On Team Decision Problems with Nonclassical Information Structures

Andreas A. Malikopoulos, Senior Member, IEEE

Abstract—In this paper, 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—Team theory, decentralized control, nonclassical information structures, Markov decision theory.

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 [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; and (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 non-sequential team problem. Formulating a well-posed non-sequential 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 non-sequential teams by using the framework of Witsenhausen’s intrinsic model [36]. In this paper, 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 (1) classical, (2) partially nested (alternatively also called overlapping, or quasiclassical), and (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 (1) n-step delayed-

This research was supported by the Sociotechnical Systems Center (SSC) at the University of Delaware.

The author is with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716 USA (email: andreas@udel.edu).
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) \textit{periodic sharing} \([54]\), where each team member has access to the information, i.e., observations and decisions, of the other members periodically; (3) \textit{delayed observation (or state)} \([55]\) \textit{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 \([56]\); and (5) \textit{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 \textit{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 \textit{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 \textit{person-by-person} approach, (2) the \textit{designer’s} approach, and (3) the \textit{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 Witten-hausen \([28]\), 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 Paper

In this paper, 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 paper are the induction of: (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 (Theorem 2 and Theorem 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 II-B, from that undertaken here is that, in this paper, 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 paper 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 (Theorem 2 and Theorem 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 (Theorem 2 and Theorem 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 Paper

The remainder of the paper proceeds 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, we provide concluding remarks and discuss potential directions for future research in Section VI.

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 \( X_{1:t}^{1:K} \) denotes the vector of random variables \( (X_t^{1}, X_t^{2}, \ldots, X_t^{K}) \), \( x_{1:t}^{1:K} \) 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 \(h^k_t(\cdot, \cdot), \ldots, h^K_t(\cdot, \cdot)\). The expectation of a random variable is denoted by \(\mathbb{E}[]\), 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 \(g\), we use \(\mathbb{E}^g[], \mathbb{P}^g(\cdot)\), and \(p^g(\cdot)\) to denote that the expectation, probability, and probability density function, respectively, depend on the choice of the control strategy \(g\). For two measurable spaces \((\mathcal{X}, \mathcal{F})\) and \((\mathcal{Y}, \mathcal{G})\), \(\mathcal{F} \otimes \mathcal{G}\) is the product \(\sigma\)-algebra on \(\mathcal{X} \times \mathcal{Y}\) generated by the collection of all measurable rectangles, i.e., \(\mathcal{F} \otimes \mathcal{G} = \sigma(\{A \times B : A \in \mathcal{F}, B \in \mathcal{G}\})\). The product of \((\mathcal{X}, \mathcal{F})\) and \((\mathcal{Y}, \mathcal{G})\) is the measurable space \((\mathcal{X} \times \mathcal{Y}, \mathcal{F} \otimes \mathcal{G})\).

B. Modeling Framework

We consider a team of \(K \in \mathbb{N}\) members with a measurable state space \((\mathcal{X}_t, \mathcal{F}_t)\), 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 \(\mathcal{F}_t\) is the associated \(\sigma\)-algebra. The state of the team is represented by a random variable \(X_t : (\Omega, \mathcal{F}) \to (\mathcal{X}_t, \mathcal{F}_t)\), defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\Omega\) is the sample space, \(\mathcal{F}\) is the associated \(\sigma\)-algebra, and \(\mathbb{P}\) is a probability measure on \((\Omega, \mathcal{F})\). The decision of each team member \(k \in \mathcal{K}\), \(\mathcal{K} = \{1, \ldots, K\}\), is represented by a random variable \(U^k_t : (\Omega, \mathcal{F}) \to (U^k_t, \mathcal{G}^k_t)\), defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and takes values in the measurable space \((U^k_t, \mathcal{G}^k_t)\), where \(U^k_t\) is team member \(k\)’s nonempty feasible set of actions at time \(t\) and \(\mathcal{G}^k_t\) is the associated \(\sigma\)-algebra. Let \(U^k_t = (U^k_t(1), \ldots, U^k_t(T))\) 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

\[
X_{t+1} = f_t \left( X_t, U^k_t, W_t \right),
\]

where \(W_t\) is a random variable defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) that corresponds to the external, uncontrollable disturbance to the team and takes values in a measurable set \((\mathcal{W}, \mathcal{G})\), i.e., \(W_t : (\Omega, \mathcal{F}) \to (\mathcal{W}, \mathcal{G})\), \(\{W_t : t = 0, 1, \ldots, T - 1\}\) 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 \(Y^k_t\), which takes values in a measurable set \((\mathcal{Y}^k, \mathcal{G}^k)\), described by the observation equation

\[
Y^k_t = h^k_t \left( X_t, Z^k_t \right),
\]

where \(Z^k_t\) is a random variable defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) that corresponds to the noise of each member’s sensor and takes values in a measurable set \((\mathcal{Z}^k, \mathcal{G}^k)\), i.e., \(Z^k_t : (\Omega, \mathcal{F}) \to (\mathcal{Z}^k, \mathcal{G}^k)\), \(\{Z^k_t : t = 0, 1, \ldots, T - 1; k = 1, \ldots, K\}\) 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:

1) \(n\)-step delayed information sharing: In this case, at time \(t\), team member \(k \in \mathcal{K}\) observes \(Y^k_t\), and the \(n\)-step, \(n \in \mathbb{R}\), past observations \(Y^n_{0:t-n}\) and decisions \(U^n_{0:t-n}\) of the entire team. Thus, at time \(t\), the data available to member \(k\) consist of the data \(\Delta^k_t\) available to all team members, i.e.,

\[
\Delta^k_t : = (Y^{1:K}_{0:t-n}, U^{1:K}_{0:t-n}),
\]

where \(Y^n_{0:t-n} = \{Y^1_{0:t-n}, \ldots, Y^K_{0:t-n}\}\), \(U^n_{0:t-n} = \{U^1_{0:t-n}, \ldots, U^K_{0:t-n}\}\), and the data \(\Delta^k_t\) known only to member \(k \in \mathcal{K}\), i.e.,

\[
\Delta^k_t : = (Y^k_{t-n+1:t}, U^k_{t-n+1:t-1}).
\]

The \(n\)-step delayed information sharing can also be asymmetric [53], i.e., for each member \(k \in \mathcal{K}\), \(Y^n_{t-n}, U^n_{t-n}, \) 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 \(\Delta^k_t\) and \(\Lambda^k_t\), \(k \in \mathcal{K}\), becomes

\[
\Delta^k_t : = (Y^{1:K}_{0:\alpha\omega}, U^{1:K}_{0:\alpha\omega}),
\]

\[
\Lambda^k_t : = (Y^k_{\alpha\omega + 1:(\alpha + 1)\omega}, U^k_{\alpha\omega + 1:(\alpha + 1)\omega}).
\]

3) \(n\)-step delayed observation sharing: In this case, \(\Delta^k_t\) and \(\Lambda^k_t\), \(k \in \mathcal{K}\), become

\[
\Delta^k_t : = (Y^{1:K}_{0:t-n}),
\]

\[
\Lambda^k_t : = (Y^k_{t-n+1:t}, U^k_{t-n+1:t-1}).
\]

4) \(n\)-step delayed control sharing: In this case, \(\Delta^k_t\) and \(\Lambda^k_t\), \(k \in \mathcal{K}\), become

\[
\Delta^k_t : = (U^{1:K}_{0:t-n}),
\]

\[
\Lambda^k_t : = (Y^k_{0:t}, U^k_{t-n+1:t-1}).
\]

5) No sharing information: In this case, \(\Delta^k_t\) and \(\Lambda^k_t\), \(k \in \mathcal{K}\), become

\[
\Delta^k_t : = 0,
\]

\[
\Lambda^k_t : = (Y^k_{0:t}, U^k_{0:t-1}).
\]

The collection \(\{(\Delta^k_t, \Lambda^k_t) : k \in \mathcal{K} : t = 0, 1, \ldots, T - 1\}\) 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 a \(n\)-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 \((\mathcal{D}_t, \mathcal{F}_t)\) and \((\mathcal{L}^k_t, \mathcal{G}^k_t)\), \(k \in \mathcal{K}\), be the measurable spaces of all possible realizations of \(\Delta^k_t\) and \(\Lambda^k_t\), respectively, where \(\mathcal{D}_t\) and \(\mathcal{L}^k_t\) are the associated \(\sigma\)-algebras. Each team member \(k\) makes a decision

\[
U^k_t = G^k_t \left( \Lambda^k_t, \Delta^k_t \right),
\]

where \(G^k_t\) is a control law of \(k \in \mathcal{K}\), which is a measurable function \(G^k_t : (\mathcal{L}^k_t \times \mathcal{D}_t, \mathcal{G}^k_t \otimes \mathcal{F}_t) \to (\mathcal{U}^k_t, \mathcal{G}^k_t)\). The control strategy of team member \(k \in \mathcal{K}\) is \(g^k = \{g^k_t : t = 0, \ldots, T - 1\}\).
For any control strategy \( g \in G \), where \( G \) is the feasible set of the control strategies for \( k \). Thus the set of feasible decentralized control strategies is \( G^{Dec} = \times_{k \in K} G^k \), i.e., \( g = \{g^1, \ldots, g^K\} \in G^{Dec} \). If \( g \in G \) is a centralized control strategy then \( \gamma = (L^k_1 \times \cdots \times L^k_{L^k} \times D_t^k) \in L^k \times \cdots \times L^k_{L^k} \times D_t^k \).

