Structural Results for Decentralized Stochastic Control with a Word-of-Mouth Communication

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

***Abstract*— In this paper, we analyze a network of agents that communicate through the "word of mouth," in which, every agent communicates only with its neighbors. We introduce the prescription approach, present some of its properties and show that it leads to a new information state. We also state preliminary structural results for optimal control strategies in systems that evolve using word-of-mouth communication. The proposed approach can be generalized to analyze several decentralized systems.**

I. INTRODUCTION

Centralized stochastic control has been the ubiquitous approach to control complex systems so far [1], [2]. A key assumption in centralized stochastic control problems is that a singular decision maker perfectly recalls all previous control actions and observations. The information available to an agent when making a decision is called the *information structure* of the system, and complete information corresponds to the *classical information structure*.

However, the classical information structure does not apply to many applications involving multiple agents [3]. In these applications, all agents simultaneously make a decision based only on their local observations and delayed, or costly communication with others [4]. Information structures where the information available to different agents is different, are classified as decentralized stochastic control problems with *non-classical information structures*.

The challenge of non-classical information structures lies in the fact that they cannot easily be analysed using dynamic programming (DP), due to a lack of separation between estimation and control. Next we describe three general approaches in the literature on decentralized stochastic control that generalize techniques from centralized stochastic control. For details on these approaches, the reader may refer to the tutorial by Mahajan et al. [5] and the references therein.

In the *person-by-person approach*, the control strategies of all agents except one are arbitrarily fixed. Then the control strategy of the chosen agent is then optimized as a centralized problem. Repeating this process for all agents allows for the derivation of *structural results* and DP for a person-by-person optimal strategy, that is not globally optimal in general. Some applications of this approach can be found in [6]–[11].

The *designer's approach* takes the point of view of a designer with knowledge of the system model and statistics. The designer's task is the selection of the globally optimal control strategy for the system by transforming the problem into a centralized planning problem. Some applications of this approach can be found in [12]–[14].

A more recent development in this field is the *common information approach* developed for problems with partial history sharing [15], and then formalized for general decentralized systems [16]. The solution is derived by reformulating the system from the viewpoint of a fictitious *coordinator* whose task is to prescribe control laws to every agent in the system. Further applications can be found for information structures with mean-field sharing [17], unreliable channels [18], and stochastic nestedness [19].

In this paper, we introduce and study a decentralized system with multiple agents who communicate with *word of mouth*. A word-of-mouth communication is characterized by a network of agents, where each agent directly communicates in a delayed manner, only with its neighbors in the network. Information originating from any agent propagates in the network through the subsequent sharing by their neighbors, and so on. This problem has a non-classical information structure because of the delays in communication.

The common information approach is considered to be the standard approach in solving decentralized control problems like the word-of-mouth information structure. However, because we let the links have asymmetric delays, the maximum delay in communication across the network can be quite large for our problem. We believe that in such problems, we can improve the structural results by taking into account the asymmetries in the system. Our primary contributions are the formulation of the problem with the word-of-mouth information structure and the introduction of the prescription approach. In addition, we state without proof, one structural result with time-invariant control policies for all agents that arises from the prescription approach.

The rest of the paper is organized as follows: In Section II, we present the information structure and problem formulation. In Section III, we present the prescription approach along with the corresponding properties. In Section IV, we state the structural results. Finally, in Section V, we draw concluding remarks, and present some ideas for future work.

II. PROBLEM FORMULATION

A. The Network of Agents

We consider a network of $K \in \mathbb{N}$ agents represented by a strongly connected graph $\mathcal{G} = (\mathcal{K}, \mathcal{E})$, where $\mathcal{K} := \{1, \dots, K\}$ is the set of agents and $\mathcal{E}$ is the set of links. A link from an agent $k \in \mathcal{K}$ to an agent $j \in \mathcal{K}$ is denoted

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

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

by $(k,j) \in \mathcal{E}$ and represents a communication link from agent k to j which is characterized by a delay of $\delta^{[k,j]} \in \mathbb{N}$ time steps. When agent k sends out information to agent j through link (k,j), we call it *transmission of information*. The information transmitted by agent k at time t is received by agent j at time $t + \delta^{[k,j]}$.

Definition 1. Let $\mathcal{N} = \{1, \ldots, m : m \in \mathcal{K}\}$ be a set of indices. For any $k, j \in \mathcal{K}$, a *path* $q_a^{[k,j]}$, $a \in \mathbb{N}$, from k to j 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 $n \in \mathcal{N} \setminus \{1\}$.

