Dynamic Games and Applications
https://doi.org/10.1007/s13235-022-00424-4

Dynamic Games Among Teams with Delayed Intra-Team Information Sharing

Dengwang Tang[1] · Hamidreza Tavafoghi[2] · Vijay Subramanian[1] · Ashutosh Nayyar[3] · Demosthenis Teneketzis[1]

Accepted: 3 January 2022

Abstract
We analyze a class of stochastic dynamic games among teams with asymmetric information, where members of a team share their observations internally with a delay of d. Each team is associated with a controlled Markov Chain, whose dynamics are coupled through the players' actions. These games exhibit challenges in both theory and practice due to the presence of signaling and the increasing domain of information over time. We develop a general approach to characterize a subset of Nash equilibria where the agents can use a compressed version of their information, instead of the full information, to choose their actions. We identify two subclasses of strategies: sufficient private information-Based (SPIB) strategies, which only compress private information, and compressed information-based (CIB) strategies, which compress both common and private information. We show that SPIB-strategy-based equilibria exist and the set of payoff profiles of such equilibria is the same as that of all Nash equilibria. On the other hand, we show that CIB-strategy-based equilibria may not exist. We develop a backward inductive sequential procedure, whose solution (if it exists) provides a CIB strategy-based equilibrium. We identify some instances where we can guarantee the existence of a

This article is part of the topical collection "Multi-agent Dynamic Decision Making and Learning" edited by Konstantin Avrachenkov, Vivek S. Borkar and U. Jayakrishnan Nair.

✉ Dengwang Tang
dwtang@umich.edu

Hamidreza Tavafoghi
tavaf@berkeley.edu

Vijay Subramanian
vgsubram@umich.edu

Ashutosh Nayyar
ashutosn@usc.edu

Demosthenis Teneketzis
teneket@umich.edu

1 Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109, USA

2 Mechanical Engineering, University of California, Berkeley, CA 94720, USA

3 Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, USA

Published online: 14 February 2022

solution to the above procedure. Our results highlight the tension among compression of information, ability of compression-based strategies to sustain all or some of the equilibrium payoff profiles, and backward inductive sequential computation of equilibria in stochastic dynamic games with asymmetric information.

Keywords Game theory · Dynamic games · Decentralized control · Stochastic dynamic systems · Compression-based equilibria · Sequential decomposition

1 Introduction

Dynamic games with asymmetric information appear in many socioeconomic contexts. In these games, multiple agents/decision makers interact repeatedly in a changing environment. Agents have different information and seek to optimize their respective long-term payoffs. For example, multiple companies may compete with each other in a market over time and each company attempts to optimize its own long-term benefits [5,7,12,32,33]; the market is also changing over time driven by the actions the companies take. Another instance of such games arises in cyberphysical systems [1,2,48,52,70]; at each time, attackers make decisions on which hosts to attack, and the system administrators/defenders choose actions to defend against the attackers, for example, by isolating some hosts from the rest of the system [52]; the system's state changes over time as a result of the attackers' and defenders' actions. In all instances of these games, when an agent takes an action, she needs to consider not only how the action will affect her current payoff but also how it will influence the system's evolution and the future actions of all agents, and hence her future payoffs.

In some settings, agents can form groups, or teams [9,49]. The agents in the same group share a common goal but may have different information available to them. This information asymmetry among teammates appears in many engineering applications. In most of these applications, the state of the system changes rapidly, and agents have to make real-time decisions. Moreover, the communication between agents is either costly, or restricted by bandwidth or delay. Examples of these settings include competing fleets of automated cars from rival companies [18] and the DARPA Spectrum Challenge [19]. In the DARPA Spectrum challenge setup, individual transceivers work in teams to maximize the sum throughput of their networks. Teams compete with other teams, and members of the same team need to coordinate and evolve their responses over time. In these settings, agents in the same team aim to choose their strategy jointly to achieve team optimality (i.e., to choose the joint strategy profile that maximizes the expected utility of the team over all joint strategy profiles) rather than just person-by-person optimality (PBPO).[1] We study a stylized model of such settings in this paper.

It is worth stating that the games among teams problems we focus on in this paper are different from cooperative games in economics research (e.g., see [36] Chapters 8–10). In cooperative game theory, the goal is to study the group formation process among agents with different objectives. In our setting, groups are assumed to be fixed and given, and we focus instead on determining the optimal actions and payoffs for each group. A unilateral deviation in our problems means one or more agents in one group deviates, but the community structure of the agents stays the same.

[1] A team strategy is person-by-person optimal (PBPO) when each team member's strategy is an optimal response given other team members' strategy profile.

The key challenges in the study of dynamic games among single agents with asymmetric information are: (i) Due to signaling[2] in many instances, an agent's assessment of the status of the game at time t, hence her strategy at time t, depends on the strategies of agents who acted before her.[3] Therefore, we cannot obtain the standard sequential decomposition (that sequentially determines the components of an equilibrium strategy profile) of the kind provided by the dynamic programming algorithm for centralized stochastic control (where the agent's optimal strategy at any time t does not depend on past strategies) [25]. (ii) The domain of the agents' strategies increases with time, as the agents acquire information over time. Thus, the computational complexity of the agents' strategies increases with time.

To address these challenges, one can look for compression-based strategies that can be sequentially computed. This creates an additional challenge: compression-based strategies could restrict the agents' ability to sustain all or some of the equilibrium payoffs of the game.

In games among teams, we have the additional challenge of coordination within asymmetrically informed team members so as to achieve team optimality instead of person-by-person optimality.

In this paper, we propose a general approach to characterize a subset of equilibrium strategy profiles of dynamic games among teams with the following goals: (i) to determine appropriate compression of information for each agent to base their decision on; (ii) to develop a sequential decomposition of the game. In addition, we would like to determine sufficient conditions for the existence of such equilibrium strategies.

1.1 Related Literature

To understand games among teams, we first examine a team's best-response strategy when other teams' strategies are fixed. Team problems, or decentralized control problems, have been extensively studied in the control literature. Researchers have developed various methodologies/approaches to decentralized control problems to determine team optimal strategies or PBPO strategies, and to determine structural results/properties for the above-mentioned strategies. These methodologies include: (i) the person-by-person approach [23,30,38,39,54–59,61–64,66] (ii) the designer's approach [28,65] (iii) the coordinator's approach [29,41,43,53]. The person-by-person approach has been used to determine qualitative/structural properties of team optimal or PBPO strategies. In this approach, the strategies of all team members/agents except one, say agent i, are assumed to be arbitrary but fixed; then, the qualitative properties of agent i's best response strategy are determined. These properties are then valid for all possible (fixed) strategies of the other agents. The designer's approach investigates the decentralized control/team problem from the point of view of a designer who knows the system model and the joint probability distributions of the primitive random variables (the system's initial state, the noise driving the system, and the noise in the agents' observations). The designer chooses the strategies of all team members at time 0 by solving an open-loop stochastic control problem, where her decision at each time is the strategy/control law for all the team members/agents. Applying stochastic control results, the designer can obtain a dynamic programming decomposition. The methodology devel-

[2] In contrast to signaling in teams, signaling in games is complicated by the fact that agents have diverging incentives.

[3] Example of such strategy dependencies appears in Ho [20] and in Nayyar and Teneketzis [40] for team problems with non-classical information structure. Since these strategy dependencies are solely due to the problem's information structure, they also appear in dynamic games with non-classical information structure (see [21]).

oped in this paper is inspired by the coordinator's approach used in Nayyar et al. [41,43], Tavafoghi et al. [53]. Similar to the designer's approach, the coordinator's approach assumes that a fictitious agent, called the coordinator, assigns instructions to agents. However, unlike the designer's approach, the coordinator is assumed to know the common information of all agents and assigns partial strategies (prescriptions) instead of full strategies to agents. The partial strategies tell an agent how to utilize her private information to generate actions. Both the designer's approach and the coordinator's approach lead to the determination of globally optimal team strategy profiles.

Research on dynamic games roughly consists of two directions: One direction focuses on repeated games or multi-stage games, where the instantaneous payoffs at each stage is only affected by actions in this stage but not by the actions in the previous stages. In these games, researchers investigated long-term interactions among agents (e.g., punishment and reward strategies) and characterized the set of equilibrium payoffs (e.g., see [31] or [36] Chapter 7). The other direction focuses on games with an underlying dynamic system, in other words, games where instantaneous payoffs can be affected by previous actions. In this more complicated setting, researchers attempted to develop methodologies for the determination of equilibria with either a general structure or a specialized structure. In this paper, we focus on the latter direction.

Games of individual agents (i.e., agents do not form teams) with an underlying dynamic system have been studied in both the economics and the control literature. Dynamic games with symmetric information have been studied extensively [4,14]. In Maskin and Tirole [34], the authors propose the concept of Markov Perfect Equilibrium (MPE) for the case where the state of the system and agents' actions are perfectly observable. The research on dynamic games with asymmetric information can be classified into two categories: zero-sum games and general (i.e., not necessarily zero-sum) games. Zero-sum games are analyzed in Renault [46,47], Zheng and Castañón [69], Gensbittel and Renault [15], Li and Shamma [26], Cardaliaguet et al. [8], Li et al. [27], Kartik and Nayyar [21]. In these works, the authors take advantage of many properties of zero-sum games, such as having a unique value and the interchangeability of equilibrium strategies. These properties do not extend to general nonzero-sum games. The literature on general dynamic games includes [16,17,35,37,42,44, 45,50,51,60]. In Nayyar et al. [42], the authors extend the MPE concept in Maskin and Tirole [34] to the case where the underlying dynamics is only partially observable. Under the crucial assumption that the common information-based (CIB) belief is strategy-independent, the authors prove that there exist equilibria where agents play *CIB strategies*, i.e., the agents choose their actions based on CIB belief and private information instead of full information. Furthermore, such equilibria can be found through a sequential decomposition of the game. In our setup, the system state is not perfectly observed; thus, our model is distinctly different from that of Maskin and Tirole [34]. Furthermore, in contrast to Nayyar et al. [42], the CIB belief in our model is strategy-dependent.

The closest work to our paper in terms of both model and approach is [45]. In Ouyang et al. [45], the authors consider a game model where, in contrast to Nayyar et al. [42], the CIB beliefs are strategy-dependent. They propose the concept of Common Information-Based Perfect Bayesian Equilibrium (CIB-PBE) as a solution concept for this game model and prove that CIB-PBE can be found through a sequential decomposition whenever this decomposition has a solution. The game model of Ouyang et al. [45] has multiple features that prevent us from directly applying their results in our analysis in Sect. 5. We will make a more detailed comparison in Sect. 3. Our work is also close in spirit to Maskin and Tirole [35]. In Maskin and Tirole [35], the authors extend their work in Maskin and Tirole [34] by considering games where actions are observable, but each agent has a fixed, private utility

type. They propose Markov Sequential Equilibrium (MSE) as a solution concept for these games, where the agents choose their actions based on a compression of their information along with their beliefs on the types of other agents. The authors show by example that MSE do not necessarily exist. As an alternative to MSE, they propose a new concept obtained from limits of ε-MSE as ε goes to 0.

Unlike either team problems or dynamic games among individual agents, games among teams (in particular, ones with an underlying dynamic system) have not been systematically studied in the literature. There are only a few works on special models of games among teams. In Farina et al. [13] and Zhang and An [68], the authors proposed algorithms to compute equilibria for zero-sum multiplayer extensive form games, where a team of players plays against an adversary. In Anantharam and Borkar [3], the authors provide an example of a zero-sum game which involves a team. However, the players in this team have symmetrical information; hence, the team is equivalent to an individual player with vector-valued actions. In Nayyar and Başar [37], the authors briefly extend their results in Nayyar et al. [42] to games among teams for a specialized model where the CIB belief is strategy independent. In both Colombino et al. [9] and Summers et al. [49], the authors solve a two-team zero-sum linear quadratic stochastic dynamic game. In Bhattacharya and Başar [6], the authors formulate and solve a game between two teams of mobile agents. The model and information structure of Bhattacharya and Başar [6] are different from ours. Additionally, games among teams have been the subject of empirical research (see, for example, [10,11]). In our work, we study analytically a model of nonzero-sum dynamic stochastic games among teams where the CIB belief is strategy dependent.

1.2 Contribution

In this paper, we consider a model of dynamic games among teams with asymmetric information. We assume that each team is associated with a dynamical system that has Markovian dynamics driven by the actions of all agents of all teams. The state of each dynamical system is assumed to be vector-valued, where each component represents an agent's local state. Agents can observe their own local states perfectly and communicate them within their respective teams with a delay of d. All actions are public, i.e., observable by every agent in every team. We also assume the presence of public noisy observations of the system's state. The instantaneous reward of a team depends on the states and actions of all teams. Our model is a generalization of the model in Ouyang et al. [45] to competing teams.

Our contributions are as follows:

- We identify appropriate compression of information for each agent. The compression is achieved in two steps: (i) the compression of team-private information that depends only on the team strategy; (ii) the compression of common information that depends on the strategy of all agents. The compression steps induce two special classes of strategies: (i) Sufficient Private Information-Based (SPIB) strategies, where agents only apply the first step of compression; and (ii) Compressed Information-Based (CIB) strategies, where agents apply both steps of compression.
- We show that SPIB-strategy-based Nash equilibria always exist, and the set of equilibrium payoff profiles of such equilibria is the same as that of all Nash equilibria.
- We develop a sequential decomposition of the game where agents play CIB strategies. We show that any solution of the sequential decomposition forms a Nash equilibrium of the game.

– We show that CIB-strategy-based Nash equilibria do not always exist. We identify some simple instances where CIB-strategy-based equilibria are guaranteed to exist.

1.3 Organization

We organize the rest of the paper as follows: In Sect. 2, we formally present our model and problem. In Sect. 3, we transform the game among teams into an equivalent game among coordinators where each coordinator represents a team. In Sect. 4, we introduce our first step of compression of information and SPIB strategies. We show the existence of SPIB-strategy-based equilibria and the equivalence of sets of payoff profiles between SPIB-strategy-based equilibria and Nash equilibria. In Sect. 5, we introduce the second step of compression and CIB strategies, and we provide a sequential decomposition of the game. We also show the general non-existence of CIB-strategy-based equilibria and provide some conditions for existence. We present some extensions and special cases of our results in Sect. 6. Then, we discuss our results in Sect. 7. We conclude in Sect. 8. Proof details are provided in the Appendix.

1.4 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 i to represent a typical team, and $-i$ represents all teams other than i. We use $t_1 : t_2$ to indicate the collection of timestamps $(t_1, t_1 + 1, \dots, t_2)$. For example, $X^1_{5:8}$ stands for the random vector $(X^1_5, X^1_6, X^1_7, X^1_8)$. For random variables or random vectors represented by Latin letters, we use the corresponding script capital 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. The products of sets in this paper are Cartesian products. We use $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ to denote probabilities and expectations, respectively. We use $\Delta(\Omega)$ to denote the set of probability distributions on a finite set Ω. When writing probabilities, we will omit the random variables when the lower case letters that represent the realizations clearly indicates the random variable it represents. For example, we will use $\mathbb{P}(y^i_t|x_t, u_t)$ as a shorthand for $\mathbb{P}(Y^i_t = y^i_t|\mathbf{X}_t = x_t, \mathbf{U}_t = u_t)$. When λ is a function from Ω_1 to $\Delta(\Omega_2)$, with some abuse of notation we write $\lambda(\omega_2|\omega_1) := (\lambda(\omega_1))(\omega_2)$ as if λ is a conditional distribution. We use $\mathbf{1}_A$ to denote the indicator random variable of an event A.

In general, probability distributions of random variables in a dynamic system are only well defined after a complete strategy profile is specified. We specify the strategy profile that defines the distribution in superscripts, e.g., $\mathbb{P}^g(x^i_t|h^0_t)$. When the conditional probability is independent of a certain part of the strategy $(g^i_t)_{(i,t)\in\Omega}$, we may omit this part of the strategy in the notation, e.g., $\mathbb{P}^{g_{1:t-1}}(x_t|y_{1:t-1}, u_{1:t-1})$, $\mathbb{P}^{g^i}(x^i_t|h^0_t)$ or $\mathbb{P}(x_{t+1}|x_t, u_t)$. We say that a realization of some random vector (for example h^0_t) is *admissible* under a partially specified strategy profile (for example g^{-i}) if the realization has strictly positive probability under some completion of the partially specified strategy profile (In this example, that means $\mathbb{P}^{g^i, g^{-i}}(h^0_t) > 0$ for some g^i). Whenever we write a conditional probability or conditional expectation, we implicitly assume that the condition has nonzero probability under the specified strategy profile. When only part of the strategy profile is specified in the superscript, we implicitly assume that the condition is admissible under the specified partial strategy profile.

2 Problem Formulation

2.1 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 $\mathcal{I} = \{1, \dots, I\}$ denote the set of teams and $\mathcal{T} = \{1, \dots, T\}$ 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 $\mathcal{N}_i = \{(i, 1), \dots, (i, N_i)\}$ 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 $U_t^{i,j} \in \mathcal{U}_t^{i,j}$, where $\mathcal{U}_t^{i,j}$ denotes the action space of agent (i, j) at time t. Each team is associated with a vector-valued dynamical system $\mathbf{X}_t^i = (X_t^{i,j})_{(i,j) \in \mathcal{N}_i}$ which evolves according to

$$\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t, W_t^{i,X}), \quad i \in \mathcal{I},$$

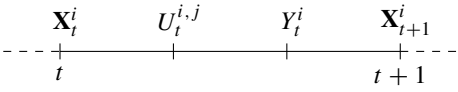
where $\mathbf{U}_t = (U_t^{k,j})_{(k,j) \in \mathcal{N}}$, and $(W_t^{i,X})_{i \in \mathcal{I}, t \in \mathcal{T}}$ is the noise in the dynamical system. We assume that $X_t^{i,j} \in \mathcal{X}_t^{i,j}$ for $(i, j) \in \mathcal{N}$ and $t \in \mathcal{T}$.

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 according to

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

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.

The order of events occurring between time steps t and $t+1$ is shown in the figure below:



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 [44,45] where each member of the oligopoly is a team.

Remark 1 Our model also captures the scenarios where a team has only one member. Such a team can be incorporated in our framework by adding a dummy agent to it and assuming a suitable internal communication delay d. If all teams are single-member teams, then d can be arbitrarily chosen.

To illustrate the key ideas of the paper without dealing with the 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 $(i, j) \in \mathcal{N}, t \in \mathcal{T}$.

2.2 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; and (2) randomized strategies where team members independently randomize; (3) randomized strategies where team members jointly randomize.

A pure strategy profile of a team is defined to be 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 by 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 2 For convenience of notation and proofs, for $t \in \{-(d-1), \ldots, -1, 0\}$, 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 $i \in \mathcal{I}$ and $(i, j) \in \mathcal{N}$.

2.3 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. 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 change their strategies. We consider randomized strategies where team members jointly randomize. To implement an arbitrary mixed strategy, a team can jointly choose a random strategy profile out of the distribution specified by the mixed strategy at the beginning of the game. Example 2 of Appendix A.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 for all $i \in \mathcal{I}$,

$$J^i(\sigma^{*i}, \sigma^{*-i}) \geq J^i(\tilde{\sigma}^i, \sigma^{*-i})$$

for any mixed strategy profile $\tilde{\sigma}^i \in \Delta(\mathcal{M}^i)$.

Since stochastic dynamic games among teams with asymmetric information is a relatively new class of dynamic games, we start with the simplest equilibrium concept, which is the Team Nash Equilibrium.

The primary objective of this paper is to characterize compression-based subclasses of Team Nash Equilibria.

3 Game Among Coordinators

In this section, we present a game among individual players that is equivalent to the game among teams formulated in Sect. 2.

We view the members of a team as being coordinated by a fictitious *coordinator* as in Nayyar et al. [43]: At each time t, team i's coordinator instructs the members of team i how to use their private information, $H_t^{i,j} \setminus H_t^i$, based on H_t^i and her past instructions up to time $t-1$ (see [43]). Using this vantage point, we can view the game among teams as a game among coordinators, where the coordinators' actions are the instructions, or *prescriptions*, provided to individual agents. Notice that unlike agents' actions, coordinators' actions (prescriptions) are not 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 map $\mathcal{X}_{t-d+1:t}^{i,j}$ 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 $\overline{H}_t^i = (H_t^i, \boldsymbol{\Gamma}_{1:t-1}^i)$, where $\boldsymbol{\Gamma}_{1:t-1}^i$ are the 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 $\nu_t^i : \overline{\mathcal{H}}_t^i \mapsto \mathcal{A}_t^i$.

Definition 4 We call two strategies $g^i, \tilde{g}^i$ of team i payoff-equivalent if the two strategies generate the same total expected reward for all agents under all pure team strategy profiles μ^{-i} of teams other than i, that is, $J^k(g^i, \mu^{-i}) = J^k(\tilde{g}^i, \mu^{-i})$ for all $k \in \mathcal{I}$ and all $\mu^{-i} \in \mathcal{M}^{-i}$.[4]

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

Lemma 1 *For every pure strategy μ^i of team i, there exists a payoff-equivalent pure coordination strategy ν^i and vice versa.*

Proof See Appendix B. □

Based on the above lemma, we can immediately conclude that a mixed strategy for a team is payoff-equivalent to a mixed coordination strategy (i.e., a distribution on the space of pure coordination strategies). As a result, Team Nash equilibria, as defined in Sect. 2.3, will be equivalent to Nash equilibria of coordinators, where the coordinators can use mixed coordination strategies. An example illustrating how a mixed team strategy can be transformed into a payoff-equivalent behavioral coordination strategy is presented in Appendix A.2.

Therefore, we can transform the game among teams to a game 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., the coordinators) in this lifted game since, unlike mixed strategies, behavioral strategies allow for independent randomizations across time and therefore, better facilitate a sequential decomposition of the dynamic game.

Definition 5 *(Behavioral coordination strategy)* A behavioral coordination strategy of team i is a collection of mappings $g^i = (g^i_t)_{t\in\mathcal{T}}$ where $g^i_t : \overline{\mathcal{H}}^i_t \mapsto \Delta(\mathcal{A}^i_t)$.

Given that the coordinators have perfect recall, that is, at any time t, the 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 [24] that behavioral coordination strategies are payoff-equivalent to mixed coordination strategies.

Lemma 2 *For any mixed coordination strategy ς^i of coordinator i, there exists a payoff-equivalent behavioral coordination strategy g^i and vice versa.*

Based on this equivalence, we can first define Nash equilibria for the coordinator's game and then restate our objective from Sect. 2.3.

Definition 6 *(Coordinators' Nash equilibrium)* For any behavioral coordination strategy profile g, define

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

A behavioral coordination strategy profile $g^* = (g^{*i}_t)_{i\in\mathcal{I},t\in\mathcal{T}}$ where $g^{*i}_t : \overline{\mathcal{H}}^i_t \mapsto \Delta(\mathcal{A}^i_t)$ is said to form a Coordinator's Nash Equilibrium (CNE) if for any $i \in \mathcal{I}$,

$$J^i(g^{*i}, g^{*-i}) \geq J^i(\tilde{g}^i, g^{*-i})$$

for any behavioral coordination strategy profile $\tilde{g}^i : \overline{\mathcal{H}}^i_t \mapsto \Delta(\mathcal{A}^i_t)$.

[4] We do not restrict the strategy types of g^i and $\tilde{g}^i$ in Definition 4. In particular, each of g^i and $\tilde{g}^i$ could be a coordination strategy or a team strategy.

In other words, a coordination strategy profile g forms a CNE if the behavioral strategies of coordinators form a Bayes–Nash equilibrium in the game of coordinators.

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., $\overline{H}_t^i \setminus H_t^0 = (\mathbf{X}_{1:t-d}^i, \boldsymbol{\Gamma}_{1:t-1}^i)$) 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 3 The game among coordinators we obtain has a few differences from the game model in Ouyang et al. [45]:

- Actions in Ouyang et al. [45] are publicly observable. As mentioned before, in our game among coordinators, the "actions" (prescriptions) of the coordinators are private information.
- The local state X_t^i in Ouyang et al. [45] 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 $t - d$.
- The transitions of local states in Ouyang et al. [45] 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.
- The public observations of local states in Ouyang et al. [45] are conditionally independent given the local states and 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.

Due to the above differences, we cannot directly apply the results of Ouyang et al. [45] to the game of coordinators.

4 Compression of Private Information

In this section, we identify a compression of a coordinator's private information that is sufficient for decision-making for the game of coordinators formulated in Sect. 3. We refer to this compression as the Sufficient Private Information (SPI). We 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 for any behavioral coordination strategy, there exists a payoff-equivalent (See Definition 4) SPIB strategy. Consequently, there always exists a Coordinator's Nash equilibrium where coordinators play SPIB strategies, and the set of equilibrium payoffs of such equilibria is the same as the set of equilibrium payoffs for CNE. Therefore, the restriction to SPIB strategies does not hurt the coordinators' ability to achieve any payoff profile that is achievable in a CNE.

We proceed as follows. We first present a preliminary result that plays an important role in the subsequent analysis. We then introduce our results. We then formally define Sufficient Private Information and Sufficient Private Information-Based (SPIB) strategies. Finally, we establish the payoff-equivalence between SPIB strategies and general behavioral coordination strategies.

4.1 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 $t \in \mathcal{T}$, $(\mathbf{X}^i_{1:t}, \boldsymbol{\Gamma}^i_{1:t})_{i\in\mathcal{I}}$ are conditionally independent across coordinators given the common information H^0_t, i.e.,*

$$\mathbb{P}^g(x_{1:t}, \gamma_{1:t}|h^0_t) = \prod_{i\in\mathcal{I}} \mathbb{P}^g(x^i_{1:t}, \gamma^i_{1:t}|h^0_t) \qquad \forall h^0_t \in \mathcal{H}^0_t.$$

Furthermore, $\mathbb{P}^g(x^i_{1:t}, \gamma^i_{1:t}|h^0_t)$ depends on g only through g^i.

Proof See Appendix C. □

As a result of Lemma 3, coordinator i's estimation of other coordinators' state 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 (history of states and prescription) is solely based on the common information.

4.2 Sufficient Private Information and SPIB Strategy

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

Recall that coordinator i's information at time t consists of $(\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}^i_{1:t-d}, \boldsymbol{\Gamma}^i_{1:t-1})$. To choose her prescriptions at time t, coordinator i needs to estimate her hidden information (i.e., $\mathbf{X}^i_{t-d+1:t}$). When $d = 1$, the belief on hidden information is simply constructed using $(\mathbf{X}^i_{t-1}, \mathbf{U}_{t-1})$ and the knowledge of the transition probabilities of the underlying system. However, when $d > 1$, more information in addition to $(\mathbf{X}^i_{t-d}, \mathbf{U}_{t-d:t-1})$ 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 $\boldsymbol{\Gamma}^i_{t-1}$ in addition to $(\mathbf{X}^i_{t-2}, \mathbf{U}_{t-2:t-1})$. This is due to the signaling effect of the action $\mathbf{U}^i_{t-1}$: since coordinator i knows $\mathbf{U}^i_{t-1}$, she can infer something about $\mathbf{X}^i_{t-1}$ through the prescription used to produce these actions (recall that $U^{i,j}_{t-1} = \Gamma^{i,j}_{t-1}(X^{i,j}_{t-2:t-1})$ for $(i, j) \in \mathcal{N}_i$). Hence at time t, coordinator i needs to take $\boldsymbol{\Gamma}^i_{t-1}$ 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 $\boldsymbol{\Gamma}^i_{t-1}$ instead of $\boldsymbol{\Gamma}^i_{t-1}$ itself. This is because at time t, coordinator i has learned $\mathbf{X}^i_{t-2}$ that she did not know at time $t-1$. The coordinator can then focus on the following essential question: given the knowledge of $\mathbf{X}^i_{t-2}$, what is the relationship between $\mathbf{X}^i_{t-1}$ and $\mathbf{U}^i_{t-1}$?

