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 [17]. 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 nontrivial 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].**

***Index Terms*—Common knowledge, stochastic optimal control, stochastic systems, team theory.**

I. Introduction

This note deals with the problem of decentralized decision making. Such problems arise in any system where multiple agents/decision makers (DMs) 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 which are as follows:

1) *Cooperative,* i.e., problems where different DMs share the same objective. Such problems are called *team problems* [3], [6], [7], [10], [11], [16], [21], [22];
2) *Stochastic,* i.e., problems where stochastic models of uncertainties are available and the goal is to minimize the expected value of the system cost;
3) *Sequential,* i.e., problems where the DMs act in a predetermined 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.

Manuscript received November 27, 2018; accepted March 24, 2019. Date of publication April 22, 2019; date of current version December 3, 2019. This work was supported in part by the ARO Award No. W911NF-17-1-0232 and in part by the NSF under Grant ECCS 1509812, Grant ECCS 1750041, and Grant CNS-1238962. Recommended by Associate Editor D. Castanon. *(Corresponding author: Ashutosh Nayyar.)*

A. Nayyar is with the Electrical and Computer Engineering Department, University of Southern California, Los Angeles, CA 90007 USA (e-mail: ashutosn@usc.edu).

D. Teneketzis is with the Department of EECS, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: teneketzis@eecs.umich.edu).

Digital Object Identifier 10.1109/TAC.2019.2912536

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 the 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], [19]. This method allows us to decompose the problem of finding optimal strategies for all agents into several smaller problems, which must be solved sequentially backward 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 [17]. 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 subproblems that must be solved sequentially backward in time. In each of these subproblems, 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 [17] and [9].

In this note, 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 DMs. We show that DMs 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 DMs and with finite probability and decision spaces. We can make three observations about our common knowledge based decomposition as follows:

1) If the underlying information structure is classical, our sequential decomposition reduces to the classical dynamic program.
2) For information structures with *nontrivial 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 subproblems than those in Witsenhausen's standard form approach.
3) For information structures with *trivial common knowledge* (see Section V), our decomposition is similar in spirit to Witsenhausen's standard form based decomposition [17]. 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 note 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 note 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], [18], [19] to present our results for general sequential team problems. Models similar to the intrinsic model have been presented in [20]. 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 tth 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 tth decision can be viewed a measurable function from the measurable space $(\Omega \times \mathbb{U}_1 \times \cdots \times \mathbb{U}_T, \mathcal{J}_t)$ to the measurable space $(\mathbb{U}_t, \mathcal{U}_t)$, where $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_T$ is a sigma-algebra that denotes the maximal information (knowledge) that can be used to select the tth decision.

Organization: The rest of the note is organized as follows. We describe the intrinsic model and information structures in Section II. We review 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 conclude in Section VII.

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 $\tilde{}$ or $\hat{}$ 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 $A_1 \times \cdots \times A_t$. $\mathbb{R}$ 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 and the Optimization Problem

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 as being chosen by a distinct DM.[1] The information available to the tth decision maker (DM t) may depend on the realization of ω and the decisions made by other DMs. In the intrinsic model [1], [18], [19], 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 $(\mathbb{U}_t, \mathcal{U}_t)$, 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 DM and the collection of all T decision strategies $\mathbf{g} = (g_1, g_2, \ldots, g_T)$ is called the *decision strategy profile*.

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 the 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 as described in (3) for each t, that achieves

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

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:

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

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 [18] 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 note.

B. Information Structures

The sigma algebras $\mathcal{J}_1, \ldots, \mathcal{J}_T$ specify the information available for making each of the T decisions and are collectively 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 $\mathcal{F} \otimes \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_T$.

Sequential and Nonsequential 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 $t = 1, \ldots, T$,

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

Otherwise, the information structure is said to be nonsequential.

The sequence $p(1), \ldots, p(T)$ 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 follows:
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 $t = 1, \ldots, T-1$.

