2021 60th IEEE Conference on Decision and Control (CDC)
December 13-15, 2021. Austin, Texas

Private Information Compression in Dynamic Games among Teams

Dengwang Tang, Hamidreza Tavafoghi, Vijay Subramanian, Ashutosh Nayyar, Demosthenis Teneketzis

2021 60th IEEE Conference on Decision and Control (CDC) | 978-1-6654-3659-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CDC45484.2021.9683479

***Abstract*— We investigate finite-horizon stochastic dynamic games among teams. Each team has its own dynamic system, whose evolution is affected by the actions of all players in all teams. Within each team, members share their local states with each other with a delay of $d > 0$. Actions are observed by all agents along with noisy observations of the systems. Such games feature the difficulties of the increasing domain of strategies and interdependence of actions and information over time. In these games, we identify a subclass of Nash Equilibria where the agents use Sufficient Private Information Based (SPIB) strategies, i.e. agents make decisions based on compressed versions of their private information along with the common information. We establish the existence of such equilibria; the proof of existence is not based on standard techniques since SPIB strategies do not feature perfect recall. Finally, we investigate a special case of our model where each agent has their own dynamic system. We show that agents can compress their private information further in this case. Our results provide a foundational step in addressing the difficulties of dynamic games among teams.**

I. INTRODUCTION

In numerous engineering and socioeconomic applications, multiple agents/players participate in dynamic games with asymmetric information. In these games, agents make decisions over time on top of a dynamically evolving physical environment in order to achieve their respective long-term goal. Examples of such applications include sensor networks, edge computing systems, transportation networks, spectrum markets, and e-commerce. For example, in transportation networks, online map services provide road and traffic information to drivers. Subsequently, drivers decide on their driving directions based on this information, then their actions in turn cause the traffic conditions to change. Another example involves dynamic markets, where multiple companies compete over time in an ever changing market. All these games have the feature that agents need to consider how their actions could influence the system evolution in addition to their current payoffs when making a decision.

In many instances of dynamic games, some of the asymmetrically informed agents have aligned interest, thus, they can form a team. Team members choose their strategy jointly in order to achieve team optimality (i.e. the joint strategy profile that maximizes the total reward over all joint strategy profiles) rather than person-by-person optimality (i.e. each member's strategy is a best response to other members' strategies). The aligned interest creates an incentive for team members to share their information with each other, but despite this information asymmetry can still persist. This is since in many applications the system is evolving fast while the communication is limited or costly. One example of a dynamic game among teams with asymmetric information is ride-sharing services with autonomous vehicles, where fleets of autonomous cars of rival companies compete with each other for customers [2]. Another example is decentralized spectrum sharing, which is exemplified in the DARPA Spectrum Challenge [3], where agents work in teams to establish wireless communication with their teammates. Each team collaborates to survey the spectrum situation, avoid interference with other teams, and utilize the spectrum resource as efficiently as possible [3].

The main/key challenges in the analysis of dynamic games of asymmetric information are: (i) interdependence of players' actions and information over time, (ii) increasing information and growing domain of strategies over time, and (iii) belief formation among players with conflicting goals. In games among teams there is an additional challenge, namely, the coordination of strategies in a team to optimize the reward for the team, especially when the members of the team have asymmetric information.

In this paper, we address the following problems/issues: (i) the coordination within teams by transforming the game among teams into an equivalent game among individuals, where the individual's strategies in the transformed game indicate how agents in the same team should coordinate with each other, and (ii) the problem of increasing information by identifying a suitable compression of private information for each agent in each team.

In the control literature, there are numerous papers related to the analysis of the non-strategic dynamic teams/dynamic decentralized stochastic control problem. These papers present approaches to obtain structural results of optimal team strategies, or methods to determine team optimal strategies or person-by-person optimal strategies. These methods include (i) the person-by-person approach [4], [5]; (ii) the designer's approach [6], [7]; and (iii) the coordinator's approach [8]–[10]. We adopt the coordinators' approach to our games among teams model in this paper. In this approach, a fictitious player called the *coordinator* is introduced into each team. The coordinator is assumed to have access to the

Full paper is at [1]. This work is supported by NSF Grant No. ECCS 1750041, ECCS 2038416, ECCS 1608361, CCF 2008130, ARO Award No. W911NF-17-1-0232, MIDAS Sponsorship Funds by General Dynamics, and Okawa Foundation Research Grant.

D. Tang, V. Subramanian, and D. Teneketzis are with the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, 48109, USA. E-mail: dwtang@umich.edu, vgsubram@umich.edu, teneket@umich.edu.

H. Tavafoghi is with the Mechanical Engineering Department, University of California, Berkeley, CA, 94720, USA. E-mail: tavaf@berkeley.edu.

A. Nayyar is with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, 90089, USA. E-mail: ashutosn@usc.edu.

team members' common information. Based on this information, the coordinator assigns instructions/prescriptions to each member. These instructions/prescriptions describe how a member should map their other part of information into actions. By introducing a coordinator, one can transform a team into a single decision making agent with full recall, where the decisions this agent make at each time are instructions instead of actions.

In the economics literature, there is a vast literature on repeated games (see [11] or [12] for a list of references). In those games, agents repeatedly interact in a static physical environment/system. Games where the agents operate in a dynamically evolving physical environment and make decisions over time have been studied by economists, control theorists, and computer science theorists (see [13] for a list of references). Our work is mainly influenced by [14]–[16]. In [14], the authors introduce the concept of a Markov Perfect Equilibrium (MPE) for games with perfectly observable actions and system states. In [15], the authors propose the concept of a Common Information Based Markov Perfect Equilibrium (CIB-MPE), which is an extension of the MPE concept to dynamic games with asymmetric information. The authors establish the existence of CIB-MPE and provide a sequential procedure to determine such equilibria. The result is obtained under a crucial assumption, namely that the CIB belief is strategy independent. In [16], the authors analyze a game model where the aforementioned assumption is not true. They introduce and analyze the concept of a Common Information Based Perfect Bayesian Equilibrium (CIB-PBE). They provide a backward induction procedure to find such equilibria whenever the procedure succeeds. They conjecture that CIB-PBE always exists. Our work is different from [14] since we focus on games with asymmetric information. Moreover, different from [15], the CIB belief in our model is strategy-dependent. Our work is closest to [16]. However, as we discuss in detail in Section III, the results of [16] cannot be directly applied in this work.