The set $\mathcal{Q}^{[k,j]} = \{q_a^{[k,j]} : a = 1, \ldots, b;\ b \in \mathbb{N}\}$ includes all paths from agent k to agent j.

Definition 2. Let agents $k, j \in \mathcal{K}$ with a path $q_a^{[k,j]}$ from k to j. The *communication delay* $d_a^{[k,j]} \in \mathbb{N}$ for $q_a^{[k,j]}$ is defined as

$$d_a^{[k,j]} = \delta^{[k,k_2]} + \cdots + \delta^{[k_{m-1},j]},$$

where $\delta^{[k_{n-1},k_n]}$ is the delay in information transfer through the link $(k_{n-1}, k_n) \in \mathcal{E}$.

The *information path*, defined formally next, from agent k to agent j in the network is the path with the least possible delay.

Definition 3. The information path from k to j denoted by $(k \to j)$ is given by a path $q_a^{[k,j]} \in \mathcal{Q}^{[k,j]}$ such that,

$$d_a^{[k,j]} = \min\left\{d_1^{[k,j]}, \ldots, d_b^{[k,j]}\right\}, \tag{1}$$

where $b := |\mathcal{Q}^{[k,j]}|$.

The strongly connected nature of the network ensures that there is always an information path $(k \to j)$ from any agent $k \in \mathcal{K}$ to any agent $j \in \mathcal{K}$. We denote the associated delay by $d^{[k,j]}$ and, by convention, set $d^{[k,k]} = 0$. In general, we allow $d^{[k,j]} \neq d^{[j,k]}$ for directed paths with delay.

B. System Description

The network of agents is considered a discrete time system that evolves up to a finite time horizon $T \in \mathbb{N}$. At time $t \in \mathcal{T}$, $\mathcal{T} = \{0, 1, \ldots, T\}$, the state of the system X_t takes values in a finite set $\mathcal{X}$ and the control variable U_t^k associated with agent $k \in \mathcal{K}$, takes values in a finite set $\mathcal{U}^k$. Let $U_t^{1:K}$ denote the vector $(U_t^1, \ldots, U_t^K)$. Starting at the initial state X_0, the evolution of the system follows the state equation

$$X_{t+1} = f_t\left(X_t, U_t^{1:K}, W_t\right), \tag{2}$$

where W_t is a random variable taking values in a finite set $\mathcal{W}$ that denoted the uncontrolled disturbance to the system. At time t the agent k makes an observation Y_t^k, given by

$$Y_t^k = h_t^k(X_t, V_t^k), \tag{3}$$

which takes values in a finite set $\mathcal{Y}^k$, where V_t^k takes values in the finite set $\mathcal{V}^k$ and denotes the noise in measurement.

Agent k selects a control action U_t^k from the set of feasible control actions $\mathcal{U}_t^k$ as a function of its information structure. The information available to each agent $k \in \mathcal{K}$ at time t is different, as discussed in Section II-C. After each agent k generates a control action U_t^k, the system incurs a cost $c_t(X_t, U_t^{1:K}) \in \mathbb{R}$. Then, we impose the following assumptions on our model:

Assumption 1. The network topology is arbitrary, known a priori, and does not change with time.

The known and invariable network topology is a part of the system dynamics.

Assumption 2. The disturbance $\{W_t : t \in \mathcal{T}\}$ and noise $\{V_t^k : t \in \mathcal{T},\ k \in \mathcal{K}\}$ are sequences of independent random variables that are also independent of each other and of the initial state X_0.

The disturbance, noise, and initial state denote the *primitive* random variables, and they have known probability distributions.

Assumption 3. The dynamics $\{f_t, h_t^k, c_t : t \in \mathcal{T},\ k \in \mathcal{K}\}$, and the set of feasible policies G are known to all agents.

These functions and the set of feasible control policies form the basis of the decision making problem.

Assumption 4. Each agent has perfect recall.

Perfect recall of the data from the memory of every agent is an essential assumption for the structural results.

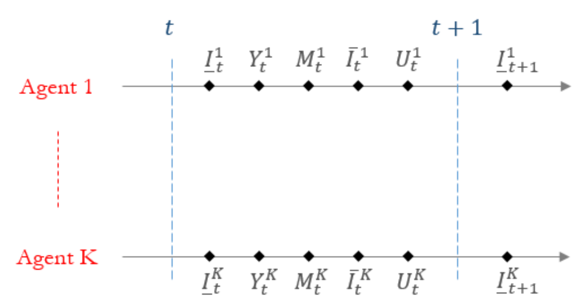

Fig. 1: Sequence of activities.

