On Team Decision Problems With Nonclassical Information Structures

Andreas A. Malikopoulos , *Senior Member, IEEE*

Abstract—In this article, we consider sequential dynamic team decision problems with nonclassical information structures. First, we address the problem from the point of view of a "manager" who seeks to derive the optimal strategy of the team in a centralized process. We derive structural results that yield an information state for the team, which does not depend on the control strategy, and thus, it can lead to a dynamic programming decomposition where the optimization problem is over the space of the team's decisions. We, then, derive structural results for each team member that yield an information state which does not depend on their control strategy, and thus, it can lead to a dynamic programming decomposition where the optimization problem for each team member is over the space of their decisions. Finally, we show that the solution of each team member is the same as the one derived by the manager. We present an illustrative example of a dynamic team with a delayed sharing information structure.

Index Terms—Decentralized control, Markov decision theory, nonclassical information structures, team theory.

I. Introduction

A. Motivation

TEAM theory [1]–[3] is a mathematical formalism for decentralized stochastic control problems [4] in which a "team," consisting of a number of members, cooperates to achieve a common objective. It was developed to provide a rigorous mathematical framework of cooperating members in which all members have the same objective yet different information. The underlying structure to model a team decision problem consists of the following [3]:

1) a number of $K \in \mathbb{N}$ members of the team;
2) the decisions of each member;
3) the information available to each member, which is different;
4) an objective, which is the same for all members;
5) the existence, or not, of communication between team members.

Team theory can be applied effectively in applications that include informationally decentralized systems such as emerging mobility systems [5], and in particular, optimal coordination of connected and automated vehicles at traffic scenarios [6]–[11], networked control systems [12], [13], mobility markets [14], smart power grids [15], [16], power systems [17], cooperative cyber-physical networks [18]–[20], social media platforms [21], cooperation of robots [22], [23], and Internet of Things [24]–[26].

B. Related Work

Team theory was established with the seminal work of Marschak [1], Radner [2], and Marschak and Radner [3] on static team problems, and with Witsenhausen [27], [28] on dynamic team problems. In static team problems [29], [30], the information received by the team members is not affected by the decisions of other team members [31], while, in dynamic team problems, the information of at least one team member is affected by the decisions of other members in the team [31]. If there is a prescribed order in which team members make decisions, then such a problem is called a sequential team problem. If, however, the team members make decisions in an order that depends on the realization of the team's uncertainty and decisions of other members, then such a problem is called a nonsequential team problem. Formulating a well-posed nonsequential team problem is more challenging as we need to ensure that the problem is causal and deadlock free [32]–[34]. Teneketzis [35] presented several results and open questions for nonsequential teams by using the framework of the Witsenhausen's intrinsic model [36]. In this article, we restrict our attention to sequential dynamic team decision problems.

The information structure in a sequential team decision problem designates who knows what about the status of the team and when [37], [38]. The information structure may designate the complexity [39]–[41] of the problem, and can lead to computational implications [42]. Witsenhausen [43] discussed several information structures and asserted some optimality results for team decision problems. Ho [44] investigated information structures within the context of team decision theory using a simple thematic example of a team consisting of two individuals who need to coordinate a meeting. More recently, Mahajan et al. [45] provided a tutorial paper with a comprehensive characterization of information structures.

Information structures are classified [37] as follows:

1) classical;

2) partially nested (alternatively also called overlapping, or quasiclassical);
3) nonclassical.

In classical information structures, all team members receive the same information and have perfect recall [46]–[49]. If there is only one team member, then such information structures are called strictly classical resulting in team decision problems that are typical centralized stochastic control problems [50], [51]. In partially nested information structures, there are some team members who have a nonempty intersection of their information structures while they have perfect recall. Any information structure that is not classical, or partially nested, is called nonclassical and can be further classified [45] as follows:

1) n -step delayed-sharing, $n \in \mathbb{R}$, where each team member has access to the information, i.e., observations and decisions, of other members after an n-step symmetric delay, i.e., same for all members [52], or asymmetric delay [53];
2) periodic sharing [54], where each team member has access to the information, i.e., observations and decisions, of the other members periodically;
3) delayed observation (or state [55]) sharing information, where each team member has access to the observations (or states if completely observable) of other members after an n-step symmetric, or asymmetric, delay;
4) delayed control sharing information, where each team member has access to the decisions of other members after an n-step symmetric, or asymmetric, delay [56];
5) no sharing information, where the team members do not share any information.

Sequential dynamic team problems with nonclassical information structures impose the following technical challenges [57]: 1) the functional optimization problem of selecting the optimal strategy is not trivial as the class of strategies is infinitely large, and 2) the data increase with time causing significant implications on storage requirements and real-time implementation. In centralized stochastic control theory, these difficulties are addressed by finding sufficient statistics to compress the growing data without loss of optimality [58] using a conditional probability of the state of the team at time t given all the data available up until time t. This conditional probability is called information state, and it takes values in a time-invariant space. Using this information state can help us derive results for optimal control strategies in a time-invariant domain. Results based on data which, even though they increase with time, are compressed to a sufficient statistic taking values in a time-invariant space are called structural results (see [59, p. 203]).

In centralized stochastic control, structural results can help us establish an information state, which does not depend on the control strategy, and thus, they are related to the concept of separation between estimation and control. An important consequence of this separation is that for any given choice of control strategies and a realization of the team's variables until time t, the information states at future times do not depend on the choice of the control strategy at time t but only on the realization

of the decision at time t (see [50, p. 84]). Thus, the future information states are separated from the choice of the current control strategy. The latter is necessary in order to formulate a classical dynamic program [60]–[62], where at each step the optimization problem is to find the optimal decision for a given realization of the information state [50].

Several structural results have been reported in the literature to date for team decision problems with nonclassical information structures [43], [52], [53], [63]–[68]. However, these results can lead to a sequential decomposition of the optimization problem over a space of functions [52], [69], [70] instead of a space of decisions to derive optimal strategies. This is due to the absence of separation between estimation and control which prevents the formulation of a classical dynamic program. There are three general approaches currently in the literature that, in conjunction with these structural results, can be used to derive optimal strategies in sequential dynamic team problems with nonclassical information structures:

1) the person-by-person approach;
2) the designer's approach;
3) the common information approach.

The person-by-person approach aims to convert the problem into a centralized stochastic control problem from the point of view of each team member. Namely, we arbitrarily fix the strategies for all team members except for one, say team member $k \in \mathcal{K}, \mathcal{K} = \{1, \ldots, K\}, K \in \mathbb{N}$. Then, we derive the optimal strategy for k given the strategies for all other members. We repeat this process for all team members until no member can improve the performance of the team by unilaterally changing their strategy. Thus, the resulting strategies are person-by-person optimal [44]. Other research efforts have taken a different approach using Girsanov's change of probability measure to transform the dynamic team problem to a static problem, in which the information structure is not affected by the members' decisions, and then applied the stochastic maximum principle to derive necessary and sufficient optimality conditions for both team and person-by-person strategies [71]–[75]. An optimal strategy of the team is necessarily person-by-person optimal (see [76, p. 195]). However, the converse is not in general true. In addition, if the team payoff function is concave polyhedral, i.e., piecewise linear and concave, then the person-by-person approach is not generally sufficient to determine an optimal strategy (see [76, p. 191]), although the problem can be reduced to a linear programming problem. Ho [77] showed that for a Gaussian team if the observation functions are linear and the cost function is quadratic, then affine control strategies are optimal. The person-by-person approach has been used in teams with broadcast information structures [78], in real-time communication using encoders and decoders [79]–[84], in quickest detection problems [85]–[88], and networked control systems [89], [90].

The designer's approach was first introduced by Witsenhausen [28], as a standard form for sequential stochastic control with a nonclassical information structure, and extended later by Mahajan [91] and Yüksel [92]. This approach addresses the team decision problem from the point of view of a "designer"

who knows the team's dynamics and statistics of all sources of uncertainties. Although sequential dynamic team decision problems are informationally decentralized, the designer's approach transforms the problem into a centralized, open-loop planning problem from the designer's point of view where the objective is to derive the strategy of the team before the team starts evolving. Therefore, no data are observed by the designer, and thus, this approach leads to a dynamic programming decomposition over a space of functions instead of decisions imposing significant computational implications [57]. The person-by-person approach has been used in conjunction with the designer's approach in real-time communication [80], [93], [94], and networked control systems [89], [90].

The common information approach [70], [95] was first presented for problems with partial history sharing [52], where the team members share a subset of their past observations and decisions to a shared memory accessible by all members of the team. The solution is derived by reformulating the problem from the viewpoint of a "coordinator" with access only to the shared information (the common information), whose task is to provide prescription strategies to each team member. At each time t, the prescription strategies of the team members map their private history of observations and decisions to their optimal decisions at t. The common information approach has been used in problems with control-sharing information structure [96], in stochastic games with asymmetric information [97], and in teams with mean-field sharing [98]. There are also some earlier papers that used similar ideas to analyze specific information structures, or structure of the team decision problem, e.g., teams with sequential decompositions [99], teams with partially nested information and common past [100], teams with delayed state sharing [55], teams with periodic sharing information structure [54], and teams with belief sharing information structure [101].

C. Contributions of This Article

In this article, we provide structural results and a classical dynamic programming decomposition of sequential dynamic team decision problems. We first address the problem from the point of view of a "manager" who seeks to derive the optimal strategy of a team in a centralized process. Then, we address the problem from the point of view of each team member, and show that the solution of each team member is the same as the one derived by the manager.

The contributions of this article are the induction of the following: 1) structural results for the team from the point of view of a manager, i.e., through a centralized process that yield an information state, which does not depend on the control strategy of the team (Theorem 1), and thus, it leads to a classical dynamic programming decomposition where the optimization problem is over the space of the team's decisions (Theorems 2 and 3); and 2) structural results for each team member that yield an information state which does not depend on their control strategy (Theorem 5), and thus, it leads to a classical dynamic programming decomposition where the optimization problem is over the space of the decisions of each team member. In addition, we show that the solution of each team member is the same as

the one derived by the manager (Theorem 7), and therefore, the team members do not need a centralized intervention.

D. Comparison With Related Work

The one feature which sharply distinguishes previous approaches, reported in Section I-B, from that undertaken here is that, in this article, we derive structural results aimed at establishing an information state that does not depend on the control strategy, and thus, we can institute separated control strategies that can lead to a classical dynamic programming decomposition. More specifically, the results in this article advance the state of the art in the following ways.

First, in contrast to the person-by-person optimal strategy [78]–[90], which is not always an optimal strategy of the team (see [76, p. 195]), our structural results for each team member (Theorem 5) guarantee that their optimal control strategies are also optimal for the team (Theorem 7).

Second, while our structural results from the point of view of a manager impose a centralized process, they yield an information state which does not depend on the control strategy of the team (Theorem 1), and thus, it can lead to a classical dynamic programming decomposition where the optimization problem is over the space of the team's decisions (Theorems 2 and 3). The designer's approach [28], [91], on the other hand, transforms the problem into a centralized, open-loop planning problem where the objective is to derive the strategy of the team before the team starts evolving. Therefore, no data are observed by the designer, and thus, this approach leads to a dynamic programming decomposition over a space of functions instead of decisions which has significant computational implications [57].

Finally, in contrast to the common information approach [52], [70], where the coordinator's problem is a centralized stochastic control problem [102] that leads to a dynamic programming decomposition where the optimization problem is over a space of functions, i.e., the prescription functions of the team members, our structural results from the manager's point of view lead to a dynamic programming decomposition where the optimization problem is over the space of the team's decisions (Theorems 2 and 3). In addition, our structural results for each team member yield an information state that leads to a dynamic programming decomposition for each team member resulting in a solution of each team member, which is the same as the one derived by the manager (Theorem 7), and thus, the team members do not need a centralized intervention.

E. Organization of This Article

The rest of this article organized as follows. In Section II, we provide the modeling framework, information structure, and the optimization problem of a team. In Section III, we derive structural results for the team from the point of view of a manager, and a dynamic programming decomposition where the optimization problem is over the space of the team's decisions. In Section IV, we derive structural results for each team member, and a dynamic programming decomposition where the optimization problem is over the space of the decisions of each team member. In Section V, we present an example of a dynamic team

with a delayed sharing information structure consisting of two members. This example was used by Varaiya and Walrand [63] to show that Witsenhausen's structural result asserted in his seminal paper [43] is suboptimal. Finally, Section VI concludes this article.

