

A Structural Result for Delayed Sharing Information Structures

Ashutosh Nayyar
Department of EECS,
University of Michigan,
Ann Arbor, MI, 48109.
Email: anayyar@umich.edu

Aditya Mahajan
Department of Electrical Engineering,
Yale University,
New Haven, CT, 06520.
Email: aditya.mahajan@yale.edu

Demosthenis Teneketzis
Department of EECS,
University of Michigan,
Ann Arbor, MI, 48109.
Email: teneket@eecs.umich.edu

Abstract—The n -step delayed sharing information structure is investigated. This information structure is a link between the classical information structure, where information is shared perfectly between the controllers, and a non-classical information structure, where there is no “lateral” sharing of information among the controllers. A structural result for optimal control policies in systems with such information structures is presented. A sequential methodology for finding optimal policies is also established. The solution approach provides an insight for identifying structural results for general decentralized stochastic control problems.

I. INTRODUCTION

A. Motivation

One of the difficulties in optimal design of decentralized control systems is handling the increase of data at the control stations with time. This increase in data means that the domain of control laws increases with time which, in turn, creates two difficulties. Firstly, the number of control strategies increases doubly exponentially with time; this makes it harder to search for an optimal strategy. Secondly, even if an optimal strategy is found, implementing functions with time increasing domain is difficult.

In centralized stochastic control [1], these difficulties can be circumvented by using the conditional probability of the state given the data available at the control station as a sufficient statistic (where the data available to a control station comprises of all observations and control actions till the current time). This conditional probability, called *information state*, takes values in a time-invariant space. Consequently, we can restrict attention to control laws with time-invariant domain. Such results, where data that is increasing with time is “compressed” to a sufficient statistic taking values in a time-invariant space, are called *structural results*. While the information state/structural result for centralized stochastic control problems is well known, no general methodology to find such information states or structural results exists for decentralized stochastic control problems.

In this paper, we find structural results for decentralized control systems with delayed sharing information structures. In a system with n -step delayed sharing, every control station knows the n -step prior observations and control actions of all other control stations. This information structure, proposed by

Witsenhausen in [2], is a link between the classical information structures, where information is shared perfectly among the controllers, and the non-classical information structures, where there is no “lateral” sharing of information among the controllers. Witsenhausen asserted a structural result for this model without any proof in his seminal paper [2]. Varaiya and Walrand [3] proved that Witsenhausen’s assertion was true for $n = 1$ but false for $n > 1$. For $n > 1$, Kurtaran [4] proposed another structural result. However, Kurtaran proved his result only for the terminal time step (that is, the last time step in a finite horizon problem); for non-terminal time steps, he gave an abbreviated argument, which we believe is incomplete.

We prove a structural result of the optimal control laws for the delayed sharing information structure. We compare our results to those conjectured by Witsenhausen and show that our structural results for n -step delay sharing information structure simplify to that of Witsenhausen for $n = 1$; for $n > 1$, our results are different from the result proposed by Kurtaran. We also present a sequential methodology to find the optimal control policies.

B. Notation

Random variables are denoted by upper case letters; their realization by the corresponding lower case letter. $X_{a:b}$ is a short hand for the vector $(X_a, X_{a+1}, \dots, X_b)$ while $X^{c:d}$ is a short hand for the vector $(X^c, X^{c+1}, \dots, X^d)$. The combined notation $X_{a:b}^{c:d}$ is a short hand for the vector $(X_i^j : i = a, a + 1, \dots, b, j = c, c + 1, \dots, d)$. $\mathbb{P}(\cdot)$ is the probability of an event, $\mathbb{E}\{\cdot\}$ is the expectation of a random variable. For a collection of functions \mathbf{g} , we use $\mathbb{P}^{\mathbf{g}}(\cdot)$ and $\mathbb{E}^{\mathbf{g}}\{\cdot\}$ to denote that the probability measure/expectation depends on the choice of functions in \mathbf{g} . $\mathbb{1}_A(\cdot)$ is the indicator function of a set A . For singleton sets $\{a\}$, we also denote $\mathbb{1}_{\{a\}}(\cdot)$ by $\mathbb{1}_a(\cdot)$. For a finite set A , $\mathcal{P}\{A\}$ denotes the space of probability mass functions on A . For convenience of exposition, we will assume all sets have finite cardinality.

C. Model

Consider a system consisting of a plant and K controllers with decentralized information. At time t , $t = 1, \dots, T$, the state of the plant X_t takes values in \mathcal{X} ; the control action U_t^k at station k , $k = 1, \dots, K$, takes values in \mathcal{U}^k . The initial

state X_0 of the plant is a random variable. With time, the plant evolves according to

$$X_t = f_t(X_{t-1}, U_t^{1:K}, V_t) \quad (1)$$

where V_t is a random variable taking values in \mathcal{V} . $\{V_t; t = 1, \dots, T\}$ is a sequence of independent random variables that are also independent of X_0 .