We summarize below the sequence of activities taken by agent $k \in \mathcal{K}$ at time t (Fig. 1):

1) The state X_t is updated based on (2).
2) Agent k receives information from all agents in $\mathcal{K}$, collectively denoted by $\underline{I}_t^k$.
3) Agent k makes an observation about the state Y_t^k based on (3).
4) Agent k updates its memory, M_t^k, defined in Section II-D, on a given protocol.
5) Agent k transmits information denoted by $\overline{I}_t^k$ to every agent $j \in \mathcal{K}$ through the shortest path $q_a^{[k,j]}$.
6) Agent k generates a control action U_t^k.

C. Information Structure of the System

The information structure of the system is characterized by the graph topology and delays along communication paths. In the word-of-mouth information structure, every agent $j \in \mathcal{K}$ at time t transmits the information $\overline{I}_t^j := \{Y_t^j, U_{t-1}^j\}$ to every other agent in the network through the relevant information paths. Agent $k \in \mathcal{K}$ receives information $\overline{I}_t^j$ at time $t + d^{[j,k]}$, where $d^{[j,k]}$ is the communication delay from j to k. Then, the *memory* of agent k at time t includes all information he received from every agent $j \in \mathcal{K}$ at time steps 0 through t.

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

$$M_t^k := \left\{ Y^j_{0:t-d^{[j,k]}}, U^j_{0:t-d^{[j,k]}-1} : j \in \mathcal{K} \right\}, \tag{4}$$

where $d^{[j,k]}$ is the delay in information transfer from every agent $j \in \mathcal{K}$ to agent k.

At time t, agent k accesses his memory M_t^k to generate a control action, namely,

$$U_t^k := g_t^k(M_t^k), \tag{5}$$

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

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

$$\textbf{Problem 1:} \qquad \mathcal{J}(\boldsymbol{g}) = \mathbb{E}^{\boldsymbol{g}} \left[\sum_{t=0}^{T} c_t(X_t, U_t^{1:K}) \right], \tag{6}$$

where the expectation is with respect to the joint probability measure on the random variables $\{X_t, U_t^1, \dots, U_t^K\}$. Then, the optimization problem is to select the optimal control policy $\boldsymbol{g}^* \in G$ that minimizes the performance criterion in (6), given the probability distributions of the primitive random variables $\{X_0, W_{0:T}, V^1_{0:T}, \dots, V^K_{0:T}\}$, and functions $\{c_t, f_t, h_t^k : t \in \mathcal{T}, k \in \mathcal{K}\}$.

III. THE PRESCRIPTION APPROACH

A. Construction of Prescriptions

For an agent $k \in \mathcal{K}$, we consider a scenario where the control action U_t^k is generated in two stages:

(1) Agent k generates a function based on information that is a subset of its memory M_t^k.

(2) This function takes as an input the compliment of the subset used to generate it, and yields the control action U_t^k.

We call these functions *prescriptions*. We show in Section III-B that they allow us to construct an optimization problem of selecting the optimal *prescription strategy* instead of the optimal control policy $\boldsymbol{g}^{*k}$. In this section, we construct the prescriptions for the agent k without changing the information structure of the system. We begin by defining the set of agents located beyond agent k to simplify the notation.

Definition 5. For an agent $k \in \mathcal{K}$, the set of agents beyond k is defined as $\mathcal{B}^k := \{j \in \mathcal{K} : j \geq k\}$.

Now we can define the information used to generate prescriptions.

Definition 6. Let $k \in \mathcal{K}$ and M_t^k be the agent's memory at time t. The *accessible information* of agent k is defined as the set A_t^k that takes values in the finite collection of sets $\mathcal{A}_t^k$ such that

$$A_t^k = \bigcap_{i=1}^{k} \left(M_t^i \right). \tag{7}$$

For example, we can write (7) for agents 1 and 2 as

$$A_t^1 = M_t^1, \tag{8}$$

$$A_t^2 = M_t^1 \cap M_t^2. \tag{9}$$

Based on Definition 6, the accessible information A_t^k has the following properties:

$$A_{t-1}^k \subset A_t^k, \tag{10}$$

$$A_t^j \subset A_t^k, \ \forall j \in \mathcal{B}^k, \tag{11}$$

where $\mathcal{B}^k$ is the set of agents beyond k. Property (10) motivates the introduction of a new term to denote the new information added to accessible information A_t^k at time t.

