2010 American Control Conference
Marriott Waterfront, Baltimore, MD, USA
June 30-July 02, 2010

Measure and cost dependent properties of information structures

Aditya Mahajan and Serdar Yüksel

Abstract— **Information structures capture how and when information is shared between the agents of a decentralized system. Traditionally, information structures are classified based on the data available at the agents. No assumptions on the probability measure or the cost function are imposed. This classification has helped in clarifying the conceptual difficulties of decentralized control. However, from a practical point of view, the classification based only on the data is restrictive as illustrated by the recent results on stochastically nested information structures. In this paper, we generalize the main idea of stochastically nested information structures to define extensions of classical and quasiclassical information structures that depend on the probability measure and the cost function. We show that the optimality results of classical and quasiclassical information structures extend to their stochastically nested counterparts.**

I. Introduction

In a decentralized multi-agent control system, the notion of *who knows what and when* is captured by information structures. Broadly speaking, information structures are classical or non-classical. If every control agent knows all the data available to agents that acted before her, the information structure is classical; otherwise, it is non-classical. Classical information structure is equivalent to a centralized system, but surprisingly, *a centralized system need not have a classical information structure*. Similar dichotomy holds for subclasses for non-classical information structure that are well understood—partially nested [1], [2], delayed sharing [3]–[5], and shared observations [6], to name a few. Since systems with identical information structures do not have similar design difficulties, using information structures to classify decentralized multi-agent control systems is fallacious. In this paper, we present generalizations of information structures that classify decentralized multi-agent systems in a more consistent manner.

First, we present an example that shows the limitations of information structures.

Example 1 (A control station with zero memory): Consider a discrete time system with a plant and a control station. At time t, let X_t denote the state of the plant, U_t the action of the control agent, W_t the process noise, and N_t the observation noise. All variables take values in compact sets. The process and the observation noises are independent across time, independent of each other, and independent of the initial state of the plant. The plant evolves as $X_{t+1} = f(X_t, U_t, W_t)$. The control station observes $Y_t = h(X_t, N_t)$ and chooses a control action $U_t = g_t(Y_t)$. (Notice that the control station has no memory). The design objective is to choose a *control strategy* $g = (g_1, \dots, g_T)$ to minimize the expected total cost $\mathbb{E}^g\big\{\sum_{t=1}^{T} c(X_t, U_t)\big\}$.

Consider two variations of this system:

V1) The observations are noiseless, *i.e.*, $Y_t = X_t$;
V2) The observations are noisy *i.e.*, the observation function h is not invertible in X.

For both variations, the information structure is non-classical because the control agent at time t does not know the observations and actions of earlier agents. Nonetheless, these variations are drastically different. V1 is a MDP (Markov decision process) and hence centralized while V2 is decentralized. Thus, information structures do not completely characterize decentralized systems.

Information structures specify only the data available to different agents, not the *usefulness* of that data. In this paper, we define generalizations of information structures, which depend on the probability measure and cost function, that capture the usefulness of information. We call these generalizations the *P-generalization* of an information structure; for any information structure, its P-generalization is denoted by adding a "P-" prefix. To illustrate the main idea of P-generalization, consider the P-classical information structure. In a P-classical information structure all control agent do not know the data available to agents that acted before them—thus, the information structure is not classical—but the missing data is redundant, *i.e.*, even if all agents had access to the smallest additional data that will make the information structure classical, the system will have the same performance. The P-generalization of an information structure is similar in spirit to stochastically-nested information structures studied in [7]. In this paper, we study P-classical and P-quasiclassical information structures and show that the solution techniques for classical and quasiclassical information structures can also be used in their P-generalizations. Similar ideas can be used to describe P-generalizations of delayed sharing and shared observations information structures. The results of this paper suggest that ***the solution technique for any information structure is also applicable to its P-generalization***. Furthermore, if we are designing the communication infrastructure of a system to achieve a particular information structure, we only need to share data to achieve the P-generalization of that information structure. This data can be significantly less than the data needed to achieve the original information structure, as shown for belief sharing information structure [7].

A. Mahajan is with the Dept. of Electrical Engineering, Yale University, New Haven, CT USA 06520. Email: aditya.mahajan@yale.edu

S. Yüksel is with the Dept. of Mathematics and Statistics, Queen's University, Kingston, ON, Canada K7L 3N6. Email: yuksel@mast.queensu.ca

The work of S. Yüksel was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Notation: Uppercase letters X, Y, Z, etc. represent random variables; the corresponding lowercase letters x, y, z, etc. represent their realizations. Blackboard bold letters $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{Z}$, etc. represent state spaces and calligraphic letters $\mathscr{F}$, $\mathscr{G}$, $\mathscr{H}$, etc. represent σ-fields. Given two σ-fields $\mathscr{F}_1$ and $\mathscr{F}_2$, $\mathscr{F}_1 \otimes \mathscr{F}_2$ denotes the product σ-field and $\mathscr{F}_1 \vee \mathscr{F}_2$ denotes the smallest σ-field containing $\mathscr{F}_1 \cup \mathscr{F}_2$.

