Decentralized Stochastic Control in Partially Nested Information Structures ⋆

Aditya Dave, Andreas A. Malikopoulos

University of Delaware, Newark, DE 19716 USA
(emails: adidave@udel.edu; andreas@udel.edu)

Abstract: In this paper, we analyze a network of agents in a partially nested information structure with a common ancestor. We present the prescription approach applied to different permutations of agents and a structural result for optimal prescriptions of control strategies. We demonstrate the proposed approach through an example that aims at establishing time-invariant domains of the prescriptions without assuming a Linear Quadratic Gaussian problem.

Keywords: Decentralized, distributed, and cooperative control; multi-agent systems.

1. INTRODUCTION

A decentralized control system comprises of multiple agents acting to control the state of the system over a finite number of stages to minimize the total cost. The defining characteristic of these problems is their *information structure*, which is characterized by the topology of the network and the information sharing pattern. We can divide information structures into three categories:

(1) *The classical information structure*: Every agent has access to the same information.

(2) *The quasi-classical information structure*: Includes partially and stochastically nested systems where control strategies for Linear Quadratic Guassian (LQG) problems are linear; see Ho and Chu (1972), Yuksel (2009).

(3) *The non-classical information structure*: Agents influence the decisions of other agents without sharing their complete observation history.

The derivation of optimal control strategies for non-LQG problems with quasi-classical and non-classical information structures is computationally challenging; see Witsenhausen (1968). A key result that can aim at addressing the associate computational challenges is finding sufficient statistics to compress the growing information available to an agent without loss of optimality. Recent advances in decentralized control have focused on finding sufficient statistics for specific information structures. Space constraints limit our literature review, and thus, for a complete discussion of the work reported in the literature to date we refer to Dave and Malikopoulos (2019) and Mahajan et al. (2012), and the references therein.

The work on partially nested information structures has primarily focused on LQG problems and specific information structures; see Mahajan and Nayyar (2015), Nayyar and Lessard (2015), Wu and Lall (2014), Wu and Lall (2010) and references therein. To the best of our knowledge, no sufficient statistics have been derived in general for partially nested information structures. The key contribution of this paper is to derive a structural result for partially nested systems with a common ancestor.

⋆ This research was supported in part by the Delaware Energy Institute (DEI) and by ARPAEs NEXTCAR program under the award number DE- AR0000796.

2. PROBLEM FORMULATION

2.1 The Network of Agents

Consider a system of $K \in \mathbb{N}$ agents represented by a partially nested network, modeled as a directed acyclic graph $\mathscr{G} = (\mathcal{K}, \mathcal{E})$. The set of agents is given by $\mathcal{K} := \{1, ..., K\}$ and every $k \in \mathcal{K}$ is a node of graph $\mathscr{G}$. If there is a link (or edge) from agent $i \in \mathcal{K}$ to agent k, it is denoted by $(i, k) \in \mathcal{E}$. Edge (i, k) represents a communication link from i to k and every agent $k \in \mathcal{K} \setminus \{1\}$ has at least one link starting at k.

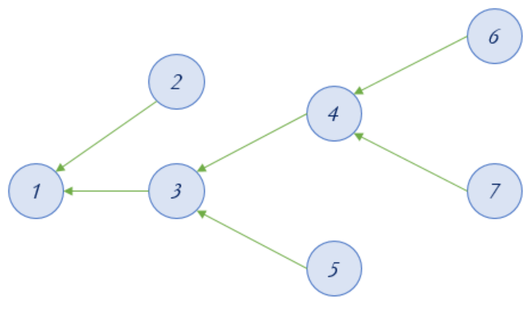

Fig. 1. A partially nested network of agents.

Definition 1. Let $j, k \in \mathcal{K}$ and $\mathcal{N} := \{1, ..., m : m \in \mathbb{N}\}$ be the set of indices. A *path* from agent k to agent j denoted by $(k \to j)$, if it exists, is given by the sequence $\{k_n\}_{n \in \mathcal{N}}$ such that: (1) $k_1 = k$ and $k_m = j$, (2) $k_n \in \mathcal{K}$ for $n \in \mathcal{N}$, and (3) there exists a link $(k_{n-1}, k_n) \in \mathcal{E}$ for $n \in \mathcal{N} \setminus \{1\}$.

For a partially nested information structure, the topology of the network of agents has to be acyclic as in Fig. 1. This implies that if there is a path $(k \to j)$ from an agent $k \in \mathcal{K}$ to another agent $j \in \mathcal{K}$, then there is no path from j to k. There is always a path from agent k to itself denoted by $(k \to k)$. The set of all paths starting at agent k is $\mathcal{Q}^k$.

By convention we index the agents so that if there is a path $(k \to j) \in \mathcal{Q}^k$ and $k \neq j$, then $k > j$. The relationship between the agents in the network can be summarized through as below.

Definition 2. The set of *ancestors* for agent $k \in \mathcal{K}$ is given by $\mathscr{A}^k := \{j \in \{1, ..., k\} : (k \to j) \in \mathcal{Q}^k\}$.

Definition 3. The set of *descendents* for agent $k \in \mathcal{K}$ is given by $\mathscr{D}^k := \{l \in \{k, ..., K\} : (l \to k) \in \mathcal{Q}^l\}$.

