Common Knowledge and Sequential Team Problems

Authors: Ashutosh Nayyar and Demosthenis Teneketzis

Computer Engineering Technical Report Number CENG-2018-02

Ming Hsieh Department of Electrical Engineering – Systems
University of Southern California
Los Angeles, California 90089-2563

November, 2018

Common Knowledge and Sequential Team Problems

Ashutosh Nayyar and Demosthenis Teneketzis

Abstract

We consider a general sequential team problem based on Witsenhausen's intrinsic model. Our formulation encompasses all teams in which the uncontrolled inputs can be viewed as random variables on a finite probability space, the number of control inputs/decisions is finite and the decisions take values in finite spaces. We define the concept of common knowledge in such teams and use it to construct a sequential decomposition of the problem of optimizing the team strategy profile. If the information structure is classical, our common knowledge based decomposition is identical to classical dynamic program. If the information structure is such that the common knowledge is trivial, our decomposition is similar in spirit to Witsenhausen's standard form based decomposition [16]. In this case, the sequential decomposition is essentially a sequential reformulation of the strategy optimization problem and appears to have limited value. For information structures with non-trivial common knowledge, our sequential decomposition differs from Witsenhausen's standard form based decomposition because of its dependence on common knowledge. Our common knowledge based approach generalizes the common information based methods of [12]–[14].

I. Introduction

This paper deals with the problem of decentralized decision-making. Such problems arise in any system where multiple agents/decision-makers have to take actions/make decisions based on their respective information. Examples of such systems include communication and power networks, sensing and surveillance systems, networked control systems and teams of autonomous robots. We focus on problems that are: (i) *Cooperative,* i.e., problems where different decision-makers share the same objective. Such problems are called *Team Problems* [3], [6], [7], [10], [11], [15], [20], [21]; (ii) *Stochastic,* i.e., problems where stochastic models of uncertainties are available and the goal is to minimize the expected value of the system cost; (iii) *Sequential,* i.e., problems where the decision-makers act in a pre-determined order that is independent of the realizations of the uncertain inputs or the choice of the decision strategy profile. Further, this order satisfies a basic causality condition: the information available to make a decision does not depend on decisions to be made in the future. Decentralized decision-making problems with the above characteristics are referred to as *sequential team* problems.

Sequential team problems can be categorized based on their information structures. *Classical* information structures have the perfect recall property, that is, the information available to make a decision includes all the information available to make all past decisions. The classical dynamic program based on Markov decision theory provides a systematic way of solving sequential team problems with classical information structure [8], [18]. This method allows us to decompose the problem of finding optimal strategies for all agents into several smaller problems which must be solved sequentially backwards in time to obtain optimal strategies. We refer to this simplification as a *sequential decomposition* of the problem.

When the information structure is not classical, a general sequential decomposition is provided by Witsenhausen's standard form based method [16]. The idea of the standard form approach is to consider the optimization problem of a designer who has to select a sequence of decision strategies, one for each agent. The designer knows the system model (including the system cost function) and the probability distributions of uncertain inputs but does not have any other information. The designer sequentially selects a decision strategy for each agent. The designer's problem can be shown to be a problem with (trivially) classical information structure. This approach can be used to decompose the designer's problem of choosing a sequence of decision strategies into several sub-problems that must be solved sequentially backwards in time. In each of these sub-problems, the designer has to optimize over one decision strategy (instead of the whole strategy profile). This approach for obtaining a sequential decomposition of sequential team problems has been described in detail in [16] and [9].

In this paper, we provide a new sequential decomposition for sequential team problems. Our approach relies on the idea of common knowledge in sequential team problems. In response to the sequential nature of the team problems we study, our definition of common knowledge is itself sequential, that is, it changes for each decision to be made. At any given time, common knowledge represents the information about uncertain inputs and agents' decisions that is available to all current and future decision-makers. We show that decision-makers can use this common knowledge to coordinate how they make decisions. Our methodology provides a sequential decomposition for any sequential team problem with finitely many decision-makers and with finite probability and decision spaces. We can make three observations about our common knowledge based decomposition: (i) If the underlying information structure is classical, our sequential decomposition reduces to the classical dynamic program. (ii) For information structures with *non-trivial common knowledge*, our sequential decomposition differs from Witsenhausen's standard from based decomposition because of its dependence on common knowledge. The use of common knowledge allows our sequential decomposition to have simpler sub-problems than those in Witsenhausen's standard form approach. (iii) For information structures with *trivial common knowledge* (see Section V), our decomposition is similar in spirit to Witsenhausen's standard form based decomposition [16]. In this case, the sequential decomposition is essentially a sequential reformulation of the strategy optimization problem and appears to have limited value.

The common knowledge approach described in this paper generalizes the common information method of [12]. The common information method has been used in [13] and [14] for studying delayed history sharing and partial history sharing models in decentralized control. In contrast to the common information based methods of [13], [14], the common knowledge approach of this paper does not require a part of agents' information to be nested over time. Further, in some cases, it can produce a sequential decomposition that is distinct from, and simpler than, the common information based decomposition.

We will adopt Witsenhausen's intrinsic model [1], [17], [18] to present our results for general sequential team problems. Models similar to the intrinsic model have been presented in [19]. The intrinsic model encompasses all systems in which (1) the uncontrolled inputs can be viewed as random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$; (2) the number of decisions to be taken is finite (T), (3) the t-th decision can be viewed as an element of a measurable space $(\mathbb{U}_t, \mathcal{U}_t)$ in which all singletons are measurable; and (4) the decision strategy for the t-th decision can be viewed a measurable function from the measurable space $(\Omega \times \mathbb{U}_1 \times ... \times \mathbb{U}_T, \mathcal{J}_t)$ to the measurable space $(\mathbb{U}_t, \mathcal{U}_t)$, where $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \ldots \otimes \mathcal{U}_T$ is a sigma-algebra that denotes the maximal information (knowledge) that can be used to select the t-th decision.

A. *Organization*

The paper is organized as follows. We describe the intrinsic model, information structures and the observations that generate a decision-maker's information sigma-algebra in Section II. We present the dynamic program for classical information structures in Section III. We define common knowledge for sequential team problems and use it to derive a sequential decomposition in Section IV. We compare common knowledge based sequential decomposition with the classical dynamic program and with Witsenhausen's standard form in Section V. We compare our common knowledge approach with the common information approach used in prior work in Section VI. We discuss the impact of sequential orders on common knowledge in Section VII. We conclude in Section VIII.

B. Notation

We denote random variables by capital letters and their realizations by corresponding small letters. Some random variables are denoted by small Greek letters (e.g., γ, ω) and we use ~ or ^ to denote a particular realization (as in $\tilde{\gamma}$, $\hat{\omega}$). For any variable $*$, we use $*_{1:t}$ as a shorthand for $(*_1, *_2, \ldots, *_t)$. For sets $A_1, \ldots, A_t$, $A_{1:t}$ denotes the product set $A_1 \times \ldots \times A_t$. $\mathbb{R}$ is the set of real numbers and $\mathbb{B}(\mathbb{R})$ is the Borel sigma-algebra on $\mathbb{R}$. If $A_1, \ldots, A_k$ form a partition of a set Ω, then $\sigma(A_1, \ldots, A_k)$ denotes the sigma-algebra generated by this partition.

II. The Intrinsic Model

Consider a stochastic system with finitely many decisions/control inputs. The decisions are denoted by U_t, $t = 1, 2, \ldots, T$, and take values in measurable spaces $(\mathbb{U}_t, \mathcal{U}_t)$, $t = 1, 2, \ldots, T$, respectively. All uncontrolled inputs to the stochastic system are modeled as a random vector $\omega = (\omega^1, \omega^2, \ldots, \omega^N)$ taking values in a measurable space $(\Omega, \mathcal{F})$. A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ specifies the probability distribution of the random vector ω . The components of ω are referred to as the *primitive random variables* of the system.

A. Decision Strategies

For $t = 1, 2, \ldots, T$, we define $U_{1:t}$ as the vector $(U_1, U_2, \ldots, U_t)$ and U_{-t} as $(U_1, \ldots, U_{t-1}, U_{t+1}, \ldots, U_T)$; we also define the product measurable space $(\mathbb{U}_{1:t}, \mathcal{U}_{1:t})$ as

$$\mathbb{U}_{1:t} := \mathbb{U}_1 \times \cdots \times \mathbb{U}_t, \qquad \mathcal{U}_{1:t} := \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_t. \tag{1}$$

It is convenient to think of each of the T decisions being chosen by a distinct decision-maker (DM)[1]. The information available to the $t-$th decision maker (DM t) may depend on the realization of ω and the decisions made by other decision-makers. In the intrinsic model [1], [17], [18], this information is represented by a sigma-algebra $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_{1:T}$. The decision U_t is chosen according to

$$U_t = g_t(\omega, U_{1:T}), \tag{2}$$

where g_t is a measurable function from the measurable space $(\Omega \times \mathbb{U}_{1:T}, \mathcal{J}_t)$ to the measurable space $(\mathbb{U}_t, \mathcal{U}_t)$ [1], [17], [18], that is,

$$g_t : (\Omega \times \mathbb{U}_{1:T}, \mathcal{J}_t) \mapsto (\mathbb{U}_t, \mathcal{U}_t). \tag{3}$$

The function g_t is called the *decision strategy* of the $t-$th decision maker and the collection of all T decision strategies $\mathbf{g} = (g_1, g_2, \ldots, g_T)$ is called the *decision strategy profile*.

[1]The fact that some of the decision-makers (DMs) may be the same physical entity is of no relevance for our purposes.

B. Cost Function and Optimization Problem

The performance of the stochastic system is measured by a cost function $c : (\Omega \times \mathbb{U}_{1:T}, \mathcal{F} \otimes \mathcal{U}_{1:T}) \mapsto (\mathbb{R}, \mathbb{B}(\mathbb{R}))$. We can now formulate the following optimization problem.

Problem 1. *Given the probability model $(\Omega, \mathcal{F}, \mathbb{P})$ for the random vector ω, the measurable decision spaces $(\mathbb{U}_t, \mathcal{U}_t)$, $t = 1, \ldots, T$, the sigma-algebras $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_{1:T}$ and a cost function $c : (\Omega \times \mathbb{U}_{1:T}, \mathcal{F} \otimes \mathcal{U}_{1:T}) \mapsto (\mathbb{R}, \mathbb{B}(\mathbb{R}))$, find a decision strategy profile $\mathbf{g} = (g_1, \ldots, g_T)$, with $g_t : \Omega \times \mathbb{U}_{1:T} \mapsto \mathbb{U}_t$ being $\mathcal{J}_t/\mathcal{U}_t$ measurable for each t, that achieves*

$$\inf_{\mathbf{g}} \mathbb{E}[c(\omega, U_1, \ldots, U_T)] \quad \textit{exactly or within } \epsilon > 0,$$

where $U_t = g_t(\omega, U_{1:T})$ for each t.

Remark 1. *A choice of strategy profile for the stochastic system creates a system of closed loop equations:*

$$u_t = g_t(\tilde{\omega}, u_{1:T}), \quad t=1,\ldots,T, \tag{4}$$

for each realization $\tilde{\omega}$ of the random vector ω. In general, there may exist $\tilde{\omega} \in \Omega$ for which this system of equations does not have a unique solution. In that case, the optimization problem is not well-posed. However, when properties C [17] or CI [1] hold, the above system of equations has a unique solution. Properties C and CI trivially hold for the sequential information structures we investigate in this paper.

C. Information Structures

The sigma algebras $\mathcal{J}_1, \ldots, \mathcal{J}_T$ together specify the information available for making each of the T decisions and are referred to as the information structure of the problem. Information structures are classified according to the relationships among the sigma algebras $\mathcal{J}_1, \ldots, \mathcal{J}_T$ and $\mathcal{F} \otimes \mathcal{U}_1 \otimes \ldots \otimes \mathcal{U}_T$.

Sequential and Non-sequential Information Structures: We say that the information structure is sequential if there exists a permutation $p : \{1, 2, \ldots, T\} \mapsto \{1, 2, \ldots, T\}$ such that for $t = 1, \ldots, T$,

$$\begin{aligned} \mathcal{J}_{p(t)} \subset \mathcal{F} \otimes \mathcal{U}_{p(1)} \otimes \mathcal{U}_{p(2)} \otimes \cdots \otimes \mathcal{U}_{p(t-1)} \otimes \\ \{\emptyset, \mathbb{U}_{p(t)}\} \otimes \ldots \otimes \{\emptyset, \mathbb{U}_{p(T)}\}. \end{aligned} \tag{5}$$

Otherwise, the information structure is said to be non-sequential.

The sequence $p(1), .., p(T)$ can be interpreted as time and (5) as a causality condition. Note that for a sequential system there may be more than one permutation satisfying the causality condition (5). In the following sections, without loss of generality, we will let p be the identity map, that is, $p(t) = t$.