$\mathbb{E}\{\cdot\}$ denotes expectation of a random variable. For a set A and an element a, $\mathbb{1}_A(a)$ is 1 if $a \in A$ and 0 otherwise. For two sets A and B, $\mathbb{1}_A(B) = \prod_{b \in B} \mathbb{1}_A(b)$, or equivalently, $\mathbb{1}_A(B)$ is 1 if $B \subset A$ and 0 otherwise.

For any natural number n, $[n]$ denotes the set $\{1, \dots, n\}$. For any subset A of natural numbers, X_A denotes the vector $(X_a : a \in A)$, $\mathbb{X}_A$ denotes the product space $\prod_{a \in A} X_a$ and $\mathscr{F}_A$ denotes product σ-field $\bigotimes_{a \in A} \mathscr{F}_a$.

II. Intrinsic model and information structures

Decentralized multi-agent control systems are typically represented in two ways:

1) *State space model:* The system consists of a plant and multiple control stations that act across multiple time stages. At each time stage, the observation of a control station depends on the *system state*. An instantaneous cost, which depends on the system state and control actions of all agents, is incurred at each stage. The objective is to choose a control policy to minimize the total expected cost over a time horizon.
2) *Intrinsic model:* The system consists of multiple control agents. Each agent acts once. The observations of an agent depends on an intrinsic event and control actions of some other agents. A cost which depends on the intrinsic event and actions of all agents is incurred. The objective is to choose the control laws of all agents to minimize the expected total cost.

These two models are equivalent. The state space model is more convenient to describe specific applications. The intrinsic model, on the other hand, is more convenient to capture the functional dependence between the control actions of the agents. For that reason, we follow the intrinsic model in this paper.

A. The intrinsic model

Consider a system with N agents. Let $(\Omega, \mathscr{B}, P)$ be the probability space for actions ω of nature, $(\mathbb{Y}_n, \mathscr{F}_n)$ and $(\mathbb{U}_n, \mathscr{G}_n)$ the measurable spaces for observations Y_n and actions U_n of agent n, $n = 1, \dots, N$. $\mathbb{U}_n$ and $\mathbb{Y}_n$ are compact metric spaces, and $\mathscr{B}$, $\mathscr{F}_n$, $\mathscr{G}_n$ are countably generated σ-fields that contain all singletons.

The observation of agent n is given by

$$Y_n = f_n(\omega, U_{D_n})$$

where $D_n \subset [n-1]$ is the set of agents that affect the observations of agent n. Define a binary relation $\prec$ on $[N]$ by $m \prec n$ if $m \in D_n$. Let $\prec^*$ denote the transitive closure of $\prec$ and $D_n^* := \{m \in [N] \setminus \{n\} : m \prec^* n\}$. Then D_n^* is the set of agents that can *signal* to agent n, *i.e.*, the set of agents whose actions affect, either directly or indirectly, the observations of agent n.

Agent n takes a control action

$$U_n = g_n(Y_n)$$

where g_n is chosen from the set G_n of functions from $\mathbb{Y}_n$ to $\mathbb{U}_n$ that are $\mathscr{G}_n/\mathscr{F}_n$ measurable.

The system incurs a cost

$$\sum_{k=1}^{K} c_k(\omega, U_{C_k}) \tag{1}$$

where C_k is a subset of $[N]$ and c_k is a bounded real-valued function, $k = 1, \dots, K$. In the sequel, we consider two following partial sums of the total cost.

$$Q_n(\omega, U_{[N]}) := \sum_{k=1}^{K} c_k(\omega, U_{C_k}) \mathbb{1}_{C_k}(n) \tag{2}$$

$$R_n(\omega, U_{[N]}) := \sum_{k=1}^{K} c_k(\omega, U_{C_k})(1 - \mathbb{1}_{C_k}([n-1])) \tag{3}$$

Any choice of the *policy* $g = (g_1, \dots, g_N)$, chosen from the set $G = G_1 \times \cdots \times G_N$, makes the variables $U_{[N]}$ measurable on $(\Omega, \mathscr{B}, P)$. The objective is to choose a policy g to minimize the expected value of (1) where the expectation is with respect to the joint measure on $U_{[N]}$ corresponding to g.