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

3) Quasi-classical (partially nested): Recall that for sequential information structures

$$\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \cdots \{\emptyset, \mathbb{U}_T\}. \quad (6)$$

For $s < t$, we say that the decision U_s *does not affect the information of the* tth *DM* if

$$\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_{s-1} \otimes \{\emptyset, \mathbb{U}_s\} \otimes \mathcal{U}_{s+1} \otimes \cdots \otimes \mathcal{U}_{t-1} \otimes \{\emptyset, \mathbb{U}_t\} \otimes \cdots \{\emptyset, \mathbb{U}_T\}. \quad (7)$$

If (7) is not true, we say that the decision U_s *affects the information of the* tth *DM*. An information structure is quasi-classical (partially nested), if $\mathcal{J}_s \subset \mathcal{J}_t$ for every $s < t$ such that U_s affects the information of the tth DM.

4) Nonclassical: An information structure that does not belong to the above three categories is called *nonclassical*.

C. Finite Spaces Assumption

In the rest of the note, 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 $\mathbb{U}_t, t = 1, \ldots, T$, are finite sets. Further, $\mathcal{F} = 2^{\Omega}$ and $\mathcal{U}_t = 2^{\mathbb{U}_t}$, for $t = 1, \ldots, T$.

D. 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 DM. Consider a collection of variables $Z_a, Z_b, \ldots, Z_k$ defined as functions from $\Omega \times \mathbb{U}^{1:T}$ to spaces $\mathbb{Z}_a, \ldots, \mathbb{Z}_k$, respectively, i.e.,

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

We will call $Z_i = \zeta_i(\omega, U_{1:T})$ an observation and ζ_i its observation map. For realizations $\tilde{\omega}$ and $u_{1:T}$ of ω and $U_{1:T}$, respectively, $z_i = \zeta_i(\tilde{\omega}, u_{1:T})$ is the corresponding realization of the observation Z_i. We will denote by $\sigma(Z_a, \ldots, Z_k)$ the smallest sigma algebra contained in $\mathcal{F} \otimes \mathcal{U}_{1:T}$ with respect to which the observation maps $\zeta_a, \ldots, \zeta_k$ are measurable. We say that observations $Z_a, \ldots, Z_k$ generate the sigma-algebra $\mathcal{J}_t$ if

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

III. Brief Discussion on Classical Information Structures

As mentioned earlier, a sequential decomposition for team problems with classical information structure is provided by the classical dynamic program. In order to compare our results with the classical sequential decomposition, we will state the classical dynamic program in Theorem 1 below. Before doing so, we would like to make a few observations about the classical information structure.

Consider Problem 1 with a classical information structure. It is straightforward to construct observations $Z_t = \zeta_t(\omega, U_1, \ldots, U_{t-1})$, $t = 1, \ldots, T$, with Z_t taking values in a finite set $\mathbb{Z}_t$, such that $\sigma(Z_{1:t}) = \mathcal{J}_t$. Further, under a classical information structure, Problem 1 can be easily transformed into the following equivalent problem where the tth DM knows $Z_{1:t}, U_{1:t-1}$.

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

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

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

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

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

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

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

In teams with classical information structures, sequentially dominant strategies always exist under Assumption 1. In fact, the classical dynamic program provides a way of constructing these sequentially dominant strategies for teams with classical information structures. The existence of such strategies is crucially dependent on the fact that in teams with classical information structures each DM's posterior belief on ω does not depend on the choice of strategy profile [8], [15]. For DM t, we denote this strategy-independent belief by π_t. It can be defined as follows:

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

$\forall \tilde{\omega} \in \Omega$ and for $z_{1:t}, u_{1:t-1}$ such that the denominator in (9) is nonzero.

Theorem 1: For a sequential team problem in Problem 2, define value functions and strategies as follows:

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

$$V_k(z_{1:k}, u_{1:k-1}) := \min_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \ldots, u_k)|z_{1:k}, u_{1:k-1}] \quad (10)$$

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

$$g_k^*(z_{1:k}, u_{1:k-1}) := \operatorname*{argmin}_{u_k \in \mathbb{U}_k} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, u_1 \ldots, u_k)|z_{1:k}, u_{1:k-1}] \quad (11)$$

where $k < T$, $Z_{k+1} = \zeta_{k+1}(\omega, u_{1:k})$ and the expectations are with respect to DMs' strategy-independent beliefs on ω. 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. An alternative proof using strategy-independent beliefs can be found in [15]. ■