In contrast to games among individual players, games among teams have not been extensively studied in the literature. There have been a few works either solving specialized models (e.g. [17]), analyzing models with restrictive assumptions (e.g. [18]), or studying games among teams empirically (e.g. [19]).

See [1] for a more extensive list of references and a detailed discussion of the literature on decentralized control, dynamic games, and dynamic games among teams.

In this paper, we investigate a family of finite-horizon stochastic dynamic games among teams with asymmetric information. Each team has its own dynamic system, whose evolution is affected by the actions of all players in all teams. Within each team, members share their local states with each other with a delay of $d > 0$. Actions are observable to all agents along with noisy observations of the systems. Our model generalizes that of [16] to games among teams.

Contributions: (i) We transform the game among teams to an equivalent game among individuals, where the strategies of the individuals in the new game indicate how agents in the same team should collaborate. (ii) We identify an appropriate way to compress the team-private information of each team. Such an information compression leads to the Sufficient Private Information Based (SPIB) strategies. (iii) We show that SPIB-strategy-based Nash Equilibria always exist. Since SPIB strategies do not feature full recall, we combine techniques in the economics and control literature to establish this result. (iv) In a special case of our model where a dynamic system is associated with each agent instead of each team, we identify an appropriate way to further compress each agent's information.

Organization: We organize the rest of the paper as follows. In Section II we formally present our model and problem. In Section III we transform the game among teams into an equivalent game among coordinators where each coordinator represents a team. In Section IV we introduce the compression of private information and SPIB strategies, and we show the existence of SPIB-strategy-based equilibria. We present a special case of our results in Section V. We conclude in Section VI.

Notation: We use capital letters to represent random variables, bold capital letters to denote random vectors, and lower case letters to represent realizations. We use superscripts to indicate teams and agents, and subscripts to indicate time. We use $t_1 : t_2$ to indicate the collection of timestamps $(t_1, t_1 + 1, \cdots, t_2)$. For example $X^1_{1:3}$ stands for the random vector (X^1_1, X^1_2, X^1_3). For random variables or random vectors, we use the corresponding script capital letters (italic capital letters for greek letters) to denote the space of values these random vectors can take. For example, $\mathcal{H}^i_t$ denotes the space of values the random vector H^i_t can take. We use $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ to denote probabilities and expectations, respectively. We use $\Delta(\mathit{\Omega})$ to denote the set of probability distributions on a finite set $\mathit{\Omega}$.

II. PROBLEM FORMULATION

A. System Model and Information Structure

We consider a finite horizon dynamic game among finitely many teams each consisting of a finite number of agents, where agents have asymmetric information. Let $\mathcal{I} = \{1, \cdots, I\}$ denote the set of teams and $\mathcal{T} = \{1, \cdots, T\}$ denote the set of time indices. We use a tuple (i, j) to indicate the j-th member of team i. For a team $i \in \mathcal{I}$, let $\mathcal{N}_i = \{(i, 1), \cdots, (i, N_i)\}$ denote team i's members. Let $\mathcal{N} = \bigcup_{i\in\mathcal{I}} \mathcal{N}_i$ denote the set of all agents. At each time $t \in \mathcal{T}$, each agent (i, j) selects an action $U^{i,j}_t \in \mathcal{U}^{i,j}_t$, where $\mathcal{U}^{i,j}_t$ denotes the action space of agent (i, j) at time t. Each team is associated with a vector-valued dynamical system $\mathbf{X}^i_t = (X^{i,j}_t)_{(i,j)\in\mathcal{N}_i}$ which evolves according to

$$\mathbf{X}^i_{t+1} = f^i_t(\mathbf{X}^i_t, \mathbf{U}_t, W^{i,X}_t), \quad i \in \mathcal{I},$$

where $\mathbf{U}_t = (U^{k,j}_t)_{(k,j)\in\mathcal{N}}$, and $(W^{i,X}_t)_{i\in\mathcal{I},t\in\mathcal{T}}$ is the noise in the dynamical system. We assume that $X^{i,j}_t \in \mathcal{X}^{i,j}_t$ for $(i, j) \in \mathcal{N}$.

We assume that the actions of all agents are publicly observed. Further, at time t, after all the agents take actions, a public observation of team i's state is generated via

$$Y_t^i = \ell_t^i(\mathbf{X}_t^i, \mathbf{U}_t, W_t^{i,Y}), \quad i \in \mathcal{I},$$

where $Y_t^i \in \mathcal{Y}_t^i$ and $(W_t^{i,Y})_{i\in\mathcal{I},t\in\mathcal{T}}$ are the observation noises.

We assume that the functions $(f_t^i)_{i\in\mathcal{I},t\in\mathcal{T}}, (\ell_t^i)_{i\in\mathcal{I},t\in\mathcal{T}}$ are common knowledge among all agents. We further assume that $(\mathbf{X}_1^i)_{i\in\mathcal{I}}, (W_t^{i,X})_{i\in\mathcal{I},t\in\mathcal{T}}$, and $(W_t^{i,Y})_{i\in\mathcal{I},t\in\mathcal{T}}$ are mutually independent primitive random variables whose distributions are also common knowledge among all agents. As a result, the teams' dynamics $(\mathbf{X}_t^i)_{t\in\mathcal{T}}, i \in \mathcal{I}$ are conditionally independent given the actions, and the public observations of different teams' systems are conditionally independent given the states and actions of all teams.

At each time t, the following information is available to all agents:

$$H_t^0 = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}),$$

where $\mathbf{Y}_t = (Y_t^i)_{i\in\mathcal{I}}, \mathbf{U}_t = (U_t^{i,j})_{(i,j)\in\mathcal{N}}$. We refer to H_t^0 as the common information among teams.

We assume that each agent (i,j) observes her own state $X_t^{i,j}$. Further, agents in the same team share their states with each other with a time delay $d \geq 1$. Thus, at time t, all agents in team i have access to H_t^i, given by

$$H_t^i = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}_{1:t-d}^i), \quad i \in \mathcal{I}.$$

We call H_t^i the common information within team i.

Finally, the information available to agent (i,j) at time t, denoted by $H_t^{i,j}$, is

$$H_t^{i,j} = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}_{1:t-d}^i, X_{t-d+1:t}^{i,j}), \quad (i,j) \in \mathcal{N}.$$

This model captures the hierarchy of information asymmetry among teams and team members. It is an abstract representation of dynamic oligopoly games [20] where each member of the oligopoly is a team.

To illustrate the key ideas of the paper without dealing with technical difficulties arising from continuum spaces, we assume that all the system random variables (i.e. all states, actions, and observations) take values in finite sets.