We assume the presence of a *common ancestor* for the system and call it agent 1, i.e., $\mathscr{D}^1 = \mathcal{K}$. Any agents that have no descendents other than themselves are known as *leaf nodes.*

2.2 System Description

We consider a finite horizon discrete time system with a given time horizon $T \in \mathbb{N}$. At time $t \in \{0, 1, ..., T\}$ the state of the system X_t and control action U_t^k of agent $k \in \mathcal{K}$ are random variables that takes values in the finite sets $\mathcal{X}$ and $\mathcal{U}$ respectively. The inital state X_0 has a known distribution and the evolution of the system is given by

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

where W_t is the uncontrolled disturbance to the system represented as a random variable with a known distribution that takes values from a finite set $\mathcal{W}$. At time t, agent $k \in \mathcal{K}$ makes an observation Y_t^k taking values in a finite set $\mathcal{Y}^k$ as

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

where V_t^k is the sensor noise represented as a random variable with a known distribution that takes values in the finite set $\mathcal{V}^k$. We assume that the random variables $\{X_0, W_{0:T}, V_{0:T}^1, ..., V_{0:T}^K\}$ are independent of one another and that their distributions are known.

Agent $k \in \mathcal{K}$ selects a control action U_t^k as a function of its memory, defined through the partially nested information structure of the system in Section 2.3. After each agent k generates a control action U_t^k, the system incurs a cost $c_t(X_t, U_t^1, ..., U_t^K)$. We assume that the functions $\{f_t, h_t^1, ..., h_t^K, c_t : t = 0, ..., T\}$ are known.

2.3 The Partially Nested Information Structure

The information structure of the system is characterized by the topology of the network (Section 2.1) and the rules of communication. We set the following rules of communication from an agent $k \in \mathcal{K}$ to an agent $j \in \mathscr{A}^k$:

(1) At time t, agent k transmits information denoted by the set $\{Y_t^k, U_{t-1}^k\}$ to every agent $j \in \mathscr{A}^k$.

(2) The communication path $(k \to j)$ has a communication delay of $\delta^{[k,j]} \in \mathbb{N}$ time steps, i.e, information $\{Y_t^k, U_{t-1}^k\}$ transmitted by k at time t is received by j at time $t + \delta^{[k,j]}$.

Note that the communication delays are deterministic and known apriori. By convention we set $\delta^{[k,k]} = 0$. The information available to an agent $k \in \mathcal{K}$ at time t is given by the following definition.

Definition 4. The *memory* of agent $k \in \mathcal{K}$ is defined as the random variable M_t^k that takes values in the finite set $\mathcal{M}_t^k$ and is given by,

$$M_t^k := \left\{Y_{0:t-\delta^{[l,k]}}^l, U_{0:t-\delta^{[l,k]}-1}^l : l \in \mathscr{D}^k\right\}. \tag{3}$$

At time t, agent k generates a control action

$$U_t^k := g_t^k(M_t^k), \tag{4}$$

where g_t^k is the control policy of agent k at time t. We define the control policy for each agent as $\boldsymbol{g}^k := (g_0^k, ..., g_T^k)$ and the strategy of the system as $\boldsymbol{g} := (\boldsymbol{g}^1, ..., \boldsymbol{g}^K)$. The set of all feasible control strategies is denoted by $\mathcal{G}$.

The performance criterion for the system is given by the total expected cost,

$$\textbf{Problem 1:} \quad \mathcal{J}(\boldsymbol{g}) = \mathbb{E}^{\boldsymbol{g}}\left[\sum_{t=0}^{T} c_t(X_t, U_t^1, ..., U_t^K)\right], \tag{5}$$

where the expectation is with respect to the joint probability measure on the random variables $X_t, U_t^1, ..., U_t^K$. The problem is to select an optimal strategy $\boldsymbol{g}^* \in \mathcal{G}$ that minimizes performance criterion (5).

3. THE PRESCRIPTION APPROACH

3.1 Permutations of the Agents

The first step in our analysis is to construct $(K-1)!$ permutations of agents in the system by fixing agent 1 as the first agent in all permutations.

Definition 5. A *permutation* of the agents in $\mathcal{K}$ is defined as a K-tuple ordered list $o_m = (o_{m,1}, ..., o_{m,K})$, with $m \in \{1, ..., (K-1)!\}$ and $o_{m,1} = 1$.

Let $\mathcal{P} = \{1, ..., K\}$ be the set of possible positions in a permutation. The component $o_{m,p}$ in permutation o_m refers to an agent with position $p \in \mathcal{P}$ in o_m. As an example, consider a system with three agents that has a permutation given by $o_m = (1, 3, 2)$. In this permutation, we say the agent 2 is located at position 3 and $o_{m,3} = 2$.

Remark 1. The lower-case letters i, j, and k refer to agents in the set $\mathcal{K}$, while the letters p, q, and r refer to the positions of agents in a permutation $o_m = (o_{m,1}, ..., o_{m,K})$.

Remark 2. When agent $k \in \mathcal{K}$ occupies position p in permutation o_m, we denote any random variables of the form U_t^k, M_t^k, etc, equivalently by $U_t^{[p]}$, $M_t^{[p]}$, etc. Similarly, we denote any function of the form g_t^k equivalently by $g_t^{[p]}$.

3.2 Construction of Prescriptions

For an agent $k \in \mathcal{K}$ located at position $p \in \mathcal{P}$ in prescription o_m, the control action $U_t^{[p]} = U_t^k$ is generated in two stages:

(1) The agent located at position p in permutation o_m generates a function based on information which is a subset of its memory $M_t^{[p]}$.

(2) This function takes as an input the compliment of the subset used to generate it and yields the control action $U_t^{[p]}$.

