T_{t+1} B)^{-1} [\hat{R} K^T_t G^T_j + \hat{R} L^T_j \Sigma] \\
+ & \sum_{s=t+1}^{T-1} B^T (A^T)^{s-t-1} P^T_{s+1} B L^T_s \]
\]

where $\Sigma = \mathbb{E}(x^1_0 (x^2_0)^T) (\mathbb{E}(x^1_0 (x^2_0)^T))^{-1}, P^T_0 = 0, G^T_0 = 1$. Moreover, the optimal cost is as follows:

\[
\begin{align*}
J_T (\gamma^*_T) & = \frac{2}{T} \mathbb{E}(x^1_0 (x^T_0))^T P^T_0 (x^T_0) + \sum_{t=0}^{T-1} \mathbb{E}(w^1_t)^T P^T_t w^1_t \\
& + \sum_{t=0}^{T-1} \mathbb{E}(x^1_0 (x^2_0)^T) [L^T_{t+1} B^T P^T_{t+1} B L^T_t] \mathbb{E}(x^2_0 | x^1_0) \\
& + \sum_{t=0}^{T-1} \mathbb{E}(x^1_0 (A^T)^t P^T_{t+1} B L^T_t \mathbb{E}(x^2_0 | x^1_0))
\]
\]

(15)

**Proof.** Following from [2] and Radner’s theorem [4], 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-Chu’s static reduction [2]. Hence, we only need to show that (10) and (11) are person-by-person optimal. In the following, we show that for the first decision maker $J_{T_{\gamma^*_T}} \leq J_{T_{(\gamma^*_T, \gamma^*_T)}}$ for all $\beta \in \Gamma$ where $(\gamma^*_T, \beta) = (\gamma^*_T \xi^1_t, \beta^* \xi^2_t)$. This implies that $(\gamma^*_T, \gamma^*_T)$ is person-by-person optimal thanks to Corollary 4 since the dynamic team is symmetrically optimal (by exchanging policies $\gamma^*_t$, $\beta^*_t$, which implies $J_{T_{(\gamma^*_T, \gamma^*_T)}} \leq J_{T_{(\gamma^*_T, \gamma^*_T)}}$ for all $\beta \in \Gamma^i$ and this implies that $\gamma^*_T$, $\gamma^*_T$ is the fixed point of the equation). For, $i = T - 1$, the optimal policy is zero since no terminal cost has been considered. For $i = 2$, we have

\[
\begin{align*}
& F(w^1_{0:T-2}, w^2_{0:T-2}, x^1_{0:T-2}) + \min \mathbb{E}\{(A x^1_{T-2} + B w^1_{T-2})^T x^1_{T-2} \\
& \times Q (A x^1_{T-2} + B w^1_{T-2}) + (A x^1_{T-2} + B (K - x^2_{T-2} + L - 2 x^1_{T-2})^T x^1_{T-2}) \\
& \times Q (A x^1_{T-2} + B (K - x^2_{T-2} + L - 2 x^1_{T-2})^T x^1_{T-2}) + (u^2_{T-2})^T \hat{R} (K - x^2_{T-2}) + L \mathbb{E}(x^1_0 (x^2_0)^T) \hat{R} (u^2_{T-2}) \\
& + (u^2_{T-2})^T \hat{Q} x^2_{T-2} + (x^2_{T-2})^T \hat{Q} x^2_{T-2} \\
& + (u^2_{T-2})^T \hat{R} (u^2_{T-2}) + (x^2_{T-2} \hat{R} u^2_{T-2} - u^2_{T-2})^T \hat{R} u^2_{T-2}) \\
& = \mathbb{E}(u^1_{0:T-1}, w^1_{0:T-1}, x^1_0 | x^1_{0:T-1}, x^2_0) \\
& \min \{(u^1_{0:T-2})^T R (u^1_{0:T-2}) + (x^1_{T-2})^T (A^T Q B) (u^1_{0:T-2}) + (u^1_{T-2})^T B^T Q A x^1_{T-2} \\
& + (u^1_{T-2})^T \hat{R} (K - x^2_{T-2}) \mathbb{E}(x^2_0 | x^1_0) \hat{R} \hat{u}^1_{T-2} \\
& + (u^1_{T-2})^T \hat{Q} x^2_{T-2} \hat{u}^1_{T-2} + (x^2_{T-2} \hat{Q} x^2_{T-2}) \hat{u}^1_{T-2} + \hat{R} \hat{Q} \hat{R} u^2_{T-2} - u^2_{T-2})^T \hat{R} u^2_{T-2}
\]
\]

hence

\[
\begin{align*}
\upsilon^*_t & = - (R + B^T Q B)^{-1} (B^T Q A x^1_{T-2} + \hat{R} (K - x^2_{T-2}) \mathbb{E}(x^2_0 | x^1_0) \hat{R} \hat{u}^1_{T-2} \\
& \hat{Q} x^2_{T-2} - \hat{u}^2_{T-2})^T \hat{R} u^2_{T-2} - u^2_{T-2})^T \hat{R} u^2_{T-2}
\]
\]
the result for $t = 0$ using the induction hypothesis and the similar calculation as above for $t = 1$. This completes the proof.

**Remark 4.** The optimal policies (10) and (11) contain two parts which can be interpreted as follows: the first part, $k^1_T x^1_t$, 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 2 to $N$-DM LQG dynamic teams. Assume that the dynamics for $i = 1, 2, ..., N$ are defined as (6), where $E(x^0_i x^0_j)^T \neq 0$ for $i, j = 1, 2, ..., N$.