**Problem 1.** The problem is to derive the optimal control strategy \( g^* \) of the team that minimizes the expected total cost

\[
J(g) = \mathbb{E}[\sum_{t=0}^{T-1} c_t(X_t, U^1_t:K) + c_T(X_T)] ,
\]

(14)

where the expectation is taken with respect to the joint probability distribution of the random variables \( X_t \) and \( U^1_t:K \) designated by the choice of \( g \), \( c_t(X_t, U^1_t:K) = (X_t \times \prod_{k \in K} U^k_t \times \xi_t \times \cdots \times \gamma_t)^T \rightarrow \mathbb{R} \) is the team’s measurable cost function, and \( c_T(X_T) = (X_T)^T \rightarrow \mathbb{R} \) is the measurable cost function at \( T \).

The statistics of the primitive random variables \( Z_0, \{W_t : t = 0, \ldots, T - 1\}, \{Z^k_t : k \in K, t = 0, \ldots, T - 1\} \), the state equations \( \{X_t : t = 0, \ldots, T - 1\} \), the observation equations \( \{h^k_t : k \in K, t = 0, \ldots, T - 1\} \), and the cost functions \( \{c_t : t = 0, \ldots, T\} \) 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 \( g \in G \) 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, \( \Pi_t \), for the team described by the state equation (1) is a function of \( (\Delta_t, \Lambda^1_t:K) \) and \( \Pi_{t+1} \) can be determined from \( \Pi_t, Y^1_{t+1}, \) and \( U^1_t:K \).

The notation is simpler if we consider densities for all probability distributions. Let \( g \in G \) be a control strategy and \( (\Delta_t, \Lambda^1_t:K) \) 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 \( g \in G \) of the team,

\[
p^\mathbb{P}(Y^1_{t+1} | X_{t+1}, \Delta_t, \Lambda^1_t:K, U^1_t:K) = p^\mathbb{P}(Y^1_{t+1} | X_{t+1}) ,
\]

(15)

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

**Proof.** The realization of \( Y^1_{t+1} \) is statistically determined by the conditional distribution of \( Y^1_{t+1} \) given \( X_{t+1} \) in (2), hence

\[
p^\mathbb{P}(Y^1_{t+1} | X_{t+1}, \Delta_t, \Lambda^1_t:K, U^1_t:K) = p^\mathbb{P}(Y^1_{t+1} | X_{t+1}) .
\]

(16)

However,

\[
p^\mathbb{P}(Y^1_{t+1} | X_{t+1}) = p^\mathbb{P}(Z^1_{t+1} \in \prod_{k \in K} B^k | X_{t+1}) ,
\]

(17)

where \( B^k \in \mathcal{Z}^k, k \in K \). Since, \( \{Z^k_t : k = 1, \ldots, K, t = 0, \ldots, T - 1\} \) is a sequence of independent random variables that are independent of \( X_{t+1} \),

\[
p^\mathbb{P}(Z^1_{t+1} \in \prod_{k \in K} B^k | X_{t+1}) = p^\mathbb{P}(Z^1_{t+1} \in \prod_{k \in K} B^k) .
\]

(18)

Thus,

\[
p^\mathbb{P}(Y^1_{t+1} | X_{t+1}) = p^\mathbb{P}(Y^1_{t+1} | X_{t+1}) .
\]

(19)

The result follows from (16) and (19).

**Lemma 2.** For any control strategy \( g \in G \) of the team,

\[
p^\mathbb{P}(X_{t+1} | X_t, \Delta_t, \Lambda^1_t:K, U^1_t:K) = p^\mathbb{P}(X_{t+1} | X_t, U^1_t:K) ,
\]

(20)

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

**Proof.** The realization of \( X_{t+1} \) is statistically determined by the conditional distribution of \( X_{t+1} \) given \( X_t \) and \( U^1_t:K \), i.e.,

\[
p^\mathbb{P}(X_{t+1} | X_t, U^1_t:K) = p^\mathbb{P}(W_t \in A | X_t, U^1_t:K) ,
\]

(21)

where \( A \in \mathcal{W} \). Since, \( \{W_t : t = 0, \ldots, T - 1\} \) is a sequence of independent random variables that are independent of \( X_t \) and \( U^1_t:K \),

\[
p^\mathbb{P}(W_t \in A | X_t, U^1_t:K) = p^\mathbb{P}(W_t \in A) .
\]

(22)

Next,

\[
p^\mathbb{P}(X_{t+1} | X_t, \Delta_t, \Lambda^1_t:K, U^1_t:K) = p^\mathbb{P}(W_t \in A | X_t, \Delta_t, \Lambda^1_t:K, U^1_t:K) = p^\mathbb{P}(W_t \in A) .
\]

(23)

The result follows from (21), (22) and (23).

**Lemma 3.** For any control strategy \( g \in G \) of the team,

\[
p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_t:K) = p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_t:K) ,
\]

(24)

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

**Proof.** We have

\[
p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_t:K) = p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_{t-2:K}, Y^1_{t-2:K}, U^1_{t-2:K}, U^1_{t-1:K}) .
\]

(25)

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

**Remark 1.** As a consequence of Lemma 3 and since \( X_t \) does not depend on \( U^1_{t:K} \), we have

\[
p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_{t:K}, U^1_{t:K}) = p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_{t:K}) .
\]

(26)

Given that the manager can observe the data \( (\Delta_t, \Lambda^1_{t:K}) \) 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 \( p^\mathbb{P}(X_t | \Delta_t, \Lambda^1_{t:K}) \). 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 \( g \in \mathcal{G} \) of the team, the conditional probability density \( p^g(X_t \mid \Delta_t, \Lambda^1_t) \) does not depend on the control strategy \( g \). It is an information state \( \Pi_t(\Delta_t, \Lambda^1_t)(X_t) \), i.e., \( \Pi_t(\Delta_t, \Lambda^1_t)(X_t) = p(X_t \mid \Delta_t, \Lambda^1_t) \) with \( \int_{\mathcal{X}} \Pi_t(\Delta_t, \Lambda^1_t) dX_t = 1 \), that can be evaluated from \( \Delta_t, \Lambda^1_t \). Moreover, there is a function \( \theta_t \), which does not depend on the control strategy \( g \), such that
\[
\Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})(X_{t+1}) = \theta_t(\Pi_t(\Delta_t, \Lambda^1_t)(X_t), Y_{t+1}^{1:t}, U_{t}^{1:t}),
\]
for all \( t = 0, 1, \ldots, T - 1 \).

Proof. See Appendix A.

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

B. Optimal Control Strategy of the Team

In what follows, to simplify notation, the information state \( \Pi_t(\Delta_t, \Lambda^1_t) \) of the team at \( t \) is denoted simply by \( \Pi_t \). We use its arguments \( \Delta_t \) and \( \Lambda^1_t \) only if it is required by our exposition.

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

In implementing a separated control strategy, we first need to compute the conditional probability \( \Pi_t(\Delta_t, \Lambda^1_t) \), 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 \( V_t(\Pi_t(\Delta_t, \Lambda^1_t)) \) be functions defined recursively for all \( g \in \mathcal{G}^* \) by
\[
V_T(\Pi_T(\Delta_T, \Lambda^1_T)) := \mathbb{E}^g[ c_T(X_T) \mid \Pi_T = \pi_T ],
\]
\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) := \inf_{U_t^{1:t}} \mathbb{E}^g[ c_t(X_t, U_t^{1:t}) + V_{t+1}(\theta_t(\Pi_t(\Delta_t, \Lambda^1_t), Y_{t+1}^{1:t}, U_t^{1:t}) \mid \Pi_t = \pi_t, U_t^{1:t} = u_t^{1:t}) ],
\]
where \( c_T(X_T) \) is the cost function at \( T \), and \( \pi_T, \pi_t, u_t^{1:t} \) are the realizations of \( \Pi_T \), \( \Pi_t \), and \( U_t^{1:t} \), respectively. Then, for any control strategy \( g \in \mathcal{G}^* \),
\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) \leq J_t(g) := \mathbb{E}^g[ \sum_{l=t}^{T-1} c_l(X_l, U_l^{1:l}) + c_T(X_T) \mid \Delta_t, \Lambda^1_t ].
\]
where \( J_t(g) \) is the cost-to-go of the team at time \( t \) corresponding to the control strategy \( g \in \mathcal{G}^* \).