We call these functions *prescriptions.* Next, we show that the prescriptions allow us to formulate an equivalent optimization problem with respect to the optimal strategy for prescriptions instead of the optimal strategy $\boldsymbol{g}^*$ in Problem 1. In this section, we construct the subset of memory $M_t^{[p]}$ and prescriptions for agents at every position p in permutation o_m without changing the information structure.

Definition 6. Let agent k be at position p in permutation o_m with memory $M_t^{[p]}$ at time t. The *accessible information* of the agent located at p is defined as the set $A_t^{[p]}$ that takes values in the finite collection of sets $\mathcal{A}_t^{[p]}$ such that,

$$A_t^{[p]} = \bigcap_{q=1}^{p} \left(M_t^{[q]}\right). \tag{6}$$

As an example for a permutation o_m, we can write (6) for the agent at position $p = 1$ as

$$A_t^{[1]} = M_t^{[1]}, \tag{7}$$

where $o_{m,1} = 1$. Similarly, for agent at position $p = 2$, we can write (6) as,

$$A_t^{[2]} = M_t^{[1]} \cap M_t^{[2]}. \tag{8}$$

The form of the accessible information motivates us to define a set of agents with positions *beyond* p.

Definition 7. For a position $p \in \mathcal{P}$ in permutation o_m, the set of positions *beyond* p is given by $\mathcal{B}^{[p]} := \{q \in \mathcal{P} : q \geq p\}$.

Then, from the definition of the accessible information $A_t^{[p]}$ we have the properties,

$$A_t^{[q]} \subset A_t^{[p]},\ \forall q \in \mathcal{B}^{[p]}, \tag{9}$$

$$A_{t-1}^{[p]} \subset A_t^{[p]}. \tag{10}$$

Property (10) motivates a new term to denote the information added to $A_t^{[p]}$ at time t.

Definition 8. The *new information* for the agent located at $p \in \mathcal{P}$ at time t is defined as the set $Z_t^{[p]}$ that takes values in a finite collection of sets $\mathcal{Z}_t^{[p]}$ such that,

$$Z_t^{[p]} := A_t^{[p]} \backslash A_{t-1}^{[p]}. \tag{11}$$

We observe in (9) that for all $p \in \mathcal{P}$ and $q \in \mathcal{B}^{[p]}$, we have $A_t^{[q]} \subset M_t^{[p]}$. Thus, we can define the *inaccessible information* of the agent at p with respect to the accessible information $A_t^{[q]}$ of any $q \in \mathcal{B}^{[p]}$.

Definition 9. The *inaccessible information* of an agent at position p in permutation o_m with respect to accessible information $A_t^{[q]}$, $q \in \mathcal{B}^{[p]}$, is defined as the set $L_t^{[p,q]}$ that takes values in the finite collection of sets $\mathcal{L}_t^{[p,q]}$ such that,

$$L_t^{[p,q]} := M_t^{[p]} \setminus A_t^{[q]}. \tag{12}$$

The pair of sets $A_t^{[q]}$ and $L_t^{[p,q]}$ forms a partition of the set $M_t^{[p]}$. We use this partition to define a *prescription* function below.

Definition 10. The *prescription* for an agent at position p is given by the function,

$$\Theta_t^{[p]} := (\Gamma_t^{[p,q]} : q \in \mathcal{P}), \tag{13}$$

where,

$$\Gamma_t^{[p,q]} : \begin{cases} \mathcal{L}_t^{[q,p]} \longmapsto \mathcal{U}_t^{[q]} \text{ if } q \notin \mathcal{B}^{[p]}, \\ \mathcal{L}_t^{[q,q]} \longmapsto \mathcal{U}_t^{[q]} \text{ if } q \in \mathcal{B}^{[p]}. \end{cases} \tag{14}$$

Each component $\Gamma_t^{[p,q]}$ of $\Theta_t^{[p]}$ is generated as,

$$\Gamma_t^{[p,q]} := \begin{cases} \psi_t^{[p,q]}(A_t^{[p]}) \text{ if } q \notin \mathcal{B}^{[p]}, \\ \psi_t^{[p,q]}(A_t^{[q]}) \text{ if } q \in \mathcal{B}^{[p]}, \end{cases} \tag{15}$$

where $\boldsymbol{\psi}^{[p,q]} := (\psi_0^{[p,q]}, ..., \psi_T^{[p,q]})$ is called the prescription policy of the agent at p for the agent at q and $\boldsymbol{\psi}^{[p]} := (\boldsymbol{\psi}^{[p,q]} : q \in \mathcal{P})$ is called the prescription strategy of the agent at p. The set of feasible prescription strategies for the agent at p is denoted by $\Psi^{[p]}$.

Remark 3. We write the vector $\Theta_t^{[p]}$ as $(\Gamma_t^{[p,q]} : q \in \mathcal{P})$ to highlight that it is defined with respect to the positions of the agents in permutation o_m.

3.3 Prescriptions and Control Strategies

The first two results state that for an agent k at position p in permutation o_m, we can use the prescription $\Theta_t^{[p]}$ to generate control action U_t^k instead of the strategy g_t^k.