B. Information fields and information structures

For $A \subset [N]$, let $\mathbb{H}_A := \Omega \times \prod_{n \in A} \mathbb{U}_n$ and $\mathbb{H} = \mathbb{H}_{[N]}$. For any σ-field $\mathscr{C}$ on $\mathbb{H}_A$, let $\langle \mathscr{C} \rangle$ denote the cylindrical extension of $\mathscr{C}$ on $\mathbb{H}$. Let $\mathscr{H}_A := \left\langle \mathscr{B} \otimes \left(\bigotimes_{n \in A} \mathscr{G}_n \right) \right\rangle$ and $\mathscr{H} = \mathscr{H}_{[N]}$.

The observations Y_n of agent n, $n = 1, \dots, N$, tell her something about the actions of nature and of agents in D_n^*. Thus, the information available to agent n can be characterized by a subfield $\mathscr{J}_n$ of $\mathscr{H}_{D_n^*}$. This subfield is called the *information field* of agent n. The collection $\{\mathscr{J}_n\}_{n=1}^N$ is called the *information structure* of the system.

There is one-to-one correspondence between the model involving ω, Y_n, and U_n, $n = 1, \dots, N$, and the model involving $\mathscr{B}$, $\mathscr{J}_n$, and $\mathscr{G}_n$, $n = 1, \dots, N$. We will use these two models interchangeably.

III. Classification of information structures

We briefly restate the classification specified in [8].

A system is *static* if $\mathscr{J}_n \subset \mathscr{H}_\emptyset$ for all $n = 1, \dots, N$, i.e., the information of all agents depends only on ω; otherwise, the system is *dynamic*.

A. Classical and quasiclassical information structures

A system is ***classical*** if each agent knows the information available to the agents that acted before her. Thus, in a classical system, agent n, $n = 1, \dots, N$, can deduce $Y_{[n-1]}$ from Y_n (without knowing the policy g). Equivalently, in a classical system $\mathscr{J}_{n-1} \subset \mathscr{J}_n$ for $n = 2, \dots, N$. A classical system is called ***strictly classical*** if each agent also knows the control action of the agents that acted before her.

Thus, in a strictly classical system, agent n, $n = 1, \dots, N$, can deduce $Y_{[n-1]}$ and $U_{[n-1]}$ from Y_n (without knowing the policy g). Equivalently, in a strictly classical system, $\mathscr{J}_{n-1} \vee \langle \mathscr{G}_{n-1} \rangle \subset \mathscr{J}_n$, $n = 2, \dots, N$. Centralized stochastic control problems like MDP (Markov decision process) and POMDP (partially observable Markov decision process) are strictly classical.

A system is ***quasiclassical*** if each agent knows the information available to agents that influence, either directly or indirectly, her observations. Thus, in a quasiclassical system, agent n, $n = 1, \dots, N$ can deduce $Y_{D_n^*}$ from Y_n (without knowing the policy g). Equivalently, in a quasiclassical system, for $n = 1, \dots, N$ and $m \in D_n^*$, $\mathscr{J}_m \subset \mathscr{J}_n$. A quasiclassical system is called ***strictly quasiclassical*** if each agent also knows the control actions of the agents that influence, either directly or indirectly, her observations. Thus, in a strictly quasiclassical system, agent n, $n = 1, \dots, N$, can deduce $(Y_{D_n^*}, U_{D_n^*})$ from Y_n (without knowing the policy g). Equivalently, in a quasiclassical system, for $n = 1, \dots, N$ and $m \in D_n^*$, $\mathscr{J}_m \vee \langle \mathscr{G}_m \rangle \subset \mathscr{J}_n$. Systems with one-step delay information structure [3], [4] and partially nested information structure [1] are strictly quasiclassical.

B. Expansions of information structures

Given a classical system, its ***strict expansion*** is the strictly classical system obtained by replacing Y_n by $(Y_n, U_{[n-1]})$, or equivalently, by replacing $\mathscr{J}_n$ by $\mathscr{J}_n \vee \left(\bigvee_{m<n} \langle G_m \rangle\right)$, $n = 1, \dots, N$. Similarly, given a quasiclassical system, its strict expansion is the strictly quasiclassical system obtained by replacing Y_n by $(Y_n, U_{D_n^*})$, or equivalently replacing $\mathscr{J}_n$ by $\mathscr{J}_n \vee \left(\bigvee_{m \in D_n^*} \langle \mathscr{G}_m \rangle\right)$, $n = 1, \dots, N$.