Proof. We prove \( (30) \) by induction. For \( t = T \),
\[
J_T(g) := \mathbb{E}^g[ c_T(X_T) \mid \Delta_T, \Lambda^1_T ]
\]
\[
= \int_{\mathcal{X}_T} c_T(X_T) \Pi_T(\Delta_T, \Lambda^1_T)(X_T) dX_T,
\]
and so \( (30) \) holds with equality.

Suppose that \( (30) \) holds for \( t + 1 \). Then,
\[
J_t(g) := \mathbb{E}^g[ \sum_{l=t}^{T-1} c_l(X_l, U_l^{1:l}) + c_T(X_T) \mid \Delta_t, \Lambda^1_t ]
\]
\[
= \mathbb{E}^g[ c_t(X_t, U_t^{1:t}) + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^{1:l}) + c_T(X_T) \mid \Delta_t, \Lambda^1_t ]
\]
\[
\geq \mathbb{E}^g[ c_t(X_t, U_t^{1:t}) + V_{t+1}(\theta_t(\Pi_t(\Delta_t, \Lambda^1_t), Y_{t+1}^{1:t}, U_t^{1:t}) \mid \Pi_t = \pi_t, U_t^{1:t} = u_t^{1:t}) ]
\]
\[
= \mathbb{E}^g[ V_t(\Pi_t(\Delta_t, \Lambda^1_t)) \mid \Delta_t, \Lambda^1_t ] = V_t(\Pi_t(\Delta_t, \Lambda^1_t)),
\]
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
\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) := \inf_{U_t^{1:t} \in \mathcal{U}_t} \mathbb{E}^g[ c_t(X_t, U_t^{1:t}) + V_{t+1}(\theta_t(\Pi_t(\Delta_t, \Lambda^1_t), Y_{t+1}^{1:t}, U_t^{1:t}) \mid \Pi_t = \pi_t, U_t^{1:t} = u_t^{1:t}) ],
\]
and let \( g \in \mathcal{G}^* \) be a separated control strategy that achieves the infimum in \( (33) \). Then \( g \in \mathcal{G}^* \) is optimal and
\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) = J_t(g),
\]
with probability 1.

**Proof.** We first prove (34) by induction. For \( t = T \),
\[
J_T(\mathbf{g}) = \mathbb{E}^\mathbf{g}\left[c_T(X_T) \mid \Delta_T, \Lambda_{t+1}^K\right] = \int_{X_T} c_T(X_T) \Pi_T(\Delta_T, \Lambda_{t+1}^K)(X_T) dX_T. \tag{35}
\]
Suppose that (33) holds for \( t + 1 \). Then
\[
\inf_{u_{t+1}^1 \in \Pi_{t+1} \subseteq K} \mathbb{E}\left[ \sum_{l=t}^{T-1} c_l(X_l, U_l^1) + c_T(X_T) \mid \Delta_t, \Lambda_t^1\right]
= \inf_{u_{t+1}^1 \in \Pi_{t+1} \subseteq K} \mathbb{E}\left[ c_t(X_t, U_t^1) + \sum_{l=t+1}^{T-1} c_l(X_l, U_l^1) + c_T(X_T) \mid \Delta_t, \Lambda_t^1\right]
+ c_T(X_T) \mid \Delta_t, \Lambda_t^1
= \inf_{u_{t+1}^1 \in \Pi_{t+1} \subseteq K} \mathbb{E}\left[ c_t(X_t, U_t^1) \right]
+ \sum_{l=t+1}^{T-1} c_l(X_l, U_l^1) + c_T(X_T) \mid \Delta_t, \Lambda_t^1
\]
where, in the third equality, we used the hypothesis and, in the fourth equality, \( u_{t+1}^1 \) achieves the infimum. Thus, (33) holds for all \( t \).

For \( t = 0 \), (34) yields \( J_0(\mathbf{g}) = V_0(\Pi_0(\Delta_0, \Lambda_0^1)) \). Taking expectations
\[
J(\mathbf{g}) = \mathbb{E}^\mathbf{g}\left[ J_0(\mathbf{g}) \right] = \mathbb{E}^\mathbf{g}\left[ V_0(\Pi_0(\Delta_0, \Lambda_0^1)) \right]. \tag{37}
\]
By Theorem 2, it follows that for any other \( \mathbf{g}' \in \mathcal{G} \),
\[
J(\mathbf{g}') \geq \mathbb{E}^\mathbf{g}\left[ V_0(\Pi_0(\Delta_0, \Lambda_0^1)) \right]. \tag{38}
\]

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

**Lemma 4.** Let \( V_t(\Pi_t(\Delta_t, \Lambda_t^1)) \) be functions defined recursively for all \( \mathbf{g} \in \mathcal{G}^t \) by
\[
V_T(\Pi_T(\Delta_T, \Lambda_T^1)) := \mathbb{E}^\mathbf{g}\left[c_T(X_T) \mid \Pi_T = \pi_T\right], \tag{39}
\]
\[
V_t(\Pi_t(\Delta_t, \Lambda_t^1)) := \inf_{u_{t+1}^1 \in \Pi_{t+1} \subseteq K} \mathbb{E}^\mathbf{g}\left[ c_t(X_t, U_t^1) + c_{t+1}(\theta_t(\Pi_t(\Delta_t, \Lambda_t^1), Y_{t+1}^1, U_{t+1}^1)) \mid \Pi_t = \pi_t, U_{t+1}^1 = u_{t+1}^1 \right]. \tag{40}
\]

**IV. STRUCTURAL RESULTS FOR THE TEAM MEMBERS**

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

**A. Information State – Team Members**

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

**Definition 3.** An information state, \( \Pi_t^k \), for member \( k \in K \) of a team described by the state equation (1) (a) is a function of \( \Delta_t, \Lambda_t^k \), and (b) \( \Pi_{t+1}^k \) can be determined from \( \Pi_t^k, Y_t^1, \Lambda_t^1, \Lambda_{t-n+1}^1, Y_{t+1}, U_t^1 \).

To proceed, we first need to prove some essential properties of the conditional probabilities densities related to the observation of team member \( k \in K \) and the team’s state.

**Lemma 5.** For any control strategy \( \mathbf{g} = \{g^1, \ldots, g^K\} \) of the team,
\[
p^\mathbf{g}(Y_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k) = p(Y_{t+1} \mid X_{t+1}), \tag{42}
\]
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
\[
p^\mathbf{g}(Y_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_t^k, U_t^k) = p^\mathbf{g}(Y_{t+1}^k \mid X_{t+1}). \tag{43}
\]
However,
\[
p^\mathbf{g}(Y_{t+1}^k \mid X_{t+1}) = p^\mathbf{g}(Z_{t+1}^k \in B^k \mid X_{t+1}), \tag{44}
\]
where \( B^k \subseteq \mathcal{B}^k \). 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
\[
p^\mathbf{g}(Z_{t+1}^k \in B^k \mid X_{t+1}) = p(Z_{t+1}^k \in B^k). \tag{45}
\]
Thus,
\[ p^k(Y_{t+1}^k \mid X_{t+1}^k) = p(Y_{t+1}^k \mid X_{t+1}). \]  
(46)

The result follows from (43) and (46).