Sequential information structures are further classified as:

1) Static: If $\mathcal{J}_t \subset \mathcal{F} \otimes \{\emptyset, \mathbb{U}_{1:T}\}$ for all t.
2) Classical: If $\mathcal{J}_t \subset \mathcal{J}_{t+1}$ for $t = 1, \dots, T-1$.
3) Quasi-classical (partially nested): Recall that for sequential information structures

$$\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \dots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \dots \{\emptyset, \mathbb{U}_T\}. \tag{6}$$

For $s < t$, we say that the decision U_s *does not affect the information of the $t-$th decision maker* if

$$\begin{aligned}\mathcal{J}_t \subset \mathcal{F} \otimes & \mathcal{U}_1 \otimes \dots \otimes \mathcal{U}_{s-1} \otimes \{\emptyset, \mathbb{U}_s\} \otimes \\ & \mathcal{U}_{s+1} \otimes \dots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \dots \{\emptyset, \mathbb{U}_T\}.\end{aligned} \tag{7}$$

If (7) is not true, we say that the decision U_s *affects the information of the $t-$th decision maker*. An information structure is Quasi-classical (partially nested) if $\mathcal{J}_s \subset \mathcal{J}_t$ for every s, t (with $s < t$) such that U_s affects the information of the $t-$th decision maker.
4) Non-classical: An information structure that does not belong to the above three categories is called *non-classical*.

D. Finite Spaces Assumption

In the rest of the paper, we will assume that the random vector ω takes values in a finite set and that the decision spaces are finite.

Assumption 1. *Ω and $\mathbb{U}_t$, $t = 1, \dots, T$, are finite sets. Further, $\mathcal{F} = 2^{\Omega}$ and $\mathcal{U}_t = 2^{\mathbb{U}_t}$, for $t = 1, \dots, T$.*

E. Information Sigma Algebra and Generating Observations

Consider a sigma-algebra $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}^{1:T}$ representing the information available to a decision-maker. Consider a collection of variables $Z_a, Z_b, \dots, Z_k$ defined as functions from $\Omega \times \mathbb{U}^{1:T}$ to spaces $\mathbb{Z}_a, \dots, \mathbb{Z}_k$ respectively,

$$\begin{aligned} & Z_i = \zeta_i(\omega, U_{1:T}), \\ \text{where } & \zeta_i : \Omega \times \mathbb{U}_{1:T} \mapsto \mathbb{Z}_i, \quad i = a, b, \dots, k. \end{aligned} \tag{8}$$

We will call $Z_i = \zeta_i(\omega, U_{1:T})$ an observation and ζ_i as its observation map. For a realization $\tilde{\omega}$ and $u_{1:T}$, $z_i = \zeta_i(\tilde{\omega}, u_{1:T})$ is the corresponding realization of the observation Z_i. We will denote by $\sigma(Z_a, \dots, Z_k)$ the smallest sigma algebra contained in $\mathcal{F} \otimes \mathcal{U}_{1:T}$ with respect to which the observation maps $\zeta_a, \dots, \zeta_k$ are measurable.

We say that observations $Z_a, \dots, Z_k$ generate the sigma-algebra $\mathcal{J}_t$ if

$$\sigma(Z_a, \dots, Z_k) = \mathcal{J}_t.$$

III. Dynamic program for classical information structures

Consider a sequential stochastic system with classical information structure, that is, a system in which the information sigma algebras of the decision-makers are nested over time: $\mathcal{J}_t \subset \mathcal{J}_{t+1}$, $t = 1, \dots, T-1$.

Lemma 1. *There exist observations $Z_t = \zeta_t(\omega, U_1, \dots, U_{t-1})$, $t = 1, \dots, T$, with Z_t taking values in a finite set $\mathbb{Z}_t$, such that $\sigma(Z_{1:t}) = \mathcal{J}_t$. The variables $Z_{1:t}$ will be collectively referred to as the observations available to the t-th decision maker. Further, any $\mathcal{J}_t/\mathcal{U}_t$ measurable decision strategy can be written as*

$$U_t = g_t(Z_{1:t}).$$

Proof. The proof follows from standard properties of finite and nested sigma algebras. A formal argument is given in Appendix A. □

Note that the observations of Lemma 1 depend on past decisions and not on future decisions. We can now re-state Problem 1 for a classical information structure under Assumption 1 as follows.

Problem 2. *Given observations $Z_t = \zeta_t(\omega, U_1, \dots, U_{t-1})$ taking values in $\mathbb{Z}_t$ for $t = 1, \dots, T$, find a decision strategy profile $\mathbf{g} = (g_1, \dots, g_T)$, where g_t maps $\mathbb{Z}_{1:t}$ to $\mathbb{U}_t$ for each t, that achieves*

$$\inf_{\mathbf{g}} \mathbb{E}[c(\omega, U_1, \dots, U_T)] \quad \textit{exactly or within } \epsilon > 0,$$

where $U_t = g_t(Z_{1:t})$ for each t.

A. Classical Information Structure with Observed Decisions

Corresponding to the classical information structure of Problem 2 , we consider an expanded information structure where, for all t, the $t-$th decision maker observes $Z_{1:t}$ as well as the past decisions $U_{1:t-1}$. Thus, the decision strategy at time t in the expanded information structure is a function of $Z_{1:t}, U_{1:t-1}$. Any strategy profile in the original information structure remains a valid strategy profile in the expanded information structure. Therefore, any achievable expected cost in the original information structure is achievable in the expanded information structure. The following lemma states that the converse is true as well, that is, any achievable expected cost in the expanded information structure is achievable in the original information structure.

Lemma 2. *For any strategy profile in the expanded information structure, there exists a strategy profile in the original information structure with the same expected cost.*

Proof. See Appendix B. □

Lemma 2 implies that the optimal expected cost under the original and the expanded information structures are the same. Therefore, we can find optimal strategies in the expanded information structure and use them to construct optimal strategies in the original information structure. In the rest of this section, we will focus on the expanded information structure where past decisions are observed by each decision-maker.

Remark 2. *In case Z_t includes U_{t-1} for each $t > 1$, the information expansion described above is redundant.*

B. Strategy-independent beliefs on Ω

Consider a time instant t and a realization $z_{1:t}, u_{1:t-1}$ of $Z_{1:t}, U_{1:t-1}$. We say that the realization $z_{1:t}, u_{1:t-1}$ is *feasible* if there exists $\hat{\omega} \in \Omega$ with $\mathbb{P}(\hat{\omega}) > 0$ such that $\zeta_k(\hat{w}, u_{1:k-1}) = z_k$ for $k = 1, \dots, t$.

For a given feasible realization $z_{1:t}, u_{1:t-1}$, we define a probability distribution on Ω as follows:

$$\pi_t(\tilde{\omega}|z_{1:t}, u_{1:t-1}) := \frac{\mathbb{P}(\tilde{\omega}) \prod_{k=1}^{t} \mathbb{1}_{\{\zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k\}}}{\sum_{\hat{\omega}} \left[\mathbb{P}(\hat{\omega}) \prod_{k=1}^{t} \mathbb{1}_{\{\zeta_k(\hat{\omega}, u_{1:k-1}) = z_k\}}\right]}, \tag{9}$$

$\forall \tilde{\omega} \in \Omega$.

It is easy to verify that $\pi_t(\cdot|z_{1:t}, u_{1:t-1})$ is indeed a probability distribution on Ω, that is,

$$\pi_t(\tilde{\omega}|z_{1:t}, u_{1:t-1}) \geq 0 \text{ and } \sum_{\tilde{\omega} \in \Omega} \pi_t(\tilde{\omega}|z_{1:t}, u_{1:t-1}) = 1.$$

Remark 3. *π_t is DM t's posterior belief on random vector ω given $z_{1:t}$ under the open loop control sequence $u_{1:t}$.*

C. Strategy-induced beliefs on Ω

Consider any arbitrary strategies $g_t(Z_{1:t}, U_{1:t-1})$, $t = 1, \dots, T$, in the expanded information structure. The strategy profile $g_1, \dots, g_T$ creates a joint probability distribution on $\omega, U_{1:T}, Z_{1:T}$ which we denote by $\mathbb{P}^g(\tilde{w}, u_{1:T}, z_{1:T})$. Focusing on the first t decisions, $g_1, \ldots, g_t$ induce a joint probability distribution on $\omega, U_{1:t}, Z_{1:t+1}$ which we denote by $\mathbb{P}^{g_{1:t}}(\tilde{w}, u_{1:t}, z_{1:t+1})$. This distribution can be explicitly written as

$$\begin{aligned}
&\mathbb{P}^{g_{1:t}}(\tilde{\omega}, z_{1:t+1}, u_{1:t}) \\
&= \mathbb{P}(\tilde{\omega}) \prod_{k=1}^{t} \left[\mathbb{1}_{\{\zeta_k(\tilde{\omega}, u_{1:k-1})=z_k\}} \mathbb{1}_{\{g_k(z_{1:k}, u_{1:k-1})=u_k\}}\right] \times \\
&\mathbb{1}_{\{\zeta_{t+1}(\tilde{\omega}, u_{1:t})=z_{t+1}\}}.
\end{aligned} \tag{10}$$

From the joint distribution $\mathbb{P}^g(\tilde{w}, u_{1:T}, z_{1:T})$, we can construct conditional probabilities (e.g., $\mathbb{P}^g(\tilde{\omega}|z_{1:t}, u_{1:t-1})$) using Bayes' rule. The following lemma identifies a crucial property of such conditional probabilities.

Lemma 3 (Strategy-independence of belief)**.** *Consider any arbitrary strategies* $g_t(Z_{1:t}, U_{1:t-1}), t = 1, \ldots, T$, *in the expanded information structure. Then,*

$$\mathbb{P}^g(\tilde{\omega}|Z_{1:t}, U_{1:t-1}) = \pi_t(\tilde{\omega}|Z_{1:t}, U_{1:t-1}),\ \mathbb{P}^g-\text{ almost surely.}$$

That is, for any realization $z_{1:t}, u_{1:t-1}$ *such that* $\mathbb{P}^g(z_{1:t}, u_{1:t-1}) > 0$,

$$\mathbb{P}^g(\tilde{\omega}|z_{1:t}, u_{1:t-1}) = \pi_t(\tilde{\omega}|z_{1:t}, u_{1:t-1}),$$

where π_t *is defined in* (9).

Proof. The proof follows from a straightforward application of Bayes' rule and (10). □

Lemma 3 is a crucial lemma for classical information structures. Consider the $t-$th decision maker. Based on the realization of its information and the strategy profile in operation, it can form a conditional belief on the uncertainty in the system. Lemma 3 states that irrespective of the strategy profile being used, DM t will have the same belief for the same realization of its information.

D. Sequentially dominant strategies and the Dynamic Program

For strategies $g_1, g_2, \ldots, g_T$ of the T decision-makers, define $J(g_1, g_2, \ldots, g_T)$ as the expected cost under these strategies. That is,

$$J(g_1, g_2, \ldots, g_T) := \mathbb{E}^{g_1,\ldots,g_T}[c(\omega, U_1, \ldots, U_T)],$$

where $\mathbb{E}^{g_1,\ldots,g_T}$ denotes that the expectation is with respect to the joint probability distribution $\mathbb{P}^{g_1,\ldots,g_T}(\tilde{\omega}, z_{1:T}, u_{1:T})$.

We say that a strategy g_T^* is a sequentially dominant strategy[2] for DM T if for any strategies $g_1, g_2, \ldots, g_T$,

$$J(g_1, g_2, \ldots, g_T) \geq J(g_1, \ldots, g_{T-1}, g_T^*).$$

[2]For DM T, sequentially dominant strategy is simply a dominant strategy.

Given a sequentially dominant strategy g_T^* for DM T, we say that g_{T-1}^* is a *sequentially dominant strategy* for DM $T-1$, if for any strategies $g_1, g_2, \dots, g_{T-1}$,

$$J(g_1, g_2, \dots, g_{T-1}, g_T^*) \geq J(g_1, \dots, g_{T-2}, g_{T-1}^*, g_T^*).$$

We can now define sequentially dominant strategies recursively: Given sequentially dominant strategies $g_{k+1}^*, g_{k+2}^*, \dots, g_T^*$ for DM $k+1$ to DMT respectively, we say that g_k^* is a sequentially dominant strategy for DM k if for any strategies $g_1, g_2, \dots, g_k$,

$$J(g_1, g_2, \dots, g_k, g_{k+1}^*, \dots, g_T^*) \geq J(g_1, \dots, g_{k-1}, g_k^*, g_{k+1}^*, \dots, g_T^*).$$

Lemma 4. *Suppose $g_1^*, \dots, g_T^*$ are sequentially dominant strategies as defined above. Then, $(g_1^*, \dots, g_T^*)$ is an optimal strategy profile.*

Proof. The proof follows directly from the definition of sequentially dominant strategies. □

For a classical information structure with observed decisions, sequentially dominant strategies can be found by a dynamic program. In the following theorem, we use $\mathbb{E}^{\pi_k}$ to denote that the expectation is with respect to the belief π_k defined in Section III-B.