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 7 The $(d-1)$-step *partially realized prescriptions*[5] (PRPs) for coordinator i at time t is a collection of functions $\boldsymbol{\Phi}_t^i := (\Phi_{t-l,l}^{i,j})_{(i,j)\in\mathcal{N}_i, 1\le l\le d-1}$, 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 $\mathcal{U}_{t-l}^{i,j}$.

Remark 4 When $d = 1$, the $(d-1)$-step PRP $\boldsymbol{\Phi}_t^i$ is empty by definition.

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 $X_{t-2}^{i,j}$.

When $d > 1$, in addition to $(\mathbf{X}_{t-d}^i, \mathbf{U}_{t-d:t-1}, \boldsymbol{\Phi}_t^i)$, 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 $(\mathbf{X}_{t-d}^i, \mathbf{U}_{t-d:t-1}, \boldsymbol{\Phi}_t^i)$ cannot necessarily provide. The belief coordinator i has on her hidden information is summarized in the following lemma.

Lemma 4 *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 $(Y_{t-d+1:t-1}^i, \mathbf{U}_{t-d:t-1}, \mathbf{X}_{t-d}^i, \boldsymbol{\Phi}_t^i)$, i.e.,*

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

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

Proof See Appendix D. □

Remark 5 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 $\boldsymbol{\Phi}_t^i$ 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 now formally define the Sufficient Private Information (SPI) and SPIB strategies which will be used in the rest of the paper.

[5] The $(d-1)$-step PRPs are the same as the partial functions defined in the second structural result in Nayyar et al. [41].

Definition 8 *(Sufficient private information)* For a given $d > 0$, the *Sufficient Private Information* (SPI) for coordinator i at time t is defined as $S_t^i = (\mathbf{X}_{t-d}^i, \boldsymbol{\Phi}_t^i)$.[6]

Definition 9 *(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

$$S_{t+1}^i = \iota_t^i(S_t^i, \mathbf{X}_{t-d+1}^i, \boldsymbol{\Gamma}_t^i). \tag{2}$$

Therefore, a coordinator does not need full recall to play an SPIB strategy.

4.3 Payoff-equivalence of SPIB Strategies with General Behavioral Coordination Strategies

To establish the payoff-equivalence between SPIB strategies and general behavioral coordination strategies, we introduce the following definition.

Definition 10 Consider a behavioral coordination strategy $g^i = (g_t^i)_{t\in\mathcal{T}}$ of coordinator i. An SPIB strategy $\rho^i = (\rho_t^i)_{t\in\mathcal{T}}$ is said to be *associated* with g^i if for all $t \in \mathcal{T}$,

$$\rho_t^i(h_t^0, s_t^i) = \sum_{\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i} g_t^i(h_t^0, \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i)\mathbb{P}^{g^i}(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | h_t^0, s_t^i)$$

for all (h_t^0, s_t^i) admissible under g^i.

Recall that due to Lemma 3, $\mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | h_t^0, s_t^i)$ depends on a behavioral coordination strategy profile g only through g^i. Hence, the above definition is independent of other coordinators' strategies.

Remark 6 The distribution $\rho_t^i(h_t^0, s_t^i)$ can be seen as the conditional distribution of $\boldsymbol{\Gamma}_t^i$ given $H_t^0 = h_t^0$, $S_t^i = s_t^i$ under the behavioral coordination strategy g^i. Similar construction is also used in Lemma 4 of Kartik et al. [22].

Lemma 5 *Let ρ^i be an SPIB strategy associated with coordinator i's strategy g^i; then ρ^i and g^i are payoff-equivalent.*

Proof See Appendix E. □

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 (see definition 6). The following theorem follows naturally from Lemma 5.

Theorem 1 *SPIB-CNE exist for the dynamic game among coordinators. Furthermore, the set of equilibrium payoff profiles of SPIB-CNEs is the same as the set of equilibrium payoff profiles for CNEs.*

[6] The compression of private information of coordinators in our model is closely related to Tavafoghi et al.'s [53] sufficient information approach. One can show that our sufficient private information $S_t^i = (\mathbf{X}_{t-d}^i, \boldsymbol{\Phi}_t^i)$ satisfies the definition of *sufficient private information* (Definition 4) in Tavafoghi et al. [53] (hence, we choose to use the same terminology).

Proof Let $g = (g^i)_{i\in\mathcal{I}}$ be a CNE. Let ρ^i be an SPIB strategy associated with g^i for each $i \in \mathcal{I}$, then $\rho = (\rho^i)_{i\in\mathcal{I}}$ is an SPIB-CNE: Using Lemma 5 iteratively, we can show that

$$J^i(\rho^i, \rho^{-i}) = J^i(g^i, g^{-i}) \geq J^i(\tilde{g}^i, g^{-i}) = J^i(\tilde{g}^i, \rho^{-i})$$

for any behavioral coordination strategy $\tilde{g}^i$ of coordinator i.

We also know that ρ has the same equilibrium payoffs as g due to Lemma 5. Therefore, the set of equilibrium payoff profiles of SPIB-CNEs is the same as the set of equilibrium payoff profiles for CNEs. □

5 Compression of Common Information and Sequential Decomposition

The SPIB strategies defined in the previous section use sufficient private information instead of the entire private information for each coordinator. If the sets $\mathcal{X}_t, \mathcal{Y}_t, \mathcal{U}_t$ are time-invariant, the set of possible values of sufficient private information used in SPIB strategies is also time-invariant. However, the common information still increases with time and this means that the domain of SPIB strategies keeps increasing with time. In order to limit the growing domain of SPIB strategies, we introduce a subclass of SPIB strategies, named Compressed Information-Based (CIB) strategies, where the coordinators use a compressed version of common information instead of the entire common information. We show that this new class of strategies satisfies a key best-response/closedness property. Based on this property, we provide a backward inductive procedure that identifies an equilibrium in this subclass of strategies if each step of this procedure has a solution. While equilibria in CIB strategies may not exist in general (see example in Sect. 5.5), we identify classes of games among teams where such equilibria do exist.

5.1 Compressed Common Information and CIB Strategy

In decentralized control problems [43,53] and games among individuals [45,51], agents can compress their common information into beliefs on hidden and (sufficient) private information for the purpose of decision-making. Similarly, we would like to consider a subclass of SPIB strategies where each coordinator compresses the common information H_t^0 to a belief on sufficient private information and hidden information, i.e., $\mathbb{P}(\mathbf{X}_{t-d:t}^k = \cdot, \boldsymbol{\Phi}_t^k = \cdot | H_t^0)$ for $k \in \mathcal{I}$. Due to Lemma 4, these beliefs can be constructed from $\mathbb{P}(\mathbf{X}_{t-d}^k = \cdot, \boldsymbol{\Phi}_t^k = \cdot | H_t^0)$ and $(Y_{t-d+1:t-1}^k, \mathbf{U}_{t-d:t-1})$. Therefore, we will consider strategies where coordinators use common information-based beliefs on the sufficient private information $S_t^k = (\mathbf{X}_{t-d}^k, \boldsymbol{\Phi}_t^k)_{k\in\mathcal{I}}$ along with the uncompressed values of $(\mathbf{Y}_{t-d+1:t-1}, \mathbf{U}_{t-d:t-1})$, instead of the whole H_t^0.

We formalize the above discussion in the rest of this subsection.

Definition 11 *(Belief generation system)* A *Belief Generation System* for coordinator i consists of a sequence of functions $\psi^i = (\psi_t^{i,k})_{k\in\mathcal{I},t\in\mathcal{T}}$ where $\psi_t^{i,k} : \left(\prod_{l\in\mathcal{I}} \Delta(\mathcal{S}_t^l)\right) \times \mathcal{Y}_{t-d+1:t} \times \mathcal{U}_{t-d:t} \mapsto \Delta(\mathcal{S}_{t+1}^k)$

Coordinator i can use this system to generate common information-based beliefs $\Pi_t^{i,k} \in \Delta(\mathcal{S}_t^k)$ for all $k \in \mathcal{I}$ as follows:

- $\Pi_1^{i,k}$ is the prior distribution of $(\mathbf{X}^k_{-(d-1)}, \boldsymbol{\Phi}^k_1)$, i.e., a measure which assigns probability 1 to the event $(\mathbf{X}^k_{-(d-1)} = 0, \boldsymbol{\Phi}^k_1 = \hat{\phi}^k_1)$, where $\hat{\phi}^k_1$ is the PRP that always produces actions $u^{k,j}_t = 0$ for all $(k, j) \in \mathcal{N}_k, t \leq 0$ (see Remark 2);
- $\Pi^{i,k}_{t+1} = \psi^{i,k}_t((\Pi^{i,l}_t)_{l\in\mathcal{I}}, \mathbf{Y}_{t-d+1:t}, \mathbf{U}_{t-d:t}), t \geq 1$.

$\Pi^{i,k}_t$ represents coordinator i's subjective belief on coordinator k's sufficient private information S^k_t. These beliefs along with $(\mathbf{Y}_{t-d+1:t-1}, \mathbf{U}_{t-d:t-1})$ will serve as coordinator i's compressed common information.

Definition 12 *(Compressed common information)* We define coordinator i's *Compressed Common Information* (CCI) at time t as

$$B^i_t = \left(\left(\Pi^{i,l}_t\right)_{l\in\mathcal{I}}, \mathbf{Y}_{t-d+1:t-1}, \mathbf{U}_{t-d:t-1}\right),$$

where $(\Pi^{i,l}_t)_{l\in\mathcal{I}}$ are generated using the belief generation system defined in Definition 11. Note that when $d = 1$, we have $B^i_t = ((\Pi^{i,l}_t)_{l\in\mathcal{I}}, \mathbf{U}_{t-1})$.

We can write the belief update using B^i_t as $\Pi^{i,k}_{t+1} = \psi^{i,k}_t(B^i_t, \mathbf{Y}_t, \mathbf{U}_t)$. With a slight abuse of notation, we use ψ^i_t to represent the collection $(\psi^{i,k}_t)_{k\in\mathcal{I}}$ and write the belief updates collectively as $(\Pi^{i,l}_{t+1})_{l\in\mathcal{I}} = \psi^i_t(B^i_t, \mathbf{Y}_t, \mathbf{U}_t)$.

We now define a subclass of strategies where coordinator i uses her CCI instead of the entire common information.

Definition 13 *(Compressed information-based strategy)* Let $\mathcal{B}_t = \left(\prod_{k\in\mathcal{I}} \Delta(\mathcal{S}^k_t)\right) \times \mathcal{Y}_{t-d+1:t-1} \times \mathcal{U}_{t-d:t-1}$. A *Compressed Information-Based* (CIB) strategy for coordinator i is a pair (λ^i, ψ^i), where $\lambda^i = (\lambda^i_t)_{t\in\mathcal{T}}$ is a collection of functions $\lambda^i_t : \mathcal{B}_t \times \mathcal{S}^i_t \mapsto \Delta(\mathcal{A}^i_t)$, and $\psi^i = (\psi^{i,k}_t)_{k\in\mathcal{I},t\in\mathcal{T}}$, $\psi^{i,k}_t : \mathcal{B}_t \times \mathcal{Y}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{S}^k_{t+1})$ is a belief generation system as defined in Definition 11.

Under a CIB strategy, coordinator i uses her belief generation system to compress common information into beliefs and then uses these beliefs along with $(\mathbf{Y}_{t-d+1:t-1}, \mathbf{U}_{t-d:t-1}, S^i_t)$ to select a randomized prescription. Thus, a CIB strategy (λ^i, ψ^i) is equivalent to an SPIB-strategy

$$\rho^i_t(h^0_t, s^i_t) = \lambda^i_t\left(\left(\pi^{i,k}_t\right)_{k\in\mathcal{I}}, y_{t-d+1:t-1}, u_{t-d:t-1}, s^i_t\right) \quad \forall h^0_t \in \mathcal{H}^0_t, \forall s^i_t \in \mathcal{S}^i_t$$

where $(\pi^{i,k}_t)_{k\in\mathcal{I}}$ is generated from h^0_t through the belief generation system defined in Definition 11.

Remark 7 One advantage of CIB strategies is that at each time coordinator i only needs to use her current CCI rather than the full common information (i.e., H^0_t) which increases with time. Thus, if the sets $\mathcal{X}_t, \mathcal{Y}_t, \mathcal{U}_t$ are time-invariant, the mappings λ^i_t, ψ^i_t in a CIB strategy have a time-invariant domain.

Remark 8 We have not imposed any restriction on the mapping ψ^i_t in coordinator i's belief generation system (see Definition 11). Intuitively, however, one can imagine that coordinator i has some prediction about others' strategies and is rationally using her prediction about others' strategies to update her beliefs through the mapping ψ^i_t. In the following discussion, our focus will be on such "rational" ψ^i_t where the notion of rationality will be captured by Bayes' rule.

We end this subsection by pointing out that coordinator i's belief generated from ψ^i can be grouped into two parts: $(\Pi_t^{i,-i})_{t\in\mathcal{T}}$ and $(\Pi_t^{i,i})_{t\in\mathcal{T}}$. The first part represents what coordinator i believes about other coordinators' SPI. The second part represents what coordinator i thinks is the other coordinators' belief on her own SPI.

5.2 Consistency and Closedness of CIB Strategies

As mentioned before, our interest in CIB strategies is motivated by the common information belief-based strategies that appeared in the solution of decentralized control problems [43, 53] or games among individuals [42,45]. The common beliefs used in these prior works are compatible with Bayes' rule (i.e., the beliefs can be obtained using Bayes' rule along with the knowledge of the system model and the strategies being used). Inspired by these observations, we are particularly interested in CIB strategies where the belief generation system is compatible with Bayes' rule, i.e., the beliefs generated by coordinator i using ψ^i agree with those generated using Bayes' rule along with the knowledge of the system model and the strategies being used.

In the following discussion, we identify a key property of such Bayes' rule compatible CIB strategies. To do so, we use the following technical definition.

Definition 14 *(Consistency)* Given $\lambda_t^i : \mathcal{B}_t \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$, a belief generation function $\psi_t^{*,i} : \mathcal{B}_t \times \mathcal{Y}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{S}_{t+1}^i)$ is said to be *consistent* with λ_t^i if the following holds: For all $b_t = ((\pi_t^l)_{l\in\mathcal{I}}, y_{t-d+1:t-1}, u_{t-d:t-1}) \in \mathcal{B}_t$, $\psi_t^{*,i}(b_t, y_t, u_t)$ is equal to the conditional distribution of S_{t+1}^i given the event $(\mathbf{Y}_t = y_t, \mathbf{U}_t = u_t)$ found using Bayes rule (whenever Bayes rule applies), assuming that $y_{t-d+1:t-1}$ and $u_{t-d:t-1}$ are the realization of recent observations and actions, S_t^i has prior distribution π_t^i, and given $S_t^i = s_t^i$, $\boldsymbol{\Gamma}_t^i$ has distribution $\lambda_t^i(b_t, s_t^i)$. That is,

$$[\psi_t^{*,i}(b_t, y_t, u_t)](s_{t+1}^i) = \frac{\Upsilon_t^i(b_t, y_t^i, u_t, s_{t+1}^i)}{\sum_{\tilde{s}_{t+1}^i} \Upsilon_t^i(b_t, y_t^i, u_t, \tilde{s}_{t+1}^i)} \tag{3}$$

whenever the denominator of (3) is nonzero, where

$$\begin{aligned}
&\Upsilon_t^i(b_t, y_t^i, u_t, s_{t+1}^i) \\
:= &\sum_{\tilde{s}_t^i} \sum_{\tilde{x}_{t-d+1:t}^i} \sum_{\tilde{\gamma}_t^i : \tilde{\gamma}_t^i(\tilde{x}_{t-d+1:t}^i) = u_t^i} \Big[\mathbb{P}(y_t^i | \tilde{x}_t^i, u_t) \mathbf{1}_{\{s_{t+1}^i = \iota_t^i(\tilde{s}_t^i, \tilde{x}_{t-d+1}^i, \tilde{\gamma}_t^i)\}} \\
&\times \lambda_t^i(\tilde{\gamma}_t^i | b_t, \tilde{s}_t^i) P_t^i(\tilde{x}_{t-d+1:t}^i | y_{t-d+1:t-1}^i, u_{t-d:t-1}, \tilde{s}_t^i) \pi_t^i(\tilde{s}_t^i) \Big]
\end{aligned}$$

for all

$$b_t = ((\pi_t^l)_{l\in\mathcal{I}}, y_{t-d+1:t-1}, u_{t-d:t-1}) \in \mathcal{B}_t, y_t^i \in \mathcal{Y}_t^i, u_t \in \mathcal{U}_t, s_{t+1}^i \in \mathcal{S}_{t+1}^i,$$

ι_t^i is defined in (2) and P_t^i is as described in Lemma 4.

For any index set $\Omega \subset \mathcal{I} \times \mathcal{T}$, we say that $\psi^{*,i} = (\psi_t^{*,i})_{(i,t)\in\Omega}$ is consistent with $\lambda^i = (\lambda_t^i)_{(i,t)\in\Omega}$ if $\psi_t^{*,i}$ is consistent with λ_t^i for all $(i,t) \in \Omega$.

A CIB strategy (λ^i, ψ^i) for coordinator i is said to be self-consistent if $\psi^{i,i}$ is consistent with λ^i. Since self-consistency can be viewed as Bayes' rule compatibility, the beliefs $(\Pi_t^{i,i})_{t\in\mathcal{T}}$ represents true conditional distributions of coordinator i's SPI given the common information under a self-consistent strategy.

Lemma 6 *Let (λ^i, ψ^i) be a self-consistent CIB strategy of coordinator i. Denote the behavioral strategy generated from (λ^i, ψ^i) as g^i. Let $h_t^0 \in \mathcal{H}_t^0$ be admissible under $g_{1:t-1}^i$, then*

$$\mathbb{P}^{g_{1:t-1}^i}(s_t^i, x_{t-d+1:t}^i | h_t^0) = \pi_t^{i,i}(s_t^i) P_t^i(x_{t-d+1:t}^i | y_{t-d+1:t-1}^i, u_{t-d:t-1}, s_t^i)$$
$$\forall s_t^i \in \mathcal{S}_t^i \ \forall x_{t-d+1:t}^i \in \mathcal{X}_{t-d+1:t}^i$$

where $\pi_t^{i,i}$ is the belief obtained using ψ^i under the realization h_t^0 of common information and P_t^i is as described in Lemma 4.

Proof See Appendix F. □

Now, consider a game with two coordinators: Suppose that coordinator 1 plays a self-consistent CIB strategy with belief generation system ψ^1. Since the belief $\Pi_t^{1,1}$ generated from ψ^1 is a true conditional distribution on coordinator 1's SPI, coordinator 2 can use $\Pi_t^{1,1}$ as her belief on coordinator 1's SPI. Further, coordinator 2 can use ψ^1 to compute coordinator 1's belief about coordinator 2's SPI. This suggests that coordinator 2 should mimic coordinator 1's belief generation system when coordinator 1's strategy is self-consistent. This observation, along with results from Markov decision theory, leads to the following crucial best-response property of CIB strategies.

Lemma 7 (Closedness of CIB strategies) *Suppose that all coordinators other than coordinator i are using self-consistent CIB strategies. Let (λ^k, ψ^k) be the CIB strategy of coordinator $k \in \mathcal{I}\backslash\{i\}$. Suppose that $\psi^j = \psi^k$ for all $j, k \in \mathcal{I}\backslash\{i\}$. Then, a best-response strategy for coordinator i is a CIB strategy with the same belief generation system as the other coordinators.*

Proof See Appendix G. □

5.3 Interpretation and Discussion of Consistency and Closedness Property

Lemma 7 imposes two conditions on the CIB strategies of coordinators other than i, namely (I) they are self-consistent, and (II) their belief generation systems are identical. In order to illustrate the significance of both conditions, we first describe how coordinator i could form her best response when all coordinators other than i are playing some generic CIB strategies that are not necessarily self-consistent or do not have an identical belief generation system.

The problem of finding coordinator i's best response to others' CIB strategies can be thought of as a stochastic control problem with partial observation. This suggests that in order to form a best response at time t, coordinator i needs to compute (or form beliefs on) the data that coordinators $-i$'s CIB strategies use, i.e., the CCI and the SPI of other coordinators. Coordinator i also needs to estimate all the hidden information in order to evaluate the payoffs. Coordinator i's estimation task can be divided into three sub-tasks: (i) to form a belief on her own hidden information $\mathbf{X}_{t-d+1:t}^i$, (ii) to recover coordinators $-i$'s CCI $(B_t^k)_{k\in\mathcal{I}\backslash\{i\}}$, and (iii) to form a belief on coordinators $-i$'s SPI and hidden information $\mathbf{X}_{t-d+1:t}^{-i}$.

For the first sub-task, coordinator i can compute the belief through the function P_t^i defined in Lemma 4 using $(Y_{t-d+1:t-1}^i, \mathbf{U}_{t-d:t-1}, S_t^i)$, without using any belief generation system. For the second sub-task, recall that B_t^k includes $(Y_{t-d+1:t-1}^i, \mathbf{U}_{t-d:t-1})$, which coordinator i already knows. Thus, to complete the second task, coordinator i can simply use $(\psi^k)_{k\in\mathcal{I}\backslash\{i\}}$ and the common information H_t^0 to compute all the beliefs in $(B_t^k)_{k\in\mathcal{I}\backslash\{i\}}$. Condition (I),

namely that the CIB strategies for coordinators other than i are self-consistent, ensures that coordinator i can also accomplish the third sub-task using the beliefs in $(B_t^k)_{k\in\mathcal{I}\setminus\{i\}}$ due to Lemma 6. By using self-consistent CIB strategies, coordinators $-i$ effectively "invite" coordinator i to use the same belief generation system as $-i$.

Thus, all of coordinator i's sub-tasks can be done if she keeps track of her own S_t^i and the CCI $(B_t^k)_{k\in\mathcal{I}\setminus\{i\}}$ used by others. Therefore, coordinator i can form a best response with a strategy that chooses prescriptions based on $(B_t^k)_{k\in\mathcal{I}\setminus\{i\}}$ and S_t^i at time t. Condition (II), namely that the belief generation systems are identical, ensures that B_t^k's are identical for all $k \in \mathcal{I}\setminus\{i\}$, and hence, the best response described above becomes a CIB strategy with the same belief generation system as the one used by all coordinators other than i.

Remark 9 Note the CIB strategy that is a best-response strategy for coordinator i in Lemma 7 may not necessarily be self-consistent. However, the equilibrium strategies in a CIB-CNE (which we will introduce later) will be self-consistent for all players.

5.4 Coordinators' Nash Equilibrium in CIB Strategies and Sequential Decomposition

The fact that one of coordinator i's best responses to others using CIB strategies (with identical and self-consistent belief generation systems) is itself a CIB strategy (with the same belief generation system as others) suggests the possibility of a Coordinators' Nash equilibrium (CNE) where all coordinators are using CIB strategies with identical and self-consistent belief generation systems. We refer to such a CNE as a CIB-CNE. More formally, a CIB-CNE is a CIB strategy profile $(\lambda^{*i}, \psi^i)_{i\in\mathcal{I}}$ where (i) all coordinators have the same belief generation system, i.e., for all $i \in \mathcal{I}$, $\psi^i = \psi^*$ for some ψ^*, (ii) for each $k \in \mathcal{I}$, $\psi^{*,k}$ is consistent with λ^k, and (iii) for each $i \in \mathcal{I}$, the CIB strategy (λ^{*i}, ψ^i) is a best response for coordinator i to $(\lambda^{*k}, \psi^k)_{k\in\mathcal{I}\setminus\{i\}}$.

Notice that in a CIB-CNE all coordinators are using the same belief generation system, hence the CCI B_t^i (as defined in Definition 12) is the same for all coordinators. We denote the identical B_t^i for all coordinators by B_t. Furthermore, when all coordinators other than i are using fixed CIB strategies, (B_t, S_t^i) can be viewed as an information state for coordinator i's stochastic control problem (see proof of Lemma 7 for details). Based on this observation, we introduce a backward inductive computation procedure for determining CIB-CNEs where B_t is used as an information state. Our procedure decomposes the game into a collection of one-stage games, one for each time t and each realization of B_t. These one-stage games are used to characterize a CIB-CNE in a backward inductive manner.

Definition 15 *(Stage game)* Given the value functions $V_{t+1} = (V_{t+1}^i)_{i\in\mathcal{I}}$, where $V_{t+1}^i : \mathcal{B}_{t+1} \times \mathcal{S}_{t+1}^i \mapsto \mathbb{R}$, a realization of the compressed common information $b_t = (\boldsymbol{\pi}_t, y_{t-d+1:t-1}, u_{t-d:t-1})$ where $\boldsymbol{\pi}_t = (\pi_t^i)_{i\in\mathcal{I}}, \pi_t^i \in \Delta(\mathcal{S}_t^i)$, and update functions $\psi_t^* = (\psi_t^{*,i})_{i\in\mathcal{I}}, \psi_t^{*,i} : \mathcal{B}_t \times \mathcal{Y}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{S}_{t+1}^i)$, we define a stage game for the coordinators dynamic game as follows:

Stage Game $G_t(V_{t+1}, b_t, \psi_t^*)$:

- There are $|\mathcal{I}|$ players, each representing a coordinator.
- (V_{t+1}, b_t, ψ_t^*) are commonly known.
- Nature chooses $\mathbf{Z}_t = (\mathbf{S}_t, \mathbf{X}_{t-d+1:t}, \mathbf{W}_t^Y)$,[7] where $\mathbf{S}_t = (S_t^k)_{k\in\mathcal{I}}$.
- Player i observes $S_t^i = s_t^i$.

[7] Since $\mathcal{X}_t, \mathcal{U}_t, \mathcal{Y}_t$ are finite sets, one can assume that $\mathbf{W}_t^Y$ also takes finite values without lost of generality.

– Player i's belief on $\mathbf{Z}_t$ is given by

$$\begin{aligned}\beta_t^i(\tilde{z}_t|s_t^i) = \mathbf{1}_{\{\tilde{s}_t^i = s_t^i\}} \prod_{k\neq i} \pi_t^k(\tilde{s}_t^k) \times \\ \times \prod_{k\in\mathcal{I}} P_t^k(\tilde{x}_{t-d+1:t}^k | y_{t-d+1:t-1}^k, u_{t-d:t-1}, \tilde{s}_t^k)\mathbb{P}(\tilde{w}_t^{k,Y}), \\ \forall \tilde{z}_t = (\tilde{s}_t, \tilde{x}_{t-d+1:t}, \tilde{w}_t^Y) \in \mathcal{S}_t \times \mathcal{X}_{t-d+1:t} \times \mathcal{W}_t^Y. \end{aligned} \tag{4}$$

where P_t^k is the belief function defined in Eq. (1).

– Player i selects a prescription $\boldsymbol{\Gamma}_t^i \in \mathcal{A}_t^i$ as her action.