**Lemma 6.** For any control strategy \( g = \{g^1, \ldots, g^K\} \) of the team, \( p^k(X_{t+1}^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) \) does not depend on the control strategy \( g^k \) of member \( k \). It depends only on the strategy \( g^k = \{g^1, \ldots, g^{k-1}, g^{k+1}, \ldots, g^K\} \), of the other team members, i.e.,
\[ p^k(X_{t+1}^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) = p^k(X_{t+1}^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k), \]  
(47)

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

**Proof.** Since \( X_{t+1}^k = f_k(X_t, U_{t+1}^k, W_t) \) and \( \Lambda_{t}^k = (Y_{t-n+1:t}^k, U_{t-n+1:t-1}^k) \), and \( Y_{t}^k = h_k(X_t, Z_t^k) \) we have
\[ p^k(X_{t+1}^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) = p^k(W_t \in A, U_{t-n}^k, Z_t^k \in B^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) = p^k(W_t \in A \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k), p^k(U_{t-n}^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) \cdot p^k(Z_t^k \in B^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k), \]  
(48)

where \( A \in \mathcal{F}, B^k \in \mathcal{X}_k, \) and \( U_{t-n}^k = (U_{t-n}^1, U_{t-n}^2, \ldots, U_{t-n}^K) \). In the last equality, the second term depends only on \( g^k \) while we dropped the superscript \( g \) in \( p(W_t \in A \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) \) and \( p(Z_t^k \in B^k \mid \Delta_{t+1}, \Lambda_{t}^k, U_t^k) \) 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 \( g = \{g^1, \ldots, g^K\} \in \mathcal{G}^{dec} \) of the team, the conditional probability density \( p(X_t \mid \Delta_t, \Lambda_t^k) \) does not depend on the control strategy \( g^k \) of member \( k \). It depends only on the strategy \( g^k = \{g^1, \ldots, g^{k-1}, g^{k+1}, \ldots, g^K\} \) of the other team members. It is an information state of the team member \( k \), i.e., \( \Pi_t^k(\Delta_t, \Lambda_t^k)(X_t) = p^k(X_t \mid \Delta_t, \Lambda_t^k) \), that can be evaluated from \( \Delta_t \) and \( \Lambda_t^k \). Moreover, there is a function \( \vartheta_t^k \), which does not depend on the control strategy \( g^k \) of member \( k \), such that
\[ \Pi_t^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) = \vartheta_t^k (\Pi_t^k(\Delta_t, \Lambda_t^k)(X_t), \Delta_{t+1}, \Lambda_{t+1}^k), \]  
(49)

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

**Proof.** See Appendix [D].

Note that the information state \( \Pi_t^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) = p(X_{t+1}^k \mid \Delta_{t+1}, \Lambda_{t+1}^k) \) 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 \( \Pi_t^k(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) \) for any particular \( X_{t+1}^k \), we need the probability density functions \( p(\cdot \mid \Delta_t, \Lambda_t^k, U_t^k) \) and \( \rho(\cdot \mid \Delta_t, \Lambda_t^k) \). This implies that the information state takes values in the space of these probability densities on the measurable space \((\mathcal{X}_t, \mathcal{F}_t)\), which is an infinite-dimensional space.

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

**Proof.** By applying Bayes’ rule, for all \( t = 0, 1, \ldots, T - 1 \), we have
\[ \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) = p(X_{t+1} \mid \Delta_{t+1}, \Lambda_{t+1}^k) \cdot p(\Lambda_{t+1}^k \mid X_{t+1}, \Delta_{t+1}, \Lambda_{t+1}^k) \cdot p(\Delta_{t+1}, \Lambda_{t+1}^k), \]  
(50)

where \( \Lambda_{t+1}^k \) is the cost function at \( \vartheta_t^k \) such that
\[ \Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}) = \vartheta_t^k (\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)(X_{t+1}), \Delta_{t+1}, \Lambda_{t+1}^k), \]  
(51)

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

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

**B. Optimal Control Strategy of the Team Members**

In view of Theorem 5, we show that the optimal separated control strategy \( g = \{g^1, \ldots, g^K\} \), i.e., \( U_t^k = g^k(\Pi_t(\Delta_t, \Lambda_t^k)) \), derived by each team member \( k \in \mathcal{K} \) yields the same solution as the one by the manager’s optimal separated control strategy \( g \in \mathcal{G} \) (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 \( g^k \) be a separated control strategy of team member \( k \in \mathcal{K} \) and \( g = \{g^1, \ldots, g^K\} \) be the team’s control strategy. We fix \( g^k = \{g^1, \ldots, g^{k-1}, g^{k+1}, \ldots, g^K\} \), and let \( V_t^k(\Pi_t(\Delta_t, \Lambda_t^k)) \) be functions defined recursively for all \( g \) by
\[ V_t^k(\Pi_t(\Delta_t, \Lambda_t^k)) = \inf_{U_t^k \in \mathcal{U}_t^k} \mathbb{E} [c_t(X_t) \mid \Pi_t^k = \pi_t^k], \]  
(52)

\[ V_t^k(\Pi_t(\Delta_t, \Lambda_t^k)) = \inf_{U_t^k \in \mathcal{U}_t^k} \mathbb{E} [c_t(X_t, U_t^k) + V_{t+1}^k(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)) \mid \Pi_t^k = \pi_t^k, \Delta_t = \delta_t, \Lambda_t^k = \lambda_t^k, U_t^k = u_t^k], \]  
(53)

where \( c_t(X_t) \) is the cost function at \( T_t \), and \( \pi_t^k, \delta_t, \lambda_t^k, u_t^k \) are the realizations of \( \Pi_t(\Delta_t, \Lambda_t^k) \) and \( \mathcal{U}_t^k \), 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 \( g^k \) be a separated control strategy of team member \( k \in K \) and \( g = \{g^1, \ldots, g^K\} \) be the team’s control strategy. We fix \( g^k = (g^{k,1}, \ldots, g^{k,k-1}, g^{k,k+1}, \ldots, g^K) \), and let \( V^k_t(\Pi_t(\Delta_t, \Lambda_t^k)) \) be functions defined recursively for all \( g \) by

\[
V^k_t(\Pi_t(\Delta_t, \Lambda_t^k)) := \inf_{u^k_t \in U^k_t} \mathbb{E}^k \left[ c_t(X_t, U^k_t, U_{t-k}) + V^k_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)) \mid \Pi_t = \pi^k_t, \Delta_t = \delta_t, \Lambda_t^k = \lambda_t^k, U_t = u_t^k \right],
\]

(54)

where \( \pi^k_t, \delta_t, \lambda_t^k, u_t^k \) are the realizations of \( \Pi^k_t, \Delta_t, \Lambda_t^k, \) and \( U_t^k \), respectively. Then, \( V^k_t(\Pi_t(\Delta_t, \Lambda_t^k)) \) 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

\[
V_T(\Pi_T(\Delta_T, \Lambda_T^{1-K})) := \mathbb{E}^T \left[ c_T(X_T) \mid \Pi_T = \pi_T \right],
\]

(55)

\[
V_t(\Pi_t(\Delta_t, \Lambda_t^{1-K})) := \inf_{u^{1-K}_t \in U^{1-K}_t} \mathbb{E}^{1-K} \left[ c_t(X_t, U^{1-K}_t) + V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^{1-K})) \mid \Pi_t = \pi_t, U^{1-K}_t = u^{1-K}_t \right],
\]

(56)

where \( \pi_t, u^{1-K}_t \) are the realizations of \( \Pi_t \) and \( U^{1-K}_t \), respectively, and let \( g^* \in G^* \) be the manager’s optimal separated control strategy which achieves the infimum in (55) for all \( t = 0, 1, \ldots, T - 1 \). Let \( g = (g^1, \ldots, g^{k-1}, g^k, g^{k+1}, \ldots, g^K) \) be the team’s strategy, where \( g^k = \{g^k_0, \ldots, g^k_{T-1}\} \) is a separated control strategy of member \( k \in K \) such that \( u^k_t = g^k_t(\Pi^k_t(\Delta_t, \Lambda_t^k)) \). Let \( V^k_t(\Pi^k_t(\Delta_t, \Lambda_t^k)) \) be functions defined recursively by each team player \( k \in K \) after fixing \( g^k = (g^{k,1}, \ldots, g^{k,k-1}, g^{k,k+1}, \ldots, g^K) \), by

\[
V^k_T(\Pi_T^k(\Delta_T, \Lambda_T^k)) := \mathbb{E}^k \left[ c_T(X_T) \mid \Pi^k_T = \pi^k_T \right],
\]

(57)

and

\[
V^k_t(\Pi^k_t(\Delta_t, \Lambda_t^k)) := \inf_{u^k_t \in U^k_t} \mathbb{E}^k \left[ c_t(X_t, U^k_t, U_{t-k}) + V^k_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda_{t+1}^k)) \mid \Pi^k_t = \pi^k_t, \Delta_t = \delta_t, \Lambda^k_t = \lambda^k_t, U_t = u^k_t \right],
\]

(58)

where \( U^k_t = (U^1_t, \ldots, U^{k-1}_t, U^{k+1}_t, \ldots, U^K_t) \), and \( \pi^k_t, \delta_t, \lambda^k_t, u^k_t \) are the realizations of \( \Pi^k_t, \Delta_t, \Lambda^k_t, \) and \( U^k_t \), respectively. Then, the solution of the manager in (55) 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.

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 \( X_t = (X^1_t, X^2_t) \), \( t = 1, 2, 3 \), of the team is two-dimensional, and the initial state (primitive random variable), \( X_0 = (X^1_0, X^2_0) \), of the team is a Gaussian random variable with zero mean, variance 1, and covariance 0.5.