II. PROBLEM FORMULATION

A. Notation

Subscripts denote time and superscripts index subsystems. We denote random variables with upper case letters, and their realizations with lower case letters, e.g., for a random variable X_t, x_t denotes its realization. The shorthand notation $X_t^{1:K}$ denotes the vector of random variables $(X_t^1, X_t^2, \ldots, X_t^K)$, $x_t^{1:K}$ denotes the vector of their realization $(x_t^1, x_t^2, \ldots, x_t^K)$, and $h_t^{1:K}(\cdot, \cdot)$ denotes the vector of functions $(h_t^1(\cdot, \cdot), \ldots, h_t^K(\cdot, \cdot))$. The expectation of a random variable is denoted by $\mathbb{E}[\cdot]$, the probability of an event is denoted by $\mathbb{P}(\cdot)$, and the probability density function is denoted by $p(\cdot)$. For a control strategy $\mathbf{g}$, we use $\mathbb{E}^{\mathbf{g}}[\cdot]$, $\mathbb{P}^{\mathbf{g}}(\cdot)$, and $p^{\mathbf{g}}(\cdot)$ to denote that the expectation, probability, and probability density function, respectively, depend on the choice of the control strategy $\mathbf{g}$. For two measurable spaces $(\mathcal{X}, \mathscr{X})$ and $(\mathcal{Y}, \mathscr{Y})$, $\mathscr{X} \otimes \mathscr{Y}$ is the product σ-algebra on $\mathcal{X} \times \mathcal{Y}$ generated by the collection of all measurable rectangles, i.e., $\mathscr{X} \otimes \mathscr{Y} := \sigma(\{A \times B : A \in \mathscr{X}, B \in \mathscr{Y}\})$. The product of $(\mathcal{X}, \mathscr{X})$ and $(\mathcal{Y}, \mathscr{Y})$ is the measurable space $(\mathcal{X} \times \mathcal{Y}, \mathscr{X} \otimes \mathscr{Y})$.

B. Modeling Framework

We consider a team of $K \in \mathbb{N}$ members with a measurable state space $(\mathcal{X}_t, \mathscr{X}_t)$, where $\mathcal{X}_t$ is the set in which the team's state takes values at time $t = 0, 1, \ldots, T - 1$, $T \in \mathbb{N}$, and $\mathscr{X}_t$ is the associated σ-algebra. The state of the team is represented by a random variable $X_t : (\Omega, \mathscr{F}) \to (\mathcal{X}_t, \mathscr{X}_t)$, defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$, where Ω is the sample space, $\mathscr{F}$ is the associated σ-algebra, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathscr{F})$. The decision of each team member $k \in \mathcal{K}$, $\mathcal{K} = \{1, \ldots, K\}$, is represented by a random variable $U_t^k : (\Omega, \mathscr{F}) \to (\mathcal{U}_t^k, \mathscr{U}_t^k)$, defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and takes values in the measurable space $(\mathcal{U}_t^k, \mathscr{U}_t^k)$, where $\mathcal{U}_t^k$ is team member ks nonempty feasible set of actions at time t and $\mathscr{U}_t^k$ is the associated σ-algebra. Let $U_t^{1:K} = (U_t^1, \ldots, U_t^K)$ be the team's decision at time t. Starting at the initial state X_0, the evolution of the team is described by the state equation

$$X_{t+1} = f_t\left(X_t, U_t^{1:K}, W_t\right) \tag{1}$$

where W_t is a random variable defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ that corresponds to the external, uncontrollable disturbance to the team and takes values in a measurable set $(\mathcal{W}, \mathscr{W})$, i.e., $W_t : (\Omega, \mathscr{F}) \to (\mathcal{W}, \mathscr{W})$. $\{W_t : t = 0, \ldots, T - 1\}$ is a sequence of independent random variables that are also independent of the initial state X_0. At time $t = 0, 1, \ldots, T - 1$, every team member $k \in \mathcal{K}$ makes an observation Y_t^k, which takes values in a measurable set $(\mathcal{Y}^k, \mathscr{Y}^k)$, described by the observation equation

$$Y_t^k = h_t^k(X_t, Z_t^k) \tag{2}$$

where Z_t^k is a random variable defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ that corresponds to the noise of each member's sensor and takes values in a measurable set $(\mathcal{Z}^k, \mathscr{Z}^k)$, i.e., $Z_t^k : (\Omega, \mathscr{F}) \to (\mathcal{Z}^k, \mathscr{Z}^k)$. $\{Z_t^k : t = 0, \ldots, T - 1; k = 1, \ldots, K\}$ is a sequence of independent random variables that are also independent of the initial state X_0 and $\{W_t : t = 0, \ldots, T - 1\}$.

C. Nonclassical Information Structures

The team has a nonclassical information structure that can be classified as follows.

1) n-Step Delayed Information Sharing: In this case, at time t, team member $k \in \mathcal{K}$ observes Y_t^k, and the n-step, $n \in \mathbb{R}$, past observations $Y_{0:t-n}^{1:K}$, and decisions $U_{0:t-n}^{1:K}$ of the entire team. Thus, at time t, the data available to member k consist of the data Δ_t available to all team members, i.e.,

$$\Delta_t := (Y_{0:t-n}^{1:K}, U_{0:t-n}^{1:K}) \tag{3}$$

where $Y_{0:t-n}^{1:K} = \{Y_{0:t-n}^1, \ldots, Y_{0:t-n}^K\}$, $U_{0:t-n}^{1:K} = \{U_{0:t-n}^1, \ldots, U_{0:t-n}^K\}$, and the data Λ_t^k known only to member $k \in \mathcal{K}$, i.e.,

$$\Lambda_t^k := (Y_{t-n+1:t}^k, U_{t-n+1:t-1}^k). \tag{4}$$

The n-step delayed information sharing can also be asymmetric [53], i.e., for each member $k \in \mathcal{K}$, $Y_{t-n_k}^k$, $U_{t-n_k}^k$, where $n_k \in \mathbb{R}$, is constant but not necessarily the same for each k.

2) Periodic Information Sharing With Period $\omega \geq 1$: In this case, for $\alpha = 1, 2, \ldots$ and $\alpha\omega < t \leq (\alpha + 1)\omega$, the pair of Δ_t and Λ_t^k, $k \in \mathcal{K}$, becomes

$$\Delta_t := (Y_{0:\alpha\omega}^{1:K}, U_{0:\alpha\omega}^{1:K}) \tag{5}$$

$$\Lambda_t^k := (Y_{\alpha\omega+1:(\alpha+1)\omega}^k, U_{\alpha\omega+1:(\alpha+1)\omega}^k). \tag{6}$$

3) n-Step Delayed Observation Sharing: In this case, Δ_t and Λ_t^k, $k \in \mathcal{K}$, become

$$\Delta_t := (Y_{0:t-n}^{1:K}) \tag{7}$$

$$\Lambda_t^k := (Y_{t-n+1:t}^k, U_{0:t-1}^k). \tag{8}$$

4) n-Step Delayed Control Sharing: In this case, Δ_t and Λ_t^k, $k \in \mathcal{K}$, become

$$\Delta_t := (U_{0:t-n}^{1:K}) \tag{9}$$

$$\Lambda_t^k := (Y_{0:t}^k, U_{t-n+1:t-1}^k). \tag{10}$$

5) No Sharing Information: In this case, Δ_t and Λ_t^k, $k \in \mathcal{K}$, become

$$\Delta_t := \emptyset \tag{11}$$

$$\Lambda_t^k := (Y_{0:t}^k, U_{0:t-1}^k). \tag{12}$$

The collection $\{(\Delta_t, \Lambda_t^k); k \in \mathcal{K}; t = 0, \ldots, T - 1\}$, is the information structure of the team and captures who knows what about the status of the team and when.

In our exposition, we consider that the team imposes an n-step delayed information sharing, which can be deemed as the general case of a nonclassical information structure. However, in what follows, the results hold for any special case (2)–(5) above and corresponding n.

D. Optimization Problem

Let $(\mathcal{D}_t, \mathscr{D}_t)$ and $(\mathcal{L}_t^k, \mathscr{L}_t^k), k \in \mathcal{K}$, be the measurable spaces of all possible realizations of Δ_t and Λ_t^k, respectively, where $\mathscr{D}_t$ and $\mathscr{L}_t^k$ are the associated σ-algebras. Each team member k makes a decision

$$U_t^k = g_t^k(\Lambda_t^k, \Delta_t) \tag{13}$$

where g_t^k is a control law of $k \in \mathcal{K}$, which is a measurable function $g_t^k : (\mathcal{L}_t^k \times \mathcal{D}_t, \mathscr{L}_t^k \otimes \mathscr{D}_t) \to (\mathcal{U}_t^k, \mathscr{U}_t^k)$. The control strategy of team member $k \in \mathcal{K}$ is $\mathbf{g}^k = \{g_t^k; t = 0, \dots, T-1\}, \mathbf{g}^k \in \mathcal{G}^k$, where $\mathcal{G}^k$ is the feasible set of the control strategies for k. Thus, the set of feasible decentralized control strategies is $\mathcal{G}^{\mathrm{Dec}} = \times_{k \in \mathcal{K}} \mathcal{G}^k$, i.e., $\mathbf{g} = \{\mathbf{g}^1, \dots, \mathbf{g}^K\} \in \mathcal{G}^{\mathrm{Dec}}$. If $\mathbf{g} \in \mathcal{G}$ is a centralized control strategy then $\mathcal{G} =: (\mathcal{L}_t^1 \times \cdots \times \mathcal{L}_t^K \times \mathcal{D}_t, \mathscr{L}_t^1 \otimes \cdots \otimes \mathscr{L}_t^K \otimes \mathscr{D}_t)$.

Problem 1: The problem is to derive the optimal control strategy $\mathbf{g}^*$ of the team that minimizes the expected total cost

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[\sum_{t=0}^{T-1} c_t(X_t, U_t^{1:K}) + c_T(X_T) \right] \tag{14}$$

where the expectation is with respect to the joint probability distribution of the random variables X_t and $U_t^{1:K}$ designated by the choice of $\mathbf{g}$, $c_t(X_t, U_t^{1:K}) : (\mathcal{X}_t \times \prod_{k \in \mathcal{K}} \mathcal{U}_t^k, \mathscr{X}_t \otimes \mathscr{U}_t^1 \otimes \dots \otimes \mathscr{U}_t^K) \to \mathbb{R}$ is the team's measurable cost function, and $c_T(\hat{X}_T) : (\mathcal{X}_T, \mathscr{X}_T) \to \mathbb{R}$ is the measurable cost function at T.

The statistics of the primitive random variables X_0, $\{W_t : t = 0, \dots, T-1\}$, $\{Z_t^k : k \in \mathcal{K}; t = 0, \dots, T-1\}$, the state equations $\{f_t : t = 0, \dots, T-1\}$, the observation equations $\{h_t^k : k \in \mathcal{K}; t = 0, \dots, T-1\}$, and the cost functions $\{c_t : t = 0, \dots, T\}$ are all known.

III. STRUCTURAL RESULTS FOR THE TEAM

We start our exposition by addressing Problem 1 from the point of view of a manager who seeks to derive the optimal strategy $\mathbf{g} \in \mathcal{G}$ of the team.

A. Information State—Team

The first step is to identify an appropriate information state for the team that can be used to formulate a classical dynamic programming decomposition for Problem 1.

Definition 1: An information state, Π_t, for the team described by the state equation (1) (a) is a function of $(\Delta_t, \Lambda_t^{1:K})$, and (b) Π_{t+1} can be determined from Π_t, $Y_{t+1}^{1:K}$, and $U_t^{1:K}$.

The notation is simpler if we consider densities for all probability distributions. Let $\mathbf{g} \in \mathcal{G}$ be a control strategy and $(\Delta_t, \Lambda_t^{1:K})$ be the information structure of the team. To proceed, we first need to prove some essential properties of the conditional probability densities related to the observations of the team members and team's state.

Lemma 1: For any control strategy $\mathbf{g} \in \mathcal{G}$ of the team

$$p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = p(Y_{t+1}^{1:K} \mid X_{t+1}) \tag{15}$$

for all $t = 0, 1, \dots, T-1$.

Proof: The realization of $Y_{t+1}^{1:K}$ is statistically determined by the conditional distribution of $Y_{t+1}^{1:K}$ given X_{t+1} in (2), hence

$$p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}). \tag{16}$$

However,

$$p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}) = p^{\mathbf{g}}(Z_{t+1}^{1:K} \in \prod_{k \in \mathcal{K}} B^k \mid X_{t+1}) \tag{17}$$

where $B^k \in \mathscr{Z}^k$, $k \in \mathcal{K}$. Since, $\{Z_t^k : k = 1, \dots, K; t = 0, \dots, T-1\}$ is a sequence of independent random variables that are independent of X_{t+1}

$$p^{\mathbf{g}}(Z_{t+1}^{1:K} \in \prod_{k \in \mathcal{K}} B^k \mid X_{t+1}) = p(Z_{t+1}^{1:K} \in \prod_{k \in \mathcal{K}} B^k). \tag{18}$$

Thus

$$p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}) = p(Y_{t+1}^{1:K} \mid X_{t+1}). \tag{19}$$

The result follows from (16) and (19). $\square$

Lemma 2: For any control strategy $\mathbf{g} \in \mathcal{G}$ of the team

$$p^{\mathbf{g}}(X_{t+1} \mid X_t, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = p(X_{t+1} \mid X_t, U_t^{1:K}) \tag{20}$$

for all $t = 0, 1, \dots, T-1$.