By successive substitution, any policy achievable in the strict expansion of a classical or quasiclassical system is achievable in the system itself, and vice versa.

Given any system, its ***classical expansion*** is the strictly classical system obtained by replacing Y_n by $(Y_{[n]}, U_{[n-1]})$, or equivalently, by replacing $\mathscr{J}_n$ by $\mathscr{J}_n \vee \left(\bigvee_{m<n} (\mathscr{J}_m \vee \langle \mathscr{G}_m \rangle)\right)$, $n = 1, \dots, N$. Similarly, given any system, its ***quasiclassical expansion*** is the strictly quasiclassical system obtained by replacing Y_n by $(Y_n, Y_{D_n^*}, U_{D_n^*})$, or equivalently replacing $\mathscr{J}_n$ by $\mathscr{J}_n \vee \left(\bigvee_{m \in D_n^*} (\mathscr{J}_m \vee \langle \mathscr{G}_m \rangle)\right)$, $n = 1, \dots, N$.

Any policy achievable in the original system is also achievable in its quasiclassical and classical expansion. Moreover, the classical expansion of the quasiclassical expansion of any system is the same as its classical expansion. Thus, the minimum expected cost of any system is no smaller than the minimum expected cost of its quasiclassical expansion, which in turn is no smaller than the minimum expected cost of its classical expansion.

C. P-classical and P-quasiclassical information structures

The above definitions of information structures do not depend on the form of the probability measure or the cost function. Such a measure-and-cost-independent classification of information structures has been extremely useful in highlighting the main conceptual difficulties in decentralized stochastic control problems. Nevertheless, in models like classical and quasiclassical information structures where these conceptual difficulties have been resolved, introducing a measure-and-cost-dependent identification of information structures is useful to establish a solution framework for a larger class of problems.

This idea has been demonstrated in [7], which defined a system that was not classical but could be reduced to one. This idea was also used in [2] to define an auxiliary problem and identify when the solution to the auxiliary problem is optimal for the original problem. However, the results in [2], [7] were restricted to either LQG (linear quadratic Gaussian) or specific state-space models, which limits their application. In this paper, we present the idea of identifying measure and cost dependent generalizations of information structures in its full generality. This extension is based on the notions of state and information state for decentralized systems [9], [10]. This generalization of classical and quasiclassical information structures is given below.

Definition 1 (P-classical information structure): A system is *P-classical* if for every $g \in G$ and all $n = 1, \dots, N$,

$$\mathbb{E}^g\left\{Q_n(\omega, U_{[N]}) \,\middle|\, Y_n, U_n\right\} = \mathbb{E}^g\left\{Q_n(\omega, U_{[N]}) \,\middle|\, Y_{[n]}, U_{[n]}\right\}. \quad (4)$$

The system of Example 1 and stochastically nested information structure [7] with two control stations are P-classical.

Definition 2 (P-quasiclassical information structure): A system is *P-quasiclassical* if for every $g \in G$ and all $n = 1, \dots, N$,

$$\mathbb{E}^g\left\{Q_n(\omega, U_{[N]}) \,\middle|\, Y_n, U_n\right\} = \mathbb{E}^g\left\{Q_n(\omega, U_{[N]}) \,\middle|\, Y_n, Y_{D_n^*}, U_n, U_{D_n^*}\right\}. \quad (5)$$

Stochastically nested information structures [7] with more than two control stations are P-quasiclassical. A similar idea was used in [2] to define an auxiliary problem and identify when the solution of the auxiliary problem is optimal for the original.

Thus, in a P-classical information structure, an agent does not necessarily know all the data that was available to the agents that act before her; nonetheless, this unknown data is redundant for the purpose of performance evaluation. Similarly, in a P-quasiclassical information structure, an agent does not necessarily know all the data available to the agents that affect her observations; nonetheless, this unknown data is redundant for the purpose of performance evaluation.

Similar idea can be used to define P-generalizations of delayed sharing pattern [3]–[5] and shared observations [6]. However, defining such an information structure needs a more specific model than what we have assumed in this paper.

In the remainder of this paper we show that an optimal solution of P-classical or P-quasiclassical information structures can be obtained from the solution of the classical or quasiclassical expansion of these systems. But first, we give an example of P-quasiclassical systems.