**Problem ($P_{1:T}^{N, \text{tree}}$):** Consider the expected cost function as

$$J^N_T(\gamma^1_N) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} E[\gamma^1_t (x^1_t x^1_t + (u^1_t)^T R u^1_t) + \sum_{j \neq i}^{N} \gamma^j_t (x^j_t x^j_t + (u^j_t)^T R u^j_t)],$$

where $\gamma^i_T = \gamma^0_{T-1}$ for $i = 1, ..., N$ and $R, \bar{R} > 0$ and $Q \geq 0$. Let $I^i_t = \{x^i_{[0:t]}, u^i_{[0:t-1]}\}$ for $i = 1, 2, ..., N$.

**Corollary 2.** For a fixed $T$ and $N$, consider a finite horizon team problem defined as ($P_{1:T}^{N, \text{tree}}$) (see (15)). If $(x^0_i, N)$ are exchangeable with an identical zero mean Gaussian distribution, and $w^i_s$ for $i = 1, ..., N$ are i.i.d. zero mean Gaussian random vectors, independent for $t = 0, ..., T - 1$, and independent of initial states, then $u^{s_1}(x^1_t, N, (T)) = K^1_T x^1_t + L^1_T(x^1_t | x^0_t)$, where $K^1_T$ and $P^1_T$ satisfy (12) and (13), and $L^1_T(x^1_t, N, (T))$ is a function of $K^1_T$.

**Proof.** The proof is similar to the one of Theorem 2.

In the following, we relax the assumption on the distribution of initial states and disturbances. Here, $x^0_i$ are not necessarily exchangeable and $w^i_s$ do not necessarily have the same distribution. We also consider LQ (not necessarily Gaussian) problems. To this end, we first define the notion of a strategy-by-strategy optimal policy which will be used in Corollary 3.

**Definition 5.** (Strategy-by-strategy optimal policy) A policy $\gamma_{1:T}^{N, N}(\gamma_{1:T-1}^{N, N}, \gamma_{0:T-1}^{N, N})$ is strategy-by-strategy optimal if the following inequalities hold: for all $i = 1, ..., N$, and for all $\beta \in \Gamma^i$,

$$J(\gamma_{1:T-1}^{N, N}, \gamma_{0:T-1}^{N, N}) \leq J(\gamma_{1:T-1}^{N, N}, \beta_{1:T}, \gamma_{1:T}^{N, N})$$

**Remark 5.** Strategy-by-strategy optimality allows for joint perturbation in policies of DMs in a given precedence graph $\gamma^i$ but person-by-person optimality (under the intrinsic model) only allows perturbation for a single DM’s policy $\gamma^i$.

**Corollary 3.** For a fixed $T$, consider a finite horizon team problem defined as ($P_{1:T}^{\text{tree}}$) (see (7)). If $x^0_i$ and $x^0_j$ have an identical zero mean Gaussian distribution and $\{w^i_s\}$s are independent random vectors for $i = 1, 2$ and $t = 0, ..., T - 1$ (not necessarily Gaussian and identical for $i = 1, 2$ and for all $t = 0, ..., T - 1$), then (10) and (11) are team optimal.

**Proof.** Following from [2] and Radner’s theorem [4], the optimal policy and person-by-person optimal policy are unique, hence the strategy-by-strategy optimal policy is also unique. We only need to show that (10) and (11) are strategy-by-strategy optimal and this implies globally optimality thanks to the uniqueness of such policies for this problem. Since $x^0_i$ and $x^0_j$ are not exchangeable, the team may not be exchangeable, hence one can not justify symmetry using Corollary 1 however, we show that by fixing policies of a DM and minimizing the cost for the other one, the optimal policy for each decision maker is identical for $t = 0, ..., T - 1$, and this implies that the team is symmetrically optimal thanks to the uniqueness of the strategy-by-strategy optimal policy for this problem. We can use dynamic programming to show (17) holds for (10) and (11), and this completes the proof.

**Remark 6.** The above result can be extended to the case where $w^1_t$ and $w^2_t$ are exchangeable for all $t = 0, ..., T - 1$ (not necessarily independent), and $\{w^i_s\}$s are independent zero mean Gaussian random vectors and independent of initial states (i.e., $(w^1_t, w^2_t) \parallel (w^1_k, w^2_k)$ for $t \neq k$, and $(w^1_t, w^2_t) \parallel (x^0_t, x^0_t)$ for all $t = 0, ..., T - 1$), then $u^{s_1}(x^1_t) = K^1_T x^1_t + L^1_T E[x^1_t | x^0_t] + \sum_{p=0}^{t-1} h^1_T(x^1_p | x^0_p)$, where $K^1_T$ and $L^1_T$ satisfy (12) and (13), respectively.

Following from Corollary 3 we present the Certainty equivalency property for symmetric LQG teams on a tree.

**Theorem 3.** For a fixed $T$, consider a finite horizon team problem defined as ($P_{1:T}^{\text{tree}}$) (see (7)). If $x^0_i$ and $x^0_j$ are exchangeable with an identical zero mean Gaussian distribution and $w^i_s$ are zero i.i.d. random vectors for $i = 1, 2$ and for all $t = 0, ..., T - 1$ and independent of initial states, then the team problem is certainty equivalent, i.e., the optimal controller’s gains $K^1_T(x^1_t)$ and $L^1_T(x^1_t)$ are independent of the distributions of disturbances $w^i_s$.

**Proof.** Let $i = 1, 2$, and $t = 0, ..., T - 1$, $x^i_{t+1} = Ax^i_t + Bu^i_t$ with $x^0_i$ and $x^0_j$ are correlated random vectors. Following from dynamic programming by fixing policies of DM2 to (11), similar to the proof of Corollary 2 we can show that the optimal policies are the identical for DMs and equivalent to (10) and (11).