Proof: The realization of X_{t+1} is statistically determined by the conditional distribution of X_{t+1} given X_t and $U_t^{1:K}$, i.e., by $p^{\mathbf{g}}(X_{t+1} \mid X_t, U_t^{1:K})$. From (1), we have

$$p^{\mathbf{g}}(X_{t+1} \mid X_t, U_t^{1:K}) = p^{\mathbf{g}}(W_t \in A \mid X_t, U_t^{1:K}) \tag{21}$$

where $A \in \mathscr{W}$. Since, $\{W_t : t = 0, \dots, T-1\}$ is a sequence of independent random variables that are independent of X_t and $U_t^{1:K}$

$$p^{\mathbf{g}}(W_t \in A \mid X_t, U_t^{1:K}) = p(W_t \in A). \tag{22}$$

Next

$$p^{\mathbf{g}}(X_{t+1} \mid X_t, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$
$$= p^{\mathbf{g}}(W_t \in A \mid X_t, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = p(W_t \in A). \tag{23}$$

The result follows from (21)–(23). $\square$

Lemma 3: For any control strategy $\mathbf{g} \in \mathcal{G}$ of the team

$$p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_t^{1:K}) = p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \tag{24}$$

for all $t = 0, 1, \dots, T-1$.

Proof: We have

$$p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_t^{1:K})$$
$$= p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_{t-2}^{1:K}, Y_{t-1}^{1:K}, Y_t^{1:K}, U_{t-2}^{1:K}, U_{t-1}^{1:K}). \tag{25}$$

However, the realization of X_t is statistically determined by the conditional distribution of X_t given X_{t-1} and $U_{t-1}^{1:K}$, which does not depend on the control strategy $\mathbf{g}$ (Lemma 2), so we can drop the superscript in (25), and thus (24) follows immediately. $\square$

Remark 1: As a consequence of Lemma 3, and since X_t does not depend on $U_t^{1:K}$, we have

$$p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = p(X_t \mid \Delta_t, \Lambda_t^{1:K}). \tag{26}$$

Given that the manager can observe the data $(\Delta_t, \Lambda_t^{1:K})$ of the team, our hypothesis is that we can compress these data to a sufficient statistic of the state of the team. This statistic is the probability density function $p(X_t \mid \Delta_t, \Lambda_t^{1:K})$. The next result proves our hypothesis and shows that such information state does not depend on the team's control strategy.

Theorem 1 (Information State—Team): For any control strategy $\mathbf{g} \in \mathcal{G}$ of the team, the conditional probability density $p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_t^{1:K})$ does not depend on the control strategy $\mathbf{g}$. It is an information state $\Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t)$, i.e., $\Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t) = p(X_t \mid \Delta_t, \Lambda_t^{1:K})$ with $\int_{\mathscr{X}_t} \Pi_t(\Delta_t,$

$\Lambda_t^{1:K})(X_t)dX_t = 1$ that can be evaluated from $\Delta_t, \Lambda_t^{1:K}$. Moreover, there is a function θ_t, which does not depend on the control strategy $\mathbf{g}$, such that

$$\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1})$$

$$= \theta_t \left[\Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t), Y_{t+1}^{1:K}, U_t^{1:K} \right] \quad (27)$$

for all $t = 0, 1, \ldots, T - 1$.

Proof: See Appendix A. □

Note that the information state $\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1}) = p(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$ of the team is the entire probability density function and not just its value at any particular realization of $(\Delta_{t+1}, \Lambda_{t+1}^{1:K})$. This is because to compute $\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1})$ for any particular X_{t+1}, we need the probability density functions $p(\cdot \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$ and $p(\cdot \mid \Delta_t, \Lambda_t^{1:K})$. This implies that the information state takes values in the space of these probability densities on the measurable space $(\mathcal{X}_t, \mathscr{X}_t)$, which is an infinite-dimensional space.

B. Optimal Control Strategy of the Team

In what follows, to simplify notation, the information state $\Pi_t(\Delta_t, \Lambda_t^{1:K})$ of the team at t is denoted simply by Π_t. We use its arguments Δ_t and $\Lambda_t^{1:K}$ only if it is required by our exposition.

Definition 2: A control strategy $\mathbf{g} = \{g_t; t = 0, \ldots, T-1\}$ is said to be *separated* if g_t depends on Δ_t and $\Lambda_t^{1:K}$ only through the information state, i.e., $U_t^{1:K} = g_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$. Let $\mathcal{G}^s \subseteq \mathcal{G}$ denote the set of all separated control strategies.

In implementing a separated control strategy, we first need to compute the conditional probability $\Pi_t(\Delta_t, \Lambda_t^{1:K})$, and then choose the control, since the task of estimation and control are separated. Next, we use the information state to define recursive functions which are analogous to the comparison principle (see [50, p. 74]).

Theorem 2: Let $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ be functions defined recursively for all $\mathbf{g} \in \mathcal{G}^s$ by

$$V_T\left(\Pi_T(\Delta_T, \Lambda_T^{1:K})\right) := \mathbb{E}^{\mathbf{g}} \left[c_T(X_T) \mid \Pi_T = \pi_T\right] \quad (28)$$

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) := \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K})\right.$$

$$+ V_{t+1}\left(\theta_t\left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K}\right]\right) \mid \Pi_t = \pi_t$$

$$\left. U_t^{1:K} = u_t^{1:K}\right] \quad (29)$$

where $c_T(X_T)$ is the cost function at T, and π_T, π_t, $u_t^{1:K}$ are the realizations of Π_T, Π_t, and $U_t^{1:K}$, respectively. Then, for any control strategy $\mathbf{g} \in \mathcal{G}$,

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) \le J_t(\mathbf{g}) := \mathbb{E}^{\mathbf{g}} \left[\sum_{l=t}^{T-1} c_l(X_l, U_l^{1:K})\right.$$

$$\left. + c_T(X_T) \mid \Delta_t, \Lambda_t^{1:K}\right] \quad (30)$$

where $J_t(\mathbf{g})$ is the cost-to-go of the team at time t corresponding to the control strategy $\mathbf{g} \in \mathcal{G}$.

Proof: We prove (30) by induction. For $t = T$

$$J_T(\mathbf{g}) := \mathbb{E}^{\mathbf{g}} \left[c_T(X_T) \mid \Delta_T, \Lambda_T^{1:K}\right]$$

$$= \int_{\mathscr{X}_T} c_T(X_T) \Pi_T(\Delta_T, \Lambda_T^{1:K})(X_T) \, dX_T \quad (31)$$

and so (30) holds with equality.

Suppose that (30) holds for $t + 1$. Then

$$J_t(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[\sum_{l=t}^{T-1} c_l(X_l, U_l^{1:K}) + c_T(X_T) \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$= \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^{1:K})\right.$$

$$\left. + c_T(X_T) \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$= \mathbb{E}^{\mathbf{g}} \left[\mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K})\right.\right.$$

$$\left. + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^{1:K}) + c_T(X_T) \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}\right]$$

$$\left. \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$\ge \mathbb{E}^{\mathbf{g}} \left[\mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K})\right.\right.$$

$$+ V_{t+1}\left(\theta_t\left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K}\right]\right) \mid \Pi_t$$

$$\left. \times (\Delta_t, \Lambda_t^{1:K}) \, U_t^{1:K}\right] \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$= \mathbb{E}^{\mathbf{g}} \left[V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$= V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) \quad (32)$$

where in the inequality, we used the hypothesis and, in the last equality, we used (29). Thus, (30) holds for all t. □

In view of Theorem 2, we show that an optimal strategy of the team is separated and obtain a classical dynamic programming decomposition where the optimization problem is over the space of the team's decisions.

Theorem 3: Let

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) := \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K})\right.$$

$$+ V_{t+1}\left(\theta_t\left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K}\right]\right) \mid \Pi_t = \pi_t$$

$$\left. U_t^{1:K} = u_t^{1:K}\right] \quad (33)$$

and let $\mathbf{g} \in \mathcal{G}^s$ be a separated control strategy that achieves the infimum in (33). Then, $\mathbf{g} \in \mathcal{G}^s$ is optimal and

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) = J_t(\mathbf{g}) \quad (34)$$

with probability 1.

Proof: We first prove (34) by induction. For $t = T$

$$J_T(\mathbf{g}) := \mathbb{E}^{\mathbf{g}} \left[c_T(X_T) \mid \Delta_T, \Lambda_T^{1:K}\right]$$

$$= \int_{\mathscr{X}_T} c_T(X_T) \Pi_T(\Delta_T, \Lambda_T^{1:K})(X_T) dX_T. \quad (35)$$

Suppose that (33) holds for $t + 1$. Then

$$\inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[\sum_{l=t}^{T-1} c_l(X_l, U_l^{1:K}) + c_T(X_T) \mid \Delta_t, \Lambda_t^{1:K}\right]$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^{1:K}) \right.$$
$$+ c_T(X_T) \,|\, \Delta_t, \Lambda_t^{1:K} \Big]$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[\mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) \right.\right.$$
$$\left. + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^{1:K}) + c_T(X_T) | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K} \right]$$
$$\left. | \Delta_t, \Lambda_t^{1:K} \right]$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[\mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) \right.\right.$$
$$+ V_{t+1} \left(\theta_t \left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K} \right] \right) | \Pi_t(\Delta_t, \Lambda_t^{1:K})$$
$$\left. U_t^{1:K} \right] | \Delta_t, \Lambda_t^{1:K} \Big]$$

$$= \mathbb{E}^{\mathbf{g}} \left[V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right) | \Delta_t, \Lambda_t^{1:K} \right] = V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right)$$
$$\tag{36}$$

where in the third equality, we used the hypothesis and, in the forth equality, $u_t^{1:K}$ achieves the infimum. Thus, (33) holds for all t.

For $t = 0$, (34) yields $J_0(\mathbf{g}) = V_0(\Pi_0(\Delta_0, \Lambda_0^{1:K}))$. Taking expectations

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[J_0(\mathbf{g}) \right] = \mathbb{E}^{\mathbf{g}} \left[V_0 \left(\Pi_0(\Delta_0, \Lambda_0^{1:K}) \right) \right]. \tag{37}$$

By Theorem 2, it follows that for any other $\mathbf{g}' \in \mathcal{G}$

$$J(\mathbf{g}') \geq \mathbb{E}^{\mathbf{g}} \left[V_0 \left(\Pi_0(\Delta_0, \Lambda_0^{1:K}) \right) \right]. \tag{38}$$

The following results are derived by using ideas from [103] and [104].

Lemma 4: Let $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ be functions defined recursively for all $\mathbf{g} \in \mathcal{G}^s$ by

$$V_T \left(\Pi_T(\Delta_T, \Lambda_T^{1:K}) \right) := \mathbb{E}^{\mathbf{g}} \left[c_T(X_T) | \Pi_T = \pi_T \right] \tag{39}$$

$$V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right) := \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) \right.$$
$$+ V_{t+1} \left(\theta_t \left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K} \right] \right) | \Pi_t = \pi_t$$
$$\left. U_t^{1:K} = u_t^{1:K} \right] \tag{40}$$

where $c_T(X_T)$ is the cost function at T, and π_T, π_t, $u_t^{1:K}$ are the realizations of Π_T, Π_t, and $U_t^{1:K}$, respectively. Then, for all $t = 0, \ldots, T$, $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ is positive homogeneous, i.e., for any $\rho > 0$, $V_t(\rho\,\Pi_t(\Delta_t, \Lambda_t^{1:K})) = \rho\,V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$.

Proof: See Appendix B. $\square$

Theorem 4: Let

$$V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right) := \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}} \left[c_t(X_t, U_t^{1:K}) \right.$$
$$+ V_{t+1} \left(\theta_t \left[\Pi_t(\Delta_t, \Lambda_t^{1:K}), Y_{t+1}^{1:K}, U_t^{1:K} \right] \right) | \Pi_t = \pi_t$$
$$\left. U_t^{1:K} = u_t^{1:K} \right] \tag{41}$$

where $c_T(X_T)$ is the cost function at T, and π_T, π_t, $u_t^{1:K}$ are the realizations of Π_T, Π_t, and $U_t^{1:K}$, respectively. Then, for

all $t = 0, \ldots, T$, $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ is concave with respect to $\Pi_t(\Delta_t, \Lambda_t^{1:K})$.

Proof: See Appendix C. $\square$

IV. STRUCTURAL RESULTS FOR THE TEAM MEMBERS

In this section, we address Problem 1 from the point of view of a team member $k \in \mathcal{K}$ who seeks to derive their optimal strategy $\mathbf{g}^k = \{g_t^k; k \in \mathcal{K}; t = 0, \ldots, T - 1\}$, which will constitute the team's strategy $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\} \in \mathcal{G}^{Dec}$.

A. Information State—Team Members

We first identify an appropriate information state for member $k \in \mathcal{K}$ that can be used to formulate a classical dynamic programming decomposition for Problem 1.

Definition 3: An information state, Π_t^k, for member $k \in \mathcal{K}$ of a team described by the state equation (1) (a) is a function of (Δ_t, Λ_t^k), and (b) Π_{t+1}^k can be determined from Π_t^k, $Y_{t-n+1}^{1:K}$, $U_{t-n+1}^{1:K}$, Y_{t+1}^k, and U_t^k.

To proceed, we first need to prove some essential properties of the conditional probabilities densities related to the observation of team member $k \in \mathcal{K}$ and the team's state.

Lemma 5: For any control strategy $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\}$ of the team

$$p^{\mathbf{g}}(Y_{t+1}^k \,|\, X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k) = p(Y_{t+1}^k \,|\, X_{t+1}) \tag{42}$$

for all $t = 0, 1, \ldots, T - 1$.