D. Example of P-quasiclassical system

Consider a system with 6 agents, and $\omega = (\omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6)$ where ω_i and U_i take values in $\mathbb{R}$. The observation of the agents are given by

$$Y_1 = \omega_0 + \omega_1 + \omega_3 + \omega_5, \quad Y_2 = \omega_0 + \omega_2 + \omega_4 + \omega_6,$$
$$Y_3 = (\omega_0 + \omega_3 + \omega_5, U_1), \quad Y_4 = (\omega_0 + \omega_4 + \omega_6, U_2),$$
$$Y_5 = (\omega_0 + \omega_5, U_1, U_3), \qquad Y_6 = (\omega_0 + \omega_6, U_2, U_4).$$

The cost function is given by

$$c(\omega, U_{[6]}) = U_{[6]}^{\mathsf{T}} \mathbf{R} \omega_0 + U_{[6]}^{\mathsf{T}} \mathbf{Q} U_{[6]}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are matrices of appropriate dimensions and $\mathbf{Q}$ is positive semi-definite. Let $Q_{i,j}$ denote the (i, j) component of $\mathbf{Q}$ and R_i denote the i component of $\mathbf{R}$. Then, the cost can also be written as

$$c(\omega, U_{[6]}) = \sum_{i=1}^{6} \Big(\omega_0 R_i U_i + Q_{i,i} U_i^2 \sum_{j=1, j \neq i}^{6} (Q_{i,j} + Q_{j,i}) U_i U_j \Big)$$

For this system

$$D_1^* = \emptyset, \qquad D_3^* = \{1\}, \qquad D_5^* = \{1, 3\},$$
$$D_2^* = \emptyset, \qquad D_4^* = \{2\}, \qquad D_6^* = \{2, 4\}.$$

The corresponding observations are not nested. Thus, the system is not quasiclassical or classical. However, it is easy to check that if the probability measure on ω is a product measure on $(\omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6)$, then the above system is P-quasiclassical.

IV. Optimal policy for strictly classical systems

An optimal (or arbitrarily close to optimal) policy of the system presented in Section II can always be obtained by a brute force search over all policies. However, a more systematic solution approach is also possible. As shown in [11], the one shot optimization problem of choosing g in G can be broken into N nested subproblems for choosing g_n in G_n. This process is called *sequential decomposition*.

When the system is classical, this decomposition simplifies. Each subproblem of choosing g_n in G_n further breaks down into subproblems for choosing $u_n = g_n(y_n)$ for each realization y_n of Y_n. Thus, in a classical system, an optimal policy g can be obtained by solving a sequence of nested *parametric* optimization subproblems (each subproblem chooses a parameter u_n). In contrast, in a general system, we need to solve nested *functional* optimization subproblems (each subproblem chooses the function g_n). Such a parametric sequential decomposition is more popularly called *dynamic programming*.

The main idea of dynamic programming is the following. From each agent's point of view, the total cost can be split into two parts: "past" cost and "future" cost. Cost terms where $C_k \subset [n-1]$ are past costs; others are future costs. For any policy g, the *future cost to go* for an agent is given by

$$J_n^g(y_n) = \mathbb{E}^g \Big\{ R_n(\omega, U_{[N]}) \,\Big|\, Y_n = y_n, U_n = g_n(y_n) \Big\}. \quad (6)$$

which can be written as

$$\mathbb{E}^g \Big\{ R_n(\omega, U_{[N]}) - R_{n+1}(\omega, U_{[N]}) + R_{n+1}(\omega, U_{[N]}) \,\Big|\, Y_n = y_n, U_n = g_n(y_n) \Big\}. \quad (7)$$

Since the system is classical, $\mathscr{J}_{n+1} \supset \mathscr{J}_n \vee \langle \mathscr{G}_n \rangle$, and the last term can be written as

$$\begin{aligned} &\mathbb{E}^g \Big\{ R_{n+1}(\omega, U_{[N]}) \,\Big|\, Y_n = y_n, U_n = g_n(y_n) \Big\} \\ &= \mathbb{E}^g \Big\{ \mathbb{E}^g \{ R_{n+1}(\omega, U_{[N]}) \,|\, Y_{n+1}, U_{n+1} = g_{n+1}(Y_{n+1}) \} \,\Big|\, Y_n = y_n, U_n = g_n(y_n) \Big\} \\ &= \mathbb{E}^g \Big\{ J_{n+1}^g(Y_{n+1}) \,\big|\, Y_n = y_n, U_n = g_n(y_n) \Big\}. \end{aligned} \quad (8)$$

Substitute (8) in (7) and write the cost to go function in a recursive form as

$$J_n^g(y_n) = \mathbb{E}^g \Big\{ \sum_{k=1}^{K} c_k(\omega, U_{C_k}) (\mathbb{1}_{C_k}([n]) - \mathbb{1}_{C_k}([n-1])) + J_{n+1}^g(Y_{n+1}) \,\Big|\, Y_n = y_n, U_n = g_n(y_n) \Big\}. \quad (9)$$