The team’s state evolves as follows:

\[
X_0 = (X^1_0, X^2_0),
\]

(59)

\[
X_1 = (X^1_1, X^2_1) = (X^1_0 + X^2_0, 0),
\]

(60)

\[
X_2 = (X^1_2, X^2_2) = (X^1_1, U^2_1) = (X^1_0 + X^2_0, U^2_2),
\]

(61)

\[
X_3 = (X^1_3, X^2_3) = (X^1_2 - X^2_2 - U^3_1, 0)
\]

(62)

\[
= (X^1_0 + X^2_0 - U^2_2 - U^3_1, 0).
\]

The observation equations are

\[
Y^k_t = X^k_{t-1}, \quad k = 1, 2; \quad t = 1, 2, 3.
\]

(63)

Each team member’s feasible sets of actions \( U^k_t \) are specified by

\[
U^k_t = \begin{cases} \mathbb{R}, & \text{if } (k, t) = (1, 3) \text{ or } (2, 2), \\ 0, & \text{otherwise}. \end{cases}
\]

(64)

Thus a control strategy \( g = \{g^k; k = 1, 2; t = 1, 2, 3\} \) is the solution of the problem consists only of the pair \( g = \{g^2_3, g^2_2 \} \) since \( g^1_2 \equiv 0 \) for the remaining \( (k, t) \). Given the modeling framework above, the information structure \( \{\Pi_t(\Delta_t, \Lambda_t^k); k = 1, 2; t = 1, 2, 3\} \) of the team is

\[
\Delta_1 = \emptyset, \quad \Delta_2 = \emptyset,
\]

(65)

\[
\Delta_3 = \{Y^1_T, Y^2_T\} = \{X^1_0, X^2_0\},
\]

(66)

Note that since \( g^1_1 \equiv 0 \) and \( g^2_2 \equiv 0 \), the realizations of \( U^1_1 \) and \( U^2_2 \) are zero, and thus \( \Delta_3 \) includes only the observations in (66). The data \( \Lambda^k_t, k = 1, 2 \), available to the team member \( k \) for the feasible control laws are

\[
\Lambda^1_3 = \{Y^1_3, Y^2_3\} = \{X^1_0, X^2_0\} = \{X^1_0\},
\]

(67)

\[
\Lambda^2_3 = \{Y^2_3\} = \{X^1_0 + X^2_0, X^1_0 + X^2_0\} = \{X^1_0 + X^2_0\}.
\]

(68)

The problem is to derive the optimal control strategy \( g^* = \{g^2_3, g^2_2\} \) which is the solution of

\[
J(g) = \min_{u^2_3 \in U^2_3, u^2_2 \in U^2_2} \frac{1}{2} \mathbb{E}^T \left[ (X^1_T)^2 + (U^3_1)^2 \right]
\]

(69)

\[
= \min_{u^2_3 \in U^2_3, u^2_2 \in U^2_2} \frac{1}{2} \mathbb{E}^T \left[ (X^1_0 + X^2_0 - U^2_2 - U^3_1)^2 + (U^3_1)^2 \right].
\]

B. Optimal Solution

The feasible set \( G \) of control strategies of the team consists of all \( g = \{g^3_3(\Lambda^1_3, \Delta_3), g^2_2(\Lambda^2_3, \Delta_2)\} \), i.e.,

\[
g^2_2; \Delta_2 \times \Lambda^2_3 \rightarrow U^2_2, \quad \text{or} \quad g^2_2; \Delta_2 \rightarrow \mathbb{R},
\]

(70)

\[
g^1_1; \Delta_1 \times \Lambda^2_3 \rightarrow U^1_3, \quad \text{or} \quad g^1_1; \{X^1_0, X^2_0\} \rightarrow \mathbb{R}.
\]

(71)

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

\[
U^2_3 = \frac{1}{2} X^2_0, \quad U^1_3 = \frac{1}{2} (X^1_0 + X^2_0) - \frac{1}{4} X^2_0.
\]

(72)
C. Solution Given by Theorem

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 step $t = 1$, each team member $k = 1, 2$ observes $Y^k_1 = X_0$ and makes a decision $U^k_1 = g^k_1(\Delta_1, \Lambda^k_1)$. The transition of the team to the next state occurs at the same time $t = 1$, i.e., $X_1 = f_1(X_0, U^1_1, U^2_1)$, and the process is repeated until $t = T$. Given that the decision of each team member $k$ at time $t$ depends on $\Delta_1$ and $\Lambda^k_1$, and that the state of the team evolves after the realization of $U^1_1, U^2_1$, 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:

$$X_0 = (X^1_0, X^2_0),$$
$$X_1 = (X^1_1, X^2_1) = (X^1_0, X^2_0),$$
$$X_2 = (X^2_2, X^1_2) = (X^1_0 + X^2_0, 0),$$
$$X_3 = (X^1_3, X^2_3) = (X^1_2 + X^2_2, U^2_2),$$
$$X_4 = (X^1_4, X^2_4) = (X^1_3 - X^2_3 - U^3_3, 0) = (X^1_0 + X^2_0, U^2_2 - U^3_3, 0),$$

where we essentially included a degenerate transition at $t = 1$.

Given the modeling framework presented in Section II, B, the information structure $\{(\Delta_t, \Lambda^k_t); k = 1, 2; t = 1, 2, 3\}$ of the team is

$$\Delta_1 = \emptyset, \quad \Delta_2 = \emptyset,$$
$$\Delta_3 = \{Y^1_0, Y^2_0, Y^1_1, Y^2_1\} = \{X^1_0, X^2_0, X^1_1, X^2_1\} = \{X^1_0, X^2_0\}. $$

Since $g^1_t = 0$ and $g^2_t = 0$, the realizations of $U^1_1$ and $U^2_1$ are zero, and thus $\Delta_3$ includes only the observations in $\{X^1_0, X^2_0\}$. The data $\Lambda^k_t, k = 1, 2$, available to the team member $k$ for the feasible control laws are

$$\Lambda^k_2 = \{Y^2_0, Y^2_1, Y^2_2\} = \{X^2_0, X^2_1, X^2_2\} = \{X^2_0\},$$
$$\Lambda^k_3 = \{Y^2_1, Y^2_2\} = \{X^2_0 + X^2_1, X^2_0\} = \{X^2_0 + X^2_1\}. $$

We solve problem (69) by using the structural results presented in Section III, considering the control strategies $g^k_t = \{g^k_1; k = 1, 2; t = 0, 1, 2, 3\}$ for each team member $k$, where the control law is of the form $g^k_1(\Pi(\Delta_1, \Lambda^k_1)) = g^k_1(\mathbb{P}(X^k_1 | \Delta_1, \Lambda^k_1))$.

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

$$V_3(\Pi_3) = \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ (X^1_0 + X^2_0 - U^2_2 - U^3_3)^2 \right] + (U^3_3)^2 | \Pi_3(\Delta_3, \Lambda^3_3), U^3_3$$

$$= \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ (X^1_0 + X^2_0 - U^2_2 - U^3_3)^2 \right] + (U^3_3)^2 | \mathbb{P}(X^1_0 + X^2_0, U^2_2 | X^1_0, X^2_0, X^1_0 + X^2_0, U^3_3),$$

where, for any given realization of $((X^1_0 + X^2_0), U^2_2)$ in the information state $\Pi_3$, the manager selects $U^3_3$ to achieve the lower bound in (82). Thus,

$$U^3_3 = \frac{1}{2}(X^1_0 + X^2_0) - \frac{1}{2}U^2_2. \quad (83)$$

Substituting (83) into (82) yields

$$V_3(\Pi_3) = \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ \frac{(X^1_0 + X^2_0 - U^2_2)^2}{2} \right] | \mathbb{P}(X^1_0 + X^2_0, U^2_2 | X^1_0, X^2_0, X^1_0 + X^2_0, U^3_3).$$

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

$$V_2(\Pi_2) = \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ V_3(\Pi_3) | \Pi_2(\Delta_2, \Lambda^2_2), U^2_2 \right]$$

$$= \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ V_3(\Pi_3) | \mathbb{P}(X^1_0 + X^2_0 | X^1_0, U^2_2) \right]$$

$$= \min_{u^2_3 \in \mathcal{U}^2_3, u^1_3 \in \mathcal{U}^1_3} \frac{1}{2} \mathbb{E}_t^f \left[ \frac{(X^1_0 + X^2_0 - U^2_2)^2}{2} \right] | \mathbb{P}(X^1_0 + X^2_0 | X^1_0, U^2_2).$$

Since

$$U^2_2 = g^2_2(\mathbb{P}(X^2_2 | \Delta_2, \Lambda^2_2)) = g^2_2(\mathbb{P}(X^1_0 + X^2_0 | X_0^2)),$$

the problem of the manager in (85) is to choose, for any given $X^2_0$, the estimate of $(X^1_0 + X^2_0)$ that minimizes the mean squared error $(X^1_0 + X^2_0 - U^2_2)^2$ in (85). Given the Gaussian statistics, the optimal solution is

$$U^2_2 = \frac{1}{2}X^2_0.$$ 

(87)

Substituting (87) into (83) yields

$$U^1_3 = \frac{1}{2}(X^1_0 + X^2_0) - \frac{1}{4}X^2_0.$$ 

(88)

Therefore, the control laws of the form $g^k_t(\Pi(\Delta_t, \Lambda^k_t)) = g^k_t(\mathbb{P}(X^k_1 | \Delta_t, \Lambda^k_t))$ yield the unique optimal solution (72) of problem (69).