Proof: The realization of Y_{t+1}^k is statistically determined by the conditional distribution of Y_{t+1}^k given X_{t+1} in (2), hence

$$p^{\mathbf{g}}(Y_{t+1}^k \,|\, X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k) = p^{\mathbf{g}}(Y_{t+1}^k \,|\, X_{t+1}). \tag{43}$$

However

$$p^{\mathbf{g}}(Y_{t+1}^k \,|\, X_{t+1}) = p^{\mathbf{g}}(Z_{t+1}^k \in B^k \,|\, X_{t+1}) \tag{44}$$

where $B^k \in \mathscr{Z}^k$. Since, $\{Z_t^k : t = 0, \ldots, T - 1; k = 1, \ldots, K\}$ is a sequence of independent random variables that are independent of X_{t+1}, we have

$$p^{\mathbf{g}}(Z_{t+1}^k \in B^k \,|\, X_{t+1}) = p(Z_{t+1}^k \in B^k). \tag{45}$$

Thus

$$p^{\mathbf{g}}(Y_{t+1}^k \,|\, X_{t+1}) = p(Y_{t+1}^k \,|\, X_{t+1}). \tag{46}$$

The result follows from (43) and (46). $\square$

Lemma 6: For any control strategy $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\}$ of the team, $p^{\mathbf{g}}(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$ does not depend on the control strategy $\mathbf{g}^k$ of member k. It depends only on the strategy $\mathbf{g}^{-k} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$, of the other team members, i.e.,

$$p^{\mathbf{g}}(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k) = p^{\mathbf{g}^{-k}}(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k) \tag{47}$$

for all $t = 0, 1, \ldots, T - 1$.

Proof: Since $X_{t+1}^k = f_t(X_t, U_t^{1:K}, W_t)$ and $\Lambda_t^k := (Y_{t-n+1:t}^k, U_{t-n+1:t-1}^k)$, and $Y_t^k = h_t^k(X_t, Z_t^k)$ we have

$$p^{\mathbf{g}}(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$= p^{\mathbf{g}}(W_t \in A, U_t^{-k}, Z_t^k \in B^k \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$= p^{\mathbf{g}}(W_t \in A \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k) \cdot p^{\mathbf{g}}(U_t^{-k} \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$\cdot p^{\mathbf{g}}(Z_t^k \in B^k \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$= p(W_t \in A \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k) \cdot p^{\mathbf{g}^{-k}}(U_t^{-k} \,|\, \Delta_{t+1})$$

$$\cdot\, p(Z_t^k \in B^k \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k) \qquad (48)$$

where $A \in \mathscr{W}$, $B^k \in \mathscr{Z}^k$, and $U_t^{-k} = (U_t^1, \ldots, U_t^{k-1},$ $U_t^{k+1}, \ldots, U_t^K)$. In the last equality, the second term depends only on $\mathbf{g}^{-k}$ while we dropped the superscript $\mathbf{g}$ in $p(W_t \in A \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$ and $p(Z_t^k \in B^k \,|\, \Delta_{t+1}, \Lambda_t^k, U_t^k)$ since both W_t and Z_t^k are two sequences of independent random variables, and the evolution of their probability measure does not depend on the control strategy.

Theorem 5 (Information State—Team Members): For any control strategy $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\} \in \mathcal{G}^{\mathrm{Dec}}$ of the team, the conditional probability density $p(X_t \,|\, \Delta_t, \Lambda_t^k)$ does not depend on the control strategy $\mathbf{g}^k$ of member k. It depends only on the strategy $\mathbf{g}^{-k} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$ of the other team members. It is an information state of the team member k, i.e., $\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t) = p^{\mathbf{g}^{-k}}(X_t \,|\, \Delta_t, \Lambda_t^k)$, that can be evaluated from Δ_t and Λ_t^k. Moreover, there is a function θ_t^k, which does not depend on the control strategy $\mathbf{g}^k$ of member k, such that

$$\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1})$$
$$= \theta_t^k\left[\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), \Delta_{t+1}, \Lambda_{t+1}^k\right] \qquad (49)$$

for all $t = 0, 1, \ldots, T-1$.

Proof: See Appendix D. □

Note that the information state $\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) = p(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_{t+1}^k)$ of member k is the entire probability density function and not just its value at any particular realization of $(\Delta_{t+1}, \Lambda_{t+1}^k)$. This is because to compute $\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1})$ for any particular X_{t+1}, we need the probability density functions $p(\cdot \,|\, \Delta_t, \Lambda_t^k, U_t^k)$ and $p(\cdot \,|\, \Delta_t, \Lambda_t^k)$. This implies that the information state takes values in the space of these probability densities on the measurable space $(\mathcal{X}_t, \mathscr{X}_t)$, which is an infinite-dimensional space.

Lemma 7: The information state of the team $\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1})$ is a function of the information state $\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1})$ of each team member $k \in \mathcal{K}$, Δ_{t+1}, and $\Lambda_{t+1}^{1:K}$ for all $t = 0, 1, \ldots, T-1$.

Proof: By applying Bayes' rule, for all $t = 0, 1, \ldots, T-1$, we have

$$\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1}) = p(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \frac{p(\Lambda_{t+1}^{-k} \,|\, X_{t+1}, \Delta_{t+1}, \Lambda_{t+1}^k)\, p(X_{t+1} \,|\, \Delta_{t+1}, \Lambda_{t+1}^k) \cdot p(\Delta_{t+1}, \Lambda_{t+1}^k)}{p(\Delta_{t+1}, \Lambda_{t+1}^{1:K})}$$

$$= \frac{p(\Lambda_{t+1}^{-k} \,|\, X_{t+1}, \Delta_{t+1}, \Lambda_{t+1}^k)\, \Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) \cdot p(\Delta_{t+1}, \Lambda_{t+1}^k)}{p(\Delta_{t+1}, \Lambda_{t+1}^{1:K})} \qquad (50)$$

where $\Lambda_t^{-k} = (\Lambda_t^1, \ldots, \Lambda_t^{k-1}, \Lambda_t^{k+1}, \ldots, \Lambda_t^K)$. Thus, for each team member $k \in \mathcal{K}$, we can select an appropriate function ϑ_t^k such that

$$\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1})$$
$$= \vartheta_{t+1}^k\left(\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}), \Delta_{t+1}, \Lambda_{t+1}^{1:K}\right). \qquad (51)$$

Corollary 1: For each team member $k \in \mathcal{K}$, the function ϑ_{t+1}^k is increasing with $\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)$.

Proof: The function ϑ_{t+1}^k is continuous and differentiable in $[0, 1]$, while the variation of ϑ_{t+1}^k with respect to $\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)$ is positive. □

B. Optimal Control Strategy of the Team Members

In view of Theorem 5, we show that the optimal separated control strategy $\mathbf{g}^k = \{g_0^k, \ldots, g_{T-1}^k\}$, i.e., $U_t^k = g_t^k(\Pi_t^k(\Delta_t, \Lambda_t^k))$, derived by each team member $k \in \mathcal{K}$ yields the same solution as the one by the manager's optimal separated control strategy $\mathbf{g} \in \mathcal{G}^s$ (Theorem 7), and thus, the team members do not need a centralized intervention. We obtain a classical dynamic programming decomposition for member $k \in \mathcal{K}$ over the space of their decisions.

Lemma 8: Let $\mathbf{g}^k$ be a separated control strategy of team member $k \in \mathcal{K}$ and $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\}$ be the team's control strategy. We fix $\mathbf{g}^{-k} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$, and let $V_t^k(\Pi_t(\Delta_t, \Lambda_t^k))$ be functions defined recursively for all $\mathbf{g}$ by

$$V_T^k\left(\Pi_T(\Delta_T, \Lambda_T^k)\right) := \mathbb{E}^{\mathbf{g}}\left[c_T(X_T) \,|\, \Pi_T^k = \pi_T^k\right] \qquad (52)$$

$$V_t^k\left(\Pi_t(\Delta_t, \Lambda_t^k)\right) := \inf_{u_t^k \in \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}}\Big[c_t(X_t, U_t^k, U_t^{-k}) + V_{t+1}^k\big(\Pi_{t+1}$$
$$(\Delta_{t+1}, \Lambda_{t+1}^k)\big) \,|\, \Pi_t^k = \pi_t^k, \Delta_t = \delta_t, \Lambda_t^k = \lambda_t^k, U_t^k = u_t^k\Big] \qquad (53)$$

where $c_T(X_T)$ is the cost function at T, and π_t^k, δ_t, λ_t^k, u_t^k are the realizations of Π_t^k, Δ_t, Λ_t^k, and U_t^k, respectively. Then, $V_t^k(\Pi_t(\Delta_t, \Lambda_t^k))$ is positive homogeneous, i.e., for any $\rho > 0$, $V_t^k(\rho\, \Pi_t(\Delta_t, \Lambda_t^k)) = \rho\, V_t^k(\Pi_t(\Delta_t, \Lambda_t^k))$.

Proof: The proof is similar to the proof of Lemma 4. □

Theorem 6: Let $\mathbf{g}^k$ be a separated control strategy of team member $k \in \mathcal{K}$ and $\mathbf{g} = \{\mathbf{g}^1, \ldots, \mathbf{g}^K\}$ be the team's control strategy. We fix $\mathbf{g}^{-k} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$, and let $V_t^k(\Pi_t(\Delta_t, \Lambda_t^k))$ be functions defined recursively for all $\mathbf{g}$ by

$$V_t^k\left(\Pi_t(\Delta_t, \Lambda_t^k)\right) := \inf_{u_t^k \in \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}}\Big[c_t(X_t, U_t^k, U_t^{-k})$$
$$+ V_{t+1}^k\left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)\right) \,|\, \Pi_t^k = \pi_t^k, \Delta_t = \delta_t$$
$$\Lambda_t^k = \lambda_t^k, U_t^k = u_t^k\Big] \qquad (54)$$

where π_t^k, δ_t, λ_t^k, u_t^k are the realizations of Π_t^k, Δ_t, Λ_t^k, and U_t^k, respectively. Then, $V_t^k(\Pi_t(\Delta_t, \Lambda_t^k))$ is concave with respect to $\Pi_t(\Delta_t, \Lambda_t^k)$.

Proof: The proof is similar to the proof of Theorem 4. □

Theorem 7: Let $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ be functions defined recursively by

$$V_T\left(\Pi_T(\Delta_T, \Lambda_T^{1:K})\right) := \mathbb{E}^{\mathbf{g}}\left[c_T(X_T) \,|\, \Pi_T = \pi_T\right] \qquad (55)$$

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) := \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}}\Big[c_t(X_t, U_t^{1:K})$$
$$+ V_{t+1}\left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})\right) \,|\, \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K}\Big] \qquad (56)$$

where π_t, $u_t^{1:K}$ are the realizations of Π_t and $U_t^{1:K}$, respectively, and let $\mathbf{g}^* \in \mathcal{G}^s$ be the manager's optimal separated control strategy which achieves the infimum in (55)–(56) for all $t = 0, 1, \ldots, T-1$. Let $\mathbf{g} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^k, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$ be the team's strategy, where $\mathbf{g}^k = \{g_0^k, \ldots, g_{T-1}^k\}$ is a separated control strategy of member $k \in \mathcal{K}$ such that $U_t^k = g_t^k(\Pi_t^k(\Delta_t, \Lambda_t^k))$. Let $V_t^k(\Pi_t^k(\Delta_t, \Lambda_t^k))$ be functions defined recursively by each team player $k \in \mathcal{K}$ after fixing $\mathbf{g}^{-k} = (\mathbf{g}^1, \ldots, \mathbf{g}^{k-1}, \mathbf{g}^{k+1}, \ldots, \mathbf{g}^K)$, by

$$V_T^k\left(\Pi_T^k(\Delta_T, \Lambda_T^k)\right) := \mathbb{E}^{\mathbf{g}^k}\left[c_T(X_T) \,|\, \Pi_T^k = \pi_T^k\right] \qquad (57)$$

and

$$V_t^k \left(\Pi_t^k(\Delta_t, \Lambda_t^k) \right) := \inf_{u_t^k \in \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}^k} \left[c_t(X_t, U_t^k, U_t^{-k}) \right.$$
$$+ V_{t+1}^k \left(\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k) \right) \mid \Pi_t^k = \pi_t^k, \Delta_t = \delta_t$$
$$\left. \Lambda_t^k = \lambda_t^k, U_t^k = u_t^k) \right] \tag{58}$$

where $U_t^{-k} = (U_t^1, \ldots, U_t^{k-1}, U_t^{k+1}, \ldots, U_t^K)$, and $\pi_t^k, \delta_t, \lambda_t^k, u_t^k$ are the realizations of $\Pi_t^k, \Delta_t, \Lambda_t^k,$ and U_t^k, respectively. Then, the solution of the manager in (55), (56) is the same as the solution derived by each player k in (57), (58) for all $t = 0, 1, \ldots, T-1$.

Proof: See Appendix E. □

V. EXAMPLE

We present an example of a delayed sharing pattern team consisting of two members ($K = 2$). This example was used by Varaiya and Walrand [63] to show that Witsenhausen's structural result asserted in his seminal paper [43] is suboptimal.