– Player i has utility

$$Q_t^i(\mathbf{Z}_t, \boldsymbol{\Gamma}_t) = r_t^i(\mathbf{X}_t, \mathbf{U}_t) + V_{t+1}^i(B_{t+1}, S_{t+1}^i), \tag{5}$$

where

$$\begin{aligned} U_t^{k,j} &= \Gamma_t^{k,j}(\mathbf{X}_t^{k,j}) \quad \forall (k,j) \in \mathcal{N}, \\ B_{t+1} &= ((\Pi_{t+1}^k)_{k\in\mathcal{I}}, (y_{t-d+2:t-1}, \mathbf{Y}_t), (u_{t-d+1:t-1}, \mathbf{U}_t)), \\ \Pi_{t+1}^k &= \psi_t^{*,k}(b_t, \mathbf{Y}_t, \mathbf{U}_t) \quad \forall k \in \mathcal{I}, \\ Y_t^k &= \ell_t^j(\mathbf{X}_t^k, \mathbf{U}_t, W_t^{k,Y}) \quad \forall k \in \mathcal{I}, \\ S_{t+1}^i &= \iota_t^i(S_t^i, \mathbf{X}_{t-d+1}^i, \boldsymbol{\Gamma}_t^i) \end{aligned}$$

Given the stage game $G_t(V_{t+1}, b_t, \psi_t^*)$, we define two associated concepts:

Definition 16 *(IBNE correspondence)* Given the value functions $V_{t+1} = (V_{t+1}^i)_{i\in\mathcal{I}}$, where $V_{t+1}^i : \mathcal{B}_{t+1} \times \mathcal{S}_{t+1}^i \mapsto \mathbb{R}$ and belief update functions $\psi_t^* = (\psi_t^{*,i})_{i\in\mathcal{I}}$, $\psi_t^{*,i} : \mathcal{B}_t \times \mathcal{Y}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{S}_{t+1}^i)$, the *Interim Bayesian Nash equilibrium correspondence* $\mathrm{IBNE}_t(V_{t+1}, \psi_t^*)$ is defined as the set of all $\lambda_t = (\lambda_t^i)_{i\in\mathcal{I}}$, $\lambda_t^i : \mathcal{B}_t \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$ such that

$$\begin{aligned}\lambda_t^i(b_t, s_t^i) \in \underset{\eta\in\Delta(\mathcal{A}_t^i)}{\arg\max} \left(\sum_{\tilde{z}_t, \tilde{\gamma}_t} \left[\eta(\tilde{\gamma}_t^i) Q_t^i(\tilde{z}_t, \tilde{\gamma}_t) \beta_t^i(\tilde{z}_t|s_t^i) \prod_{k\neq i} \lambda_t^k(\tilde{\gamma}_t^k|b_t, \tilde{s}_t^k) \right] \right) \\ \forall b_t \in \mathcal{B}_t, s_t^i \in \mathcal{S}_t^i, \forall i \in \mathcal{I}, \end{aligned}$$

where β_t^i and Q_t^i are defined using $(V_{t+1}^i, b_t, \psi_t^*)$ in (4) and (5), respectively.

Definition 17 *(DP operator)* Given a value function $V_{t+1}^i : \mathcal{B}_{t+1} \times \mathcal{S}_{t+1}^i \mapsto \mathbb{R}$ and a CIB strategy profile (λ_t^*, ψ_t^*) at time t, where $\lambda_t^* = (\lambda_t^{*i})_{i\in\mathcal{I}}$, $\lambda_t^{*i} : \mathcal{B}_t \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$ and $\psi_t^* = (\psi_t^{*,i})_{i\in\mathcal{I}}$, $\psi_t^{*,i} : \mathcal{B}_t \times \mathcal{Y}_t \times \mathcal{U}_t \mapsto \Delta(\mathcal{S}_{t+1}^i)$, the *dynamic programming operator* DP_t^i defines the value function at time t through

$$[\mathrm{DP}_t^i(V_{t+1}^i, \lambda_t^*, \psi_t^*)](b_t, s_t^i) := \sum_{\tilde{z}_t, \tilde{\gamma}_t} Q_t^i(\tilde{z}_t, \tilde{\gamma}_t) \beta_t^i(\tilde{z}_t|s_t^i) \prod_{k\in\mathcal{I}} \lambda_t^{*k}(\tilde{\gamma}_t^k|b_t, \tilde{s}_t^k),$$

where β_t^i and Q_t^i are defined using $(V_{t+1}^i, b_t, \psi_t^*)$ in (4) and (5), respectively.

Theorem 2 (Sequential decomposition) *Let $(\lambda^{*i}, \psi^*)_{i\in\mathcal{I}}$ be a CIB strategy profile with identical belief generation system ψ^* for all $i \in \mathcal{I}$. If this strategy profile satisfies the dynamic program defined below:*

$$V_{T+1}^i(\cdot,\cdot) = 0 \quad \forall i \in \mathcal{I};$$

and for $t \in \mathcal{T}$

$$\lambda_t^* \in \mathrm{IBNE}_t(V_{t+1}, \psi_t^*); \tag{6}$$
$$\psi_t^* \textit{ is consistent with } \lambda_t^*;$$
$$V_t^i := \mathrm{DP}_t^i(V_{t+1}^i, \lambda_t^*, \psi_t^*) \quad \forall i \in \mathcal{I}, \tag{7}$$

then $(\lambda^{*i}, \psi^*)_{i\in\mathcal{I}}$ *forms a CIB-CNE.*

Proof See Appendix H. □

Remark 10 Note that (6) and (7) can be verified for each realization $b_t \in \mathcal{B}_t$ separately, i.e., one can check that $\lambda_t^*(b_t, \cdot)$ is an IBNE of the stage game $G_t(V_{t+1}, b_t, \psi_t^*(b_t, \cdot))$, and that $\psi_t^*(b_t, \cdot)$ is consistent with $\lambda_t^*(b_t, \cdot)$ for each b_t.

5.5 Existence of CIB-CNE

We have shown in Theorem 1 that an SPIB-CNE always exists. However, a CIB-CNE does not necessarily exist, even when each team contains only one member (i.e., in games among individuals). We present below one example where CIB-CNEs do not exist.

Example 1 Consider a 3-stage dynamic game (i.e., $\mathcal{T} = \{1, 2, 3\}$) with two players: Alice (A) and Bob (B). Each player forms a one-person team. Let $X_t^A \in \{-1, 1\}$ and $X_t^B \equiv \varnothing$, i.e., Bob is not associated with a state. Let $\mathbf{Y}_t = \varnothing$, i.e., there is no public observation of the states. The initial state X_1^A is uniformly distributed on $\{-1, 1\}$. At $t = 1$, (a) Alice can choose an action $U_1^A \in \{-1, 1\}$ and Bob has no actions to take; (b) the next state is given by $X_2^A = X_1^A \cdot U_1^A$; (c) the instantaneous reward is given by

$$r_1^A(\mathbf{X}_1, \mathbf{U}_1) = -r_1^B(\mathbf{X}_1, \mathbf{U}_1) = \varepsilon \cdot \mathbf{1}_{\{U_1^A = +1\}},$$

where $\varepsilon \in (0, \frac{1}{3})$.

At $t = 2$, (a) neither player has any action to take; (b) the state at next time is given by $X_3^A = X_2^A$; (c) the instantaneous rewards are 0 for both players; (This stage is a dummy stage inserted in the game to alter the definition of the CCI at the beginning of the last stage.)

At $t = 3$, (a) Alice has no action to take, and Bob chooses $U_3^B \in \{\mathrm{L}, \mathrm{R}\}$; (b) The instantaneous reward $r_3^A(\mathbf{X}_3, \mathbf{U}_3)$ for Alice is given by

$$r_3^A(-1, \mathrm{L}) = 0, \quad r_3^A(-1, \mathrm{R}) = 1$$
$$r_3^A(+1, \mathrm{L}) = 2, \quad r_3^A(+1, \mathrm{R}) = 0$$

and $r_3^B(\mathbf{X}_3, \mathbf{U}_3) = -r_3^A(\mathbf{X}_3, \mathbf{U}_3)$.

In a game where each team contains only one person, we can assume the delay d to be any number (see Remark 1). In the next proposition, we view Example 1 as a game among teams with internal delay $d = 1$.

Proposition 1 *There exist no CIB-CNE in the game described in Example 1.*

Proof See Appendix I. □

Remark 11 One can provide an example for non-existence of CIB-CNE for any $d > 0$ by inserting $d-1$ additional dummy stages (analogous to stage 2) into Example 1, and viewing it as a game among teams with internal delay d. Example 1 can also be used to show that the CIB-PBE concept defined in Ouyang et al. [45] for games among individuals does not exist in general, hence the conjecture in Ouyang et al. [45] that a CIB-PBE always exists is not true.

Intuitively, the reason that a CIB-CNE does not exist in this game is that at $t = 3$, a CIB strategy requires Bob to choose his action based only on a compressed version of his information rather than the full information. This compression does not hurt Bob's ability to form a best response. However, in an equilibrium, Bob needs to carefully choose from the set of optimal responses to induce Alice to play the predicted mixed strategy. Being unable to choose different actions under different histories due to information compression makes Bob unable to sustain an equilibrium. In this game, as in the example in Maskin and Tirole [35], payoff irrelevant information plays an essential role in sustaining the equilibrium.

In the remainder of this section, we present two subclasses of the dynamic games described in Sect. 2 where CIB-CNEs exist.

5.5.1 Signaling-Neutral Teams

In this subsection, we consider $d = 1$. One subclass of games where CIB-CNEs exist is when the teams are *signaling-neutral*. In these games, the agents are indifferent in terms of signaling to other teams, i.e., revealing more or less information about their private information to the other teams does not affect their utility. (Note that agents can always actively reveal information to their teammates through their actions.)

We shall now describe the game:

Definition 18 A team i whose state $\mathbf{X}_t^i$ can be recovered from $(\mathbf{Y}_t^i, \mathbf{U}_t)$ (i.e., for every fixed u_t, $\ell_t^i(x_t^i, u_t, W_t^{i,Y})$ has disjoint support for different $x_t^i \in \mathcal{X}_t^i$) is called a *public* team. Otherwise, it is called *private* team.

For a public team i, the private state $\mathbf{X}_{t-1}^i$ is effectively part of the common information of all members of all teams at time t.

Definition 19 *(Information dependency graph)* The information dependency graph $\mathcal{G}$ of a dynamic game is a directed graph defined as follows: The vertices represent the teams. A directed edge $i \leftarrow j$ is present if either the state transition, the observation, or the instantaneous reward of team i at some time t depends directly on either the state or the actions of team j. In other words, there is no directed edge from j to i if and only if $\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t^{-j}, W_t^{i,X})$, $\mathbf{Y}_t^i = \ell_t^i(\mathbf{X}_t^i, \mathbf{U}_t^{-j}, W_t^{i,Y})$ and $r_t^i(\mathbf{X}_t, \mathbf{U}_t) = r_t^i(\mathbf{X}_t^{-j}, \mathbf{U}_t^{-j})$ for some functions f_t^i, ℓ_t^i, r_t^i for all t. Self-loops are not considered in this graph.

Theorem 3 *Let $d = 1$. If every strongly connected component of the information dependency graph $\mathcal{G}$ of a dynamic game consists of either (I) a single team, or (II) multiple public teams, then a CIB-CNE exists.*

Proof See Appendix J. □

Remark 12 The precedence relation among teams considered in Theorem 3 is similar to the s-partition of teams that was presented and analyzed in Yoshikawa [67].

When the condition in Theorem 3 is satisfied, all teams will be neutral in signaling: When a private team i sends information, this information is only useful to those teams whose actions do not affect team i's utility. Public players are always neutral in signaling since their state history is publicly available.

Notice that in Example 1, Alice (as a one-person team) is a private team while Bob is a public team. The instantaneous reward of Bob at $t = 3$ depends on Alice's state X_2^A, while Alice's instantaneous reward at $t = 3$ depends on Bob's action. Hence, Alice and Bob form a strongly connected component in the information dependency graph.

5.5.2 Signaling-Free Equilibria

In this section, we introduce another class of games where CIB-CNE exists. These games are games-among-teams extensions of *Game M* defined in Ouyang et al. [45]. We present the result for a general $d > 0$.

Theorem 4 *A dynamic game that satisfies all of the following conditions has a CIB-CNE:*

(i) States are uncontrolled, i.e., $\mathbf{X}^i_{t+1} = f^i_t(\mathbf{X}^i_t, W^{i,X}_t)$.

(ii) Observations are uncontrolled, i.e., $Y^i_t = \ell^i_t(\mathbf{X}^i_t, W^{i,Y}_t)$.

(iii) Instantaneous rewards of team i can be expressed as $r^i_t(\mathbf{X}^{-i}_t, \mathbf{U}_t)$.

Proof See Appendix K for a direct proof. Alternatively, one can first assume that the teams share information with a delay of $d = 0$, then we can view a team as one individual since team members have the same information. One can then apply results for *Game M* in Ouyang et al. [45] to obtain an equilibrium where each player/team plays a public strategy (i.e., a strategy that does not use private information), in particular, a strategy where actions are solely based on the common information-based belief. Since public strategies can also be played when $d > 0$, we conclude that the equilibrium we obtained is also an equilibrium for the original game. □

6 Additional Result

Consider a special case of the model in Sect. 2 where both the evolution and the observations of the local states of each member of each team are conditionally independent given the actions, i.e.,

$$\begin{aligned} X^{i,j}_{t+1} &= f^{i,j}_t(X^{i,j}_t, \mathbf{U}_t, W^{i,j}_t), \\ \mathbf{Y}^i_t &= (Y^{i,j}_t)_{(i,j)\in\mathcal{N}_i}, \\ Y^{i,j}_t &= \ell^{i,j}_t(X^{i,j}_t, \mathbf{U}_t, W^{i,j,Y}_t), \end{aligned}$$

where $(W^{i,j,X}_t, W^{i,j,Y}_t)_{t\in\mathcal{T},(i,j)\in\mathcal{N}}$ are mutually independent primitive random variables.

In this case, we show that the independence among team members' state dynamics enables us to consider equilibria where the coordinators assign prescriptions that map $X^{i,j}_t$ to $U^{i,j}_t$ (instead of mapping $X^{i,j}_{t-d+1:t}$ to $U^{i,j}_t$). This is because, given H^i_t, the belief of member (i, j) about her teammates' states is independent of $X^{i,j}_{t-d+1:t}$. In other words, one can replace the hidden information $\mathbf{X}^i_{t-d+1:t}$ with the *sufficient hidden information* $\mathbf{X}^i_t$.[8]

[8] The compression of hidden information to sufficient hidden information is similar to the shedding of irrelevant information in Mahajan [29].

Definition 20 *(Simple prescriptions)* A *simple prescription* for coordinator i at time t is a collections of functions $\bar{\gamma}_t^i = (\bar{\gamma}_t^{i,j})_{(i,j)\in\mathcal{N}_i}$, $\bar{\gamma}_t^{i,j} : \mathcal{X}_t^{i,j} \mapsto \mathcal{U}_t^{i,j}$.

Lemma 8 *Let g^i be a behavioral coordination strategy of coordinator i. Then, there exists a behavioral coordination strategy $\bar{g}^i$ payoff-equivalent to g^i such that $\bar{g}^i$ only assigns simple prescriptions.*

Proof See Appendix L. □

Given the above result, one can restrict attention to *sufficient hidden information-based* strategies where each coordinator i assigns simple prescriptions based on $\overline{H}_t^i$. With this restriction, results analogous to that of Sects. 4 and 5 can be derived considering similar compression of private and common information.

7 Discussion

7.1 Implementation of Behavioral Coordination Strategies

One can interpret behavioral coordination strategies as strategies with coordinated randomization, i.e., the strategies are randomized, but all the team members know exactly how this randomization is done. We note that one can view the main purpose of randomization as to "confuse" other teams. As such, it is best to use coordinated randomization where every team member knows what partial mapping their teammate is using; such coordinated randomization is superior to private and independent randomization by each individual member in a team: This is because individual randomization can create information that are unknown to teammates, while the same "confuse" effect to other teams can be achieved with coordinated randomization.

To implement behavioral coordination strategies, a team can utilize a correlation device which generates a random seed at each time t. Then, each member (i, j) of the team i can choose an action based on $H_t^{i,j}$ and present and past random seeds generated by the correlation device, or equivalently, choose an action based on $(H_t^{i,j}, \boldsymbol{\Gamma}_{1:t-1}^i)$ and the current random seed, where $\boldsymbol{\Gamma}_{1:t-1}^i$ is sequentially updated. If the behavioral coordination strategy is a CIB strategy, then member (i, j) needs to use $(B_t, \mathbf{X}_{t-d}^i, \boldsymbol{\Phi}_t^i, X_{t-d+1:t}^{i,j})$ and current random seed to chose an action, where $(B_t, \boldsymbol{\Phi}_t^i)$ are sequentially updated.

In the absence of correlation devices accessible at every time, a behavioral coordination strategy can also be implemented as its equivalent mixed strategy (recall Lemma 1 and Lemma 2): Before the beginning of the game, the team can jointly pick a strategy profile in $\mathcal{G}^i$ randomly, according to a distribution induced from the behavioral coordination strategy.

7.2 Stage Game: IBNE Versus BNE

One can observe that the belief of the agents defined in the stage game (Definition 15) can be seen as a conditional distribution derived from the common prior

$$\beta_t(\tilde{z}_t) = \prod_{k\in\mathcal{I}} \left[\pi_t^k(\tilde{s}_t^k) P_t^k(\tilde{x}_t^k | y_{t-d+1:t-1}^k, u_{t-d:t-1}, \tilde{s}_t^k) \mathbb{P}(\tilde{w}_t^{k,Y}) \right]. \tag{8}$$

However, in the aforementioned stage game we focus on the beliefs of agents instead of a common prior, and we use *Interim Bayesian Nash equilibrium* (IBNE) as the equilibrium

concept instead of BNE. This is because, unlike a standard Bayesian game with a common prior, the true prior of the stage game is dependent on the actual strategy played in previous stages. The prior β_t described in (8) may not be a true prior, since some coordinator i may have already deviated from the strategy prediction which π_t^i's were relying on. However, coordinator i is always trying to optimize her reward given (b_t, s_t^i), no matter $\pi_t^i(s_t^i) = 0$ or not. Hence, in this stage game, we must consider the player's belief and strategy for all possible realizations s_t^i under *any* strategy profile, not just those with positive probability under the prior in (8). The corresponding equilibrium concept is Interim Bayesian Nash Equilibrium instead of Bayes–Nash equilibrium. IBNE strengthens BNE by requiring the strategy of an agent to be optimal under *all* private information realizations, including those with zero probability under the common prior.

7.3 Choice of Compressed Common Information

In decentralized control [43] and certain settings of games among individuals [42,45], a common information-based belief Π_t on the state is usually enough to serve as an information state, or compression of common information. However, in our setting we use a subset of actions and observations in addition to the CIB belief as the compressed common information. We argue below that this is necessary for our setting.

To illustrate the point, consider the case $d = 1$ and assume that all coordinators use the same belief generation system and hence, the same CCI (denoted by B_t^*). An alternative for the CCI $B_t^* = ((\Pi_t^{*,i})_{i\in\mathcal{I}}, \mathbf{U}_{t-1})$ is the CIB belief $\tilde{\boldsymbol{\Pi}}_t^* = (\tilde{\Pi}_t^{*,i})_{i\in\mathcal{I}}, \tilde{\Pi}_t^{*,i} \in \Delta(\mathcal{X}_{t-1:t}^i)$ where $\tilde{\Pi}_t^{*,i}$ represents the belief on $\mathbf{X}_{t-1:t}^i$ based on common information. One might argue that we can use $\tilde{\boldsymbol{\Pi}}_t^*$ instead of B_t^* through the following argument: After we transform the game into games among coordinators, because of the full recall of coordinator i, coordinator i's belief (on other coordinators' private information and all hidden information) is independent of her behavioral coordination strategy $\tilde{g}^i$. Hence, coordinator i can always form this belief as if she was using the strategy prediction g^{*i} no matter what strategy she is actually using.

However, this argument can run into technical problems: A crucial step for Lemma 7 is Eq. (26), which establishes that coordinator i's belief can be expressed as a function of $(B_t^*, \mathbf{X}_{t-1}^i)$ for any behavioral coordination strategy $\tilde{g}^i$ coordinator i might use. To use $\tilde{\boldsymbol{\Pi}}_t^*$ alone as the information state, one needs to argue that coordinator i's belief on her hidden information, $\mathbb{P}(X_t^i = \cdot | x_{t-1}^i, u_{t-1})$, can be computed solely through $(\tilde{\pi}_t^{*,i}, x_{t-1}^i)$ without using u_{t-1}. Through belief independence of strategy, one may argue that

$$\begin{aligned}
\mathbb{P}(x_t^i | x_{t-1}^i, u_{t-1}) &= \mathbb{P}^{g^{*i}, g^{*-i}}(x_t^i | x_{t-1}^i, u_{t-1}) \\
&= \mathbb{P}^{g^{*,i}, g^{*,-i}}(x_t^i | x_{t-1}^i, y_{1:t-1}, u_{1:t-1}) \\
&= \frac{\mathbb{P}^{g^{*,i}, g^{*,-i}}(x_t^i, x_{t-1}^i | y_{1:t-1}, u_{1:t-1})}{\mathbb{P}^{g^{*,i}, g^{*,-i}}(x_{t-1}^i | y_{1:t-1}, u_{1:t-1})} \\
&= \frac{\tilde{\pi}_t^{*,i}(x_{t-1}^i, x_t^i)}{\sum_{\tilde{x}_t^i} \tilde{\pi}_t^{*,i}(x_{t-1}^i, \tilde{x}_t^i)}.
\end{aligned} \tag{9}$$

However, the above argument is not always valid. It is only valid when the denominator of (9) is nonzero, but it can be zero. One simple example is the following: Let $\hat{x}_{t-1}^i \in \mathcal{X}_{t-1}^i$ be some fixed state and $\hat{u}_{t-1}^i \in \mathcal{X}_{t-1}^i$ be some fixed action profile. Let $\hat{\mathcal{A}}_{t-1}^i$ be the set of prescriptions that maps $\hat{x}_{t-1}^i$ to $\hat{u}_{t-1}^i$. Suppose that the strategy prediction g^{*i} is a behavioral

coordination strategy satisfying the following:

$$g^{*i}_{t-1}(\overline{h}^i_{t-1})(\gamma^i_{t-1}) = 0 \qquad \forall \overline{h}^i_{t-1} \in \overline{\mathcal{H}}^i_{t-1}, \gamma^i_{t-1} \in \hat{\mathcal{A}}^i_{t-1},$$

i.e., under g^{*i}, coordinator i never assigns any prescription that maps $\hat{x}^i_{t-1}$ to $\hat{u}^i_{t-1}$. If $\tilde{\pi}^{*,i}_t$ is consistent with the strategy prediction g^{*i}, then

$$\sum_{\tilde{x}^i_t} \tilde{\pi}^{*,i}_t(\hat{x}^i_{t-1}, \tilde{x}^i_t) = \mathbb{P}^{g^i, g^{-i}}(\hat{x}^i_{t-1} | h^0_t) = 0$$

if $u^i_{t-1} = \hat{u}^i_{t-1}$. When coordinator i uses a strategy $\tilde{g}^i$ such that $\mathbf{X}^i_{t-1} = \hat{x}^i_{t-1}$, $\mathbf{U}^i_{t-1} = \hat{u}^i_{t-1}$ could happen with nonzero probability, coordinator i cannot use $\tilde{\pi}^{*,i}_t$ to form her belief on her hidden information. This is contrary to what we need in Eq. (26) in the proof of Lemma 7, which states that the belief function is compatible with *any* behavioral coordination strategy $\tilde{g}^i$.

8 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 controlled by the actions of all agents. We developed a general approach to characterize a subset of Nash equilibria with the following feature: At each time, each agent can make their decision based on a compressed version of their information, instead of the full information. We identified two subclasses of strategies: sufficient private information-based (SPIB) strategies, which only compresses private information, and compressed information-based (CIB) strategies, which compresses both common and private information. We showed that SPIB-strategy-based equilibria always exist and can attain all the payoff profiles of Nash Equilibria. On the other hand, CIB strategy-based equilibria do not always exist. We developed a backward inductive sequential procedure, whose solution (if it exists) is a CIB strategy-based equilibrium. We characterized certain game environments where the solution exists. Our results highlight the discord between compression of information, ability of compression-based strategies to sustain all or some of the equilibrium payoff profiles, and backward inductive sequential computation of equilibria in stochastic dynamic games.

Moving forward, there are a few research problems arising from this work: (i) discovering broader conditions for the existence of CIB-CNE in the model of this paper; (ii) developing an efficient algorithm which solves the dynamic program of CIB-CNE (when a solution exists); (iii) determining minimal additional information needed to be added to the CCI such that CIB-CNE (under the new CCI) is guaranteed to attain some or all of the equilibrium payoff profiles; (iv) defining a notion of ϵ-CIB-CNE, analyzing its existence, and developing sequential computation procedures to find them; (v) characterizing compression-based subclasses of equilibrium refinements for games among teams.

Other future research directions include identifying a suitable compression of information and developing a sequential decomposition for other models of games among teams, for

example (i) games with continuous state and action spaces (e.g., linear quadratic Gaussian settings), and (ii) general models with non-observable actions.

Funding This work is supported by National Science Foundation (NSF) Grant No. ECCS 1750041, ECCS 2038416, ECCS 1608361, CCF 2008130, Army Research Office (ARO) Award No. W911NF-17-1-0232, and Michigan Institute for Data Science (MIDAS) Sponsorship Funds by General Dynamics.

Availability of Data and Material Not applicable.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Code Availability Not applicable.

A Two Examples

A.1 A Motivating Example for Sect. 2

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 contains the minimum richness of strategies that ensures an equilibrium exists in games among teams. In particular, if we restrict the teams to use independently randomized strategies, i.e., type 1 and type 2 strategies described in Sect. 2.2, then an equilibrium may not exist. This example is similar to the examples in Farina et al. [13], Zhang and An [68], Anantharam and Borkar [3] in spirit, despite the fact that in our example the players in the same team have asymmetric information.

Example 2 *(Guessing game)* Consider a two-stage 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 are given by

$$r_1^B(\mathbf{X}_1, \mathbf{U}_1) = 0,$$
$$r_2^B(\mathbf{X}_2, \mathbf{U}_2) = \mathbf{1}_{\{X_2^{A,1} = U_2^{B,1}\}} + \mathbf{1}_{\{X_2^{A,2} = U_2^{B,2}\}}.$$

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.

It can be verified (see Appendix A.2.1 for a detailed derivation) that if we restrict both teams to use independently randomized strategies (including deterministic strategies), then there exists no equilibria. However, there does exist an equilibrium where Team A randomizes in a correlated manner, specifically, the following strategy profile σ^*: 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 σ^*, each agent (A, j) chooses a uniform random action irrespective of their states. It is important to have $(A, 1)$ and $(A, 2)$ choose these actions in a correlated way to ensure that they obtain the full instantaneous reward while not revealing any information.

A.2 An Illustrative Example for Sect. 3

The following example illustrates how to visualize games among teams from the coordinators' viewpoint.

Example 3 Consider a variant of the Guessing Game in Example 2 with the same system model and information structure but different action sets and reward functions. In the new game, Team A moves at both $t = 1$ and $t = 2$, with $\mathcal{U}_t^{A,j} = \{-1, 1\}$ for $t = 1, 2$ and $j = 1, 2$. Team B moves only at time $t = 2$ as in the original game. The new reward functions are given by

$$r_1^A(\mathbf{X}_1, \mathbf{U}_1) = 0,$$
$$r_2^A(\mathbf{X}_2, \mathbf{U}_2) = \mathbf{1}_{\{X_2^{A,2} = U_2^{A,1}, X_2^{A,1} = U_2^{A,2}\}} + \mathbf{1}_{\{\mathbf{X}_2^A \neq \mathbf{U}_2^B\}},$$
$$r_1^B(\mathbf{X}_1, \mathbf{U}_1) = 0,$$
$$r_2^B(\mathbf{X}_2, \mathbf{U}_2) = \mathbf{1}_{\{\mathbf{X}_2^A = \mathbf{U}_2^B\}}.$$

In this example, Team A's task is to guess its own state after a round of publicly observable communication while not leaking information to Team B.