Theorem 1. *For a sequential team problem that has a classical information structure with observed decisions,*

1) *Define value functions $V_k(z_{1:k}, u_{1:k-1})$ recursively as follows:*

$$\begin{aligned} &V_T(z_{1:T}, u_{1:T-1}) := \\ &\min_{u_T \in \mathbb{U}_T} \mathbb{E}^{\pi_T}[c(\omega, u_1 \dots, u_T)|z_{1:T}, u_{1:T-1}], \\ &V_k(z_{1:k}, u_{1:k-1}) := \\ &\min_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \dots, u_k)|z_{1:k}, u_{1:k-1}], \end{aligned} \tag{11}$$

$k = T-1, \dots, 2, 1$, *where $Z_{k+1} = \zeta_{k+1}(\omega, u_{1:k})$ and the expectations are with respect to strategy-independent beliefs on ω (as defined in* (9)*) .*

2) *Further, define*

$$\begin{aligned} &g_T^*(z_{1:T}, u_{1:T-1}) := \\ &argmin_{u_T \in \mathbb{U}_T} \mathbb{E}^{\pi_T}[c(\omega, u_1 \dots, u_{T-1}, u_T)|z_{1:T}, u_{1:T-1}] \\ &g_k^*(z_{1:k}, u_{1:k-1}) := \\ &argmin_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \dots, u_k)|z_{1:k}, u_{1:k-1}], \end{aligned} \tag{12}$$

$k = T-1, \ldots, 2, 1$.

The strategies $g_T^*, \ldots, g_1^*$ are sequentially dominant strategies for DM T to DM 1 respectively. Consequently, $(g_1^*, \ldots, g_T^*)$ is an optimal strategy profile.

Proof. The theorem can be proved using a standard dynamic programming argument [8] with $S_k = (Z_{1:k}, U_{1:k-1})$ as the state at time k. We provide an alternative proof in Appendix E that highlights the role of strategy-independent beliefs in finding sequentially dominant strategies. □

E. Non-classical Information Structure and Absence of Sequentially Dominant strategies

Consider a simple sequential problem with two DMs. ω takes values in the measurable space $(\Omega = \{0,1\}, \mathcal{F} = 2^{\Omega})$ with equal probabilities. The decision spaces of the two DMs are $\mathbb{U}_1 = \{0,1\}$ and $\mathbb{U}_2 = \{0,1,2\}$ respectively. The decision spaces are associated with sigma algebras $\mathcal{U}_1 = 2^{\mathbb{U}_1}$ and $\mathcal{U}_2 = 2^{\mathbb{U}_2}$ respectively. The cost function is $c(\omega, U_1, U_2) = (\omega + U_1 - U_2)^2$. We consider two information structures:

1) Classical case: Consider the classical information structure:

$$\mathcal{J}_1 = 2^{\Omega} \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}, \tag{13}$$

$$\mathcal{J}_2 = 2^{\Omega} \times 2^{\mathbb{U}_1} \times \{\emptyset, \mathbb{U}_2\}. \tag{14}$$

In other words, DM 1's strategy can be any function of ω, while DM 2's can be any function of ω and U_1.

The expected cost $\mathbb{E}[(\omega + U_1 - U_2)^2]$ is clearly non-negative. Moreover, irrespective of DM 1's strategy, DM 2 can achieve the optimal cost (equal to 0) by using the strategy

$$U_2 = g_2^*(\omega, U_1) = \omega + U_1.$$

The strategy g_2^* defined above is a (sequentially) dominant strategy for DM 2. Thus, we have determined an optimal strategy for the last DM *without any consideration of the strategy(ies) used by DM(s) who acted before.* This sequential strategy dominance is the fundamental reason why Theorem 1 allows us to sequentially find optimal strategies in a backward inductive manner for classical information structures.

2) Non-classical case: Consider now the non-classical information structure:

$$\mathcal{J}_1 = 2^{\Omega} \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}, \tag{15}$$

$$\mathcal{J}_2 = \{\emptyset, \Omega\} \times 2^{\mathbb{U}_1} \times \{\emptyset, \mathbb{U}_2\}. \tag{16}$$

In other words, DM 1's strategy can be any function of ω, while DM 2's can be any function U_1.

Consider now two possible strategy pairs:

$$U_1 = g_1^*(\omega) = \omega, \quad U_2 = g_2^*(U_1) = 2U_1,$$

$$U_1 = h_1^*(\omega) = 1 - \omega, \quad U_2 = h_2^*(U_1) = 1.$$

It is easy to verify that both strategy pairs achieve the optimal expected cost equal to 0 whereas strategy pairs (g_1^*, h_2^*) and (h_1^*, g_2^*) result in positive expected costs. This illustrates that neither g_2^* or h_2^* are sequentially dominant strategies for DM 2. In fact, for any strategy λ of DM 2, at least one of $J(g_1^*, \lambda)$ and $J(h_1^*, \lambda)$ is positive. This implies that there is no sequentially dominant strategy for DM 2. Therefore, we can not fix DM 2's strategy without taking into account the strategy of the DM who acted before. Thus, we cannot expect to obtain a sequential decomposition for this problem of the kind in Theorem 1 where we could obtain the last DM's strategy without considering the strategies of earlier DMs. This lack of sequential strategy dominance is the fundamental reason why the results of Theorem 1 do not extend to general non-classical information structures.

IV. Common Knowledge and Sequential Team Problems

For classical information structures, Theorem 1 provides a sequential decomposition for obtaining an optimal strategy profile. Since Theorem 1 is limited to classical information structures, we need a new methodology to obtain a similar decomposition for non-classical information structures. We will use common knowledge to construct such a decomposition. We will refer to this decomposition as the *common knowledge based dynamic program.*

A. Common Knowledge in Sequential Problems

Recall that the information of the $t-$th decision maker is characterized by $\mathcal{J}_t \subset \mathcal{F} \times \mathcal{U}_1 \times \cdots \times \mathcal{U}_{t-1} \times \{\emptyset, \mathbb{U}_t\} \times \ldots \{\emptyset, \mathbb{U}_T\} \subset \mathcal{F} \times \mathcal{U}_{1:T}$. We define the common knowledge at time t as the intersection of sigma algebras associated with DM t and with decision-makers *that have yet to make their decision*. That is, we define

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s. \tag{17}$$

Common knowledge was first defined in [2] in the context of static decision problems. A related definition of "common information" and "private information" for static decision problems was presented and discussed in [4], [5].

Lemma 5 (Properties of Common Knowledge)**.**

1) *Coarsening property:* $\mathcal{C}_t \subset \mathcal{J}_t$.
2) *Nestedness property:* $\mathcal{C}_t \subset \mathcal{C}_{t+1}$.
3) *Common observations: There exist observations* $Z_1, Z_2, \ldots, Z_T$*, with* Z_t *taking values in a finite measurable space* $(\mathbb{Z}_t, 2^{\mathbb{Z}_t})$ *and*

$$Z_t := \zeta_t(\omega, U_1, \ldots, U_{t-1}), \tag{18}$$

such that $\sigma(Z_{1:t}) = \mathcal{C}_t$*. These variables will be referred to as common observations.*

4) *Private Observations: There exist observations* $Y_1, Y_2, \ldots, Y_T$*, with* Y_t *taking values in a finite measurable space* $(\mathbb{Y}_t, 2^{\mathbb{Y}_t})$ *and*

$$Y_t := \eta_t(\omega, U_1, \ldots, U_{t-1}), \tag{19}$$

such that $\sigma(Z_{1:t}, Y_t) = \mathcal{J}_t$*. These variables will be referred to as private observations. Further, any* $\mathcal{J}_t/\mathcal{U}_t$ *measurable decision strategy can be written as*

$$U_t = g_t(Z_{1:t}, Y_t).$$

Proof. See Appendix C. □

Remark 4. *The proof of the Lemma 5 also describes a method of constructing the common and private observations. Without loss of generality, we will assume that* $\mathbb{Y}_t$ *is the set of first* $|\mathbb{Y}_t|$ *positive integers, that is,* $\mathbb{Y}_t = \{1, 2, \ldots, |\mathbb{Y}_t|\}$.

We can now state Problem 1 for a general information structure under Assumption 1 in terms of common and private observations as follows.

Problem 3. *Given common observations* $Z_t = \zeta_t(\omega, U_1, \ldots, U_{t-1})$*, taking values in* $\mathbb{Z}_t$ *for* $t = 1, \ldots, T$*, and private observations* $Y_t = \eta_t(\omega, U_1, \ldots, U_{t-1})$*, taking values in* $\mathbb{Y}_t$ *for* $t = 1, \ldots, T$*, find a decision strategy profile* $\mathbf{g} = (g_1, \ldots, g_T)$*, where* g_t *maps* $\mathbb{Z}_{1:t} \times \mathbb{Y}_t$ *to* $\mathbb{U}_t$ *for each* t*, that achieves*

$$\inf_{\mathbf{g}} \mathbb{E}[c(\omega, U_1, \ldots, U_T)] \quad \textit{exactly or within } \epsilon > 0,$$

where $U_t = g_t(Z_{1:t}, Y_t)$ *for each* t*.*

B. Common Knowledge based Dynamic Program

We now proceed as follows:

1) First, we formulate a new sequential decision-making problem from the point of view of a coordinator whose information at time t is described by the common knowledge sigma algebra $\mathcal{C}_t = \sigma(Z_{1:t})$ at time t.
2) Next, we show that for any strategy profile in Problem 3, we can construct an equivalent strategy in the coordinator's problem that achieves the same cost (with probability 1). Conversely, for any strategy in the coordinator's problem we can construct an equivalent decision strategy profile in Problem 3 that achieves the same cost (with probability 1).
3) Finally, we obtain a dynamic program for the coordinator's problem. This provides a sequential decomposition for Problem 3 due to the equivalence between the two problems established in Step 2.

We elaborate on these steps below.

Step 1: We consider the following modified problem. We start with the model of Problem 3 and introduce a coordinator who has the following features:

1) The coordinator's information at time t is characterized by the sigma algebra $\mathcal{C}_t = \sigma(Z_{1:t})$.
2) At each time t, the coordinator's decision space is the set of all functions from the space of DM t's private observation, $\mathbb{Y}_t$, to DM t's decision space $\mathbb{U}_t$. Note that the space of all function from $\mathbb{Y}_t$ to $\mathbb{U}_t$ can be identified with the product space $\mathbb{U}_t^{|\mathbb{Y}_t|} = \mathbb{U}_t \times \ldots \times \mathbb{U}_t$ (where the number of terms in the product is $|\mathbb{Y}_t|$). In case $\mathbb{Y}_t$ is a singleton, $\mathbb{U}_t^{|\mathbb{Y}_t|} = \mathbb{U}_t$. In the rest of this section, we will simply say that the coordinator selects an element from the set $\mathbb{U}_t^{|\mathbb{Y}_t|}$.

We use γ_t to denote the element from the set $\mathbb{U}_t^{|\mathbb{Y}_t|}$ selected by the coordinator at time t. Clearly, γ_t is a tuple of size $|\mathbb{Y}_t|$. For $y = 1, \ldots, |\mathbb{Y}_t|$, $\gamma_t(y)$ is denotes the yth component of this tuple.

Interpretation of γ_t: $\gamma_t(y)$ is to be interpreted as the decision prescribed by the coordinator to the $t-$th decision-maker if its private observation takes the value y. Thus, γ_t can be seen as a *prescription* to the $t-$th decision-maker that specifies for each value of DM t's private observation a prescribed decision. Given the prescription γ_t from the coordinator and the private observations Y_t of DM t, the decision taken by the $t-$th decision maker can be written as

$$U_t = \gamma_t(Y_t). \tag{20}$$

Procedure for selecting prescriptions: The coordinator chooses its prescription at time t, γ_t, as a function of the realization of the common observations up until time t. That is, the coordinator uses a sequence of functions $\boldsymbol{\psi} := (\psi_1, \psi_2, \ldots, \psi_T)$, where

$$\psi_t : \mathbb{Z}_{1:t} \mapsto \mathbb{U}_t^{|\mathbb{Y}_t|}, \tag{21}$$

to choose the prescription. The sequence of functions $\boldsymbol{\psi} := (\psi_1, \psi_2, \ldots, \psi_T)$ is referred to as the *coordinator's strategy*. If the realization of common observations by time t is $z_{1:t}$, the prescription chosen using the strategy $\boldsymbol{\psi}$ is $\psi_t(z_{1:t})$.

The optimization problem for the coordinator is to find a strategy $\boldsymbol{\psi}$ that achieves

$$\inf_{\boldsymbol{\psi}} \mathbb{E}[c(\omega, U_1, \ldots, U_T)] \;\; \text{exactly or within } \epsilon > 0,$$

where $U_t = \gamma_t(Y_t)$ and $\gamma_t = \psi_t(Z_{1:t})$.

Step 2: The key idea of this step is to establish an equivalence between Problem 3 and the coordinator's problem defined above. Consider a strategy profile $\mathbf{g} = (g_1, g_2 \ldots, g_T)$ in Problem 3. Under this strategy profile, $U_t = g_t(Z_{1:t}, Y_t)$, $t = 1\ldots, T$. This strategy profile induces a joint probability distribution on $\omega, U_{1:T}, Y_{1:T}, Z_{1:T}$. We denote this distribution by $\mathbb{P}^g(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T})$.

We will now construct a strategy for the coordinator using the strategy profile $\mathbf{g}$. Recall that DM t's strategy in Problem 3, g_t, maps $\mathbb{Z}_{1:t} \times \mathbb{Y}_t$ to $\mathbb{U}_t$. We will think of g_t as a collection of partial functions from $\mathbb{Y}_t$ to $\mathbb{U}_t$, one for each $z_{1:t} \in \mathbb{Z}_{1:t}$. For each $z_{1:t}$, the corresponding partial function from $\mathbb{Y}_t$ to $\mathbb{U}_t$ can be identified with an element of the set $\mathbb{U}_t^{|\mathbb{Y}_t|}$.