A. Problem Formulation

In this example, the team evolves for a time horizon $T = 3$ while there is a delay $n = 2$ on information sharing between the two team members. The state $X_t = (X_t^1, X_t^2)$, $t = 1, 2, 3$, of the team is two-dimensional, and the initial state (primitive random variable), $X_0 = (X_0^1, X_0^2)$, of the team is a Gaussian random variable with zero mean, variance 1, and covariance 0.5.

The team's state evolves as follows:

$$X_0 = (X_0^1, X_0^2) \tag{59}$$
$$X_1 = (X_1^1, X_1^2) = (X_0^1 + X_0^2, 0) \tag{60}$$
$$X_2 = (X_2^1, X_2^2) = (X_1^1, U_2^2) = (X_0^1 + X_0^2, U_2^2) \tag{61}$$
$$X_3 = (X_3^1, X_3^2) = (X_2^1 - X_2^2 - U_3^1, 0)$$
$$= (X_0^1 + X_0^2 - U_2^2 - U_3^1, 0). \tag{62}$$

The observation equations are

$$Y_t^k = X_{t-1}^k, \quad k = 1, 2; t = 1, 2, 3. \tag{63}$$

Each team member's feasible sets of actions $\mathcal{U}_t^k$ are specified by

$$\mathcal{U}_t^k = \begin{cases} \mathbb{R}, & \text{if } (k, t) = (1, 3) \text{ or } (2, 2) \\ 0, & \text{otherwise} \end{cases}. \tag{64}$$

Thus, a control strategy $\mathbf{g} = \{g_t^k; k = 1, 2; t = 1, 2, 3\}, \mathbf{g} \in \mathcal{G}$, of the team consists only of the pair $\mathbf{g} = \{g_3^1, g_2^2\}$ since $g_t^k \equiv 0$ for the remaining (k, t). Given the modeling framework above, the information structure $\{(\Delta_t, \Lambda_t^k); k = 1, 2; t = 1, 2, 3\}$ of the team is

$$\Delta_1 = \emptyset, \ \Delta_2 = \emptyset \tag{65}$$
$$\Delta_3 = \{Y_1^1, Y_1^2\} = \{X_0^1, X_0^2\}. \tag{66}$$

Note that since $g_1^1 \equiv 0$ and $g_1^2 \equiv 0$, the realizations of U_1^1 and U_1^2 are zero, and thus Δ_3 includes only the observations in (66). The data $\Lambda_t^k, k = 1, 2$, available to the team member k for the feasible control laws are

$$\Lambda_2^2 = \{Y_1^2, Y_2^2\} = \{X_0^2, X_1^2\} = \{X_0^2\} \tag{67}$$

$$\Lambda_3^1 = \{Y_2^1, Y_3^1\} = \{X_0^1 + X_0^2, X_0^1 + X_0^2\} = \{X_0^1 + X_0^2\}. \tag{68}$$

The problem is to derive the optimal control strategy $\mathbf{g}^* = \{g_3^{1*}, g_2^{2*}\}$, which is the solution of

$$J(\mathbf{g}) = \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2} \mathbb{E}^{\mathbf{g}} \left[(X_3^1)^2 + (U_3^1)^2 \right]$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2} \mathbb{E}^{\mathbf{g}} \left[(X_0^1 + X_0^2 - U_2^2 - U_3^1)^2 + (U_3^1)^2 \right]. \tag{69}$$

B. Optimal Solution

The feasible set $\mathcal{G}$ of control strategies of the team consists of all $\mathbf{g} = \{g_3^1(\Lambda_3^1, \Delta_3), g_2^2(\Lambda_2^2, \Delta_2)\}$, i.e.,

$$g_2^2 : \Delta_2 \times \Lambda_2^2 \to U_2^2, \text{ or } g_2^2 : X_0^2 \to \mathbb{R} \tag{70}$$
$$g_3^1 : \Delta_3 \times \Lambda_3^1 \to U_3^1, \text{ or } g_3^1 : \{X_0^1, X_0^2\} \to \mathbb{R}. \tag{71}$$

The problem (69) has a unique optimal solution (see [63])

$$U_2^2 = \frac{1}{2} X_0^2, \quad U_3^1 = \frac{1}{2}(X_0^1 + X_0^2) - \frac{1}{4} X_0^2. \tag{72}$$

C. Solution Given by Theorem 3

Varaiya and Walrand [63] adopted the notation used by Witsenhausen [43] to describe chronologically the evolution of the team in their example, which proceeds as follows: The initial state X_0 of the team is generated at $t = 0$. Then, at the next time step $t = 1$, each team member $k = 1, 2$ observes $Y_1^k = X_0$ and makes a decision $U_1^k = g_1^k(\Delta_1, \Lambda_1^k)$. The transition of the team to the next state occurs at the same time $t = 1$, i.e., $X_1 = f_1(X_0, U_1^{1:K})$, and the process is repeated until $t = T$. Given that the decision of each team member k at time t depends on Δ_t and Λ_t^k, and that the state of the team evolves after the realization of $U_t^{1:2}$, to be consistent, from a notation point of view, with the state equation (1) in Section II-B, the evolution of the state of the team needs to be revised as follows:

$$X_0 = (X_0^1, X_0^2) \tag{73}$$
$$X_1 = (X_1^1, X_1^2) = (X_0^1, X_0^2) \tag{74}$$
$$X_2 = (X_2^1, X_2^2) = (X_0^1 + X_0^2, 0) \tag{75}$$
$$X_3 = (X_3^1, X_3^2) = (X_0^1 + X_0^2, U_2^2) \tag{76}$$
$$X_4 = (X_4^1, X_4^2) = (X_3^1 - X_3^2 - U_3^1, 0)$$
$$= (X_0^1 + X_0^2 - U_2^2 - U_3^1, 0) \tag{77}$$

where we essentially included a degenerate transition at $t = 1$. Given the modeling framework presented in Section II-B, the information structure $\{(\Delta_t, \Lambda_t^k); k = 1, 2; t = 1, 2, 3\}$ of the team is

$$\Delta_1 = \emptyset, \ \Delta_2 = \emptyset \tag{78}$$
$$\Delta_3 = \{Y_0^1, Y_0^2, Y_1^1, Y_1^2\} = \{X_0^1, X_0^2, X_1^1, X_1^2\} = \{X_0^1, X_0^2\}. \tag{79}$$

Since $g_1^1 \equiv 0$ and $g_1^2 \equiv 0$, the realizations of U_1^1 and U_1^2 are zero, and thus, Δ_3 includes only the observations in (68). The data

$\Lambda_t^k, k = 1, 2$, available to the team member k for the feasible control laws are

$$\Lambda_2^2 = \{Y_0^2, Y_1^2, Y_2^2\} = \{X_0^2, X_1^2, X_2^2\} = \{X_0^2\} \tag{80}$$

$$\Lambda_3^1 = \{Y_2^1, Y_3^1\} = \{X_0^1 + X_0^2, X_0^1 + X_0^2\} = \{X_0^1 + X_0^2\}. \tag{81}$$

We solve problem (69) by using the structural results presented in Section III considering the control strategies $\mathbf{g}^k = \{g_t^k; \ k = 1, 2; \ t = 0, 1, 2, 3\}$ for each team member k, where the control law is of the form $g_t^k(\Pi(\Delta_t, \Lambda_t^k)) = g_t^k(\mathbb{P}(X_t \mid \Delta_t, \Lambda_t^k))$.

For $t = 3$, the manager formulates the dynamic program

$$V_3(\Pi_3)$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^1}\left[(X_0^1 + X_0^2 - U_2^2 - U_3^1)^2\right.$$

$$\left. + (U_3^1)^2 \mid \Pi_3(\Delta_3, \Lambda_3^1), U_3^1\right]$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^1}\left[(X_0^1 + X_0^2 - U_2^2 - U_3^1)^2\right.$$

$$\left. + (U_3^1)^2 \mid \mathbb{P}(X_0^1 + X_0^2, U_2^2 \mid X_0^1, X_0^2, X_0^1 + X_0^2), U_3^1\right] \tag{82}$$

where for any given realization of $((X_0^1 + X_0^2), U_2^2)$ in the information state Π_3, the manager selects U_3^1 to achieve the lower bound in (82). Thus

$$U_3^1 = \frac{1}{2}(X_0^1 + X_0^2) - \frac{1}{2}U_2^2. \tag{83}$$

Substituting (83) into (82) yields

$$V_3(\Pi_3) = \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^1}\left[\frac{\left(X_0^1 + X_0^2 - U_2^2\right)^2}{2}\right|$$

$$\mathbb{P}(X_0^1 + X_0^2, U_2^2 \mid X_0^1, X_0^2, X_0^1 + X_0^2), U_3^1\right]. \tag{84}$$

For $t = 2$, the manager formulates the following dynamic program:

$$V_2(\Pi_2)$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^2}\left[V_3(\Pi_3) \mid \Pi_2(\Delta_2, \Lambda_2^2), U_2^2\right]$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^2}\left[V_3(\Pi_3) \mid \mathbb{P}(X_0^1 + X_0^2 \mid X_0^2), U_2^2\right]$$

$$= \min_{u_2^2 \in \mathcal{U}_2^2, u_3^1 \in \mathcal{U}_3^1} \frac{1}{2}\mathbb{E}^{\mathbf{g}^2}\left[\frac{\left(X_0^1 + X_0^2 - U_2^2\right)^2}{2} \mid \mathbb{P}(X_0^1 + X_0^2 \mid\right.$$

$$\left. X_0^2), U_2^2\right]. \tag{85}$$

Since

$$U_2^2 = g_2^2\left(\mathbb{P}(X_2 \mid \Delta_2, \Lambda_2^2)\right) = g_2^2\left(\mathbb{P}(X_0^1 + X_0^2 \mid X_0^2)\right) \tag{86}$$

the problem of the manager in (85) is to choose, for any given X_0^2, the estimate of $(X_0^1 + X_0^2)$ that minimizes the mean squared error $(X_0^1 + X_0^2 - U_2^2)^2$ in (85). Given the Gaussian statistics, the optimal solution is

$$U_2^2 = \frac{1}{2}X_0^2. \tag{87}$$

Substituting (87) into (83) yields

$$U_3^1 = \frac{1}{2}(X_0^1 + X_0^2) - \frac{1}{4}X_0^2. \tag{88}$$

Therefore, the control laws of the form $g_t^k(\Pi(\Delta_t, \Lambda_t^k)) = g_t^k(\mathbb{P}(X_t \mid \Delta_t, \Lambda_t^k))$ yield the unique optimal solution (72) of problem (69).

VI. CONCLUDING REMARKS AND DISCUSSION

In this article, we provided structural results and a classical dynamic programming decomposition of sequential dynamic team decision problems. We first addressed the problem from the point of view of a manager who seeks to derive the optimal strategy of a team in a centralized process. Then, we addressed the problem from the point of view of each team member, and showed that their solutions is the same as the ones derived by the manager. The key contributions of this article are 1) the structural results for the team from the point of view of a manager that yield an information state which does not depend of the control strategy of the team, and 2) the structural results for each team member that yield an information state which does not depend on their control strategy. These results allow us to formulate two dynamic programming decompositions: (a) one for the team where the manager's optimization problem is over the space of the team's decisions, and (b) one for each team member where the optimization problem is over the space of the decision of each member. Finally, we showed that the solution of each team member is the same as the one derived by the manager. Therefore, each team member can derive their optimal strategy, which is also optimal for the team, without the manager's intervention. One particular limitation of the proposed approach is that the solution of the dynamic programming decompositions might become computationally intensive since the cost-to-go functions are defined on an infinite dimensional space. This a typical challenge in problems of partial observed Markov decision processes. However, given the characterization of the value functions (Theorems 4 and 6) computationally efficient algorithms can be found.

A potential direction for future research should explore the intersection of learning and control for team decision problems with nonclassical information structures. For example, cyber-physical systems, in most instances, represent systems of systems with informationally decentralized structure. In such systems, however, there is typically a large volume of data with a dynamic nature which is added to the system gradually and not altogether in advance. Therefore, neither traditional supervised (or unsupervised) learning nor typical model-based control approaches can effectively facilitate feasible solutions with performance guarantees. These challenges could be circumvented at the intersection of learning and control [105]. Given that the control strategies presented here are separated, a similar separation could be established between learning and control, and thus, combine the online and offline advantages of both traditional supervised (or unsupervised) learning and typical model-based control approaches.

APPENDIX A

PROOF OF THEOREM 1

By applying Bayes' rule, we have

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \frac{p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})}$$

$$= \frac{p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})}$$

$$= \frac{p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})}$$

$$\tag{89}$$

where in the second equality we used Lemma 1.

Next

$$p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) = p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^{1:K}, Y_{t+1}^{1:K}, U_t^{1:K})$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, Y_{t+1}^{1:K}, U_t^{1:K}) \, dX_{t+1}$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(Y_{t+1}^{1:K} \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}$$

where by Lemma 1, the last equation becomes

$$p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \int_{\mathscr{X}_{t+1}} p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}. \tag{90}$$

Note that $p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) = p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$ since $Y_{t-n+1}^{1:K}$ and $U_{t-n+1}^{1:K}$ are already included in $\Lambda_t^{1:K}$, hence, we can write (90) as

$$p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \int_{\mathscr{X}_{t+1}} p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}. \tag{91}$$

Substituting (91) into (89), we have

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \frac{p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})}{\int_{\mathscr{X}_{t+1}} p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K} U_t^{1:K}) \, dX_{t+1}} \tag{92}$$

which we can write as

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \phi_t \left[p^{\mathbf{g}}(\cdot \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}), Y_{t+1}^{1:K} \right] (X_{t+1}) \tag{93}$$

with the function ϕ_t chosen appropriately.