The system has K observation posts. At time t , $t = 1, \dots, T$, the observation Y_t^k of post k , $k = 1, \dots, K$, takes values in \mathcal{Y}^k . These observations are generated according to

$$Y_t^k = h_t^k(X_{t-1}, W_t^k) \quad (2)$$

where W_t^k are random variables taking values in \mathcal{W}^k . $\{W_t^k; t = 1, \dots, T; k = 1, \dots, K\}$ are independent random variables that are also independent of X_0 and $\{V_t; t = 1, \dots, T\}$.

The system has n -step delayed sharing. This means that at time t , control station k observes the current observation Y_t^k of observation post k , the n steps old observations $Y_{t-n}^{1:K}$ of all posts, and the n steps old actions $U_{t-n}^{1:K}$ of all stations. Each station has perfect recall; so, it remembers everything that it has seen and done in the past. Thus, at time t , data available at station k can be written as (Δ_t, Λ_t^k) , where

$$\Delta_t := (Y_{1:t-n}^{1:K}, U_{1:t-n}^{1:K})$$

is the data known to all stations and

$$\Lambda_t^k := (Y_{t-n+1:t}^k, U_{t-n+1:t-1}^k)$$

is the additional data known at station k , $k = 1, \dots, K$. Let \mathcal{D}_t be the space of all possible realizations of Δ_t ; and \mathcal{L}^k be the space of all possible realizations of Λ_t^k . Station k chooses action U_t^k according to a control law g_t^k , i.e.,

$$U_t^k = g_t^k(\Lambda_t^k, \Delta_t). \quad (3)$$

The choice of $\mathbf{g} = \{g_t^k; k = 1, \dots, K; t = 1, \dots, T\}$ is called a *design* or a *control strategy*. \mathcal{G} denotes the class of all possible designs. At time t , a cost $c_t(X_t, U_t^1, \dots, U_t^K)$ is incurred. The performance $\mathcal{J}(\mathbf{g})$ of a design is given by the expected total cost under it, i.e.,

$$\mathcal{J}(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left\{ \sum_{t=1}^T c_t(X_t, U_t^{1:K}) \right\} \quad (4)$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of \mathbf{g} . We consider the following problem.

Problem 1: Given the statistics of the primitive random variables X_0 , $\{V_t; t = 1, \dots, T\}$, $\{W_t^k; k = 1, \dots, K; t = 1, \dots, T\}$, the plant functions $\{f_t; t = 1, \dots, T\}$, the observation functions $\{h_t^k; k = 1, \dots, K; t = 1, \dots, T\}$, and the cost functions $\{c_t; t = 1, \dots, T\}$ choose a design \mathbf{g}^* from \mathcal{G} that minimizes the expected cost given by (4).

D. The structural results

Witsenhausen [2] asserted the following structural result for Problem 1.

Structural Result (Witsenhausen [2]): In Problem 1, without loss of optimality we can restrict attention to control strategies of the form

$$U_t^k = g_t^k(\Lambda_t^k, \mathbb{P}(X_{t-n} | \Delta_t)). \quad (5)$$

Witsenhausen's result claims that all control stations can "compress" the common information Δ_t to a sufficient statistic $\mathbb{P}(X_{t-n} | \Delta_t)$. Unlike Δ_t , the size of $\mathbb{P}(X_{t-n} | \Delta_t)$ does not increase with time.

As mentioned earlier, Witsenhausen asserted this result without a proof. Varaiya and Walrand [3] proved that the above separation result is true for $n = 1$ but false for $n > 1$. Kurtaran [4] proposed an alternate structural result for $n > 1$.

Structural Result (Kurtaran [4]): In Problem 1, without loss of optimality we can restrict attention to control strategies of the form

$$U_t^k = g_t^k(Y_{t-n+1:t}^k, \mathbb{P}^{g_{1:t-1}^{1:K}}(X_{t-n}, U_{t-n+1:t-1}^{1:K} | \Delta_t)). \quad (6)$$

Kurtaran used a different labeling of the time indices, so the statement of the result in his paper is slightly different from what we have stated above.

Kurtaran's result claims that all control stations can "compress" the common information Δ_t to a sufficient statistic $\mathbb{P}^{g_{1:t-1}^{1:K}}(X_{t-n}, U_{t-n+1:t-1}^{1:K} | \Delta_t)$, whose size does not increase with time.