A Team Nash equilibrium $(\sigma^{*A}, \sigma^{*B})$ of this game is as follows: Team A chooses one of the four pure strategy profiles listed below with equal probability:

- $\mu_1^{A,1}(x_1^{A,1}) = -x_1^{A,1}, \mu_1^{A,2}(x_1^{A,2}) = x_1^{A,2},$
  $\mu_2^{A,1}(\mathbf{u}_1, x_{1:2}^{A,1}) = u_1^{A,2}, \mu_2^{A,2}(\mathbf{u}_1, x_{1:2}^{A,2}) = -u_1^{A,1};$
- $\mu_1^{A,1}(x_1^{A,1}) = -x_1^{A,1}, \mu_1^{A,2}(x_1^{A,2}) = -x_1^{A,2},$
  $\mu_2^{A,1}(\mathbf{u}_1, x_{1:2}^{A,1}) = -u_1^{A,2}, \mu_2^{A,2}(\mathbf{u}_1, x_{1:2}^{A,2}) = -u_1^{A,1};$
- $\mu_1^{A,1}(x_1^{A,1}) = x_1^{A,1}, \mu_1^{A,2}(x_1^{A,2}) = x_1^{A,2},$
  $\mu_2^{A,1}(\mathbf{u}_1, x_{1:2}^{A,1}) = u_1^{A,2}, \mu_2^{A,2}(\mathbf{u}_1, x_{1:2}^{A,2}) = u_1^{A,1};$

$$\bullet\ \mu_1^{A,1}(x_1^{A,1}) = x_1^{A,1},\ \mu_1^{A,2}(x_1^{A,2}) = -x_1^{A,2},$$
$$\mu_2^{A,1}(\mathbf{u}_1, x_{1:2}^{A,1}) = -u_1^{A,2},\ \mu_2^{A,2}(\mathbf{u}_1, x_{1:2}^{A,2}) = u_1^{A,1};$$

while Team B choose $\mathbf{U}_2^B$ uniformly at random independent of $\mathbf{U}_1$. In words, from Team B's point of view, Team A chooses $\mathbf{U}_1^A$ to be a uniform random vector independent of $\mathbf{X}_1^A$. However, the randomization is done in a coordinated manner: Before the game starts, both members of team A randomly draw a card from two cards, where one card says "lie" and the other says "tell the truth." Both players then tell each other what card they have drawn before the game starts. At time $t = 1$, both players in Team A play the strategy indicated by their cards. At time $t = 2$, Team A can then perfectly recover $\mathbf{X}_1^A$ from $\mathbf{U}_1^A$ and the knowledge about the strategy being used at $t = 1$.

Now, we describe Team A's equilibrium strategy by the equivalent coordinator A's behavioral strategy. Use **ng** to denote the prescription that maps -1 to 1 and 1 to -1. Use **id** to denote the identity map prescription, i.e., the prescription that maps -1 to -1 and 1 to 1. Use $\mathbf{cp}_b$ to denote the constant prescription that always instruct individuals to play $b \in \{-1, 1\}$. The mixed strategy profile σ^{*A} is equivalent to the following behavioral coordination strategy: At time $t = 1$, $g_1^A(\varnothing) \in \Delta(\mathcal{A}_1^{A,1} \times \mathcal{A}_1^{A,2})$ satisfies

$$g_1^A(\varnothing)(\gamma_1^{A,1}, \gamma_1^{A,2}) = \frac{1}{4} \qquad \forall \gamma_1^{A,1}, \gamma_1^{A,2} \in \{\mathbf{ng}, \mathbf{id}\}.$$

At time $t = 2$, $g_2^A : \mathcal{U}_1^{A,1} \times \mathcal{U}_1^{A,2} \times \mathcal{A}_1^{A,1} \times \mathcal{A}_1^{A,2} \mapsto \Delta(\mathcal{A}_2^{A,1} \times \mathcal{A}_2^{A,2})$ is a deterministic strategy that satisfies

$$g_2^A(u^1, u^2, \mathbf{ng}, \mathbf{id}) = \mathrm{DM}(\mathbf{cp}_{u^2}, \mathbf{cp}_{-u^1}),$$
$$g_2^A(u^1, u^2, \mathbf{ng}, \mathbf{ng}) = \mathrm{DM}(\mathbf{cp}_{-u^2}, \mathbf{cp}_{-u^1}),$$
$$g_2^A(u^1, u^2, \mathbf{id}, \mathbf{id}) = \mathrm{DM}(\mathbf{cp}_{u^2}, \mathbf{cp}_{u^1}),$$
$$g_2^A(u^1, u^2, \mathbf{id}, \mathbf{ng}) = \mathrm{DM}(\mathbf{cp}_{-u^2}, \mathbf{cp}_{u^1}),$$

where $\mathrm{DM} : \mathcal{A}_2^{A,1} \times \mathcal{A}_2^{A,2} \mapsto \Delta(\mathcal{A}_2^{A,1} \times \mathcal{A}_2^{A,2})$ represents the delta measure. In words, the coordinator of Team A randomly chooses one of all four possible prescription profiles at time $t = 1$. At time $t = 2$, based on the observed action and the prescriptions chosen before, the coordinator of Team A directly assign actions to agents to instruct them to recover the state from the actions at $t = 1$. Note that the behavioral coordination strategy at $t = 2$ depends explicitly on the past prescription $\boldsymbol{\Gamma}_1^A$ in addition to the realization of past actions. This is because the coordinator needs to remember not only the agents' actions, but also the rationale behind those actions in order to interpret the signals sent through the actions.

A.2.1 Proof of Claim in Example 2

Define two pure strategies μ^A and $\tilde{\mu}^A$ of Team A as follows:

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

Now, assume that Team A and Team B are restricted to use independently randomized strategies (type 2 strategies defined in Sect. 2.2). We will show in two steps that there exist no equilibria within this class of strategies.

Step 1: If Team A and Team B's type 2 strategies form an equilibrium, then Team A is playing either μ^A or $\tilde{\mu}^A$.

Let $p_j(x)$ denote the probability that player (A, j) plays $U_1^{A,j} = -x$ given $X_1^{A,j} = x$. Define

$$q_j = \frac{1}{2}p_j(-1) + \frac{1}{2}p_j(+1),$$

i.e., the ex-ante probability that player (A, j) "lies."

Then, we have

$$\mathbb{E}[r_1^A(\mathbf{X}_1, \mathbf{U}_1)] = q_1(1-q_2) + q_2(1-q_1).$$

Under an equilibrium, Team B will optimally respond to Team A strategy's described through (p_1, p_2). We can find a lower bound of Team B's reward by fixing a strategy: Consider the "random guess" strategy of Team B, where each of (B, j) (for $j = 1, 2$) chooses $U_2^{B,j}$ uniformly at random irrespective of $\mathbf{U}_1^A$ and independent of the other team member. Team B can thus guarantee an expected reward of $\frac{1}{2} + \frac{1}{2} = 1$ given any strategy of Team A. Since $r_2^A(\mathbf{X}_2, \mathbf{U}_2) = -r_2^B(\mathbf{X}_2, \mathbf{U}_2)$, we conclude that Team A's total reward in an equilibrium is upper bounded by

$$q_1(1-q_2) + q_2(1-q_1) - 1 = -q_1q_2 - (1-q_1)(1-q_2) \leq 0$$

Let σ^B denote the strategy of Team B. Let $\pi_j(u^1, u^2)$ denote the probability that player (B, j) plays $U_2^{B,j} = -u^j$ given $U_1^{A,1} = u^1$, $U_1^{A,2} = u^2$ (i.e., the probability that player (B, j) believes that (A, j) was "lying" hence guesses the opposite of what was signaled). If Team A plays μ^A, then the total reward of Team A is

$$\begin{aligned} J^A(\mu^A, \sigma^B) &= 1 - \mathbb{E}[1 - \pi_1(X_1^{A,1}, -X_1^{A,2}) + \pi_2(X_1^{A,1}, -X_1^{A,2})] \\ &= \frac{1}{4}\sum_{\mathbf{x}\in\{-1,1\}^2} (-\pi_1(\mathbf{x}) + \pi_2(\mathbf{x})). \end{aligned}$$

If Team A plays $\tilde{\mu}^A$, then the total reward of Team A is

$$\begin{aligned} J^A(\tilde{\mu}^A, \sigma^B) &= 1 - \mathbb{E}[\pi_1(-X_1^{A,1}, X_1^{A,2}) + 1 - \pi_2(-X_1^{A,1}, X_1^{A,2})] \\ &= \frac{1}{4}\sum_{\mathbf{x}\in\{-1,1\}^2} (\pi_1(\mathbf{x}) - \pi_2(\mathbf{x})). \end{aligned}$$

Observe that $J^A(\mu^A, \sigma^B) + J^A(\tilde{\mu}^A, \sigma^B) = 0$. Hence, for any σ^B, either $J^A(\mu^A, \sigma^B) \geq 0$ or $J^A(\tilde{\mu}^A, \sigma^B) \geq 0$. In particular, we can conclude that Team A's total reward is at least 0 in any equilibrium.

We have established both an upper bound and lower bound for Team A's total reward in an equilibrium. Hence, we must have

$$-q_1q_2 - (1-q_1)(1-q_2) = 0,$$

which implies $q_1 = 0, q_2 = 1$ or $q_1 = 1, q_2 = 0$. The former case corresponds to Team A playing the pure strategy μ^A, and the latter to playing $\tilde{\mu}^A$.

Step 2: There does not exist equilibria where Team A plays μ^A or $\tilde{\mu}^A$.

Suppose that Team A plays μ^A. Then, the only best response of Team B is to play $U_2^{B,1} = U_1^{A,1}$, $U_2^{B,2} = -U_1^{A,2}$. Then, Team A's total reward is $J^A(\mu^A, \sigma^B) = 1 - 1 - 1 = -1$. If

Team A deviate to $\tilde{\mu}^A$, then Team A can obtain a total reward of $+1$ (recall that $J^A(\mu^A, \sigma^B) + J^A(\tilde{\mu}^A, \sigma^B) = 0$ for any σ^B). Hence, Team A does not play μ^A at equilibrium.

Similar arguments apply to $\tilde{\mu}^A$, which completes the proof.

B Proof of Lemma 1

Given a pure strategy profile μ^i of team i, define a pure coordination strategy profile ν^i by

$$\nu_t^i(h_t^i, \gamma_{1:t-1}^i) = (\mu_t^{i,j}(h_t^i, \cdot))_{(i,j)\in\mathcal{N}_i} \quad \forall h_t^i \in \mathcal{H}_t^i, \gamma_{1:t-1}^i \in \mathcal{A}_{1:t-1}^i.$$

We first prove that for every pure strategy profile μ^i, there exist a payoff-equivalent coordination strategy profile ν^i by coupling two systems. In one of the systems, we assume that team i uses a pure strategy. In the other system, we assume that team/coordinator i uses the corresponding pure coordination strategies. We assume that all teams other than i use the same pure strategy profile $\mu^{-i} = (\mu^k)_{k\in\mathcal{I}\setminus\{i\}}$ in both systems. The realizations of primitive random variables (i.e., $(X_1^i)_{i\in\mathcal{I}}$, $(W_t^{i,X}, W_t^{i,Y})_{i\in\mathcal{I},t\in\mathcal{T}}$) are assumed to be the same for two systems. We proceed to show that the realizations of all system variables (i.e., $(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{U}_t)_{t\in\mathcal{T}}$) will be the same for both systems. As a result, the expected payoffs are the same for both systems.

We prove that the realizations of $(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{U}_t)_{t\in\mathcal{T}}$ are the same by induction on time t.

Induction Base: At $t = 1$, the realizations of $\mathbf{X}_1$ are the same for two systems by assumption. For the first system, we have

$$U_1^{i,j} = \mu_1^{i,j}(X_1^{i,j}) \quad \forall (i,j) \in \mathcal{N}_i,$$

and for the second system we have

$$\begin{aligned} \boldsymbol{\Gamma}_1^i &= \nu_t^i(H_1^i) = (\mu_t^{i,j}(\cdot))_{(i,j)\in\mathcal{N}_i}, \\ U_1^{i,j} &= \Gamma_1^{i,j}(X_1^{i,j}) \quad \forall (i,j) \in \mathcal{N}_i, \end{aligned}$$

which means that $U_1^{i,j} = \mu_1^i(X_1^{i,j})$ also holds in the second system for all $(i,j) \in \mathcal{N}_i$.

It is clear that $\mathbf{U}_1^{-i}$ are the same for both systems since in both systems,

$$U_1^{k,j} = \mu_1^{k,j}(X_1^{k,j}) \quad \forall (k,j) \in \mathcal{N}\setminus\mathcal{N}_i.$$

We conclude that $\mathbf{U}_1$ are the same for both systems. Since $(W_1^{k,Y})_{k\in\mathcal{I}}$ are the same for both systems, $Y_1^k = \ell_1^i(X_1^k, \mathbf{U}_1, W_1^{k,Y}), k \in \mathcal{I}$ are the same for both systems.

Induction Step: Suppose that $\mathbf{X}_s, \mathbf{Y}_s, \mathbf{U}_s$ are the same for both systems for all $s < t$. Now, we prove it for t.

First, since the realizations of $\mathbf{X}_{t-1}^i, \mathbf{U}_{t-1}, W_{t-1}^{i,X}$ are the same for both systems and

$$\mathbf{X}_t^k = f_t^k(\mathbf{X}_{t-1}^k, \mathbf{U}_{t-1}, W_{t-1}^{k,X}) \quad \forall k \in \mathcal{I},$$

$\mathbf{X}_t$ are the same for both systems.

Consider the actions taken by the members of team i at time t. For the first system,

$$U_t^{i,j} = \mu_t^{i,j}(H_t^{i,j}) = \mu_t^{i,j}(H_t^i, X_{t-d+1:t}^{i,j}) \quad \forall (i,j) \in \mathcal{N}_i.$$

In the second system,

$$\boldsymbol{\Gamma}_t^i = \nu_t^i(H_t^i) = (\mu_t^{i,j}(H_t^i, \cdot))_{(i,j)\in\mathcal{N}_i}$$

$$U_t^{i,j} = \Gamma_t^{i,j}(X_{t-d+1:t}^{i,j}) \quad \forall (i,j) \in \mathcal{N}_i,$$

which means that

$$U_t^{i,j} = \mu_t^{i,j}(H_t^i, X_{t-d+1:t}^{i,j}) \quad \forall (i,j) \in \mathcal{N}_i.$$

The actions taken by the members of other teams at time t are

$$U_t^{k,j} = \mu_t^{k,j}(H_t^k, X_{t-d+1:t}^{k,j}) \quad \forall (k,j) \in \mathcal{N} \backslash \mathcal{N}_i.$$

for both systems.

We conclude that $\mathbf{U}_t$ has the same realization for two systems since $(H_t^k, X_{t-d+1:t}^{k,j})_{k\in\mathcal{I}}$ have the same realization by the induction hypothesis and the argument above. Since $(W_t^{i,Y})_{i\in\mathcal{I}}$ are the same for both systems, $Y_t^k = \ell_t^k(X_t^k, \mathbf{U}_t, W_t^{k,Y}), k \in \mathcal{I}$ are same for both systems.

Therefore, we have established the induction step, proving that μ^i and ν^i generate the same realization of $(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{U}_t)_{t\in\mathcal{T}}$ under the same realization of the primitive random variables. Therefore, ν^i is a payoff-equivalent pure coordination strategy profile of μ^i.

To complete the other half of the proof, for each given coordination strategy ν^i of team/coordinator i we define a pure team strategy $\mu^i = (\mu_t^{i,j})_{(i,j)\in\mathcal{N}_i, t\in\mathcal{T}}$ through

$$\mu_t^{i,j}(h_t^{i,j}) = \gamma_t^{i,j}(x_{t-d+1:t}^{i,j}) \qquad \forall h_t^{i,j} \in \mathcal{H}_t^{i,j} \qquad \forall (i,j) \in \mathcal{N}_i,$$

where $\gamma_t^i = (\gamma_t^{i,j})_{(i,j)\in\mathcal{N}_i}$ is recursively defined by $\nu_{1:t}^i$ and h_t^i through

$$\gamma_t^i = \nu_t^i(h_t^i, \gamma_{1:t-1}^i) \quad \forall t \in \mathcal{T}.$$

Then using an argument similar to the one for the proof of the first half we can show that μ^i is payoff-equivalent to ν^i.

C Proof of Lemma 3

Induction on time t.

Induction Base: At $t = 1$, we have $\mathbf{X}_1^k$ to be independent for different k because of the assumption on primitive random variables. Furthermore, since H_1^k is a deterministic random vector (see Remark 2) and the randomization of different coordinators are independent, we conclude that $(\mathbf{X}_1^k, \boldsymbol{\Gamma}_1^k)$ are mutually independent for different k. The distribution of $(\mathbf{X}_1^k, \boldsymbol{\Gamma}_1^k)$ depends on g only through g^k.

Induction Step: Suppose that $(\mathbf{X}_{1:t}^k, \boldsymbol{\Gamma}_{1:t}^k)$ are conditionally independent given H_t^0 and $\mathbb{P}^g(\mathbf{X}_{1:t}^k, \boldsymbol{\Gamma}_{1:t}^k | H_t^0)$ depends on g only through g^k. Now, we have

$$\begin{aligned}
&\mathbb{P}^g(x_{1:t+1}, \gamma_{1:t+1} | h_{t+1}^0) \\
&= \mathbb{P}^g(x_{t+1} | h_{t+1}^0, x_{1:t}, \gamma_{1:t+1}) \mathbb{P}^g(\gamma_{t+1} | h_{t+1}^0, x_{1:t}, \gamma_{1:t}) \mathbb{P}^g(x_{1:t}, \gamma_{1:t} | h_{t+1}^0) \\
&= \left(\prod_{k\in\mathcal{I}} \mathbb{P}(x_{t+1}^k | x_t^k, u_t) g_{t+1}^k(\gamma_{t+1}^k | h_{t+1}^0, x_{1:t-d+1}^k, \gamma_{1:t}^k) \right) \mathbb{P}^g(x_{1:t}, \gamma_{1:t} | h_{t+1}^0).
\end{aligned}$$

We then claim that

$$\mathbb{P}^g(x_{1:t}, \gamma_{1:t}, y_t, u_t | h_t^0) = \prod_{k\in\mathcal{I}} F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0)$$

where for each $k \in \mathcal{I}$, F_t^k is a function that depends only on g^k.

To establish the claim, we note that

$$\begin{aligned}
&\mathbb{P}^g(x_{1:t}, \gamma_{1:t}, y_t, u_t | h_t^0) \\
&= \mathbb{P}^g(y_t, u_t | h_t^0, x_{1:t}, \gamma_{1:t}) \mathbb{P}^g(x_{1:t}, \gamma_{1:t} | h_t^0) \\
&= \left(\prod_{k \in \mathcal{I}} \mathbb{P}(y_t^k | x_t^k, u_t) \mathbf{1}_{\{u_t^k = \gamma_t^k(x_{t-d+1:t}^k)\}} \right) \mathbb{P}^g(x_{1:t}, \gamma_{1:t} | h_t^0) \\
&= \left(\prod_{k \in \mathcal{I}} \mathbb{P}(y_t^k | x_t^k, u_t) \mathbf{1}_{\{u_t^k = \gamma_t^k(x_{t-d+1:t}^k)\}} \right) \left(\prod_{k \in \mathcal{I}} \mathbb{P}^{g_k}(x_{1:t}^k, \gamma_{1:t}^k | h_t^0) \right) \\
&= \prod_{k \in \mathcal{I}} F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0),
\end{aligned}$$

where in the third step we have used the induction hypothesis.

Given the claim, we have

$$\begin{aligned}
\mathbb{P}^g(x_{1:t}, \gamma_{1:t} | h_{t+1}^0) &= \frac{\mathbb{P}^g(x_{1:t}, \gamma_{1:t}, y_t, u_t | h_t^0)}{\sum_{\tilde{x}_{1:t}, \tilde{\gamma}_{1:t}} \mathbb{P}^g(\tilde{x}_{1:t}, \tilde{\gamma}_{1:t}, y_t, u_t | h_t^0)} \\
&= \frac{\prod_{k \in \mathcal{I}} F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0)}{\sum_{\tilde{x}_{1:t}, \tilde{\gamma}_{1:t}} \prod_{k \in \mathcal{I}} F_t^k(\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k, h_{t+1}^0)} \\
&= \frac{\prod_{k \in \mathcal{I}} F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0)}{\prod_{k \in \mathcal{I}} \left(\sum_{\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k} F_t^k(\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k, h_{t+1}^0) \right)} \\
&= \prod_{k \in \mathcal{I}} \left(\frac{F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0)}{\sum_{\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k} F_t^k(\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k, h_{t+1}^0)} \right)
\end{aligned}$$

and then

$$\mathbb{P}^g(x_{1:t+1}, \gamma_{1:t+1} | h_{t+1}^0) = \prod_{k \in \mathcal{I}} G_t^k(x_{1:t+1}^k, \gamma_{1:t+1}^k, h_{t+1}^0),$$

where G_t^k is given by

$$\begin{aligned}
G_t^k(x_{1:t+1}^k, \gamma_{1:t+1}^k, h_{t+1}^0) =& \mathbb{P}(x_{t+1}^k | x_t^k, u_t) g_{t+1}^k(\gamma_{t+1}^k | h_{t+1}^0, x_{1:t-d+1}^k, \gamma_{1:t}^k) \times \\
& \times \frac{F_t^k(x_{1:t}^k, \gamma_{1:t}^k, h_{t+1}^0)}{\sum_{\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k} F_t^k(\tilde{x}_{1:t}^k, \tilde{\gamma}_{1:t}^k, h_{t+1}^0)}.
\end{aligned}$$

One can check that G_t^k depends on g only through g^k and

$$\sum_{\tilde{x}_{1:t+1}^k, \tilde{\gamma}_{1:t+1}^k} G_t^k(\tilde{x}_{1:t+1}^k, \tilde{\gamma}_{1:t+1}^k, h_{t+1}^0) = 1,$$

therefore

$$G_t^k(x_{1:t+1}^k, \gamma_{1:t+1}^k, h_{t+1}^0) = \mathbb{P}^{g^k}(x_{1:t+1}^k, \gamma_{1:t+1}^k | h_{t+1}^0).$$

Hence, we establish the induction step.

D Proof of Lemma 4

Assume that $\overline{h}_t^i \in \overline{\mathcal{H}}_t^i$ is admissible under g. From Lemma 3, we know that $\mathbb{P}^g(x_{1:t}^i, \gamma_{1:t}^i | h_t^0)$ does not depend on g^{-i}. As a conditional distribution obtained from $\mathbb{P}^g(x_{1:t}^i, \gamma_{1:t}^i | h_t^0)$, $\mathbb{P}^g(x_{t-d+1:t}^i | \overline{h}_t^i)$ does not depend on g^{-i} either.

Therefore, we can compute the belief of coordinator i by replacing g^{-i} with $\hat{g}^{-i}$, which is an open-loop strategy profile that always generates the actions $u_{1:t-1}^{-i}$.

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

Note that we always have $\mathbb{P}^{g^i, \hat{g}^{-i}}(\overline{h}_t^i) > 0$ for all $\overline{h}_t^i$ admissible under g.

Furthermore, we can also introduce additional random variables into the condition that are conditionally independent according to Lemma 3, i.e.,

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

where $x_{t-d:t}^{-i} \in \mathcal{X}_{t-d:t}^{-i}$ is such that $\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{t-d:t}^{-i} | \overline{h}_t^i) > 0$.

Let $\tau = t - d + 1$. By Bayes' rule

$$\begin{aligned} &\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{\tau:t}^i | \overline{h}_t^i, x_{\tau-1:t}^{-i}) \\ &= \frac{\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{\tau:t}, y_{\tau:t-1}, u_{\tau:t-1}, \gamma_{\tau:t-1}^i | h_\tau^{*i})}{\sum_{\tilde{x}_{\tau:t}^i} \mathbb{P}^{g^i, \hat{g}^{-i}}(\tilde{x}_{\tau:t}^i, x_{\tau:t}^{-i}, y_{\tau:t-1}, u_{\tau:t-1}, \gamma_{\tau:t-1}^i | h_\tau^{*i})}, \end{aligned} \tag{10}$$

where

$$h_\tau^{*i} = (y_{1:\tau-1}, u_{1:\tau-1}, x_{1:\tau-1}^i, x_{\tau-1}^{-i}, \gamma_{1:\tau-1}^i).$$

We have

$$\begin{aligned} &\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{\tau:t}, y_{\tau:t-1}, u_{\tau:t-1}, \gamma_{\tau:t-1}^i | h_\tau^{*i}) \\ &= \prod_{l=1}^{d-1} \Big[\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{t-l+1}, y_{t-l} | h_\tau^{*i}, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l}, \gamma_{\tau:t-l}^i) \\ &\quad \times \mathbb{P}^{g^i, \hat{g}^{-i}}(u_{t-l}^i | h_\tau^{*i}, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l-1}, \gamma_{\tau:t-l}^i) \\ &\quad \times \mathbb{P}^{g^i, \hat{g}^{-i}}(\gamma_{t-l}^i | h_\tau^{*i}, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l-1}, \gamma_{\tau:t-l-1}^i) \Big] \\ &\quad \times \mathbb{P}^{g^i, \hat{g}^{-i}}(x_\tau | h_\tau^{*i}). \end{aligned} \tag{11}$$

The first three terms in the above product are

$$\begin{aligned} &\mathbb{P}^{g^i, \hat{g}^{-i}}(x_{t-l+1}, y_{t-l} | h_\tau^{*i}, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l}, \gamma_{\tau:t-l}^i) \\ &= \prod_{k \in \mathcal{I}} [\mathbb{P}(x_{t-l+1}^k | x_{t-l}^k, u_{t-l}) \mathbb{P}(y_{t-l}^k | x_{t-l}^k, u_{t-l})], \\ &\mathbb{P}^{g^i, \hat{g}^{-i}}(u_{t-l}^i | h_\tau^{*i}, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l-1}, \gamma_{\tau:t-l}^i) \\ &= \prod_{(i,j) \in \mathcal{N}_i} \mathbf{1}_{\{u_{t-l}^{i,j} = \gamma_{t-l}^{i,j}(x_{t-l-d+1:t-l}^{i,j})\}} \\ &= \prod_{(i,j) \in \mathcal{N}_i} \mathbf{1}_{\{u_{t-l}^{i,j} = \phi_{t-l,l}^i(x_{\tau:t-l}^i)\}}, \end{aligned}$$

$$\begin{aligned}&\mathbb{P}^{g^i,\hat{g}^{-i}}(\gamma^i_{t-l}|h^{*i}_\tau, x_{\tau:t-l}, y_{\tau:t-l-1}, u_{\tau:t-l-1}, \gamma^i_{\tau:t-l-1})\\ &= g^i_{t-l}(\gamma^i_{t-l}|y_{1:t-l-1}, u_{1:t-l-1}, x^i_{1:t-d-l}, \gamma^i_{1:t-l-1}),\end{aligned} \tag{12}$$

respectively.