The main idea of constructing the coordinator's strategy from $\mathbf{g}$ is the following: for each time t,

1) For each realization $z_{1:t}$ of common observations, $g_t(z_{1:t}, \cdot) : \mathbb{Y}_t \mapsto \mathbb{U}_t$. This mapping from $\mathbb{Y}_t$ to $\mathbb{U}_t$ can be identified with an element in the product space $\mathbb{U}_t^{|\mathbb{Y}_t|}$.
2) For each realization $z_{1:t}$ of common observations, the coordinator will select the prescription (that is, an element from $\mathbb{U}_t^{|\mathbb{Y}_t|}$) identified with the mapping $g_t(z_{1:t}, \cdot) : \mathbb{Y}_t \mapsto \mathbb{U}_t$.
3) With a slight abuse of notation, we can describe the coordinator's strategy as

$$\psi_t(z_{1:t}) := g_t(z_{1:t}, \cdot).$$

The above expression is to be interpreted as follows: Recall that $\psi_t(z_{1:t})$ is an element of $\mathbb{U}_t^{|\mathbb{Y}_t|}$, that is, it is a tuple of size $|\mathbb{Y}_t|$. The above expression says that for $y = 1, \ldots, |\mathbb{Y}_t|$, the yth component of $\psi_t(z_{1:t})$ is given by $g_t(z_{1:t}, y)$.

The coordinator's strategy defined above induces a joint probability distribution on $\omega, U_{1:T}, Y_{1:T}, Z_{1:T}$. We denote this distribution by $\mathbb{P}^{\psi}(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T})$.

Lemma 6. *The probability distributions $\mathbb{P}^g$ and $\mathbb{P}^{\psi}$ are identical, that is, for any $\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}$,*

$$\mathbb{P}^{\psi}(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}) = \mathbb{P}^g(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}).$$

Consequently,

$$\mathbb{E}^{\psi}[c(\omega, U_1, \ldots, U_T)] = \mathbb{E}^{g}[c(\omega, U_1, \ldots, U_T)]$$

Proof. See Appendix D. □

We now go in the reverse direction: given a strategy for the coordinator $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_T)$, we will construct a strategy profile $\mathbf{h} = (h_1, \ldots, h_T)$ in Problem 3. The main idea of constructing $\mathbf{h}$ from $\boldsymbol{\phi}$ is the following: for each time t,

1) For each realization $z_{1:t}$ of common observations, $\phi_t(z_{1:t})$ is an element of $\mathbb{U}_t^{|\mathbb{Y}_t|}$, that is, it is a tuple of size $|\mathbb{Y}_t|$.
2) For each realization $z_{1:t}$ of common observations and realization y_t of the private observation in Problem 3, DM t's decision is the y_t-th component of $\phi_t(z_{1:t})$.
3) With a slight abuse of notation, we can describe DM t's strategy as

$$h_t(z_{1:t}, \cdot) := \phi_t(z_{1:t}).$$

The above expression is to be interpreted as follows: for $y = 1, \ldots, |\mathbb{Y}_t|$, $h_t(z_{1:t}, y)$ is the yth component of $\phi_t(z_{1:t})$.

The following lemma can be established using an argument identical to the one used to prove Lemma 6.

Lemma 7. *The probability distributions $\mathbb{P}^h$ and $\mathbb{P}^\phi$ are identical, that is, for any $\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}$,*

$$\mathbb{P}^h(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}) = \mathbb{P}^\phi(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}).$$

Consequently,

$$\mathbb{E}^h[c(\omega, U_1, \ldots, U_T)] = \mathbb{E}^\phi[c(\omega, U_1, \ldots, U_T)]$$

Lemmas 6 and 7 imply that we can first find an optimal strategy for the coordinator and then use it to construct optimal strategies in Problem 3.

Step 3: The key idea of this step is to show that the problem of finding an optimal strategy for the coordinator is a sequential decision-making problem with a classical information structure.

Recall that the coordinator at time t knows $Z_{1:t}$ and selects γ_t. Also recall that for each time t, we have

$$Z_t = \zeta_t(\omega, U_{1:t-1}) \tag{22}$$

$$Y_t = \eta_t(\omega, U_{1:t-1}) \tag{23}$$

$$U_t = \gamma_t(Y_t). \tag{24}$$

By eliminating $Y_{1:T}$ and $U_{1:T}$ from the above system of equations, we can construct functions $\Theta_1, \Theta_2, \ldots, \Theta_T$ such that

$$Z_t = \Theta_t(\omega, \gamma_{1:t-1}), \quad t = 1, \ldots, T. \tag{25}$$

Similarly eliminating $U_{1:T}$ from the cost, we can construct function C such that

$$c(\omega, U_{1:T}) = C(\omega, \gamma_{1:T}). \tag{26}$$

With these transformations, the coordinator's problem can now be written as follows.

Problem 4. *Given observations $Z_t = \zeta_t(\omega, \gamma_1, \ldots, \gamma_{t-1})$ taking values in $\mathbb{Z}_t$ for $t = 1, \ldots, T$, find a strategy $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_T)$ for the coordinator, where ψ_t maps $\mathbb{Z}_{1:t}$ to $\mathbb{U}_t^{|\mathbb{Y}_t|}$ for each t, that achieves*

$$\inf_{\boldsymbol{\psi}} \mathbb{E}[C(\omega, \gamma_1, \ldots, \gamma_T)] \quad \textit{exactly or within } \epsilon > 0,$$

where $\gamma_t = \psi_t(Z_{1:t})$ for each t.

Comparing Problem 4 with Problem 2, it is evident that Problem 4 is a sequential decision-making problem with classical information structure with the prescription γ_t as the coordinator's decision and $Z_{1:t}$ as its information at time t. Hence, we can use the analysis of Section III (in particular, Lemma 2 and Theorem 1) to find an optimal strategy for the coordinator.

As in Section III, we say that the realization $z_{1:t}, \tilde{\gamma}_{1:t-1}$ of the coordinator's observations and decisions is *feasible* if there exists $\hat{\omega} \in \Omega$ with $\mathbb{P}(\hat{\omega}) > 0$ such that $\Theta_k(\hat{w}, \tilde{\gamma}_{1:k-1}) = z_k$ for $k = 1, \ldots, t$. For a given feasible realization $z_{1:t}, \tilde{\gamma}_{1:t-1}$ in the coordinator's problem, the strategy-independent belief on Ω is given as

$$\pi_t(\tilde{\omega} | z_{1:t}, \tilde{\gamma}_{1:t-1}) := \frac{\mathbb{P}(\tilde{\omega}) \prod_{k=1}^{t} \mathbb{1}_{\{\Theta_k(\tilde{\omega}, \tilde{\gamma}_{1:k-1}) = z_k\}}}{\sum_{\hat{\omega}} \left[\mathbb{P}(\hat{\omega}) \prod_{k=1}^{t} \mathbb{1}_{\{\Theta_k(\hat{\omega}, \tilde{\gamma}_{1:k-1}) = z_k\}} \right]} \ . \tag{27}$$

We can now use the dynamic program of Theorem 1 for the coordinator's problem and obtain the following result.

Theorem 2. *For the coordinator's problem (Problem 4), an optimal strategy is given by the following dynamic program:*

1) *Define value functions $V_k(z_{1:k}, \tilde{\gamma}_{1:k-1})$ recursively as follows:*

$$V_T(z_{1:T}, \tilde{\gamma}_{1:T-1}) :=$$
$$\min_{\tilde{\gamma}_T \in \mathbb{U}_T^{|\mathbb{Y}_T|}} \mathbb{E}^{\pi_T}[C(\omega, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_T)|z_{1:T}, \tilde{\gamma}_{1:T-1}],$$
$$V_k(z_{1:k}, \tilde{\gamma}_{1:k-1}) :=$$
$$\min_{\tilde{\gamma}_k \in \mathbb{U}_k^{|\mathbb{Y}_k|}} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_k)|z_{1:k}, \tilde{\gamma}_{1:k-1}], \tag{28}$$

$k = T-1, \ldots, 2, 1$*, where* $Z_{k+1} = \Theta_{k+1}(\omega, \tilde{\gamma}_{1:k})$ *and the expectations are with respect to coordinator's strategy independent beliefs on* ω *(as defined in* (27)*).*

2) *The optimal strategy for the coordinator as a function of its observations and past decisions is given as*

$$\psi_T^*(z_{1:T}, \tilde{\gamma}_{1:T-1}) :=$$
$$argmin_{\tilde{\gamma}_T \in \mathbb{U}_T^{|\mathbb{Y}_T|}} \mathbb{E}^{\pi_T}[C(\omega, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_T)|z_{1:T}, \tilde{\gamma}_{1:T-1}],$$
$$\psi_k^*(z_{1:k}, \tilde{\gamma}_{1:k-1}) :=$$
$$argmin_{\tilde{\gamma}_k \in \mathbb{U}_k^{|\mathbb{Y}_k|}} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_k)|z_{1:k}, \tilde{\gamma}_{1:k-1}], \tag{29}$$

$k = T-1, \ldots, 2, 1.$

The dynamic program of Theorem 2 identifies an optimal strategy for the coordinator as a function of its common observations and past prescriptions. We can construct an equivalent strategy $\boldsymbol{\phi}$ for the coordinator by eliminating past prescriptions in a manner identical to Lemma 2 so that for each t,

$$\gamma_t = \phi_t(Z_{1:t}) = \psi_t^*(Z_{1:t}, \gamma_{1:t-1}).$$

Construction of optimal strategies in Problem 3: We can now construct an optimal strategy profile in Problem 3 (without the coordinator) using the construction of Lemma 7: For each realization $z_{1:t}$ of common observations and realization y_t of the private observation in Problem 3, DM t's decision is the y_t-th component of $\phi_t(z_{1:t})$ and we denote this by

$$h_t^*(z_{1:t}, \cdot) := \phi_t(z_{1:t}), \quad t = 1, \ldots, T. \tag{30}$$

Because $\boldsymbol{\phi}$ is an optimal strategy for the coordinator, Lemmas 6 and 7 imply that $\mathbf{h}^* = (h_1^*, \ldots, h_T^*)$ is an optimal strategy profile for the DMs in Problem 3.

Remark 5. *It should be clear that the constructed strategy h_t^* in* (30) *uses both common and private observations to decide DM t's decision. For each realization $z_{1:t}$ of the common observations at time t, the partial function $h_t^*(z_{1:t}, \cdot) : \mathbb{Y}_t \mapsto \mathbb{U}_t$ is precisely the prescription the coordinator would have selected under its optimal strategy if it observed $z_{1:t}$. One could say that in Problem 3, DM t first uses its common observations to figure out the prescription the coordinator would have selected had it been present and then uses its private observation to pick the prescribed action under the coordinator's prescription.*

C. Discussion

Theorem 2 provides a sequential decomposition for the coordinator's problem and, due to the equivalence established in Lemmas 6 and 7, for Problem 3 with a general (in particular, non-classical) information structure. We call this decomposition the common knowledge based dynamic program. It is important to emphasize some key differences between the common knowledge based dynamic program and the dynamic program for classical information structures given in Theorem 1: (i) At time k, the dynamic program in Theorem 1 involves a minimization over the set of decisions available to DM k, namely $\mathbb{U}_k$. The decomposition in Theorem 2, on the other hand, involves a minimization over the space of functions from $\mathbb{Y}_k$ to $\mathbb{U}_k$. (ii) For each realization of DM k's observations, the minimizing decision in Theorem 1 is an optimal decision for DM k for that realization of observations. In the decomposition of Theorem 2, for each realization of the *common observations* at time k, the minimizing $\tilde{\gamma}_k$ identifies an optimal mapping from private observation to decision for DM k.

We believe that the existence of a dynamic program in general sequential teams is an interesting result for the following reason: Given such a dynamic program, one can then start investigating whether the specific form of the information and cost structure in the given team problem may be exploited to simplify it. We believe this has to be done on a case-by-case basis as in classical dynamic program.

Finally, we can make a brief comment about the computational benefit of the common knowledge based dynamic program over a brute force search over all strategy profiles. Let $|\mathbb{Z}_t| = z, |\mathbb{Y}_t| = y$ and $|\mathbb{U}_t| = u$. Then, the number of possible strategy profiles for the team is $\prod_{k=1}^{T} u^{z^k y}$. In the common knowledge based dynamic program, the minimization at time k is over a set of size u^y. The total number of such minimization problems to be solved in the dynamic program is $\sum_{k=1}^{T} z^k u^{y(k-1)}$. Thus, the approximate complexity of the dynamic program can be taken to be $u^y \sum_{k=1}^{T} z^k u^{y(k-1)}$. Note that the time index appears as exponent of an exponent in the brute force complexity whereas it appears as an exponent in the dynamic program complexity. This indicates computational benefits from the dynamic program.