Definition 7. The *new information* for agent k at time t is defined as the set Z_t^k that takes values in a finite collection of sets $\mathcal{Z}_t^k$ such that

$$Z_t^k := A_t^k \backslash A_{t-1}^k. \tag{12}$$

We observe in (11) that the accessible information A_t^j of any agent $j \in \mathcal{B}^k$ is a subset of the memory M_t^k. Thus, we can define the *inaccessible information* of the agent k with respect to the accessible information A_t^j for every $j \in \mathcal{B}^k$.

Definition 8. The *inaccessible information* of agent k with respect to accessible information A_t^j, $j \in \mathcal{B}^k$, is defined as the set of random variables $L_t^{[k,j]}$ that takes values in the finite collection of sets $\mathcal{L}_t^{[k,j]}$ such that

$$L_t^{[k,j]} := M_t^k \setminus A_t^j. \tag{13}$$

The pair of sets A_t^j and $L_t^{[k,j]}$ forms a partition of the set M_t^k, such that

$$M_t^k = \{L_t^{[k,j]}, A_t^j\}, \ \forall j \in \mathcal{B}^k. \tag{14}$$

As an example, consider a system with three agents Fig. 2. Here, we have the following relationships for agent 1:

$$\begin{aligned} &A_t^1 = M_t^1, \\ &A_t^3 \subset A_t^2 \subset M_t^1, \\ &M_t^1 = \{A_t^2, L_t^{[1,2]}\} = \{A_t^3, L_t^{[1,3]}\}. \end{aligned} \tag{15}$$

Now we use these partitions of the memory to define the *prescription function*.

Definition 9. The *prescription function* $\Gamma_t^{[k,j]}$ of an agent $k \in \mathcal{K}$ for the agent $j \in \mathcal{K}$ is defined as follows

$$\Gamma_t^{[k,j]} : \begin{cases} \mathcal{L}_t^{[j,k]} \longrightarrow \mathcal{U}_t^j, \ \text{if } j \notin \mathcal{B}^k, \\ \mathcal{L}_t^{[j,j]} \longrightarrow \mathcal{U}_t^j, \ \text{if } j \in \mathcal{B}^k, \end{cases} \tag{16}$$

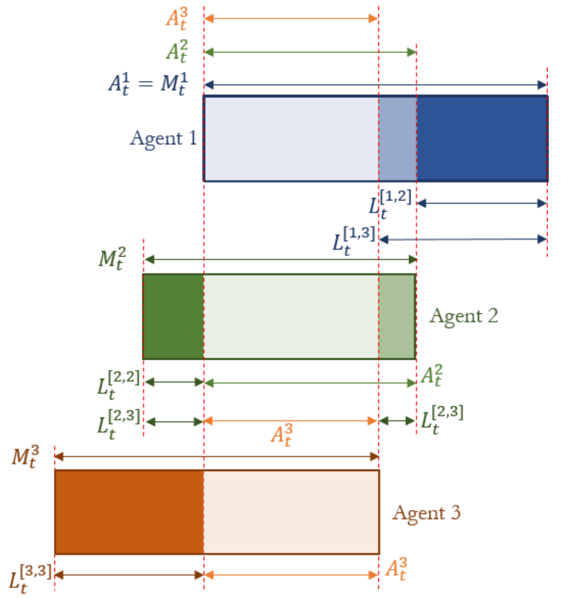

Fig. 2: Memory partitions of three agents.

and takes values in the set of feasible prescription functions $\mathscr{G}_t^{[k,j]}$.

Remark 1. In Definition 9, the inaccessible information of agent k is defined with respect to the accessible information A_t^j for $j \in \mathcal{B}^k$. Note that in the first part of (16), we have $k \in \mathcal{B}^j$, and thus (16) holds.

Every prescription function $\Gamma^{[k,j]}$ is generated as follows

$$\Gamma_t^{[k,j]} := \begin{cases} \psi_t^{[k,j]}(A_t^k), & \text{if } j \notin \mathcal{B}^k, \\ \psi_t^{[k,j]}(A_t^j), & \text{if } j \in \mathcal{B}^k, \end{cases} \tag{17}$$

where we call $\psi_t^{[k,j]}$ the *prescription strategy* of the agent k for the agent j given by the mapping

$$\psi_t^{[k,j]} : \begin{cases} \mathcal{A}_t^k \longrightarrow \mathscr{G}_t^{[k,j]}, & \text{if } j \notin \mathcal{B}^k, \\ \mathcal{A}_t^j \longrightarrow \mathscr{G}_t^{[k,j]}, & \text{if } j \in \mathcal{B}^k. \end{cases} \tag{18}$$

We call $\boldsymbol{\psi}^k := (\boldsymbol{\psi}^{[k,1]}, \ldots, \boldsymbol{\psi}^{[k,K]})$ the *prescription strategy* of the agent k. The set of feasible prescription strategies for the agent k is denoted by Ψ^k. The *complete prescription* of an agent k is defined next.