Kurtaran proved his result for only the terminal time-step and gave an abbreviated argument for non-terminal time-steps. In this paper, we prove a new structural result.

Structural Result (this paper): In Problem 1, without loss of optimality we can restrict attention to control strategies of the form

$$U_t^k = g_t^k(\Lambda_t^k, \mathbb{P}^{g_{1:t-1}^{1:K}}(X_{t-1}, \Lambda_t^{1:K} | \Delta_t)). \quad (7)$$

This result claims that all control stations can "compress" the common information Δ_t to a sufficient statistic $\mathbb{P}^{g_{1:t-1}^{1:K}}(X_{t-1}, \Lambda_t^{1:K} | \Delta_t)$, whose size does not increase with time.

Our structural results cannot be derived from Kurtaran's result and vice-versa. At present, we are not sure of the correctness of Kurtaran's result. As we mentioned before, we believe that the proof given by Kurtaran is incomplete. We have not been able to complete Kurtaran's proof; neither have we been able to find a counterexample to his result.

II. STRUCTURAL RESULT

In this section, we prove the structural result (7) for optimal strategies of the K control stations. For the ease of notation, we first prove the result for $K = 2$, and then show how to extend it for general K . We refer the reader to [5] for detailed proofs of the results.

A. Two Controller system ($K = 2$)

The proof for $K = 2$ proceeds as follows:

- 1) First, we formulate a centralized stochastic control problem from the point of view of a coordinator who observes the shared information Δ_t , but does not observe the private information $(\Lambda_t^1, \Lambda_t^2)$ of the two controllers.
- 2) Next, we argue that any strategy for the coordinator's problem can be implemented in the original problem and vice versa. Hence, the two problems are equivalent.
- 3) Then, we identify states sufficient for input-output mapping for the coordinator's problem.
- 4) Finally, we transform the coordinator's problem into a MDP (Markov decision process), and obtain a structural result for the coordinator's problem. This structural result is also a structural result for the delayed sharing information structures due to the equivalence between the two problems.

Below, we elaborate on each of these stages.

Stage 1

We consider the following modified problem. In the model described in Section I-C, in addition to the two controllers, a coordinator that knows the common (shared) information Δ_t available to both controllers at time t is present. At time t , the coordinator decides the *partial functions*

$$\gamma_t^k : \mathcal{L}^k \mapsto \mathcal{U}^k$$

for each controller k , $k = 1, 2$. The choice of the partial functions at time t is based on the realization of the common (shared) information and the partial functions selected before time t . These functions map each controller's *private information* Λ_t^k to its control action U_t^k at time t . The coordinator then informs all controllers of all the partial functions it selected at time t . Each controller then uses its assigned partial function to generate a control action as follows.

$$U_t^k = \gamma_t^k(\Lambda_t^k). \quad (8)$$

The system dynamics and the cost are same as in the original problem. At next time step, the coordinator observes the new common observation

$$Z_{t+1} := \{Y_{t-n+1}^1, Y_{t-n+1}^2, U_{t-n+1}^1, U_{t-n+1}^2\}. \quad (9)$$

Thus at the next time, the coordinator knows $\Delta_{t+1} = Z_{t+1} \cup \Delta_t$ and its choice of all past partial functions and it selects the next partial functions for each controller. The system proceeds sequentially in this manner until time horizon T .

In the above formulation, the only decision maker is the coordinator: the individual controllers simply carry out the necessary evaluations prescribed by (8). At time t , the coordinator knows the common (shared) information Δ_t and all past partial functions $\gamma_{1:t-1}^1$ and $\gamma_{1:t-1}^2$. The coordinator uses a decision rule ψ_t to map this information to its decision, that is,

$$(\gamma_t^1, \gamma_t^2) = \psi_t(\Delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2), \quad (10)$$

or equivalently,

$$\gamma_t^k = \psi_t^k(\Delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2), \quad k = 1, 2. \quad (11)$$

The choice of $\psi = \{\psi_t; t = 1, \dots, T\}$ is called a *coordination strategy*. Ψ denotes the class of all possible coordination strategies. The performance of a coordinating strategy is given by the expected total cost under that strategy, that is,

$$\hat{\mathcal{J}}(\psi) = \mathbb{E}^\psi \left\{ \sum_{t=1}^T c_t(X_t, U_t^1, U_t^2) \right\} \quad (12)$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of ψ . The coordinator has to solve the following optimization problem.

Problem 2 (The Coordinator's Optimization Problem):

Given the system model of Problem 1, choose a coordination strategy ψ^* from Ψ that minimizes the expected cost given by (12).