2) Symmetric output feedback partially nested LQG dynamic teams on a graph: In the following, we develop a structural result to the case where the information structure of each decision maker over time satisfies a structure which is identical for all DMs and is partially nested. An example of such a graph structure has been depicted in Fig. 2.
Theorem 4. For a fixed $T$ and $N$, consider a finite horizon team problem defined as $(P^N_T)$ (see [16]). If $x^i_0$ and $z^i_0$ are exchangeable with an identical zero mean Gaussian distribution and $w_i$s for $i = 1, 2$ are i.i.d. zero mean Gaussian random vectors, independent for all $t = 0, \ldots, T - 1$ and independent of initial states, then

$$ u^i_{t}(T,N) = K^i_{t}(T) y^i_t + L^i_{t}(T,N) \sum_{j=1, j \neq i}^{N} \mathbb{E}(y^j_0 | y^i_0) $$

where $K^i_{t}(T)$ are obtained by considering only one decision maker and ignoring other decision makers.

Proof. The proof is similar to that of Theorem 2 using the results in [2] and Corollary 1.

Remark 7. A related work is [20], 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 decision makers who have no common ancestors and no common descendants have either 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 [20], where a multi-tree structure has been imposed on a graph.

IV. 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. Unlike the existing results in mean field games, mean field teams, and social optima mean field control problems where provided that initial states and disturbances are independent, for games the limit of the sequence of Nash equilibrium has been investigated when the number of decision makers goes to infinity, and for teams and social optima control problems the centralized performance has been achieved asymptotically by a decentralized controller [21], [22], [23], [24], [25], [26], here, through concentration of measure arguments, we show the convergence of optimal policies for $N$-DM convex mean field teams with a symmetric information structure to the corresponding optimal policies of mean field teams.

Remark 8. We note that a related paper which considered games is [27] where a 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. This means that the information of each DM is not affected by any of the actions of DMs. For dynamic teams, there are two difficulties: (i) obtaining variational equations is challenging since fixing policies of DMs and perturbing only one DM’s policies, perturbs the observation of other DMs and hence the controls $u^{-i} = (y^1, \ldots, \gamma_0^{t-1} (y^{i-1}), \gamma_1^{t+1} (y^{i+1}), \ldots, \gamma_N^{t} (y^{N}))$. (ii) solutions of variational equations which give person-by-person optimal policies are inconclusive for global optimality due to the lack of convexity in general since convexity of the cost function in $(u^1, \ldots, u^N)$ does not necessarily imply the convexity of the team problem in policies, hence person-by-person optimality does not necessarily imply the global optimality (for sufficient conditions of convexity in policies see [14, Section 3.3]).

Problem $(P^{MF}_T)$: Consider the expected cost function for $N$-DM teams as

$$ J_T^N (\gamma^1_{T-1}) = \frac{1}{N} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}_{\gamma^i_t} \left[ c(x^i_t, u^i_t, \frac{1}{N} \sum_{p=1}^{N} u^p_t, \frac{1}{N} \sum_{p=1}^{N} x^p_t) \right], $$

where $\gamma^i_{T-1} = \gamma^0_{T-1}$ and the cost function is continuous and non-negative.

Problem $(P^{\infty, MF}_T)$: Consider the expected cost for mean field teams as

$$ J_T^{\infty} (\gamma^1_{T-1}) = \limsup_{N \to \infty} J_T^N (\gamma^1_{T-1}), $$

where $\gamma^i_{T-1} = \gamma^i_{0,T-1}$ for $i \in \mathbb{N}$. Let

$$ x^i_{t+1} = f(x^i_t, u^i_t, w^i_t), $$

$$ y^i_t = h(c^i_t, u^i_t), $$

where the function $f$ and $h$ are continuous in all arguments and $I^i_t = \{ y^i_t \}$ for $i \in \mathbb{N}$.

Lemma 3. Consider a team defined as $(P^{N,MF}_T)$ (see (18)) with a symmetric information structure. Assume the problem is convex in policies. Let the action space be compact and convex for each decision maker. Assume $(x^i_0)$ are exchangeable zero mean random vectors with an identical distribution, and $w_i$s are i.i.d. zero mean random vectors for $i = 1, \ldots, N$, independent for $t = 0, \ldots, T - 1$, and independent of initial states. Assume $\gamma$’s are independent argument in the proof of Corollary 1 since (18) satisfies (1).

Theorem 5. Consider a team defined as $(P^{N,MF}_T)$ (see (18)) with a symmetric information structure. Assume the problem is convex in policies. Let the action space be compact and convex for each decision maker. Assume $\omega = \{ x^i_0 \} \in \mathcal{N}$ are exchangeable zero mean random vectors with an identical distribution, and $w_i$s are i.i.d. zero mean random vectors for $i = 1, \ldots, N$, independent for $t = 0, \ldots, T - 1$, and independent of initial states. Assume $\gamma$’s are independent for $i = 1, \ldots, N$. If there exists a sequence of optimal policies for $(P^{N,MF}_T)$, $\{ \gamma^i_N \}$, which converges (for every decision maker due to the symmetry) pointwise to $\gamma^{\infty}$ as $N \to \infty$, then $\gamma^{\infty}$ (which is identically symmetric) is an optimal policy for $(P^{\infty, MF}_T)$.

