

On Decentralized Minimax Control with Nested Subsystems

Aditya Dave, *Student Member, IEEE*, Nishanth Venkatesh, *Student Member, IEEE*,
Andreas A. Malikopoulos, *Senior Member, IEEE*

Abstract—In this paper, we investigate a decentralized control problem with nested subsystems, which is a general model for one-directional communication amongst many subsystems. The noises in our dynamics are modeled as uncertain variables which take values in finite sets. The objective is to minimize a worst-case shared cost. We demonstrate how the prescription approach can simplify the information structure and derive a structural form for optimal control strategies. The structural form allows us to restrict attention to control strategies whose domains do not grow in size with time, and thus, this form can be utilized in systems with long time horizons. Finally, we present a dynamic program to derive the optimal control strategies and validate our results with a numerical example.

I. INTRODUCTION

Modern engineering systems with interdependent subsystems and limited communication have directed attention towards decentralized control problems with cooperative agents. Such systems include social media platforms [1], large hierarchical organizations [2], and vehicle platoons [3]. Typically, in decentralized systems, different agents must coordinate their actions with access to different information. Such systems are characterized by their *information structure*, which describes the information available to each agent at each time [4]. The literature in decentralized control has predominantly focused on minimizing the expected cost shared by all agents [5]–[11], or a given norm of a shared cost in linear time-invariant systems [12]–[14], for specific information structures. In these problems, the system noises are modeled as random variables with known distributions.

More recently, there has been an interest in decentralized control problems with a worst-case cost [15], [16]. In these problems, we assume no knowledge of the distributions of the noises but consider that their feasible sets are known. Such noises are called *uncertain variables*, which are non-stochastic analogues of random variables [17], [18]. The objective is to minimize a maximum possible cost given only the feasible sets for all uncertain variables. This conservative measure of system performance is appropriate for (1) safety-critical applications that require concrete performance guarantees or (2) systems where only feasible sets can be determined a priori for all noises. For centralized worst-case problems, dynamic programs (DPs) can be used to compute the optimal control strategies [19], [20], but these DPs cannot directly be applied to decentralized problems

because of information asymmetry among agents [4]. This issue has been resolved for systems with partial history sharing among all agents using the common information approach. While this approach is valid for a large number of models [15], its computational tractability suffers when the private information of agents increases with time, because of the concurrent growth in domains of the optimal control strategies [7]. Increasing private information may occur in systems with one-directional communication [9], [21] like large organizations, where management may not share all observations with employees [2] and vehicle platoons [3]. However, with one-directional communication, the information available to a group of agents may be nested within the information available to another group of agents.

In this paper, we analyze a general decentralized system with *nested subsystems*, which allows one-directional communication among agents. We consider multiple subsystems with multiple agents, each of whom partially observes a global state. Agents sequentially share data with all agents within their subsystem. Thus, there is common information within each subsystem and private information available only to individual agents. In our information structure, we consider that the common information in each subsystem is also available to all agents in preceding subsystems, but the common information in preceding subsystems may contain data which is unavailable to agents in a given subsystem. For example, consider a problem with 3 subsystems. Then, the common information of all agents in subsystem 3 is a subset of the common information in subsystem 2, which in turn, is a subset of the common information in subsystem 1. Note that partial history sharing is a special case of this model which occurs when the problem has just one subsystem.

Our main contribution in this paper is the derivation of a structural form for optimal control strategies in decentralized minimax problems with nested subsystems (Theorem 4), which can be used to compute the strategies using a DP (Subsection IV-E). We model all noises using uncertain variables with known feasible sets. In our exposition, we use a combination of *the person-by-person approach* [5], and *the prescription approach* [8], [9] to iteratively derive time-invariant domains for optimal strategies. Our analysis is more involved than simply applying the common information approach, which sets up a centralized problem with information available to *all* agents in the system. However, in comparison with the common information approach: (1) our approach requires milder conditions on private information to achieve time invariant domains for optimal control strategies (Remark 6), and (2) thus, our DP improves the computational

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

The authors are with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716 USA (email: adidave@udel.edu; nish@udel.edu; andreas@udel.edu).

tractability of a larger class of systems. As it is standard in the literature [15], [19], we first analyze the terminal cost problem and then extend our results to allow additive costs.

The remainder of the paper proceeds as follows. In Section II, we formulate the problem. In Section III, we construct an equivalent problem with a simpler information structure. In Section IV, we analyze the problem and derive our main results, and in Section V, we present a numerical example validating our results. Finally, in Section VI, we present concluding remarks and discuss ongoing work.

II. PROBLEM FORMULATION

We consider a decentralized system with a set of subsystems $\mathcal{N} = \{1, \dots, N\}$, $N \in \mathbb{N}$. Each subsystem $n \in \mathcal{N}$ contains $K^n \in \mathbb{N}$ agents in a set $\mathcal{K}^n = \{1, \dots, K^n\}$, and each agent acts across $T \in \mathbb{N}$ discrete time steps. For each $t = 0, \dots, T$, the state of the system is denoted by an uncertain variable X_t which takes values in a finite set \mathcal{X}_t and the control action of each agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, is denoted by an uncertain variable $U_t^{k,n}$ which takes values in a finite set $\mathcal{U}_t^{k,n}$. Recall that uncertain variables are non-stochastic analogues of random variables with known feasible sets but unknown distributions [17]. We denote all actions in a subsystem $n \in \mathcal{N}$ collectively by $U_t^n := (U_t^{k,n} : k \in \mathcal{K}^n)$ and all actions in the system by $U_t^{1:N} := (U_t^1, \dots, U_t^N)$. Starting at X_0 , the state of the system evolves as

$$X_{t+1} = f_t(X_t, U_t^{1:N}, W_t), \quad t = 0, \dots, T-1, \quad (1)$$

where W_t is the disturbance of the system at time t which takes values in a finite set \mathcal{W}_t . At each $t = 0, \dots, T$, each agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, has an observation $Y_t^{k,n} := h_t^{k,n}(X_t, V_t^{k,n})$, which takes values in a finite set $\mathcal{Y}_t^{k,n}$. Here, $V_t^{k,n}$ is an observation noise which takes values in a finite set $\mathcal{V}_t^{k,n}$. We denote all noises in subsystem $n \in \mathbb{N}$ collectively by $V_t^n := (V_t^{k,n} : k \in \mathcal{K}^n)$ with a feasible set $\mathcal{V}_t^n := \prod_{k \in \mathcal{K}^n} \mathcal{V}_t^{k,n}$. The disturbances $\{W_t : t = 0, \dots, T\}$, measurement noises $\{V_t^{k,n} : k \in \mathcal{K}^n, n \in \mathcal{N}, t = 0, \dots, T\}$, and initial state X_0 are collectively called the *primitive variables*. Each primitive variable is independent from all other primitive variables. This ensures that the system is Markovian in a non-stochastic sense [17], [19]. Next, we describe the information structure of the system.

Definition 1. The *memory* of agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, at each t , is a set $M_t^{k,n} \subseteq \{Y_{0:t-1}^{i,m}, U_{0:t-1}^{i,m} : i \in \mathcal{K}^m, m \geq n\}$, which takes values in a finite collection of sets $\mathcal{M}_t^{k,n}$ and satisfies *perfect recall*, i.e., $M_t^{k,n} \subseteq M_{t+1}^{k,n}$ for all $t = 0, \dots, T-1$.

Remark 1. To be consistent with the exposition in the literature [5], [15], we consider that for all $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, and $t = 0, \dots, T$, the memory $M_t^{k,n}$ is updated before the observation $Y_t^{k,n}$ is realized.