Stage 2

We now show that the Problem 2 is equivalent to Problem 1. Specifically, we will show that any design g for Problem 1 can be implemented by the coordinator in Problem 2 with the same value of the problem objective. Conversely, any coordination strategy ψ in Problem 2 can be implemented in Problem 1 with the same value of the performance objective.

Any design g for Problem 1 can be implemented by the coordinator in Problem 2 as follows. At time t the coordinator selects partial functions (γ_t^1, γ_t^2) using the common (shared) information δ_t as follows.

$$\gamma_t^k(\cdot) = g_t^k(\cdot, \delta_t) =: \psi_t^k(\delta_t), \quad k = 1, 2. \quad (13)$$

Consider Problems 1 and 2. Use design g in Problem 1 and coordination strategy ψ given by (13) in Problem 2. Fix a specific realization of the initial state X_0 , the plant disturbance $\{V_t; t = 1, \dots, T\}$, and the observation noise $\{W_t^1, W_t^2; t = 1, \dots, T\}$. Then, the choice of ψ according to (13) implies that the realization of the state $\{X_t; t = 1, \dots, T\}$, the observations $\{Y_t^1, Y_t^2; t = 1, \dots, T\}$, and the control actions $\{U_t^1, U_t^2; t = 1, \dots, T\}$ are identical in Problem 1 and 2. Thus, any design g for Problem 1 can be implemented by the coordinator in Problem 2 by using a coordination strategy given by (13) and the total expected cost under g in Problem 1 is same as the total expected cost under the coordination strategy given by (13) in Problem 2.

By a similar argument, any coordination strategy ψ for Problem 2 can be implemented by the control stations in Problem 1 as follows. At time 1, both stations know δ_1 ; so, all of them can compute $\gamma_1^1 = \psi_1^1(\delta_1)$, $\gamma_1^2 = \psi_1^2(\delta_1)$. Then station k chooses action $u_1^k = \gamma_1^k(\lambda_1^k)$. Thus,

$$g_1^k(\lambda_1^k, \delta_1) = \psi_1^k(\delta_1)(\lambda_1^k), \quad k = 1, 2. \quad (14a)$$

At time 2, both stations know δ_2 and γ_1^1, γ_1^2 , so both of them can compute $\gamma_2^k = \psi_2^k(\delta_2, \gamma_1^1, \gamma_1^2)$, $k = 1, 2$. Then station k chooses action $u_2^k = \gamma_2^k(\lambda_2^k)$. Thus,

$$g_2^k(\lambda_2^k, \delta_2) = \psi_2^k(\delta_2, \gamma_1^1, \gamma_1^2)(\lambda_2^k), \quad k = 1, 2. \quad (14b)$$

Proceeding this way, at time t both stations know δ_t and $\gamma_{1:t-1}^1$ and $\gamma_{1:t-1}^2$, so both of them can compute $(\gamma_{1:t}^1, \gamma_{1:t}^2) = \psi_t(\delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2)$. Then, station k chooses action $u_t^k = \gamma_t^k(\lambda_t^k)$. Thus,

$$g_t^k(\lambda_t^k, \delta_t) = \psi_t^k(\delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2)(\lambda_t^k), \quad k = 1, 2. \quad (14c)$$

Now consider Problems 2 and 1. Use coordinator strategy ψ in Problem 2 and design \mathbf{g} given by (14) in Problem 1. Fix a specific realization of the initial state X_0 , the plant disturbance $\{V_t; t = 1, \dots, T\}$, and the observation noise $\{W_t^1, W_t^2; t = 1, \dots, T\}$. Then, the choice of \mathbf{g} according to (14) implies that the realization of the state $\{X_t; t = 1, \dots, T\}$, the observations $\{Y_t^1, Y_t^2; t = 1, \dots, T\}$, and the control actions $\{U_t^1, U_t^2; t = 1, \dots, T\}$ are identical in Problem 2 and 1. Hence, any coordination strategy ψ for Problem 2 can be implemented by the stations in Problem 1 by using a design given by (14) and the total expected cost under ψ in Problem 2 is same as the total expected cost under the design given by (14) in Problem 1.

Since Problems 1 and 2 are equivalent, we derive structural results for the latter problem. Unlike, Problem 1, where we have multiple control stations, the coordinator is the only decision maker in Problem 2.

Stage 3

We now look at Problem 2 as a controlled input-output system from the point of view of the coordinator and identify a state sufficient for input-output mapping. From the coordinator's viewpoint, the input at time t has two components: a stochastic input that consists of the plant disturbance V_t and observation noises W_t^1, W_t^2 ; and a controlled input that consists of the partial functions γ_t^1, γ_t^2 . The output is the observations Z_{t+1} given by (9). The cost is given by $c_t(X_t, U_t^1, U_t^2)$. We want to identify a state sufficient for input-output mapping for this system.