Nonclassical Information Structure and Absence of Sequentially Dominant strategies: We present a simple example to show that if the information structure is nonclassical, sequentially dominant strategies may not exist. 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, each associated with the corresponding power set sigma algebra. The cost function is $c(\omega, U_1, U_2) = (\omega + U_1 - U_2)^2$. Consider a classical information structure where DM 2 knows ω and U_1. In this case, it is easy to verify that $g_2^*(\omega, U_1) = \omega + U_1$ is a sequentially dominant strategy for DM 2. Consider now the following nonclassical information structure:

$$\mathcal{J}_1 = 2^{\Omega} \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\}, \ \mathcal{J}_2 = \{\emptyset, \Omega\} \times 2^{\mathbb{U}_1} \times \{\emptyset, \mathbb{U}_2\}. \quad (12)$$

In other words, DM 1's strategy can be any function of ω, while DM 2's strategy can be any function U_1. Consider now two possible strategy pairs

$$(i)\ U_1 = g_1^*(\omega) = \omega,\ U_2 = g_2^*(U_1) = 2U_1$$
$$(ii)\ U_1 = h_1^*(\omega) = 1 - \omega,\ U_2 = h_2^*(U_1) = 1.$$

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^* nor h_2^* is a sequentially dominant strategy 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 cannot find an optimal strategy for DM 2 without taking into account the strategy of the DM who acted before.

More generally, the absence of sequentially dominant strategies in nonclassical information structures means that we cannot find DM t's optimal strategy without taking into account the strategy of the DMs who acted before it. Thus, for nonclassical teams, we cannot expect to obtain a sequential decomposition of the kind in Theorem 1, where we could obtain DM t's strategy without considering the strategies of earlier DMs.

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 nonclassical information structures. We will use common knowledge to construct a sequential decomposition and refer to it as the *common knowledge based dynamic program.*

Recall that the information available to DM t is described by a sigma-algebra $\mathcal{J}_t \subset \mathcal{F} \otimes \mathcal{U}_{1:T}$. We define the common knowledge at time t as the intersection of sigma algebras associated with DMs t to T. That is, we define

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s. \tag{13}$$

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] and [5].

Lemma 1 (Properties of Common Knowledge):

1) Coarsening property: $\mathcal{C}_t \subset \mathcal{J}_t$ for all t.
2) Nestedness property: $\mathcal{C}_t \subset \mathcal{C}_{t+1}$ for $t < T$.
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
$$Z_t := \zeta_t(\omega, U_1, \ldots, U_{t-1}) \tag{14}$$
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
$$Y_t := \eta_t(\omega, U_1, \ldots, U_{t-1}) \tag{15}$$
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: The proof is by construction and can be found in [15]. ■

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 $\mathbb{Z}_t$ 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 $\mathbf{g} = (g_1, \ldots, g_T)$, with $g_t : \mathbb{Z}_{1:t} \times \mathbb{Y}_t \mapsto \mathbb{U}_t$ for each t, that achieves

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

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

A. 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 that time.
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 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) At each time t, the coordinator's information is described 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 ($\mathbb{U}_t$). The set of all functions from $\mathbb{Y}_t$ to $\mathbb{U}_t$ can be identified with the product space $\mathbb{U}_t^{|\mathbb{Y}_t|} = \mathbb{U}_t \times \cdots \times \mathbb{U}_t$ (where the number of terms in the product is $|\mathbb{Y}_t|$).

We use γ_t to denote the element from the set $\mathbb{U}_t^{|\mathbb{Y}_t|}$ selected by the coordinator at time t. γ_t is a tuple of size $|\mathbb{Y}_t|$. $\gamma_t(y)$ 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 DM t if DM t's private observation takes the value y. Thus, γ_t can be seen as a *prescription* to DM t 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, the decision taken by DM t can be written as follows:

$$U_t = \gamma_t(Y_t). \tag{16}$$

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

$$\psi_t : \mathbb{Z}_{1:t} \mapsto \mathbb{U}_t^{|\mathbb{Y}_t|} \tag{17}$$

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, $U_t = g_t(Z_{1:t}, Y_t),\ t = 1\dots,T$. This strategy profile induces a joint probability distribution on $\omega, U_{1:T}, Y_{1:T}, Z_{1:T}$. $\mathbb{P}^g(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T})$ denotes the probability of the realization $\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}$ under the probability distribution induced by $\mathbf{g}$.