The last term satisfies

$$\mathbb{P}^{g^i,\hat{g}^{-i}}(x_\tau|h^{*i}_\tau) = \prod_{k\in\mathcal{I}} \mathbb{P}(x^k_\tau|x^k_{\tau-1}, u_{\tau-1}).$$

Substituting (11) - (12) into (10), we obtain

$$\mathbb{P}^{g^i,\hat{g}^{-i}}(x^i_{\tau:t}|\overline{h}^i_t, x^{-i}_{\tau-1:t}) = \frac{F^i_t(x^i_{\tau:t}, y^i_{\tau:t-1}, u_{\tau-1:t-1}, x^i_{\tau-1}, \phi^i_t)}{\sum_{\tilde{x}^i_{\tau:t}} F^i_t(\tilde{x}^i_{\tau:t}, y^i_{\tau:t-1}, u_{\tau-1:t-1}, x^i_{\tau-1}, \phi^i_t)}$$

where

$$\begin{aligned}F^i_t(x^i_{\tau:t}, y^i_{\tau:t-1}, u_{\tau-1:t-1}, \phi^i_t) &:= \mathbb{P}(x^i_\tau|x^i_{\tau-1}, u_{\tau-1})\\ &\times \prod_{l=1}^{d-1}\left(\mathbb{P}(x^i_{t-l+1}|x^i_{t-l}, u_{t-l})\mathbb{P}(y^i_{t-l}|x^i_{t-l}, u_{t-l}) \prod_{(i,j)\in\mathcal{N}_i} \mathbf{1}_{\{u^{i,j}_{t-l}=\phi^{i,j}_{t-l,l}(x^{i,j}_{\tau:t-l})\}}\right)\end{aligned}$$

Therefore, we have proved that

$$\begin{aligned}\mathbb{P}^g(x^i_{t-d+1:t}|\overline{h}^i_t) &= P^i_t(x^i_{t-d+1:t}|y^i_{t-d+1:t-1}, u_{t-d:t-1}, x^i_{t-d}, \phi^i_t)\\ &:= \frac{F^i_t(x^i_{t-d+1:t}, y^i_{t-d+1:t-1}, u_{t-d:t-1}, x^i_{t-d}, \phi^i_t)}{\sum_{\tilde{x}^i_{t-d+1:t}} F^i_t(\tilde{x}^i_{t-d+1:t}, y^i_{t-d+1:t-1}, u_{t-d:t-1}, x^i_{t-d}, \phi^i_t)}\end{aligned}$$

where P^i_t is independent of g.

E Proof of Lemma 5

For notational convenience, define

$$\overline{H}^{-i}_t = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}^{-i}_{1:t-d}, \boldsymbol{\Gamma}^{-i}_{1:t-1})$$

Claim 1 $\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, s^i_t, \overline{h}^{-i}_t)$ does not depend on g.[9]

Claim 2 $\mathbb{P}^g(\gamma_t, s^i_t, \overline{h}^{-i}_t) = \mathbb{P}^{\rho^i,g^{-i}}(\gamma_t, s^i_t, \overline{h}^{-i}_t)$ for all $\gamma_t \in \mathcal{A}_t, s^i_t \in \mathcal{S}^i_t, \overline{h}^{-i}_t \in \mathcal{H}^{-i}_t$.

Given Claims 1 and 2, we conclude that

$$\mathbb{P}^g(x_{t-d+1:t}, \gamma_t, s^i_t, \overline{h}^{-i}_t) = \mathbb{P}^{\rho^i,g^{-i}}(x_{t-d+1:t}, \gamma_t, s^i_t, \overline{h}^{-i}_t) \tag{13}$$

for all $x_{t-d+1:t} \in \mathcal{X}_{t-d+1:t}, \gamma_t \in \mathcal{A}_t, s^i_t \in \mathcal{S}^i_t, \overline{h}^{-i}_t \in \mathcal{H}^{-i}_t$.

[9] We claim that the value of this conditional probability is the same for g and $\tilde{g}$ whenever the conditional probability is well-defined under both g and $\tilde{g}$. However, whether or not the the conditional probability is well-defined does depend on g. In the lemma, we always apply Claim 1 by multiplying $\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, s^i_t, \overline{h}^{-i}_t)$ with some other terms. Those terms will be 0 whenever $\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, s^i_t, \overline{h}^{-i}_t)$ is not well defined.

Marginalizing (13), we obtain

$$\mathbb{P}^g(x_{t-d+1:t}, \gamma_t) = \mathbb{P}^{\rho^i, g^{-i}}(x_{t-d+1:t}, \gamma_t). \tag{14}$$

Since $U_t^{k,j} = \Gamma_t^{k,j}(X_{t-d+1:t}^{k,j})$ for all $(k, j) \in \mathcal{N}$, we can write $r_t^k(\mathbf{X}_t, \mathbf{U}_t) = \tilde{r}_t^k(\mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t)$ for some function $\tilde{r}_t^k$ that does not depend on the strategy profile. Then, using linearity of expectation and (14) we obtain

$$\begin{aligned} J^k(g) &= \mathbb{E}^g\left[\sum_{t\in\mathcal{T}} \tilde{r}_t^k(\mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t)\right] \\ &= \sum_{t\in\mathcal{T}} \mathbb{E}^g[\tilde{r}_t^k(\mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t)] = \sum_{t\in\mathcal{T}} \mathbb{E}^{\rho^i, g^{-i}}[\tilde{r}_t^k(\mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t)] \\ &= \mathbb{E}^{\rho^i, g^{-i}}\left[\sum_{t\in\mathcal{T}} \tilde{r}_t^k(\mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t)\right] = J^k(\rho^i, g^{-i}) \end{aligned}$$

for all behavioral coordination strategy profile g^{-i}. Hence, g^i and ρ^i are payoff-equivalent.

Proof of Claim 1 For notational convenience, define

$$\overline{H}_t = \bigcup_{i\in\mathcal{I}} \overline{H}_t^i = (\mathbf{Y}_{1:t-1}, \mathbf{U}_{1:t-1}, \mathbf{X}_{1:t-d}, \boldsymbol{\Gamma}_{1:t-1}).$$

Consider $\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, \overline{h}_t)$ first. Since $\boldsymbol{\Gamma}_t$ is a randomized prescription generated based on $\overline{H}_t$ which enters the system after $\mathbf{X}_{t-d+1:t}$ are realized, we have

$$\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, \overline{h}_t) = \mathbb{P}^g(x_{t-d+1:t}|\overline{h}_t). \tag{15}$$

Due to Lemma 3, we have

$$\mathbb{P}^g(x_{t-d+1:t}|\overline{h}_t) = \prod_{k\in\mathcal{I}} \mathbb{P}^g(x_{t-d+1:t}^k|\overline{h}_t^k). \tag{16}$$

By Lemma 4,

$$\mathbb{P}^g(x_{t-d+1:t}^k|\overline{h}_t^k) = P_t^k(x_{t-d+1:t}^k|y_{t-d+1:t-1}^k, u_{t-d+1:t-1}, s_t^i) \tag{17}$$

where P_t^k is a function that does not depend on g.

Combining (15), (16), (17), we have

$$\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, \overline{h}_t) = \prod_{k\in\mathcal{I}} P_t^k(x_{t-d+1:t}^k|y_{t-d+1:t-1}^k, u_{t-d+1:t-1}, s_t^k).$$

Since $(S_t^i, \overline{H}_t^{-i})$ is a function of $\overline{H}_t$, by Smoothing Property of Conditional Probability we conclude that

$$\mathbb{P}^g(x_{t-d+1:t}|\gamma_t, s_t^i, \overline{h}_t^{-i}) = \prod_{k\in\mathcal{I}} P_t^k(x_{t-d+1:t}^k|y_{t-d+1:t-1}^k, u_{t-d+1:t-1}, s_t^k)$$

where the right-hand side does not depend on g. □

Proof of Claim 2 Proof by induction on t.

Induction Base: The claim is true at $t = 1$ since ρ_1^i and g_1^i are the same strategies.

Induction Step: Suppose that the claim is true for time $t - 1$. Prove the result for t.

First,

$$
\begin{aligned}
&\mathbb{P}^g(\gamma_t, s_t^i, \overline{h}_t^{-i}) \\
&= \sum_{(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i):\tilde{s}_t^i = s_t^i} \mathbb{P}^g(\gamma_t | s_t^i, \overline{h}_t^{-i}, \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i) \mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i, s_t^i, \overline{h}_t^{-i}) \\
&= \sum_{(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i):\tilde{s}_t^i = s_t^i} g_t^i(\gamma_t^i | h_t^0, \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i) \left(\prod_{k \neq i} g_t^k(\gamma_t^k | \overline{h}_t^k) \right) \mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i, s_t^i, \overline{h}_t^{-i}) \\
&= \sum_{(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i):\tilde{s}_t^i = s_t^i} g_t^i(\gamma_t^i | h_t^0, \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i) \left(\prod_{k \neq i} g_t^k(\gamma_t^k | \overline{h}_t^k) \right) \mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | s_t^i, \overline{h}_t^{-i}) \\
&\quad \times \mathbb{P}^g(s_t^i, \overline{h}_t^{-i}) \\
&= \left(\sum_{(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i):\tilde{s}_t^i = s_t^i} g_t^i(\gamma_t^i | \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i, h_t^0) \mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | s_t^i, \overline{h}_t^{-i}) \right) \\
&\quad \times \left(\prod_{k \neq i} g_t^k(\gamma_t^k | \overline{h}_t^k) \right) \mathbb{P}^g(s_t^i, \overline{h}_t^{-i}).
\end{aligned}
$$

Using Lemma 3, we have

$$
\mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | s_t^i, \overline{h}_t^{-i}) = \mathbb{P}^{g^i}(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | s_t^i, h_t^0).
$$

Therefore,

$$
\sum_{(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i):\tilde{s}_t^i = s_t^i} g_t^i(\gamma_t^i | \tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i, h_t^0) \mathbb{P}^g(\tilde{x}_{1:t-d}^i, \tilde{\gamma}_{1:t-1}^i | s_t^i, \overline{h}_t^{-i}) = \rho_t^i(\gamma_t^i | h_t^0, s_t^i)
$$

for all (h_t^0, s_t^i) admissible under g. Notice that $\mathbb{P}^g(h_t^0, s_t^i) = 0$ implies that $\mathbb{P}^g(s_t^i, \overline{h}_t^{-i}) = 0$, hence we conclude that

$$
\mathbb{P}^g(\gamma_t, s_t^i, \overline{h}_t^{-i}) = \rho_t^i(\gamma_t^i | h_t^0, s_t^i) \left(\prod_{k \neq i} g_t^k(\gamma_t^k | \overline{h}_t^k) \right) \mathbb{P}^g(s_t^i, \overline{h}_t^{-i})
$$

for all $\gamma_t \in \mathcal{A}_t, s_t^i \in \mathcal{S}_t^i, \overline{h}_t^{-i} \in \mathcal{H}_t^{-i}$.

Similarly,

$$
\mathbb{P}^{\rho^i, g^{-i}}(\gamma_t, s_t^i, \overline{h}_t^{-i}) = \rho_t^i(\gamma_t^i | h_t^0, s_t^i) \left(\prod_{k \neq i} g_t^k(\gamma_t^k | \overline{h}_t^k) \right) \mathbb{P}^{\rho^i, g^{-i}}(s_t^i, \overline{h}_t^{-i})
$$

for all $\gamma_t \in \mathcal{A}_t, s_t^i \in \mathcal{S}_t^i, \overline{h}_t^{-i} \in \mathcal{H}_t^{-i}$. Hence, it suffices to show that $\mathbb{P}^g(s_t^i, \overline{h}_t^{-i}) = \mathbb{P}^{\rho^i, g^{-i}}(s_t^i, \overline{h}_t^{-i})$.

Given the induction hypothesis, it suffices to show that

$$
\mathbb{P}^g(s_t^i, \overline{h}_t^{-i} | \gamma_{t-1}, s_{t-1}^i, \overline{h}_{t-1}^{-i}) = \mathbb{P}^{\rho^i, g^{-i}}(s_t^i, \overline{h}_t^{-i} | \gamma_{t-1}, s_{t-1}^i, \overline{h}_{t-1}^{-i})
$$

for all $(\gamma_{t-1}, s^i_{t-1}, \overline{h}^{-i}_{t-1})$ admissible under g (or admissible under (ρ^i, g^{-i}), which is an equivalent condition because of the induction hypothesis).

Given that

$$\begin{aligned} S^i_t &= \iota^i_{t-1}(S^i_{t-1}, \mathbf{X}^i_{t-d}, \boldsymbol{\Gamma}^i_{t-1}), \\ \overline{H}^{-i}_t &= (\overline{H}^{-i}_{t-1}, \mathbf{Y}_{t-1}, \mathbf{U}_{t-1}, \mathbf{X}^{-i}_{t-d}, \boldsymbol{\Gamma}^{-i}_{t-1}), \\ Y^k_{t-1} &= \ell^k_{t-1}(\mathbf{X}^k_{t-1}, \mathbf{U}^k_{t-1}, W^{k,Y}_{t-1}) \quad \forall k \in \mathcal{I}, \\ U^{k,j}_{t-1} &= \Gamma^{k,j}_{t-1}(X^{k,j}_{t-d:t-1}) \quad \forall (k,j) \in \mathcal{N}, \end{aligned}$$

it follows that $(S^i_t, \overline{H}^{-i}_t)$ is a strategy-independent function of $(\boldsymbol{\Gamma}_{t-1}, S^i_{t-1}, \overline{H}^{-i}_{t-1}, \mathbf{X}_{t-d:t-1}, \mathbf{W}^Y_{t-1})$. Since $\mathbf{W}^Y_{t-1} = (W^{k,Y}_{t-1})_{k\in\mathcal{I}}$ is a primitive random vector independent of $(\boldsymbol{\Gamma}_{t-1}, S^i_{t-1}, \overline{H}^{-i}_{t-1})$, it suffices to show that

$$\mathbb{P}^g(x_{t-d:t-1}|\gamma_{t-1}, s^i_{t-1}, \overline{h}^{-i}_{t-1}) = \mathbb{P}^{\rho^i, g^{-i}}(x_{t-d:t-1}|\gamma_{t-1}, s^i_{t-1}, \overline{h}^{-i}_{t-1}). \tag{18}$$

We know that (18) is true due to Claim 1. Hence, we established the induction step. □

Remark 13 In general, a behavioral coordination strategy profile yields different distributions on the trajectory of the system in comparison with the distributions generated from its associated SPIB strategy profile. It is the equivalence of marginal distributions that allows us to establish the equivalence of payoffs using linearity of expectation. This (payoff) equivalence between behavioral coordination strategies and their associated SPIB strategy profiles is different from the equivalence of behavioral coordination strategies with mixed team strategies where not only are the payoffs equivalent, but distributions on the trajectory of the system are also the same.

F Proof of Lemma 6

We will prove a stronger result.

Lemma 9 *Let (λ^{*k}, ψ^*) be a CIB strategy such that $\psi^{*,k}$ is consistent with λ^{*k}. Let g^{*k} be the behavioral strategy profile generated from (λ^{*k}, ψ^*). Let π^k_t represent the belief on S^k_t generated by ψ^* at time t based on h^0_t. Let $t < \tau$. Consider a fixed $h^0_\tau \in \mathcal{H}^0_\tau$ and some $\tilde{g}^k_{1:t-1}$ (not necessarily equal to $g^{*k}_{1:t-1}$). Assume that h^0_τ is admissible under $(\tilde{g}^k_{1:t-1}, g^{*k}_{t:\tau-1})$. Suppose that*

$$\begin{aligned} \mathbb{P}^{\tilde{g}^k_{1:t-1}}(s^k_t, x^k_{t-d+1:t}|h^0_t) =& \pi^k_t(s^k_t) P^k_t(x^k_{t-d+1:t}|y^k_{t-d+1:t-1}, u_{t-d:t-1}, s^k_t) \\ & \forall s^k_t \in \mathcal{S}^k_t \ \forall x^k_{t-d+1:t} \in \mathcal{X}^k_{t-d+1:t}. \end{aligned} \tag{19}$$

Then,

$$\begin{aligned} \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_{t:\tau-1}}(s^k_\tau, x^k_{\tau-d+1:\tau}|h^0_\tau) =& \pi^k_\tau(s^k_\tau) P^k_\tau(x^k_{\tau-d+1:\tau}|y^k_{\tau-d+1:\tau-1}, u_{\tau-d:\tau-1}, s^k_\tau) \\ & \forall s^k_\tau \in \mathcal{S}^k_\tau \ \forall x^k_{\tau-d+1:\tau} \in \mathcal{X}^k_{\tau-d+1:\tau}. \end{aligned}$$

The assertion of Lemma 6 follows from Lemma 9 and the fact that (19) is true for $t = 1$.

Proof of Lemma 9 We only need to prove the result for $\tau = t + 1$.

Since h^0_{t+1} is admissible under $(\tilde{g}^k_{1:t-1}, g^{*k}_t)$, we have

$$\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(h^0_{t+1}) > 0 \tag{20}$$

where $\hat{g}^{-k}_{1:t}$ is the open-loop strategy where all coordinators except k choose prescriptions that generate the actions $u^{-k}_{1:t}$.

From Lemma 3, we know that $\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, g^{-k}}(s^k_{t+1}|h^0_{t+1})$ is independent of g^{-k}. Therefore,

$$\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t}(s^k_{t+1}|h^0_{t+1}) = \frac{\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(s^k_{t+1}, y_t, u_t|h^0_t)}{\sum_{\tilde{s}^k_{t+1}} \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{s}^k_{t+1}, y_t, u_t|h^0_t)}, \tag{21}$$

and the denominator of (21) is nonzero due to (20).

We have

$$\begin{aligned}
&\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(s^k_{t+1}, y_t, u_t|h^0_t) \\
&= \sum_{\tilde{s}^k_t} \sum_{\tilde{x}^k_{t-d+1:t}} \sum_{\tilde{x}^{-k}_t} \sum_{\tilde{\gamma}^k_t : \tilde{\gamma}^k_t(\tilde{x}^k_{t-d+1:t}) = u^k_t} \Big[\mathbb{P}(y^k_t|\tilde{x}^k_t, u_t) \mathbb{P}(y^{-k}_t|\tilde{x}^{-k}_t, u_t) \\
&\quad \times \mathbf{1}_{\{s^k_{t+1} = \iota^k_t(\tilde{s}^k_t, \tilde{x}^k_{t-d+1}, \tilde{\gamma}^k_t)\}} \lambda^{*k}_t(\tilde{\gamma}^k_t|b_t, \tilde{s}^k_t) \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^k_{t-d+1:t}, \tilde{x}^{-k}_t, \tilde{s}^k_t|h^0_t) \Big] \\
&= \sum_{\tilde{s}^k_t} \sum_{\tilde{x}^k_{t-d+1:t}} \sum_{\tilde{x}^{-k}_t} \sum_{\tilde{\gamma}^k_t : \tilde{\gamma}^k_t(\tilde{x}^k_{t-d+1:t}) = u^k_t} \Big[\mathbb{P}(y^k_t|\tilde{x}^k_t, u_t) \mathbb{P}(y^{-k}_t|\tilde{x}^{-k}_t, u_t) \\
&\quad \times \mathbf{1}_{\{s^k_{t+1} = \iota^k_t(\tilde{s}^k_t, \tilde{x}^k_{t-d+1}, \tilde{\gamma}^k_t)\}} \lambda^{*k}_t(\tilde{\gamma}^k_t|b_t, \tilde{s}^k_t) \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^k_{t-d+1:t}, \tilde{s}^k_t|h^0_t) \\
&\quad \times \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^{-k}_t|h^0_t) \Big] \\
&= \left(\sum_{\tilde{x}^{-k}_t} \mathbb{P}(y^{-k}_t|\tilde{x}^{-k}_t, u_t) \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^{-k}_t|h^0_t) \right) \\
&\quad \times \sum_{\tilde{s}^k_t} \sum_{\tilde{x}^k_{t-d+1:t}} \sum_{\tilde{\gamma}^k_t : \tilde{\gamma}^k_t(\tilde{x}^k_{t-d+1:t}) = u^k_t} \Big[\mathbb{P}(y^k_t|\tilde{x}^k_t, u_t) \mathbf{1}_{\{s^k_{t+1} = \iota^k_t(\tilde{s}^k_t, \tilde{x}^k_{t-d+1}, \tilde{\gamma}^k_t)\}} \\
&\quad \times \lambda^{*k}_t(\tilde{\gamma}^k_t|b_t, \tilde{s}^k_t) \mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^k_{t-d+1:t}, \tilde{s}^k_t|h^0_t) \Big].
\end{aligned} \tag{22}$$

where $b_t = (\boldsymbol{\pi}_t, y_{t-d+1:t-1}, u_{t-d:t-1})$ and $\boldsymbol{\pi}_t = (\pi^l_t)_{l \in \mathcal{I}}$ is generated from ψ^*.

Recall that we assume

$$\begin{aligned}
&\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t, \hat{g}^{-k}_{1:t}}(\tilde{x}^k_{t-d+1:t}, \tilde{s}^k_t|h^0_t) \\
&= \pi^k_t(\tilde{s}^k_t) P^k_t(\tilde{x}^k_{t-d+1:t}|y^k_{t-d+1:t-1}, u_{t-d:t-1}, \tilde{s}^k_t).
\end{aligned} \tag{23}$$

Using (21), (22), and (23), we obtain

$$\mathbb{P}^{\tilde{g}^k_{1:t-1}, g^{*k}_t}(s^k_{t+1}|h^0_{t+1}) = \frac{\Upsilon^k_t(b_t, y^k_t, u_t, s^k_{t+1})}{\sum_{\tilde{s}^k_{t+1}} \Upsilon^k_t(b_t, y^k_t, u_t, \tilde{s}^k_{t+1})}$$

where

$$\Upsilon^k_t(b_t, y^k_t, u_t, s^k_{t+1})$$

$$= \sum_{\tilde{s}_t^k} \sum_{\tilde{x}_{t-d+1:t}^k} \sum_{\tilde{\gamma}_t^k : \tilde{\gamma}_t^k(\tilde{x}_{t-d+1:t}^k) = u_t^k} \Big[\mathbb{P}(y_t^k | \tilde{x}_t^k, u_t) \mathbf{1}_{\{s_{t+1}^k = l_t^k(\tilde{s}_t^k, \tilde{x}_{t-d+1}^k, \tilde{\gamma}_t^k)\}}$$
$$\times \lambda_t^{*k}(\tilde{\gamma}_t^k | b_t, \tilde{s}_t^k) \pi_t^k(\tilde{s}_t^k) P_t^k(\tilde{x}_{t-d+1:t}^k | y_{t-d+1:t-1}^k, u_{t-d:t-1}, \tilde{s}_t^k) \Big],$$

Therefore by the definition of consistency of $\psi^{*,k}$ with respect to λ^{*k}, we conclude that

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(s_{t+1}^k | h_{t+1}^0) = \pi_{t+1}^k(s_{t+1}^k).$$

Now, consider $\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+2:t+1}^k, s_{t+1}^k | h_{t+1}^0)$.

– If $\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(s_{t+1}^k | h_{t+1}^0) = 0$, then we have $\pi_{t+1}^k(s_{t+1}^k) = 0$ and

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+2:t+1}^k, s_{t+1}^k | h_{t+1}^0) = 0.$$

– If $\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(s_{t+1}^k | h_{t+1}^0) > 0$, then

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+2:t+1}^k, s_{t+1}^k | h_{t+1}^0)$$
$$= \mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+1:t}^k | h_{t+1}^0, s_{t+1}^k) \pi_{t+1}^k(s_{t+1}^k).$$

We have shown in Lemma 4 that

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+2:t+1}^k | \overline{h}_{t+1}^k)$$
$$= P_{t+1}^k(\tilde{x}_{t-d+2:t+1}^k | y_{t-d+2:t}^k, u_{t-d+1:t}, s_{t+1}^k)$$

and (h_{t+1}^0, s_{t+1}^k) is a function of $\overline{h}_{t+1}^k$. By the law of iterated expectation, we have

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}, \hat{g}_{1:t}^{-k}}(\tilde{x}_{t-d+2:t+1}^k | h_{t+1}^0, s_{t+1}^k)$$
$$= P_{t+1}^k(\tilde{x}_{t-d+2:t+1}^k | y_{t-d+2:t}^k, u_{t-d+1:t}, s_{t+1}^k).$$

We conclude that

$$\mathbb{P}^{\tilde{g}_{1:t-1}^k, g_t^{*k}}(\tilde{x}_{t-d+2:t+1}^k, s_{t+1}^k | h_{t+1}^0)$$
$$= P_t^k(\tilde{x}_{t-d+2:t+1}^k | y_{t-d+2:t}^k, u_{t-d+1:t}, s_{t+1}^k) \pi_{t+1}^k(s_{t+1}^k)$$

for all $s_{t+1}^k \in \mathcal{S}_{t+1}^k$ and all $x_{t-d+2:t+1}^k \in \mathcal{X}_{t-d+2:t+1}^k$. □

G Proof of Lemma 7

Let g^{-i} denote the behavioral strategy profile of all coordinators other than i generated from the CIB strategy profile $(\lambda^k, \psi^k)_{k \in \mathcal{I} \setminus \{i\}}$. Let $(\overline{h}_t^i, \gamma_t^i)$ be admissible under g^{-i}.

Let $\tilde{g}^i$ denote coordinator i's behavioral coordination strategy. Because of Lemma 3, we have

$$\mathbb{P}^{\tilde{g}^i, g^{-i}}(x_{t-d+1:t}, \gamma_t^{-i} | \overline{h}_t^i, \gamma_t^i)$$
$$= \mathbb{P}^{\tilde{g}^i, g^{-i}}(x_{t-d+1:t}, \gamma_t^{-i} | h_t^0, x_{1:t-d}^i, \gamma_{1:t}^i)$$
$$= \mathbb{P}^{\tilde{g}^i}(x_{t-d+1:t}^i | h_t^0, x_{1:t-d}^i, \gamma_{1:t}^i) \prod_{k \neq i} \mathbb{P}^{g^k}(x_{t-d+1:t}^k, \gamma_t^k | h_t^0).$$

We know that $\boldsymbol{\Gamma}_t^i$ and $\mathbf{X}_{t-d+1:t}^i$ are conditionally independent given $\overline{H}_t^i$ since $\boldsymbol{\Gamma}_t^i$ is chosen as a randomized function of $\overline{H}_t^i$ at a time when $\mathbf{X}_{t-d+1:t}^i$ are already realized. Therefore,

$$\begin{aligned}\mathbb{P}^{\tilde{g}^i,g^{-i}}(x_{t-d+1:t}^i|h_t^0,x_{1:t-d}^i,\gamma_{1:t}^i)&=\mathbb{P}^{\tilde{g}^i,g^{-i}}(x_{t-d+1:t}^i|h_t^0,x_{1:t-d}^i,\gamma_{1:t-1}^i)\\&=P_t^i(x_{t-d:t}^i|y_{t-d+1:t-1}^i,u_{t-d:t-1},s_t^i),\end{aligned}$$

where $s_t^i=(x_{t-d}^i,\phi_t^i)$ and P_t^i is the belief function defined in Eq. (1).

We conclude that

$$\begin{aligned}&\mathbb{P}^{g^{-i}}(x_{t-d+1:t},\gamma_t^{-i}|\overline{h}_t^i,\gamma_t^i)\\&=P_t^i(x_{t-d:t}^i|y_{t-d+1:t-1}^i,u_{t-d:t-1},s_t^i)\prod_{k\neq i}\mathbb{P}^{g^k}(x_{t-d+1:t}^k,\gamma_t^k|h_t^0).\end{aligned}\tag{24}$$

Since all coordinators other than coordinator i are using the same belief generation systems, we have $B_t^j=B_t^k$ for $j,k\neq i$. Denote $B_t=B_t^k$ for all $k\in\mathcal{I}\backslash\{i\}$. Let $b_t=\left(\left(\pi_t^{*,l}\right)_{l\in\mathcal{I}},y_{t-d+1:t-1},u_{t-d:t-1}\right)$ be a realization of B_t. Also define $\psi^*=\psi^k$ for all $k\neq i$.

Consider $k\neq i$. Coordinator k's strategy g^k is a self-consistent CIB strategy. We also have h_t^0 admissible under g^k since $(\overline{h}_t^i,\gamma_t^i)$ is admissible under g^{-i}. Hence, applying Lemma 6 we have

$$\mathbb{P}^{g^k}(\tilde{s}_t^k,x_{t-d+1:t}^k|h_t^0)=\pi_t^{*,k}(\tilde{s}_t^k)P_t^k(x_{t-d+1:t}^k|y_{t-d+1:t-1}^k,u_{t-d:t-1},\tilde{s}_t^k).$$

Hence, the second term of the right hand side of (24) satisfies

$$\begin{aligned}&\mathbb{P}^{g^k}(x_{t-d+1:t}^k,\gamma_t^k|h_t^0)=\sum_{\tilde{s}_t^k}\mathbb{P}^{g^k}(\tilde{s}_t^k,x_{t-d+1:t}^k,\gamma_t^k|h_t^0)\\&=\sum_{\tilde{s}_t^k}\left[\pi_t^{*,k}(\tilde{s}_t^k)P_t^k(x_{t-d+1:t}^k|y_{t-d+1:t-1}^k,u_{t-d:t-1},\tilde{s}_t^k)\lambda_t^k(\gamma_t^k|b_t,\tilde{s}_t^k)\right],\end{aligned}\tag{25}$$

where P_t^k is the belief function defined in Eq. (1).