Proof. Following from Lemma 1 and an analogous argument in the proof of Corollary 1 since (18) satisfies (1).
where we denote 

\[ Q_N^*(B) := \frac{1}{N} \sum_{i=1}^{N} 1_{\beta_i \in B} \text{ where } \beta_i := (\gamma_i^*, \omega_i) \]

where \( B \in \mathcal{Z} := \mathbb{U} \times \mathbb{Y} \), and \( \mathbb{U} = (\prod_{k=0}^{T-1} \mathbb{U}^k), \mathbb{Y} = (\prod_{k=0}^{T-1} \mathbb{Y}^k) \). In the following, we follow two steps analogous to the proof of [12, Theorem 8] to complete the proof. Since the setup is different, we give details below. First, we show that \( Q_N^* \) converges \( \mathbb{P} \)-almost sure weakly to \( Q^* \), then we show \( \limsup N \rightarrow \infty J_N^* (\gamma_{\tau}^*) = J_\mathbb{P}^* (\gamma^*) \) and invoke [12, Theorem 5].

(Step 1): In this step, we show that for every \( g \in C_b (\mathcal{Z}) \), where we denote \( C_b (\mathcal{Z}) \) as the space of continuous and bounded functions in \( \mathcal{Z} \), \( \mathbb{P} \)-almost surely,

\[
\lim_{N \rightarrow \infty} \left| \int g dQ_N^* - \int g dQ^* \right| \leq \lim_{N \rightarrow \infty} \left( \left| \int g dQ_N^* - \int g dQ_N \right| + \left| \int g dQ_N - \int g dQ^* \right| \right) = 0.
\]

We have, for every \( g \in C_b (\mathcal{Z}) \),

\[
\lim_{N \rightarrow \infty} \mathbb{P} \left( \left| \int g dQ_N^* - \int g dQ_N \right| \geq \epsilon \right) \leq \epsilon^{-1} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left| g(\gamma_i^*, \omega_i) - g(\gamma_i^*, \omega_i) \right| \right) \]

\[= \epsilon^{-1} \lim_{N \rightarrow \infty} \mathbb{E} \left( \left| g(\gamma_i^*, \omega_i) - g(\gamma_i^*, \omega_i) \right| \right) \]

where (23) follows from Markov’s inequality, the triangle inequality and the definition of the empirical measure, and (24) follows from the fact that \( \gamma_i^* \)'s are identical random variables (this follows from symmetry of the information structure and Lemma 3). Since \( g \) is bounded and continuous, the dominated convergence theorem implies (25). Following from (25), for every subsequence there exists a further subsequence such that \( \int g dQ_N^* \rightarrow \int g dQ^* \) converges to zero \( \mathbb{P} \)-almost surely as \( l \rightarrow \infty \). Now, we show that \( Q_N^* \) converges \( \mathbb{P} \)-almost sure weakly to \( Q^* = \delta_{\mathbb{E}(\gamma_{\tau}^*)} \) in the space of probability measures in \( \mathcal{Z} \), \( \mathcal{P}(\mathcal{Z}) \), that is \( \int g dQ_N^* \rightarrow \int g dQ^* \) converges to zero \( \mathbb{P} \)-almost surely for every \( g \in C_b (\mathcal{Z}) \). Let

\[ L(\gamma^*, \omega, y) := g(\gamma^*, \omega, y) - \mathbb{E} \left( g(\gamma^*, \omega, y) \right) \]

We have,

\[
\lim_{N \rightarrow \infty} \mathbb{P} \left( \left| \int g dQ_N^* - \int g dQ^* \right| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \lim_{N \rightarrow \infty} \mathbb{E} \left( \left| \sum_{i=1}^{N} g(\gamma_i^*, \omega_i) - \sum_{i=1}^{N} g(\gamma_i^*, \omega_i) \right| \right)^2 = \lim_{N \rightarrow \infty} \frac{1}{(N \epsilon)^2} \mathbb{E} \left( \left( \sum_{i=1}^{N} L(\gamma_i^*, \omega_i, y_i) \right)^2 \right) \]

(26)

\[= \frac{1}{(N \epsilon)^2} \mathbb{E} \left( \left( \sum_{i=1}^{N} L(\gamma_i^*, \omega_i, y_i) \right)^2 \mid \omega_0 \right) \]

(27)

(28)

\[= \lim_{N \rightarrow \infty} \frac{1}{(N \epsilon)^2} \sum_{i=1}^{N} \mathbb{E} \left( \left( L(\gamma_i^*, \omega_i, y_i) \right)^2 \mid \omega_0 \right) \]

where (26) follows from Chebyshev’s inequality, and (27) follows from the law of iterated expectations. Given \( \omega_0, y_i \)'s are independent, and also identical following from symmetry of the information structure and Lemma 3. Hence, this implies that the second term of (28) is zero, and (29) is true using the law of iterated expectations and since \( g \in C_b (\mathcal{Z}) \). Hence, (22) holds through choosing a suitable subsequence, and \( Q_N^* \) converges \( \mathbb{P} \)-almost sure weakly to \( Q^* \).