A variable is a state sufficient for input output mapping of a control system if it satisfies the following properties (see [6]).

- P1) The next state is a function of the current state and the current inputs.
- P2) The current output is function of the current state and the current inputs.
- P3) The instantaneous cost is a function of the current state, the current control inputs, and the next state.

We claim that such a state for Problem 2 is the following.

Definition 1: For each t define

$$S_t := (X_{t-1}, \Lambda_t^1, \Lambda_t^2) \quad (15)$$

Next we show that $S_t, t = 1, 2, \dots, T+1$, satisfy properties (P1)–(P3). Specifically, we have the following.

Proposition 1:

- 1) There exist functions $\hat{f}_t, t = 2, \dots, T$ such that

$$S_{t+1} = \hat{f}_{t+1}(S_t, V_t, W_{t+1}^1, W_{t+1}^2, \gamma_t^1, \gamma_t^2). \quad (16)$$

- 2) There exist functions $\hat{h}_t, t = 2, \dots, T$ such that

$$Z_t = \hat{h}_t(S_{t-1}). \quad (17)$$

- 3) There exist functions $\hat{c}_t, t = 1, \dots, T$ such that

$$c_t(X_t, U_t^1, U_t^2) = \hat{c}_t(S_t, \gamma_t^1, \gamma_t^2, S_{t+1}). \quad (18)$$

Proof: The proposition is an immediate consequence of the definitions of S_t, Z_t, Λ_t^k , the dynamics and observation equations of the system given by (1) and (2) and the evaluations carried out by the control stations according to (8). ■

Stage 4

Proposition 1 establishes S_t as the state sufficient for input-output mapping for the coordinator's problem. We now define information states for the coordinator.

Definition 2 (Information States): For a coordination strategy ψ , define information states Π_t as

$$\Pi_t(s_t) := \mathbb{P}^\psi(S_t = s_t \mid \Delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2). \quad (19)$$

As shown in Proposition 1, the state evolution of S_t depends on the controlled inputs (γ_t^1, γ_t^2) and the random noise $(V_t, W_{t+1}^1, W_{t+1}^2)$. This random noise is independent across time. Consequently, Π_t evolves in a controlled Markovian manner as below.

Proposition 2: For $t = 1, \dots, T-1$, there exists functions F_t (which do not depend on the coordinator's strategy) such that

$$\Pi_{t+1} = F_{t+1}(\Pi_t, \gamma_t^1, \gamma_t^2, Z_{t+1}). \quad (20)$$

At $t = 1$, since there is no shared information, Π_1 is simply the unconditional probability $\mathbb{P}(S_1) = \mathbb{P}(X_0, Y_1^1, Y_1^2)$. Thus, Π_1 is fixed a priori from the joint distribution of the primitive random variables and does not depend on the choice of coordinator's strategy ψ . Proposition 2 shows that at $t = 2, \dots, T$, Π_t depends on the strategy ψ only through the choices of $\gamma_{1:t-1}^1$ and $\gamma_{1:t-1}^2$. Moreover, as shown in Proposition 1, the instantaneous cost at time t can be written in terms of the current and next states (S_t, S_{t+1}) and the control inputs (γ_t^1, γ_t^2) . Combining the above two properties, we get the following:

Proposition 3: The process $\Pi_t, t = 1, 2, \dots, T$ is a controlled Markov chain with γ_t^1, γ_t^2 as the control actions at time t , i.e.,

$$\mathbb{P}(\Pi_{t+1} \mid \Delta_t, \Pi_{1:t}, \gamma_{1:t}^1, \gamma_{1:t}^2) = \mathbb{P}(\Pi_{t+1} \mid \Pi_t, \gamma_t^1, \gamma_t^2). \quad (21)$$

Furthermore, there exists a deterministic function C_t such that

$$\mathbb{E}\{\hat{c}_t(S_t, \gamma_t^1, \gamma_t^2, S_{t+1}) \mid \Delta_t, \Pi_{1:t}, \gamma_{1:t}^1, \gamma_{1:t}^2\} = C_t(\Pi_t, \gamma_t^1, \gamma_t^2). \quad (22)$$

The controlled Markov property of the process $\{\Pi_t, t = 1, \dots, T\}$ immediately gives rise to the following structural result.