In an information structure with nested subsystems, we partition the memory $M_t^{k,n}$ of each agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, into two components, C_t^n and $L_t^{k,n}$, with the following properties:

1) *The common information* of all agents in a subsystem $n \in \mathcal{N}$ at each $t = 0, \dots, T$ is the information available to all agents in \mathcal{K}^n . We define it as the set $C_t^n := \bigcap_{k \in \mathcal{K}^n} M_t^{k,n}$ which takes values in a finite collection of sets \mathcal{C}_t^n , and satisfies the following properties: (1) *nestedness*, i.e., $C_t^m \subseteq C_t^n$ for all $m \in \mathcal{N}$ with $m > n$ and $t = 0, \dots, T$, and (2) *perfect recall*, i.e., $C_t^n \subseteq C_{t+1}^n$ for all t . The property of nestedness implies that the common information of a subsystem is also available to all preceding subsystems.

2) *The private information* of an agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, at each $t = 0, \dots, T$ is the information available only to agent k . We define it as the set $L_t^{k,n} := M_t^{k,n} \setminus C_t^n$ which takes values in a finite collection of sets $\mathcal{L}_t^{k,n}$. For any pair of subsystems $n, m \in \mathcal{N}$ with $m > n$, we impose the condition that $L_t^{k,m} \cap C_t^n = \emptyset$ for all $k \in \mathcal{K}^m$, i.e., no element in the private information of an agent can be available to *all* agents of a preceding subsystem.

For each $n \in \mathcal{N}$, we also define the *new information* added to C_t^n at each $t = 0, \dots, T$ as a set $Z_t^n := C_t^n \setminus C_{t-1}^n$, which takes values in a finite collection of sets \mathcal{Z}_t^n , where $C_{-1}^n := \emptyset$. The new information is also nested, i.e., $Z_t^m \subseteq Z_t^n$, for all $m, n \in \mathcal{N}$, $m > n$, and t . Thus, the new information available to any subsystem is a subset of the new information available to any preceding subsystem at each t . As an illustration, consider a system with three agents divided into two subsystems in Fig. 1. Subsystem 1 contains two agents and subsystem 2 contains one agent. The red dashed arrows indicate the communication between the agents at time t . The one-directional, dashed arrow indicates sharing of common information from subsystem 2 to all agents in subsystem 1 at time t , which ensures that $C_t^2 \subseteq C_t^1$.

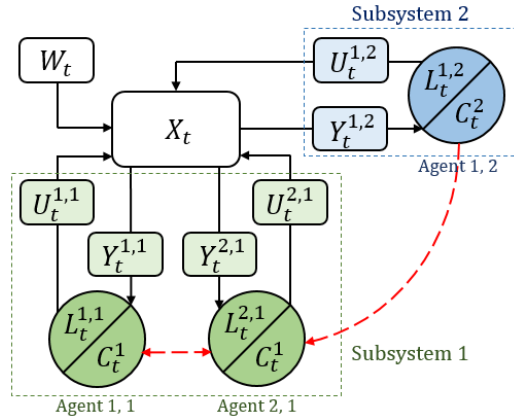


Fig. 1. A system with three agents in two nested subsystems at time t

After updating their memory and realizing their observations at each $t = 0, \dots, T$, each agent $k \in \mathcal{K}^n$, $n \in \mathcal{N}$, selects an action

$$U_t^{k,n} = g_t^{k,n}(Y_t^{k,n}, M_t^{k,n}) = g_t^{k,n}(Y_t^{k,n}, L_t^{k,n}, C_t^n), \quad (2)$$

where $M_t^{k,n} = \{L_t^{k,n}, C_t^n\}$, and $g_t^{k,n} : \mathcal{Y}_t^{k,n} \times \mathcal{M}_t^{k,n} \rightarrow \mathcal{U}_t^{k,n}$ is the *control law* of agent k in subsystem n at time t . The *control strategy* of subsystem $n \in \mathcal{N}$ is $\mathbf{g}^n := (g_{0:T}^{k,n} : k \in \mathcal{K}^n)$ and the control strategy of the system

is $\mathbf{g} := (\mathbf{g}^1, \dots, \mathbf{g}^N)$. The set of all feasible strategy profiles is \mathcal{G} . All agents collectively incur a shared terminal cost $d_T(X_T) \in \mathbb{R}_{\geq 0}$ at the time horizon T . The system performance is measured by the worst-case terminal cost

$$\mathcal{J}(\mathbf{g}) := \max_{X_0, W_{0:T}, V_{0:T}^{1:N}} d_T(X_T). \quad (3)$$

Problem 1. The optimization problem is $\min_{\mathbf{g} \in \mathcal{G}} \mathcal{J}(\mathbf{g})$, given the feasible sets $\{\mathcal{X}_0, \mathcal{W}_t, \mathcal{V}_t^n : n \in \mathcal{N}, t = 0, \dots, T\}$, the cost function d_T , and the system dynamics $\{f_t, h_t^{k,n} : k \in \mathcal{K}^n, n \in \mathcal{N}, t = 0, \dots, T\}$.

Our aim is to develop a DP which can derive an optimal strategy profile $\mathbf{g}^* \in \mathcal{G}$ for Problem 1. Note that an optimal strategy is guaranteed to exist because all variables take values in finite sets.

Remark 2. Variations of many existing information structures are special cases of our model. For example, a system with partial history sharing [7], [15], emerges from our model when all agents belong to a single subsystem, i.e., we set $N = 1$. Similarly, a system with one-directional delayed communication among a chain of $K \in \mathbb{N}$ agents is the special case where we have one agent per subsystem, i.e., $N = K$ and $|\mathcal{K}^n| = 1$ for all $n \in \mathcal{N}$. Simpler cases of instantaneous one-directional communication among just two agents were considered in [9], [21] for stochastic control.

III. EQUIVALENT SYSTEM WITH NESTED INFORMATION

In this section, we construct an equivalent system with a simpler information structure, where each subsystem acts as the decision maker. For example, consider two distinct agents $k, i \in \mathcal{K}^n$ in a subsystem $n \in \mathcal{N}$. In general, given a strategy \mathbf{g} , agent k cannot generate the action $U_t^{i,n}$ of agent i since they cannot access the private information $L_t^{i,n}$. However, all agents in subsystem $n \in \mathcal{N}$ can access the common information C_t^n . This motivates us to consider a two-stage mechanism for action generation: (1) each subsystem n generates a *partial action* for each agent in \mathcal{K}^n using only the common information C_t^n , and then (2) each agent $k \in \mathcal{K}^n$ uses the partial action along with their private information $L_t^{k,n}$ to generate their action U_t^k .

Definition 2. A *partial action* for agent $k \in \mathcal{K}^n, n \in \mathcal{N}$, at each $t = 0, \dots, T$ is a mapping $\hat{U}_t^{k,n} : \mathcal{Y}_t^{k,n} \times \mathcal{L}_t^{k,n} \rightarrow \mathcal{U}_t^{k,n}$ where $\hat{\mathcal{U}}_t^{k,n}$ is a finite set.

At each $t = 0, \dots, T$, a subsystem $n \in \mathcal{N}$ generates a partial action for an agent $k \in \mathcal{K}^n$ using a partial control law $\hat{g}_t^{k,n} : C_t^n \rightarrow \hat{\mathcal{U}}_t^{k,n}$, which yields $\hat{U}_t^{k,n} = \hat{g}_t^{k,n}(C_t^n)$. We define the partial control law for subsystem n at time t as $\hat{g}_t^n := (\hat{g}_t^{k,n} : k \in \mathcal{K}^n)$ and the partial control strategy for the system as $\hat{\mathbf{g}} := (\hat{g}_{0:T}^1, \dots, \hat{g}_{0:T}^N)$. To simplify the notation, we also define $\hat{U}_t^n := (\hat{U}_t^{k,n} : k \in \mathcal{K}^n)$ and compactly write $\hat{U}_t^n = \hat{g}_t^n(C_t^n)$ for each $t = 0, \dots, T$. The partial action \hat{U}_t^n is generated using only the common information in subsystem n and thus is available to all agents in \mathcal{K}^n . Furthermore, given a partial action \hat{U}_t^n , each agent $k \in \mathcal{K}^n$ must generate their action as $U_t^k = \hat{U}_t^{k,n}(Y_t^{k,n}, L_t^{k,n})$.