VI. CONCLUDING REMARKS AND DISCUSSION

In this paper, 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 are the same as the ones derived by the manager. The key contributions of the paper 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 on 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 (Theorem 4 and Theorem 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, cyberphysical 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.

VII. ACKNOWLEDGMENTS

The author would like to thank Aditya Mahajan, Ashutosh Nayyar, and Serdar Yüksel for several helpful discussions that led to improving the exposition in Section IV.

APPENDIX A

PROOF OF THEOREM 1

By applying Bayes’ rule, we have

\[ p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}) = \int_{\mathcal{X}_{t+1}} p^\theta(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, Y_t^{1:K}, U_t^{1:K}) \, dX_{t+1} \]

Next,

\[ p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}) = p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}, Y_t^{1:K}, U_t^{1:K}) \]

\[ = \int_{\mathcal{X}_{t+1}} p^\theta(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, Y_t^{1:K}, U_t^{1:K}) \, dX_{t+1} \]

\[ = \int_{\mathcal{X}_{t+1}} p^\theta(Y_t^{1:K} | X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^\theta(X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1} \]

\[ = \int_{\mathcal{X}_{t+1}} p^\theta(Y_t^{1:K} | X_{t+1}, \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}, \]

where by Lemma 1 the last equation becomes

\[ p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}) = \int_{\mathcal{X}_{t+1}} p(Y_t^{1:K} | X_{t+1}) p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}. \]

Note that \( p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) = p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \) since \( Y_t^{1:K} \) and \( U_t^{1:K} \) are already included in \( \Lambda_t^{1:K} \), hence we can write (90) as

\[ p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}) = \int_{\mathcal{X}_{t+1}} p(Y_t^{1:K} | X_{t+1}) p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^\theta(\Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}. \]

Substituting (91) into (89), we have

\[ p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}) = p(Y_t^{1:K} | X_{t+1}) p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \int_{\mathcal{X}_{t+1}} p(Y_t^{1:K} | X_{t+1}) p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t+1}, \]

which we can write as

\[ p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}) = \phi_t[p(\cdot | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}), Y_t^{1:K}] \, (X_{t+1}), \]

with the function \( \phi_t \) chosen appropriately.

Next,

\[ p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) = \int_{\mathcal{X}_{t}} p^\theta(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \cdot p^\theta(X_{t} | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \, dX_{t}. \]

By Lemma 2 and Remark 1 (92) becomes

\[ p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) = \int_{\mathcal{X}_{t}} p(X_{t+1} | X_{t}, U_t^{1:K}) p(X_{t} | \Delta_t, \Lambda_t^{1:K}) \, dX_{t}, \]

which we can write as

\[ p^\theta(X_{t+1} | \Delta_{t+1}, \Lambda_t^{1:K}, U_t^{1:K}) = \psi_t[p(\cdot | \Delta_t, \Lambda_t^{1:K}, U_t^{1:K}) \, (X_{t+1}), \]

where in the second equality we used Lemma 1.
with the function \( \psi_t \) chosen appropriately.
Substituting (90) into (93) yields

\[
p^R(X_{t+1} \mid \Delta_{t+1}, \Lambda^1_{t+1}) = \phi_t[\psi_t[p(\cdot \mid \Delta_t, \Lambda^1_{t+1}, U^1_{t+1}], Y^1_{t+1}]](X_{t+1}). \tag{97}
\]

Therefore \( p^R(X_{t+1} \mid \Delta_{t+1}, \Lambda^1_{t+1}) \) does not depend on the control strategy \( g \), so we can drop the superscript. Moreover, we can choose appropriate function \( \theta_t \) such that

\[
p(X_{t+1} \mid \Delta_{t+1}, \Lambda^1_{t+1}) = \theta_t[\Pi_t(\Delta_t, \Lambda^1_{t+1})(X_t), Y^1_{t+1}, U^1_{t+1}]. \tag{98}
\]

**APPENDIX B
PROOF OF LEMMA 4**

Obviously, for \( t = T \),

\[
V_T(\rho \Pi_T(\Delta_T, \Lambda^1_T)) = \int_{\mathcal{X}_T} c_T(X_T) \rho \Pi_T(\Delta_T, \Lambda^1_T)(X_T) \, dX_T = \rho V_T(\Pi_T(\Delta_T, \Lambda^1_T)). \tag{99}
\]

For \( t = 0, \ldots, T-1 \), by assigning \( \Pi_t = \rho \Pi_t \) [recall \( p(X_t \mid \Delta_t, \Lambda^1_t) = \Pi_t(\Delta_t, \Lambda^1_t) \), (40) becomes

\[
V_t(\rho \Pi_t(\Delta_t, \Lambda^1_t)) = \inf_{u^1_t \in \Pi_t} \int_{\mathcal{X}_t} c_t(X_t, U^1_t) \rho \Pi_t(\Delta_t, \Lambda^1_t)(X_t) \, dX_t + \int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} V_{t+1}(\rho \Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \cdot p(Y^1_{t+1} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1} \, dY^1_{t+1}, \tag{100}
\]

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

Next, from (27),

\[
\rho \Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1}) p(Y^1_{t+1} \mid X_{t+1}) = \int_{\mathcal{X}_{t+1}} p(X_{t+1} \mid X_t, U^1_t) \rho \Pi_t(\Delta_t, \Lambda^1_t) \, dX_t = \Pi_t(\Delta_{t+1}, \Lambda^1_{t+1}). \tag{101}
\]

Substituting (101) into (100), we have

\[
V_t(\rho \Pi_t(\Delta_t, \Lambda^1_t)) = \inf_{u^1_t \in \Pi_t} \int_{\mathcal{X}_t} c_t(X_t, U^1_t) \rho \Pi_t(\Delta_t, \Lambda^1_t)(X_t) \, dX_t + \int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \cdot p(Y^1_{t+1} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1} \, dY^1_{t+1} = V_t(\Pi_t(\Delta_t, \Lambda^1_t)). \tag{102}
\]

**APPENDIX C
PROOF OF THEOREM 4**

We start with (41)

\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) = \inf_{u^1_t \in \Pi_t} \int_{\mathcal{X}_t} c_t(X_t, U^1_t) \Pi_t(\Delta_t, \Lambda^1_t)(X_t) \, dX_t + \int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \cdot p(Y^1_{t+1} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1} \, dY^1_{t+1}, \tag{103}
\]

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

Choosing

\[
\rho = \int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} p(Y^1_{t+1} \mid X_{t+1}) \cdot p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1}. \tag{104}
\]

we can use the positive homogeneity of \( V_t(\Pi_t(\Delta_t, \Lambda^1_t)) \) (Lemma 4) to write the second part of (103) as follows

\[
\int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \cdot p(Y^1_{t+1} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1} \, dY^1_{t+1} \tag{105}
\]

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

Thus, we can write (103) as

\[
V_t(\Pi_t(\Delta_t, \Lambda^1_t)) = \inf_{u^1_t \in \Pi_t} \int_{\mathcal{X}_t} c_t(X_t, U^1_t) \Pi_t(\Delta_t, \Lambda^1_t)(X_t) \, dX_t + \int_{\mathcal{X}_t} \int_{\mathcal{X}_{t+1}} V_{t+1}(\Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \cdot p(Y^1_{t+1} \mid X_{t+1}) \, p(X_{t+1} \mid X_t, U^1_t) \rho \rho(X_t \mid \Delta_t, \Lambda^1_t) \, dX_t \, dX_{t+1} \, dY^1_{t+1}. \tag{106}
\]

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

\[
V_{t+1}(\rho \Pi_{t+1}(\Delta_{t+1}, \Lambda^1_{t+1})) \tag{107}
\]

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
\[
\int_{\mathcal{X}} V_{t+1} \left( p(Y_{t+1}^{1:K} | X_{t+1}) \int_{\mathcal{X}} p(X_{t+1} | X_t, U_t^{1:K}) \right. \\
\left. p(X_t | \Delta_t, \Lambda_t^{1:K}) \right) dX_t
\]
\[dY_t^{1:K} \] (108)
is concave. Since the pointwise infimum of concave functions is concave, ([106] is concave.