Assumption 1. $\mathcal{X}_t^{i,j}, \mathcal{Y}_t^i, \mathcal{U}_t^{i,j}$ are finite sets for all $(i,j) \in \mathcal{N}, t \in \mathcal{T}$.

B. Strategies and Reward Functions

For games among teams, there are three possible types of team strategies one could consider: (1) pure strategies, i.e. deterministic strategies; (2) randomized strategies where team members independently randomize; and (3) randomized strategies where team members jointly randomize.

A pure strategy profile of a team is a collection of functions $\mu^i = (\mu_t^{i,j})_{(i,j)\in\mathcal{N}_i,t\in\mathcal{T}}$, where $\mu_t^{i,j} : \mathcal{H}_t^{i,j} \mapsto \mathcal{U}_t^{i,j}$. Define $\mathcal{M}_t^{i,j}$ as the space of functions from $\mathcal{H}_t^{i,j}$ to $\mathcal{U}_t^{i,j}$. Let $\mathcal{M}^i := \prod_{t\in\mathcal{T}} \prod_{(i,j)\in\mathcal{N}_i} \mathcal{M}_t^{i,j}$. Any randomized strategy of a team, either of type 2 or type 3, can be described through a mixed strategy $\sigma^i \in \Delta(\mathcal{M}^i)$. In particular, if team members independently randomize, the mixed strategy σ^i being used to describe the strategy profile will be a product of measures on $\mathcal{M}^{i,j} := \prod_{t\in\mathcal{T}} \mathcal{M}_t^{i,j}$ for $(i,j) \in \mathcal{N}_i$.

Team i's total reward under a pure strategy profile $\mu = (\mu_t^{i,j})_{(i,j)\in\mathcal{N},t\in\mathcal{T}}$ is

$$J^i(\mu) = \mathbb{E}^\mu \left[\sum_{t\in\mathcal{T}} r_t^i(\mathbf{X}_t, \mathbf{U}_t) \right],$$

where the functions $(r_t^i)_{i\in\mathcal{I},t\in\mathcal{T}}, r_t^i : \mathcal{X}_t \times \mathcal{U}_t \mapsto \mathbb{R}$, representing the instantaneous rewards, are common knowledge among all agents. Team i's total reward under a mixed strategy profile $\sigma = (\sigma^i)_{i\in\mathcal{I}}, \sigma^i \in \Delta(\mathcal{M}^i)$, is then an average of the total rewards under pure strategy profiles, i.e. $J^i(\sigma) = \sum_{\mu\in\mathcal{M}} \left[\prod_{i\in\mathcal{I}} \sigma^i(\mu^i)\right] J^i(\mu)$.

Note that while members of the same team may jointly randomize their strategies, the randomizations of different teams are independent of each other.

Remark 1. For convenience of notation and proofs, for $t \in \{-(d-1), \cdots, -1, 0\}$, we define $\mathcal{X}_t^{i,j} = \mathcal{U}_t^{i,j} = \mathcal{Y}_t^i = \{0\}$ and $r_t^i(\mathbf{X}_t, \mathbf{U}_t) = 0$ for all $i \in \mathcal{N}$ and $(i,j) \in \mathcal{N}$.

C. Solution Concept

In this work, a team refers to a group of agents that have asymmetric information and the same objective. Because of the shared objective, members of the same team can jointly decide on the strategy to use before the start of the game for the collective benefit of the team. Hence, we can assume that every member of the team knows the strategy of the others in the team. Therefore, when considering an equilibrium concept, we should consider team deviations rather than individual deviations, i.e. multiple members of the same team may decide to play a different strategy than the equilibrium strategy. We consider randomized strategies where team members jointly randomize. Example 1 of Section II-C.1 illustrates why such strategies must be considered when we study games among teams.

The above discussion motivates the definition of a Team Nash Equilibrium.

Definition 1 (Team Nash Equilibrium)**.** A mixed strategy profile $\sigma^* = (\sigma^{*i})_{i\in\mathcal{I}}, \sigma^{*i} \in \Delta(\mathcal{M}^i)$, is said to form a Team Nash Equilibrium (TNE) if $J^i(\sigma^{*i}, \sigma^{*-i}) \geq J^i(\tilde{\sigma}^i, \sigma^{*-i})$ for any mixed strategy profile $\tilde{\sigma}^i \in \Delta(\mathcal{M}^i)$ for all $i \in \mathcal{I}$.

The primary objective of this paper is to characterize a subclass of Team NE and establish the existence of these Team NE.

1) A Motivating Example: The following example illustrates the importance of considering jointly randomized mixed strategies when we study games among teams. Similar to the role mixed strategies play in games among individual players, the space of jointly randomized mixed strategies is rich enough to ensure that an equilibrium exists in games among teams. In particular, if we restrict the teams to use pure or independently randomized strategies, i.e. type 1 and type 2 strategies described in Section II-B, then an equilibrium may not exist. This example is similar to the example in Section 2 of [21] in spirit, with the main difference that in our example the players in the same team have asymmetric information.

Example 1 (Guessing Game)**.** Consider a two-stage zero-sum game (i.e. $\mathcal{T} = \{1, 2\}$) of two teams $\mathcal{I} = \{A, B\}$, each consisting of two players. The set of all agents is given by $\mathcal{N} = \{(A, 1), (A, 2), (B, 1), (B, 2)\}$. Let $\mathbf{X}_t^A = (X_t^{A,1}, X_t^{A,2}) \in \{-1, 1\}^2$ and Team B does not have a state, i.e. $\mathbf{X}_t^B = \varnothing$. Assume $\mathcal{U}_t^{i,j} = \{-1, 1\}$ for $t = 1, i = A$ or $t = 2, i = B$ and $\mathcal{U}_t^{i,j} = \varnothing$ otherwise, i.e. Team A moves at time 1, and Team B moves at time 2. At time 1, $X_1^{A,1}$ and $X_1^{A,2}$ are independently uniformly distributed on $\{-1, 1\}$. Team A's system is assumed to be static, i.e. $\mathbf{X}_2^A = \mathbf{X}_1^A$.