Lemma 1. *Let agent $k \in \mathcal{K}$ be located at position p in permutation o_m and $\Theta_t^{[p]}$ be its prescription. For every strategy $\boldsymbol{g}$, there exists a prescription strategy $\boldsymbol{\psi}^{[p]}$ such that control action $U_t^{[p]}$ is given by*

$$U_t^{[p]} = \Gamma_t^{[p,p]} \left(L_t^{[p,p]}\right). \tag{16}$$

Proof. Let $A_t^{[p]}$ and $L_t^{[p,p]}$ be the accessible and inaccessible information respectively of agent k located at position p in permutation o_m. We select a strategy $\boldsymbol{g}$ that can generate control action $U_t^{[p]}$ through (4). Then, we define

$$\Gamma_t^{[p,p]}(\cdot) := g_t^{[p,p]}(A_t^{[p]}, \cdot). \tag{17}$$

Hence

$$\begin{aligned} U_t^{[p]} =& \Gamma_t^{[p,p]}(L_t^{[p,p]}) \\ =& g_t^{[p]}(A_t^{[p]}, L_t^{[p,p]}) = g_t^{[p]}(M_t^{[p]}) = g_t^k(M_t^k). \end{aligned} \tag{18}$$

□

Thus, any strategy $\boldsymbol{g}$ can be implemented through an appropriate prescription strategy $\boldsymbol{\psi}^{[p]}$ to generate the same control actions $U_t^{[p]}$ for agents at any position $p \in \mathcal{P}$. Next, we show that the reverse is also true.

Lemma 2. *Let agent $k \in \mathcal{K}$ be located at position p in permutation o_m and let $\Theta_t^{[p]}$ be its prescription. For every prescription strategy $\boldsymbol{\psi}^{[p]} \in \Psi^{[p]}$, there exists a strategy $\boldsymbol{g} \in \mathcal{G}$ such that*

$$\Gamma_t^{[p,p]} \left(L_t^{[p,p]}\right) = g_t^{[p]}(M_t^{[p]}). \tag{19}$$

Proof. We select a prescription strategy $\boldsymbol{\psi}^{[p]}$ and construct the corresponding control policy $\boldsymbol{g}^{[p]}$. Hence

$$\begin{aligned} U_t^{[p]} &= g_t^{[p]}(M_t^{[p]}) \\ &= g_t^{[p]}(A_t^{[p]}, L_t^{[p,p]}) := \psi_t^{[p,p]}(A_t^{[p]})(L_t^{[p,p]}). \end{aligned} \tag{20}$$

□

Our next result establishes the equivalence between prescriptions $\Theta_t^{[p]}$ and $\Theta_t^{[q]}$ generated by agents at positions p and q respectively in permutation o_m.

Lemma 3. *Let k and j be two agents at positions $p \in \mathcal{P}$ and $q \in \mathcal{B}^{[p]}$ and let $\boldsymbol{\psi}^{[p]}$ be a given prescription strategy for the agent located at p. Then, there exists a positional* *relationship function* $e^{[q,p]} = (e_1^{[q,p]}, ..., e_T^{[q,p]})$ *such that the prescription strategy of the agent at q constructed as*

$$\psi_t^{[q,r]} = e_t^{[q,p]}\left(\psi_t^{[p,r]}\right),\ \forall r \in \mathcal{P}, \tag{21}$$

yields

$$\begin{aligned}
&\Gamma_t^{[p,r]}(L_t^{[r,r]}) = \Gamma_t^{[q,r]}(L_t^{[r,r]}) \text{ if } r \in \mathcal{B}^{[q]},\\
&\Gamma_t^{[p,r]}(L_t^{[r,r]}) = \Gamma_t^{[q,r]}(L_t^{[r,q]}) \text{ if } r \in \mathcal{B}^{[p]}, r \notin \mathcal{B}^{[q]},\\
&\Gamma_t^{[p,r]}(L_t^{[r,p]}) = \Gamma_t^{[q,r]}(L_t^{[r,q]}) \text{ if } r \notin \mathcal{B}^{[p]}.
\end{aligned} \tag{22}$$

Proof. Let $g_t^{[r]}$ be the control policy for a third agent $i \in \mathcal{K}$ at position r at time t. We prove the result by constructing the control policy $g_t^{[r]}$ and the prescription policy $\psi_t^{[q]}$ in three cases, given a prescription strategy $\boldsymbol{\psi}^{[p]}$.

(1) For $r \in \mathcal{B}^{[q]}$, let $g_t^{[r]} : \mathcal{M}_t^{[r]} \mapsto \mathcal{U}_t^{[r]}$ and $\psi_t^{[q,r]}(A_t^{[r]}) : \mathcal{L}_t^{[r,r]} \mapsto \mathcal{U}_t^{[r]}$ be given by,

$$\psi_t^{[p,r]}(A_t^{[r]})(L_t^{[r,r]}) =: g_t^{[r]}(A_t^{[r]}, L_t^{[r,r]}), \tag{23}$$

$$\psi_t^{[p,r]} =: \psi_t^{[q,r]}. \tag{24}$$

(2) For $r \in \mathcal{B}^{[p]} \cap (\mathcal{K} \setminus \mathcal{B}^{[q]})$, let $g_t^{[r]} : \mathcal{M}_t^{[r]} \mapsto \mathcal{U}_t^{[r]}$ and $\psi_t^{[q,r]}(A_t^{[q]}) : \mathcal{L}_t^{[r,q]} \mapsto \mathcal{U}_t^{[r]}$ be given by,

$$\begin{aligned}
&\psi_t^{[p,r]}(A_t^{[r]})(L_t^{[r,r]}) =: g_t^{[r]}(A_t^{[r]}, L_t^{[r,r]})\\
&= g_t^{[r]}(A_t^{[q]}, L_t^{[r,q]}) =: \psi_t^{[q,r]}(A_t^{[q]})(L_t^{[r,q]}).
\end{aligned} \tag{25}$$