Theorem 1: In Problem 2, without loss of optimality we can restrict attention to coordination strategies of the form

$$(\gamma_t^1, \gamma_t^2) = \psi_t(\Pi_t), \quad t = 1, \dots, T. \quad (23)$$

Proof: From Proposition 3, we conclude that the optimization problem for the coordinator is to control the evolution

of the controlled Markov process $\{\Pi_t, t = 1, 2, \dots, T\}$ by selecting the partial functions $\{\gamma_t^1, \gamma_t^2, t = 1, 2, \dots, T\}$ in order to minimize $\sum_{t=1}^T \mathbb{E} \{C_t(\Pi_t, \gamma_t^1, \gamma_t^2)\}$. This is an instance of the well-known Markov decision problems where it is known that the optimal strategy is a function of the current state. Thus, the structural result follows from Markov decision theory [1]. ■

The above result can also be stated in terms of the original problem.

Theorem 2 (Structural Result): In Problem 1 with $K = 2$, without loss of optimality we can restrict attention to coordination strategies of the form

$$U_t^k = g_t^k(\Lambda_t^k, \Pi_t), \quad k = 1, 2. \quad (24)$$

where

$$\Pi_t = \mathbb{P}^{(g_{1:t-1}^1, g_{1:t-1}^2)}(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t) \quad (25)$$

where $\Pi_1 = \mathbb{P}(X_0, Y_1^1, Y_1^2)$ and for $t = 2, \dots, T$, Π_t is evaluated as follows:

$$\Pi_{t+1} = F_{t+1}(\Pi_t, g_t^1(\cdot, \Pi_t), g_t^2(\cdot, \Pi_t), Z_{t+1}) \quad (26)$$

Proof: Theorem 1 established the structure of the optimal coordination strategy. As we argued in Stage 2, this optimal coordination strategy can be implemented in Problem 1 and is optimal for the objective (4). At $t = 1$, $\Pi_1 = \mathbb{P}(X_0, Y_1^1, Y_1^2)$ is known to both controllers and they can use the optimal coordination strategy to select partial functions according to:

$$(\gamma_1^1, \gamma_1^2) = \psi_1(\Pi_1)$$

Thus,

$$U_1^k = \gamma_1^k(\Lambda_1^k) = \psi_1^k(\Pi_1)(\Lambda_1^k) =: g_1^k(\Lambda_1^k, \Pi_1), \quad k = 1, 2. \quad (27)$$

At time instant $t + 1$, both controllers know Π_t and the common observations $Z_{t+1} = (Y_{t-n+1}^1, Y_{t-n+1}^2, U_{t-n+1}^1, U_{t-n+1}^2)$; they use the partial functions $(g_t^1(\cdot, \Pi_t), g_t^2(\cdot, \Pi_t))$ in equation (20) to evaluate Π_{t+1} . The control actions at time $t + 1$ are given as:

$$\begin{aligned} U_{t+1}^k &= \gamma_{t+1}^k(\Lambda_{t+1}^k) = \psi_{t+1}(\Pi_{t+1})(\Lambda_{t+1}^k) \\ &=: g_{t+1}^k(\Lambda_{t+1}^k, \Pi_{t+1}), \quad k = 1, 2. \end{aligned} \quad (28)$$

Moreover, using the design \mathbf{g} defined according to (28), the coordinator's information state Π_t can also be written as:

$$\begin{aligned} \Pi_t &= \mathbb{P}^\psi(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2) \\ &= \mathbb{P}^{\mathbf{g}}(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t, g_{1:2}^1(\cdot, \Pi_1), \dots, g_{t-1}^{1:2}(\cdot, \Pi_{t-1})) \\ &= \mathbb{P}^{(g_{1:t-1}^1, g_{1:t-1}^2)}(X_{t-1}, \Lambda_t^1, \Lambda_t^2 \mid \Delta_t) \end{aligned} \quad (29)$$

where we dropped the partial functions from the conditioning terms in (29) because under the given control laws $(g_{1:t-1}^1, g_{1:t-1}^2)$, the partial functions used from time 1 to $t - 1$ can be evaluated from Δ_t (by using Proposition 2 to evaluate $\Pi_{1:t-1}$). ■

Theorem 2 establishes the first structural result stated in Section I-D for $K = 2$. In the next section, we show how to extend the result for general K .

B. Extension to General K

Theorem 2 for two controllers ($K = 2$) can be easily extended to general K by following the same sequence of arguments as in stages 1 to 4 above. Thus, at time t , the coordinator introduced in Stage 1 now selects partial functions $\gamma_t^k : \mathcal{L}^k \mapsto \mathcal{U}^k$, for $k = 1, 2, \dots, K$. The state sufficient for input output mapping from the coordinator's perspective is given as $S_t := (X_{t-1}, \Lambda_t^{1:K})$ and the information state Π_t for the coordinator is

$$\Pi_t(s_t) := \mathbb{P}^\psi(S_t = s_t \mid \Delta_t, \gamma_{1:t-1}^{1:K}). \quad (30)$$

Results analogous to Propositions 1–3 can now be used to conclude the structural result of Theorem 2 for general K .