Next, we construct a new state which is controlled only using the partial actions of all subsystems. For each $n \in \mathcal{N}$, let $L_t^n := \{Y_t^{k,n}, L_t^{k,n} : k \in \mathcal{K}^n\}$. Then, the new state is

$$\hat{X}_t := \{X_t, L_t^1, \dots, L_t^N\}, \quad t = 0, \dots, T, \quad (4)$$

and it takes values in a finite collection of sets $\hat{\mathcal{X}}_t$.

Lemma 1. For each $t = 0, \dots, T-1$, we can construct a state evolution function $\hat{f}_t(\cdot)$ and an observation rule $\hat{h}_t^n(\cdot), n \in \mathcal{N}$, such that $\hat{X}_{t+1} = \hat{f}_t(\hat{X}_t, \hat{U}_t^{1:N}, W_t, V_{t+1}^{1:N})$ and $Z_{t+1}^n = \hat{h}_t^n(\hat{X}_t, \hat{U}_t^{1:N})$, respectively. In addition, we can construct a terminal cost function $\hat{d}_T(\cdot)$ which yields a shared cost $\hat{d}_T(\hat{X}_T) = d_T(X_T)$.

Proof. We first show how $\hat{f}_t(\cdot)$ can be constructed at each $t = 0, \dots, T$ by establishing that each component in $\hat{X}_{t+1} = \{X_{t+1}, L_{t+1}^1, \dots, L_{t+1}^N\}$ can be written in terms of the variables in the RHS. Note that $X_{t+1} = f_t(X_t, U_t^{1:N}, W_t)$, where $U_t^{k,n} = \hat{U}_t^{k,n}(Y_t^{k,n}, L_t^{k,n})$ for all $k \in \mathcal{K}^n, n \in \mathcal{N}$. For each $n \in \mathcal{N}$, $L_{t+1}^n \subseteq \{L_t^m, Y_{t+1}^{k,m}, U_{t+1}^{k,m} : k \in \mathcal{K}^m, m \geq n\}$ and for each $k \in \mathcal{K}^n$, $Y_{t+1}^{k,n} = h_{t+1}^{k,n}(X_{t+1}, V_{t+1}^{k,n})$. Using these equations, each term in \hat{X}_{t+1} can be written using the variables in the RHS. The proof is complete by defining an appropriate function $\hat{f}_t(\cdot)$. The other functions can also be constructed in a similar manner. \square

Lemma 1 yields a “new” decentralized system with a shared state \hat{X}_t , action \hat{U}_t^n , and observation Z_t^n for each subsystem $n \in \mathcal{N}$ and $t = 0, \dots, T$. The cost is $\hat{d}_T(\hat{X}_T)$ and the performance criterion is $\hat{\mathcal{J}}(\hat{\mathbf{g}}) := \max_{X_0, W_{0:T}, V_{0:T}^{1:N}} \hat{d}_T(\hat{X}_T)$. This new decentralized system has a nested information structure [9] since at each t , a partial action \hat{U}_t^n is generated using only $C_t^n = Z_{0:t}^n$, while $Z_t^m \subseteq Z_t^n$ for all $n, m \in \mathcal{N}$ with $m > n$. Next, we define a new optimization problem and show that it is equivalent to Problem 1.

Problem 2. The new problem is $\min_{\hat{\mathbf{g}} \in \hat{\mathcal{G}}} \hat{\mathcal{J}}(\hat{\mathbf{g}})$, given the sets $\{\mathcal{X}_0, \mathcal{W}_t, \mathcal{V}_t^n : n \in \mathcal{N}, t = 0, \dots, T\}$, the cost function \hat{d}_T , and the dynamics $\{\hat{f}_t, \hat{h}_t^n : k \in \mathcal{K}^n, n \in \mathcal{N}, t = 0, \dots, T\}$.

Lemma 2. For any given control strategy \mathbf{g} , consider a partial control strategy $\hat{\mathbf{g}}$

$$\hat{g}_t^{k,n}(C_t^n)(\cdot, \cdot) := g_t^{k,n}(\cdot, \cdot, C_t^n), \quad t = 0, \dots, T, \quad (5)$$

for all $k \in \mathcal{K}^n, n \in \mathcal{N}$. Then, $\mathcal{J}(\mathbf{g}) = \hat{\mathcal{J}}(\hat{\mathbf{g}})$. Moreover, for any given partial control strategy $\hat{\mathbf{g}}$, we can construct a control strategy \mathbf{g} such that

$$g_t^{k,n}(\cdot, \cdot, C_t^n) := \hat{g}_t^{k,n}(C_t^n)(\cdot, \cdot), \quad t = 0, \dots, T, \quad (6)$$

for all $k \in \mathcal{K}^n, n \in \mathcal{N}$. Then, $\hat{\mathcal{J}}(\hat{\mathbf{g}}) = \mathcal{J}(\mathbf{g})$.

Proof. For the first part, given a control strategy \mathbf{g} and the partial control strategy $\hat{\mathbf{g}}$, the action at each t for any agent $k \in \mathcal{K}^n, n \in \mathcal{N}$, is $U_t^k = g_t^{k,n}(Y_t^{k,n}, L_t^{k,n}, C_t^n) = \hat{g}_t^{k,n}(C_t^n)(Y_t^{k,n}, L_t^{k,n})$. Thus, the control law and partial control law result in the same action U_t^k for any given $\{Y_t^{k,n}, L_t^{k,n}, C_t^n\}$. Subsequently, for any given realizations

of all primitive variables $\{X_0, W_{0:T}, V_{0:T}^{1:3}\}$, the two strategies yield the same terminal state X_T and cost $d_T(X_T)$. The proof of the second part follows using arguments similar to the first part. \square

Lemma 2 implies that given an optimal partial strategy \hat{g}^* for Problem 2, a control strategy g^* constructed by (6) is an optimal solution to Problem 1, with performance $\mathcal{J}(g^*)$. Similarly, an optimal strategy for Problem 1 yields an optimal solution to Problem 2. Thus, the two problems are equivalent.

Remark 3. It is easier to analyze the equivalent system with a nested information structure than the original system. In the next section, we use this new information structure to derive results for optimal partial strategies, and then use Lemma 2 to characterize the optimal control strategies.

IV. MAIN RESULTS

In this section, we derive a structural form for optimal partial strategies in Problem 2. To simplify the exposition, we carry out the detailed derivation for a system with only 3 subsystems. This illustrates all the key arguments required to prove the results for different cases. Later, in subsection IV-D, we present the results for N subsystems.

A. Analysis for Subsystem 1

In this subsection, we analyze the optimal partial strategy of subsystem 1 in a system with 3 subsystems, i.e., $\mathcal{N} = \{1, 2, 3\}$. We use the person-by-person approach and thus, arbitrarily fix the partial control strategies \hat{g}^2 and \hat{g}^3 , such that $\hat{U}_t^n = \hat{g}_t^n(C_t^n)$, for all $n = 2, 3$ and $t = 0, \dots, T$, and set up a centralized problem for subsystem 1. Recall that $C_t^n \subseteq C_t^1$, and thus, any agent in subsystem 1 can derive the partial action \hat{U}_t^n for all $n = 2, 3$ and $t = 0, \dots, T$. Next, we construct a centralized problem for subsystem 1 with a new state $S_t^1 := \{\hat{X}_t, C_t^2\}$ for each $t = 0, \dots, T$, which takes values in a finite collection of sets \mathcal{S}_t^1 .