The rewards of Team A are given by

$$r_1^A(\mathbf{X}_1, \mathbf{U}_1) = \mathbf{1}_{\{X_1^{A,1}U_1^{A,1}X_1^{A,2}U_1^{A,2}=-1\}},$$
$$r_2^A(\mathbf{X}_2, \mathbf{U}_2) = -\mathbf{1}_{\{X_2^{A,1}=U_2^{B,1}\}} - \mathbf{1}_{\{X_2^{A,2}=U_2^{B,2}\}},$$

and the rewards of Team B satisfies $r_t^B(\mathbf{X}_t, \mathbf{U}_t) = -r_t^A(\mathbf{X}_t, \mathbf{U}_t)$ for $t = 1, 2$.

Assume that there are no additional common observations other than past actions, i.e. $\mathbf{Y}_t = \varnothing$. We set the delay $d = 2$, i.e. agent (A, 1) does not know $X_t^{A,2}$ throughout the game and a similar property is true for agent (A, 2). In this game, the task of Team A is to choose actions according to their states at $t = 1$ in order to earn a positive reward, while not revealing too much information through their actions to Team B. The task of Team B is to guess Team A's state.

An equilibrium where Team A randomizes in a correlated manner is given in the following: At $t = 1$, Team A plays $\gamma^A = (\gamma^{A,1}, \gamma^{A,2})$ with probability 1/2, and $\tilde{\gamma}^A = (\tilde{\gamma}^{A,1}, \tilde{\gamma}^{A,2})$ with probability 1/2, where

$$\gamma^{A,1}(x_1^{A,1}) = x_1^{A,1}, \quad \gamma^{A,2}(x_1^{A,2}) = -x_1^{A,2},$$
$$\tilde{\gamma}^{A,1}(x_1^{A,1}) = -x_1^{A,1}, \quad \tilde{\gamma}^{A,2}(x_1^{A,2}) = x_1^{A,2},$$

and at $t = 2$, the two members of Team B choose independent and uniformly distributed actions on $\{-1, 1\}$, independent of their action and observation history. In this equilibrium, each agent (A, j) chooses a uniform random action irrespective of their states. However, $(A, 1)$ and $(A, 2)$ choose these actions in a correlated way to ensure that they obtain the full instantaneous reward while not revealing any information.

It can be verified that if we restrict both teams to use independently randomized strategies (including deterministic strategies), then there exists no equilibria: Since the game is a zero-sum game, it can be easily verified that if an independently randomized equilibrium existed, it would achieve the same expected payoff for each team as the jointly randomized equilibrium described above. One can obtain the Team A's value by computing Team A's expected payoff at the equilibrium given in the example. Then one can verify that no independently randomized strategies of Team A can attain this value.

III. GAME OF COORDINATORS

In this section we present a game among individual players that is equivalent to the game among teams formulated in Section II.

We view the agents of a team as being coordinated by a fictitious *coordinator* as in [9]: At each time t, team i's coordinator instructs the members of team i how to use their private information $H_t^{i,j} \backslash H_t^i$. The coordinator's instructions are based on H_t^i and her past instructions up to time $t-1$ (see [9]). Using this vantage point, we can view the games among teams as games among coordinators, where the coordinators' actions are the instructions, or *prescriptions*, provided to individual agents. Notice that unlike agents' actions, coordinators' actions (prescriptions) cannot be publicly observed. To proceed further we formally define coordinators' actions and strategies, and prove Lemma 1.

Definition 2 (Prescription)**.** Coordinator i's *prescriptions* at time t is a collection of functions $\gamma_t^i = (\gamma_t^{i,j})_{(i,j)\in\mathcal{N}_i}$ where $\gamma_t^{i,j} : \mathcal{X}_{t-d+1:t}^{i,j} \mapsto \mathcal{U}_t^{i,j}$.

Define $\mathcal{A}_t^{i,j}$ to be the space of functions that maps $\mathcal{X}_{t-d+1:t}^{i,j}$ to $\mathcal{U}_t^{i,j}$. Define $\mathcal{A}_t^i = \prod_{(i,j)\in\mathcal{N}_i} \mathcal{A}_t^{i,j}$.

Definition 3 (Pure Coordination Strategy)**.** Define the augmented team-common information of team i to be $\overline{H}_t^i = (H_t^i, \mathbf{\Gamma}_{1:t-1}^i)$, where $\mathbf{\Gamma}_{1:t-1}^i$ are past prescriptions assigned by the coordinator of team i. A pure coordination strategy of team i is a collection of mappings $\nu^i = (\nu_t^i)_{t\in\mathcal{T}}$ where $\nu_t^i : \overline{\mathcal{H}}_t^i \mapsto \mathcal{A}_t^i$.

The next lemma establishes the equivalence between pure coordination strategies and pure strategies of a team.

Lemma 1. *For every pure coordination strategy profile ν, there exists a pure strategy profile μ that yields the same payoffs for all teams and vice versa.*

Proof. Follows from [9]. □

Based on the above lemma, we can immediately conclude that a mixed strategy profile for the teams in the original game is equivalent to a mixed coordination strategy (i.e. a distribution on the space of pure coordination strategies) profile . As a result, Team Nash Equilibria, as defined in Section II-C, will be equivalent to Nash Equilibria of coordinators, where the coordinators can use mixed coordination strategies.

Therefore, we can transform the games among teams to games among individual players, where each player is a (team) coordinator whose actions are prescriptions. Following the standard approach in game theory, we now consider behavioral strategies of the individuals (i.e. coordinators) in this lifted game since, unlike mixed strategies, behavioral strategies allow for independent randomizations across time and are therefore more amenable to a sequential analysis.

Definition 4 (Behavioral Coordination Strategy)**.** A behavioral coordination strategy of team i is a collection of mappings $g^i = (g_t^i)_{t\in\mathcal{T}}$ where $g_t^i : \overline{\mathcal{H}}_t^i \mapsto \Delta(\mathcal{A}_t^i)$.

Given that the coordinators have perfect recall, that is, at any time t, each coordinator remembers all her observations up to time t, and all her "actions" (prescriptions) up to time $t-1$, we can conclude from Kuhn's theorem [22] that behavioral coordination strategies are equivalent to mixed coordination strategies in the following sense.

Lemma 2. *For any behavioral coordination strategy profile, there exists a mixed coordination strategy profile with the same expected payoffs and vice versa.*

Based on the above equivalence we can first define Coordinator's Nash Equilibria in behavioral strategies and then restate our objective from Section II-C.