Remark 6. *Assumption 1 is important for the analysis presented in Section IV. For a general sequential* *team with infinite spaces, one may need additional technical conditions to ensure that common and private observations of Lemma 5 can be constructed and that the coordinator's strategies are well-defined measurable functions.*

D. Example

Consider a team problem with two decision-makers. The probability space we will consider is: $\Omega = \{1, 2, 3, 4, 5\}, \mathcal{F} = 2^{\Omega}$ with equal probabilities for all outcomes in Ω. The decision spaces of the two decision-makers are finite sets $\mathbb{U}_1$ and $\mathbb{U}_2$ respectively, each associated with the respective power-set sigma-algebra. The objective is to find strategies for the two DMs to minimize the expected value of $c(\omega, U_1, U_2)$. We consider the following information structure:

$$\mathcal{J}_1 = \sigma(\{1, 2\}, \{3, 4\}, \{5\}) \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}, \tag{31}$$

$$\mathcal{J}_2 = \sigma(\{1, 3\}, \{2, 4\}, \{5\}) \times 2^{\mathbb{U}_1} \times \{\emptyset, \mathbb{U}_2\}. \tag{32}$$

This information structure is non-classical since $\mathcal{J}_1 \not\subset \mathcal{J}_2$. As discussed before, we cannot obtain a classical dynamic program for such an information structure. Is there another way to obtain a sequential decomposition for this information structure? Our results in Section IV provide a positive answer to this question. For this example, this common knowledge based sequential decomposition can be obtained as follows:

(i) The common knowledge sigma-algebras are:

$$\mathcal{C}_1 = \bigcap_{s=1}^{2} \mathcal{J}_s, \quad \mathcal{C}_2 = \mathcal{J}_2. \tag{33}$$

(ii) We can now define common observations Z_1, Z_2 and private observations Y_1, Y_2 such that $\mathcal{C}_t = \sigma(Z_{1:t})$ and $\mathcal{J}_t = \sigma(Y_t, Z_{1:t})$ for $t = 1, 2$. Each observation (common or private) is a function from $\Omega \times \mathbb{U}_1 \times \mathbb{U}_2$ to a (suitably chosen) finite set. For our example, the following definitions will meet the requirements:

$$Z_1 = \zeta_1(\omega, U_1, U_2) = \zeta_1(\omega) := \mathbb{1}_{\{\omega \in \{5\}\}}, \tag{34}$$

$$Z_2 = \zeta_2(\omega, U_1, U_2) = \zeta_2(\omega, U_1) := \begin{cases} (1, U_1) \ if \ \omega \in \{1, 3\}, \\ (2, U_1) \ if \ \omega \in \{2, 4\}, \\ (5, U_1) \ if \ \omega \in \{5\}, \end{cases} \tag{35}$$

$$\begin{aligned} Y_1 &= \eta_1(\omega, U_1, U_2) = \eta_1(\omega) := \mathbb{1}_{\{\omega \in \{3,4\}\}}, \\ Y_2 &= \eta_2(\omega, U_1, U_2) := 0. \end{aligned} \tag{36}$$

(iii) In addition to the common and private observations, our sequential decomposition makes use of prescriptions that map private observations to decisions. For our example, we will use the prescription at $t = 1$: $\gamma_1 : \{0, 1\} \mapsto \mathbb{U}_1$. The space of all prescriptions can be written as $\mathbb{U}_1 \times \mathbb{U}_1$.

(iv) Based on common observations and prescriptions, we define the following beliefs on ω:

$$\pi_1(\tilde{\omega}|Z_1 = 1) = \begin{cases} 1 \ if \ \tilde{\omega} = 5, \\ 0 \ if \ \tilde{\omega} \neq 5 \end{cases} \tag{37}$$

$$\pi_1(\omega|Z_1 = 0) = \begin{cases} 1/4 \ if \ \omega \neq 5, \\ 0 \ if \ \omega = 5 \end{cases} \tag{38}$$

For each function $\tilde{\gamma}_1 : \{0, 1\} \mapsto \mathbb{U}_1$, define

$$\pi_2(\tilde{\omega}|Z_1 = 1, Z_2 = (5, u_1), \tilde{\gamma}_1) := \begin{cases} 1 \ if \ \tilde{\omega} = 5, \\ 0 \ if \ \tilde{\omega} \neq 5 \end{cases} \tag{39}$$

For $i = 1, 2$,

$$\pi_2(\tilde{\omega}|Z_1 = 0, Z_2 = (i, u_1), \tilde{\gamma}_1) \tag{40}$$

$$:= \frac{\mathbb{1}_{\{\zeta_2(\tilde{\omega},\tilde{\gamma}_1(\eta_1(\tilde{\omega})))=(1,u_1)\}}}{\sum_{\hat{\omega}\neq 5}\left[\mathbb{1}_{\{\zeta_2(\hat{\omega},\tilde{\gamma}_1(\eta_1(\hat{\omega})))=(i,u_1)\}}\right]} \tag{41}$$

if $\tilde{\omega} \neq 5$; $\pi_2(\tilde{\omega}|Z_1 = 0, Z_2 = (i, u_1)) := 0$ if $\tilde{\omega} = 5$.

(iv) We can now define value functions based on common observations and prescriptions:

$$V_2(z_{1:2}, \tilde{\gamma}_1) := \min_{u_2 \in \mathbb{U}_2} \mathbb{E}^{\pi_2}[c(\omega, \tilde{\gamma}_1(Y_1(\omega)), u_2)|z_{1:2}, \gamma_1],$$

$$V_1(z_1) := \min_{\tilde{\gamma}_1 \in \mathbb{U}_1 \times \mathbb{U}_1} \mathbb{E}^{\pi_1}[V_2(z_1, Z_2, \tilde{\gamma}_1)|z_1], \tag{42}$$

where the expectations at $t = 2, 1$ are with respect to the beliefs defined in (37)-(40). Our result in Section IV show that optimal strategies for the two DMs can be obtained from the value functions defined above in a straightforward manner. Thus, even though the information structure of the team was non-classical, we can still obtain a sequential decomposition of the strategy optimization problem.

V. Comparison with classical dynamic program and Witsenhasuen's standard form

The analysis of Section IV and Theorem 2 apply to all sequential team problems under Assumption 1. We consider two special cases below.

A. Classical information structure

We show that Theorem 2 is equivalent to the classical dynamic program of Theorem 1 when the sequential team problem has a classical information structure. The nestedness of information sigma-algebras in classical information structures (i.e., $\mathcal{J}_t \subset \mathcal{J}_{t+1}$ for all t) implies that the common knowledge sigma-algebra at time t is the same as $\mathcal{J}_t$:

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s = \mathcal{J}_t. \tag{43}$$

We can construct common observations as in Lemma 5 such that $\mathcal{C}_t = \sigma(Z_{1:t})$. Since $\mathcal{C}_t = \mathcal{J}_t$, the private observation can be defined as a constant

$$Y_t := \eta_t(\omega, U_1, \dots, U_{t-1}) := 1. \tag{44}$$

The implication of (43) is that the coordinator's information at time t is the same as DM t's information. Moreover, since Y_t is a constant, the coordinator's decision space $\mathbb{U}_t^{|\mathbb{Y}_t|} = \mathbb{U}_t$. The prescription γ_t is simply the decision to be taken at time t. Thus, in the classical information structure case, the coordinator prescribes decisions to DM t based on the observations $Z_{1:t}$.

Substituting $\gamma_t = U_t$ and using the fact that $|\mathbb{Y}_t| = 1$ for all t, it is easy to check that the result of Theorem 2 reduces to the result of Theorem 1. Thus, the dynamic program of Theorem 1 for classical information structures can be viewed as a special case of the common knowledge based dynamic program of Theorem 2.

B. Trivial common knowledge

In some information structures, the common knowledge among agents may just be the trivial sigma algebra:

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s = \{\emptyset, \Omega \times \mathbb{U}^{1:T}\}. \tag{45}$$

In this case, the common observations of Lemma 5 can be defined as constants:

$$Z_t := \zeta_t(\omega, U_1, \dots, U_{t-1}) := 1, \tag{46}$$

and the private observation at time t describes all the information of DM t. The coordinator's prescription at time t can be interpreted as DM t's strategy — it provides a decision for each possible realization of DM t's observations. Moreover, since the common observations are constants, the coordinator's problem can be viewed as an open loop control problem with the associated dynamic program given by Theorem 2. This is similar to the sequential decomposition of team problems in [16].

VI. CONNECTIONS WITH THE COMMON INFORMATION APPROACH

In the intrinsic model, the information of DM t is represented by a sigma-algebra $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_{1:T}$. Alternatively, the information of DM t could be described in terms of observations it has access to. Consider a team problem where for each t DM t has access to the following observations: $\tilde{Z}_{1:t}$, and $\tilde{Y}_t$. For each t, $\tilde{Z}_t$ and $\tilde{Y}_t$ are functions of $\omega, U_{1:t-1}$. We will refer to $\tilde{Z}_{1:t}$ as the *common information* at time t and $\tilde{Y}_t$ as the *private information* at time t.

Given the above information structure, we can follow the steps of Section IV-B, using $\tilde{Z}_{1:T}, \tilde{Y}_{1:T}$ instead of the common and private observations $Z_{1:T}, Y_{1:T}$ described in Lemma 5, to construct a coordinator's problem. The coordinator now knows the common information $\tilde{Z}_{1:t}$ at time t and selects prescriptions that map the private information $\tilde{Y}_t$ to decision U_t. Since this new version of the coordinator's problem is still a sequential decision-making problem with classical information structure, we can find its dynamic program in the same way as in Section IV-B. Such an approach for sequential team problems that uses common information among decision-makers to construct the coordinator's problem and its associated dynamic program was described in [12]. It was used in [13] and [14] for studying delayed history sharing and partial history sharing models in decentralized stochastic control problems.

We can make the following observations about the relationship between the common information approach summarized above and the common knowledge based approach of this paper:

Firstly, the common information approach for sequential teams requires the information structure described above: for each t DM t has access to $\tilde{Z}_{1:t}$, and $\tilde{Y}_t$. Thus, it requires that there is a part of the decision-makers' information that is nested over time. If no such part exists, one can still use the common information approach by creating degenerate observations $\tilde{Z}_t = 0$ for each t. As mentioned earlier, the common knowledge approach of this paper applies to any sequential information structure.

Secondly, the common information based dynamic program may be different from the common knowledge based dynamic program obtained in Section IV-B. To see why, note that the sigma-algebras associated with DM t in the above information structure is $\mathcal{J}_t = \sigma(\tilde{Z}_{1:t}, \tilde{Y}_t)$ and the common knowledge sigma-algebra at time t is $\mathcal{C}_t = \bigcap_{s=t}^{T} \mathcal{J}_s$. It is straightforward to see that the common information at time t, $\tilde{Z}_{1:t}$, is measurable with respect to $\mathcal{C}_t$. In other words, $\sigma(\tilde{Z}_{1:t}) \subset \mathcal{C}_t$. However, it may be the case that $\sigma(\tilde{Z}_{1:t})$ is a strict subset of $\mathcal{C}_t$ (see the first example in Section VI-A). Thus, the coordinator based on common knowledge may be more informed (i.e, it may be associated with a larger sigma-algebra) than a coordinator based only on common information. This difference between the two coordinators' information implies that the associated dynamic programs may be different.

Furthermore, the common information based dynamic program may be computationally more demanding than its common knowledge based counterpart. To see why, recall that we construct common and private observations in the common knowledge approach to ensure that $\sigma(Z_{1:t}) = \mathcal{C}_t$ and $\sigma(Z_{1:t}, Y_t) = \mathcal{J}_t$. Thus, we have that $\sigma(\tilde{Z}_{1:t}) \subset \sigma(Z_{1:t})$ but $\sigma(\tilde{Z}_{1:t}, \tilde{Y}_t) = \sigma(Z_{1:t}, Y_t)$. This implies that the private observations Y_t can take values in a smaller space than the original private information $\tilde{Y}_t$. This, in turn, implies that, the prescriptions in the common knowledge dynamic program take values in a smaller space ($\mathbb{U}_t^{|\mathbb{Y}_t|}$) than the space of prescriptions in the common information dynamic program ($\mathbb{U}_t^{|\tilde{\mathbb{Y}}_t|}$). Thus, the common information approach may result in a more complicated dynamic program than that resulting from the common knowledge approach.

Finally, we note that if the observations in common information were all constants (or if there were no common information), then $\sigma(\tilde{Z}_{1:t}) = \{\emptyset, \Omega \times \mathbb{U}^{1:T}\}$ and the common information based dynamic program will be identical to Witsenhausen's sequential decomposition [16]. Such an instance is presented in the second example of Section VI-A below.

A. Examples

1. To illustrate the difference between common information and common knowledge, we consider a team problem with two decision-makers. The probability space is: $\Omega = \{1, 2, 3\}, \mathcal{F} = 2^{\Omega}$ with equal probabilities for all outcomes in Ω. The information structure is given in terms of the observations each decision-maker has access to.

(i) DM 1 knows $\tilde{X}_1^1 = \mathbb{1}_{\{\omega=1\}}$ and $\tilde{X}_1^2 = \mathbb{1}_{\{\omega\in\{1,2\}\}}$.

(ii) DM 2 knows $\tilde{X}_1^1 = \mathbb{1}_{\{\omega=1\}}$ and $\tilde{X}_2^2 = \mathbb{1}_{\{\omega=3\}}$.

We can identify the common information at $t = 1$ to be $\tilde{Z}_1 = \tilde{X}_1^1$.