C. One-Step Delay

In this section, we focus on the one-step delay information structure, that is, we take $n = 1$. For a two-controller system with $n = 1$, we have, $\Delta_t = (Y_{1:t-1}^1, Y_{1:t-1}^2, U_{1:t-1}^1, U_{1:t-1}^2)$, $\Lambda_t^1 = (Y_t^1)$ and $\Lambda_t^2 = (Y_t^2)$.

The partial functions γ_t^k , $k = 1, 2$ to be selected by the coordinator are therefore mappings from \mathcal{Y}^k to \mathcal{U}^k , $k = 1, 2$. Also, the new shared observation at time $t + 1$ is $Z_{t+1} = (Y_t^1, Y_t^2, U_t^1, U_t^2)$. Following the arguments in Section II, the state sufficient for input-output mapping from the coordinator's perspective is

$$S_t = (X_{t-1}, Y_t^1, Y_t^2),$$

and the coordinator's information state is

$$\begin{aligned} \Pi_t(x_{t-1}, y_t^1, y_t^2) \\ := \Pr^\psi(X_{t-1} = x_{t-1}, Y_t^1 = y_t^1, Y_t^2 = y_t^2 \mid \\ \Delta_t, \gamma_{1:t-1}^1, \gamma_{1:t-1}^2). \end{aligned} \quad (31)$$

At $t = 1$, Π_1 is simply the unconditional probability $\mathbb{P}(S_1) = \mathbb{P}(X_0, Y_1^1, Y_1^2)$. Thus, Π_1 is fixed a priori from the joint distribution of the primitive random variables and does not depend on the choice of coordinator's strategy ψ . Moreover, the update equation for the information states in Proposition 2 can be refined for $n = 1$ as follows:

Proposition 4: For $t = 1, \dots, T - 1$, there exists functions G_t (which do not depend on the coordinator's strategy) such that

$$\Pi_{t+1} = G_{t+1}(\Pi_t, Z_{t+1}). \quad (32)$$

Proposition 4 and the fact Π_1 is fixed a priori imply that, for all time instants t , the realization of Π_t does not depend on the choice of partial functions but only on the realization of the shared observations, Z_1, Z_2, \dots, Z_t , which is Δ_t . Thus, we can write Π_t as:

$$\Pi_t(x_{t-1}, y_t^1, y_t^2) := \mathbb{P}(X_{t-1} = x_{t-1}, Y_t^1 = y_t^1, Y_t^2 = y_t^2 \mid \Delta_t) \quad (33)$$

The result of Theorem 2 can now be restated for this case as follows:

Corollary 1: In Problem 1 with $K = 2$ and $n = 1$, there is no loss of optimality in restricting to control laws of the form:

$$U_t^k = g_t^k(Y_t^k, \Pi_t), \quad k = 1, 2. \quad (34)$$

where

$$\Pi_t := \mathbb{P}(X_{t-1}, Y_t^1, Y_t^2 \mid \Delta_t) \quad (35)$$

We can now compare our result for one-step delay with the structural result proved in [3]. The result in [3] states that there is no loss of optimality in using control laws of the form:

$$U_t^k = g_t^k(Y_t^k, \mathbb{P}(X_{t-1} \mid \Delta_t)), \quad k = 1, 2. \quad (36)$$

One can recover the above structural result from (34) by observing that there is a one-to-one correspondence between Π_t and the belief $\mathbb{P}(X_{t-1} \mid \Delta_t)$. Clearly, the belief $\mathbb{P}(X_{t-1} \mid \Delta_t)$ is a marginal of Π_t and therefore can be evaluated from Π_t . Moreover, given the belief $\mathbb{P}(X_{t-1} \mid \Delta_t)$, one can evaluate Π_t as:

$$\begin{aligned} \Pi_t(x_{t-1}, y_t^1, y_t^2 \mid \Delta_t) &= \mathbb{P}(Y_t^1 = y_t^1 \mid X_{t-1} = x_{t-1}) \\ &\quad \times \mathbb{P}(Y_t^2 = y_t^2 \mid X_{t-1} = x_{t-1}) \\ &\quad \times \mathbb{P}(X_{t-1} = x_{t-1} \mid \Delta_t). \end{aligned} \quad (37)$$

The fact that there is a one-to-one correspondence between Π_t and $\mathbb{P}(X_{t-1} \mid \Delta_t)$ means that the separation result proposed in this paper for $n = 1$ is effectively equivalent to the one proved in [3].