Definition 5 (Coordinators' Nash Equilibrium)**.** For any behavioral coordination strategy profile g, define $J^i(g) = \mathbb{E}^g\left[\sum_{t\in\mathcal{T}} r_t^i(\mathbf{X}_t, \mathbf{U}_t)\right]$. A behavioral coordination strategy profile $g^* = (g_t^{*i})_{i\in\mathcal{I},t\in\mathcal{T}}$ where $g_t^{*i} : \overline{\mathcal{H}}_t^i \mapsto \Delta(\mathcal{A}_t^i)$ is said to form a Coordinator's Nash Equilibrium (CNE) if $J^i(g^{*i}, g^{*-i}) \geq J^i(\tilde{g}^i, g^{*-i})$ for any behavioral coordination strategy $\tilde{g}^i$ for each team $i \in \mathcal{I}$.

Given that we have lifted the game among teams to a game among coordinators, we adjust the terminology for the information structure accordingly. From now on, we will refer to the common information among all teams (i.e. H_t^0) as simply the *common information*, while the information that members of team i share but is not known to other teams (i.e. $\overline{H}_t^i \backslash H_t^0 = (X_{1:t-d}^i, \mathbf{\Gamma}_{1:t-1}^i)$) will be referred to as the *private information* of coordinator i. The information that is private to an agent (i.e. $X_{t-d+1:t}^{i,j}$) will be referred to as *hidden information*, since none of the coordinators observe this information.

Remark 2. The games among coordinators we obtain have similarities to the dynamic games considered in [16]. However, the results of [16] do not apply here because of a few key differences between the two classes of games: (i) Actions in [16] are publicly observable. As mentioned before, in our game among coordinators, the "actions" (prescriptions) of the coordinators are private information. (ii) The local state X_t^i in [16] is perfectly observable by player i without delay. In our game among coordinators, at time t, a coordinator can only observe her local state up to time $t - d$. (iii) The transitions of local states in [16] are conditionally independent given the actions, i.e. $\mathbb{P}(x_{t+1}|x_t, u_t) = \prod_i \mathbb{P}(x_{t+1}^i|x_t^i, u_t)$. In our game among coordinators, transition of local states are not independent given the prescriptions. (iv) The public observation process of local states in [16] is conditionally independent given the actions, i.e. $\mathbb{P}(y_t|x_t, u_t) = \prod_i \mathbb{P}(y_t^i|x_t^i, u_t)$. In our game among coordinators, public observations of local states are not independent given the prescriptions and local states.

IV. Compression of Private Information

In this section, we identify a subset of a coordinator's private information that is sufficient for decision-making for the game of coordinators formulated in Section III. We refer to this subset of private information as the Sufficient Private Information (SPI). We then restrict attention to Sufficient Private Information Based (SPIB) strategies, where coordinators choose prescriptions based on their sufficient private information along with the common information. As a result, the coordinators do not need full recall to play SPIB strategies. We show that there always exists a Coordinator's Nash Equilibrium where coordinators play SPIB strategies. Therefore, the restriction to SPIB strategies does not hurt the existence of equilibria.

We proceed as follows. We first present a structural result on the coordinators' beliefs that plays an important role in the subsequent analysis. We then separately treat the $d = 1$ and $d > 1$ cases, where d is the delay in information sharing within the same team. The SPI for $d = 1$ case turns out to be structurally simpler than that for $d > 1$ case. Finally, for both $d = 1$ and $d > 1$ cases, we show that CNEs where coordinators play SPIB strategies always exist.

A. A Preliminary Result

We show that the states and prescriptions of different coordinators are conditionally independent given the common information.

Lemma 3 (Conditional Independence)**.** *Under any behavioral coordination strategy profile g and for each time $t \in \mathcal{T}$, $(\mathbf{X}_{1:t}^k, \mathbf{\Gamma}_{1:t}^k)_{k\in\mathcal{I}}$ are conditionally independent given the common information H_t^0. Furthermore, the conditional distribution of $(\mathbf{X}_{1:t}^k, \mathbf{\Gamma}_{1:t}^k)$ given H_t^0 depends on g only through g^k.*

Proof. Can be shown through induction on time t, where the induction step is established via Bayes rule. See [1, Appendix C] for all the details. □

As a result of Lemma 3, coordinator i's estimation of other coordinators' states and prescriptions is independent of her own strategy and private information. In other words, while coordinator i has access to both the common information and her private information, her belief on the other coordinators' private information (i.e. history of states and prescriptions) is solely based on the common information.

B. Result for $d = 1$

While coordinator i's private information consists of $(\mathbf{X}_{1:t-1}^i, \mathbf{\Gamma}_{1:t-1}^i)$, she does not have to use all of it to form a best response.

Lemma 4. *Under $d = 1$, for any behavioral coordination strategy profile g^{-i} of all coordinators other than i, there exists a best response behavioral coordination strategy g^i for coordinator i that chooses randomized prescriptions based solely on $(H_t^0, \mathbf{X}_{t-1}^i)$.*

Proof. Deferred to the proof of Lemma 6. □

Lemma 4 shows that the coordinators can ignore much of their private information without compromising their objective.

C. Result for $d > 1$

We now identify a compressed version of private information for $d > 1$ that is sufficient for decision-making.

Recall that coordinator i's information at time t consists of $\overline{H}_t^i = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}_{1:t-d}^i, \mathbf{\Gamma}_{1:t-1}^i)$. To choose her prescriptions at time t, coordinator i needs to estimate her hidden information (i.e. $\mathbf{X}_{t-d+1:t}^i$). When $d = 1$, the belief on hidden information is simply constructed using $(\mathbf{X}_{t-1}^i, \mathbf{U}_{t-1})$ and the knowledge of the transition probabilities of the underlying system. However, when $d > 1$, more information in addition to $(\mathbf{X}_{t-d}^i, \mathbf{U}_{t-d:t-1})$ is needed to form the belief.

To illustrate this, we start with the case $d = 2$. When $d = 2$, the belief of coordinator i on her hidden information would depend on the last prescription $\mathbf{\Gamma}_{t-1}^i$ in addition to $(\mathbf{X}_{t-2}^i, \mathbf{U}_{t-2:t-1})$. This is due to the signaling effect of the action $\mathbf{U}_{t-1}^i$: since coordinator i knows $\mathbf{U}_{t-1}^i$, she can infer something about $\mathbf{X}_{t-1}^i$ through the prescription used to produce these actions (recall that $U_{t-1}^{i,j} = \Gamma_{t-1}^{i,j}(X_{t-2:t-1}^{i,j})$ for $(i,j) \in \mathcal{N}_i$). Hence at time t, coordinator i needs to take $\mathbf{\Gamma}_{t-1}^i$ into account when forming her belief on the hidden information.