Lemma 3. *For any given partial strategies \hat{g}^2 and \hat{g}^3 , we can construct a state evolution function $\hat{f}_t^1(\cdot)$ and an observation rule $\hat{h}_t^1(\cdot)$, such that $S_{t+1}^1 = \hat{f}_t^1(S_t^1, \hat{U}_t^1, W_t, V_{t+1}^{1:3})$, and $Z_{t+1}^1 = \hat{h}_t^1(S_t^1, \hat{U}_t^1)$, respectively, for all $t = 0, \dots, T-1$. In addition, we can construct a terminal cost function $\hat{d}_T^1(\cdot)$, which yields a shared cost $\hat{d}_T^1(S_T^1) = \hat{d}_T(\hat{X}_T)$.*

Proof. We first show how $\hat{f}_t^1(\cdot)$ can be constructed at each $t = 0, \dots, T$ by establishing that each component in $S_{t+1}^1 = \{\hat{X}_{t+1}, C_{t+1}^2\}$ can be written in terms of the variables in the RHS. Note that $\hat{X}_{t+1} = \hat{f}_t(\hat{X}_t, \hat{U}_t^1, \hat{g}_t^2(C_t^2), \hat{g}_t^3(C_t^3), W_t, V_{t+1}^{1:3})$, where $\hat{g}_t^2(\cdot)$ and $\hat{g}_t^3(\cdot)$ are given and $C_t^2 \subseteq C_t^1$. Furthermore, $C_{t+1}^2 = C_t^2 \cup Z_{t+1}^2$, where $Z_{t+1}^2 = \hat{h}_t^2(\hat{X}_t, \hat{U}_t^1, \hat{g}_t^2(C_t^2), \hat{g}_t^3(C_t^3))$. Thus, given \hat{g}^2 and \hat{g}^3 , each term in S_{t+1}^1 can be written as a function of the variables in the argument of the RHS. The proof is complete by defining an appropriate function $\hat{f}_t^1(\cdot)$. The other functions can be constructed similarly by rewriting the LHS as a function of the variables in the argument in the RHS. \square

Given the partial control strategies \hat{g}^2 and \hat{g}^3 , Lemma 3 yields a new *centralized* system with state S_t^1 , action \hat{U}_t^1 , and observation Z_t^1 for all $t = 0, \dots, T$. The terminal cost incurred by the system is $\hat{d}_T^1(S_T^1)$ and the worst-case performance criterion is $\mathcal{J}^1(\hat{g}^1) := \max_{X_0, W_{0:T}, V_{0:T}^{1:3}} \hat{d}_T^1(S_T^1)$.

Problem 3. The problem for subsystem 1 is $\min_{\hat{g}^1} \mathcal{J}^1(\hat{g}^1)$, given the control partial strategies \hat{g}^2 and \hat{g}^3 , the feasible sets $\{\mathcal{X}_0, \mathcal{W}_t, \mathcal{V}_t^n : n = 1, 2, 3, t = 0, \dots, T\}$, the cost function \hat{d}_T^1 , and the dynamics $\{\hat{f}_t^1, \hat{h}_t^1 : t = 0, \dots, T\}$.

Remark 4. Using Lemma 3, we construct a cost $\hat{d}_T^1(S_T^1)$ for the centralized problem such that $\hat{d}_T^1(S_T^1) = \hat{d}_T(\hat{X}_T)$ for all realizations of the primitive variables $\{X_0, W_{0:T}, V_{0:T}^{1:3}\}$. Thus, the performance $\mathcal{J}^1(\hat{g}^1)$ for any given \hat{g}^2 and \hat{g}^3 is equal to $\mathcal{J}(\hat{g}^1, \hat{g}^2, \hat{g}^3)$.

Note that Problem 3 is a partially observed centralized problem with subsystem 1 as the sole decision maker. Specifically, at each $t = 0, \dots, T$, the component C_t^2 of the state S_t^1 is completely observed by subsystem 1, whereas, the component \hat{X}_t must be estimated using the common information C_t^1 and history of partial actions $\hat{U}_{0:t-1}^{1:3}$. For partially observed centralized problems, e.g., Problem 3, a DP was provided in [19] and its complexity was reduced using a *sufficiently informative function*. We simply call this an *information state* to be consistent with its stochastic counterpart. In minimax problems, an information state at time t is a set of feasible values for the partially observed state which are compatible with the information available to a subsystem at time t . Since subsystem 1 completely observes C_t^2 , we need to define an information state only for the component \hat{X}_t . For all $t = 0, \dots, T$, given the realizations c_t^1 and $\hat{u}_{0:t-1}^{1:3}$ of C_t^1 and $\hat{U}_{0:t-1}^{1:3}$, respectively, the realized information state of subsystem 1 is the set

$$P_t^1 := \{\hat{x}_t \in \hat{\mathcal{X}}_t \mid \exists (x_0 \in \mathcal{X}_0, w_{0:t-1} \in \prod_{\ell=0}^{t-1} \mathcal{W}_\ell, v_{0:t}^n \in \prod_{\ell=0}^t \mathcal{V}_\ell^n, n = 1, 2, 3) \text{ such that } s_{\ell+1}^1 = \hat{f}_\ell^1(s_\ell^1, \hat{u}_\ell^1, w_\ell, v_{\ell+1}^{1:3}), \\ z_{\ell+1}^1 = \hat{h}_\ell^1(s_\ell^1, \hat{u}_\ell^1), \ell = 0, \dots, t-1\}. \quad (7)$$

The information state each t is a function of uncertain variables and thus, in general, we denote it with the set-valued uncertain variable Π_t^1 which takes values in a finite collection of feasible sets $\mathcal{P}_t^1 \subseteq 2^{\hat{\mathcal{X}}_t}$, where $2^{\hat{\mathcal{X}}_t}$ denotes the power set of $\hat{\mathcal{X}}_t$. Next, we show that the information state Π_t^1 does not depend on the choice of partial strategy \hat{g} .

Lemma 4. *At each $t = 0, \dots, T-1$, there exists a function $\hat{f}_t^1(\cdot)$ independent of \hat{g} , such that*

$$\Pi_{t+1}^1 = \hat{f}_t^1(\Pi_t^1, \hat{U}_{t+1}^{1:3}, Z_{t+1}^1). \quad (8)$$

Proof. The proof follows the same arguments as the ones of Lemma 5 in Section IV-B. \square

Lemma 4 establishes the Markovian and strategy independent evolution of Π_t^1 at each t . Thus, we state the following result for Problem 3 using the centralized DP in [19].

Theorem 1. For any given partial strategies \hat{g}^2 and \hat{g}^3 of subsystems 2 and 3, respectively, in Problem 3, without loss of optimality, we can restrict attention to partial strategies \hat{g}^{*1} with the structural form

$$\hat{U}_t^1 = \hat{g}_t^{*1}(\Pi_t^1, C_t^2), \quad t = 0, \dots, T. \quad (9)$$

Proof. This result follows from standard arguments for centralized minimax control problems in [19]. \square

Theorem 1 establishes a structural form for an optimal partial strategy \hat{g}^{*1} of subsystem 1 in Problem 3, which holds for all \hat{g}^2 and \hat{g}^3 . From Remark 4, we note that any globally optimal partial strategy $(\hat{g}^{*1}, \hat{g}^{*2}, \hat{g}^{*3})$ for Problem 2, must necessarily be a person-by-person optimal solution of Problem 3, i.e., after fixing \hat{g}^{*2} and \hat{g}^{*3} for subsystems 2 and 3, respectively, the strategy \hat{g}^{*1} is optimal for Problem 3. Thus, there exists a globally optimal partial strategy $(\hat{g}^{*1}, \hat{g}^{*2}, \hat{g}^{*3})$ where \hat{g}^{*1} takes the structural form in (9).

B. Analysis for Subsystem 2

In this subsection, we restrict attention to partial control strategies \hat{g}^1 of agent 1 which satisfy (9), and derive a structural form for the partial strategy of subsystem 2. Note that given \hat{g}^1 , agents in subsystem 2 cannot generate \hat{U}_t^1 at any t because they cannot access Π_t^1 . Thus, before using the person-by-person approach for subsystem 2, we consider that \hat{U}_t^1 is generated in two stages for all $t = 0, \dots, T$: (1) subsystem 2 generates a *prescription* for subsystem 1 using only C_t^2 , and then (2) subsystem 1 uses the prescription and its information state Π_t^1 to generate its partial action \hat{U}_t^1 .