The common-knowledge sigma-algebra at $t = 1$ is given as $\mathcal{C}_1 = \sigma(\tilde{X}_1^1, \tilde{X}_1^2) \cap \sigma(\tilde{X}_1^1, \tilde{X}_2^2)$. It is easy to see that $\mathcal{C}_1 = 2^{\Omega} \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}$. A common observation at $t = 1$ that generates this common knowledge sigma algebra can be written as $Z_1 = \omega$. Clearly, a coordinator who knows Z_1 is more informed than a coordinator who knows only $\tilde{Z}_1$. In other words, $\sigma(\tilde{Z}_1)$ is a strict subset of $\sigma(Z_1)$.

2. To illustrate that the common information based dynamic program may be different from the one obtained using common knowledge, we consider a team problem with three decision-makers. The probability space is: $\Omega = \{1, 2, 3, 4, 5\}, \mathcal{F} = 2^{\Omega}$ with equal probabilities for all outcomes in Ω. The decision spaces of the three decision-makers are finite sets $\mathbb{U}_1, \mathbb{U}_2$ and $\mathbb{U}_3$ respectively, each associated with the respective power-set sigma-algebra. The objective is to find strategies for the DMs to minimize the expected value of $c(\omega, U_1, U_2, U_3)$. The information structure is given in terms of the observations each decision-maker has access to.

(i) DM 1 knows

$$\tilde{X}_1 := \begin{cases} 1 \; if \; \omega \in \{1,2\}, \\ 3 \; if \; \omega \in \{3,4\}, \\ 5 \; if \; \omega \in \{5\}. \end{cases}$$

(ii) DM 2 knows

$$\tilde{X}_2 := \begin{cases} 1 \; if \; \omega \in \{1,3\}, \\ 2 \; if \; \omega \in \{2,4\}, \\ 5 \; if \; \omega \in \{5\}. \end{cases}$$

(iii) DM 3 knows $\tilde{X}_1$ and

$$\tilde{X}_3 := \begin{cases} 1 \; if \; \omega \in \{1,4\}, \\ 2 \; if \; \omega \in \{2,5\}, \\ 3 \; if \; \omega \in \{3\}. \end{cases}$$

For this information structure there is no common information at $t = 1, 2$. In particular, there is no observation $\tilde{Z}_1$ that is available to all three DMs and there is no observation $\tilde{Z}_2$ that is available to DMs 2 and 3. The private informations are $\tilde{Y}_1 = \tilde{X}_1$, $\tilde{Y}_2 = \tilde{X}_2$ and $\tilde{Y}_3 = (\tilde{X}_1, X_3)$. Thus, the coordinator in the common information approach for this example will have no observations and the resulting dynamic program will be identical to Witsenhausen's sequential decomposition.

If we consider the sigma-algebras $\sigma(\tilde{X}_1), \sigma(\tilde{X}_2), \sigma(\tilde{X}_1, \tilde{X}_3)$ associated with the DMs, then it can be easily seen that the common knowledge sigma-algebras are non-trivial and given as:

(i) $\mathcal{C}_1 = \sigma(Z_1)$, where $Z_1 = \mathbb{1}_{\{\omega=5\}}$. In other words, $\mathcal{C}_1 = \sigma(\{1,2,3,4\},\{5\}) \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}$.

(ii) $\mathcal{C}_2 = \sigma(\tilde{X}_2)$. That is, $\mathcal{C}_2 = \sigma(\{1,3\},\{2,4\},\{5\}) \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}$.

(iii) $\mathcal{C}_3 = \sigma(\tilde{X}_1, \tilde{X}_3)$. In other words, $\mathcal{C}_3 = 2^{\Omega} \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}$.

Thus, in this example, the coordinator in the common knowledge based dynamic program will have non-trivial observations and the corresponding dynamic program will be distinct from Witsenhausen's sequential decomposition.

VII. Sequential orders and common knowledge

In Section II-C, we mentioned that an information structure is sequential if there exists a permutation $p : \{1, 2, \dots, T\} \mapsto \{1, 2, \dots, T\}$ such that

$$\mathcal{J}_{p(t)} \subset \mathcal{F} \otimes \mathcal{U}_{p(1)} \otimes \mathcal{U}_{p(2)} \otimes \cdots \otimes \mathcal{U}_{p(t-1)} \otimes \{\emptyset, \mathbb{U}_{p(t)}\} \otimes \\ \ldots \otimes \{\emptyset, \mathbb{U}_{p(T)}\}. \tag{47}$$

In some cases, there may be more than one permutation that satisfies the above causality condition. Let p, q be two distinct permutations satisfying the causality condition. Following the analysis of Section IV, the two permutations would result in two different definitions of common knowledge: For the permutation p, the common knowledge is

$$\mathcal{C}_t^p := \bigcap_{s=t}^{T} \mathcal{J}_{p(s)},$$

while for permutation q, the common knowledge is

$$\mathcal{C}_t^q := \bigcap_{s=t}^{T} \mathcal{J}_{q(s)}.$$

The two versions of common knowledge would produce two different dynamic programs. In general, it may not be clear apriori which permutation should be preferred for obtaining the dynamic program. Suppose it is the case that for all time t, $\mathcal{C}_t^q \subset \mathcal{C}_t^p$. Because the common knowledge under permutation q is a subset of the common knowledge under permutation p, the private observations under q may have to take values in a larger space than the private observations under p. This, in turn, implies that, the prescriptions in the dynamic program for permutation q take values in a larger space than the space of prescriptions in the dynamic program obtained using permutation p. Thus, one could argue that the permutation resulting in more common knowledge, that is the permutation p, should be preferred for obtaining the sequential decomposition of the problem.

Consider, for example, the information structure of the second example in Section VI-A. Consider the following two permutations of the three decision-makers: $p =$ (DM 1, DM 2, DM 3) and $q =$ (DM 2, DM 1, DM 3). Both permutations will satisfy the causality condition. However, it is not the case that the common knowledge sigma-algebras under one permutation are contained in the common knowledge sigma-algebras under the other permutation for all $t = 1, 2, 3$. Let's modify the observation of DM 2 in this example to the following:

$$\tilde{X}_2 := \begin{cases} 1 \; if \; \omega \in \{1, 2\}, \\ 3 \; if \; \omega \in \{3\}, \\ 4 \; if \; \omega \in \{4\}, \\ 5 \; if \; \omega \in \{5\}. \end{cases}$$

For this modified example, it can be easily verified that for all time $t = 1, 2, 3$, $\mathcal{C}_t^q \subset \mathcal{C}_t^p$. Thus, in this case the permutation $p =$ (DM 1, DM 2, DM 3) should be preferred for obtaining the sequential decomposition of the problem.

VIII. CONCLUSION

We considered sequential team problems based on Witsenhausen's intrinsic model. We started with the case of classical information structures and presented the dynamic program for this case. We then defined the concept of common knowledge in sequential team problems with general information structures. We showed how common knowledge can be used to construct a sequential decomposition of sequential team problems by means of an equivalent sequential decision-making problem that has a classical information structure. This equivalent problem was formulated from the perspective of a coordinator who knows the common knowledge. This common knowledge based sequential decomposition unifies the dynamic programming results of classical information structures and Witsenhausen's sequential decomposition of general sequential problems. In addition to providing an analytical and computational benefit, the development of sequential decomposition for problems with non-classical information structures opens up the possibility of systematic methods to find structural results and information states for decision-makers in such problems.

REFERENCES

[1] M. S. Andersland and D. Teneketzis. Information structures, causality, and nonsequential stochastic control i: Design-independent properties. *SIAM Journal on Control and Optimization*, 30(6):1447–1475, 1992.

[2] R. J. Aumann. Agreeing to disagree. *Annals of Statistics*, (4):1236–1239, 1976.

[3] P. R. de Waal and J. H. van Schuppen. A class of team problems with discrete action spaces: Optimality conditions based on multimodularity. *SIAM Journal on Control and Optimization*, 38(3):875–892, 2000.

[4] R. Dick, G. Hexner, and Y. Ho. On common information in decision problems. *IEEE Transactions on Information Theory*, 23(3):393–396, May 1977.

[5] G. Hexner and Y. Ho. Information structure: Common and private (corresp.). *IEEE Transactions on Information Theory*, 23(3):390–393, May 1977.

[6] Y. C. Ho and K. C. Chu. Team decision theory and information structures in optimal control problems-part 1. *IEEE Transaction on Automatic Control*, 17(1):15–22, 1972.

[7] Y. C. Ho, M. Kastner, and E. Wong. Teams, signaling and information theory. *IEEE Transaction on Automatic Control*, 23(2):305–312, 1978.

[8] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation Identification and Adaptive Control*. Prentice Hall, 1986.

[9] A. Mahajan. *Sequential decomposition of sequential dynamic teams: applications to real-time communication and networked control systems*. PhD thesis, University of Michigan, Ann Arbor, MI, Sept. 2008.

[10] A. Mahajan, N. Martins, M. Rotkowitz, and S. Yüksel. Information structures in optimal decentralized control. In *Proc. 51st IEEE Conf. Decision and Control*, pages 1291 – 1306, Maui, Hawaii, Dec. 2012.

[11] J. Marschak and R. Radner. *Economic Theory of Teams*. Yale University Press, New Haven, 1972.

[12] A. Nayyar. *Sequential decision making in decentralized systems*. PhD thesis, University of Michigan, Ann Arbor, MI, 2011.

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

[14] A. Nayyar, A. Mahajan, and D. Teneketzis. Decentralized stochastic control with partial history sharing: A common information approach. *IEEE Transactions on Automatic Control*, 58(7):1644–1658, July 2013.

[15] R. Radner. Team decision problems. *The Annals of Math. Statistics*, 33(3):857–881, Sept. 1962.

[16] H. Witsenhausen. A standard form for sequential stochastic control. *Mathematical Systems Theory*, 7(1):5–11, 1973.

[17] H. S. Witsenhausen. On information structures, feedback and causality. *SIAM Journal of Control*, 9(2):149–160, May 1971.

[18] H. S. Witsenhausen. The intrinsic model for discrete stochastic control: Some open problems. In A. Bensoussan and J. L. Lions, editors, *Control Theory, Numerical Methods and Computer System Modelling*, volume 107 of *Lecture Notes in Economics and Mathematical Systems*, pages 322–335. Springer Verlag, 1975.

[19] H. S. Witsenhausen. Equivalent stochastic control problems. *Mathematics of Control Signals and Systems*, 1:3–11, 1988.

[20] Y. C. Ho. Team decision theory and information structures. In *Proceedings of the IEEE*, volume 68, pages 644–654, 1980.

[21] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Birkhäuser, Boston, MA, 2013.

APPENDIX A
PROOF OF LEMMA 1

Since $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}^{1:T}$ is a finite sigma-algebra, it is completely characterized by a partition P_t : $\{P_t^1, P_t^2, \dots, P_t^{n_t}\}$ of the finite set $\Omega \times \mathbb{U}^{1:T}$. Moreover, since

$$\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \dots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \ldots \otimes \{\emptyset, \mathbb{U}_T\},$$

any set P_t^k in the partition P_t is of the form

$$P_t^k = A \times B_1 \times \ldots \times B_{t-1} \times \mathbb{U}_t \times \ldots \times \mathbb{U}_T,$$

where $A \subset \Omega$, $B_1 \subset \mathbb{U}_1, \dots, B_{t-1} \subset \mathbb{U}_{t-1}$.