Furthermore, for $d = 2$, when making a decision at time t, coordinator i can use a compressed version of the prescription $\mathbf{\Gamma}_{t-1}^i$ instead of $\mathbf{\Gamma}_{t-1}^i$ itself. This is because at time t, coordinator i has learned $\mathbf{X}_{t-2}^i$ that she didn't know at time $t-1$. The coordinator can then focus on the following essential question: given the knowledge of $\mathbf{X}_{t-2}^i$, what is the relationship between $\mathbf{X}_{t-1}^i$ and $\mathbf{U}_{t-1}^i$?

Similarly, for a general $d > 1$, to estimate the hidden information, each coordinator needs to utilize her past $(d-1)$ prescriptions. Again, a coordinator can use a compressed version of the past $(d-1)$ prescriptions, since she can incorporate the additional information she knows at time t that she did not know back when the prescriptions were chosen. Each coordinator can now focus on the relationship between the unknown states and the known actions, given what is already known. This motivates the definition of $(d-1)$-step *partially realized prescriptions* (PRPs).

Definition 6. The $(d-1)$-step *partially realized prescriptions*[1] (PRPs) for coordinator i at time t is a collection of functions $\mathbf{\Phi}_t^i := (\Phi_{t-l,l}^{i,j})_{(i,j)\in\mathcal{N}_i, 1\le l\le d-1}$, where

$$\Phi_{t-l,l}^{i,j} = \Gamma_{t-l}^{i,j}(X_{t-l-d+1:t-d}^{i,j}, \cdot)$$

is a function from $\mathcal{X}_{t-d+1:t-l}^{i,j}$ to $\mathcal{U}_{t-l}^{i,j}$.

PRPs have smaller dimension than prescriptions. To illustrate this point, consider the case where $d = 2$: A prescription $\gamma_{t-1}^{i,j}$ can be represented as a table, where the rows represent $x_{t-2}^{i,j} \in \mathcal{X}_{t-2}^{i,j}$, the columns represent $x_{t-1}^{i,j} \in \mathcal{X}_{t-1}^{i,j}$, and the entries represent the corresponding action $u_{t-1}^{i,j} = \gamma_{t-1}^{i,j}(x_{t-2:t-1}^{i,j})$ to take. On the other hand, the 1-step partially realized prescription $\phi_t^{i,j} = \gamma_{t-1}^{i,j}(x_{t-2}^{i,j}, \cdot)$ can be represented by one row of the table of $\gamma_{t-1}^{i,j}$ chosen based on the realization of $X_{t-2}^{i,j}$.

In addition to $(\mathbf{X}_{t-d}^i, \mathbf{U}_{t-d:t-1}, \mathbf{\Phi}_t^i)$, coordinator i also needs to use $Y_{t-d+1:t-1}^i$ to form a belief on her hidden information since $Y_{t-d+1:t-1}^i$ can provide additional insight on $\mathbf{X}_{t-d+1:t-1}^i$ that $(\mathbf{X}_{t-d}^i, \mathbf{U}_{t-d:t-1}, \mathbf{\Phi}_t^i)$ cannot necessarily provide. The belief coordinator i has on her hidden information is summarized in the following lemma.

Lemma 5. *Suppose that the behavioral coordination strategy profile $g = (g^i)_{i\in\mathcal{I}}$ is being played. Then the conditional distribution of $\mathbf{X}_{t-d+1:t}^i$ given $\overline{H}_t^i$ under g can be expressed as a fixed function of $(Y_{t-d+1:t-1}^i, \mathbf{U}_{t-d:t-1}, \mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$, i.e.*

$$\begin{aligned}&\mathbb{P}^g(x_{t-d+1:t}^i|\overline{h}_t^i)\\ &= P_t^i(x_{t-d+1:t}^i|y_{t-d+1:t-1}^i, u_{t-d:t-1}, x_{t-d}^i, \phi_t^i)\ \forall \overline{h}_t^i \in \overline{\mathcal{H}}_t^i\end{aligned}$$

for some function P_t^i that does not depend on g.

Proof. Using Lemma 3, we can compute the belief of coordinator i by replacing g^{-i} with surrogate strategies and adding conditionally independent random variables into the condition, i.e.

$$\mathbb{P}^{g^i,g^{-i}}(x_{t-d+1:t}^i|\overline{h}_t^i) = \mathbb{P}^{g^i,\hat{g}^{-i}}(x_{t-d+1:t}^i|\overline{h}_t^i, x_{t-d:t}^{-i}),$$

where $\hat{g}^{-i}$ is an open-loop strategy profile that always generates the actions $u_{1:t-1}^{-i}$, and $x_{t-d:t}^{-i} \in \mathcal{X}_{t-d:t}^{-i}$ is such that $\mathbb{P}^{g^i,\hat{g}^{-i}}(x_{t-d:t}^{-i}|\overline{h}_t^i) > 0$. The proof then follows by computing $\mathbb{P}^{g^i,\hat{g}^{-i}}(x_{t-d+1:t}^i|\overline{h}_t^i, x_{t-d:t}^{-i})$ with Bayes rule and the chain rule. See [1, Appendix D] for a detailed proof. ☐

Remark 3. The above result can be interpreted in the following way: $\mathbf{X}_{t-d}^i$ is perfectly observed, hence coordinator i can discard $\mathbf{X}_{1:t-d-1}^i$ which are irrelevant information due to the Markov property. Since $\mathbf{X}_{t-d+1:t-1}^i$ are not perfectly observed by coordinator i, every public observation and action based upon $\mathbf{X}_{t-d+1:t-1}^i$ are important to coordinator i since it can help in estimating the state $\mathbf{X}_{t-d+1:t-1}^i$. Note that $\mathbf{\Phi}_t^i$ encodes the essential information coordinator i needs to remember at time t about her previous signaling strategy: how does $\mathbf{X}_{t-d+1:t-1}^i$ (unknown) map to $\mathbf{U}_{t-d+1:t-1}^i$ (known)? With this piece of information, coordinator i can fully interpret the signals sent through $\mathbf{U}_{t-d+1:t-1}^i$.

We claim that while coordinator i's private information consists of $(\mathbf{X}_{1:t-d}^i, \mathbf{\Gamma}_{1:t-1}^i)$, she only needs to use $(\mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$ along with the common information to choose prescriptions.