(3) For $r \notin \mathcal{B}^{[p]}$, let $g_t^{[r]} : \mathcal{M}_t^{[r]} \mapsto \mathcal{U}_t^{[r]}$ and $\psi_t^{[q,r]}(A_t^{[q]}) : \mathcal{L}_t^{[r,q]} \mapsto \mathcal{U}_t^{[r]}$ be given by,

$$\begin{aligned}
&\psi_t^{[p,r]}(A_t^{[p]})(L_t^{[r,p]}) =: g_t^{[r]}(A_t^{[p]}, L_t^{[r,p]})\\
&= g_t^{[r]}(A_t^{[q]}, L_t^{[r,q]}) =: \psi_t^{[q,r]}(A_t^{i})(L_t^{[r,q]}).
\end{aligned} \tag{26}$$

Then, through (24), (25) and (26), we can define a function $e^{[q,p]} : \Psi^{[p]} \mapsto \Psi^{[q]}$ such that (21) satisfies (22). ☐

Lemma 4. *Let k and j be two agents at positions* $p \in \mathcal{P}$ *and* $q \notin \mathcal{B}^{[p]}$ *in permutation* o_m *and let* $\boldsymbol{\psi}^{[p]}$ *be a given prescription strategy for the agent at p. Then, there exists a positional relationship function* $e^{[q,p]} = (e_1^{[q,p]}, ..., e_T^{[q,p]})$ *such that the prescription strategy of the agent at q constructed through* (21) *leads to*

$$\begin{aligned}
&\Gamma_t^{[p,r]}(L_t^{[r,r]}) = \Gamma_t^{[q,r]}(L_t^{[r,r]}) \text{ if } r \in \mathcal{B}^{[p]},\\
&\Gamma_t^{[p,r]}(L_t^{[r,p]}) = \Gamma_t^{[q,r]}(L_t^{[r,r]}) \text{ if } r \in \mathcal{B}^{[q]}, r \notin \mathcal{B}^{[p]},\\
&\Gamma_t^{[p,r]}(L_t^{[r,p]}) = \Gamma_t^{[q,r]}(L_t^{[r,q]}) \text{ if } r \notin \mathcal{B}^{[q]}.
\end{aligned} \tag{27}$$

Proof. The proof is similar to the proof of Lemma 3, and thus it is omitted due to space constraints. ☐

Corollary 1. *Let* $\boldsymbol{\psi}^{[p]}$ *be the prescription strategy of the agent at* $p \in \mathcal{P}$, *and* $a_t^{[p]}$ *the realization of* $A_t^{[p]}$. *For any* $q \in \mathcal{B}^{[p]}$, *the realization* $\theta_t^{[q]}$ *of the prescription* $\Theta_t^{[q]}$ *is given by*

$$\gamma_t^{[q,r]} = \begin{cases} e_t^{[q,p]}(\psi_t^{[p,r]})(a_t^{[q]}) & \text{if } r \notin \mathcal{B}^{[q]},\\ \psi_t^{[p,r]}(a_t^{[r]}) & \text{if } r \in \mathcal{B}^{[q]}. \end{cases} \tag{28}$$

For every pair of agents at positions p and q in permutation o_m, we define the function $e^{[q,p]}$ satisfying Lemmas 3 and 4. Then, for any two agents at arbitrary positions $p, q \in \mathcal{P}$ in permutation o_m, we have

$$U_t^{[q]} = \Gamma_t^{[q,q]}(L_t^{[q,q]}) = \begin{cases} \Gamma_t^{[p,q]}(L_t^{[q,p]}) \text{ if } q \notin \mathcal{B}^{[p]},\\ \Gamma_t^{[p,q]}(L_t^{[q,q]}) \text{ if } q \in \mathcal{B}^{[p]}. \end{cases} \tag{29}$$

The next result relates prescription strategies for the common ancestor derived with respect to two different permutations.

Lemma 5. *Let* $o_m, o_n \in \mathcal{O}$ *be two permutations on the set* $\mathcal{K}$ *such that an agent* $k \in \mathcal{K}$ *occupies position p in* o_m *and position* p' *in* o_n. *Let* $\psi_t^{[1,p]}$ *be the prescription policy of agent 1 for prescription* $\Gamma_t^{[1,p]}$ *in permutation* o_m. *There exists a permutation relationship function* $d_t^{[o_m,o_n]}$ *such that the prescription policy* $\psi_t^{[1',p']}$ *in* o_n *given by*

$$\psi_t^{[1',p']} = d_t^{[o_n,o_m]}\left(\psi_t^{[1,p]}\right), \tag{30}$$

leads to the same control action U_t^k *as* $\psi_t^{[1,p]}$.

Proof. We construct the control policy g_t^k and a prescription policy $\psi_t^{[1',p']}$ of agent 1 for agent k located at position p' in permutation o_n as

$$\begin{aligned}
\psi_t^{[1,p']}(A_t^{p'})(L_t^{[p',p']}) &=: g_t^k(M_t^k)\\
&=: \psi_t^{[1,p]}(A_t^{[p]})(L_t^{[p,p]}).
\end{aligned} \tag{31}$$

From Corollary 1, we know that for any p in a permutation o_m there exists an invertible function $e^{[p,1]}$ such that,

$$\psi^{[p,p]} = e^{[p,1]}\left(\psi^{[1,p]}\right). \tag{32}$$

The result follows by substitution and similar arguments to ones made in Lemma 2. A detailed proof can be found in Dave and Malikopoulos (2019). ☐

Lemma 5 establishes that all permutations are equivalent for the purpose of deriving prescription strategies. To this end, we use the trivial permutation $o_1 = (1, 2, ..., K)$ for subsequent sections. In o_1, agent $k \in \mathcal{K}$ is located at position $k \in \mathcal{P}$.