III. SEQUENTIAL DECOMPOSITION

In addition to obtaining the structural result of Theorem 2, the coordinator's problem also allows us to write a dynamic program for finding the optimal control strategies as shown below. We focus on the two controller case ($K = 2$). The result can be easily extended to general K .

Theorem 3: The optimal coordination strategy can be found by the following dynamic program: For $t = 1, \dots, T$, define the functions $J_t : \mathcal{P}\{S\} \mapsto \mathbb{R}$ as follows. For $\pi \in \mathcal{P}\{S\}$ let

$$J_T(\pi) = \inf_{\tilde{\gamma}^1, \tilde{\gamma}^2} \mathbb{E} \left\{ C_T(\Pi_T, \gamma_T^1, \gamma_T^2) \mid \Pi_T = \pi, \begin{array}{l} \gamma_T^1 = \tilde{\gamma}^1, \\ \gamma_T^2 = \tilde{\gamma}^2 \end{array} \right\}. \quad (38)$$

For $t = 1, \dots, T - 1$, and $\pi \in \mathcal{P}\{S\}$ let

$$J_t(\pi) = \inf_{\tilde{\gamma}^1, \tilde{\gamma}^2} \mathbb{E} \left\{ C_t(\Pi_t, \gamma_t^1, \gamma_t^2) + J_{t+1}(\Pi_{t+1}) \mid \begin{array}{l} \Pi_t = \pi, \\ \gamma_t^1 = \tilde{\gamma}^1, \\ \gamma_t^2 = \tilde{\gamma}^2 \end{array} \right\}. \quad (39)$$

The $\arg \inf(\gamma_t^{*,1}, \gamma_t^{*,2})$ in the RHS of $J_t(\pi)$ is the optimal action for the coordinator at time t then $\Pi_t = \pi$. Thus,

$$(\gamma_t^{*,1}, \gamma_t^{*,2}) = \phi_t^*(\pi_t)$$

The corresponding control strategy for Problem 1, given by (14) is optimal for Problem 1.

Proof: As in Theorem 1, we use the fact that the coordinator's optimization problem can be viewed as a Markov decision problem with Π_t as the state of the Markov process. The dynamic program follows from standard results in Markov decision theory [1]. The optimality of the corresponding control strategy for Problem 1 follows from the equivalence between the two problems. ■

The dynamic program of Theorem 3 can be extended to general K in a manner similar to Section II-B.

IV. CONCLUSION

We studied the stochastic control problem with n -step delay sharing information structure and established a structural result for it. We also provided a sequential methodology to identify the optimal control strategies. To derive our results, we formulated an alternative problem from the point of a coordinator of the system. We believe that this idea of formulating an alternative problem from the point of view of a coordinator which has access to information common to all controllers is also useful for general decentralized control problems, as is illustrated by [7] and [8].

Our structural results differ from those asserted by Witsenhausen in a fundamental way. The sufficient statistic (also called information state) $\mathbb{P}(X_{t-n} \mid \Delta_t)$ of Witsenhausen's assertion does not depend on the control strategy. The sufficient statistic $\mathbb{P}^{g_{1:t-1}^{1:K}}(X_{t-1}, \Lambda_t^{1:K} \mid \Delta_t)$ of our first result *depend on the control laws used before time t* . An alternative structural result that highlights the dependence of the coordinator's sufficient statistic at time t and the controls laws used before time t is presented in [5].

ACKNOWLEDGMENTS

This research was supported in part by NSF Grant CCR-0325571 and NASA Grant NNX06AD47G.

REFERENCES

- [1] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Prentice Hall, 1986.
- [2] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.
- [3] P. Varaiya and J. Walrand, "On delayed sharing patterns," *IEEE Trans. Autom. Control*, vol. 23, no. 3, pp. 443–445, 1978.
- [4] B. Kurtaran, "Corrections and extensions to "decentralized stochastic control with delayed sharing information pattern"," *IEEE Trans. Autom. Control*, vol. 24, no. 4, pp. 656–657, Aug. 1979.
- [5] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies for delayed sharing information structures," *submitted to the IEEE Trans. Autom. Control*, Feb. 2010.
- [6] H. S. Witsenhausen, "Some remarks on the concept of state," in *Directions in Large-Scale Systems*, Y. C. Ho and S. K. Mitter, Eds. Plenum, 1976, pp. 69–75.
- [7] A. Nayyar and D. Teneketzis, "On the structure of real-time encoders and decoders in a multi-terminal communication system," *IEEE Trans. Info. Theory*, submitted.
- [8] A. Mahajan, A. Nayyar, and D. Teneketzis, "Identifying tractable decentralized control problems on the basis of information structures," in *proceedings of the 46th Allerton conference on communication, control and computation*, Sep. 2008, pp. 1440–1449.