Recall that $b_t=\left(\left(\pi_t^{*,l}\right)_{l\in\mathcal{I}},y_{t-d+1:t-1},u_{t-d:t-1}\right)$. From (24) and (25), we conclude that

$$\mathbb{P}^{g^{-i}}(x_{t-d+1:t},\gamma_t^{-i}|\overline{h}_t^i,\gamma_t^i)=F_t^i(x_{t-d+1:t},\gamma_t^{-i}|b_t,s_t^i)\tag{26}$$

for some function F_t^i for all $(\overline{h}_t^i,\gamma_t^i)$ admissible under g^{-i}.

Consider the total reward of coordinator i. By the law of iterated expectation, we can write

$$J^i(\tilde{g}^i,g^{-i})=\mathbb{E}^{\tilde{g}^i,g^{-i}}\left[\sum_{t\in\mathcal{T}}\mathbb{E}^{g^{-i}}[r_t^i(\mathbf{X}_t,\mathbf{U}_t)|\overline{H}_t^i,\boldsymbol{\Gamma}_t^i]\right].$$

For $(\overline{h}_t^i,\gamma_t^i)$ admissible under g^{-i},

$$\begin{aligned}&\mathbb{E}^{g^{-i}}[r_t^i(\mathbf{X}_t,\mathbf{U}_t)|\overline{h}_t^i,\gamma_t^i]\\&=\sum_{\tilde{x}_{t-d+1:t}}\sum_{\tilde{\gamma}_t^{-i}}r_t^i(\tilde{x}_t,(\gamma_t^i(\tilde{x}_{t-d+1:t}^i),\tilde{\gamma}_t^{-i}(\tilde{x}_{t-d+1:t}^{-i})))F_t^i(\tilde{x}_{t-d+1:t},\tilde{\gamma}_t^{-i}|b_t,s_t^i)\\&=\overline{r}_t^i(b_t,s_t^i,\gamma_t^i),\end{aligned}$$

for some function $\overline{r}_t^i$ that depends on g^{-i} (specifically, on λ_t^{-i}) but not on $\tilde{g}^i$.

We claim that (B_t, S_t^i) is a controlled Markov process controlled by coordinator i's prescriptions, given that other coordinators are using the strategy profile g^{-i}. Let $\tilde{g}^i$ denote an arbitrary strategy for coordinator i (not necessarily a CIB strategy). We need to prove that

$$\mathbb{P}^{\tilde{g}^i, g^{-i}}(b_{t+1}, s_{t+1}^i | b_{1:t}, s_{1:t}^i, \gamma_{1:t}^i) = \Xi_t^i(b_{t+1}, s_t^i | b_t, s_t^i, \gamma_t^i) \quad \forall (b_{1:t}, s_{1:t}^i, \gamma_{1:t}^i) \text{ s.t. } \mathbb{P}^{\tilde{g}^i, g^{-i}}(b_{1:t}, s_{1:t}^i, \gamma_{1:t}^i) > 0$$

for some function Ξ_t^i independent of $\tilde{g}^i$.

We know that

$$\begin{aligned}
B_{t+1} &= (\boldsymbol{\Pi}_{t+1}, \mathbf{Y}_{t-d+2:t}, \mathbf{U}_{t-d+1:t}),\\
\boldsymbol{\Pi}_{t+1} &= \psi_t^*(B_t, \mathbf{Y}_t, \mathbf{U}_t),\\
Y_t^k &= \ell_t^k(\mathbf{X}_t^k, \mathbf{U}_t, W_t^{k,Y}) \quad \forall k \in \mathcal{I},\\
U_t^{k,j} &= \Gamma_t^{k,j}(X_{t-d+1:t}^{k,j}) \quad \forall (k,j) \in \mathcal{N},\\
S_{t+1}^i &= \iota_t^i(S_t^i, \mathbf{X}_{t-d+1}^i, \boldsymbol{\Gamma}_t^i).
\end{aligned}$$

Hence, (B_{t+1}, S_t^i) is a fixed function of $(B_t, S_t^i, \mathbf{X}_{t-d+1:t}, \boldsymbol{\Gamma}_t, \mathbf{W}_t^Y)$, where $\mathbf{W}_t^Y$ is a primitive random vector independent of $(B_{1:t}, S_{1:t}^i, \boldsymbol{\Gamma}_{1:t}^i, \mathbf{X}_{t-d+1:t})$. Therefore, it suffices to prove that

$$\mathbb{P}^{\tilde{g}^i, g^{-i}}(x_{t-d+1:t}, \gamma_t^{-i} | b_{1:t}, s_{1:t}^i, \gamma_{1:t}^i) = \Xi_t^i(x_{t-d+1:t}, \gamma_t^{-i} | b_t, s_t^i, \gamma_t^i)$$

for some function Ξ_t^i independent of $\tilde{g}^i$.

$(B_{1:t}, S_{1:t}^i, \boldsymbol{\Gamma}_{1:t}^i)$ is a function of $(\overline{H}_t^i, \boldsymbol{\Gamma}_t^i)$. Therefore, by applying smoothing property of conditional expectations to both sides of (26) we obtain

$$\mathbb{P}^{\tilde{g}^i, g^{-i}}(x_{t-d+1:t}, \gamma_t^{-i} | b_{1:t}, s_{1:t}^i, \gamma_{1:t}^i) = F_t^i(x_{t-d+1:t}, \gamma_t^{-i} | b_t, s_t^i),$$

where we know that F_t^i, as defined in (26), is independent of $\tilde{g}^i$.

We conclude that coordinator i faces a Markov Decision Problem where the state process is (B_t, S_t^i), the control action is $\boldsymbol{\Gamma}_t^i$, and the total reward is

$$\mathbb{E}\left[\sum_{t \in \mathcal{T}} \overline{r}_t^i(B_t, S_t^i, \boldsymbol{\Gamma}_t^i)\right].$$

By standard MDP theory, coordinator i can form a best response by choosing $\boldsymbol{\Gamma}_t^i$ as a function of (B_t, S_t^i).

H Proof of Theorem 2

Let (λ^*, ψ^*) be a pair that solves the dynamic program defined in the statement of the theorem. Let g^{*k} denote the behavioral coordination strategy corresponding to (λ^{*k}, ψ^*) for $k \in \mathcal{I}$. We only need to show the following: Suppose that the coordinators other than coordinator i play g^{*-i}, then g^{*i} is a best response to g^{*-i}.

Let $h_t^0 \in \mathcal{H}_t^0$ be admissible under g^{*-i}. Then,

$$\mathbb{P}^{g^{*k}}(s_t^k, x_{t-d+1:t}^k | h_t^0) = \pi_t^k(s_t^k) P_t^k(x_{t-d+1:t}^k | y_{t-d+1:t-1}^k, u_{t-d:t-1}, s_t^k) \tag{27}$$

for all $k \neq i$ by Lemma 6, where π_t^k is the belief generated by ψ^* when h_t^0 occurs.

By Lemma 4, we also have

$$\mathbb{P}(\tilde{s}_t^i, \tilde{x}_{t-d+1:t}^i | h_t^0, s_t^i) = P_t^i(\tilde{x}_{t-d+1:t}^i | y_{t-d+1:t-1}^i, u_{t-d:t-1}, \tilde{s}_t^i) \tag{28}$$

Combining (27) and (28), the belief for coordinator i defined in the stage game according to Definition 15 satisfies

$$\begin{aligned}
&\beta_t^i(\tilde{z}_t | s_t^i) \\
&= \mathbf{1}_{\{\tilde{s}_t^i = s_t^i\}} \prod_{k \neq i} \pi_t^k(\tilde{s}_t^k) \left(\prod_{k \in \mathcal{I}} P_t^k(\tilde{x}_{t-d+1:t}^k | y_{t-d+1:t-1}^k, u_{t-d:t-1}, \tilde{s}_t^k) \right) \mathbb{P}(\tilde{w}_t^Y) \\
&= \mathbb{P}(\tilde{s}_t^i, \tilde{x}_{t-d+1:t}^i | h_t^0, s_t^i) \left(\prod_{k \neq i} \mathbb{P}^{g^{*k}}(\tilde{s}_t^k, \tilde{x}_{t-d+1:t}^k | h_t^0) \right) \mathbb{P}(\tilde{w}_t^Y) \\
&= \mathbb{P}^{g^{*-i}}(\tilde{s}_t, \tilde{x}_{t-d+1:t} | h_t^0, s_t^i) \mathbb{P}(\tilde{w}_t^{k,Y}) = \mathbb{P}^{g^{*-i}}(\tilde{z}_t | h_t^0, s_t^i)
\end{aligned}$$

for all (h_t^0, s_t^i) admissible under g^{*-i}, i.e., the belief represents a true conditional distribution. Since $\beta_t^i(\cdot | s_t^i)$ is a fixed function of (b_t, s_t^i), by applying smoothing property on both sides of the above equation we can obtain

$$\beta_t^i(\tilde{z}_t | s_t^i) = \mathbb{P}^{g^{*-i}}(\tilde{z}_t | b_t, s_t^i).$$

for all (b_t, s_t^i) admissible under g^{*-i}.[10]

Then, the interim expected utility considered in the definition of IBNE correspondences (Definition 16) can be written as

$$\begin{aligned}
&\sum_{\tilde{z}_t, \tilde{\gamma}_t} \eta(\tilde{\gamma}_t^i) Q_t^i(\tilde{z}_t, \tilde{\gamma}_t) \beta_t^i(\tilde{z}_t | x_{t-1}^i) \prod_{k \neq i} \lambda_t^{*k}(\tilde{\gamma}_t^k | b_t, \tilde{s}_t^k) \\
&= \sum_{\tilde{\gamma}_t^i} \eta(\tilde{\gamma}_t^i) \mathbb{E}^{g_{1:t}^{*-i}}[Q_t^i(\mathbf{Z}_t, \boldsymbol{\Gamma}_t) | b_t, s_t^i, \tilde{\gamma}_t^i].
\end{aligned}$$

for all (b_t, s_t^i) admissible under g^{*-i}.

The condition of Theorem 2 then implies

$$\lambda_t^{*i}(b_t, s_t^i) \in \arg\max_{\eta \in \Delta(\mathcal{A}_t^i)} \sum_{\tilde{\gamma}_t} \eta(\tilde{\gamma}_t^i) \mathbb{E}^{g^{*-i}} \left[r_t^i(\mathbf{X}_t, \mathbf{U}_t) + V_{t+1}^i(B_{t+1}, S_{t+1}^i) | b_t, s_t^i, \tilde{\gamma}_t^i \right]; \tag{29}$$

$$V_t^i(b_t, s_t^i) = \sum_{\tilde{\gamma}_t^i} \left[\lambda_t^{*i}(\tilde{\gamma}_t^i | b_t, s_t^i) \mathbb{E}^{g_{1:t}^{*-i}}[r_t^i(\mathbf{X}_t, \mathbf{U}_t) + V_{t+1}^i(B_{t+1}, S_{t+1}^i) | b_t, s_t^i, \tilde{\gamma}_t^i] \right] \tag{30}$$

for all (b_t, s_t^i) admissible under g^{*-i}.

Recall that in the proof of Lemma 7, we have already proved that fixing (λ^{*-i}, ψ^*), (B_t, S_t^i) is a controlled Markov process controlled by $\boldsymbol{\Gamma}_t^i$. Hence, (29) and (30) show that λ_t^{*i}

[10] Note that $\mathbb{P}^{g^{-i}}(\tilde{z}_t | b_t, s_t^i)$ is different from $\beta_t^i(\tilde{z}_t | s_t^i)$. Since B_t is just a compression of the common information based on an predetermined update rule ψ, which may or may not be consistent with the actually played strategy, B_t may not represent the true belief. $\mathbb{P}^{g^{-i}}(\tilde{z}_t | b_t, s_t^i)$ is the belief an agent *inferred from* the event $B_t = b_t, S_t^i = s_t^i$. The agent knows that b_t might not contain the true belief, but it is useful anyway in inferring the true state. $\beta_t^i(\tilde{z}_t | s_t^i)$ is a conditional distribution *computed with* b_t, *pretending* that b_t contains the true belief.

is a dynamic programming solution of the MDP with instantaneous reward

$$\overline{r}_t^i(B_t, S_t^i, \boldsymbol{\Gamma}_t^i) := \mathbb{E}^{g^{*-i}}[r_t^i(\mathbf{X}_t, \mathbf{U}_t)|B_t, S_t^i, \boldsymbol{\Gamma}_t^i].$$

Therefore, λ^{*i} maximizes

$$\mathbb{E}^{\lambda^i, \lambda^{*-i}}\left[\sum_{t\in\mathcal{T}} \overline{r}_t^i(B_t, S_t^i, \boldsymbol{\Gamma}_t^i)\right]$$

over all $\lambda^i = (\lambda_t^i)_{t\in\mathcal{T}}$, $\lambda_t^i : \mathcal{B}_t \times \mathcal{S}_t^i \mapsto \Delta(\mathcal{A}_t^i)$.

Notice that for any λ^i, if g^i is the behavioral coordination strategy corresponding to the CIB strategy (λ^i, ψ_t^*), then by Law of Iterated Expectation

$$\mathbb{E}^{\lambda^i, \lambda^{*-i}}\left[\sum_{t\in\mathcal{T}} \overline{r}_t^i(B_t, S_t^i, \boldsymbol{\Gamma}_t^i)\right] = \mathbb{E}^{g^i, g^{*-i}}\left[\sum_{t\in\mathcal{T}} r_t^i(\mathbf{X}_t, \mathbf{U}_t)\right].$$

Hence, we know that g^{*i} maximizes

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

over all g^i generated from a CIB strategy with the belief generation system ψ^*.

By the closedness property of CIB strategies (Lemma 7), we conclude that g^{*i} is a best response to g^{*-i} over all behavioral coordination strategies of coordinator i, proving the result.

I Proof of Proposition 1

We will characterize all the Bayes–Nash equilibria of Example 1 in terms of individual players' behavioral strategies. Then, we will show that none of the BNE correspond to a CIB-CNE.

Let $p = (p_1, p_2) \in [0, 1]^2$ describe Alice's behavioral strategy: p_1 is the probability that Alice plays $U_1^A = -1$ given $X_1^A = -1$; p_2 is the probability that Alice plays $U_1^A = +1$ given $X_1^A = +1$. Let $q = (q_1, q_2) \in [0, 1]^2$ denote Bob's behavioral strategy: q_1 is the probability that Bob plays $U_3^B = \mathrm{L}$ when observing $U_1^A = -1$, q_2 is the probability that Bob plays $U_3^B = \mathrm{L}$ when observing $U_1^A = +1$.

Claim

$$p^* = \left(\frac{1}{3}, \frac{1}{3}\right), \quad q^* = \left(\frac{1}{3} + \varepsilon, \frac{1}{3} - \varepsilon\right)$$

is the unique BNE of Example 1.

Given the claim, one can conclude that a CIB-CNE does not exist in this game: Suppose that (λ^*, ψ^*) forms a CIB-CNE, then by the definition of CIB strategies, at $t = 1$ the team of Alice chooses a prescription (which maps $\mathcal{X}_1^A$ to $\mathcal{U}_1^A$) based on no information. At $t = 3$, the team of Bob chooses a prescription (which is equivalent to an action since Bob has no state) based solely on B_3. Define the induced behavioral strategy of Alice and Bob through

$$p_1 = \lambda_1^{*A}(\mathbf{id}|\varnothing) + \lambda_1^{*A}(\mathbf{cp}_{-1}|\varnothing),$$

$$p_2 = \lambda_1^{*A}(\mathbf{id}|\varnothing) + \lambda_1^{*A}(\mathbf{cp}_{+1}|\varnothing),$$
$$q_1 = \lambda_3^{*B}(\mathbf{L}|b_3[-1]),$$
$$q_2 = \lambda_3^{*B}(\mathbf{L}|b_3[+1]),$$

where $b_3[u]$ is the CCI under belief generation system ψ^* when $U_1^A = u$. $\mathbf{id}$ is the prescription that chooses $U_1^A = X_1^A$; $\mathbf{cp}_u$ is the prescription that chooses $U_1^A = u$ irrespective of X_1^A; $\mathbf{L}$ is Bob's prescription that chooses $U_3^B = \mathrm{L}$.

The consistency of ψ_1^* with respect to λ_1^* implies that

$$\Pi_2(-1) = \frac{p_1}{p_1 + 1 - p_2} \quad \text{if } p \neq (0,1), U_1 = -1,$$
$$\Pi_2(+1) = \frac{p_2}{p_2 + 1 - p_1} \quad \text{if } p \neq (1,0), U_1 = +1,$$

The consistency of ψ_2^* with respect to λ_2^* implies that

$$\Pi_3(+1) = \Pi_2(U_1^A).$$

If a CIB-CNE induces behavioral strategy $p^* = \left(\frac{1}{3}, \frac{1}{3}\right)$, then the CIB belief $\Pi_3 \in \Delta(\mathcal{X}_2)$ will be the same for both $U_1 = +1$ and $U_1 = -1$ under any consistent belief generation system ψ^*. Then, $B_3 = (\Pi_3, \mathbf{U}_2)$ will be the same for both $U_1 = +1$ and $U_1 = -1$ since $\mathbf{U}_2$ only takes one value. Hence, Bob's-induced stage behavioral strategy q should satisfy $q_1 = q_2$. However, $q^* = \left(\frac{1}{3} + \varepsilon, \frac{1}{3} - \varepsilon\right)$ is such that $q_1^* \neq q_2^*$; hence, (p^*, q^*) cannot be induced from any CIB-CNE.

Since the induced behavioral strategy of any CIB-CNE should form a BNE in the game among individuals, we conclude that a CIB-CNE does not exist in Example 1.

Proof of Claim Denote Alice's total expected payoff to be $J(p,q)$. Then,

$$\begin{aligned}
&J(p,q)\\
&= \frac{1}{2}\varepsilon(1 - p_1 + p_2) + \frac{1}{2}\left((1-p_1)(1-q_2) + p_1 \cdot 2q_1\right) +\\
&\quad + \frac{1}{2}\left((1-p_2)(1-q_1) + p_2 \cdot 2q_2\right)\\
&= \frac{1}{2}\varepsilon(1 - p_1 + p_2) + \frac{1}{2}(2 - p_1 - p_2) + \frac{1}{2}(2p_1 + p_2 - 1)q_1 + \frac{1}{2}(2p_2 + p_1 - 1)q_2.
\end{aligned}$$

Since this is a zero-sum game, Alice's expected payoff at equilibrium can be characterized as

$$J^* = \max_p \min_q J(p,q).$$

Alice plays p at some equilibrium if and only if $\min_q J(p,q) = J^*$. Define $J^*(p) = \min_q J(p,q)$. We compute

$$\begin{aligned}
J^*(p) = &\frac{1}{2}\varepsilon(1 - p_1 + p_2) + \frac{1}{2}(2 - p_1 - p_2)\\
&+ \begin{cases}
\frac{1}{2}(3p_1 + 3p_2) - 1 & 2p_1 + p_2 \leq 1, 2p_2 + p_1 \leq 1\\
\frac{1}{2}(2p_2 + p_1 - 1) & 2p_1 + p_2 > 1, 2p_2 + p_1 \leq 1\\
\frac{1}{2}(2p_1 + p_2 - 1) & 2p_1 + p_2 \leq 1, 2p_2 + p_1 > 1\\
0 & 2p_1 + p_2 > 1, 2p_2 + p_1 > 1
\end{cases}
\end{aligned}$$

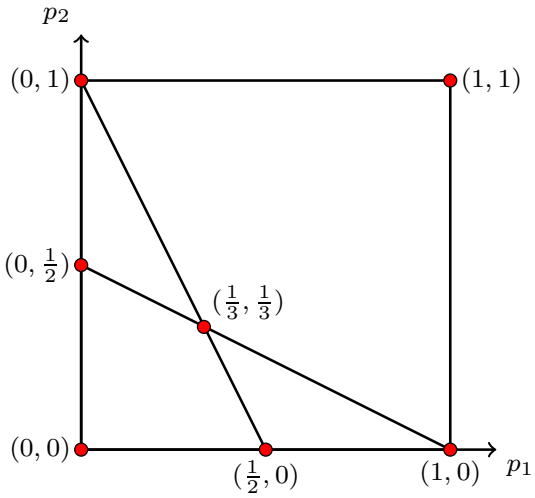


Fig. 1 The pieces (polygons) for which $J^*(p)$ is linear on. The extreme points of the pieces are labeled

The set of equilibrium strategies for Alice is the set of maximizers of $J^*(p)$. Since $J^*(p)$ is a continuous piecewise linear function, the set of maximizers can be found by comparing the values at the extreme points of the pieces.

We have

$$\begin{aligned}
J^*(0,0) &= \frac{1}{2}\varepsilon + 1 - 1 = \frac{1}{2}\varepsilon;\\
J^*\left(\frac{1}{2}, 0\right) &= \frac{1}{2}\varepsilon \cdot \frac{1}{2} + \frac{1}{2}\cdot\frac{3}{2} + \frac{1}{2}\cdot\frac{3}{2} - 1 = \frac{1}{4}\varepsilon + \frac{1}{2};\\
J^*\left(0, \frac{1}{2}\right) &= \frac{1}{2}\varepsilon \cdot \frac{3}{2} + \frac{1}{2}\cdot\frac{3}{2} + \frac{1}{2}\cdot\frac{3}{2} - 1 = \frac{3}{4}\varepsilon + \frac{1}{2};\\
J^*(1,0) &= \frac{1}{2}\varepsilon \cdot 0 + \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 0 = \frac{1}{2};\\
J^*(0,1) &= \frac{1}{2}\varepsilon \cdot 2 + \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 0 = \varepsilon + \frac{1}{2};\\
J^*\left(\frac{1}{3}, \frac{1}{3}\right) &= \frac{1}{2}\varepsilon + \frac{1}{2}\cdot\frac{4}{3} + \frac{1}{2}\cdot 0 = \frac{1}{2}\varepsilon + \frac{2}{3};\\
J^*(1,1) &= \frac{1}{2}\varepsilon + \frac{1}{2}\cdot 0 + 0 = \frac{1}{2}\varepsilon.
\end{aligned}$$

Since $\varepsilon < \frac{1}{3}$, we have $(\frac{1}{3}, \frac{1}{3})$ to be the unique maximum among the extreme points. Hence, we have $\arg\max_p J^*(p) = \{(\frac{1}{3}, \frac{1}{3})\}$, i.e., Alice always plays $p^* = (\frac{1}{3}, \frac{1}{3})$ in any BNE of the game.

Now, consider Bob's equilibrium strategy. q^* is an equilibrium strategy of Bob only if $p^* \in \arg\max_p J(p, q^*)$.

For each q, $J(p, q)$ is a linear function of p and

$$\nabla_p J(p,q) = \left(-\frac{1}{2}\varepsilon - \frac{1}{2} + q_1 + \frac{1}{2}q_2, \frac{1}{2}\varepsilon - \frac{1}{2} + \frac{1}{2}q_1 + q_2\right) \quad \forall p \in (0,1)^2.$$

We need $\nabla_p J(p, q^*)\Big|_{p=p^*} = (0, 0)$. Hence,

$$-\frac{1}{2}\varepsilon - \frac{1}{2} + q_1^* + \frac{1}{2}q_2^* = 0;$$
$$\frac{1}{2}\varepsilon - \frac{1}{2} + \frac{1}{2}q_1^* + q_2^* = 0,$$

which implies that $q^* = (\frac{1}{3} + \varepsilon, \frac{1}{3} - \varepsilon)$, proving the claim. □

J Proof of Theorem 3

We use Theorem 2 to establish the existence of CIB-CNE: We show that for each t there always exists a pair (λ_t^*, ψ_t^*) such that λ_t^* forms an equilibrium at t given ψ_t^*, and ψ_t^* is consistent with λ_t^*. We provide a constructive proof of existence of CIB-CNE by proceeding backwards in time.

Since $d = 1$, we have $S_t^i = \mathbf{X}_{t-1}^i$. The CCI consists of the beliefs along with $\mathbf{U}_{t-1}$.

Consider the condensation of the information graph into a directed acyclic graph (DAG) whose nodes are strongly connected components. Each node may contain multiple teams. Consider one topological ordering of this DAG. Denote the nodes by $[1], [2], \cdots$ ($[j]$ is reachable from $[k]$ only if $k < j$.) We use the notation $X_t^{[k]}$, $\Pi_t^{[k]}$ to denote the vector of the system variables of the teams in a node. In particular, following Definition 15, we define $\mathbf{Z}_t^{[k]} = (\mathbf{X}_{t-1:t}^{[k]}, \mathbf{W}_t^{[k],Y})$. We also use $[1:k]$ as a short hand for the set $[1] \cup [2] \cup \cdots \cup [k]$. Define $B_t^{[1:k]} = (\Pi_t^{[1:k]}, \mathbf{U}_{t-1}^{[1:k]})$. (Note that the usage of superscript here is different from the CCI B_t^i defined in Definition 12.)

We construct the solution first backwards in time, then in the order of the node for each stage. For this purpose, we an some induction invariant on the value functions V_t^i (as defined in Theorem 2) for the solution we are going to construct.

Induction Invariant: For each time t and each node index k,

- $V_t^i(b_t, x_{t-1}^i)$ depends on b_t only through $(b_t^{[1:k-1]}, u_{t-1}^i)$ for all teams $i \in [k]$, if $[k]$ consists of only one team. (With some abuse of notation, we write $V_t^i(b_t, x_{t-1}^i) = V_t^i(b_t^{[1:k-1]}, u_{t-1}^i, x_{t-1}^i)$ in this case.)
- $V_t^i(b_t, x_{t-1}^i)$ depends on b_t only through $b_t^{[1:k]}$ for all teams $i \in [k]$, if $[k]$ consists of multiple public teams. (We write $V_t^i(b_t, x_{t-1}^i) = V_t^i(b_t^{[1:k]}, x_{t-1}^i)$ in this case.)

Induction Base: For $t = T + 1$, we have $V_{T+1}^i(\cdot) \equiv 0$ for all coordinators $i \in \mathcal{I}$ hence the induction invariant is true.

Induction Step: Suppose that the induction invariant is true at time $t + 1$ for all nodes. We construct the solution so that it is also true at time t.

To complete this step, we provide a procedure to solve the stage game. We argue that one can solve a series of optimization problems or finite games following the topological order of the nodes through an inner induction step.

Inner Induction Step: Suppose that the first $k - 1$ nodes has been solved, and the equilibrium strategy $\lambda_t^{*[1:k-1]}$ uses only $b_t^{[1:k-1]}$ along with private information. Suppose that the update rules $\psi_t^{*,[1:k-1]}$ have also been determined, and they use only $(b_t^{[1:k-1]}, y_t^{[1:k-1]}, u_t^{[1:k-1]})$. We now establish the same property for $(\lambda_t^{[k]}, \psi_t^{[k]})$.