3.4 The Designer's Problem

In the previous section, Lemmas 1 through 4 imply that the control action $U_t^j = U_t^{[q]}$ of an agent j located at position $q \in \mathcal{P}$ in permutation o_m can be generated equivalently through the prescription $\Gamma_t^{[p,q]}$ of an agent k located at $p \in \mathcal{P}$ in permutation o_m. This relationship is given in (29).

This allows us to write the cost to the system at time t for all $k \in \mathcal{P}$ in the trivial permutation o_1 as

$$\begin{aligned}
&c_t(X_t, U_t^1, ..., U_t^K)\\
&=: c_t^{o_1}\big(X_t, \Gamma_t^{[k,1]}(L_t^{[1,k]}), ..., \Gamma_t^{[k,k]}(L_t^{[k,k]}),\\
&\qquad \Gamma_t^{[k,k+1]}(L_t^{[k+1,k+1]}), ..., \Gamma_t^{[k,K]}(L_t^{[K,K]})\big),
\end{aligned} \tag{33}$$

where the function $c_t^{o_1}(.)$ is same as function $c_t(.)$ with its inputs after X_t taken in the order of permutation o_1. This can be verified through substitution.

Then, we can reformulate Problem 1 from the point of view of a designer with access to memory $M_t^{[k]}$ that must select an optimal prescription strategy $\boldsymbol{\psi}^{*[k]}$ that minimizes the performance criterion.

Problem 2: $\mathcal{J}^{[k]}(\boldsymbol{\psi}^{[k]}) =$

$$\mathbb{E}^{\boldsymbol{\psi}^{[k]}}\Big[\sum_{t=0}^{T} c_t^{o_1}\big(X_t, \Gamma_t^{[k,1]}(L_t^{[1,k]}), ..., \Gamma_t^{[k,k]}(L_t^{[k,k]}), \Gamma_t^{[k,k+1]}(L_t^{[k+1,k+1]}), ..., \Gamma_t^{[k,K]}(L_t^{[K,K]})\big)\Big], \quad (34)$$

and select strategies $\boldsymbol{\psi}^{*[j]}$ for all $j \in \mathcal{P}$ through (21) and (30). We call this Problem 2 for position $k \in \mathcal{P}$. A detailed proof for the equivalence between Problem 1 and Problem 2 can be found in Dave and Malikopoulos (2019). Next we give a state sufficient for input-output mapping.

Lemma 6. *A state sufficient for input-output mapping for Problem 2 for position $k \in \mathcal{P}$ is*

$$S_t^{[k]} := \{X_t, L_t^{[1,k]}, ..., L_t^{[k,k]}, L_t^{[k+1,k]}, ..., L_t^{[K,K]}\}. \quad (35)$$

Proof. The state $S_t^{[k]}$ satisfies the three properties given in Mahajan (2008):

(1) There exist functions $\hat{f}_t^{[k]}$, $t = 0, ..., T$ such that

$$S_{t+1}^{[k]} = \hat{f}_t^{[k]}(S_t^{[k]}, W_t, V_{t+1}^{1:K}, \Theta_t^{[k]}). \quad (36)$$

(2) There exist functions $\hat{h}_t^{[k]}$, $t = 0, ..., T$ such that

$$Z_{t+1}^{[k]} = \hat{h}_t^{[k]}(S_t^{[p]}, \Theta_t^{[k]}, V_{t+1}^{1:K}). \quad (37)$$

(3) There exist functions $\hat{c}_t^{[k]}$, $t = 0, ..., T$ such that

$$c_t(X_t, U_t^1, ..., U_t^K) = \hat{c}_t^{[k]}(S_t^{[k]}, \Theta_t^{[k]}). \quad (38)$$

The proof can be found in Dave and Malikopoulos (2019). □

Then, from the point of view of the designer the system behaves like a Partially Observed Markov Decision Process (POMDP) with state $S_t^{[k]}$, control input $\Theta_t^{[k]}$, output $Z_t^{[k]}$ and cost $\hat{c}_t^{[k]}(S_t^{[k]}, \Theta_t^{[k]})$ at time t. The deviation from standard theory is that different prescriptions $\Gamma_t^{[k,j]}$ for $j \in \mathcal{B}^{[k]}$ are generated as functions of accessible information $A_t^{[j]}$ as opposed to $A_t^{[k]}$. We address this concern in Section 4.2. We first define the information state k.

Definition 11. Let $S_t^{[k]}$, $A_t^{[k]}$ and $\Theta_{0:t-1}^{[k]}$ be the state, history of outputs and control inputs respectively at time t for prescription problem k. The *information state* is defined as a probability distribution $\Pi_t^{[k]}$ that takes values in the feasible realizations $\mathscr{P}_t^{[k]} := \Delta(\mathcal{S}_t^{[k]})$ such that

$$\Pi_t^{[k]}(s_t^{[k]}) := \mathbb{P}^{\boldsymbol{\psi}^{[k]}}(S_t^{[k]} = s_t^{[k]} | A_t^{[k]}, \Theta_{0:t-1}^{[k]}). \quad (39)$$

4. RESULTS

4.1 Properties of the Information States

In this section, we present results that establish that the information state $\Pi_t^{[k]}$ for all $k \in \mathcal{K}$ is independent from the prescription strategy $\boldsymbol{\psi}^{[k]}$.