We will now construct a strategy $\boldsymbol{\psi}$ for the coordinator using the strategy profile $\mathbf{g}$. Recall that g_t, DM t's strategy in Problem 3, 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 $\mathbb{U}_t^{|\mathbb{Y}_t|}$.

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

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 $y = 1, \dots, |\mathbb{Y}_t|$, the yth component of $\psi_t(z_{1:t})$ is given by $g_t(z_{1:t}, y)$.

The coordinator's strategy $\boldsymbol{\psi}$ constructed above induces a joint probability distribution on $\omega, U_{1:T}, Y_{1:T}, Z_{1:T}$. $\mathbb{P}^\psi(\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T})$ denotes the probability of the realization $\tilde{\omega}, u_{1:T}, y_{1:T}, z_{1:T}$ under the probability distribution induced by $\boldsymbol{\psi}$.

Lemma 2: 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, $\mathbb{E}^\psi[c(\omega, U_1, \dots, U_T)] = \mathbb{E}^g[c(\omega, U_1, \dots, U_T)]$.

Proof: The proof can be found in [15]. ■

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

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 will be the y_tth 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 $y = 1, \dots, |\mathbb{Y}_t|$, $h_t(z_{1:t}, y)$ is the yth component of $\phi_t(z_{1:t})$.

Lemma 3: 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, $\mathbb{E}^h[c(\omega, U_1, \dots, U_T)] = \mathbb{E}^\phi[c(\omega, U_1, \dots, U_T)]$.

Proof: The proof can be found in [15]. ■

Lemmas 2 and 3 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

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

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

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

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

$$Z_t = \Theta_t(\omega, \gamma_{1:t-1}), \quad t = 1, \dots, T. \tag{21}$$

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

$$c(\omega, U_{1:T}) = C(\omega, \gamma_{1:T}). \tag{22}$$

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

Problem 4: Given observations $Z_t = \Theta_t(\omega, \gamma_1, \dots, \gamma_{t-1})$ taking values in $\mathbb{Z}_t$ for $t = 1, \dots, T$, find a strategy $\boldsymbol{\psi} = (\psi_1, \dots, \psi_T)$ for the coordinator, with $\psi_t : \mathbb{Z}_{1:t} \mapsto \mathbb{U}_t^{|\mathbb{Y}_t|}$ for each t, that achieves

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

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

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

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 $k = 1, \dots, t$. For a given feasible realization $z_{1:t}, \tilde{\gamma}_{1:t-1}$, the coordinator's belief on ω is given as

$$\pi_t(\tilde{\omega}|z_{1:t}, \tilde{\gamma}_{1:t-1}) := \frac{\mathbb{P}(\tilde{\omega}) \prod_{k=1}^t \mathbb{1}_{\{\Theta_k(\tilde{\omega}, \tilde{\gamma}_{1:k-1}) = z_k\}}}{\sum_{\hat{\omega}} \left[\mathbb{P}(\hat{\omega}) \prod_{k=1}^t \mathbb{1}_{\{\Theta_k(\hat{\omega}, \tilde{\gamma}_{1:k-1}) = z_k\}}\right]}. \tag{23}$$

(23) defines the coordinator's posterior belief on ω after observing $z_{1:t}, \tilde{\gamma}_{1:t-1}$. Since the coordinator's problem has a classical information structure, this belief does not depend on the choice of coordinator's strategy.

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:

$$\begin{aligned} &V_T(z_{1:T}, \tilde{\gamma}_{1:T-1}) \\ &:= \min_{\tilde{\gamma}_T \in \mathbb{U}_T^{|\mathbb{Y}_T|}} \mathbb{E}^{\pi_T}[C(\omega, \tilde{\gamma}_1 \dots, \tilde{\gamma}_T)|z_{1:T}, \tilde{\gamma}_{1:T-1}] \\ &V_k(z_{1:k}, \tilde{\gamma}_{1:k-1}) \\ &:= \min_{\tilde{\gamma}_k \in \mathbb{U}_k^{|\mathbb{Y}_k|}} \mathbb{E}^{\pi_k}[V_{k+1}(z_{1:k}, Z_{k+1}, \tilde{\gamma}_1 \dots, \tilde{\gamma}_k)|z_{1:k}, \tilde{\gamma}_{1:k-1}] \end{aligned} \tag{24}$$

for $k = T-1, \dots, 2, 1$, 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 (23)].