(Step 2): We have

\[
\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E} \left[ c(\omega_0, \gamma_j^*(\omega, y), \frac{1}{N} \sum_{i=1}^{N} \gamma_i^*(\omega, y)) \right] = \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E} \left[ c(\omega_0, \gamma_j^*(\omega, y), \frac{1}{N} \sum_{i=1}^{N} \gamma_i^*(\omega, y)) \right] \]

(30)

where following from (20) and (18), (30) is true for some \( \mathbb{P} \)-almost sure continuous \( \tilde{c} : \Omega \times \mathbb{U} \times \mathbb{U} \times \mathbb{Y} \rightarrow \mathbb{R} \), and the first inequality follows from the definition of \( V_N^* \) and replacing limsup by liminf. The second inequality follows from Fatou’s lemma. In the following, we justify (31). Since \( Q_N^* \) converges \( \mathbb{P} \)-almost sure weakly to \( Q^* \), using continuous mapping theorem [28, page 20], we have \( Q_N^* \rightarrow Q^* \) \( \mathbb{P} \)-almost sure weakly to \( Q^* \) \( \mathcal{U} \times \mathcal{Y} \), hence the compactness of \( \mathcal{U} \) implies \( \int_{\mathcal{U}} u Q_N^* \rightarrow \int_{\mathcal{U}} u Q^* \) \( \mathbb{P} \)-almost sure for all \( \omega \in \Omega \). Since given \( \omega_0, y_i \)'s are i.i.d., hence we can show that \( \int_{\mathcal{U}} y Q_N^* \rightarrow \int_{\mathcal{U}} y Q^* \) using the law of iterated expectations. Since the cost function is \( \mathbb{P} \)-almost sure continuous, \( \tilde{c}(\omega_0, u, \int_{\mathcal{U}} u Q_N^* \rightarrow \int_{\mathcal{U}} u Q^* \) \( \mathbb{P} \)-almost sure for all \( \omega \in \Omega \). Define a non-negative bounded sequence \( G_N := \min \{ M, c(\omega_0, u, \int_{\mathcal{U}} u Q_N^* \rightarrow \int_{\mathcal{U}} y Q_N^* \} \) \( \mathbb{P} \)-almost sure. Where the sequence \( \{ G_N \} \) converges as \( M \rightarrow \infty \) to \( G := \tilde{c}(\omega_0, u, \int_{\mathcal{U}} y Q_N^* \rightarrow \int_{\mathcal{U}} y Q_N^* \) \( \mathbb{P} \)-almost sure, then we have \( \mathbb{P} \)-almost surely

\[
\liminf_{N \rightarrow \infty} \int_{\mathcal{U}} c(\omega_0, u, \int_{\mathcal{U}} y Q_N^* \rightarrow \int_{\mathcal{U}} y Q_N^*) \]

(31)
where the first inequality follows from the definition of $G^M_N$ and the second equality is true using [29, Theorem 3.5] or [30, Theorem 3.1], since $G^M_N$ is bounded (hence is uniformly $Q^c_N$-integrable) and continuously converges to $G^M$, and the monotone convergence theorem implies the last equality. Hence, (31) holds which implies $\limsup_{T \to \infty} J^N_F(\gamma^N,\infty) = J^F(\gamma^\infty,\infty)$ and [12, Theorem 5] completes the proof.

**Remark 9.** Analogous to [12, Theorem 9], we can relax the hypothesis that action spaces are compact; this is particularly important for LQG models.

### A. LQG mean field teams with a symmetric partially nested information structure

In the following, we present results for LQG teams with a mean field coupling through the cost function. First, using Corollary 2 we obtain globally optimal policies for N-DM LQG teams with a mean field coupling with correlated initial states and disturbances. We note however that in our model, the coupling is only in the cost function and not in the dynamics since this makes the problem non-classical. Next, as an implication of Theorem 3 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.

**Problem (P$_{N,MF}$):** Consider the expected cost function for N-DM LQG teams with a mean field coupling as

$$J^N_F(\gamma^N,\infty) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}^i [x^i_t]^TQx^i_t + (u^i_t)^TRu^i_t + \frac{1}{N-1} \sum_{j=1, j \neq i}^{N} (u^j_t)^TRu^j_t + \frac{1}{N-1} \sum_{j=1, j \neq i}^{N} (x^j_t)^T \tilde{Q}x^j_t + (x^i_t)^T \tilde{Q}x^i_t],$$

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

**Problem (P$_{\infty,N,MF}$):** Consider the expected cost function for LGQ mean field teams as

$$J^\infty_F(\gamma^\infty,\infty) = \limsup_{N \to \infty} J^N_F(\gamma^N,\infty)$$

Let $I^i_t = \{x^i_{[0,t]}, u^i_{[0,t-1]}\}$ for $i \in \mathbb{N}$, and dynamics be as (6).

**Corollary 4.** For a fixed $T$ and $N$, consider a finite horizon team problem defined as (P$_{T,LQG}$) (see (32)). If $(x^1_{0,T})$ are exchangeable zero mean Gaussian random vectors with an identical distribution, and $w_i$'s are i.i.d. zero mean Gaussian random vectors for $i = 1, \ldots, N$, independent for $t = 0, \ldots, T - 1$, and independent of initial states, then

$$u^{i,*(T), (N)}_t = K^{(T)}_i x^i_t + \frac{L^{(N), (T)}}{N-1} \sum_{j=1, j \neq i}^{N} \mathbb{E}(x^j_0|x^i_0),$$

where $K^{(T)}_i$ and $P^{(T)}_i$ satisfy (12) and (13), and $L^{(N), (T)}$ is a function of $K^{(T)}_i$. Proof. The proof is similar to the one in Theorem 2.