Dynamic programming optimizes these cost-to-go functions by backward induction. For that matter, define

$$V_N(y_N) = \inf_{u_N \in \mathbb{U}_N} \mathbb{E} \Big\{ \sum_{k=1}^{K} c_k(\omega, U_{C_k}) (\mathbb{1}_{[N]}(C_k) - \mathbb{1}_{[N-1]}(C_k)) \,\Big|\, Y_N = y_N, U_N = u_N \Big\} \quad (10)$$

and for $n = N-1, \dots, 1$, define

$$V_n(y_n) = \inf_{u_n \in \mathbb{U}_n} \mathbb{E} \Big\{ \sum_{k=1}^{K} c_k(\omega, U_{C_k}) (\mathbb{1}_{C_k}([n]) - \mathbb{1}_{C_k}([n-1])) + V_{n+1}(Y_{n+1}) \,\Big|\, Y_n = y_n, U_n = u_n \Big\}. \quad (11)$$

By induction, we can show that for any $g \in G$

$$V_n(y_n) \leq J_n^g(y_n)$$

with equality only if g is an optimal policy. This suggests a method to find an optimal policy. Recursively solve V_n defined by (10) and (11) and choose $g_n(y_n)$ to be the arg inf of the RHS of $V_n(y_n)$. Since $\mathbb{U}_n$ is compact and c_k is bounded, the arg inf lies in $\mathbb{U}_n$ and is a valid control action. The equations (10) and (11) are called the dynamic programming equations.

A critical step in the above decomposition is (8), which follows from the law of iterated expectations. However, for non-classical systems, $\mathscr{J}_n \not\subset \mathscr{J}_{n+1}$ so the argument in (8) breaks down, and a parametric decomposition of the form (10) and (11) is not possible. Nonetheless, when the system is P-classical, an optimal policy can be identified from the solution of (10) and (11) corresponding to its classical extension. This is demonstrated in the next section. First, we prove a result about the classical expansion of a system.

Consider any system (S) and its classical expansion (SC). Let Y_n' be the observations and g_n' be the control laws of agent n, $n = 1, \dots, N$, in (SC). Let G'_n be the set of all control laws g'_n and G' set of all policies in (SC).

Definition 3: A policy g' of (SC) is *achievable at agent n in (S)* if g'_n is $\mathscr{G}_n/\mathscr{F}^*_n$ measurable, where $\mathscr{F}^*_n$ is the cylindrical extension of $\mathscr{F}_n$ to $\mathbb{Y}_{[n]} \times U_{[n-1]}$. Another way of stating this property is that for any $(y^{(1)}_{[n]}, u^{(1)}_{[n-1]})$ and $(y^{(2)}_{[n]}, u^{(2)}_{[n-1]})$ in $\mathbb{Y}_{[n]} \times U_{[n-1]}$ such that for $y^{(1)}_n = y^{(2)}_n$, we have $g'_n(y^{(1)}_{[n]}, u^{(1)}_{[n-1]}) = g'_n(y^{(2)}_{[n]}, u^{(2)}_{[n-1]})$. A policy is *achievable in (S)* if it is achievable at all agents in (S).

Proposition 1: Given any policy g' of (SC) that is achievable in (S), we can find a *corresponding* policy g of (S) such that for any realization of ω, using g in (S) gives the same realization of the random variables as using g' is (SC).

Proof: Define policy g as follows. For any $(y_{[n]}, u_{[n-1]})$ in $\mathbb{Y}_{[n]} \times U_{[n-1]}$,

$$g_n(y_n) = g'_n(y_n, y_{[n-1]}, u_{[n-1]}). \tag{12}$$

The values of $(y_{[n-1]}, u_{[n-1]})$ do not matter because $g'_n(y_n, y_{[n-1]}, u_{[n-1]})$ takes the same value for all $(y_{[n-1]}, u_{[n-1]})$. Proceeding sequentially, we can show that for any realization of ω, the using g in (S) leads to the same realization of $U_{[N]}$ as using g' in (SC). ■

V. Optimal Policy for P-Classical Systems

For a P-classical system (S), consider its classical expansion (SC). According to Proposition 1, if an optimal policy of (SC) is achievable in (S), then the corresponding policy is optimal in (S).

The system (SC) is classical, so the dynamic programming equations (10) and (11) give an optimal policy g'. *However, such a policy g' need not be achievable in (S)*,[1]. Nonetheless, it is a stepping stone to find an optimal policy g'^* for (SC) that is achievable in (S).