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

$$\begin{aligned}
&\psi_T^*(z_{1:T}, \tilde{\gamma}_{1:T-1}) \\
&:= \underset{\tilde{\gamma}_T \in \mathbb{U}_T^{|\mathbb{Y}_T|}}{\operatorname{argmin}} \ \mathbb{E}^{\pi_T}\left[C(\omega, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_T) | z_{1:T}, \tilde{\gamma}_{1:T-1}\right] \\
&\psi_k^*(z_{1:k}, \tilde{\gamma}_{1:k-1}) \\
&:= \underset{\tilde{\gamma}_k \in \mathbb{U}_k^{|\mathbb{Y}_k|}}{\operatorname{argmin}} \ \mathbb{E}^{\pi_k}\left[V_{k+1}(z_{1:k}, Z_{k+1}, \tilde{\gamma}_1 \ldots, \tilde{\gamma}_k)\right. \\
&\quad \left.| z_{1:k}, \tilde{\gamma}_{1:k-1}\right] \qquad (25)
\end{aligned}$$

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

The dynamic program of Theorem 2 identifies an optimal strategy for the coordinator as a function of its observations and past prescriptions. We can construct an equivalent strategy ϕ for the coordinator by eliminating past prescriptions 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 using the construction of Lemma 3: 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_tth component of $\phi_t(z_{1:t})$ and we denote this by

$$h_t^*(z_{1:t}, \cdot) := \phi_t(z_{1:t}), \quad t = 1, \ldots, T. \qquad (26)$$

Because ϕ is an optimal strategy for the coordinator, Lemmas 2 and 3 imply that $\mathbf{h}^* = (h_1^*, \ldots, h_T^*)$ is an optimal strategy profile for the DMs in Problem 3.

Remark 2: It should be clear that the constructed strategy h_t^* in (26) 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.

B. Discussion

Theorem 2 provides a sequential decomposition for the coordinator's problem and, due to the equivalence established in Lemmas 2 and 3, for Problem 3 with a general (in particular, nonclassical) 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: First, 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 $\mathbb{U}_k$. Second, 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 3: 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 1 can be constructed and that the coordinator's strategies are well-defined measurable functions.

C. Example

Consider a team problem with two DMs. 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 DMs 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:

$$\mathcal{J}_1 = \sigma(\{1,2\},\{3,4\},\{5\}) \times \{\emptyset, \mathbb{U}_1\} \times \{\emptyset, \mathbb{U}_2\} \qquad (27)$$

$$\mathcal{J}_2 = \sigma(\{1,3\},\{2,4\},\{5\}) \times 2^{\mathbb{U}_1} \times \{\emptyset, \mathbb{U}_2\}. \qquad (28)$$

This information structure is nonclassical, since $\mathcal{J}_1 \not\subset \mathcal{J}_2$. As discussed before, we cannot obtain a classical dynamic program for such an information structure. For this example, the common knowledge based dynamic program can be obtained as follows:

i) The common knowledge sigma-algebras are:

$$\mathcal{C}_1 = \bigcap_{s=1}^{2} \mathcal{J}_s, \ \mathcal{C}_2 = \mathcal{J}_2. \qquad (29)$$

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 $t = 1, 2$. For our example, the following definitions will meet the requirements:

$$Z_1 = \zeta_1(\omega) := \mathbb{1}_{\{\omega \in \{5\}\}} \qquad (30)$$

$$Z_2 = \zeta_2(\omega, U_1) := \begin{cases} (1, U_1)\, if\, \omega \in \{1,3\} \\ (2, U_1)\, if\, \omega \in \{2,4\} \\ (5, U_1)\, if\, \omega \in \{5\} \end{cases} \qquad (31)$$

$$\begin{aligned} Y_1 &= \eta_1(\omega) := \mathbb{1}_{\{\omega \in \{3,4\}\}} \\ Y_2 &= \eta_2(\omega, U_1) := 0. \end{aligned} \qquad (32)$$

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 need the prescription at $t = 1$: $\gamma_1 : \{0,1\} \mapsto \mathbb{U}_1$. The space of all such prescriptions can be written as $\mathbb{U}_1 \times \mathbb{U}_1$.