Definition 3. A *prescription* by subsystem 2 for subsystem 1 at any $t = 0, \dots, T$ is a mapping $\Gamma_t^{[2,1]} : \mathcal{P}_t^1 \rightarrow \mathcal{U}_t^1$ which takes values in a finite set $\mathcal{F}_t^{[2,1]}$.

At each $t = 0, \dots, T$, subsystem 2 generates a prescription for subsystem 1 using a prescription law $\psi_t^{[2,1]} : C_t^2 \rightarrow \mathcal{F}_t^{[2,1]}$, which yields $\Gamma_t^{[2,1]} = \psi_t^{[2,1]}(C_t^2)$. We define the prescription strategy for subsystem 2 as $\psi^2 := (\psi_t^{[2,1]} : t = 0, \dots, T)$. Since the prescription $\Gamma_t^{[2,1]}$ is generated only using $C_t^2 \subseteq C_t^1$, it is available to both subsystems 1 and 2. Then, subsystem 1 must generate its partial action as $\hat{U}_t^1 = \Gamma_t^{[2,1]}(\Pi_t^1)$. Note that in this formulation, the prescription $\Gamma_t^{[2,1]}$ is generated by subsystem 2 and subsystem 1 simply utilizes the given prescription to obtain \hat{U}_t^1 . Thus, we can consider that at each $t = 0, \dots, T$, subsystem 2 selects a *complete action* $\Theta_t^2 := (\Gamma_t^{[2,1]}, U_t^2)$. This motivates us to simultaneously derive a structural form for the optimal prescription and partial strategies of subsystem 2. To this end, we use the person-by-person approach by fixing the partial control strategy \hat{g}^3 of subsystem 3 and noting that subsystem 2 can derive \hat{U}_t^3 for all t because $C_t^3 \subseteq C_t^2$. Next, we construct a centralized problem for subsystem 2 with a state $S_t^2 := \{\hat{X}_t, \Pi_t^1, C_t^3\}$ at each t , which takes values in a finite collection of sets \mathcal{S}_t^2 . Using the same arguments as Lemma 3, we can construct a state evolution function $\tilde{f}_t^2(\cdot)$ such that $S_{t+1}^2 = \tilde{f}_t^2(S_t^2, \Theta_t^2, W_t, V_{t+1}^{1:3})$, and an observation rule $\tilde{h}_t^2(\cdot)$ which yields $Z_{t+1}^2 = \tilde{h}_t^2(S_t^2, \Theta_t^2)$ to obtain

a centralized problem with state S_t^2 , observation Z_t^2 and complete action $\Theta_t^2 = (\Gamma_t^{[2,1]}, U_t^2)$ for all t . Furthermore, we can construct a cost function $\tilde{d}_T^2(\cdot)$ which yields the terminal cost $\tilde{d}_T^2(S_T^2) = \hat{d}_T(\hat{X}_T)$ and a corresponding performance criterion $\mathcal{J}^2(\psi^2, \hat{g}^2) := \max_{X_0, W_{0:T}, V_{0:T}^{1:3}} \tilde{d}_T^2(S_T^2)$.

Problem 4. The problem for subsystem 2 is $\inf_{\psi^2, \hat{g}^2} \mathcal{J}^2(\psi^2, \hat{g}^2)$, given a partial control strategy \hat{g}^3 , the feasible sets $\{\mathcal{X}_0, \mathcal{W}_t, \mathcal{V}_t^n : n = 1, 2, 3, t = 0, \dots, T\}$, the cost function \tilde{d}_T^2 , and the dynamics $\{\tilde{f}_t^2, \tilde{h}_t^2 : t = 0, \dots, T\}$.

Remark 5. Consider a fixed partial control strategy \hat{g}^3 for agent 3. Using the same sequence of arguments as Lemma 2, for each prescription strategy ψ^2 , we can construct a partial control strategy \hat{g}^1 such that the performance criterion $\mathcal{J}(\hat{g}^1, \hat{g}^2, \hat{g}^3)$ for Problem 2 is equal to $\mathcal{J}^2(\psi^2, \hat{g}^2)$. Similarly, for each partial control strategy \hat{g}^1 we can construct a prescription strategy ψ^2 which yields the same performance. Thus, it is equivalent to select either a partial control strategy \hat{g}^1 or a prescription strategy ψ^2 .

Problem 4 is a partially observed centralized problem with subsystem 2 as the sole decision maker. Specifically, at each $t = 0, \dots, T$, the component C_t^3 in state S_t^2 is perfectly observed by subsystem 2, whereas \hat{X}_t and Π_t^1 must be estimated using an information state like (7). Given the realizations $c_t^2, \gamma_{0:t-1}^{[2,1]}$, and $\hat{u}_{0:t-1}^{2:3}$ of $C_t^2, \Gamma_{0:t-1}^{[2,1]}$, and $\hat{U}_{0:t-1}^{2:3}$, respectively, the realized information state of subsystem 2 is

$$P_t^2 := \{\hat{x}_t \in \hat{\mathcal{X}}_t, P_t^1 \in \mathcal{P}_t^1 \mid \exists (x_0 \in \mathcal{X}_0, w_{0:t-1} \in \prod_{\ell=0}^{t-1} \mathcal{W}_\ell, v_{0:t}^n \in \prod_{\ell=0}^t \mathcal{V}_\ell^n, n = 1, 2, 3) \text{ s.t. } s_{\ell+1}^2 = \tilde{f}_\ell^2(s_\ell^2, \theta_\ell^2, w_\ell, v_{\ell+1}^{1:3}), z_{\ell+1}^2 = \tilde{h}_\ell^2(s_\ell^2, \theta_\ell^2), \ell = 0, \dots, t-1\}, \quad (10)$$

Thus, the information state of subsystem 2 is a set-valued uncertain variable Π_t^2 which takes values in a finite collection of feasible sets $\mathcal{P}_t^2 \subseteq 2^{\hat{\mathcal{X}}_t} \times 2^{\mathcal{P}_t^1}$. Next, we show that the information state Π_t^2 is independent of the choice of \hat{g} .

Lemma 5. At each $t = 0, \dots, T-1$, there exists a function $\tilde{f}_t^2(\cdot)$ independent from \hat{g} , such that

$$\Pi_{t+1}^2 = \tilde{f}_t^2(\Pi_t^2, \Gamma_t^{[2,1]}, \hat{U}_t^{2:3}, Z_{t+1}^2). \quad (11)$$

Proof. Let $P_t^2 \in \mathcal{P}_t^2$ be a given set-valued realization of the information state Π_t^2 at time t . Then, for the realizations $\gamma_t^{[2,1]}, \hat{u}_t^{2:3}$ and z_{t+1}^2 of $\Gamma_t^{[2,1]}, \hat{U}_t^{2:3}$ and Z_{t+1}^2 , subsystem 2 can eliminate some possible elements in P_t^2 at the start of time $t+1$. Specifically, we define an interim set

$$Q_t^2 := \{(\hat{x}_t, P_t^1) \in P_t^2 \mid z_{t+1}^2 = \hat{h}_t^2(\hat{x}_t, \hat{u}_t^{1:3}), \hat{u}_t^1 = \gamma_t^{[2,1]}(\pi_t^1)\}, \quad (12)$$

which is completely determined given the set P_t^2 and the realizations $\{\gamma_t^{[2,1]}, \hat{u}_t^{2:3}, z_{t+1}^2\}$. Next, we derive the realization