Lemma 6. *Given an arbitrary $d > 1$, for any behavioral coordination strategy profile g^{-i} of all coordinators other than i, there exists a best response behavioral coordination strategy g^i for coordinator i that chooses randomized prescriptions based solely on $(H_t^0, \mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$.*

Proof. Using Lemma 5, one can show that with a fixed g^{-i}, coordinator i faces a Markov Decision Process (MDP) with $(H_t^0, \mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$ as the state. The lemma then follows from standard results on MDPs. ☐

Remark 4. Lemmas 5 and 6 and their proofs also apply to $d = 1$, in which case the $(d-1)$-step PRP $\mathbf{\Phi}_t^i$ is empty by definition.

[1]The $(d-1)$-step PRPs are the same as the partial functions defined in the second structural result in [8].

From now on, we unify the results for $d=1$ and $d>1$. We formally define the Sufficient Private Information (SPI) and SPIB strategies which will be used in the rest of the paper.

Definition 7 (Sufficient Private Information)**.** For a given $d>0$, the *Sufficient Private Information* (SPI) for coordinator i at time t is defined as $S_t^i = (\mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$.

Definition 8 (Sufficient Private Information Based Strategy)**.** A *Sufficient Private Information Based* (SPIB) strategy for coordinator i is a collection of functions $\rho^i = (\rho_t^i)_{t\in\mathcal{T}}, \rho_t^i : \mathcal{H}_t^0 \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$.

It can be easily verified that S_t^i can be sequentially updated, i.e., there exists a fixed, strategy-independent function ι_t^i such that $S_{t+1}^i = \iota_t^i(S_t^i, \mathbf{X}_{t-d+1}^i, \mathbf{\Gamma}_t^i)$. Therefore, a coordinator does not need full recall to play an SPIB strategy.

D. Coordinators' Nash Equilibrium in SPIB Strategies and its Existence

Since the coordinators have perfect recall, we know from standard results for dynamic games that a CNE, as defined in Definition 5, exists (see Chapter 11 of [23], for example). However, in those CNEs, coordinators do not necessarily play SPIB strategies, hence the standard arguments that guarantee the existence of CNE cannot be used to establish the existence of CNE in SPIB strategies. Moreover, SPIB strategies do not feature full recall, hence one cannot directly apply standard arguments to establish the existence of CNE in SPIB strategies.

An SPIB strategy profile $\rho = (\rho_t^i)_{i\in\mathcal{I}, t\in\mathcal{T}}, \rho_t^i : \mathcal{H}_t^0 \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$ is called a *Sufficient Private Information Based Coordinators' Nash Equilibrium* (SPIB-CNE) if ρ, seen as a profile of behavioral coordination strategies, forms a Coordinator's Nash Equilibrium.

Theorem 1. *There exists at least one SPIB-CNE for the dynamic game among coordinators.*

Proof Sketch. The complete proof is presented in [1, Appendix F]. We sketch the key steps of the proof here.

Step 1: Restrict attention to ϵ-trembling SPIB strategies of coordinators, where all prescriptions have probability greater or equal to ϵ.

Step 2: Recall the proof of Lemma 6: Fixing g^{-i}, coordinator i faces a Markov Decision Process (MDP) with state $(H_t^0, \mathbf{X}_{t-d}^i, \mathbf{\Phi}_t^i)$. Define a best response correspondence using the dynamic program that solves the above MDP.

Step 3: Use Kakutani's fixed point theorem to argue the existence of equilibria in the restricted game.

Step 4: Take the limit as ϵ goes to 0. □

E. Implementing SPIB Strategies in a Team

We have established the existence of CNE in the game of coordinators when all coordinators play SPIB strategies. In this subsection, we return our attention to teams and discuss how a team can implement SPIB strategies that are defined in the context of coordinators. Given an SPIB strategy ρ^i for coordinator i, members of team i can implement the strategy in the following way: They can utilize a correlation device which generates a random seed R_t^i at each time t. In the absence of a correlation device that can be used in real-time, a team can also collectively generate all the random seeds $R_1^i, \cdots, R_T^i$ before the game begins. The members of team i agree on a common deterministic procedure (called `ChooseRandom`) to choose random prescriptions based on the random seed R_t^i and a given distribution. During the game, each member of team i acts as the coordinator on their own, and chooses actions by applying prescriptions on their d most recent private states. The implementation of the above procedure is described in Procedure 1.

Procedure 1: Agent (i,j)'s implementation of an SPIB strategy

input : An SPIB strategy $\rho^i = (\rho_t^i)_{t\in\mathcal{T}}$ at begining; $(\mathbf{Y}_{t-1}, \mathbf{U}_{t-1}, X_t^{i,j}, \mathbf{X}_{t-d}^{i,-j}, R_t^i)$ at each time t.
output: Actions $U_t^{i,j}$ at each time t.

```
// Initialize according to Remark 1:
CI ← H_0^0;               // Common Information
SPI ← S_0^i;              // Sufficient Private Information
HI ← X_{-(d-1):0}^{i,j};  // Hidden Information
Pres ← Γ_0^i;             // Prescriptions
// During the game:
for t ← 1 to T do
    // At time t:
    CI ← Concatenate(CI, Y_{t-1}, U_{t-1});
    SPI ← ι_t^i(SPI, HI[1], X_{t-d}^{i,-j}, Pres);
    HI ← Concatenate(HI[2:end], X_t^{i,j});
    DistPres ← ρ_t^i (CI, SPI);
    Pres ← ChooseRandom(DistPres, R_t^i);
    U_t^{i,j} ← Pres[j](HI);
end
```

V. SPECIAL CASE: SEPARATED DYNAMICS

Consider a special case of the model in Section II where the state of each member of each team evolves independently given the actions, i.e.

$$X_{t+1}^{i,j} = f_t^{i,j}(X_t^{i,j}, \mathbf{U}_t, W_t^{i,j}), \tag{1}$$

where $(W_t^{i,j})_{t\in\mathcal{T},(i,j)\in\mathcal{N}}$ are mutually independent primitive random variables. In this case, we can consider equilibria where the coordinators assign prescriptions that map $X_t^{i,j}$ to $U_t^{i,j}$ (instead of mapping $X_{t-d+1:t}^{i,j}$ to $U_t^{i,j}$); this is because, given H_t^i, the belief of member (i,j) about her teammates' states is independent of $X_{t-d+1:t}^{i,j}$. In other words, one can replace the hidden information $\mathbf{X}_{t-d+1:t}^i$ with the *sufficient hidden information* $\mathbf{X}_t^i$.[2] We elaborate on this point below.