- If the k-th node contains a single coordinator i, the value to go is $V_{t+1}^i(B_{t+1}^{[1:k-1]}, \mathbf{U}_t^i, \mathbf{X}_t^i)$ by the induction hypothesis. The instantaneous reward for a coordinator i in the k-th

node can be expressed by $r_t^i(\mathbf{X}_t^{[1:k]}, \mathbf{U}_t^{[1:k]})$ by the information graph. In the stage game, coordinator i chooses a prescription to maximize the expected value of

$$Q_t^i(b_t^{[1:k-1]}, \mathbf{Z}_t^{[1:k]}, \boldsymbol{\Gamma}_t^{[1:k]}) := r_t^i(\mathbf{X}_t^{[1:k]}, \mathbf{U}_t^{[1:k]}) + V_{t+1}^i(B_{t+1}^{[1:k-1]}, \mathbf{U}_t^i, \mathbf{X}_t^i),$$

where

$$\begin{aligned} B_{t+1}^{[1:k-1]} &= (\Pi_{t+1}^{[1:k-1]}, \mathbf{U}_t^{[1:k-1]}), \\ \Pi_{t+1}^j &= \psi_t^{*,j}(b_t^{[1:k-1]}, \mathbf{Y}_t^j, \mathbf{U}_t^{[1:k-1]}) \quad \forall j \in [1:k-1], \\ \mathbf{Y}_t^j &= \ell_t^j(\mathbf{X}_t^j, \mathbf{U}_t^{[1:k-1]}, \mathbf{W}_t^{j,Y}) \quad \forall j \in [1:k-1], \\ \mathbf{U}_t^j &= \boldsymbol{\Gamma}_t^j(\mathbf{X}_t^j) \quad \forall j \in [1:k]. \end{aligned}$$

The expectation is computed using the belief β_t^i (defined through Eq. (4) in Definition 15) along with $\lambda_t^{*[1:k-1]}$ that has already been determined. It can be written as

$$\begin{aligned} &\sum_{\tilde{s}_t, \tilde{\gamma}_t^{[1:k-1]}} \beta_t^i(\tilde{s}_t | x_{t-1}^i) Q_t^i(b_t^{[1:k-1]}, \tilde{s}_t^{[1:k]}, (\tilde{\gamma}_t^{[1:k-1]}, \gamma_t^i)) \\ &\quad \times \prod_{j \in [1:k-1]} \lambda_t^j(\tilde{\gamma}_t^j | b_t^{[1:k-1]}, \tilde{x}_{t-1}^j) \\ &\quad = \sum_{\tilde{s}_t^{[1:k]}, \tilde{\gamma}_t^{[1:k-1]}} \mathbf{1}_{\{\tilde{x}_{t-1}^i = x_{t-1}^i\}} \mathbb{P}(\tilde{w}_t^{[1:k],Y}) \left(\prod_{j \in [1:k-1]} \pi_t^j(\tilde{x}_{t-1}^j) \mathbb{P}(\tilde{x}_t^j | \tilde{x}_{t-1}^j, u_{t-1}^{[1:k-1]}) \right) \\ &\quad \times \left(\prod_{j \in [1:k-1]} \lambda_t^{*j}(\tilde{\gamma}_t^j | b_t^{[1:k-1]}, x_{t-1}^j) \right) \\ &\quad \times \mathbb{P}(\tilde{x}_t^i | x_{t-1}^i, u_{t-1}^{[1:k]}) Q_t^i(b_t^{[1:k-1]}, \tilde{s}_t^{[1:k]}, (\tilde{\gamma}_t^{[1:k]}, \gamma_t^i)). \end{aligned}$$

Therefore, the expected reward of coordinator i depends on b_t through $(b_t^{[1:k-1]}, u_{t-1}^i)$. Coordinator i can choose the optimal prescription based on $(b_t^{[1:k-1]}, u_{t-1}^i, x_{t-1}^i)$, i.e., $\lambda_t^{*i}(b_t, x_{t-1}^i) = \lambda_t^{*i}(b_t^{[1:k-1]}, u_{t-1}^i, x_{t-1}^i)$. We then have $V_t^i(b_t, x_{t-1}^i) = V_t^i(b_t^{[1:k-1]}, u_{t-1}^i, x_{t-1}^i)$. The update rule $\psi_t^{*,[k]} = \psi_t^{*,i}$ is then determined to be an arbitrary update rule consistent with $\lambda_t^{*,i}$, which can be chosen as a function from $\mathcal{B}_t^{[1:k]} \times \mathcal{Y}_t^{[k]} \times \mathcal{U}_t^{[1:k]}$ (instead of $\mathcal{B}_t \times \mathcal{Y}_t^{[k]} \times \mathcal{U}_t$) to $\Pi_{t+1}^{[k]}$.

– If the k-th node contains a group of public teams, then update rules $\hat{\psi}_t^{*,[k]}$ are fixed, irrespective of the stage game strategies, i.e., there exist a unique update rule $\hat{\psi}_t^{*,i}$ that is compatible with any $\lambda_t^{*,i}$ for a public team i. This update rule is a map from $\mathcal{Y}_t^{[k]} \times \mathcal{U}_t^{[1:k]}$ to a vector of delta measures on $\prod_{i \in [k]} \Delta(\mathcal{X}_{t-1}^i)$, i.e., the map to recover $\mathbf{X}_{t-1}^{[k]}$ from the observations (see Definition 18). The function takes $\mathbf{U}_t^{[1:k]}$ as its argument due to the fact that the observations of the k-th node depends on $\mathbf{U}_t$ only through $\mathbf{U}_t^{[1:k]}$.
The value to go for each coordinator i can be expressed as $V_{t+1}^i(B_t^{[1:k]}, \mathbf{X}_{t-1}^i)$ by induction hypothesis. The instantaneous reward can be written as $r_t^i(\mathbf{X}_t^{[1:k]}, \mathbf{U}_t^{[1:k]})$ by the definition of the information dependency graph.
In the stage game, coordinator i in the k-th node chooses a distribution η_t^i on prescriptions to maximize the expected value of

$$Q_t^i(b_t^{[1:k]}, \mathbf{Z}_t^{[1:k]}, \boldsymbol{\Gamma}_t^{[1:k]}) := r_t^i(\mathbf{X}_t^{[1:k]}, \mathbf{U}_t^{[1:k]}) + V_{t+1}^i(B_{t+1}^{[1:k]}, \mathbf{X}_t^i),$$

where

$$
\begin{aligned}
B_{t+1}^{[1:k]} &= (\Pi_{t+1}^{[1:k]}, \mathbf{U}_t^{[1:k]}),\\
\Pi_{t+1}^{j} &= \psi_t^{*,j}(b_t^{[1:k-1]}, \mathbf{Y}_t^{j}, \mathbf{U}_t^{[1:k-1]}) \quad \forall j \in [1:k-1],\\
\Pi_{t+1}^{[k]} &= \hat{\psi}_t^{*,[k]}(b_t^{[1:k]}, \mathbf{Y}_t^{[1:k]}, \mathbf{U}_t^{[1:k]}),\\
\mathbf{Y}_t^{j} &= \ell_t^{j}(\mathbf{X}_t^{j}, \mathbf{U}_t^{[1:k]}, \mathbf{W}_t^{j,Y}) \quad \forall j \in [1:k],\\
\mathbf{U}_t^{j} &= \boldsymbol{\Gamma}_t^{j}(\mathbf{X}_t^{j}) \quad \forall j \in [1:k].
\end{aligned}
$$

The expectation is taken with respect to the belief β_t^i (defined through Eq. (4) in Definition 15) and the strategy prediction $\lambda_t^{[1:k]}$. This expectation can be written as

$$
\begin{aligned}
&\sum_{\tilde{s}_t, \tilde{\gamma}_t^{[1:k]}} \beta_t^i(\tilde{s}_t | x_{t-1}^i) Q_t(b_t^{[1:k]}, \tilde{s}_t^{[1:k]}, \tilde{\gamma}_t^{[1:k]}) \eta_t^i(\tilde{\gamma}_t^i) \times \prod_{\substack{j\in[1:k]\\ j\neq i}} \lambda_t^j(\tilde{\gamma}_t^j | b_t^{[1:k-1]}, \tilde{x}_{t-1}^j)\\
&= \sum_{\tilde{s}_t^{[1:k]}, \tilde{\gamma}_t^{[1:k]}} \mathbf{1}_{\{\tilde{x}_{t-1}^i = x_{t-1}^i\}} \mathbb{P}(\tilde{w}_t^{[1:k],Y})\\
&\quad \times \left(\prod_{\substack{j\in[1:k]\\ j\neq i}} \pi_t^j(\tilde{x}_{t-1}^j) \mathbb{P}(\tilde{x}_t^j | \tilde{x}_{t-1}^j, u_{t-1}^{[1:k]}) \lambda_t^{*j}(\tilde{\gamma}_t^j | b_t^{[1:k]}, \tilde{x}_{t-1}^j) \right)\\
&\quad \times \mathbb{P}(\tilde{x}_t^i | x_{t-1}^i, u_{t-1}^{[1:k]}) \eta_t^i(\tilde{\gamma}_t^i) Q_t^i(b_t^{[1:k]}, \tilde{s}_t^{[1:k]}, \tilde{\gamma}_t^{[1:k]}),
\end{aligned}
$$

which dependents only on b_t only through $b_t^{[1:k]}$. Therefore, the stage game defined in Definition 15 induces a finite game between the coordinators in the k-th node (instead of all coordinators) with parameter $(b_t^{[1:k]}, (\psi_t^{*,[1:k-1]}, \hat{\psi}_t^{*,[k]}))$ (instead of (b_t, ψ_t)), where $\lambda_t^{*[1:k-1]}$ has been fixed. Teams in the k-th node form/play a stage game where the first $k-1$ nodes act like nature, while the coordinators after k-th node have no effect in the payoffs of the coordinators in the k-th node. Hence, a coordinator i in the k-th node can based their decision on $(b_t^{[1:k]}, x_{t-1}^i)$, i.e., $\lambda_t^{*i}(b_t, x_{t-1}^i) = \lambda_t^{*i}(b_t^{[1:k]}, x_{t-1}^i)$. We also have $V_t^i(b_t, x_{t-1}^i) = V_t^i(b_t^{[1:k]}, x_{t-1}^i)$. The update rule is determined by $\psi_t^{*,[k]} = \hat{\psi}_t^{*,[k]}$, which is guaranteed to be consistent with $\lambda_t^{*[k]}$.

In summary, we determine (λ_t^*, ψ_t^*) using a node-by-node approach. If the k-th node consists of one team, then we first determine $\lambda_t^{*[k]}$ from an optimization problem dependent on $(\lambda_t^{*[1:k-1]}, \psi_t^{*,[1:k-1]})$ and then determine $\psi_t^{*,[k]}$. If the k-th node consists of multiple public players, then we first determine $\psi_t^{*,[k]}$ and then solve $\lambda_t^{*[k]}$ from a finite game dependent on $(\lambda_t^{*[1:k-1]}, \psi_t^{*,[1:k]})$. Hence, we have constructed the solution and established both inner and outer induction steps, proving the theorem.

K Proof of Theorem 4

We prove the Theorem for $d = 1$. The proof idea for $d > 1$ is similar.

We will prove a stronger result. For each $\Pi_t^i \in \Delta(\mathcal{X}_{t-1}^i)$, define the corresponding $\hat{\Pi}_t^i \in \Delta(\mathcal{X}_t)$ by

$$\overline{\Pi}_t^i(x_t^i) := \sum_{\tilde{x}_{t-1}^i} \Pi_t^i(\tilde{x}_{t-1}^i)\mathbb{P}(x_t^i|\tilde{x}_{t-1}^i).$$

Define $\hat{\psi}_t^i$ to be the signaling-free update function, i.e., the belief update function such that

$$\Pi_{t+1}^i(x_t^i) = \hat{\psi}_t^i(\overline{\Pi}_t^i, \mathbf{Y}_t^i) = \frac{\overline{\Pi}_t^i(x_t^i)\mathbb{P}(\mathbf{Y}_t^i|x_t^i)}{\sum_{\tilde{y}_t^i} \overline{\Pi}_t^i(x_t^i)\mathbb{P}(\tilde{y}_t^i|x_t^i)}.$$

Define *open-loop* prescriptions as the prescriptions that simply instruct members of a team to take a certain action irrespective their private information. We will show that there exist an equilibrium where each team plays a common information-based signaling-free (CIBSF) strategy, i.e., the common belief generation system for all coordinators is given by the signaling-free update functions $\hat{\psi}$, and coordinator i chooses randomized open-loop prescriptions based on $\overline{\boldsymbol{\Pi}}_t = (\overline{\Pi}_t^i)_{i\in\mathcal{I}}$ instead of $(B_t, \mathbf{X}_{t-1}^i)$.

Induction Invariant: $V_t^i(B_t, \mathbf{X}_{t-1}^i) = V_t^i(\overline{\boldsymbol{\Pi}}_t, \mathbf{X}_{t-1}^i)$.

Induction Base: The induction variant is true for $t = T+1$ since $V_{T+1}^i(\cdot) \equiv 0$ for all $i \in \mathcal{I}$.

Induction Step: Suppose that the induction variant is true for $t+1$, prove it for time t.

Let $\hat{\psi}_t$ be the signaling-free update rule. We solve the stage game $G_t(V_{t+1}, \hat{\psi}_t, b_t)$. In the stage game, coordinator i chooses a prescription to maximize the expectation of

$$r_t^i(\mathbf{X}_t^{-i}, \mathbf{U}_t) + V_{t+1}^i(\overline{\boldsymbol{\Pi}}_{t+1}, \mathbf{X}_t^i),$$

where

$$\begin{aligned}
\overline{\Pi}_{t+1}^k(x_{t+1}^k) &= \sum_{\tilde{x}_t^k} \Pi_{t+1}^k(\tilde{x}_t^k)\mathbb{P}(x_{t+1}^k|\tilde{x}_t^k) \quad \forall x_{t+1}^k \in \mathcal{X}_{t+1}^k,\\
\Pi_{t+1}^k &= \hat{\psi}_t^k(\overline{\Pi}_t^k, \mathbf{Y}_t^k) \quad \forall k \in \mathcal{I},\\
\mathbf{Y}_t^k &= \ell_t^k(\mathbf{X}_t^k, \mathbf{W}_t^{k,Y}) \quad \forall k \in \mathcal{I},\\
U_t^{k,j} &= \Gamma_t^{k,j}(X_t^{k,j}) \quad \forall (k,j) \in \mathcal{N}.
\end{aligned}$$

Since $V_{t+1}^i(\overline{\boldsymbol{\Pi}}_{t+1}, \mathbf{X}_t^i)$ does not depend on coordinator i's prescriptions, coordinator i only need to maximize the expectation of $r_t^i(\mathbf{X}_t^{-i}, \mathbf{U}_t)$, which is

$$\sum_{\tilde{x}_{t-1:t}^{-i}, \tilde{\gamma}_t^{-i}} \left(\prod_{j\neq i} \pi_t^j(\tilde{x}_{t-1}^j)\mathbb{P}(\tilde{x}_t^j|\tilde{x}_{t-1}^j)\lambda_t^j(\tilde{\gamma}_t^j|b_t, \tilde{x}_{t-1}^j) \right) r_t^i(\tilde{x}_t^{-i}, (\tilde{\gamma}_t^{-i}(\tilde{x}_t^{-i}), \gamma_t^i(x_t^i))).$$

Claim In the stage game, if all coordinators $-i$ use CIBSF strategy, then coordinator i can respond with a CIBSF strategy.

Proof of Claim Let $\eta_t^k : \overline{\Pi}_t \mapsto \Delta(\mathcal{U}_t^k)$ be the CIBSF strategy of coordinator $k \neq i$. Then, coordinator i's expected payoff given γ_t^i can be written as

$$\sum_{\tilde{x}_{t-1:t}^{-i}, \tilde{u}_t^{-i}} \left(\prod_{j\neq i} \pi_t^j(\tilde{x}_{t-1}^j)\mathbb{P}(\tilde{x}_t^j|\tilde{x}_{t-1}^j)\eta_t^j(\tilde{u}_t^j|\overline{\pi}_t) \right) r_t^i(\tilde{x}_t^{-i}, (\tilde{u}_t^{-i}, \gamma_t^i(x_t^i)))$$

$$= \sum_{\tilde{x}_t^{-i}, \tilde{u}_t^{-i}} \left(\prod_{j \neq i} \left(\sum_{\tilde{x}_{t-1}^j} \pi_t^j(\tilde{x}_{t-1}^j) \mathbb{P}(\tilde{x}_t^j | \tilde{x}_{t-1}^j) \right) \eta_t^j(\tilde{u}_t^j | \overline{\pi}_t) \right) \times r_t^i(\tilde{x}_t^{-i}, (\tilde{u}_t^{-i}, \gamma_t^i(x_t^i)))$$

$$= \sum_{\tilde{x}_t^{-i}, \tilde{u}_t^{-i}} \left(\prod_{j \neq i} \overline{\pi}_t^j(\tilde{x}_t^j) \eta_t^j(\tilde{u}_t^j | \overline{\pi}_t) \right) r_t^i(\tilde{x}_t^{-i}, (\tilde{u}_t^{-i}, \gamma_t^i(x_t^i)))$$

$$=: \overline{r}_t^i(\overline{\pi}_t, \eta_t^{-i}, \gamma_t^i(x_t^i)).$$

Hence, coordinator i can respond with a prescription γ_t^i such that $\gamma_t^i(x_t^i) = u_t^i$ for all x_t^i, where

$$u_t^i \in \arg\max_{\tilde{u}_t^i} \overline{r}_t^i(\overline{\pi}_t, \eta_t^{-i}, \tilde{u}_t^i),$$

can be chosen based on $(\overline{\pi}_t, \eta_t^{-i})$, proving the claim. □

Given the claim, we conclude that there exist a stage game equilibrium where all coordinators play CIBSF strategies: Define a new stage game where we restrict each coordinator to CIBSF strategies. A best response in the restricted stage game will be also a best response in the original stage game due to the claim. The restricted game is a finite game (It is a game of symmetrical information with parameter $\overline{\pi}_t$ where coordinator i's action is u_t^i and its payoff is a function of $\overline{\pi}_t$ and u_t.) that always has an equilibrium. The equilibrium strategy will be consistent with $\hat{\psi}_t$ due to Lemma 10.

Lemma 10 *The signaling-free update rule $\hat{\psi}_t^i$ is consistent with any $\lambda_t^i : \mathcal{B}_t \times \mathcal{X}_{t-1}^i \mapsto \Delta(\mathcal{A}_t^i)$ that corresponds to a CIBSF strategy at time t.*

Proof It follows from standard arguments related to strategy independence of belief (See Chapter 6 of Kumar and Varaiya [25]). □

Let $\eta_t^* = (\eta_t^{*j})_{j \in \mathcal{I}}$, $\eta_t^{*j} : \overline{\Pi}_t \mapsto \Delta(\mathcal{U}_t^j)$ be a CIBSF strategy profile that is a stage game equilibrium. Then, the value function

$$V_t^i(b_t, x_{t-1}^i) = \left(\max_{\tilde{u}_t^i} \overline{r}_t^i(\overline{\pi}_t, \eta_t^{*-i}, \tilde{u}_t^i) \right) + \sum_{\tilde{x}_t, \tilde{y}_t} V_{t+1}^i(\hat{\psi}_t(\overline{\pi}_t, \tilde{y}_t), \tilde{x}_t^i) \mathbb{P}(\tilde{y}_t | \tilde{x}_t) \mathbb{P}(\tilde{x}_t^i | x_{t-1}^i) \overline{\pi}_t^{-i}(\tilde{x}_t^{-i})$$

depends on (b_t, x_{t-1}^i) only through $(\overline{\pi}_t, x_{t-1}^i)$, establishing the induction step.

L Proof of Lemma 8

In this appendix, when we specify a team's strategy through a profile of individual strategies, for example $\varphi^i = (\varphi^{i,l})_{(i,l) \in \mathcal{N}_i}$, we assume that members of team i apply these strategies independent of their teammates.

We first show three auxiliary results, Lemmas 11–13 that forms the basis of our proof of Lemma 8.

Lemma 11 (Conditional independence among teammates) *Suppose that members of team i use behavioral strategies $\varphi^i = (\varphi^{i,l})_{(i,l)\in\mathcal{N}_i}$ where $\varphi^{i,l} = (\varphi_t^{i,l})_{t\in\mathcal{T}}$, $\varphi_t^{i,l} : \mathcal{H}_t^{i,l} \mapsto \Delta(\mathcal{U}_t^{i,l})$. Suppose that all teams other than i use a behavioral coordination strategy profile g^{-i}. Then, $(\mathbf{X}_{t-d+1:t}^{i,l})_{(i,l)\in\mathcal{N}_i}$ are conditionally independent given the common information H_t^i. Furthermore, the conditional distribution of $\mathbf{X}_{t-d+1:t}^{i,j}$ given H_t^i depends on (φ^i, g^{-i}) only through $\varphi^{i,j}$.*

Lemma 12 *Let $\mu^{i,j}$ be a pure strategy of agent (i,j). Let $\varphi_t^{i,-j} = (\varphi_t^{i,l})_{(i,l)\in\mathcal{N}_i\setminus\{(i,j)\},t\in\mathcal{T}}$, $\varphi_t^{i,l}:\mathcal{H}_t^{i,l} \mapsto \Delta(\mathcal{U}_t^{i,l})$ be behavioral strategies of all members of team i except (i,j). Then, there exist a behavioral strategy $\bar{\varphi}^{i,j} = (\bar{\varphi}_t^{i,j})_{t\in\mathcal{T}}$, $\bar{\varphi}_t^{i,j} : \mathcal{H}_t^i \times \mathcal{X}_t^{i,j} \mapsto \Delta(\mathcal{U}_t^{i,j})$ such that $(\mu^{i,j}, \varphi^{i,-j})$ is payoff-equivalent to $(\bar{\varphi}^{i,j}, \varphi^{i,-j})$.*

Lemma 13 *Let μ^i be a pure strategy of team i. There exists a payoff-equivalent behavioral strategy profile $\bar{g}^i$ that only assigns simple prescriptions.*

Based on Lemmas 11–13, we proceed to complete the proof of Lemma 8 via the following steps.

1. Let σ^i be a payoff-equivalent mixed team strategy to g^i. (See Sect. 3).
2. For each $\mu^i \in \text{supp}(\sigma^i)$, let $\bar{g}^i[\mu^i]$ be a payoff-equivalent behavioral strategy profile $\bar{g}^i$ that only assigns simple prescriptions (Lemma 13)
3. Let $\bar{\varsigma}^i[\mu^i]$ be a payoff-equivalent mixed coordination strategy of $\bar{g}^i[\mu^i]$ constructed from Kuhn's Theorem [24].
4. Define a new mixed coordination strategy $\bar{\varsigma}^i$ by

$$\bar{\varsigma}^i = \sum_{\mu^i\in\text{supp}(\sigma^i)} \sigma^i(\mu^i)\cdot\bar{\varsigma}^i[\mu^i].$$

5. Let $\bar{g}^i$ be a payoff-equivalent behavioral coordination strategy profile to $\bar{\varsigma}^i$ constructed from Kuhn's Theorem [24].

It is clear that $\bar{g}^i$ will be payoff-equivalent to σ^i. Furthermore, $\bar{g}^i$ always assigns simple prescriptions since the construction in Kuhn's Theorem does not change the set of possible prescriptions.

Proof of Lemma 11 Assume that h_t^i is admissible under φ^i. Let g^i be a behavioral coordination strategy defined by

$$g_t^i(\gamma_t^i|\overline{h}_t^i) = \prod_{(i,j)\in\mathcal{N}_i}\prod_{x_{t-d+1:t}^{i,j}} \varphi_t^{i,j}(\gamma_t^{i,j}(x_{t-d+1:t}^{i,j})|h_t^i, x_{t-d+1:t}^{i,j}),$$

i.e., at time t, the coordinator generate independent prescriptions for each member of the team. If we view the prescription $\Gamma_t^{i,j}$ as a table of actions, then it is determined as follows: Each entry of the table is determined independently, where the entry corresponding to $x_{t-d+1:t}^{i,j}$ is randomly drawn with distribution $\varphi_t^{i,j}(h_t^i, x_{t-d+1:t}^{i,j})$.

Using arguments similar to those in the proof of Lemma 1 one can show that (g^i, g^{-i}) and (φ^i, g^{-i}) generate the same distributions of $(\mathbf{Y}_{1:t}, \mathbf{U}_{1:t}, \mathbf{X}_{1:t})$, hence

$$\mathbb{P}^{\varphi^i,g^{-i}}(x_{t-d+1:t}^i|h_t^i) = \mathbb{P}^{g^i,g^{-i}}(x_{t-d+1:t}^i|h_t^i).$$

By Lemma 3, we know that $\mathbb{P}^g(x^i_{1:t}, \gamma^i_{1:t}|h^0_t)$ does not depend on g^{-i}. As a conditional distribution obtained from $\mathbb{P}^g(x^i_{1:t}, \gamma^i_{1:t}|h^0_t)$, $\mathbb{P}^g(x^i_{t-d+1:t}|h^i_t)$ does not depend on g^{-i} either. Therefore, we have

$$\mathbb{P}^{g^i,g^{-i}}(x^i_{t-d+1:t}|h^i_t) = \mathbb{P}^{g^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t)$$

where $\hat{g}^{-i}$ is an open-loop strategy profile that always generates the actions $u^{-i}_{1:t-1}$.

Again, $(g^i, \hat{g}^{-i})$ and $(\varphi^i, \hat{g}^{-i})$ generate the same distributions on $(\mathbf{Y}_{1:t}, \mathbf{U}_{1:t}, \mathbf{X}_{1:t})$, hence

$$\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t) = \mathbb{P}^{g^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t).$$

We now have

$$\mathbb{P}^{\varphi^i,g^{-i}}(x^i_{t-d+1:t}|h^i_t) = \mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t).$$

Due to Lemma 3, we also have

$$\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t) = \mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{t-d+1:t}|h^i_t, x^{-i}_{t-d:t})$$

where $x^{-i}_{t-d:t} \in \mathcal{X}^{-i}_{t-d:t}$ is such that $\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^{-i}_{t-d:t}|h^i_t) > 0$.