**APPENDIX D**

**PROOF OF THEOREM**

By applying Bayes’ rule, we have
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = \frac{p^k(X_{t+1}, \Delta_t, \Lambda_t^{1:K}, U_t^k)}{p^k(\Delta_t, \Lambda_t^{1:K})}
\]
\[= \frac{p^k(Y_{t+1}^k | X_{t+1}, \Delta_t, \Lambda_t^k, U_t^k)}{p^k(\Delta_t, \Lambda_t^{1:K})} p^k(X_{t+1} | \Delta_t, \Lambda_t^k, U_t^k)
\]
(109)

where in the second equality we used Lemma 5.

Next,
\[
p^k(\Delta_t, \Lambda_t^{1:K}) = \int_{\mathcal{X}^1} p^k(X_{t+1}, \Delta_t, \Lambda_t^k, Y_{t+1}^1, U_t^k)
\]
\[= \int_{\mathcal{X}^1} p^k(X_{t+1}, \Delta_t, \Lambda_t^k, Y_{t+1}^1, U_t^k) dX_{t+1}
\]
\[= \int_{\mathcal{X}^1} p^k(Y_{t+1}^1 | X_{t+1}, \Delta_t, \Lambda_t^k, U_t^k)
\cdot \ p^k(X_{t+1} | \Delta_t, \Lambda_t^k, U_t^k) dX_{t+1}
\]
\[= \int_{\mathcal{X}^1} p^k(Y_{t+1}^1 | X_{t+1}, \Delta_t, \Lambda_t^k, U_t^k)
\cdot \ p^k(X_{t+1} | \Delta_t, \Lambda_t^k, U_t^k) dX_{t+1},
\]
where by Lemma 5 the last equation becomes
\[
p^k(\Delta_t, \Lambda_t^{1:K}) = \int_{\mathcal{X}^1} p^k(Y_{t+1}^1 | X_{t+1}, \Delta_t, \Lambda_t^k, U_t^k)
\cdot \ p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) dX_{t+1}.
\]
(111)

Substituting (111) into (109), we have
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = \frac{p^k(Y_{t+1}^k | X_{t+1}, \Delta_t, \Lambda_t^k, U_t^k)}{p^k(Y_{t+1}^1 | X_{t+1}, \Delta_t, \Lambda_t^{1:K})} p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) dX_{t+1},
\]
(112)

which we can write as
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = \phi^k \left[ p^k(X_{t+1} | \Delta_t, \Lambda_t^k, U_t^k) \right] (X_{t+1}),
\]
(113)

with the function \(\phi^k\) chosen appropriately.

By Lemma 6, \(p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K})\) depends only on the control strategy \(g^k\), so we can write \(p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K})\) as
\[
\int_{\mathcal{X}} p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^k) dX_t = \int_{\mathcal{X}} p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}, U_t^k) dX_t,
\]
(114)

where the last term in (114) can be written as
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = p^k(X_{t+1} | \Delta_t, Y_t^{1:K}, U_t^{1:K}, \Lambda_t^k) = p^k(X_{t+1} | \Delta_t, Y_t^{1:K}, U_t^{1:K}, \Lambda_t^k)
\]
\[= p^k(X_{t+1} | \Delta_t, \Lambda_t^k) \cdot p^k(Y_t^{1:K}, U_t^{1:K} | \Delta_t, \Lambda_t^k),
\]
(115)

where \(Y_t^{1:K} = (Y_t^{1:n} + 1, Y_t^{1:n} + 2, \ldots, Y_t^K + 1, \ldots, Y_t^{K + 1})\), and \(U_t^{1:K} = (U_t^{1:n} + 1, U_t^{1:n} + 2, \ldots, U_t^{K + 1})\). In the second equality, we dropped \(Y_t^{1:K}, U_t^{1:K}\) from conditioning since they both are included in \(\Lambda_t^k\). Substituting (113) into (114) yields
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = \phi^k \left[ p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) \right] (X_{t+1}).
\]
(117)

Therefore, \(p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K})\), does not depend on the control strategy \(g^k\) of the team member \(k\), so we can adjust the subscript accordingly. Moreover, we can choose appropriate function \(\phi^k\) such that
\[
p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) = \phi^k \left[ p^k(X_{t+1} | \Delta_t, \Lambda_t^{1:K}) \right] (X_{t+1}).
\]
(118)

**APPENDIX E**

**PROOF OF THEOREM**

Let \(g^* = \{g^{1:1}, g^{1:2}, \ldots, g^{k+1}, g^{k+2}, \ldots, g^{K}\}\) be the optimal separated control strategy of the manager which achieves the infimum in (55)-(56). Starting with (56), we have
Continuing backward in time, it follows that for all \( t = 0, 1, \ldots, T \), and for any \( u_t^{-k} \),
\[
\arg \inf_{u_t \in \mathcal{U}_t} \left[ \int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \, dX_t + \mathbb{E}^\pi \left[ V_{t+1} \left( \Pi_{t+1}^{k=1} (\Delta_{t+1}, A_t^{1:K}) (X_{t+1}), \Delta_t, A_t^{1:K} \right) \right] \right] \tag{123}
\]

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

\[V_t(\Pi_t(\Delta_t, A_t^{1:K})) = \inf_{u_t^k \in \mathcal{U}_t^k} \mathbb{E}^\pi \left[ c_t(X_t, U_t^k, U_t^{-k}) \right] \]

\[+ V_{t+1}(\Pi_{t+1}^{k=1} (\Delta_{t+1}, A_t^{1:K})) \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \]

\[= \inf_{u_t^k \in \mathcal{U}_t^k} \inf_{u_t^{-k} \in \mathcal{U}_t^{-k}} \mathbb{E}^\pi \left[ c_t(X_t, U_t^k, U_t^{-k}) \right] \cdot \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \]

\[+ V_{t+1}(\Pi_{t+1}^{k=1} (\Delta_{t+1}, A_t^{1:K})) \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \]

\[= \inf_{u_t^k \in \mathcal{U}_t^k} \inf_{u_t^{-k} \in \mathcal{U}_t^{-k}} \left[ \int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \, dX_t + \Pi_t = \pi_t, \Delta_t = \delta_t, A_t^{1:K} = \lambda_t^{1:K}, U_t^{k=1} = u_t^k \right] \tag{119}\]

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

\[V_t(\Pi_t(\Delta_t, A_t^{1:K})) = \inf_{u_t^k \in \mathcal{U}_t^k} \left[ \int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \right] \cdot \Pi_t = \pi_t, \Delta_t = \delta_t, A_t^{1:K} = \lambda_t^{1:K}, U_t^{k=1} = u_t^k \]

\[+ \mathbb{E}^\pi \left[ V_{t+1} \left( \Pi_{t+1}^{k=1} (\Delta_{t+1}, A_t^{1:K}) (X_{t+1}), \Delta_t, A_t^{1:K} \right) \right] \mid \Pi_t = \pi_t, U_t^{1:K} = u_t^{1:K} \]

\[= \inf_{u_t^k \in \mathcal{U}_t^k} \left[ \int_{\mathcal{X}_t} c_t(X_t, U_t^k, U_t^{-k}) \, dX_t + \Pi_t = \pi_t, \Delta_t = \delta_t, A_t^{1:K} = \lambda_t^{1:K}, U_t^{k=1} = u_t^k \right] \tag{120}\]

where \( \pi_t^{1:k} \), \( \Delta_t \), \( \lambda_t^{1:K} \), and \( U_t = u_t^k \) are the realizations of \( \Pi_t^{k=1} \), \( \Delta_t \), \( A_t^{1:K} \), and \( U_t^k \), respectively. Given that the function \( \vartheta_t^k \) is increasing with \( \Pi_t^{k=1}(\Delta_t, A_t^{1:K}) \) for all \( k \) and \( t \) (Corollary 1), from (120), it follows that at \( t = T - 1 \) and for any \( u_{T-1}^{-k} \)

\[\arg \inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[ \int_{\mathcal{X}_{T-1}} c_{T-1}(X_{T-1}, U_{T-1}^k, U_{T-1}^{-k}) \, dX_{T-1} \right] \cdot \vartheta_{T-1}^k \left( \Pi_{T-1}^{k=1} (\Delta_{T-1}, A_{T-1}^{1:K}) (X_{T-1}), \Delta_{T-1}, A_{T-1}^{1:K} \right) \, dX_{T-1} \]

\[+ \mathbb{E}^\pi \left[ V_{T-1} \left( \Pi_{T-1}^{k=1} (\Delta_{T-1}, A_{T-1}^{1:K}) \right) \right] \mid \Pi_{T-1}^{k=1} (\Delta_{T-1}, A_{T-1}^{1:K}) \]

\[= \arg \inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[ \int_{\mathcal{X}_{T-1}} c_{T-1}(X_{T-1}, U_{T-1}^k, U_{T-1}^{-k}) \, dX_{T-1} \right] \cdot \Pi_{T-1}^{k=1} (\Delta_{T-1}, A_{T-1}^{1:K}) \, dX_{T-1} \]

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

\[\arg \inf_{u_{T-1}^k \in \mathcal{U}_{T-1}^k} \left[ V_{T-1} \left( \Pi_{T-1}^{k=1} (\Delta_{T-1}, A_{T-1}^{1:K}) \right) \right] \]
16

[68] ——, “Structural results for decentralized stochastic control with a
word-of-mouth communication,” in 2020 American Control Conference
(ACC). IEEE, 2020, pp. 2796–2801.