Definition 10. The *complete prescription* for agent k is given by the function

$$\Theta_t^k : \mathcal{L}_t^{[1,k]} \times \cdots \times \mathcal{L}_t^{[k,k]} \times \mathcal{L}_t^{[k+1,k+1]} \times \cdots \times \mathcal{L}_t^{[K,K]} \longrightarrow \mathcal{U}_t^1 \times \cdots \times \mathcal{U}_t^k, \tag{19}$$

which takes values in the set of functions $\mathscr{G}_t^k$.

The complete prescription for agent k is constructed as $\Theta_t^k = (\Gamma_t^{[k,1]}, \ldots, \Gamma_t^{[k,K]})$.

Remark 2. The prescription $\Gamma_t^{[k,j]}$ of agent k for agent j is only available to agent k. The equivalent prescription available to agent j is $\Gamma_t^{[j,j]}$. The relationship between the two is given in Lemmas 3 and 4 in Section III-B.

Remark 3. Every agent needs to generate prescriptions corresponding to every other agent in the system so that we can define the *information state* in Section IV-B.

B. Properties of the Prescription Approach

In this section, we present the relationships between the different prescriptions and control policies. The first result states that for an agent $k \in \mathcal{K}$ we can use the complete prescription Θ_t^k to generate control action U_t^k.

Lemma 1. *Let agent $k \in \mathcal{K}$ and let Θ_t^k be its complete prescription. For any given control policy $\boldsymbol{g} \in G$, there exists a prescription strategy $\boldsymbol{\psi}^k \in \Psi^k$ such that*

$$U_t^k = \Gamma_t^{[k,k]} \left(L_t^{[k,k]} \right). \tag{20}$$

Proof. Due to space limitation, the proof has been omitted but can be found in the extended version of the paper [20]. □

Similarly, for any prescription strategy $\boldsymbol{\psi}^k$, we can construct an appropriate control policy $\boldsymbol{g}$ that generates the same control actions U_t^k for all agents in $\mathcal{K}$.

Lemma 2. *Let agent $k \in \mathcal{K}$ and let Θ_t^k be its complete prescription. For any given prescription strategy $\boldsymbol{\psi}^k \in \Psi^k$, there exists a control policy $\boldsymbol{g} \in G$ such that*

$$U_t^k = \Gamma_t^{[k,k]}(L_t^{[k,k]}) = g_t^k(M_t^k). \tag{21}$$

Proof. Due to space limitation, the proof has been omitted but can be found in the extended version of the paper [20]. □

Lemmas 1 and 2 imply that the control action U_t^k of every agent $k \in \mathcal{K}$ generated through a prescription strategy $\boldsymbol{\psi}^k$, can also be generated through an appropriate policy $\boldsymbol{g}$ and vice versa.

Definition 11. Given two agents $k, j \in \mathcal{K}$, a *positional relationship* from agent k to agent j is given by the function

$$e^{[j,k]} : \Psi^k \longrightarrow \Psi^j. \tag{22}$$

Next we show the existence of a positional relationship $e^{[j,k]}$ from any agent $k \in \mathcal{K}$ to every agent $j \in \mathcal{K}$ with desirable properties that allow us to construct optimal control policies of *all* agents from the optimal prescription strategy of just *one* agent. The following result establishes that using a positional relationship $e^{[j,k]} = (e_1^{[j,k]}, \ldots, e_T^{[j,k]})$, the agent j can derive the prescription strategy for agent $i \in \mathcal{K}$, when given the prescription strategy of agent k for agent i, namely

$$\psi_t^{[j,i]} := e_t^{[j,k]} \left(\psi_t^{[k,i]} \right), \ \forall i \in \mathcal{K}. \tag{23}$$

Lemma 3. *Let agent $k \in \mathcal{K}$ and agent $j \in \mathcal{B}^k$. For any given prescription strategy $\boldsymbol{\psi}^k$ of agent k, there exists a positional relationship $e^{[j,k]}$ such that a prescription strategy $\boldsymbol{\psi}^j$ of agent j generated using (23) yields:*