P_{t+1}^2 of the information state Π_{t+1}^2 using set Q_t^2 as

$$\begin{aligned} P_{t+1}^2 &= \{\hat{x}_{t+1} \in \hat{\mathcal{X}}_{t+1}, P_{t+1}^1 \in \mathcal{P}_{t+1}^1 \mid \hat{x}_{t+1} = \hat{f}_t(\hat{x}_t, \\ &\hat{u}_t^{1:3}, w_t, v_{t+1}^{1:3}), P_{t+1}^1 = \hat{f}_t^1(P_t^1, \hat{u}_t^{1:3}, z_{t+1}^1), z_{t+1}^1 = \hat{h}_t^1(\hat{x}_t, \\ &\hat{u}_t^{1:3}), \hat{u}_t^1 = \gamma_t^{[2,1]}(P_t^1) \text{ for all } [(\hat{x}_t, P_t^1) \in Q_t^2, \\ &w_t \in \mathcal{W}_t, v_{t+1}^n \in \mathcal{V}_{t+1}^n, n = 1, 2, 3]\}. \end{aligned} \quad (13)$$

Thus, we can define an appropriate function $\tilde{f}_t^2(\cdot)$ using (13) such that that $P_{t+1}^2 = \tilde{f}_t^2(P_t^2, \gamma_t^{[2,1]}, \hat{u}_t^{2:3}, z_{t+1}^2)$. \square

Lemma 5 establishes that the evolution of Π_t^2 is Markovian. Thus, we can state the following result for Problem 4 using the standard centralized DP in [19].

Theorem 2. *For a given partial strategy \hat{g}^3 of subsystem 3 in Problem 4, without loss of optimality, we can restrict attention to prescription strategies ψ^{*2} and partial strategies \hat{g}^{*2} with the structural form*

$$\Gamma_t^{[2,1]} = \psi_t^{*[2,1]}(\Pi_t^2, C_t^3), \quad (14)$$

$$\hat{U}_t^2 = \hat{g}_t^{*2}(\Pi_t^2, C_t^3), \quad t = 0, \dots, T. \quad (15)$$

Proof. This result follows from standard arguments for centralized minimax control problems [19]. \square

Using an optimal prescription strategy which satisfies (14), we construct an optimal partial strategy \hat{g}^{*1} for Problem 2 such as $\hat{g}_t^{*1}(\cdot, \Pi_t^2, C_t^3) := \psi_t^{*[2,1]}(\Pi_t^2, C_t^3)(\cdot)$ and recall from Remark 5 that they yield the same performance. Thus, we can derive an optimal partial strategy \hat{g}^{*1} in the form

$$\hat{U}_t^1 = \hat{g}_t^{*1}(\Pi_t^{1:2}, C_t^3), \quad t = 0, \dots, T. \quad (16)$$

C. Analysis for Subsystem 3

In this subsection, we restrict attention to partial strategies \hat{g}^1 and \hat{g}^2 of the form in (16) and (15), respectively, and derive a structural form for the partial strategy of subsystem 3. Given \hat{g}^1 and \hat{g}^2 , subsystem 3 cannot generate the control actions $\hat{U}_t^{1:2}$ for any $t = 0, \dots, T$, because it cannot access the information states Π_t^1 and Π_t^2 . Instead, subsystem 3 generates a prescription for both subsystems 1 and 2 in a manner similar to Section III-B. Specifically, for each t , a prescription of subsystem 3 for subsystem 1 is a mapping $\Gamma_t^{[3,1]} : \mathcal{P}_t^1 \times \mathcal{P}_t^2 \rightarrow \hat{U}_t^1$ which takes values in a finite set $\mathcal{F}_t^{[3,1]}$ and that for subsystem 2 is a mapping $\Gamma_t^{[3,2]} : \mathcal{P}_t^2 \rightarrow \hat{U}_t^2$ which takes values in a finite set $\mathcal{F}_t^{[3,2]}$. At each $t = 0, \dots, T$, subsystem 3 generates a prescription for each $n = 1, 2$ using a prescription law $\psi_t^{[3,n]} : C_t^3 \rightarrow \mathcal{F}_t^{[3,n]}$, which yields $\Gamma_t^{[3,n]} = \psi_t^{[3,n]}(C_t^3)$. We define the prescription strategy of subsystem 3 as $\psi^3 := (\psi_t^{[3,n]} : n = 1, 2, t = 0, \dots, T)$. In this formulation, after receiving the prescriptions of subsystem 3, subsystem 1 generates its partial action as $\hat{U}_t^1 = \Gamma_t^{[3,1]}(\Pi_t^{1:2})$ and subsystem 2 generates its partial action as $\hat{U}_t^2 = \Gamma_t^{[3,2]}(\Pi_t^2)$. Thus, subsystem 3 is the sole decision maker with a complete action $\Theta_t^3 := (\Gamma_t^{[3,1]}, \Gamma_t^{[3,2]}, \hat{U}_t^3)$. Next, we construct a new state for subsystem 3 as $S_t^3 := \{\hat{X}_t, \Pi_t^1, \Pi_t^2\}$ for all t , which takes values in a finite collection of sets S_t^3 . Using the same

arguments as Lemma 3, we can construct a state evolution function $\bar{f}_t^3(\cdot)$ such that $S_{t+1}^3 = \bar{f}_t^3(S_t^3, \Theta_t^3, W_t, V_{t+1}^{1:3})$, and an observation rule $\bar{h}_t^3(\cdot)$ which yields $Z_{t+1}^3 = \bar{h}_t^3(S_t^3, \Theta_t^3)$ to obtain a centralized problem with state S_t^3 , observation Z_t^3 and complete action Θ_t^3 for all t . Furthermore, we can construct a cost function $\bar{d}_T^3(\cdot)$ which yields a terminal cost $\bar{d}_T^3(S_T^3) := \hat{d}_T(\hat{X}_T)$ and a performance criterion $\mathcal{J}^3(\psi^3, \hat{g}^3) := \max_{X_0, W_{0:T}, V_{0:T}^{1:3}} \bar{d}_T^3(S_T^3)$.

Problem 5. The problem for subsystem 3 is $\inf_{\psi^3, \hat{g}^3} \mathcal{J}^3(\psi^3, \hat{g}^3)$, given the feasible sets $\{\mathcal{X}_0, \mathcal{W}_t, \mathcal{V}_t^n : n = 1, 2, 3, t = 0, \dots, T\}$, the cost function \bar{d}_T^3 and the dynamics $\{\bar{f}_t^3, \bar{h}_t^3 : t = 0, \dots, T\}$.

Using the same arguments as in Lemma 2 and Remark 5, we conclude that we can generate partial actions \hat{U}_t^n , for $n = 1, 2$ and $t = 0, \dots, T$, equivalently using either a prescription strategy ψ^3 or an appropriate partial control strategy \hat{g}^n . In Problem 5, at each t , subsystem 3 must estimate the entire state S_t^3 . Thus, for the realizations c_t^3 and $\theta_{0:t-1}^3$ of C_t^3 and $\Theta_{0:t-1}^3$, respectively, the realized information state for subsystem 3 is

$$\begin{aligned} P_t^3 &:= \{\hat{x}_t \in \hat{\mathcal{X}}_t, P_t^1 \in \mathcal{P}_t^1, P_t^2 \in \mathcal{P}_t^2 \mid \exists (x_0 \in \mathcal{X}_0, w_{0:t-1} \\ &\in \prod_{\ell=0}^{t-1} \mathcal{W}_\ell, v_{0:t}^n \in \prod_{\ell=0}^t \mathcal{V}_\ell^n, n = 1, 2, 3) \text{ s.t. } s_{\ell+1}^3 = \bar{f}_\ell^3(s_\ell^3, \theta_\ell^3, \\ &w_\ell, v_{\ell+1}^{1:3}), z_{\ell+1}^3 = \bar{h}_\ell^3(s_\ell^3, \theta_\ell^3), \ell = 0, \dots, t-1\}. \end{aligned} \quad (17)$$

We define the information state of subsystem 3 as a set-valued uncertain variable which takes values in a finite collection of feasible sets $\mathcal{P}_t^3 \subseteq 2^{\hat{\mathcal{X}}_t} \times 2^{\mathcal{P}_t^1} \times 2^{\mathcal{P}_t^2}$. Next, we show that Π_t^2 is independent of \hat{g} .