[69] A. Nayyar and D. Teneketzis, “Common knowledge and sequential
team problems,” IEEE Transactions on Automatic Control, vol. 64,
no. 12, pp. 5108–5115, 2019.

[70] A. Nayyar, A. Mahajan, and D. Teneketzis, “Decentralized stochastic
control with partial history sharing: A common information approach,”
IEEE Transactions on Automatic Control, vol. 58, no. 7, pp. 1644–
1658, 2013.

[71] C. D. Charalambous and N. U. Ahmed, “Equivalence of decentralized
stochastic dynamic decision systems via girsanov’s measure transforma-
tion,” in 53rd IEEE Conference on Decision and Control, 2014, pp.
439–444.

[72] C. D. Charalambous, “Decentralized optimality conditions of stochastic
differential decision problems via girsanov’s measure transformation,”
Mathematics of Control, Signals, and Systems, 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. Springer
International Publishing, 2015, pp. 155–163.

[74] ——, “Centralized versus decentralized optimization of distributed
stochastic differential decision systems with different information
structures—part i: A general theory,” IEEE Transactions on Automatic
Control, vol. 62, no. 3, pp. 1194–1209, 2017.

[75] ——, “Centralized versus decentralized optimization of distributed
stochastic differential decision systems with different information
structures—part ii: Applications,” IEEE Transactions on Automatic
Control, vol. 63, no. 7, pp. 1913–1928, 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). North-Holland Pub. Co, 1972.

[77] Yu-Chi Ho and K’ai-Ching Chu, “Team decision theory and informa-
tion,” in Managerial Economics) Volume in Honor of Jacob Marschak (Studies in Mathematical and
Managerial Economics). North-Holland Pub. Co, 1972.

[78] S. Yüksel, “On optimal causal coding of partially observed markov
sources in single and multiterminal settings,” IEEE Transactions on
Automatic Control, vol. 23, no. 4, pp. 627–632, 1978.

[79] G. Casalino, F. Davoli, R. Minciardi, P. Puliafito, and R. Zoppoli,
“Partially nested information structures with a common past,” IEEE
transactions on automatic control, vol. 29, no. 9, pp. 846–850, 1984.

[80] Y. Kaspi and N. Merhav, “Structure theorem for real-time variable-
rate lossy source encoders and memory-limited decoders with side
information,” in IEEE International Symposium on Information Theory -
Proceedings, 2010, pp. 86–90.

[81] S. Yüksel, “On optimal causal coding of partially observed markov
sources in single and multiterminal settings,” IEEE Transactions on
Information Theory, vol. 59, no. 1, pp. 424–437, 2013.

[82] D. Teneketzis and P. Varaiya, “The Decentralized Quickest Detection
Problem,” IEEE Transactions on Automatic Control, vol. 29, no. 7, pp.
641–644, 1984.

[83] V. V. Veeravalli, T. Başar, and H. V. Poor, “Decentralized Sequential
Detection with a Fusion Center Performing the Sequential Test,” IEEE
Transactions on Information Theory, vol. 39, no. 2, pp. 433–442, 1993.

[84] J. N. Tsitsiklis, “Decentralized detection,” in Advances in Statistical
Signal Processing, vol. 2, pp. 297–344, 1993.

[85] V. V. Veeravalli, “Decentralized quickest change detection,” IEEE
Transactions on Information Theory, vol. 47, no. 4, pp. 1657–1665,
2001.

[86] P. Varaiya and J. Walrand, “Causal coding and control for Markov
chains,” Systems and Control Letters, vol. 3, no. 4, pp. 189–192, 1983.

[87] A. Mahajan and D. Teneketzis, “Optimal performance of networked
control systems with nonclassical information structures,” SIAM Jour-
nal on Control and Optimization, vol. 48, no. 3, pp. 1377–1404, 2009.

[88] A. Mahajan, “Sequential decomposition of sequential dynamic teams:
Applications to real-time communication and networked control sys-
tems,” Ph.D. dissertation, University of Michigan, 2008.

[89] S. Yüksel, “A universal dynamic program and refined existence results
for decentralized stochastic control,” SIAM Journal on Control and
Optimization, vol. 58, no. 5, pp. 2711–2739, 2020.

[90] A. Mahajan and D. Teneketzis, “On the design of globally optimal
communication strategies for real-time noisy communication systems
with noisy feedback,” IEEE Journal on Selected Areas in Communi-
cations, vol. 26, no. 4, pp. 1440–1449, 2008.

[91] ——, “Optimal design of sequential real-time communication systems,”
IEEE Transactions on Information Theory, vol. 55, no. 11, pp. 5317–
5338, 2009.

[92] A. Mahajan, A. Nayyar, and D. Teneketzis, “Identifying tractable
decentralized control problems on the basis of information structure,”
in 2008 46th Annual Allerton Conference on Communication, Control,
and Computing, 2008, pp. 1440–1449.

[93] A. Mahajan, “Optimal decentralized control of coupled subsystems
with control sharing,” IEEE Transactions on Automatic Control, vol.
58, no. 9, pp. 2377–2382, 2013.

[94] A. Nayyar, A. Gupta, C. Langbort, and T. Basar, “Common information
based markov perfect equilibria for stochastic games with asymmetric
information: Finite games,” IEEE Transactions on Automatic Control,
vol. 59, no. 3, pp. 555–570, 2014.

[95] J. Arabneydi and A. Mahajan, “Team optimal control of coupled
subsystems with mean-field sharing,” in 53rd IEEE Conference on
Decision and Control, Dec 2014, pp. 1669–1674.

[96] T. Yoshikawa, “Decomposition of dynamic team decision problems,”
IEEE Transactions on Automatic Control, vol. 23, no. 4, pp. 627–632,
1978.

[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 Transactions on Automatic Control,
vol. 59, no. 3, pp. 850–570, 2014.

[98] J. Arabneydi and A. Mahajan, “Team optimal control of coupled
subsystems with mean-field sharing,” in 53rd IEEE Conference on
Decision and Control, Dec 2014, pp. 1669–1674.

[99] A. Nayyar, A. Mahajan, and D. Teneketzis, The Common-Information
Approach to Decentralized Stochastic Control. Cham: Springer
International Publishing, 2014, pp. 123–156.

[100] E. J. Sondik, “The optimal control of partially observed markov
processes,” Dissertation. Stanford University, 1971.

[101] R. D. Smallwood and E. J. Sondik, “The Optimal Control of Partially
Observable Markov Processes Over a Finite Horizon,” Operations
Research, vol. 21, no. 5, pp. 1071–1088, Sept. 1973.

[102] A. A. Malikopoulos, “Separation of learning and control for cyber-
physical systems,” 2022 arXiv:2107.06379.

[103] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge
University Press, 2004.

Andreas A. Malikopoulos (S’06–M’09–SM’17) received the Diploma in mechanical engineering
from the National Technical University of Athens,
Greece, in 2000. He received M.S. and Ph.D.
degrees from the department of mechanical en-
ingineering at the University of Michigan, Ann
Arbor, Michigan, USA, in 2004 and 2008, re-
spectively. He is the Terri Connor Kelly and John
Kelly Career Development Associate Professor in
the Department of Mechanical Engineering at the
University of Delaware, 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 at Oak Ridge National Laboratory,
and a Senior Researcher with General Motors Global Research &
Development. His research spans several fields, including analysis, opti-
mization, 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. He has
been an Associate Editor of the IEEE Transactions on Intelligent Vehi-
cles and IEEE Transactions on Intelligent Transportation Systems from
2017 to 2020. He is currently an Associate Editor of Automatica
and IEEE Transactions on Automatic Control. He is a member of SIAM,
AAAS, and a Fellow of the ASME.