1. $\Gamma_t^{[k,i]}(L_t^{[i,i]}) = \Gamma_t^{[j,i]}(L_t^{[i,i]}), \text{ if } i \in \mathcal{B}^j,$
2. $\Gamma_t^{[k,i]}(L_t^{[i,i]}) = \Gamma_t^{[j,i]}(L_t^{[i,j]}), \text{ if } i \in \mathcal{B}^k, i \notin \mathcal{B}^j,$
3. $\Gamma_t^{[k,i]}(L_t^{[i,k]}) = \Gamma_t^{[j,i]}(L_t^{[i,j]}), \text{ if } i \notin \mathcal{B}^k.$ (24)

Proof. Due to space limitation, the proof has been omitted but can be found in the extended version of the paper [20]. □

Lemma 4. *Let agents $k, j \in \mathcal{K}$ with $j \notin \mathcal{B}^k$. For any given prescription strategy $\boldsymbol{\psi}^k$ of agent k, there exists a positional relationship $e^{[j,k]}$ such that a prescription strategy $\boldsymbol{\psi}^j$ of agent j generated from* (23) *yields:*

$$\begin{aligned} &1.\ \Gamma_t^{[k,i]}(L_t^{[i,i]}) = \Gamma_t^{[j,i]}(L_t^{[i,i]}),\ \text{if } i \in \mathcal{B}^k, \\ &2.\ \Gamma_t^{[k,i]}(L_t^{[i,k]}) = \Gamma_t^{[j,i]}(L_t^{[i,i]}),\ \text{if } i \in \mathcal{B}^j, i \notin \mathcal{B}^k, \\ &3.\ \Gamma_t^{[k,i]}(L_t^{[i,k]}) = \Gamma_t^{[j,i]}(L_t^{[i,j]}),\ \text{if } i \notin \mathcal{B}^j. \end{aligned} \tag{25}$$

Proof. The proof is very similar to the proof of Lemma 3. It is omitted due to space limitations. □

To this end, we consider a positional relationship function $e^{[j,k]}$ from every $k \in \mathcal{K}$ to every position $j \in \mathcal{K}$ which satisfies the properties in Lemmas 3 and 4. This implies that for any two agents k and j, we have the relation,

$$U_t^j = \Gamma_t^{[j,j]}(L_t^{[j,j]}) = \begin{cases} \Gamma_t^{[k,j]}(L_t^{[j,k]}), \text{ if } j \notin \mathcal{B}^k, \\ \Gamma_t^{[k,j]}(L_t^{[j,j]}), \text{ if } j \in \mathcal{B}^k. \end{cases} \tag{26}$$

IV. RESULTS

A. Equivalent Prescription Problems

Lemmas 1 through 4 lead to (26). This implies that the control action U_t^j for an agent $j \in \mathcal{K}$ can be equivalently obtained through the prescription function $\Gamma_t^{[k,j]}$ of another agent $k \in \mathcal{K}$, if the inaccessible information is available. Using (26), the cost to the system at time t is given by

$$\begin{aligned} &c_t(X_t, U_t^1, \dots, U_t^K) \\ &=: c_t\big(X_t, \Gamma_t^{[k,1]}(L_t^{[1,k]}), \dots, \Gamma_t^{[k,k]}(L_t^{[k,k]}), \\ &\qquad \Gamma_t^{[k,k+1]}(L_t^{[k+1,k+1]}), \dots, \Gamma_t^{[k,K]}(L_t^{[K,K]})\big). \end{aligned} \tag{27}$$

We then reformulate Problem 1 in terms of the prescription strategy of the agent k. The optimization problem is to select the optimal prescription strategy $\boldsymbol{\psi}^{*k} \in \Psi^k$ that minimizes the performance criterion, given by

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

$$\begin{aligned} &\mathbb{E}^{\boldsymbol{\psi}^k}\Big[\sum_{t=0}^{T} c_t\big(X_t, \Gamma_t^{[k,1]}(L_t^{[1,k]}), \dots, \Gamma_t^{[k,k]}(L_t^{[k,k]}), \\ &\qquad \Gamma_t^{[k,k+1]}(L_t^{[k+1,k+1]}), \dots, \Gamma_t^{[k,K]}(L_t^{[K,K]})\big)\Big]. \end{aligned} \tag{28}$$

The task of deriving optimal prescription strategy $\boldsymbol{\psi}^{*k}$, and subsequently, the complete prescription Θ_t^k for agent k is to be achieved with access only to the memory M_t^k. This maintains the decentralized nature of the problem as the real-time implementation of the strategies is still decentralized. Now, we show the equivalence between the two problems.