iv) Based on common observations and prescriptions, we define the following strategy-independent beliefs on ω.

$$\pi_1(\tilde{\omega} | Z_1 = 1) = \begin{cases} 1\, if\, \tilde{\omega} = 5 \\ 0\, if\, \tilde{\omega} \neq 5 \end{cases} \qquad (33)$$

$$\pi_1(\tilde{\omega} | Z_1 = 0) = \begin{cases} 1/4\, if\, \tilde{\omega} \neq 5 \\ 0\, if\, \tilde{\omega} = 5. \end{cases} \qquad (34)$$

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

$$\pi_2(\tilde{\omega}|Z_1 = 1, Z_2 = (5, u_1), \tilde{\gamma}_1) := \begin{cases} 1\, if\, \tilde{\omega} = 5 \\ 0\, if\, \tilde{\omega} \neq 5. \end{cases} \tag{35}$$

For $i = 1, 2$

$$\pi_2(\tilde{\omega}|Z_1 = 0, Z_2 = (i, u_1), \tilde{\gamma}_1) \tag{36}$$

$$:= \frac{\mathbb{1}_{\{\zeta_2(\tilde{\omega}, \tilde{\gamma}_1(\eta_1(\tilde{\omega})))=(i,u_1)\}}}{\sum_{\hat{\omega}\neq 5}\left[\mathbb{1}_{\{\zeta_2(\hat{\omega}, \tilde{\gamma}_1(\eta_1(\hat{\omega})))=(i,u_1)\}}\right]} \text{ if } \tilde{\omega} \neq 5; \tag{37}$$

$\pi_2(\tilde{\omega}|Z_1 = 0, Z_2 = (i, u_1), \tilde{\gamma}_1) := 0$ if $\tilde{\omega} = 5$.

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

$$V_2(z_{1:2}, \tilde{\gamma}_1) := \min_{u_2\in\mathbb{U}_2} \mathbb{E}^{\pi_2}[c(\omega, \tilde{\gamma}_1(\eta_1(\omega)), u_2)|z_{1:2}, \gamma_1]$$

$$V_1(z_1) := \min_{\tilde{\gamma}_1\in\mathbb{U}_1\times\mathbb{U}_1} \mathbb{E}^{\pi_1}[V_2(z_1, \zeta_2(\omega, \tilde{\gamma}_1(\eta_1(\omega))), \tilde{\gamma}_1)|z_1] \tag{38}$$

where the expectations at $t = 2, 1$ are with respect to the beliefs defined in (33)–(36). Our result in Section IV shows 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 nonclassical, we can still obtain a sequential decomposition of the strategy optimization problem.

V. Comparison With Classical Dynamic Program and Witsenhausen'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 in this section.

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 $\mathcal{J}_t$:

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s = \mathcal{J}_t. \tag{39}$$

We can construct common observations as in Lemma 1 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

$$Y_t := \eta_t(\omega, U_1, \ldots, U_{t-1}) := 1. \tag{40}$$

The implication of (39) 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 a decision 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 DMs may just be the trivial sigma algebra.

$$\mathcal{C}_t := \bigcap_{s=t}^{T} \mathcal{J}_s = \{\emptyset, \Omega \times \mathbb{U}^{1:T}\}. \tag{41}$$

In this case, the common observations of Lemma 1 can be defined as constants

$$Z_t := \zeta_t(\omega, U_1, \ldots, U_{t-1}) := 1 \tag{42}$$

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 [17].

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 the 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-A, 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 1, 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-A. Such an approach for sequential team problems that uses common information among DMs 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 note:

First, 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 DMs' 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 note applies to any sequential information structure.

Second, the common information based dynamic program may be different from the common knowledge based dynamic program obtained in Section IV-A. To see why, note that the sigma-algebra 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 $\mathcal{C}_t$. 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 observation 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.

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 threeDMs. 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 DMs 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 DM has access to.