[2]The compression of hidden information to sufficient hidden information is similar to the shedding of irrelevant information in [24].

Definition 9 (Simple Prescriptions)**.** A *simple prescription* for coordinator i at time t is a collections of functions $\theta_t^i = (\theta_t^{i,j})_{(i,j)\in\mathcal{N}_i}, \theta_t^{i,j} : \mathcal{X}_t^{i,j} \mapsto \mathcal{U}_t^{i,j}$.

Lemma 7. *Suppose that g^{-i} is a behavioral coordination strategy profile for coordinators other than coordinator i, then there exists a best response behavioral coordination strategy g^i for coordinator i that chooses randomized simple prescriptions based on $\overline{H}_t^i$.*

Proof Sketch. A more detailed proof is presented in [1, Appendix M]. We sketch the key steps here.

Step 1: Fixing the strategy profile for all agents of all teams other than agent (i,j), show that agent (i,j) faces an POMDP with state $Z_t = (H_t^0, \mathbf{X}_{1:t}^{-i}, \Gamma_{1:t-1}^{-i}, \mathbf{X}_{1:t}^{i,-j}, X_t^{i,j})$ and observation $H_t^{i,j}$.

Step 2: Show that the conditional distribution of Z_t given $H_t^{i,j}$ does not depend on $X_{t-d+1:t-1}^{i,j}$. Conclude that agent (i,j) can optimize team i's payoff by choosing a strategy where her action at time t does not depend on $X_{t-d+1:t-1}^{i,j}$.

Step 3: Conclude that team i can form a best response to g^{-i} with strategies that always assign simple prescriptions. □

Given the above result, one can restrict the set of feasible prescriptions for each coordinator to contain only the simple prescriptions. With this modification, results analogous to that of Sections IV can be derived. In particular, we have the following analogue of Theorem 1.

Theorem 2. *Consider the game formulated in Section II but with dynamics given by* (1)*. Then there exist at least one SPIB-CNE for the dynamic game among coordinators where each coordinator uses only simple prescriptions.*

VI. Conclusion and Future Work

We studied a model of dynamic games among teams with asymmetric information, where agents in each team share their observations with a delay of d. Each team is associated with a controlled Markov Chain, whose dynamics are influenced by the actions of all agents. We proposed Sufficient Private Information Based (SPIB) strategies, where agents can use a compression of their information rather than full information to make decisions. Such strategies do not feature full recall. Nevertheless, we showed that SPIB-strategy-based equilibria are guaranteed to exist. We also analyzed a special case of our model where the state of each member of each team evolves independently given the actions and showed that the agents can compress their information further in this case. Our results provide an important first step in addressing all the difficulties associated with dynamic games of teams.

Moving forward, there are a few research problems arising from this work: (i) identifying a non-empty subset of SPIB strategy-based equilibria where agents compress the common information among all agents as well; (ii) investigating the payoff properties of SPIB-strategy-based equilibria in comparison to general Nash Equilibria; (iii) developing efficient algorithms to find SPIB-strategy-based equilibria.

References

[1] D. Tang, H. Tavafoghi, V. Subramanian, A. Nayyar, and D. Teneketzis, "Dynamic games among teams with delayed intra-team information sharing," *arXiv preprint arXiv:2102.11920*, 2021.

[2] P. A. Hancock, I. Nourbakhsh, and J. Stewart, "On the future of transportation in an era of automated and autonomous vehicles," *Proceedings of the National Academy of Sciences*, vol. 116, no. 16, pp. 7684–7691, 2019.

[3] T. Harbert. (2014) Radio wrestlers fight it out at the DARPA Spectrum Challenge. [Online]. Available: https://spectrum.ieee.org/telecom/wireless/radio-wrestlers-fight-it-out-at-the-darpa-spectrum-challenge

[4] H. Witsenhausen, "On the structure of real-time source coders," *Bell System Technical Journal*, vol. 58, no. 6, pp. 1437–1451, 1979.

[5] J. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 814–820, 1983.

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

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

[8] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1606–1620, 2011.

[9] ——, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

[10] H. Tavafoghi, Y. Ouyang, and D. Teneketzis, "A unified approach to dynamic decision problems with asymmetric information: Non-strategic agents," *IEEE Transactions on Automatic Control*, pp. 1–1, 2021.

[11] R. B. Myerson, *Game theory*. Harvard university press, 2013.

[12] G. J. Mailath and L. Samuelson, *Repeated games and reputations: long-run relationships*. Oxford university press, 2006.

[13] H. Tavafoghi, Y. Ouyang, D. Teneketzis, and M. Wellman, "Game theoretic approaches to cyber security: Challenges, results, and open problems," in *Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Control-and Game-theoretic Approaches to Cyber Security*, S. Jajodia, G. Cybenko, P. Liu, C. Wang, and M. Wellman, Eds. Springer Nature, 2019, vol. 11830, pp. 29–53.

[14] E. Maskin and J. Tirole, "Markov perfect equilibrium: I. observable actions," *Journal of Economic Theory*, vol. 100, no. 2, pp. 191–219, 2001.

[15] A. Nayyar, A. Gupta, C. Langbort, and T. Başar, "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, 2013.

[16] Y. Ouyang, H. Tavafoghi, and D. Teneketzis, "Dynamic games with asymmetric information: Common information based perfect Bayesian equilibria and sequential decomposition," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 222–237, 2016.

[17] S. Bhattacharya and T. Başar, "Multi-layer hierarchical approach to double sided jamming games among teams of mobile agents," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 5774–5779.

[18] A. Nayyar and T. Başar, "Dynamic stochastic games with asymmetric information," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 7145–7150.

[19] C. A. Cox and B. Stoddard, "Strategic thinking in public goods games with teams," *Journal of Public Economics*, vol. 161, pp. 31–43, 2018.

[20] Y. Ouyang, H. Tavafoghi, and D. Teneketzis, "Dynamic oligopoly games with private Markovian dynamics," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 5851–5858.

[21] V. Anantharam and V. Borkar, "Common randomness and distributed control: A counterexample," *Systems & control letters*, vol. 56, no. 7-8, pp. 568–572, 2007.

[22] H. Kuhn, "Extensive games and the problem of information, in: H.W. Kuhn and A.W. Tucker (eds.)," *Contributions to the Theory of Games*, vol. 2, pp. 193–216, 1953.

[23] M. J. Osborne and A. Rubinstein, *A course in game theory*. The MIT Press, 1994.

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