Lemma 7. *(Dave and Malikopoulos (2019)) At time t, there exists a function $F_t^{[k]}$ independent of the prescription strategy $\boldsymbol{\psi}^{[k]}$ such that*

$$\Pi_{t+1}^{[k]} = F_{t+1}^{[k]}(\Pi_t^{[k]}, \Theta_t^{[k]}, Z_{t+1}^{[k]}). \quad (40)$$

Lemma 8. *(Dave and Malikopoulos (2019)) The evolution of the information state Π_t at time t is as a controlled Markov Chain,*

$$\mathbb{P}(\Pi_{t+1}^{[k]} | D_t^{[k]}, \Theta_{0:t}^{[k]}, \Pi_{0:t}^{[k]}) = \mathbb{P}(\Pi_{t+1}^{[k]} | \Pi_t^{[k]}, \Theta_t^{[k]}). \quad (41)$$

Lemma 9. *(Dave and Malikopoulos (2019)) At time t, there exists a function $C_t^{[k]}$, independent of the prescription strategy $\boldsymbol{\psi}^{[k]}$, such that,*

$$\mathbb{E}^{\boldsymbol{\psi}^{[k]}}\big[\hat{c}_t^{[k]}(S_t^{[k]}, \Theta_t^{[k]}) | A_t^{[k]}, \Theta_{0:t}^{[k]}\big] = C_t^{[k]}(\Pi_t^{[k]}, \Theta_t^{[k]}). \quad (42)$$

Together these results establish that the information state $\Pi_t^{[k]}$ evolves as a controlled Markov chain with control inputs $\Theta_t^{[k]}$.

4.2 Structural Result

We start with a structural result for Problem 2 for position K. Position K is the last position in the trivial permutation and thus $\mathcal{B}^{[K]} = \emptyset$. Then, the prescription component $\Gamma_t^{[K,k]}$ is a function of $A_t^{[K]}$ for every $k \in \mathcal{P}$ and we have the following structural result from Nayyar et al. (2011).

Lemma 10. *Consider Problem 2 for position K. There exists an optimal prescription strategy $\boldsymbol{\psi}^{*[K]}$ of the form*

$$\Gamma_t^{*[K,k]} = \psi_t^{*[K,k]}(\Pi_t^{[K]}). \quad (43)$$

For every other position k for Problem 2, we observe that in the trivial permutation $o_1 = (1, ..., K)$ for agent $j \in \mathcal{B}^k$, property (9) gives us $A_t^{[j]} \subset A_t^{[k]}$. Thus, given $A_t^{[k]}$ and the optimal prescription strategy $\boldsymbol{\psi}^{*[k]}$, agent k can derive optimal prescriptions $\Theta_t^{*[j]}$ for $j \in \mathcal{B}^{[k]}$ through (28). This leads us to the following structural result for the optimal prescription strategy $\boldsymbol{\psi}^{*[k]}$.

Theorem 1. *For Problem 2 for position k in permutation o_1, there exists an optimal prescription strategy $\boldsymbol{\psi}^{*[k]}$ of the form*

$$\Gamma_t^{*[k,j]} := \begin{cases} \psi_t^{*[k,j]}(\Pi_t^{[k]}, ..., \Pi_t^{[K]}) \text{ if } j \notin \mathcal{B}^{[k]}, \\ \psi_t^{*[k,j]}(\Pi_t^{[j]}, ..., \Pi_t^{[K]}) \text{ if } j \in \mathcal{B}^{[k]}. \end{cases} \quad (44)$$

Proof. Due to space constraints, we provide just a sketch of the proof. The proof by mathematical induction is written in three steps:

(1) We assume that the structural result holds for the prescription $\Theta_t^{[k+1]}$ for position $k+1$ in permutation o_m.

(2) Starting with time T, we can show for time steps $T-1, T-2, ..., 0$ that the structural result holds for position k, given that it holds for position $k+1$.

(3) We note that for the last position K, structural result (43) in Lemma 10 is equivalent to (44).

A detailed proof can be found in Dave and Malikopoulos (2019). □

The structural result in Theorem 1 is valid for every permutation o_m. For agents with positions p and q in o_m, the structural result can be stated as,

$$\Gamma_t^{*[p,q]} := \begin{cases} \psi_t^{*[p,q]}(\Pi_t^{[r]} : r \in \mathcal{B}^{[p]}) \textit{ if } q \notin \mathcal{B}^{[p]}, \\ \psi_t^{*[p,q]}(\Pi_t^{[r]} : r \in \mathcal{B}^{[q]}) \textit{ if } q \in \mathcal{B}^{[p]}. \end{cases} \quad (45)$$

In practice, we wish to use dynamic programming to derive the optimal strategy $\boldsymbol{\psi}^{*[1]}$ for agent 1 at position 1 since the prescriptions $\Gamma_t^{*[1,k]}$ have the smallest domain size among equivalent prescriptions generated by other agents in the same permutation, namely

$$L_t^{[k,k]} \subset L_t^{[k,j]},\ \forall j \in \mathcal{B}^{[k]}. \tag{46}$$

However, if we compare permutations, every agent $k \in \mathcal{K}$ minimizes the domain size of the optimal prescription $\Gamma_t^{*[1,k]}$ of agent 1 for agent k in a permutation o_m where $o_{m,2} = k$. An example of how this can be achieved is presented below and the arguments made there can also be made for other systems.

4.3 An Example with 3 Agents

Consider a system with 3 agents: the common ancestor denoted by agent 1 and leaf agents 2 and 3. There are no delays in communication. Memories of the agents at time t are given by

$$M_t^3 = \{Y_{0:t}^3, U_{0:t-1}^3\}, \tag{47}$$

$$M_t^2 = \{Y_{0:t}^2, U_{0:t-1}^2\}, \tag{48}$$

$$M_t^1 = \{Y_{0:t}^1, U_{0:t-1}^1, M_t^2, M_t^3\}. \tag{49}$$

The two permutations for this system are $o = (1, 2, 3)$ and $o' = (1, 3, 2)$. When referring to positions in permutation o, we will write them as $p \in \mathcal{P}$ and when referring to positions in permutation o', we will write them as $p' \in \mathcal{P}$.