1) DM 1 knows

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

2) DM 2 knows

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

3) 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 can be taken to be $\tilde{Y}_1 = \tilde{X}_1$, $\tilde{Y}_2 = \tilde{X}_2$, 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 similar 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 nontrivial and are given as follows:

1) $\mathcal{C}_1 = \sigma(Z_1)$, where $Z_1 = \mathbb{1}_{\{\omega=5\}}$. In other words, $\mathcal{C}_1 = \sigma(\{1, 2, 3, 4\}, \{5\}) \otimes \{\emptyset, \mathbb{U}_1\} \otimes \{\emptyset, \mathbb{U}_2\}$.
2) $\mathcal{C}_2 = \sigma(\tilde{X}_2)$. That is, $\mathcal{C}_2 = \sigma(\{1, 3\}, \{2, 4\}, \{5\}) \otimes \{\emptyset, \mathbb{U}_1\} \otimes \{\emptyset, \mathbb{U}_2\}$.
3) $\mathcal{C}_3 = \sigma(\tilde{X}_1, \tilde{X}_3)$. In other words, $\mathcal{C}_3 = 2^\Omega \otimes \{\emptyset, \mathbb{U}_1\} \otimes \{\emptyset, \mathbb{U}_2\}$.

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

VII. Conclusion

We considered sequential team problems based on Witsenhausen's intrinsic model with finite probability and decision spaces. We started with the case of classical information structures and reviewed the classical 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 team problems with nonclassical information structures opens up the possibility of systematic methods for finding structural results and information states for DMs in such problems.

References

[1] M. S. Andersland and D. Teneketzis, "Information structures, causality, and nonsequential stochastic control I: Design-independent properties," *SIAM J. Control Optim.*, vol. 30, no. 6, pp. 1447–1475, 1992.
[2] R. J. Aumann, "Agreeing to disagree," *Ann. Statist.*, vol. 4, no. 6, pp. 1236–39, 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 J. Control Optim.*, vol. 38, no. 3, pp. 875–892, 2000.
[4] R. Dick, G. Hexner, and Y. Ho, "On common information in decision problems," *IEEE Trans. Inf. Theory*, vol. IT-23 no. 3, pp. 393–396, May 1977.
[5] G. Hexner and Y. Ho, "Information structure: Common and private (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-23 no. 3, pp. 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 Trans. Autom. Control*, vol. AC-17, no. 1, pp. 15–22, Feb. 1972.
[7] Y. C. Ho, M. Kastner, and E. Wong, "Teams, signaling and information theory," *IEEE Trans. Autom. Control*, vol. AC-23, no. 2, pp. 305–312, Apr. 1978.
[8] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Englewood Cliffs, NJ, USA: Prentice Hall, 1986.
[9] A. Mahajan, "Sequential Decomposition of Sequential Dynamic Teams: Applications to Real-time Communication and Networked Control Systems," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, USA, Sep. 2008.
[10] A. Mahajan, N. Martins, M. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *Proc. 51st IEEE Conf. Decis. Control*, Dec. 2012, pp. 1291–1306.
[11] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven, NY, USA: Yale Univ. Press, 1972.
[12] A. Nayyar, "Sequential Decision Making in Decentralized Systems," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, USA, 2011.
[13] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1606–1620, Jul. 2011.
[14] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.
[15] A. Nayyar and D. Teneketzis, "Common knowledge and sequential team problems," Tech. Rep. CENG-2018-02, Nov. 2018. [Online]. Available: http://ceng.usc.edu/techreports/2018/Nayyar%20CENG-2018-02.pdf
[16] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 857–881, Sep. 1962.
[17] H. Witsenhausen, "A standard form for sequential stochastic control," *Math. Syst. Theory*, vol. 7, no. 1, pp. 5–11, 1973.
[18] H. S. Witsenhausen, "On information structures, feedback and causality," *SIAM J. Control*, vol. 9, no. 2, pp. 149–160, May 1971.
[19] H. S. Witsenhausen, "The intrinsic model for discrete stochastic control: Some open problems," in *Control Theory, Numerical Methods and Computer System Modelling*, vol. 107, A. Bensoussan and J. L. Lions, Eds., Lecture Notes in Economics and Mathematical Systems, New York, NY, USA: Springer-Verlag, 1975, pp. 322–335.
[20] H. S. Witsenhausen, "Equivalent stochastic control problems," *Math. Control Signals Syst.*, vol. 1, pp. 3–11, 1988.
[21] Y. C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644–654, Jun. 1980.
[22] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA, USA: Birkhäuser, 2013.