**Corollary 5.** For a fixed $T$, consider a finite horizon team problem defined as (P$_{T,LQG}$) (see (33)). Assume $(x^1_{0,T})$ are exchangeable zero mean Gaussian random vectors with an identical distribution, and $w_i$'s are i.i.d. zero mean Gaussian random vectors for $i = 1, \ldots, N$, independent for $t = 0, \ldots, T - 1$, and independent of initial states. If $L^{(N), (T)}_i$ converges pointwise as $N \to \infty$ to $L^{(\infty), (T)}_i$, then

$$u^{i,*(T), (\infty)}_t = K^{(T)}_i x^i_t + \frac{L^{(\infty), (T)}_i \Sigma x^i_0}{\Sigma},$$

where $K^{(T)}_i$ and $P^{(T)}_i$ satisfy (12) and (13), and $\Sigma = \mathbb{E}(x^1_0|x^i_0)(\mathbb{E}(x^1_0|x^i_0)^T)^{-1}$.

**Proof.** Invoking Theorem 5 Corollary 4 using Remark 9 complete the proof.

### V. AVERAGE COST INFINITE HORIZONS PROBLEMS FOR PARTIALLY NESTED DYNAMIC TEAMS

In the following, we consider average cost problems with symmetric partially nested information structure as well as LQG dynamic teams with the one step delay sharing pattern and sparsity constraints.

**A. Average cost horizon problems for symmetric partially nested LQG dynamic teams on a tree**

Now, consider infinite horizon team problems.

**Problem (P$_{\infty,\infty}$):** Consider the following expected cost function

$$J(\gamma^{1:2}) = \lim_{T \to \infty} \mathbb{E}(x^1_0|x^i_0; T - 1, u^{1:2}_{T-1}]),$$

where the cost function is defined as (6) and the information of each DM is defined as $I^i_t = \{x^i_{[0,t]}, u^i_{[0,t-1]}\}$.

First, we introduce a lemma essential for Theorem 6.

**Lemma 4.** Consider the sequence $(a^i_0)_{T=1}^{T}$. Assume $\lim_{T \to \infty} a^i_T = a$ for $i = 0, \ldots, T - 1$. If for every fixed $T \in \mathbb{N}$, $a^i_{l+1}$ for all $l = 0, \ldots, T - 1$, then

$$\lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^{T} a^i_T = a.$$

**Proof.** We have

$$\lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^{T} a^i_T = \lim_{T \to \infty} \frac{1}{T} \sum_{i=0}^{T-1} a^i_{T-1} = \lim_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} a^i_k = a,$$

where the second equality follows from the assumption $a^i_T = a^i_{T+1}$ and the last equality follows from the Cesàro mean argument.

**Theorem 6.** Consider average cost infinite horizon team problems defined as (P$_{\infty,\infty}$) (see (35)). Assume $(A, B)$ are
stabilizable and \((A,Q^2)\) are detectable. Assume \(x_0^1\) and \(x_0^2\) are exchangeable with an identical zero mean Gaussian distribution and \(w_t^i\) are i.i.d. zero mean Gaussian random variables for \(i = 1, 2\) and for all \(t = 0, \ldots, T - 1\) and independent of initial states. If \(L_t^{(T)}\) (see (14)) converges pointwise to \(L_t^{(\infty)}\) as \(T \to \infty\), then the pointwise limit of the sequence of optimal policies for \((P_t^{T,\infty})\) is team optimal for \((P_t^{T,\infty})\) and stabilizes the closed-loop system,

\[
  u_t^{1\ast(\infty)} = K x_t^1 + L_t^{(\infty)} E(X_0^2|x_0^1),
  \quad u_t^{2\ast(\infty)} = K x_t^2 + L_t^{(\infty)} E(X_0^1|x_0^2),
\]

where

\[
  K = -(R + B^T P B)^{-1} B^T P A,
  P = Q + A^T P A - A^T P B^T (R + B^T P B)^{-1} P B A,
  L_t^{(\infty)} = -(R + B^T P B)^{-1} \{ RKG_t^{(\infty)} + \bar{R}L_t^{(\infty)} \Sigma \}
  + \sum_{s=t+1}^{\infty} B^T (A^T)^{t-s} P B L_s^{(\infty)}
  G_t^{(\infty)} = (A + BK)^t + \sum_{s=1}^{t} (A + BK)^{t-s} B L_s^{(\infty)} \Sigma
\]

Proof. We show \(\limsup_{T \to \infty} J_T(\gamma_{t}^+; \gamma_{t}^-) = J(\gamma_{t}^+; \gamma_{t}^-)\) and invoke [12, Theorem 5] or [31, Theorem 1] to complete the proof. From (14), we have

\[
  \lim_{T \to \infty} |J_T(\gamma_{t}^-) - J_T(\gamma_{t}^+)| = \lim_{T \to \infty} \frac{T}{T} \left( \sum_{t=0}^{T-1} E((x_t^1)^T(P_t^T - P)x_0^1) + \sum_{t=0}^{T-1} E((x_t^2)^T(P_t^T - P)x_0^2) \right)
  \leq \lim_{T \to \infty} \frac{T}{T} \left( \sum_{t=0}^{T-1} E((x_t^1)^T(P_t^T - P)x_0^1) + \sum_{t=0}^{T-1} E((x_t^2)^T(P_t^T - P)x_0^2) \right)
  \leq \sum_{t=1}^{T-1} \left( E((x_t^1)^T(P_t^T - P)x_0^1) + E((x_t^2)^T(P_t^T - P)x_0^2) \right)
  = 0,
\]

where (36) is zero since \(P_t^{(T)} \rightarrow P\), and (37) converges to zero using Lemma 4 since \(P_{t+1}^{(T)} = P_t^{(T)}\). Expression (39) converges to zero since \(L_{t}^{(\infty)}\) (see (14)) converges pointwise to \(L_t^{(\infty)}\) as \(T \to \infty\), we have \(\sum_{t=0}^{\infty} T B^T (A_t^T)^{t-s} P B L_{t+s}^{(\infty)} < \infty\), and this implies that \(\lim_{t \to \infty} L_t^{(\infty)} = 0\). Hence, we have for every \(\epsilon > 0\), there exists \(N > T\) such that for ever \(t > N\), \(|Tr[L_t^{(\infty)}(L_t^{(\infty)})^T]| < \epsilon\). We define \(L_t^{(T)} = 0\) for \(t > T\). Expression (38) is equal to zero since