Lemma 5. *For any agent $k \in \mathcal{K}$, Problem 2 is equivalent to Problem 1.*

Proof. The proof follows from Lemmas 1 and 2. Due to space limitation, the proof has been omitted but can be found in the extended version of the paper [20]. □

Next, we present a state sufficient for input-output mapping in Problem 2 for agent k following the exposition presented in [21].

Lemma 6. *A state sufficient for input-output mapping for agent $k \in \mathcal{K}$ is*

$$S_t^k := \Big\{X_t, L_t^{[1,k]}, \dots, L_t^{[k-1,k]}, L_t^{[k,k]}, \dots, L_t^{[K,K]}\Big\}. \tag{29}$$

Proof. The state S_t^k satisfies the three sufficient properties:

1) There exist functions $\{\hat{f}_t^k\colon t \in \mathcal{T}\}$ such that

$$S_{t+1}^k = \hat{f}_t^k(S_t^k, W_t, V_{t+1}^{1:K}, \Theta_t^k). \tag{30}$$

2) There exist functions $\{\hat{h}_t^k\colon t \in \mathcal{T}\}$ such that

$$Z_{t+1}^k = \hat{h}_t^k(S_t^k, \Theta_t^k, V_{t+1}^{1:K}). \tag{31}$$

3) There exist functions $\{\hat{c}_t^k\colon t \in \mathcal{T}\}$ such that

$$c_t(X_t, U_t^{1:K}) = \hat{c}_t^k(S_t^k, \Theta_t^k). \tag{32}$$

The three equations above can each be verified by substitution of variables on the LHS. The complete proof can be found in [22]. □

B. The Information States

For an agent $k \in \mathcal{K}$, in Problem 2 the system is characterized by the state S_t^k, control input Θ_t^k, output Z_t^k (with $Z_{0:t}^k = A_t^k$) and cost $\hat{c}_t^k(S_t^k, \Theta_t^k)$ at time t. The prescription functions $\Gamma_t^{[k,j]}$, $j \in \mathcal{B}^k$, are generated as functions of the accessible information A_t^j. Then, we define the *information state* for agent k below.

Definition 12. Let S_t^k be the state, A_t^k the accessible information, and $\Theta_{0:t-1}^k$ the control inputs at time t an agent $k \in \mathcal{K}$. The *information state* is defined as a probability distribution Π_t^k that takes values in the possible realizations $\mathscr{P}_t^k := \Delta(\mathcal{S}_t^k)$ such that,

$$\Pi_t^k(s_t^k) := \mathbb{P}^{\boldsymbol{\psi}^k}(S_t^k = s_t^k \,|\, A_t^k, \Theta_{0:t-1}^k). \tag{33}$$

The information state Π_t^k is independent from the prescription strategy $\boldsymbol{\psi}^k$, its evolution is Markovian, and it is sufficient along with the prescription Θ_t^k to express the expected cost to the system at time t. Thus, the information state Π_t^k evolves as a controlled Markov chain with control inputs Θ_t^k. Due to space limitation, the proofs of these three properties are omitted, but can be found in [22].

C. Structural Results

We start by presenting a structural result for agent K. The set of agents beyond agent K is $\mathcal{B}^K = \{K\}$. Using (17), this implies that for all agents $k \in \mathcal{K}$, the prescription component $\Gamma_t^{[K,k]}$ is a function of the accessible information $A_t^{[K]}$. This leads to the structural result of the common information approach in [16].

Lemma 7. *Consider agent K. There exists an optimal prescription strategy $\boldsymbol{\psi}^{*K}$ of the form*

$$\Gamma_t^{*[K,k]} = \psi_t^{*[K,k]}(\Pi_t^K), \tag{34}$$

that optimizes the performance criterion (28) *in Problem 2.*

We know that for any two agents $k \in \mathcal{K}$ and $j \in \mathcal{B}^k$, we have $A_t^j \subset A_t^k$. Then, the prescription approach leads to the following structural result, proved in [22].