Next

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$

$$= \int_{\mathscr{X}_t} p^{\mathbf{g}}(X_{t+1} \mid X_t, \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$

$$\cdot p^{\mathbf{g}}(X_t \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_t. \tag{94}$$

By Lemma 2 and Remark 1, (94) becomes

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$

$$= \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \, p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \tag{95}$$

which we can write as

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_t, \Lambda_t^{1:K}, U_t^{1:K})$$

$$= \psi_t \left[p(\cdot \mid \Delta_t, \Lambda_t^{1:K}), U_t^{1:K} \right] (X_{t+1}) \tag{96}$$

with the function ψ_t chosen appropriately.

Substituting (96) into (93) yields

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \phi_t \left[\psi_t \left[p(\cdot \mid \Delta_t, \Lambda_t^{1:K}), U_t^{1:K} \right], Y_{t+1}^{1:K} \right] (X_{t+1}). \tag{97}$$

Therefore, $p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K})$ does not depend on the control strategy $\mathbf{g}$, so we can drop the superscript. Moreover, we can choose appropriate function θ_t such that

$$p(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^{1:K}) = \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})(X_{t+1})$$

$$= \theta_t \left[\Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t), Y_{t+1}^{1:K}, U_t^{1:K} \right]. \tag{98}$$

APPENDIX B

PROOF OF LEMMA 4

Obviously, for $t = T$

$$V_T \left(\rho \, \Pi_T(\Delta_T, \Lambda_T^{1:K}) \right)$$

$$= \int_{\mathscr{X}_T} c_T(X_T) \, \rho \, \Pi_T(\Delta_T, \Lambda_T^{1:K})(X_T) \, dX_T$$

$$= \rho \, V_T \left(\Pi_T(\Delta_T, \Lambda_T^{1:K}) \right). \tag{99}$$

For $t = 0, \ldots, T - 1$, by assigning $\Pi_t = \rho \, \Pi_t$ [recall $p(X_t \mid \Delta_t, \Lambda_t^{1:K}) = \Pi_t(\Delta_t, \Lambda_t^{1:K})$], (40) becomes

$$V_t \left(\rho \, \Pi_t(\Delta_t, \Lambda_t^{1:K}) \right)$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^{1:K}) \, \rho \, \Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t) \, dX_t \right.$$

$$+ \int_{\mathscr{Y}_{t+1}} \int_{\mathscr{X}_{t+1}} \int_{\mathscr{X}_t} V_{t+1} \left(\rho \, \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) \right)$$

$$\cdot p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U_t^{1:K}) \, \rho \, p(X_t \mid \Delta_t, \Lambda_t^{1:K})$$

$$\left. dX_t \, dX_{t+1} \, dY_{t+1}^{1:K} \right] \tag{100}$$

where $\mathscr{Y}_{t+1} = \otimes_{k \in \mathcal{K}} \mathscr{Y}^k$.

Next, from (27) Substituting (101) shown at the bottom of the next page, into (100), we have

$$V_t \left(\rho \, \Pi_t(\Delta_t, \Lambda_t^{1:K}) \right)$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^{1:K}) \, \rho \, \Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t) \, dX_t \right.$$

$$+ \int_{\mathscr{Y}_{t+1}} \int_{\mathscr{X}_{t+1}} \int_{\mathscr{X}_t} V_{t+1} \left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) \right)$$

$$\cdot p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U_t^{1:K}) \, \rho \, p(X_t \mid \Delta_t, \Lambda_t^{1:K})$$

$$\left. dX_t \, dX_{t+1} \, dY_{t+1}^{1:K} \right]$$

$$= \rho \, V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right). \tag{102}$$

APPENDIX C

PROOF OF THEOREM 4

We start with (41)

$$V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right)$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^{1:K}) \, \Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t) \, dX_t \right.$$

$$+ \int_{\mathscr{Y}_{t+1}} \int_{\mathscr{X}_{t+1}} \int_{\mathscr{X}_t} V_{t+1} \left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) \right)$$

$$\cdot p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U_t^{1:K}) \, p(X_t \mid \Delta_t, \Lambda_t^{1:K})$$

$$\left. dX_t \, dX_{t+1} \, dY_{t+1}^{1:K} \right] \tag{103}$$

where $\mathscr{Y}_{t+1} = \otimes_{k \in \mathcal{K}} \mathscr{Y}^k$.

Choosing

$$\rho = \int_{\mathscr{X}_{t+1}} \int_{\mathscr{X}_t} p(Y_{t+1}^{1:K} \mid X_{t+1})$$

$$\cdot p(X_{t+1} \mid X_t, U_t^{1:K}) \, p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \, dX_{t+1} \tag{104}$$

we can use the positive homogeneity of $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ (Lemma 4) to write the second part of (103) as follows:

$$\int_{\mathscr{Y}_{t+1}} \int_{\mathscr{X}_{t+1}} \int_{\mathscr{X}_t} V_{t+1} \left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) \right)$$

$$\cdot p(Y_{t+1}^{1:K} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U_t^{1:K}) \, p(X_t \mid \Delta_t, \Lambda_t^{1:K})$$

$$dX_t \, dX_{t+1} \, dY_{t+1}^{1:K}$$

$$= \int_{\mathscr{Y}_{t+1}} V_{t+1} \left(\rho \, \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}) \right) \, dY_{t+1}^{1:K}$$

$$= \int_{\mathscr{Y}_{t+1}} V_{t+1} \left(p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \right.$$

$$\left. p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \right) dY_{t+1}^{1:K} \tag{105}$$

where in the last equality, we substituted (104) and (27).

Thus, we can write (103) as

$$V_t \left(\Pi_t(\Delta_t, \Lambda_t^{1:K}) \right)$$

$$= \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^{1:K}) \, \Pi_t(\Delta_t, \Lambda_t^{1:K})(X_t) \, dX_t \right.$$

$$+ \int_{\mathscr{Y}_{t+1}} V_{t+1} \left(p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \right.$$

$$\left. \left. p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \right) dY_{t+1}^{1:K} \right]. \tag{106}$$

The remainder of the proof follows by induction. Suppose that $V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}))$ is concave. Since

$$V_{t+1} \left(p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \right.$$

$$\left. p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \right) \tag{107}$$

is the composition of a concave function and increasing linear function, it follows that it is concave. However, concavity is preserved by integration (see [106, p. 79]), hence

$$\int_{\mathscr{Y}_{t+1}} V_{t+1} \left(p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \right.$$

$$\left. p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \right) dY_{t+1}^{1:K} \tag{108}$$

is concave. Since the pointwise infimum of concave functions is concave, (106) is concave.

APPENDIX D

PROOF OF THEOREM 5

By applying Bayes' rule, we have

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k)$$

$$= \frac{p^{\mathbf{g}}(Y_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k) \cdot p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k)}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^k)}$$

$$= \frac{p(Y_{t+1}^k \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k)}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^k)}$$

$$= \frac{p(Y_{t+1}^k \mid X_{t+1}) \, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k) \cdot p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^k, U_t^k)}{p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^k)} \tag{109}$$

where in the second equality, we used Lemma 5.

Next

$$p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^k) = p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^k, Y_{t+1}^k, U_t^k)$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^k, Y_{t+1}^k, U_t^k) \, dX_{t+1}$$

$$\rho \, \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})$$

$$= \frac{p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \, \wp \, p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t}{\int_{\mathscr{X}_{t+1}} p(Y_{t+1}^{1:K} \mid X_{t+1}) \int_{\mathscr{X}_t} p(X_{t+1} \mid X_t, U_t^{1:K}) \cdot \wp \, p(X_t \mid \Delta_t, \Lambda_t^{1:K}) \, dX_t \, dX_{t+1}}$$

$$= \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K}). \tag{101}$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(Y_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$\cdot\, p^{\mathbf{g}}(X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k)\, dX_{t+1}$$

$$= \int_{\mathscr{X}_{t+1}} p^{\mathbf{g}}(Y_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$\cdot\, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)\, p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^k, U_t^k)\, dX_{t+1} \quad (110)$$

where by Lemma 5, the last equation becomes

$$p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_{t+1}^k)$$

$$= \int_{\mathscr{X}_{t+1}} p(Y_{t+1}^k \mid X_{t+1})\, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$\cdot\, p^{\mathbf{g}}(\Delta_{t+1}, \Lambda_t^k, U_t^k)\, dX_{t+1}. \quad (111)$$

Substituting (111) into (109), we have

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k)$$

$$= \frac{p(Y_{t+1}^k \mid X_{t+1})\, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)}{\int_{\mathscr{X}_{t+1}} p(Y_{t+1}^k \mid X_{t+1})\, p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)\, dX_{t+1}} \quad (112)$$

which we can write as

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$= \phi_t^k \left[p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k), Y_{t+1}^k \right] (X_{t+1}) \quad (113)$$

with the function ϕ_t^k chosen appropriately.

By Lemma 6, $p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)$ depends only on the control strategy $\mathbf{g}^{-k}$, so we can write $p^{\mathbf{g}^{-k}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)$. Next

$$p^{\mathbf{g}^{-k}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$= \int_{\mathscr{X}_t} p^{\mathbf{g}^{-k}}(X_{t+1} \mid X_t, \Delta_{t+1}, \Lambda_t^k, U_t^k)$$

$$\cdot\, p^{\mathbf{g}^{-k}}(X_t \mid \Delta_{t+1}, \Lambda_t^k, U_t^k)\, dX_t \quad (114)$$

where the last term in (114) can be written as

$$p^{\mathbf{g}^{-k}}(X_t \mid \Delta_{t+1}, \Lambda_t^k, U_t^k) = p^{\mathbf{g}^{-k}}(X_t \mid \Delta_t, Y_{t-n+1}^{1:K}, U_{t-n+1}^{1:K}$$

$$\Lambda_t^k) = p^{\mathbf{g}^{-k}}(X_t \mid \Delta_t, Y_{t-n+1}^{-k}, U_{t-n+1}^{-k}, \Lambda_t^k)$$

$$= p^{\mathbf{g}^{-k}}(X_t \mid \Delta_t, \Lambda_t^k) \cdot p^{\mathbf{g}^{-k}}(Y_{t-n+1}^{-k}, U_{t-n+1}^{-k} \mid \Delta_t, \Lambda_t^k) \quad (115)$$

where $Y_{t-n+1}^{-k} = (Y_{t-n+1}^1, \ldots, Y_{t-n+1}^{k-1}, Y_{t-n+1}^{k+1}, \ldots, Y_{t-n+1}^K)$, and $U_{t-n+1}^{-k} = (U_{t-n+1}^1, \ldots, U_{t-n+1}^{k-1}, U_{t-n+1}^{k+1}, \ldots, U_{t-n+1}^K)$. In the second equality, we dropped Y_{t-n+1}^k, U_{t-n+1}^k from conditioning since they both are included in Λ_t^k. Substituting (115) into (114) yields

$$p^{\mathbf{g}^{-k}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_t^k, U_t^k) = \int_{\mathscr{X}_t} p^{\mathbf{g}^{-k}}(X_{t+1} \mid X_t, \Delta_{t+1}$$

$$\Lambda_t^k, U_t^k) \cdot p^{\mathbf{g}^{-k}}(X_t \mid \Delta_t, \Lambda_t^k) \cdot p^{\mathbf{g}^{-k}}(Y_{t-n+1}^{-k}, U_{t-n+1}^{-k} \mid \Delta_t, \Lambda_t^k)$$

$$\cdot\, dX_t = \int_{\mathscr{X}_t} p^{\mathbf{g}^{-k}}(X_{t+1} \mid X_t, \Delta_t, Y_{t-n+1}^{1:K}, U_{t-n+1}^{1:K}, \Lambda_t^k, U_t^k)$$

$$\cdot\, p^{\mathbf{g}^{-k}}(X_t \mid \Delta_t, \Lambda_t^k) \cdot p^{\mathbf{g}^{-k}}(Y_{t-n+1}^{-k}, U_{t-n+1}^{-k} \mid \Delta_t, \Lambda_t^k)\, dX_t. \quad (116)$$

Substituting (116) into (113) yields

$$p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k)$$

$$= \phi_t^k \left[\psi_t^k \left[p^{\mathbf{g}^{-k}}(\cdot \mid \Delta_t, \Lambda_t^k), Y_{t-n+1}^{1:K}, U_{t-n+1}^{1:K}, \Delta_t, \Lambda_t^k, U_t^k \right] \right.$$

$$\left. Y_{t+1}^k \right] (X_{t+1}). \quad (117)$$

Therefore, $p^{\mathbf{g}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k)$, does not depend on the control strategy $\mathbf{g}^k$ of the team member k, so we can adjust the subscript accordingly. Moreover, we can choose appropriate function θ_t^k such that

$$p^{\mathbf{g}^{-k}}(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k) = \Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1})$$

$$= \theta_t^k \left[\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), Y_{t-n+1}^{1:K}, U_{t-n+1}^{1:K}, \Delta_t, \Lambda_t^k, Y_{t+1}^k, U_t^k \right]$$

$$= \theta_t^k \left[\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), \Delta_{t+1}, \Lambda_{t+1}^k \right]. \quad (118)$$