Lemma 6. *At each $t = 0, \dots, T-1$, there exists a function $\tilde{f}_t^3(\cdot)$ independent from \hat{g} , such that*

$$\Pi_{t+1}^3 = \tilde{f}_t^3(\Pi_t^3, \Gamma_t^{[3,1]}, \Gamma_t^{[3,2]}, \hat{U}_t^3, Z_{t+1}^3). \quad (18)$$

Proof. The proof is the same as the proof of Lemma 5. Due to space constraints, it is omitted. \square

Lemma 6 establishes that the evolution of the Π_t^3 is Markovian. Thus, we can use the standard centralized DP decomposition from [19] to state the following result.

Theorem 3. *In Problem 5, without loss of optimality, we can restrict attention to prescription strategies ψ^{*3} and partial strategies \hat{g}^{*3} with the structural form*

$$\Gamma_t^{[3,n]} = \psi_t^{*[3,n]}(\Pi_t^3), \quad n = 1, 2, \quad (19)$$

$$\hat{U}_t^3 = \hat{g}_t^{*3}(\Pi_t^3), \quad t = 0, \dots, T. \quad (20)$$

Proof. This proof follows similar arguments for centralized minimax control problems [19]. Due to space constraints, it is omitted. \square

As a result of Theorem 3, we conclude that we can derive an optimal partial strategy \hat{g}^* for Problem 2 with the form

$$\hat{U}_t^n = \hat{g}_t^{*n}(\Pi_t^{n:3}), \quad n = 1, 2, 3, \quad t = 0, \dots, T. \quad (21)$$

D. Result for N Subsystems

The results equivalent to Theorems 1 - 3 for $N \in \mathbb{N}$ subsystems can be proven using arguments similar to those for 3 subsystems and mathematical induction. The steps for the mathematical induction are detailed in Appendix A of our online preprint [22]. Here, due to space constraints, we simply state the main result for Problem 2 and Problem 1.

Theorem 4. *In Problem 2, without loss of optimality, we can restrict attention to partial strategies $\hat{\mathbf{g}}^*$ with the form*

$$\hat{U}_t^n = \hat{g}_t^{*n}(\Pi_t^{n:N}), \quad t = 0, \dots, T. \quad (22)$$

Furthermore, in Problem 1, there exists an optimal control strategy $\mathbf{g}^{*k,n}$ for each $k \in \mathcal{K}^n$ and $n \in \mathcal{N}$, with the structural form

$$U_t^{k,n} = g_t^{*k,n}(Y_t^{k,n}, L_t^{k,n}, \Pi_t^{n:N}), \quad t = 0, \dots, T. \quad (23)$$

Proof. The proof follows from the arguments detailed in Appendix A of [22]. \square

Remark 6. The structural form of optimal control strategies in (23) cannot be obtained by a direct application of the common information approach [15]. Note that the domains of our optimal strategies do not grow in size with time when the feasible sets \mathcal{X}_t and $\mathcal{L}_t^{k,n}$ do not grow in size with time for all $k \in \mathcal{K}^n$ and $n \in \mathcal{N}$. This improves the computational tractability of our approach for larger values of horizon T . In contrast, the common information approach considers C_t^1 as a part of the private information of all agents in \mathcal{K}^1 and C_t^1 grows in size with time. Thus, the domains of optimal strategies increase in size with time for $N \geq 2$.

E. Dynamic Programming Decomposition

In this subsection, we construct a DP decomposition for the optimal prescription and partial strategy of subsystem N . Let $\theta_t^N = (\gamma_t^{[N,1]}, \dots, \gamma_t^{[N,N-1]}, \hat{u}_t^N)$ be the realization of Θ_t^N for all $t = 0, \dots, T$. Then, for each possible information state $P_t^N \in \mathcal{P}_t^N$, we define a value function at time T as $V_T(P_t^N) := \max_{s_T^N \in P_t^N} \bar{d}_T^N(s_T^N)$. Furthermore, for each possible information state $P_t^N \in \mathcal{P}_t^N$, for all $t = 0, \dots, T-1$, we iteratively define the value functions

$$V_t(P_t^N) := \min_{\theta_t^N} \max_{z_{t+1}^N \in \tilde{\mathcal{Z}}_{t+1}^N(P_t^N, \theta_t^N)} V_{t+1}[\tilde{f}_t^N(P_t^N, \theta_t^N, z_{t+1}^N)], \quad (24)$$

where the set of feasible values of z_{t+1}^N is given by

$$\tilde{\mathcal{Z}}_{t+1}^N(P_t^N, \theta_t^N) := \{z_{t+1}^N \in \mathcal{Z}_{t+1}^N \mid z_{t+1}^N = \bar{h}_t^N(s_t, \theta_t^N, w_t, v_{t+1}^{1:N}) \text{ for all } s_t^N \in P_t^N, w_t \in \mathcal{W}_t, v_{t+1}^n \in \mathcal{V}_{t+1}^n, n \in \mathcal{N}\}.$$

The prescription law at time t for each $n < N$ is $\gamma_t^{*[N,n]} = \psi_t^{*[N,n]}(P_t^N)$ and the partial control law for subsystem N is $\hat{u}_t^{*N} = \hat{g}_t^{*N}(P_t^N)$, i.e., they are the arginf in the RHS of (24). The corresponding prescription strategy ψ^{*N} and partial strategy $\hat{\mathbf{g}}^{*N}$ can be shown to be optimal for $\mathcal{J}^N(\psi^N, \hat{\mathbf{g}}^N)$ using standard arguments [15], [19].

Remark 7. After solving the DP, we can construct optimal control strategies for Problem 1 in a manner similar to Lemma 2.

Remark 8. When applying our results to an arbitrary decentralized system, there may be multiple feasible ways to allocate the agents to subsystems. For example, consider a three agent system with one-directional communication from agent 3 to agent 2 and from agent 2 to agent 1. We can either allocate all agents to one subsystem or allocate each agent to a unique subsystem. Both of these are valid options because they both ensure that the common information of each subsystem is nested within the common information of all preceding subsystems. However, Theorem 4 leads to a different DP for different allocations. Thus, a system designer must decide on an allocation before deriving the optimal control strategies. In general, our DP performs better when more agents have private information which does not grow in size with time. Thus, in the three agent system with one-directional communication, allocating each agent to a unique subsystem is better than allocating all agents to one subsystem.

F. Extension to Additive Cost

Consider a variation of our problem where the system incurs a cost $d_t(X_t, U_t^{1:N})$ at each $t = 0, \dots, T$ and the performance is measured by the worst-case total cost

$$\Xi(\mathbf{g}) := \max_{X_0, W_{0:T}, V_{0:T}^{1:N}} \sum_{t=0}^{T-1} d_t(X_t, U_t^{1:N}) + d_T(X_T). \quad (25)$$

We can transform the additive cost in (25) into a terminal cost using a technique from [15], [19]. At each $t = 0, \dots, T$, we define an uncertain variable $A_t := \sum_{\ell=0}^{t-1} d_\ell(X_\ell, U_\ell^{1:N}) \in \mathcal{A}_t$ which tracks the cost incurred by the system up to time t . Note that $A_0 = 0$. Then, at each t , we consider an augmented state for the system, $\{X_t, A_t\}$ and note that the augmented state at time $t+1$, $\{X_{t+1}, A_{t+1}\}$, evolves as a function of $\{X_t, A_t\}$, the control actions $U_t^{1:N}$ and the disturbance W_t . Then the total cost is simply given by the terminal cost $A_T + d_T(X_T)$ and thus, we can apply our results to this equivalent terminal cost problem. This extension to total cost problems does present a challenge because the information state at time t takes values in a subset of $2^{\mathcal{X}_t \times \mathcal{A}_t}$, which may grow in size with time as \mathcal{A}_t generally grows in size with time. We plan to address this in the future by exploring alternative augmented states.