\[
  \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left( Tr[(L_t^{(T)})(L_t^{(T)})^T B^T P_{t+1}^{(T)} B L_t^{(T)}] - (L_t^{(\infty)})(L_t^{(\infty)})^T B^T P_{t+1}^{(T)} B L_t^{(T)} \right)
  = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left( Tr[L_t^{(T)}(L_t^{(T)})^T B^T P_{t+1}^{(T)} B L_t^{(T)}] - (L_t^{(\infty)})(L_t^{(\infty)})^T B^T P_{t+1}^{(T)} B L_t^{(T)} \right)
  \leq \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left( M_1 T |Tr[L_t^{(T)}(L_t^{(T)})^T B^T P_{t+1}^{(T)} B L_t^{(T)}] - (L_t^{(\infty)})(L_t^{(\infty)})^T B^T P_{t+1}^{(T)} B L_t^{(T)} \right)
  + \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left( M_2 T |Tr[B^T (P_{t+1}^{(T)} - P) B] | \right)
  \leq \sum_{t=0}^{T-1} \left( M_1 T |Tr[L_t^{(T)}(L_t^{(T)})^T B^T P_{t+1}^{(T)} B L_t^{(T)}] - (L_t^{(\infty)})(L_t^{(\infty)})^T B^T P_{t+1}^{(T)} B L_t^{(T)} \right)
  \leq \sum_{t=0}^{T-1} \epsilon \leq \epsilon,
\]

where the result follows from Lemma 4 and the fact that \(|Tr[L_t^{(\infty)}(L_t^{(\infty)})^T]| < \epsilon\) for every \(t > N\). Hence, equality (40) is true and global optimality follows from [12, Theorem 5]. Now, we show the stability of the closed-loop system. We have,

\[
  \lim_{t \to \infty} E(||x_t^1||^2) \leq \left( ||A + BK|| \right)^2 E(||x_0^1||^2) + \sum_{i=0}^{T-1} \left( ||A + BK|| \right)^2 (t-1-i) \times E(||B L_t^{(\infty)}(x_0^2|x_0^1)||^2) + ||w_t^1||^2 \right) < \infty,
\]

where the last inequality follows from the fact that \(||A + BK|| < 1\) (all the eigenvalues of \(A + BK\) are inside of the unit circle), and \(||L_t^{(\infty)}|| < M\).

**B. Average cost team problems with one step delayed sharing pattern and sparsity constraints**

In [11], the problem of finding a steady state solution for state feedback average cost LQG team problems with a delay and sparsity has been considered. Under the independence of disturbances assumption, convergence of the solution to the stabilizing solutions for finite horizon of the corresponding algebraic Riccati equation has been shown [11]; however, the global optimality of the limit solution has not been established. We establish the global optimality of the limit solution. Consider the following dynamics for \(i = 1, \ldots, N\),

\[
  x_{t+1}^i = \sum_{j} A_{ij} x_t^j + B_{ij} u_t^j + \omega_t^i,
\]

where for a fixed \(i = 1, \ldots, N\), \(A_{ij}\) and \(B_{ij}\) are zero if sum of the delays along the directed path from \(j\) to \(i\) with the shortest path, \(D_{ij}\), is more than one, i.e., \(D_{ij} > 1\). Assume \(x_0^i\)s and \(\omega_0^i\)s are mutually independent Gaussian random vectors for all \(i = 1, \ldots, N\) and \(\omega_0^i\)s are independent for \(t = 0, \ldots, T - 1\) with covariance matrix \(\Sigma^i\). The information structure is defined as

\[
  I_t^i = \{ x_k^i | j = 1, \ldots, N, k = 0, \ldots, t - D_{ij} \}.
\]

The above condition along with the assumption on the dynamics of the system implies that the information structure is partially nested. Since the disturbances and initial states are mutually independent, the information structure can be
decomposed into the independent sets [11]. To this end, an information graph, $G(\mathcal{M}, \mathcal{F})$, has been defined for a graph [11]. In the following, we follow the notation in [11]. Define $s^j_k$ as a set of all nodes reachable from node $j$ within $k$ steps. Define $\mathcal{M} := \{s^j_k | j = 1, \ldots, N \text{ and } k \geq 0\}$ and $\mathcal{F} := \{(s^j_k, s^{j+1}_k) | j = 1, \ldots, N \text{ and } k \geq 0\}$. The following theorem is from [11], and gives an optimal policy for finite horizon problems. Assume that the delay of sharing between each decision maker is not greater than one. Assume there are no directed cycles with a total delay of zero.