Further, since $\mathcal{J}_t \subset \mathcal{J}_{t+1}$, the partition P_{t+1} associated with $\mathcal{J}_{t+1}$ is a finer partition than P_t. That is, P_t^k in the partition P_t can be written as union of sets from partition P_{t+1}:

$$P_t^k = P_{t+1}^{k(1)} \cup P_{t+1}^{k(2)} \cup \ldots \cup P_{t+1}^{k(m_k)},$$

for some indices $k(1), \dots, k(m_k)$. Define the observations $Z_1, \dots, Z_T$ as follows:

1) Note that each set in the partition at time 1 is of the form:

$$P_1^k = A \times \mathbb{U}_1 \times \ldots \times \mathbb{U}_T,$$

where $A \subset \Omega$. We define $Z_1 = \zeta_1(\omega)$ as follows:

$$\zeta_1(\tilde{\omega}) = \quad k \quad \text{ if } (\tilde{\omega}, u_{1:T}) \in P_1^k \ \forall u_{1:T} \tag{48}$$

2) Recall that any set in the partition at time 1 can be written as union of some sets in partition at time 2:

$$P_1^k = P_2^{k(1)} \cup P_2^{k(2)} \cup \ldots \cup P_2^{k(m_k)},$$

and that any set P_2^j is of the form:

$$P_2^j = A \times B_1 \times \mathbb{U}_2 \times \ldots \times \mathbb{U}_T,$$

where $A \subset \Omega$, $B_1 \subset \mathbb{U}_1$. We define $Z_2 = \zeta_2(\omega, U_1)$ as follows:

$$\begin{aligned} &\zeta_2(\tilde{\omega}, u_1) = i, \\ &\text{if } (\tilde{\omega}, u_{1:T}) \in P_1^k \text{ and } \ (\tilde{\omega}, u_1, u_{2:T}) \in P_2^{k(i)} \ \forall u_{2:T}. \end{aligned} \tag{49}$$

3) Proceeding in the same manner, we note that any set in the partition at time $t-1$ can be written as union of some sets in partition at time t:

$$P_{t-1}^k = P_t^{k(1)} \cup P_t^{k(2)} \cup \ldots \cup P_t^{k(m_k)},$$

and that any set P_t^j is of the form:

$$P_t^j = A \times B_1 \times \ldots \times B_{t-1} \times \mathbb{U}_t \times \ldots \times \mathbb{U}_T,$$

where $A \subset \Omega$, $B_1 \subset \mathbb{U}_1, \ldots, B_{t-1} \subset \mathbb{U}_{t-1}$. We define $Z_t = \zeta_t(\omega, U_1, U_2, \ldots, U_{t-1})$ as follows:

$$\begin{aligned} &\zeta_t(\tilde{\omega}, u_{1:t-1}) = i, \text{ if } (\tilde{\omega}, u_{1:t-2}, u_{t-1:T}) \in P_{t-1}^k \\ &\text{and } \ (\tilde{\omega}, u_{1:t-1}, u_{t:T}) \in P_t^{k(i)} \ \forall u_{t:T}. \end{aligned} \tag{50}$$

We now argue that for any realization $z_{1:t}$ of $Z_{1:t}$, the set

$$\begin{aligned} &\zeta^{-1}(z_{1:t}) := \\ &\{(\tilde{\omega}, u_{1:T}) \in \Omega \times \mathbb{U}^{1:T} \mid \zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k, k = 1, \ldots, t\} \end{aligned} \tag{51}$$

is either empty or equal to one of the sets in the partition P_t associated with the sigma-algebra $\mathcal{J}_t$. Suppose $\zeta^{-1}(z_{1:t})$ is not empty and includes $(\tilde{\omega}, u_{1:T})$. Since $P_1, \ldots, P_t$ are partitions of $\Omega \times \mathcal{U}^{1:T}$, we must have

$$(\tilde{\omega}, u_{1:T}) \in P_k^{a(k)}, \quad k = 1, \ldots, t, \tag{52}$$

for some indices $a(k), k = 1, \ldots, t$. Further, since the partitions get finer with time, we must that

$$P_t^{a(t)} \subset P_{t-1}^{a(t-1)} \subset \ldots \subset P_1^{a(1)}.$$

Consider any other vector $(\hat{\omega}, u_{1:T}) \in P_t^{a(t)}$. Evaluating the maps $\zeta_1, \ldots, \zeta_t$ on this vector, we see that

$$\zeta_k(\hat{\omega}, u_{1:k-1}) = \zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k, \quad k = 1, \ldots, t.$$

Further, if $(\hat{\omega}, u_{1:T}) \notin P_t^{a(t)}$, then let $s \leq t$ be the smallest time index such that $(\hat{\omega}, u_{1:T}) \notin P_s^{a(s)}$. Then $\zeta_s(\hat{\omega}, u_{1:T}) \neq \zeta_s(\tilde{\omega}, u_{1:T}) = z_s$. Hence, $(\hat{\omega}, u_{1:T}) \notin \zeta^{-1}(z_{1:t})$. Thus, we can conclude that $\zeta^{-1}(z_{1:t})$ is either empty or equal to $P_t^{a(t)}$ for some index $a(t)$.

Thus, the collection of non-empty inverse images $\{\zeta^{-1}(z_{1:t})\}$, for all possible realizations $z_{1:t}$, is equal to the collection the sets in partition associated with $\mathscr{J}_t$: $\{P_t^1, P_t^2, \ldots, P_t^{n_t}\}$. Thus,

$$\sigma(Z_{1:t}) = \mathscr{J}_t.$$

Finally, if f is a $\mathscr{J}_t/\mathcal{U}_t$ measurable function, then it is completely characterized by the values it takes for each set in the partition P_t. Since, the realization $z_{1:t}$ identifies a unique set in partition P_t, we can define

$$g_t(z_{1:t}) := \begin{cases} f(\zeta^{-1}(z_{1:t})), & \text{if } \zeta^{-1}(z_{1:t}) \neq \emptyset, \\ \text{arbitrary} & \text{if } \zeta^{-1}(z_{1:t}) = \emptyset \end{cases}$$

With the above definition, it is easy to check that for any $(\tilde{\omega}, u_{1:T})$, if the corresponding realization of the observations is $z_{1:t}$, then $f(\tilde{\omega}, u_{1:T}) = g_t(z_{1:t})$.

APPENDIX B
PROOF OF LEMMA 2

The proof uses a simple repeated substitution argument: Consider a strategy profile $g_1, \ldots, g_T$ in the expanded information structure. Under this strategy we have the following T equations: $U_t = g_t(Z_{1:t}, U_{1:t-1}), \quad t = 1, \ldots, T$. We can eliminate $U_{1:t-1}$ from each equation by repeated substitution to obtain an equivalent system of equations: $U_t = h_t(Z_{1:t}), \quad t = 1, \ldots, T$. (For example, $U_1 = h_1(Z_1) := g_1(Z_1); U_2 = h_2(Z_1, Z_2) := g_2(Z_1, Z_2, h_1(Z_1))$ and so on.) It is straightforward to establish that for any realization $\tilde{\omega}$ of the random vector ω, the strategy profiles $g_{1:T}$ and $h_{1:T}$ generate the same realizations of observations and decisions and hence incur the same cost.

APPENDIX C
PROOF OF LEMMA 5

Parts 1 and 2 follow directly from the definition of common knowledge in (17).

Because $\mathcal{C}_t, t = 1, \ldots, T$ satisfy the nestedness property and $\mathcal{C}_t$, being a subset of $\mathscr{J}_t$ satisfies

$$\mathcal{C}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \ldots \otimes \{\emptyset, \mathbb{U}_T\},$$

we can follow the construction of observations used in proving Lemma 1 to construct common observations $Z_t = \zeta_t(\omega, U_1, \ldots, U_{t-1})$ such that $\sigma(Z_{1:t}) = \mathcal{C}_t$.

To construct the private observation at time t, we consider the partitions P_t and Q_t of the finite set $\Omega \times \mathbb{U}^{1:T}$ that generate $\mathcal{J}_t$ and $\mathcal{C}_t$ respectively. Since $\mathcal{C}_t \subset \mathcal{J}_t$, the partition P_t associated the $\mathcal{J}_t$ is a finer partition than Q_t. Thus, a set Q_t^k in the partition Q_t can be written as union of sets from partition P_t:

$$Q_t^k = P_t^{k(1)} \cup P_t^{k(2)} \cup \ldots \cup P_t^{k(m_k)},$$

for some indices $k(1), \ldots, k(m_k)$. We now define $Y_t = \eta_t(\omega, U_1, U_2, \ldots, U_{t-1})$ as follows:

$$\begin{aligned} &\eta_t(\tilde{\omega}, u_{1:t-1}) = i, \\ &\text{ if } (\tilde{\omega}, u_{1:t-1}, u_{t:T}) \in Q_t^k \text{ and } \ (\tilde{\omega}, u_{1:t-1}, u_{t:T}) \in P_t^{k(i)} \ \forall u_{t:T}. \end{aligned} \tag{53}$$

To prove that $\sigma(Z_{1:t}, Y_t) = \mathcal{J}_t$, we first define the inverse map

$$\begin{aligned} &\xi^{-1}(z_{1:t}, y_t) := \\ &\{(\tilde{\omega}, u_{1:T}) \in \Omega \times \mathbb{U}^{1:T} \mid \zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k, k = 1, \ldots, t, \\ &\ \eta_t(\tilde{\omega}, u_{1:t-1}) = y_t\}. \end{aligned} \tag{54}$$

We now argue that $\xi^{-1}(z_{1:t}, y_t)$ is either empty or equal to one of the sets in the partition P_t associated with the sigma-algebra $\mathcal{J}_t$. Suppose $\xi^{-1}(z_{1:t}, y_t)$ is not empty and includes $(\tilde{\omega}, u_{1:T})$. Since Q_t, P_t are partitions of $\Omega \times \mathcal{U}^{1:T}$, we must have

$$(\tilde{\omega}, u_{1:T}) \in Q_t^a, \quad (\tilde{\omega}, u_{1:T}) \in P_t^b \tag{55}$$

for some indices a, b. Further, we must have that

$$P_t^b \subset Q_t^a.$$

Consider any other vector $(\hat{\omega}, u_{1:T}) \in P_t^b$. Since $(\hat{\omega}, u_{1:T})$ must also belong to Q_t^a and the partition Q_t generates $\sigma(Z_{1:t})$, it follows that $(\hat{\omega}, u_{1:T})$ and $(\tilde{\omega}, u_{1:T})$ will produce the same realization of $Z_{1:t}$, that is,

$$\zeta_k(\hat{\omega}, u_{1:k-1}) = \zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k, \quad k = 1, \ldots, t.$$

Further, by definition, the private observation map η_t evaluated on $(\hat{\omega}, u_{1:T})$ and $(\tilde{\omega}, u_{1:T})$ would give the same value, that is,

$$\eta_t(\hat{\omega}, u_{1:t-1}) = \eta_t(\tilde{\omega}, u_{1:t-1}) = y_t.$$

Finally, if $(\hat{\omega}, u_{1:T}) \notin P_t^b$, then it would necessarily result in a different realization of either the common observations $z_{1:t}$ (if $(\hat{\omega}, u_{1:T}) \notin Q_t^a$) or the private observation y_t (if $(\hat{\omega}, u_{1:T}) \in Q_t^a$, but $(\hat{\omega}, u_{1:T}) \notin P_t^b$).

Thus, we can conclude that $\xi^{-1}(z_{1:t}, y_t)$ is either empty or equal to P_t^b for some index b. Therefore, the collection of non-empty inverse images $\{\xi^{-1}(z_{1:t}, y_t)\}$, for all possible realizations of $z_{1:t}, y_t$, is equal to the collection the sets in the partition associated with $\mathcal{J}_t$: $\{P_t^1, P_t^2, \ldots, P_t^{n_t}\}$. Thus,

$$\sigma(Z_{1:t}, Y_t) = \mathcal{J}_t.$$

Finally, if f is a $\mathcal{J}_t/\mathcal{U}_t$ measurable function, then it is completely characterized by the values it takes for each set in the partition P_t. Since, the realization $z_{1:t}, y_t$ identifies a unique set in partition P_t, we can define

$$g_t(z_{1:t}, y_t) := \begin{cases} f(\xi^{-1}(z_{1:t}, y_t)), & \text{if } \xi^{-1}(z_{1:t}, y_t) \neq \emptyset, \\ \text{arbitrary} & \text{if } \xi^{-1}(z_{1:t}, y_t) = \emptyset \end{cases}$$

With the above definition, it is easy to check that for any $(\tilde{\omega}, u_{1:T})$, if the corresponding realization of the common and private observations is $z_{1:t}, y_t$, then $f(\tilde{\omega}, u_{1:T}) = g_t(z_{1:t}, y_t)$.

APPENDIX D
PROOF OF LEMMA 6

The probability distribution under $\mathbf{g}$ can be factorized as

$$\mathbb{P}^g(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}) = \mathbb{P}(\tilde{\omega}) \times \prod_{k=1}^{T} \left[\mathbb{1}_{\{\eta_k(\tilde{\omega}, u_{1:k-1}) = y_k\}} \mathbb{1}_{\{\zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k\}} \mathbb{1}_{\{g_k(z_{1:k}, y_k) = u_k\}} \right] \tag{56}$$

Recall that $\psi_t(z_{1:t})$ is an element of $\mathbb{U}_t^{|\mathbb{Y}_t|}$, that is, it is a tuple of size $|\mathbb{Y}_t|$. Let us denote by $\psi_t(z_{1:t})(y)$ the yth component of this tuple. The definition of $\psi_t(z_{1:t})$ implies that $\psi_t(z_{1:t})(y) = g_t(z_{1:t}, y)$.

The probability distribution $\mathbb{P}^\psi$ is

$$\mathbb{P}^\psi(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}) = \mathbb{P}(\tilde{\omega}) \times \prod_{k=1}^{T} \left[\mathbb{1}_{\{\eta_k(\tilde{\omega}, u_{1:k-1}) = y_k\}} \mathbb{1}_{\{\zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k\}} \mathbb{1}_{\{\psi_k(z_{1:k})(y_k) = u_k\}} \right] \tag{57}$$

Since for any time t and any y, $\psi_t(z_{1:t})(y) = g_t(z_{1:t}, y)$, it follows that

$$\begin{aligned} &\mathbb{P}^\psi(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}) = \mathbb{P}(\tilde{\omega}) \times \\ &\prod_{k=1}^{T} \left[\mathbb{1}_{\{\eta_k(\tilde{\omega}, u_{1:k-1}) = y_k\}} \mathbb{1}_{\{\zeta_k(\tilde{\omega}, u_{1:k-1}) = z_k\}} \mathbb{1}_{\{g_k(z_{1:k}, y_k) = u_k\}} \right] \\ &= \mathbb{P}^g(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}). \end{aligned} \tag{58}$$

APPENDIX E
PROOF OF THEOREM 1

We will first identify a sequentially dominant strategy for DM T using the strategy-independent belief $\pi_T(\cdot|z_{1:T}, u_{1:T-1})$ on Ω.