APPENDIX E

PROOF OF THEOREM 7

Let $\mathbf{g}^* = \{\mathbf{g}^{*1}, \ldots, \mathbf{g}^{*k-1}, \mathbf{g}^{*k}, \mathbf{g}^{*k+1}, \ldots, \mathbf{g}^{*K}\}$ be the optimal separated control strategy of the manager, which achieves the infimum in (55) and (56). Starting with (56), we have

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) = \inf_{u_t^{1:K} \in \prod_{k \in \mathcal{K}} \mathcal{U}_t^k} \mathbb{E}^{\mathbf{g}^*} \left[c_t(X_t, U_t^{1:K}) \right.$$

$$\left. + V_{t+1}\left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})\right) \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \right]$$

$$= \inf_{u_t^k \in \mathcal{U}_t^k} \inf_{u_t^{-k} \in \prod_{i \in \mathcal{K} \setminus \{k\}} \mathcal{U}_t^i} \mathbb{E}^{\mathbf{g}^*} \left[c_t(X_t, U_t^k, U_t^{-k}) \right.$$

$$\left. + V_{t+1}\left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})\right) \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \right]$$

$$= \inf_{u_t^k \in \mathcal{U}_t^k} \inf_{u_t^{-k} \in \prod_{i \in \mathcal{K} \setminus \{k\}} \mathcal{U}_t^i} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \Pi_t(\Delta_t \right.$$

$$\Lambda_t^{1:K})(X_t)\, dX_t + \mathbb{E}^{\mathbf{g}^*} \left[V_{t+1}\left(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1:K})\right) \right.$$

$$\left. \left. \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \right] \right]. \quad (119)$$

The function $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ is concave with respect to $\Pi_t(\Delta_t, \Lambda_t^{1:K})$ (Theorem 4) for all $t = 0, 1, \ldots, T$. Since $\Pi_t(\Delta_t, \Lambda_t^{1:K})$ is increasing with $\Pi_t^k(\Delta_t, \Lambda_t^k)$ for all k (Corollary 1), it follows that $V_t(\Pi_t(\Delta_t, \Lambda_t^{1:K}))$ is also increasing with respect to $\Pi_t^k(\Delta_t, \Lambda_t^k)$ for all k. Thus, the manager can solve (119) for each member $k \in \mathcal{K}$ separately by fixing $u_t^{-k} = (u_t^1, \ldots, u_t^{k-1}, u_t^{k+1}, \ldots, u_t^K)$. By substituting (51) into (119) and for any arbitrary u_t^{-k}, we have

$$V_t\left(\Pi_t(\Delta_t, \Lambda_t^{1:K})\right) = \inf_{u_t^k \in \mathcal{U}_t^k} \left[\int_{\mathscr{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \right.$$

$$\cdot\, \vartheta_t^k\left(\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), \Delta_t, \Lambda_t^{1:K}\right) dX_t$$

$$+ \mathbb{E}^{\mathbf{g}^*} \left[V_{t+1}\left(\vartheta_{t+1}^k\left(\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}), \Delta_{t+1}, \Lambda_{t+1}^{1:K}\right)\right) \right|$$

$$\left. \left. \Pi_t^k = \pi_t^k, \Delta_t = \delta_t, \Lambda_t^{1:K} = \lambda_t^{1:K}, U_t^k = u_t^k \right] \right] \quad (120)$$

where $\pi_t^k, \delta_t, \lambda_t^{1:K}$, and $U_t^k = u_t^k$ are the realizations of Π_t^k, $\Delta_t, \Lambda_t^{1:K}$, and U_t^k, respectively. Given that the function ϑ_t^k is increasing with $\Pi_t^k(\Delta_t, \Lambda_t^k)$ for all k and t (Corollary 1), from

(120), it follows that at $t = T - 1$ and for any u_{T-1}^{-k}

$$\arg\inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[\int_{\mathcal{X}_{T-1}} c_{T-1}(X_{T-1}, U_{T-1}^k, U_{T-1}^{-k}) \right.$$
$$\cdot \vartheta_{T-1}^k \left(\Pi_{T-1}^k(\Delta_{T-1}, \Lambda_{T-1}^k)(X_{T-1}), \Delta_{T-1}, \Lambda_{T-1}^{1:K} \right) dX_{T-1}$$
$$\left. + \int_{\mathcal{X}_T} c_T(X_T) \cdot \vartheta_T^k \left(\Pi_T^k(\Delta_T, \Lambda_T^k)(X_T), \Delta_T, \Lambda_T^{1:K} \right) dX_T \right]$$

$$= \arg\inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[\int_{\mathcal{X}_{T-1}} c_{T-1}(X_{T-1}, U_{T-1}^k, U_{T-1}^{-k}) \right.$$
$$\cdot \Pi_{T-1}^k(\Delta_{T-1}, \Lambda_{T-1}^k)(X_{T-1}) dX_{T-1}$$
$$\left. + \int_{\mathcal{X}_T} c_T(X_T) \cdot \Pi_T^k(\Delta_T, \Lambda_T^k)(X_T) dX_T \right] \quad (121)$$

or, alternatively, from (55) to (56) and (57) to (58), (121) can be written as

$$\arg\inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[V_{T-1} \left(\Pi_{T-1}(\Delta_{T-1}, \Lambda_{T-1}^{1:K}) \right) \right]$$
$$= \arg\inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[V_{T-1}^k \left(\Pi_{T-1}^k(\Delta_{T-1}, \Lambda_{T-1}^k) \right) \right] \quad (122)$$

Continuing backward in time, it follows that for all $t = 0, 1, \ldots, T$, and for any u_t^{-k}

$$\arg\inf_{u_t^k \in \mathcal{U}_t^k} \left[\int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \vartheta_t^k \right.$$
$$\times \left(\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), \Delta_t, \Lambda_t^{1:K} \right)$$
$$\cdot dX_t + \mathbb{E}^{\mathbf{g}^*} \left[V_{t+1} \left(\vartheta_{t+1}^k \left(\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}), \Delta_{t+1} \right. \right. \right.$$
$$\left. \left. \Lambda_{t+1}^{1:K} \right) \right) \Big| \Pi_t^k = \pi_t^k, \Delta_t = \delta_t, \Lambda_t^{1:K} = \lambda_t^{1:K}, U_t^k = u_t^k \right] \Bigg]$$

$$= \arg\inf_{u_t^k \in \mathcal{U}_t^k} \left[\int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \Pi_t^k(\Delta_t, \Lambda_t^k)(X_t) dX_t \right.$$
$$+ \mathbb{E}^{\mathbf{g}^k} \left[V_{t+1}^k \left(\Pi_{t+1}^k(\Delta_{t+1}, \Lambda_{t+1}^k) \right) \mid \Pi_t^k = \pi_t^k, \Delta_t = \delta_t \right.$$
$$\left. \left. \Lambda_t^k = \lambda_t^k, U_t^k = u_t^k \right) \right] \Bigg]. \quad (123)$$

Therefore, the solution of the manager in (55) and (56) is the same as the solution derived by each member k in (57) and (58) for all $t = 0, 1, \ldots, T - 1$.

ACKNOWLEDGMENT

REFERENCES

[1] J. Marshak, *Elements for a Theory of Teams* (1955). Dordrecht, The Netherlands: Springer, 1974, pp. 63–76.

[2] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, pp. 857–881, 1962.

[3] J. Marschak and R. Radner, *Economic Theory of Teams*. London, U.K.: Yale Univ. Press, 1972.

[4] N. Sandell, P. Varaiya, M. Athans, and M. Safonov, "Survey of decentralized control methods for large scale systems," *IEEE Trans. Autom. Control*, vol. AC-23, no. 2, pp. 108–128, Apr. 1978.

[5] L. Zhao and A. A. Malikopoulos, "Enhanced mobility with connectivity and automation: A review of shared autonomous vehicle systems," *IEEE Intell. Transp. Syst. Mag.*, vol. 14, no. 1, pp. 87–102, Jan./Feb. 2022.

[6] A. A. Malikopoulos, C. G. Cassandras, and Y. Zhang, "A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections," *Automatica*, vol. 93, pp. 244–256, 2018.

[7] C. G. Cassandras, "Automating mobility in smart cities," *Annu. Rev. Control*, vol. 44, pp. 1–8, 2017.

[8] A. I. Mahbub, A. A. Malikopoulos, and L. Zhao, "Decentralized optimal coordination of connected and automated vehicles for multiple traffic scenarios," *Automatica*, vol. 117, 2020, Art. no. 108958.

[9] I. A. Ntousakis, I. K. Nikolos, and M. Papageorgiou, "Optimal vehicle trajectory planning in the context of cooperative merging on highways," *Transp. Res. Part C: Emerg. Technol.*, vol. 71, pp. 464–488, 2016.

[10] A. A. Malikopoulos, L. E. Beaver, and I. V. Chremos, "Optimal time trajectory and coordination for connected and automated vehicles," *Automatica*, vol. 125, 2021, Art. no. 109469.

[11] B. Chalaki and A. A. Malikopoulos, "Optimal control of connected and automated vehicles at multiple adjacent intersections," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 3, pp. 972–984, May 2022.

[12] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[13] X. Zhang et al., "Networked control systems: A survey of trends and techniques," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 1, pp. 1–17, Jan. 2020.

[14] I. V. Chremos and A. A. Malikopoulos, "Socioeconomic Impact of Emerging Mobility Markets and Implementation Strategies," in *AI-enabled Technologies for Autonomous and Connected Vehicles*, I. Kolmanovsky, Y. Murphey, and P. Watta, Eds. Berlin, Germany: Springer, 2021.

[15] S. K. Khaitan and J. D. McCalley, "Cyber physical system approach for design of power grids: A survey," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2013, pp. 1–5.

[16] A. M. Howlader, N. Urasaki, and A. Y. Saber, "Control strategies for wind-farm-based smart grid system," *IEEE Trans. Ind. Appl.*, vol. 50, no. 5, pp. 3591–3601, Sep.–Oct. 2014.

[17] W. Du, J. Bi, and H. F. Wang, "Small-signal angular stability of power system as affected by grid-connected variable speed wind generators—A survey of recent representative works," *CSEE J. Power Energy Syst.*, vol. 3, no. 3, pp. 223–231, 2017.

[18] F. Pasqualetti, F. Dorfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 110–127, Feb. 2015.

[19] M. Sami, N. K. Noordin, M. Khabazian, F. Hashim, and S. Subramaniam, "A survey and taxonomy on medium access control strategies for cooperative communication in wireless networks: Research issues and challenges," *IEEE Commun. Surv. Tut.*, vol. 18, no. 4, pp. 2493–2521, Fourthquarter 2016.

[20] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, "Submodularity in input node selection for networked linear systems: Efficient algorithms for performance and controllability," *IEEE Control Syst. Mag.*, vol. 37, no. 6, pp. 52–74, Dec. 2017.

[21] A. Dave, I. V. Chremos, and A. A. Malikopoulos, "Social media and misleading information in a democracy: A mechanism design approach," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2633–2639, May 2022.

[22] K. Saulnier, D. Saldaña, A. Prorok, G. J. Pappas, and V. Kumar, "Resilient flocking for mobile robot teams," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 1039–1046, Apr. 2017.

[23] L. E. Beaver and A. A. Malikopoulos, "An overview on optimal flocking," *Annu. Rev. Control*, vol. 51, pp. 88–99, 2021.

[24] Z. Li, R. Chen, L. Liu, and G. Min, "Dynamic resource discovery based on preference and movement pattern similarity for large-scale social Internet of Things," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 581–589, Aug. 2016.

[25] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial Internet of Things: A cyber-physical systems perspective," *IEEE Access*, vol. 6, pp. 78238–78259, 2018.

[26] J. A. Ansere, G. Han, L. Liu, Y. Peng, and M. Kamal, "Optimal resource allocation in energy-efficient internet-of-things networks with imperfect CSI," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5401–5411, Jun. 2020.

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

[28] H. S. Witsenhausen, "A standard form for sequential stochastic control," *Math. Syst. theory*, vol. 7, no. 1, pp. 5–11, 1973.

[29] J. Krainak, J. Speyer, and S. Marcus, "Static team problems–Part I: Sufficient conditions and the exponential cost criterion," *IEEE Trans. Autom. Control*, vol. AC-27, no. 4, pp. 839–848, Aug. 1982.

[30] J. Krainak, J. Speyer, and S. Marcus, "Static team problems—Part II: Affine control laws, projections, algorithms, and the legt problem," *IEEE Trans. Autom. Control*, vol. AC-27, no. 4, pp. 848–859, Aug. 1982.

[31] S. Yüksel and T. Basar, *Stochastic Networked Control Systems*, 2013 ed. Basel, Switzerland: Birkhäuser, 2013.

[32] M. S. Andersland, "Decoupling non-sequential stochastic control problems," *Syst. Control Lett.*, vol. 16, no. 1, pp. 65–69, 1991.