Let $\tau = t - d + 1$. By Bayes' rule,

$$\begin{aligned}
&\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{\tau:t}|h^i_t, x^{-i}_{\tau-1:t}) \\
&= \frac{\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x_{\tau:t}, y_{\tau:t-1}, u_{\tau:t-1}|h^0_\tau, x_{1:\tau-1}, x^{-i}_{\tau-1})}{\sum_{\tilde{x}^i_{\tau:t}} \mathbb{P}^{\varphi^i,\hat{g}^{-i}}(\tilde{x}^i_{\tau:t}, x^{-i}_{\tau:t}, y_{\tau:t-1}, u_{\tau:t-1}|h^0_\tau, x_{1:\tau-1}, x^{-i}_{\tau-1})}
\end{aligned} \tag{31}$$

We have

$$\begin{aligned}
&\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x_{\tau:t}, y_{\tau:t-1}, u_{\tau:t-1}|h^0_\tau, x_{1:\tau-1}, x^{-i}_{\tau-1}) \\
&= \prod_{l=1}^{d-1} \Big[\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x_{t-l+1}, y_{t-l}|y_{1:t-l-1}, u_{1:t-l}, x^i_{1:t-l}, x^{-i}_{\tau-1:t-l}) \\
&\quad \times \mathbb{P}^{\varphi^i,\hat{g}^{-i}}(u^i_{t-l}|y_{1:t-l-1}, u_{1:t-l-1}, x^i_{1:t-l}, x^{-i}_{\tau-1:t-l}) \Big] \\
&\quad \times \mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x_\tau|h^0_\tau, x_{1:\tau-1}, x^{-i}_{\tau-1}) \\
&= \prod_{l=1}^{d-1} \Big[\Big(\prod_{(i,j)\in\mathcal{N}_i} \mathbb{P}(x^{i,j}_{t-l+1}|x^{i,j}_{t-l}, u_{t-l}) \mathbb{P}(y^{i,j}_{t-l}|x^{i,j}_{t-l}, u_{t-l}) \\
&\quad \times \varphi^{i,j}_{t-l}(u^{i,j}_{t-l}|h^{i,j}_{t-l}) \Big) \mathbb{P}(x^{-i}_{t-l+1}|x^{-i}_{t-l}, u_{t-l}) \mathbb{P}(y^{-i}_{t-l}|x^{-i}_{t-l}, u_{t-l}) \Big] \\
&\quad \times \left(\prod_{(i,j)\in\mathcal{N}_i} \mathbb{P}(x^{i,j}_\tau|x^{i,j}_{\tau-1}, u_{\tau-1}) \right) \mathbb{P}(x^{-i}_\tau|x^{-i}_{\tau-1}, u_{\tau-1}).
\end{aligned} \tag{32}$$

Substituting (32) into (31), we obtain

$$\begin{aligned}
\mathbb{P}^{\varphi^i,\hat{g}^{-i}}(x^i_{\tau:t}|h^i_t, x^{-i}_{\tau-1:t}) &= \frac{\prod_{(i,j)\in\mathcal{N}_i} F^{i,j}_t(x^{i,j}_{\tau:t}, h^i_t)}{\sum_{\tilde{x}^i_{\tau:t}} \prod_{(i,j)\in\mathcal{N}_i} F^{i,j}_t(\tilde{x}^{i,j}_{\tau:t}, h^i_t)} \\
&= \prod_{(i,j)\in\mathcal{N}_i} \frac{F^{i,j}_t(x^{i,j}_{\tau:t}, h^i_t)}{\sum_{\tilde{x}^{i,j}_{\tau:t}} F^{i,j}_t(\tilde{x}^{i,j}_{\tau:t}, h^i_t)}
\end{aligned}$$

where

$$F_t^{i,j}(x_{\tau:t}^{i,j}, h_t^i) = \prod_{s=1}^{d-1}\left[\mathbb{P}(x_{t-l+1}^{i,j}|x_{t-l}^{i,j}, u_{t-l})\mathbb{P}(y_{t-l}^{i,j}|x_{t-l}^{i,j}, u_{t-l})\varphi_{t-l}^{i,j}(u_{t-l}^{i,j}|h_{t-l}^{i,j})\right] \times \mathbb{P}(x_\tau^{i,j}|x_{\tau-1}^{i,j}, u_{\tau-1})$$

is a function that depends on $\varphi^{i,j}$ but not $\varphi^{i,-j}$.

Therefore, we have proved that

$$\mathbb{P}^{\varphi^i, g^{-i}}(x_{\tau:t}^i|h_t^i) = \prod_{(i,j)\in\mathcal{N}_i} \frac{F_t^{i,j}(x_{\tau:t}^{i,j}, h_t^i)}{\sum_{\tilde{x}_{\tau:t}^{i,j}} F_t^{i,j}(\tilde{x}_{\tau:t}^{i,j}, h_t^i)}. \tag{33}$$

Marginaling (33) we have

$$\mathbb{P}^{\varphi^i, g^{-i}}(x_{\tau:t}^{i,j}|h_t^i) = \frac{F_t^{i,j}(x_{\tau:t}^{i,j}, h_t^i)}{\sum_{\tilde{x}_{\tau:t}^{i,j}} F_t^{i,j}(\tilde{x}_{\tau:t}^{i,j}, h_t^i)}$$

which depends on (φ^i, g^{-i}) only through $\varphi^{i,j}$.

Hence, we conclude that

$$\mathbb{P}^{\varphi^i, g^{-i}}(x_{t-d+1:t}^i|h_t^i) = \prod_{(i,j)\in\mathcal{N}_i} \mathbb{P}^{\varphi^i, g^{-i}}(x_{t-d+1:t}^{i,j}|h_t^i),$$

and $\mathbb{P}^{\varphi^i, g^{-i}}(x_{t-d+1:t}^{i,j}|h_t^i)$ depends on (φ^i, g^{-i}) only through $\varphi^{i,j}$.

Remark 14 In general, the conditional independence among teammates is not true when team members jointly randomize.

□

Proof of Lemma 12 For notational convenience, define

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

Due to Lemma 11, $\mathbb{P}^{\varphi^i, g^{-i}}(\tilde{x}_{t-d+1:t-1}^{i,j}|h_t^i, x_t^{i,j})$ depends on the strategy profile only through $\varphi^{i,j}$.

Set

$$\overline{\varphi}_t^{i,j}(u_t^{i,j}|h_t^i, x_t^{i,j}) = \sum_{\tilde{x}_{t-d+1:t-1}^{i,j}} \mathbf{1}_{\{u_t^{i,j}=\mu_t^{i,j}(h_t^i, \tilde{x}_{t-d+1:t-1}^{i,j}, x_t^{i,j})\}} \mathbb{P}^{\mu^{i,j}}(\tilde{x}_{t-d+1:t-1}^{i,j}|h_t^i, x_t^{i,j})$$

for all $(h_t^i, x_t^{i,j})$ admissible under $\mu^{i,j}$. Otherwise, $\overline{\varphi}_t^{i,j}(h_t^i, x_t^{i,j})$ is set arbitrarily.

Let μ^{-i} be a pure team strategy profile of teams other than i. Let the superscript $-(i,j)$ denote all agents (of all teams) other than (i,j). We will prove by induction that

$$\mathbb{P}^{\mu^{i,j}, \varphi_t^{i,-j}, \mu^{-i}}(u_t, x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) = \mathbb{P}^{\overline{\varphi}^{i,j}, \varphi_t^{i,-j}, \mu^{-i}}(u_t, x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) \tag{34}$$

Given (34), the claim can be established with linearity of expectation similar to the proof of Lemma 5.

Induction Base: (34) is true for $t = 1$ since $\overline{\varphi}_1^{i,j}$ is the same strategy as $\mu_1^{i,j}$.

Induction Step: Suppose that (34) is true for time $t-1$. Prove the result for time t.

First,

$$
\begin{aligned}
&\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(u_t, x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) \\
&= \sum_{\tilde{x}_{t-d+1:t-1}^{i,j}} \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(u_t | x_t, \tilde{x}_{t-d+1:t-1}^{i,j}, x_{t-d+1:t-1}^{-(i,j)}, h_t) \\
&\quad \times \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, \tilde{x}_{t-d+1:t-1}^{i,j}, x_{t-d+1:t-1}^{-(i,j)}, h_t) \\
&= \sum_{\tilde{x}_{t-d+1:t-1}^{i,j}} \mathbf{1}_{\{u_t^{i,j}=\mu_t^{i,j}(h_t^0, \tilde{x}_{t-d+1:t-1}^{i,j}, x_t^{i,j})\}} \left(\prod_{(i,l)\in\mathcal{N}_i\setminus\{(i,j)\}} \varphi_t^{i,l}(u_t^{i,l}|h_t^{i,l}) \right) \\
&\quad \times \left(\prod_{(k,j)\in\mathcal{N}_{-i}} \mathbf{1}_{\{u_t^{k,j}=\mu_t^{k,j}(h_t^{k,j})\}} \right) \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, \tilde{x}_{t-d+1:t-1}^{i,j}, x_{t-d+1:t-1}^{-(i,j)}, h_t) \\
&= G_t^{i,j} \times \left(\prod_{(i,l)\in\mathcal{N}_i\setminus\{(i,j)\}} \varphi_t^{i,l}(u_t^{i,l}|h_t^{i,l}) \right) \left(\prod_{(k,j)\in\mathcal{N}_{-i}} \mathbf{1}_{\{u_t^{k,j}=\mu_t^{k,j}(h_t^{k,j})\}} \right) \\
&\quad \times \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t)
\end{aligned}
$$

where

$$
\begin{aligned}
G_t^{i,j} := \sum_{\tilde{x}_{t-d+1:t-1}^{i,j}} \Big[&\mathbf{1}_{\{u_t^{i,j}=\mu_t^{i,j}(h_t^0, \tilde{x}_{t-d+1:t-1}^{i,j}, x_t^{i,j})\}} \\
&\times \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(\tilde{x}_{t-d+1:t-1}^{i,j} | x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) \Big].
\end{aligned}
$$

From Lemmas 3 and 11, we know that

$$
\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(\tilde{x}_{t-d+1:t-1}^{i,j} | x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) = \mathbb{P}^{\mu^{i,j}}(\tilde{x}_{t-d+1:t-1}^{i,j} | x_t^{i,j}, h_t^i)
$$

for all $(x_t^{i,j}, h_t^i)$ admissible under $\mu^{i,j}$. Note that $\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) = 0$ for $(x_t^{i,j}, h_t^i)$ not admissible under $\mu^{i,j}$.

Hence, we conclude that

$$
\begin{aligned}
G_t^{i,j} &= \sum_{\tilde{x}_{t-d+1:t-1}^{i,j}} \mathbf{1}_{\{u_t^{i,j}=\mu_t^{i,j}(h_t^0, \tilde{x}_{t-d+1:t-1}^{i,j}, x_t^{i,j})\}} \mathbb{P}^{\mu^{i,j}}(\tilde{x}_{t-d+1:t-1}^{i,j} | x_t^{i,j}, h_t^i) \\
&= \bar{\varphi}_t^{i,j}(u_t^{i,j} | h_t^i, x_t^{i,j})
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(u_t, x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) \\
&= \bar{\varphi}_t^{i,j}(u_t^{i,j} | h_t^i, x_t^{i,j}) \left(\prod_{(i,l)\in\mathcal{N}_i\setminus\{(i,j)\}} \varphi_t^{i,l}(u_t^{i,l}|h_t^{i,l}) \right) \left(\prod_{(k,j)\in\mathcal{N}_{-i}} \mathbf{1}_{\{u_t^{k,j}=\mu_t^{k,j}(h_t^{k,j})\}} \right) \\
&\quad \times \mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t).
\end{aligned}
$$

Similarly, we have

$$
\mathbb{P}^{\bar{\varphi}^{i,j},\varphi_t^{i,-j},\mu^{-i}}(u_t, x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t)
$$

$$= \bar{\varphi}_t^{i,j}(u_t^{i,j}|h_t^i, x_t^{i,j}) \left(\prod_{(i,l)\in\mathcal{N}_i\setminus\{(i,j)\}} \varphi_t^{i,l}(u_t^{i,l}|h_t^{i,l}) \right) \left(\prod_{(k,j)\in\mathcal{N}_{-i}} \mathbf{1}_{\{u_t^{k,j}=\mu_t^{k,j}(h_t^{k,j})\}} \right)$$
$$\times \mathbb{P}^{\bar{\varphi}^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t).$$

Hence, it suffices to prove that

$$\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t) = \mathbb{P}^{\bar{\varphi}^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t).$$

Given the induction hypothesis, it suffices to show that

$$\begin{aligned} &\mathbb{P}^{\mu^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t|u_{t-1}, x_{t-1}, x_{t-d:t-2}^{-(i,j)}, h_{t-1}) \\ =&\mathbb{P}^{\bar{\varphi}^{i,j},\varphi_t^{i,-j},\mu^{-i}}(x_t, x_{t-d+1:t-1}^{-(i,j)}, h_t|u_{t-1}, x_{t-1}, x_{t-d:t-2}^{-(i,j)}, h_{t-1}) \end{aligned} \tag{35}$$

for all $(x_{t-1}, x_{t-d:t-2}^{-(i,j)}, h_{t-1})$ admissible under $(\mu^{i,j}, \varphi_t^{i,-j}, \mu^{-i})$ (or admissible under $(\bar{\varphi}^{i,j}, \varphi_t^{i,-j}, \mu^{-i})$, which is the same condition due to the induction hypothesis).

Since

$$\begin{aligned} \mathbf{X}_t^k &= f_{t-1}^k(\mathbf{X}_{t-1}^k, \mathbf{U}_t, W_t^{k,X}), \quad k \in \mathcal{I} \\ H_t &= (H_{t-1}, \mathbf{Y}_{t-1}, \mathbf{U}_{t-1}), \\ \mathbf{Y}_{t-1}^k &= \ell_{t-1}^k(\mathbf{X}_{t-1}^k, \mathbf{U}_{t-1}, W_{t-1}^{k,Y}), \quad k \in \mathcal{I} \end{aligned}$$

we have $(\mathbf{X}_t, \mathbf{X}_{t-d+1:t-1}^{-(i,j)}, H_t)$ to be a strategy-independent function of the random vector $(\mathbf{U}_{t-1}, \mathbf{X}_{t-1}, \mathbf{X}_{t-d:t-2}^{-(i,j)}, H_{t-1}, \mathbf{W}_{t-1}^X, \mathbf{W}_{t-1}^Y)$, where $(\mathbf{W}_{t-1}^X, \mathbf{W}_{t-1}^Y)$ is a primitive random vector independent of $(\mathbf{U}_{t-1}, \mathbf{X}_{t-1}, \mathbf{X}_{t-d:t-2}^{-(i,j)}, H_{t-1})$. Therefore, (35) is true and we established the induction step. □

Proof of Lemma 13 Through iterative application of Lemma 12, we conclude that for every pure strategy μ^i, there exist a payoff-equivalent behavioral strategy profile $\bar{\varphi}^i = (\bar{\varphi}_t^{i,j})_{(i,j)\in\mathcal{N}_i, t\in\mathcal{T}}$, where $\bar{\varphi}_t^{i,j} : \mathcal{H}_t^i \times \mathcal{X}_t^{i,j} \mapsto \Delta(\mathcal{U}_t^{i,j})$. Define $\bar{g}^i$ by

$$\bar{g}_t^i(\gamma_t^i|h_t^i) = \begin{cases} \prod_{(i,j)\in\mathcal{N}_i} \prod_{x_t^{i,j}} \bar{\varphi}_t^{i,j}(\gamma_t^{i,j}(x_t^{i,j})|h_t^i, x_t^{i,j}) & \gamma_t^i \in \bar{\mathcal{A}}_t^i \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{\mathcal{A}}_t^i \subset \mathcal{A}_t^i$ is the set of simple prescriptions. Then, using arguments similar to those in the proof of Lemma 1 one can show that $\bar{g}^i$ is payoff-equivalent to $\bar{\varphi}^i$, and hence payoff-equivalent to μ^i. □

References

1. Amin S, Litrico X, Sastry S, Bayen AM (2013) Cyber security of water SCADA systems—Part I: analysis and experimentation of stealthy deception attacks. IEEE Trans Control Syst Technol 21(5):1963–1970. https://doi.org/10.1109/tcst.2012.2211873
2. Amin S, Schwartz GA, Cárdenas AA, Sastry SS (2015) Game-theoretic models of electricity theft detection in smart utility networks: providing new capabilities with advanced metering infrastructure. IEEE Control Syst Mag 35(1):66–81. https://doi.org/10.1109/mcs.2014.2364711
3. Anantharam V, Borkar V (2007) Common randomness and distributed control: a counterexample. Syst Control Lett 56(7–8):568–572. https://doi.org/10.1016/j.sysconle.2007.03.010
4. Başar T, Olsder GJ (1999) Dynamic noncooperative game theory, vol 23. SIAM, Philadelphia

5. Bergemann D, Välimäki J (2006) Dynamic price competition. J Econ Theory 127(1):232–263. https://doi.org/10.1016/j.jet.2005.01.002
6. Bhattacharya S, Başar T (2012) 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, pp 5774–5779. https://doi.org/10.1109/cdc.2012.6426411
7. Cabral L (2011) Dynamic price competition with network effects. Rev Econ Stud 78(1):83–111. https://doi.org/10.1093/restud/rdq007
8. Cardaliaguet P, Rainer C, Rosenberg D, Vieille N (2016) Markov games with frequent actions and incomplete information-the limit case. Math Oper Res 41(1):49–71. https://doi.org/10.1287/moor.2015.0715
9. Colombino M, Smith RS, Summers TH (2017) Mutually quadratically invariant information structures in two-team stochastic dynamic games. IEEE Trans Autom Control 63(7):2256–2263. https://doi.org/10.1109/tac.2017.2772020
10. Cooper DJ, Kagel JH (2005) Are two heads better than one? Team versus individual play in signaling games. Am Econ Rev 95(3):477–509. https://doi.org/10.1257/0002828054201431
11. Cox CA, Stoddard B (2018) Strategic thinking in public goods games with teams. J Public Econ 161:31–43. https://doi.org/10.1016/j.jpubeco.2018.03.007
12. Doganoglu T (2003) Dynamic price competition with consumption externalities. NETNOMICS Econ Res Electron Netw 5(1):43–69. https://doi.org/10.1023/A:1024994117734
13. Farina G, Celli A, Gatti N, Sandholm T (2018) Ex-ante coordination and collusion in zero-sum multiplayer extensive-form games. In: Conference on neural information processing systems (NIPS)
14. Filar J, Vrieze K (2012) Competitive Markov decision processes. Springer, New York
15. Gensbittel F, Renault J (2015) The value of Markov chain games with incomplete information on both sides. Math Oper Res 40(4):820–841. https://doi.org/10.1287/moor.2014.0697
16. Gupta A, Nayyar A, Langbort C, Başar T (2014) Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric information. SIAM J Control Optim 52(5):3228–3260. https://doi.org/10.1137/140953514
17. Gupta A, Langbort C, Başar T (2016) Dynamic games with asymmetric information and resource constrained players with applications to security of cyberphysical systems. IEEE Trans Control Netw Syst 4(1):71–81. https://doi.org/10.1109/tcns.2016.2584183
18. Hancock PA, Nourbakhsh I, Stewart J (2019) On the future of transportation in an era of automated and autonomous vehicles. Proc Natl Acad Sci 116(16):7684–7691. https://doi.org/10.1073/pnas.1805770115
19. Harbert T (2014) Radio wrestlers fight it out at the DARPA Spectrum Challenge. https://spectrum.ieee.org/telecom/wireless/radio-wrestlers-fight-it-out-at-the-darpa-spectrum-challenge
20. Ho YC (1980) Team decision theory and information structures. Proc IEEE 68(6):644–654. https://doi.org/10.1109/proc.1980.11718
21. Kartik D, Nayyar A (2020) Upper and lower values in zero-sum stochastic games with asymmetric information. Dyn Games Appl 1–26. https://doi.org/10.1007/s13235-020-00364-x
22. Kartik D, Nayyar A, Mitra U (2021) Common information belief based dynamic programs for stochastic zero-sum games with competing teams. arXiv preprint arXiv:2102.05838
23. Kaspi Y, Merhav N (2010) Structure theorem for real-time variable-rate lossy source encoders and memory-limited decoders with side information. In: 2010 IEEE international symposium on information theory (ISIT), pp 86–90. https://doi.org/10.1109/isit.2010.5513283
24. Kuhn H (1953) Extensive games and the problem of information. In: Contributions to the theory of games (AM-28), volume II. Princeton University Press, pp 193–216. https://doi.org/10.1515/9781400881970-012
25. Kumar PR, Varaiya P (1986) Stochastic systems: estimation, identification and adaptive control. Prentice-Hall, Inc, Englewood Cliffs
26. Li L, Shamma J (2014) LP formulation of asymmetric zero-sum stochastic games. In: 2014 53rd IEEE conference on decision and control. IEEE, pp 1930–1935. https://doi.org/10.1109/cdc.2014.7039680
27. Li L, Langbort C, Shamma J (2019) An LP approach for solving two-player zero-sum repeated Bayesian games. IEEE Trans Autom Control 64(9):3716–3731. https://doi.org/10.1109/tac.2018.2885644
28. Mahajan A (2008) Sequential decomposition of sequential dynamic teams: Applications to real-time communication and networked control systems. PhD thesis, University of Michigan, Ann Arbor
29. Mahajan A (2013) Optimal decentralized control of coupled subsystems with control sharing. IEEE Trans Autom Control 58(9):2377–2382. https://doi.org/10.1109/cdc.2011.6160970
30. Mahajan A, Teneketzis D (2009) Optimal performance of networked control systems with nonclassical information structures. SIAM J Control Optim 48(3):1377–1404. https://doi.org/10.1137/060678130
31. Mailath GJ, Samuelson L (2006) Repeated games and reputations: long-run relationships. Oxford University Press, Oxford

32. Maskin E, Tirole J (1988) A theory of dynamic oligopoly. I: Overview and quantity competition with large fixed costs. Econom J Econom Soc 549–569. https://doi.org/10.2307/1911700
33. Maskin E, Tirole J (1988) A theory of dynamic oligopoly. II: Price competition, kinked demand curves, and edgeworth cycles. Econom J Econom Soc. https://doi.org/10.2307/1911701
34. Maskin E, Tirole J (2001) Markov perfect equilibrium: I. observable actions. J Econ Theory 100(2):191–219. https://doi.org/10.1006/jeth.2000.2785
35. Maskin E, Tirole J (2013) Markov equilibrium. In: J. F. Mertens memorial conference. https://youtu.be/UNtLnKJzrhs
36. Myerson RB (2013) Game theory. Harvard University Press, Harvard
37. Nayyar A, Başar T (2012) Dynamic stochastic games with asymmetric information. In: 2012 IEEE 51st IEEE conference on decision and control (CDC). IEEE, pp 7145–7150. https://doi.org/10.1109/cdc.2012.6426857
38. Nayyar A, Teneketzis D (2011) On the structure of real-time encoding and decoding functions in a multiterminal communication system. IEEE Trans Inf Theory 57(9):6196–6214. https://doi.org/10.1109/tit.2011.2161915
39. Nayyar A, Teneketzis D (2011) Sequential problems in decentralized detection with communication. IEEE Trans Inf Theory 57(8):5410–5435. https://doi.org/10.1109/tit.2011.2158478
40. Nayyar A, Teneketzis D (2019) Common knowledge and sequential team problems. IEEE Trans Autom Control 64(12):5108–5115. https://doi.org/10.1109/tac.2019.2912536
41. Nayyar A, Mahajan A, Teneketzis D (2011) Optimal control strategies in delayed sharing information structures. IEEE Trans Autom Control 56(7):1606–1620. https://doi.org/10.1109/tac.2010.2089381
42. Nayyar A, Gupta A, Langbort C, Başar T (2013) Common information based Markov perfect equilibria for stochastic games with asymmetric information: finite games. IEEE Trans Autom Control 59(3):555–570. https://doi.org/10.1109/tac.2013.2283743
43. Nayyar A, Mahajan A, Teneketzis D (2013) Decentralized stochastic control with partial history sharing: a common information approach. IEEE Trans Autom Control 58(7):1644–1658. https://doi.org/10.1109/tac.2013.2239000
44. Ouyang Y, Tavafoghi H, Teneketzis D (2015) Dynamic oligopoly games with private Markovian dynamics. In: 2015 54th IEEE conference on decision and control (CDC). IEEE, pp 5851–5858. https://doi.org/10.1109/cdc.2015.7403139
45. Ouyang Y, Tavafoghi H, Teneketzis D (2016) Dynamic games with asymmetric information: common information based perfect Bayesian equilibria and sequential decomposition. IEEE Trans Autom Control 62(1):222–237. https://doi.org/10.1109/tac.2016.2544936
46. Renault J (2006) The value of Markov chain games with lack of information on one side. Math Oper Res 31(3):490–512. https://doi.org/10.1287/moor.1060.0199
47. Renault J (2012) The value of repeated games with an informed controller. Math Oper Res 37(1):154–179. https://doi.org/10.1287/moor.1110.0518
48. Shelar D, Amin S (2017) Security assessment of electricity distribution networks under DER node compromises. IEEE Trans Control Netw Syst 4(1):23–36. https://doi.org/10.1109/tcns.2016.2598427
49. Summers T, Li C, Kamgarpour M (2017) Information structure design in team decision problems. IFAC-PapersOnLine 50(1):2530–2535. https://doi.org/10.1016/j.ifacol.2017.08.067
50. Tavafoghi H (2017) On design and analysis of cyber-physical systems with strategic agents. PhD thesis, University of Michigan, Ann Arbor
51. Tavafoghi H, Ouyang Y, Teneketzis D (2016) On stochastic dynamic games with delayed sharing information structure. In: 2016 IEEE 55th conference on decision and control (CDC). IEEE, pp 7002–7009. https://doi.org/10.1109/cdc.2016.7799348
52. Tavafoghi H, Ouyang Y, Teneketzis D, Wellman M (2019) Game theoretic approaches to cyber security: challenges, results, and open problems. In: Jajodia S, Cybenko G, Liu P, Wang C, Wellman M (eds) Adversarial and uncertain reasoning for adaptive cyber defense: control-and game-theoretic approaches to cyber security, vol 11830. Springer, New York, pp 29–53. https://doi.org/10.1007/978-3-030-30719-6_3
53. Tavafoghi H, Ouyang Y, Teneketzis D (March 2022) A unified approach to dynamic decision problems with asymmetric information: non-strategic agents. IEEE Trans Autom Control. https://doi.org/10.1109/tac.2021.3060835, to appear
54. Teneketzis D (2006) On the structure of optimal real-time encoders and decoders in noisy communication. IEEE Trans Inf Theory 52(9):4017–4035. https://doi.org/10.1109/tit.2006.880067
55. Teneketzis D, Ho YC (1987) The decentralized Wald problem. Inf Comput 73(1):23–44. https://doi.org/10.1016/0890-5401(87)90038-1
56. Teneketzis D, Varaiya P (1984) The decentralized quickest detection problem. IEEE Trans Autom Control 29(7):641–644. https://doi.org/10.1109/tac.1984.1103601

57. Tenney RR, Sandell NR (1981) Detection with distributed sensors. IEEE Trans Aerosp Electron Syst AES 17(4):501–510. https://doi.org/10.1109/taes.1981.309178
58. Tsitsiklis JN (1993) Decentralized detection. Adv Stat Signal Process 297–344
59. Varaiya P, Walrand J (1983) Causal coding and control for Markov chains. Syst Control Lett 3(4):189–192. https://doi.org/10.1016/0167-6911(83)90012-9
60. Vasal D, Sinha A, Anastasopoulos A (2019) A systematic process for evaluating structured perfect Bayesian equilibria in dynamic games with asymmetric information. IEEE Trans Autom Control 64(1):81–96. https://doi.org/10.1109/tac.2018.2809863
61. Veeravalli VV (2001) Decentralized quickest change detection. IEEE Trans Inf Theory 47(4):1657–1665. https://doi.org/10.1109/18.923755
62. Veeravalli VV, Başar T, Poor HV (1993) Decentralized sequential detection with a fusion center performing the sequential test. IEEE Trans Inf Theory 39(2):433–442. https://doi.org/10.1109/18.212274
63. Veeravalli VV, Başar T, Poor HV (1994) Decentralized sequential detection with sensors performing sequential tests. Math Control Signals Syst 7(4):292–305. https://doi.org/10.1007/bf01211521
64. Walrand J, Varaiya P (1983) Optimal causal coding-decoding problems. IEEE Trans Inf Theory 29(6):814–820. https://doi.org/10.1109/tit.1983.1056760
65. Witsenhausen HS (1973) A standard form for sequential stochastic control. Math Syst Theory 7(1):5–11. https://doi.org/10.1007/bf01824800
66. Witsenhausen HS (1979) On the structure of real-time source coders. Bell Syst Tech J 58(6):1437–1451. https://doi.org/10.1002/j.1538-7305.1979.tb02263.x
67. Yoshikawa T (1978) Decomposition of dynamic team decision problems. IEEE Trans Autom Control 23(4):627–632. https://doi.org/10.1109/tac.1978.1101791
68. Zhang Y, An B (2020) Computing team-maxmin equilibria in zero-sum multiplayer extensive-form games. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, no. 02, pp 2318–2325. https://doi.org/10.1609/aaai.v34i02.5610
69. Zheng J, Castañón DA (2013) Decomposition techniques for Markov zero-sum games with nested information. In: 2013 52nd IEEE conference on decision and control. IEEE, pp 574–581. https://doi.org/10.1109/cdc.2013.6759943
70. Zhu Q, Başar T (2015) Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: games-in-games principle for optimal cross-layer resilient control systems. IEEE Control Syst Mag 35(1):46–65. https://doi.org/10.1109/mcs.2014.2364710