Definition 1. *Define*

$$g_T^*(z_{1:T}, u_{1:T-1}) := argmin_{u_T \in \mathbb{U}_T} \mathbb{E}^{\pi_T}[c(\omega, u_1 \ldots, u_{T-1}, u_T)|z_{1:T}, u_{1:T-1}], \tag{59}$$

where $\mathbb{E}^{\pi_T}[c(\omega, u_1 \ldots, u_{T-1}, u_T)|z_{1:T}, u_{1:T-1}]$ *is the expectation with respect to strategy-independent beliefs on* ω *(as defined in* (9)*). Also define*

$$\begin{aligned} V_T(z_{1:T}, u_{1:T-1}) &:= \min_{u_T \in \mathbb{U}_T} \mathbb{E}^{\pi_T}[c(\omega, u_1 \ldots, u_T)|z_{1:T}, u_{1:T-1}] \\ &= \mathbb{E}^{\pi_T}[c(\omega, u_1 \ldots, u_{T-1}, g_T^*(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}]. \end{aligned} \tag{60}$$

We will refer to the function $V_T(\cdot)$ defined in (60) as the value function at time T.

Lemma 8. g_T^* *is a sequentially dominant strategy for DM* T*.*

Proof.

$$\begin{aligned} & J(g_1, g_2, \ldots, g_T) \\ &= \mathbb{E}^{g_1,\ldots,g_T}[c(\omega, U_1, \ldots, U_T)] \\ &= \mathbb{E}^{g_1,\ldots,g_T}[c(\omega, U_1, \ldots, U_{T-1}, g_T(Z_{1:T}, U_{1:T-1}))] \\ &= \mathbb{E}^{g_1,\ldots,g_T}[\mathbb{E}^{g_{1:T}} c(\omega, U_1, \ldots, U_{T-1}, g_T(Z_{1:T}, U_{1:T-1}))|Z_{1:T}, U_{1:T-1}] \\ &= \sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) \mathbb{E}^{g_{1:T-1}}[c(\omega, u_{1:T-1}, g_T(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}] && (61) \\ &= \sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) \mathbb{E}^{\pi_T}[c(\omega, u_{1:T-1}, g_T(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}] && (62) \\ &\geq \sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) \mathbb{E}^{\pi_T}[c(\omega, u_{1:T-1}, g_T^*(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}] && (63) \end{aligned}$$

where we use the fact that $g_{1:T-1}$ determine the distribution on $Z_{1:T}, U_{1:T-1}$ in (61), Lemma 3 in (62) and the definition of $g_T^*(z_{1:T}, u_{1:T-1})$ in (63).

Repeating the above steps for $g_1, \ldots, g_{T-1}, g_T^*$ gives

$$\begin{aligned} & J(g_1, g_2, \ldots, g_T^*) = \mathbb{E}^{g_1,\ldots,g_T^*}[c(\omega, U_1, \ldots, U_T)] \\ &= \sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) \mathbb{E}^{\pi_T}[c(\omega, u_{1:T-1}, g_T^*(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}] && (64) \end{aligned}$$

Combining (63) and (64) gives

$$J(g_1, g_2, \dots, g_T) \geq J(g_1, \dots, g_{T-1}, g_T^*),$$

hence proving the lemma. □

Because Lemma 8 establishes g_T^* as a dominant strategy for DM T, we can fix g_T^* as DM T's strategy without compromising optimality. That is, there is an optimal strategy profile where DM T is using strategy g_T^*.

Remark 7. *The strategy independence of belief is crucial for proving that g_T^* is a dominant strategy. A central difficulty in non-classical information structures is the non-existence of a dominant strategy at the last stage.*

Having found a dominant strategy at the last stage, we can essentially eliminate the last stage from our decision problem and view $T-1$ stage as the new last stage. After the $T-1$ stage, DM T and the stochastic system behave according to fixed functions and result in a cost that can be viewed as depending only on ω and the decisions in the first $T-1$ stages.

With the strategy of DM T fixed to g_T^*, we can express the expected cost of a strategy profile in terms of value function at time T using the following lemma.

Lemma 9. *For any $g_1, \dots, g_{T-1}$,*

$$J(g_1, \dots, g_{T-1}, g_T^*) = \mathbb{E}^{g_1,\dots,g_{T-1},g_T^*}[c(\omega, U_1, \dots, U_T)] = \mathbb{E}^{g_1,\dots,g_{T-1}}[V_T(Z_{1:T}, U_{1:T-1})].$$

Proof. From (64), we know that

$$\begin{aligned} &J(g_1, g_2, \dots, g_{T-1}, g_T^*) = \mathbb{E}^{g_1,\dots,g_T^*}[c(\omega, U_1, \dots, U_T)] \\ &= \sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) \mathbb{E}^{\pi_T}[c(\omega, u_{1:T-1}, g_T^*(z_{1:T}, u_{1:T-1}))|z_{1:T}, u_{1:T-1}], \end{aligned} \tag{65}$$

Using the definition of $V_T(\cdot)$ from (60), the right hand side in (65) can be written as

$$\sum_{z_{1:T}, u_{1:T-1}} \mathbb{P}^{g_{1:T-1}}(z_{1:T}, u_{1:T-1}) V_T(z_{1:T}, u_{1:T-1}) = \mathbb{E}^{g_1,\dots,g_{T-1}}[V_T(Z_{1:T}, U_{1:T-1})], \tag{66}$$

which proves the lemma. □

General Induction Step: We now proceed inductively by first assuming the following induction hypothesis for DMs $k+1$ to DM T and then using it to prove the analogous statement for DMs k to T.

Induction Hypothesis: Suppose there exist sequentially dominant strategies $g^*_{k+1}, g^*_{k+2}, \ldots, g^*_T$ *for DMs* $k+1$ *to* T *respectively and there exists a function* $V_{k+1}(Z_{1:k+1}, U_{1:k})$ *such that for any* $g_1, \ldots, g_k$*,*

$$J(g_1, \ldots, g_k, g^*_{k+1}, \ldots, g^*_T) \\ = \mathbb{E}^{g_1, \ldots, g_k, g^*_{k+1}, \ldots, g^*_T}[c(\omega, U_1, \ldots, U_T)] = \mathbb{E}^{g_1, \ldots, g_k}[V_{k+1}(Z_{1:k+1}, U_{1:k})]. \tag{67}$$

The induction hypothesis implies that for any strategies $g_1, \ldots, g_T$

$$J(g_1, \ldots, g_T) \geq J(g_1, \ldots, g_k, g^*_{k+1}, \ldots, g^*_T).$$

We have already established the hypothesis for $k+1 = T$ in Lemmas 8 and 9 above.

We now focus on DM k. We say that h_k is a sequentially dominant strategy for DM k if for any strategies $g_1, g_2, \ldots, g_k$,

$$J(g_1, g_2, \ldots, g_k, g^*_{k+1}, \ldots, g^*_T) \geq J(g_1, \ldots, g_{k-1}, h_k, g^*_{k+1}, \ldots, g^*_T).$$

Definition 2. *Define*

$$g^*_k(z_{1:k}, u_{1:k-1}) := argmin_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \ldots, u_k) | z_{1:k}, u_{1:k-1}],$$

where $Z_{k+1} = \zeta_{k+1}(\omega, u_{1:k})$ *and the expectation on the right hand side is with respect to strategy-independent beliefs on* ω *(as defined in* (9)*). Also, define*

$$V_k(z_{1:k}, u_{1:k-1}) := \min_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \ldots, u_k) | z_{1:k}, u_{1:k-1}]. \tag{68}$$

Lemma 10. g^*_k *is a sequentially dominant strategy for DM* k*. That is, for any strategies* $g_1, g_2, \ldots, g_k$*,*

$$J(g_1, g_2, \ldots, g_k, g^*_{k+1}, \ldots, g^*_T) \geq J(g_1, \ldots, g_{k-1}, g^*_k, \ldots, g^*_T).$$

Proof. See Appendix F. □

Lemma 11. *For any* $g_1, \dots, g_{k-1}$,

$$J(g_1, \dots, g_{k-1}, g_k^*, \dots, g_T^*) = \mathbb{E}^{g_1,\dots,g_{k-1},g_k^*,\dots,g_T^*}[c(\omega, U_1, \dots, U_T)]$$
$$= \mathbb{E}^{g_1,\dots,g_{k-1}}[V_k(Z_{1:k}, U_{1:k-1})]. \tag{69}$$

Proof. See Appendix G. □

The induction hypothesis and Lemma 10 shows that $g_k^*, \dots, g_T^*$ are sequentially dominant strategies for DMs k to T respectively. Lemma 11 shows that there exists a function $V_k(Z_{1:k}, U_{1:k-1})$ such that for any $g_1, \dots, g_{k-1}$,

$$J(g_1, \dots, g_{k-1}, g_k^*, \dots, g_T^*) = \mathbb{E}^{g_1,\dots,g_{k-1},g_k^*,\dots,g_T^*}[c(\omega, U_1, \dots, U_T)]$$
$$= \mathbb{E}^{g_1,\dots,g_{k-1}}[V_k(Z_{1:k}, U_{1:k-1})]. \tag{70}$$

In other words, if the induction hypothesis is true for DMs $k+1$ to T, then it must also hold for DMs k to T. Since we have already established the induction hypothesis when $k+1 = T$, it follows that it is true for all $k = T, \dots, 1$. Hence, $g_T^*, \dots, g_1^*$ defined according to Definitions 1 and 2 are sequentially dominant strategies for DMs T to 1 respectively.

APPENDIX F
PROOF OF LEMMA 10

We first note that $V_k(z_{1:k}, u_{1:k-1})$ can also be written as

$$V_k(z_{1:k}, u_{1:k-1}) :=$$
$$\min_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k}), u_1 \dots, u_k)|z_{1:k}, u_{1:k-1}], \tag{71}$$

Now,

$$J(g_1, g_2, \ldots, g_k, g_{k+1}^*, \ldots, g_T^*) = \mathbb{E}^{g_1, g_2, \ldots, g_k}[V_{k+1}(Z_{1:k+1}, U_{1:k})]$$
$$= \mathbb{E}^{g_1, \ldots, g_k}[\mathbb{E}^{g_{1:k}}[V_{k+1}(Z_{1:k+1}, U_{1:k}) | Z_{1:k}, U_{1:k-1}]$$
$$= \sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times$$
$$\mathbb{E}^{g_{1:k}}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k-1}, g_k(z_{1:k}, u_{1:k-1})), u_{1:k-1},$$
$$g_k(z_{1:k}, u_{1:k-1})) | z_{1:k}, u_{1:k-1}] \tag{72}$$
$$= \sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times$$
$$\mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k-1}, g_k(z_{1:k}, u_{1:k-1})), u_{1:k-1},$$
$$g_k(z_{1:k}, u_{1:k-1})) | z_{1:k}, u_{1:k-1}] \tag{73}$$
$$\geq \sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times$$
$$\mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k-1}, g_k^*(z_{1:k}, u_{1:k-1})), u_{1:k-1},$$
$$g_k^*(z_{1:k}, u_{1:k-1})) | z_{1:k}, u_{1:k-1}] \tag{74}$$

where we use the fact that $g_{1:k-1}$ determine the distribution on $Z_{1:k}, U_{1:k-1}$ in (72), Lemma 3 in (73) and the definition of $g_k^*(z_{1:k}, u_{1:k-1})$ in (74).

Repeating the above steps for $g_1, \ldots, g_{k-1}, g_k^*, \ldots, g_T^*$ gives

$$J(g_1, g_2, \ldots, g_{k-1}, g_k^*, g_{k+1}^*, \ldots, g_T^*)$$
$$= \mathbb{E}^{g_1, g_2, \ldots, g_k^*}[V_{k+1}(Z_{1:k+1}, U_{1:k})]$$
$$= \sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times$$
$$\mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k-1}, g_k^*(z_{1:k}, u_{1:k-1})), u_{1:k-1},$$
$$g_k^*(z_{1:k}, u_{1:k-1})) | z_{1:k}, u_{1:k-1}] \tag{75}$$

Combining (74) and (75) establishes the result.

APPENDIX G
PROOF OF LEMMA 11

From (75), we know that

$$\begin{aligned} &J(g_1, g_2, \ldots, g_{k-1}, g_k^*, \ldots, g_T^*) \\ &= \sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times \\ &\quad \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, \zeta_{k+1}(\omega, u_{1:k-1}, g_k^*(z_{1:k}, u_{1:k-1})), u_{1:k-1}, \\ &\quad\quad g_k^*(z_{1:k}, u_{1:k-1}))|z_{1:k}, u_{1:k-1}]. \end{aligned} \tag{76}$$

Using the definition of $V_k(\cdot)$ from (68), the right hand side in (76) can be written as

$$\begin{aligned} &\sum_{z_{1:k}, u_{1:k-1}} \mathbb{P}^{g_{1:k-1}}(z_{1:k}, u_{1:k-1}) \times V_k(z_{1:k}, u_{1:k-1}) \\ &= \mathbb{E}^{g_1, \ldots, g_{k-1}}[V_k(Z_{1:k}, U_{1:k-1})], \end{aligned} \tag{77}$$

which proves the lemma.