Theorem 1: If (S) is P-classical, there exists an optimal policy g'^* of (SC) that is achievable in (S). The policy corresponding to g'^* is optimal for (S).

Proof: Let g' be an optimal policy for (SC). Pick an agent n. Fix the policy g'_{-n} of all other agents. We will show that we can find a policy (g'^*_n, g'_{-n}) which is optimal for (SC) and achievable for (S) at agent n. Repeating this process for all agents will give us a policy that is achievable at all agents in (S).

In (SC), consider the "best response" policy of agent n to g'_{-n}. Since, g' is optimal g'_n is a best response. If g'_n is achievable at agent n in (S), we are done. Otherwise, choose

$$\begin{aligned} u^*_n &= g^*_n(y_{[n]}, u_{[n-1]}) \\ &= \arg\inf_{u_n \in \mathbb{U}_n} \mathbb{E}^{g'_{-n}}\{R_n(\omega, U_{[N]})) \,|\, Y_{[n]} = y_{[n]}, U_{[n]} = u_{[n]}\} \\ &\overset{(a)}{=} \arg\inf_{u_n \in \mathbb{U}_n} \mathbb{E}^{g'_{-n}}\{Q_n(\omega, U_{[N]})) \,|\, Y_{[n]} = y_{[n]}, U_{[n]} = u_{[n]}\} \\ &\overset{(b)}{=} \arg\inf_{u_n \in \mathbb{U}_n} \mathbb{E}^{g'_{-n}}\{Q_n(\omega, U_{[N]})) \,|\, Y_n = y_n, U_n = u_n\} \end{aligned} \tag{13}$$

where (a) is true because $R_n(\omega, U_{[N]}) - Q_n(\omega, U_{[N]})$ does not depend of U_n, and (b) is true because (S) is P-classical. The expectation in the first two equations do not depend on g' due to policy independence of conditional expectation [12]. Relation (13) shows that g^*_n is achievable at agent n in (S). This completes the proof of the result. ■

The above result suggests the following methodology to find an optimal design for a P-classical system. First, find consider its classical expansion (SC), and find an optimal policy g' for (SC) using dynamic programming. Then, one by one, find the best response of all agents in (SC) to remaining components of g'. This resultant policy g'^* is achievable in the original system. Therefore, the policy corresponding to g'^* is optimal in the original system.

VI. Optimal Policy for Strictly Quasiclassical Systems

Quasiclassical systems are not necessarily classical. Hence, the dynamic program of Section IV is not applicable to them. In this section, we present coupled dynamic programs to find person-by-person optimal solutions for quasiclassical systems. Our approach is based on the approach of [13] for the decentralized Wald problem.

Coupled dynamic programs are obtained by restricting attention to classical subsystem and finding optimal policy for such subsystems by dynamic programming.

A subset $A := \{\alpha_1, \dots, \alpha_m\}$ of agents forms a *classical subsystem* if $\alpha_i \prec \alpha_{i+1}$ (or, equivalently, $\alpha_i \in D_{\alpha_{i+1}}$), $i = 1, \dots, m-1$.

Given a classical subsystem A, a policy g can be partitioned into two parts, g_A that depends on control laws of agents in A and g_{-A} that depends on control laws of agents not in A. For a fixed g_{-A}, the system is classical and the "best response" g_A can be determined by dynamic programming. This suggests the following process to obtain a person-by-person optimal solution for the entire system.

Partition the set $[N]$ of agents into mutually disjoint and collectively exhaustive subsets $A_1, \dots, A_M$[2] such that A_m is a classical subsystem, $m = 1, \dots, M$. Orthogonal search over these subsystems gives a person-by-person optimal policy. Such a search proceeds as follows. Arbitrarily fix the control policies of all subsystems. Now pick one subsystem,

[1]The definition of P-classical is not strong enough to ensure that g' is achievable in (S). If, however, we assume that

$$\mathbb{E}^g\Big\{R_n(\omega, U_{[N]}) \,\Big|\, Y_n, U_n\Big\} = \mathbb{E}^g\Big\{R_n(\omega, U_{[N]}) \,\Big|\, Y_{[n]}, U_{[n]}\Big\},$$

then g' obtained by solving (10) and (11) is also achievable in (S). An example where the above condition holds is the stochastically nested information structure with two control stations considered in [7].