The accessible information of every agent with respect to permutations o and o' is given by

$$A_t^{[1]} = M_t^1, \qquad A_t^{[1']} = M_t^1, \tag{50}$$

$$A_t^{[2]} = M_t^2, \qquad A_t^{[2']} = M_t^3, \tag{51}$$

$$A_t^{[3]} = \emptyset, \qquad A_t^{[3']} = \emptyset. \tag{52}$$

The corresponding information states are given by $\Pi_t^{[p]} = \mathbb{P}(S_t^{[p]}|A_t^{[p]})$ and $\Pi_t^{[p']} = \mathbb{P}(S_t^{[p']}|A_t^{[p']})$ in permutations o and o' respectively. The optimal prescription strategy $\boldsymbol{\psi}^{*[1]}$ of agent 1 for permutation o through Theorem 1 is

$$\Gamma_t^{*[1,p]} = \psi_t^{*[1,p]}(\Pi_t^{[p]}, ..., \Pi_t^{[3]}). \tag{53}$$

The size of the domain of prescription $\Gamma_t^{*[1,p]}$ is given by $|L_t^{[p,p]}|$. Since $A_t^{[3]} = \emptyset$, the domain of prescription $\Gamma_t^{*[1,3]}$ of agent 1 for agent 3 in permutation o is given by $L_t^{[3,3]} = M_t^3$. This domain grows larger with time.

In contrast, the optimal prescription $\Gamma_t^{*[1',2']}$ of agent 1 for agent 3 in permutation o' is given by

$$\Gamma_t^{*[1',2']} = \psi_t^{*[1',2']}(\Pi_t^{[2']}, \Pi_t^{[3']}), \tag{54}$$

where the domain of $\Gamma_t^{*[1',2']}$ is $\emptyset$ and does not grow with time. In order to exploit the advantage of certain permutations for certain agents, we can define a new mixed cost function by invoking Lemma 5 as

$$\bar{C}_t^1(X_t, L_t^{[1,1]}, L_t^{[2,2]}, L_t^{[2',2']}, \Gamma_t^{[1,1]}, \Gamma_t^{[1,2]}, \Gamma_t^{[1',2']}) := c_t(X_t, U_t^1, U_t^2, U_t^3). \tag{55}$$

Then the optimal mixed prescription of agent 1 is given by a combination of prescriptions generated through permutations o and o' as $\Theta_t^{*1} = (\Gamma_t^{*[1,1]}, \Gamma_t^{*[1,2]}, \Gamma_t^{*[1',2']})$ and the corresponding mixed prescription strategy is given by $(\boldsymbol{\psi}^{*[1,1]}, \boldsymbol{\psi}^{*[1,2]}, \boldsymbol{\psi}^{*[1',2']})$.

5. DISCUSSION AND CONCLUSIONS

In this paper, we presented a structural result for decentralized control in partially nested information structures with a common ancestor through the prescription approach. We showed that prescriptions can equivalently be made with respect to various different permutations of agents. This allows us to combine the structural results obtained from different permutations to reap the maximum benefit as illustrated through an example. We believe that a similar approach can be implemented in other systems as well.

We note that the prescription approach and structural results do not depend on the information structure of the system beyond defining the memory of the agents and the permutations. Thus, ongoing work includes simplifying these results and applying the prescription approach to other decentralized control problems.

REFERENCES

Dave, A. and Malikopoulos, A. (2019). The prescription approach to decentralized stochastic control with word-of-mouth communication. *arXiv e-prints*, arXiv:1907.12125.

Ho, Y.C. and Chu, K.C. (1972). Team decision theory and information structures in optimal control problems–Part I. *IEEE Transactions on Automatic Control*, 17(1), 15–22. doi:10.1109/TAC.1972.1099850.

Mahajan, A. (2008). Sequential decomposition of sequential dynamic teams: Applications to real-time communication and networked control systems.

Mahajan, A., Martins, N.C., Rotkowitz, M.C., and Yüksel, S. (2012). Information structures in optimal decentralized control. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 1291–1306. IEEE.

Mahajan, A. and Nayyar, A. (2015). Sufficient statistics for linear control strategies in decentralized systems with partial history sharing. *IEEE Transactions on Automatic Control*, 60(8), 2046–2056.

Nayyar, A. and Lessard, L. (2015). Structural results for partially nested lqg systems over graphs. In *2015 American Control Conference (ACC)*, 5457–5464. IEEE.

Nayyar, A., Mahajan, A., and Teneketzis, D. (2011). Optimal control strategies in delayed sharing information structures. *IEEE Transactions on Automatic Control*, 56(7), 1606–1620.

Witsenhausen, H.S. (1968). A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1), 131–147.

Wu, J. and Lall, S. (2010). A dynamic programming algorithm for decentralized markov decision processes with a broadcast structure. In *49th IEEE Conference on Decision and Control (CDC)*, 6143–6148. IEEE.

Wu, J. and Lall, S. (2014). A theory of sufficient statistics for teams. In *53rd IEEE Conference on Decision and Control*, 2628–2635. IEEE.

Yuksel, S. (2009). Stochastic nestedness and the belief sharing information pattern. *IEEE Transactions on Automatic Control*, 54(12), 2773–2786.