Theorem 1. *Consider agent $k \in \mathcal{K}$. There exists an optimal prescription strategy $\boldsymbol{\psi}^{*k}$ of the form*

$$\Gamma_t^{*[k,j]}(\cdot) = \begin{cases} \psi_t^{*[k,j]}(\Pi_t^k, \ldots, \Pi_t^K), \text{ if } j \notin \mathcal{B}^k, \\ \psi_t^{*[k,j]}(\Pi_t^j, \ldots, \Pi_t^K), \text{ if } j \in \mathcal{B}^k, \end{cases} \tag{35}$$

that optimizes the performance criterion (28) *in Problem 2.*

D. A Comparison with Existing Approaches

In the existing literature, the graphical approach presented in [23] has similarities with the prescription approach. However, it applies only to problems where agents have perfect observations. Meanwhile, the common information approach in [16] can be applied to the our system, but it leads to the result in Lemma 7. Then the control action for an agent $k \in \mathcal{K}$ is given by

$$U_t^{*k} = \Gamma_t^{*[K,k]}(L_t^{[k,K]}). \tag{36}$$

In contrast, we see that when we consider Problem 2 for agent k, the control action of agent k is given by

$$U_t^{*k} = \Gamma_t^{*[k,k]}(L_t^{[k,k]}). \tag{37}$$

Now, from (14), we note that,

$$A_t^k \cup L_t^{[k,k]} = A_t^K \cup L_t^{[k,K]}, \tag{38}$$

and from (11) we have the relation,

$$A_t^K \subset A_t^k, \tag{39}$$

because $K \in \mathcal{B}^k$ for all $k \in \mathcal{K}$. Then, (38) and (39) imply,

$$L_t^{[k,k]} \subset L_t^{[k,K]}. \tag{40}$$

Thus, the prescription functions generated through Theorem 1 have an an equal or smaller domain when compared with those generated through Lemma 10.

V. CONCLUSIONS

In this paper, we introduce a network of agents with a word-of-mouth communication structure, and analyze it using the prescription approach, which yielded some desired properties. We showed that the structural result derived through the common information approach can be considered as the outcome of one reformulations using the prescription approach. Finally, we provided, without proof, a preliminary structural result arising from the prescription approach. A direction for future research should seek to extend these results for a broader class of decentralized systems.

REFERENCES

[1] A. A. Malikopoulos, V. Maroulas, and J. XIong, "A multiobjective optimization framework for stochastic control of complex systems," in *Proceedings of the 2015 American Control Conference*, pp. 4263–4268, 2015.

[2] A. A. Malikopoulos, "A duality framework for stochastic optimal control of complex systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2756–2765, 2016.

[3] A. A. Malikopoulos, C. G. Cassandras, and Y. J. Zhang, "A decentralized energy-optimal control framework for connected automated vehicles at signal-free intersections," *Automatica*, vol. 93, no. April, pp. 244–256, 2018.

[4] A. A. Malikopoulos, C. Charalambous, and I. Tzortzis, "The average cost of markov chains subject to total variation distance uncertainty," in *Systems & Control Letters*, vol. 120, pp. 29–35, 2018.

[5] 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.

[6] H. Witsenhausen, "On the structure of real time source coders," *Bell Syst. Tech. J*, vol. 58, no. 6, pp. 1437–1451, 1979.

[7] A. Nayyar and D. Teneketzis, "On jointly optimal real-time encoding and decoding strategies in multi-terminal communication systems," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 1620–1627, 2008.

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

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

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

[11] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems–Part I," *IEEE Transactions on Automatic Control*, vol. 17, no. 1, pp. 15–22, 1972.

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

[13] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal Strategies for Communication and Remote Estimation With an Energy Harvesting Sensor," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2246–2260, 2013.

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

[15] 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.

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

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

[18] S. M. Asghari, Y. Ouyang, and A. Nayyar, "Optimal local and remote controllers with unreliable uplink channels," *IEEE Transactions on Automatic Control*, vol. 64, no. 5, pp. 1816–1831, 2018.

[19] S. Yuksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2773–2786, 2009.

[20] A. Dave and A. Malikopoulos, "Structural results for decentralized stochastic control with a word-of-mouth communication," *arXiv e-prints*, p. arXiv:1809.08712, Sep 2018.

[21] A. Mahajan, *Sequential Decomposition of Sequential Dynamic Teams: Applications to Real-Time Communication and Networked Control Systems.* PhD thesis, University of Michigan, 2008.

[22] A. Dave and A. Malikopoulos, "The prescription approach to decentralized stochastic control with word-of-mouth communication," *arXiv e-prints*, p. arXiv:1907.12125, Sep 2019.

[23] A. Mahajan and S. Tatikonda, "An algorithmic approach to identify irrelevant information in sequential teams," *Automatica*, vol. 61, pp. 178–191, 2015.