**Theorem 7.** [11] Consider the LQG team problem with

$$J_T(\tau_T) := \frac{1}{T} \mathbb{E}_\tau T \left[ \sum_{t=0}^{T-1} \left( Q S^T R S^T u_t \right) \right]$$

where $\tau > 0$, $Q \geq 0$, and $[Q S^T R S^T] \geq 0$, $u_t = (u^1_t, \ldots, u^N_t)$. Then, the optimal controller is

$$u^*_t = \sum_{r \in \mathcal{M}} I^{(i),r}_{t} K^r_{t} \tilde{\zeta}^r_{t},$$

where $r \in \mathcal{M}$, $I^{(i),r}$ is the identity matrix partition according to $\{t\}$ and $r$, and

$$K^r_{t} = -(R^{rr} + (B^{rr}S^T X^r_{t+1}B^{rr})^{-1}(S^{rr} + (A^{rr}S^T X^r_{t+1}B^{rr})$$

$$- (K^r_{t})^T (R^{rr} + (B^{rr}S^T X^r_{t+1}B^{rr})K^r_{t}$$

$$X^r_{t} = Q^{rr} + (A^{rr}S^T X^r_{t+1}A^{rr})$$

and

$$\tilde{\zeta}^r_{t+1} = \sum_{s \rightarrow r} (A^{rr} + B^{rr}K^r_{t}) \tilde{\zeta}^s_{t} + \sum_{s \rightarrow r} I^{s,1}_{t} \omega^{s}_{t},$$

where $\tilde{\zeta}^s_{t}$ is stabilizable and $\sum_{s \rightarrow r} I^{s,1}_{t} \omega^{s}_{t}$, where $s \in \mathcal{M}$ is the unique node such that there is an edge between $r$ and $s$, i.e., $r \rightarrow s$. Moreover, the optimal cost is obtained as

$$J_T(\gamma^*_T) = \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} \left[ (X^i_{t+1})^{(i),1} \mathbb{E} (x^i_{t+1}(x^i_0)^T) \right]$$

$$+ \sum_{t=0}^{T-1} \sum_{s \rightarrow r} \mathbb{E} \left[ (X^i_{t+1})^{(i),1} \Sigma^i_{t} \right],$$

where $(X^i_{t+1})^{(i),1}$ denotes the sub-matrix $(X^i_{t+1})$ corresponds to the $i$-th array.

In the following, we refine a related result in [11].

**Theorem 8.** Consider the LQG team problem with

$$J(\gamma) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=0}^{T-1} \left( Q S^T R S^T u_t \right) \right]$$

Assume for self loops $s \rightarrow s$ in the information graph, $(A^{ss}, B^{ss})$ is stabilizable and $[A^{ss} - e^{\theta} B^{ss} R^{ss}]$ has a full column rank for every $\theta \in [0, 2\pi]$, where $C^{ss}$ and $D^{ss}$ are the matrix of the form

$$[C^{ss} \ D^{ss}]^T [C^{ss} \ D^{ss}] = \left[ \begin{array}{cc} C^{ss} & D^{ss} \\ D^{ss} & C^{ss} \end{array} \right].$$

Then,

$$u^*_t = \sum_{r \in \mathcal{M}} I^{(i),r} K^{(\infty),r} \tilde{\zeta}^r_{t},$$

where $K^{(\infty),r} = \lim_{T \to \infty} K^r_{T}$, $X^{(\infty)} = \lim_{T \to \infty} X^r_{T}$ for all $t$ and $\zeta^r_{(\infty)}$ is obtained from (42) by replacing $K^r_{T}$ with $K^{(\infty),r}$.

**Proof.** As $T \to \infty$, we have $X^{(\infty)} = \lim_{T \to \infty} X^r_{T}$ for (41) since the the recursion for $s \to s$ corresponds to the classical Riccati equation and $X^{(\infty)}$ is a continuous function of $X^r_{T}$ [11, Corollary 7]. We have,

$$\lim_{T \to \infty} |J_T(\gamma^*_T) - J_T(\gamma^*_T)|$$

$$= \lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^{N} \mathbb{E} \left[ (X^i_{0})^{(i),1} \mathbb{E} (x^i_{0}(x^i_0)^T) \right]$$

$$+ \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E} \left[ (X^i_{t+1})^{(i),1} \Sigma^i_{t} \right]$$

$$= 0,$$

where (43) and Lemma 4 implies the last equality. This is because (41) implies that $X^{(\infty)} = X^r_{T}$; hence, invoking [12, Theorem 5] implies the global optimality of $\gamma^*_T$. The stability argument follows from [11, Corollary 7] and [32].

**VI. CONCLUSION**

In this paper, we studied dynamic teams with symmetric partially nested information structure. We presented a characterization for symmetrically optimal teams for convex exchangeable team problems. We obtained global optimal solutions for average cost finite and infinite horizon dynamic team problems with symmetric partially nested information structure. For mean field teams with symmetric partially nested information structure, we show the convergence of optimal polices for mean field teams with $N$ decision makers to the corresponding optimal policy of mean-field teams. Moreover, we obtained globally optimal policies for average cost finite horizon state feedback problems for LQG dynamic teams with sparsity and delay constraints.

**REFERENCES**

[1] H. S. Witsenhausen, “Separation of estimation and control for discrete time systems,” Proceedings of the IEEE, vol. 59, pp. 1557–1566, November 1971.

[2] Y. C. Ho and K. C. Chu, “Team decision theory and information structures in optimal control problems - part I,” IEEE Transactions on Automatic Control, vol. 17, pp. 15–22, February 1972.

[3] H. Witsenhausen, “Equivalent stochastic control problems,” Math. Control, Signals, and Systems, vol. 1, pp. 3–11, 1988.

[4] R. Radner, “Team decision problems,” Annals of Mathematical Statistics, vol. 33, pp. 857–881, 1962.

[5] M. Rotkowitz and S. Lall, “A characterization of convex problems in decentralized control,” IEEE Transactions on Automatic Control, vol. 51, pp. 274–286, February 2006.

[6] M. Rotkowitz, “On information structures, convexity, and linear optimality,” in IEEE 47th Annual Conference on Decision and Control (CDC), 2008, pp. 1642–1647.