V. NUMERICAL EXAMPLE

In this subsection, we validate our results with a simple example. We consider two agents who seek to surround a target. The agents and target can each move along a linear grid with $\Lambda \in \mathbb{N}$ points. Starting at X_0^0 , the position of the target is updated at each $t = 0, \dots, T$ as

$$X_{t+1}^0 = \begin{cases} X_t^0 + W_t^0, & \text{if } 1 \leq X_t^0 + W_t^0 \leq \Lambda, \\ \Lambda, & \text{if } X_t^0 + W_t^0 \geq \Lambda, \\ 1, & \text{if } X_t^0 + W_t^0 \leq 1, \end{cases} \quad (26)$$

with a disturbance $W_t^0 \in \{-1, 0, 1\}$. Each agent occupies a separate subsystem, and thus, we refer to an agent simply by its subsystem $n = 1, 2$. At each t , each agent n selects an action $U_t^n \in \{-1, 0, 1\}$, and updates its position as

$$X_{t+1}^n = \begin{cases} X_t^n + U_t^n, & \text{if } 1 \leq X_t^n + U_t^n \leq \Lambda, \\ \Lambda, & \text{if } X_t^n + U_t^n \geq \Lambda, \\ 1, & \text{if } X_t^n + U_t^n \leq 1. \end{cases} \quad (27)$$

The positions of both agents are observed perfectly by both the agents. Additionally, agent 2 perfectly observes the position of the target, but can only communicate this to agent 1 with a delay of 1 time step. Agent 1 also receives a noisy observation of the target's position $Y_t^1 = \max\{1, \min\{X_t^0 + V_t^1, \Lambda\}\}$, with a noise $V_t^1 \in \{-1, 0\}$. Agent 1 has faulty equipment and cannot transmit their observations to agent 2. Thus, for all $t = 0, \dots, T$, the information structure is given by $C_t^1 = \{X_{0:t-1}^0, Y_{0:t-1}^1, U_{0:t-1}^{1:2}, X_{0:t}^{1:2}\}$ and $L_t^1 = \{Y_t^1\}$ for agent 1, and $C_t^2 = \{X_{0:t-1}^0, U_{0:t-1}^{1:2}, X_{0:t}^{1:2}\}$ and $L_t^2 = \{X_t^2\}$ for agent 2. At the onset, both agents receive a common observation Y_0 implying that the target's starting position is in the set $\{Y_0 - 1, Y_0, Y_0 + 1\}$. The terminal cost is

$$d_T(X_T^{0:3}) = \begin{cases} \sum_{n=1}^2 |X_T^0 - X_T^n|, & \text{if } I^s = 1, \\ D + \sum_{n=1}^2 |X_T^0 - X_T^n|, & \text{if } I^s = 0, \end{cases} \quad (28)$$

where $I^s \in \{0, 1\}$ indicates if the agents have successfully surrounded the target, $\sum_{n=1}^2 |X_T^0 - X_T^n|$ penalizes the distance of the agents from the target, and $D > 0$ is a penalty for failing to surround the target. We compare in Table I the optimal worst-case performance achieved for different realizations of the initial conditions $\{x_0^1, x_0^2, y_0\}$, with and without using information states in the DP.

TABLE I
OPTIMAL COST FOR $\Lambda = 8$, $T = 3$, AND $D = 10$.

$\{x_0^1, x_0^2, y_0\}$	With Information States	Without Information States
$\{8, 8, 2\}$	18	18
$\{3, 6, 7\}$	4	4
$\{3, 3, 4\}$	14	14
$\{3, 5, 8\}$	4	4

VI. CONCLUSION

In this paper, we introduced a general model for decentralized control with nested subsystems and a worst-case shared cost. The disturbances to the system were modeled using uncertain variables. We derived a structural form for optimal control strategies in systems with a terminal cost criterion and observed that it can be used to improve computational tractability for long time horizons. Finally, we presented a DP to derive the optimal strategies and validated our results with an example. A potential direction for future research includes specializing our results for systems with specific dynamics. Another important direction of future research is identifying approximation techniques which can improve computational tractability with a bounded loss in optimal

performance. Finally, we believe that our results can also be utilized in applications of decentralized reinforcement learning problems with a worst-case cost.

REFERENCES

- [1] A. Dave, I. V. Chremos, and A. A. Malikopoulos, "Social Media and Misleading Information in a Democracy: A Mechanism Design Approach," *IEEE Transactions on Automatic Control*, 2022 (in press).
- [2] J. Zabochnik, "Centralized and decentralized decision making in organizations," *Journal of Labor Economics*, vol. 20, no. 1, pp. 1–22, 2002.
- [3] A. M. I. Mahbub and A. A. Malikopoulos, "A Platoon Formation Framework in a Mixed Traffic Environment," *IEEE Control Systems Letters (LCCS)*, vol. 6, pp. 1370–1375, 2021.
- [4] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 1291–1306, IEEE, 2012.
- [5] A. Nayyar and L. Lessard, "Structural results for partially nested lqg systems over graphs," in *2015 American Control Conference (ACC)*, pp. 5457–5464, IEEE, 2015.
- [6] A. Dave and A. A. Malikopoulos, "Decentralized Stochastic Control in Partially Nested Information Structures," in *IFAC-PapersOnLine*, vol. 52, (Chicago, IL, USA), pp. 97–102, 2019.
- [7] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.
- [8] A. Dave and A. A. Malikopoulos, "Structural results for decentralized stochastic control with a word-of-mouth communication," in *2020 American Control Conference (ACC)*, pp. 2796–2801, IEEE, 2020.
- [9] A. Dave and A. A. Malikopoulos, "A dynamic program for a team of two agents with nested information," in *60th IEEE Conference on Decision and Control (CDC)*, pp. 3768–3773, IEEE, 2021.
- [10] A. A. Malikopoulos, "On team decision problems with nonclassical information structures," *IEEE Transactions on Automatic Control*, 2022 arXiv:2101.10992.
- [11] A. Dave, N. Venkatesh, and A. A. Malikopoulos, "On decentralized control of two agents with nested accessible information," *2022 American Control Conference (ACC)*, 2022 (to appear).
- [12] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.
- [13] P. Shah and P. A. Parrilo, " \mathcal{H}_2 -optimal decentralized control over posets: A state-space solution for state-feedback," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3084–3096, 2013.
- [14] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," in *2012 American Control Conference (ACC)*, pp. 6314–6321, IEEE, 2012.
- [15] M. Gagrani and A. Nayyar, "Decentralized minimax control problems with partial history sharing," in *2017 American Control Conference (ACC)*, pp. 3373–3379, IEEE, 2017.
- [16] M. Gagrani, Y. Ouyang, M. Rasouli, and A. Nayyar, "Worst-case guarantees for remote estimation of an uncertain source," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1794–1801, 2020.
- [17] G. N. Nair, "A nonstochastic information theory for communication and state estimation," *IEEE Transactions on automatic control*, vol. 58, no. 6, pp. 1497–1510, 2013.
- [18] A. Rangi and M. Franceschetti, "Towards a non-stochastic information theory," in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 997–1001, IEEE, 2019.
- [19] D. Bertsekas and I. Rhodes, "Sufficiently informative functions and the minimax feedback control of uncertain dynamic systems," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 117–124, 1973.
- [20] P. Bernhard, "Minimax - or feared value - L_1 / L_∞ control," *Theoretical computer science*, vol. 293, no. 1, pp. 25–44, 2003.
- [21] Y. Xie, J. Dibangoye, and O. Buffet, "Optimally solving two-agent decentralized pomdps under one-sided information sharing," in *International Conference on Machine Learning*, pp. 10473–10482, PMLR, 2020.
- [22] A. Dave, N. Venkatesh, and A. A. Malikopoulos, "On decentralized minimax control with nested subsystems," *arXiv:2109.06328*, 2022.