[33] M. S. Andersland and D. Teneketzis, "Information structures, causality, and nonsequential stochastic control I: Design-independent properties," *SIAM J. Control Optim.*, vol. 30, no. 6, pp. 1447–1475, 1992.

[34] M. S. Andersland and D. Teneketzis, "Information structures, causality, and nonsequential stochastic control II: Design-dependent properties," *SIAM J. Control Optim.*, vol. 32, no. 6, pp. 1726–1751, 1994.

[35] D. Teneketzis, "On information structures and nonsequential stochastic control," *CWI Quart.*, vol. 9, no. 3, pp. 241–260, 1996.

[36] H. S. Witsenhausen, "The Intrinsic Model for Discrete Stochastic Control: Some Open Problems," in *Control Theory, Numerical Methods and Computer Systems Modelling*, A. Bensoussan and J. L. Lions, Eds. Berlin, Germany: Springer, 1975, pp. 322–335.

[37] J. H. van Schuppen and T. Villa, *Coordination Control of Distributed Systems*. Berlin, Germany: Springer, 2015.

[38] N. Saldi and S. Yüksel, "Geometry of information structures, strategic measures and associated control topologies," 2021, *arXiv:2010.07377*.

[39] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control*, vol. 6, no. 1, pp. 131–147, 1968.

[40] C. H. Papadimitriou and J. Tsitsiklis, "On the complexity of designing distributed protocols," *Inf. Control*, vol. 53, no. 3, pp. 211–218, 1982.

[41] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Trans. Autom. Control*, vol. AC-30, no. 5, pp. 440–446, May 1985.

[42] C. H. Papadimitriou and J. N. Tsitsiklis, "Intractable problems in control theory," in *Proc. 24th IEEE Conf. Decis. Control*, 1985, pp. 1099–1103.

[43] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[44] Y.-C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644–654, Jun. 1980.

[45] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *Proc. IEEE 51st IEEE Conf. Decis. Control*, 2012, pp. 1291–1306.

[46] A. A. Malikopoulos, "A duality framework for stochastic optimal control of complex systems," *IEEE Trans. Autom. Control*, vol. 61, no. 10, pp. 2756–2765, Oct. 2016.

[47] A. A. Malikopoulos, "Centralized stochastic optimal control of complex systems," in *Proc. Eur. Control Conf.*, 2015, pp. 721–726.

[48] A. A. Malikopoulos, "Equilibrium control policies for Markov chains," in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 7093–7098.

[49] A. A. Malikopoulos, V. Maroulas, and J. XIong, "A multiobjective optimization framework for stochastic control of complex systems," in *Proc. Amer. Control Conf.*, 2015, pp. 4263–4268.

[50] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1986.

[51] H. J. Kushner, *Introduction to Stochastic Control*. New York, NY, USA: Holt, Rinehart and Winston, 1971.

[52] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1606–1620, Jul. 2011.

[53] A. Dave and A. A. Malikopoulos, "The prescription approach for decentralized stochastic control with word-of-mouth communication," 2021, *arXiv:1907.12125*.

[54] J. M. Ooi, S. M. Verbout, J. T. Ludwig, and G. W. Wornell, "A separation theorem for periodic sharing information patterns in decentralized control," *IEEE Trans. Autom. Control*, vol. 42, no. 11, pp. 1546–1550, Nov. 1997.

[55] M. Aicardi, F. Davoli, and R. Minciardi, "Decentralized optimal control of Markov chains with a common past information set," *IEEE Trans. Autom. Control*, vol. AC-32, no. 11, pp. 1028–1031, Nov. 1987.

[56] J. Bismut, "An example of interaction between information and control: The transparency of a game," *IEEE Trans. Autom. Control*, vol. AC-18, no. 5, pp. 518–522, Oct. 1973.

[57] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Math. Oper. Res.*, vol. 12, no. 3, pp. 441–450, Aug. 1987.

[58] C. Striebel, "Sufficient statistics in the optimum control of stochastic systems," *J. Math. Anal. Appl.*, vol. 12, no. 3, pp. 576–592, 1965.

[59] V. Krishnamurthy, *Partially Observed Markov Decision Processes (From Filtering to Controlled Sensing)*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[60] D. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed. Nashua, NH, USA: Athena Sci., 2017.

[61] A. Hordjik, "Dynamic programming and Markov potential theory," *Math. Centre Tracts*, vol. 51, pp. 514–514, 1974.

[62] R. A. Howard, *Dynamic Programming and Markov Process*. Cambridge, MA, USA: MIT Press, 1960.

[63] P. Varaiya and J. Walrand, "On delayed sharing patterns," *IEEE Trans. Autom. Control*, vol. AC-23, no. 3, pp. 443–445, Jun. 1978.

[64] B. Kurtaran, "Corrections and extensions to "Decentralized stochastic control with delayed sharing information pattern"," *IEEE Trans. Autom. Control*, vol. AC-24, no. 4, pp. 656–657, Aug. 1979.

[65] J. Wu and S. Lall, "A theory of sufficient statistics for teams," in *Proc. 53rd IEEE Conf. Decis. Control*, 2014, pp. 2628–2635.

[66] A. Gupta, S. Yüksel, T. Başar, and C. Langbort, "On the existence of optimal policies for a class of static and sequential dynamic teams," *SIAM J. Control Optim.*, vol. 53, no. 3, pp. 1681–1712, 2015.

[67] A. Dave and A. A. Malikopoulos, "Decentralized stochastic control in partially nested information structures," *IFAC-PapersOnLine*, Chicago, IL, USA, vol. 52, no. 20, pp. 97–102, 2019.

[68] A. Dave and A. A. Malikopoulos, "Structural results for decentralized stochastic control with a word-of-mouth communication," in *Proc. 2020 Amer. Control Conf.*, 2020, pp. 2796–2801.

[69] A. Nayyar and D. Teneketzis, "Common knowledge and sequential team problems," *IEEE Trans. Autom. Control*, vol. 64, no. 12, pp. 5108–5115, Dec. 2019.

[70] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.

[71] C. D. Charalambous and N. U. Ahmed, "Equivalence of decentralized stochastic dynamic decision systems via Girsanov's measure transformation," in *Proc. 53rd IEEE Conf. Decis. Control*, 2014, pp. 439–444.

[72] C. D. Charalambous, "Decentralized optimality conditions of stochastic differential decision problems via Girsanov's measure transformation," *Math. Control, Signals, Syst.*, vol. 28, no. 3, p. 19, 2016.

[73] C. D. Charalambous and N. U. Ahmed, *Team Theory and Information Structures of Stochastic Dynamic Decentralized Decision*. Berlin, Germany: Springer, 2015, pp. 155–163.

[74] C. D. Charalambous and N. U. Ahmed, "Centralized versus decentralized optimization of distributed stochastic differential decision systems with different information structures—Part I: A general theory," *IEEE Trans. Autom. Control*, vol. 62, no. 3, pp. 1194–1209, Mar. 2017.

[75] C. D. Charalambous and N. U. Ahmed, "Centralized versus decentralized optimization of distributed stochastic differential decision systems with different information structures—Part II: Applications," *IEEE Trans. Autom. Control*, vol. 63, no. 7, pp. 1913–1928, Jul. 2018.

[76] C. B. McGuire and R. Radner, Eds., *Decision and Organization: A Volume in Honor of Jacob Marschak (Studies in Mathematical and Managerial Economics)*. Amsterdam, The Netherlands: North-Holland Pub. Co., 1972.

[77] Y.- C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems—Part I," *IEEE Trans. Autom. Control*, vol. AC-17, no. 1, pp. 15–22, Feb. 1972.

[78] J. Wu and S. Lall, "A dynamic programming algorithm for decentralized Markov decision processes with a broadcast structure," in *Proc. IEEE Conf. Decis. Control*, 2010, pp. 6143–6148.

[79] H. Witsenhausen, "On the structure of real time source coders," *Bell Syst. Tech. J*, vol. 58, no. 6, pp. 1437–1451, 1979.

[80] P. Varaiya and J. Walrand, "Optimal causal coding-decoding problems," *Math. Oper. Res. Bell Syst. Tech. J. Bell Syst. Tech. J*, vol. 59, no. 8, pp. 814–820, 1983.

[81] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4017–4035, Sep. 2006.

[82] A. Nayyar and D. Teneketzis, "On jointly optimal real-time encoding and decoding strategies in multi-terminal communication systems," in *Proc. IEEE Conf. Decis. Control*, 2008, pp. 1620–1627.

[83] Y. Kaspi and N. Merhav, "Structure theorem for real-time variable-rate lossy source encoders and memory-limited decoders with side information," in *Proc. IEEE Int. Symp. Inf. Theory*, 2010, pp. 86–90.

[84] S. Yüksel, "On optimal causal coding of partially observed Markov sources in single and multiterminal settings," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 424–437, Jan. 2013.

[85] D. Teneketzis and P. Varaiya, "The decentralized quickest detection problem," *IEEE Trans. Autom. Control*, vol. AC-29, no. 7, pp. 641–644, Jul. 1984.

[86] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 433–442, Mar. 1993.

[87] J. N. Tsitsiklis, "Decentralized detection," *Adv. Stat. Signal Process.*, vol. 2, pp. 297–344, 1993.

[88] V. V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1657–1665, May 2001.

[89] P. Varaiya and J. Walrand, "Causal coding and control for Markov chains," *Syst. Control Lett.*, vol. 3, no. 4, pp. 189–192, 1983.

[90] A. Mahajan and D. Teneketzis, "Optimal performance of networked control systems with nonclassical information structures," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1377–1404, 2009.

[91] A. Mahajan, "Sequential decomposition of sequential dynamic teams: Applications to real-time communication and networked control systems." Ph.D. dissertation, Dept. Elect. Eng., Univ. Michigan, Ann Arbor, Ann Arbor, MI, USA, 2008.

[92] S. Yüksel, "A universal dynamic program and refined existence results for decentralized stochastic control," *SIAM J. Control Optim.*, vol. 58, no. 5, pp. 2711–2739, 2020.

[93] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication systems with noisy feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 580–595, May 2008.

[94] A. Mahajan and D. Teneketzis, "Optimal design of sequential real-time communication systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5317–5338, Nov. 2009.

[95] A. Mahajan, A. Nayyar, and D. Teneketzis, "Identifying tractable decentralized control problems on the basis of information structure," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, 2008, pp. 1440–1449.

[96] A. Mahajan, "Optimal decentralized control of coupled subsystems with control sharing," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2377–2382, Sep. 2013.

[97] A. Nayyar, A. Gupta, C. Langbort, and T. Basar, "Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 555–570, Mar. 2014.

[98] J. Arabneydi and A. Mahajan, "Team optimal control of coupled subsystems with mean-field sharing," in *Proc. 53rd IEEE Conf. Decis. Control*, 2014, pp. 1669–1674.

[99] T. Yoshikawa, "Decomposition of dynamic team decision problems," *IEEE Trans. Autom. Control*, vol. AC-23, no. 4, pp. 627–632, Aug. 1978.

[100] G. Casalino, F. Davoli, R. Minciardi, P. Puliafito, and R. Zoppoli, "Partially nested information structures with a common past," *IEEE Trans. Autom. control*, vol. AC-29, no. 9, pp. 846–850, Sep. 1984.

[101] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Autom. Control*, vol. 54, no. 12, pp. 2773–2786, Dec. 2009.

[102] A. Nayyar, A. Mahajan, and D. Teneketzis, *The Common-Information Approach to Decentralized Stochastic Control*. Berlin, Germany: Springer, 2014, pp. 123–156.

[103] E. J. Sondik, "The optimal control of partially observed Markov processes," Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 1971.

[104] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations Res.*, vol. 21, no. 5, pp. 1071–1088, Sep. 1973.

[105] A. A. Malikopoulos, "Separation of learning and control for cyber-physical systems," 2022, *arXiv:2107.06379*.

[106] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Andreas A. Malikopoulos (Senior Member, IEEE) received the Diploma degree in mechanical engineering from the National Technical University of Athens, Athens, Greece, in 2000, and the M.S. and Ph.D. degrees in mechanical engineering from the Department of Mechanical Engineering, University of Michigan, Ann Arbor, Ann Arbor, MI, USA, in 2004 and 2008, respectively.

He is currently the Terri Connor Kelly and John Kelly Career Development Associate Professor with the Department of Mechanical Engineering, University of Delaware, Newark, DE, USA, the Director of the Information and Decision Science (IDS) Laboratory, and the Director of the Sociotechnical Systems Center. Prior to these appointments, he was the Deputy Director and the Lead of the Sustainable Mobility Theme of the Urban Dynamics Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA, and a Senior Researcher with General Motors Global Research and Development. His research spans several fields, including analysis, optimization, and control of cyber-physical systems; decentralized systems, stochastic scheduling and resource allocation problems, and learning in complex systems. The emphasis is on applications related to smart cities, emerging mobility systems, and sociotechnical systems.

Dr. Malikopoulos has been an Associate Editor for IEEE TRANSACTIONS ON INTELLIGENT VEHICLES and IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS from 2017 to 2020. He is currently an Associate Editor for *Automatica* and IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is a member of SIAM, AAAS, and a Fellow of the ASME.