[2]Such a partitioning is different from the i-partitions and the s-partitions considered in [14]. The i- and s-partitions break the set of agents into groups such that the agents within a group do not interact directly, while we break the set of agents into groups such that the groups have classical information structure.

say A_m, and determine the best response g_{A_m} assuming g_{-A_m} is fixed. Subsystem A_m uses this best response policy in the future. Now pick another subsystem, say $A_{m'}$, $m' \neq m$, and determine $g_{A_{m'}}$ assuming $g_{-A_{m'}}$ is fixed. Subsystem $A_{m'}$ uses this best response policy in the future. Continue in this manner by cyclically changing the policy of the subsystems one-by-one. If this algorithm converges, the resultant policy is person-by-person optimal. If the cost is convex in policy space, then a person-by-person optimal policy is also globally optimal.

The best response of each subsystem is determined by a dynamic program. Thus, effectively we have M dynamic programs, one for each subsystem, and we are cyclically solving them one by one. These dynamic programs are coupled through the observations and the cost.

A critical step in the above decomposition is breaking the system into a collection of classical subsystems. If the system is not quasiclassical, then such a decomposition is not possible. Nonetheless, if the system is P-quasiclassical, a person-by-person-optimal policy can be identified in a similar manner.

VII. Optimal policy of strictly P-quasiclassical systems

A subset $A := \{\alpha_1, \dots, \alpha_m\}$ of agents is *a P-classical subsystems* if for all $i \in [m]$

$$\mathbb{E}\{Q_{\alpha_i}(\omega, U_{[N]}) \,|\, Y_{\alpha_i}, U_{\alpha_i}\} = \mathbb{E}\{Q_{\alpha_i}(\omega, U_{[N]}) \,|\, Y_{\alpha_{[i]}}, U_{\alpha_{[i]}}\}$$

For a P-quasiclassical system (S) consider its quasiclassical expansion (SQ). Partition the set $[N]$ of agents into mutually disjoint and collectively exhaustive subsets $A_1, \dots, A_M$ such that A_m is classical subsystem of (SQ), $m = 1, \dots, M$. Then, by construction, A_m is a P-classical subsystem of (S). Follow the orthogonal search procedure of Section VI. At each step, pick a subsystem A_m. Fix g_{-A_m}. A_m is P-classical, so determine the best response g_{A_m} using the method of Section V. Proceed as in Section VI to find a person-by-person optimal policy.

VIII. Concluding Remarks

In this paper, we introduced P-classical and P-quasiclassical information structures that are characterized by σ-fields and the underlying probability measure and cost function. These information structures are inspired by stochastically nested information structures and the notion of state in decentralized system. Systems with such information structures can achieve the same performance as their classical and quasiclassical extensions. Thus, all for practical purposes, P-classical and P-quasiclassical information structures are the same as classical and quasiclassical information structures.

References

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

[2] K.-C. Chu, "Team decision theory and information structures in optimal control problems–Part II," *Automatic Control, IEEE Transactions on*, vol. 17, no. 1, pp. 22–28, 1972.

[3] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[4] P. Varaiya and J. Walrand, "On delayed sharing patterns," *IEEE Trans. Autom. Control*, vol. 23, no. 3, pp. 443–445, 1978.

[5] A. Nayyar, A. Mahajan, and D. Teneketzis, "A separation result for delayed sharing information structures," in *the proceedings of American control conference*, 2010.

[6] A. Mahajan, A. Nayyar, and D. Teneketzis, "Identifying tractable decentralized control problems on the basis of information structures," in *proceedings of the 46th Allerton conference on communication, control and computation*, Sep. 2008, pp. 1440–1449.

[7] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Autom. Control*, pp. 2773–2786, Dec. 2009.

[8] H. S. Witsenhausen, "The intrinsic model for discrete stochastic control: Some open problems," in *Control Theory, Numerical Methods and Computer System Modelling*, ser. Lecture Notes in Economics and Mathematical Systems, A. Bensoussan and J. L. Lions, Eds. Springer Verlag, 1975, vol. 107, pp. 322–335.

[9] ——, "Some remarks on the concept of state," in *Directions in Large-Scale Systems*, Y. C. Ho and S. K. Mitter, Eds. Plenum, 1976, pp. 69–75.

[10] A. Mahajan, "Sequential decomposition of sequential dynamic teams: applications to real-time communication and networked control systems," Ph.D. dissertation, University of Michigan, Ann Arbor, MI, 2008.

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

[12] ——, "On policy independence of conditional expectation," *Information and Control*, vol. 28, pp. 65–75, 1975.

[13] D. Teneketzis and Y. Ho, "The decentralized Wald problem," *Information and Computation (formerly Information and Control)*, vol. 73, no. 1, pp. 23–44, Apr. 1987.

[14] T. Yoshikawa, "Decomposition of dynamic team decision problems," *IEEE Trans. Autom. Control*, vol. 23, no. 4, pp. 627–632, Aug. 1978.