Stochastic Nestedness and Information Analysis of Tractability in Decentralized Control

Serdar Yüksel[1]

Abstract— **Communication requirements for nestedness conditions require exchange of very large data noiselessly, hence these assumptions are generally impractical. In this paper, we present a weaker notion of nestedness, which we term as stochastic nestedness. Stochastic nestedness is characterized with a sequence of Markov chain conditions. It is shown that if the information structure of two decision makers satisfy a stochastically nested structure, then the optimization admits a dynamic programming recursion and the optimization is tractable; and in particular for the LQG problems, the team optimal solution is linear, despite the lack of deterministic nestedness or partial nestedness. It is also shown that the common state required need not be consisting of observations and it suffices to share beliefs on the state and applied control actions; a pattern we refer to as *k-step belief sharing pattern*. In case stochastic nestedness is absent, we can evaluate a precise expression for the minimum amount of information required to achieve belief sharing. The information exchange needed is generally strictly less than the information exchange needed for deterministic nestedness (even under optimal coders) and is zero whenever stochastic nestedness applies. We provide explicit examples of stochastically nested information structures and exhibit the benefit of belief sharing on information exchange requirements and discuss the monotone value of information channels.**

I. INTRODUCTION

In a decentralized system, different information is available to different decision makers who try to act on a common system towards a common goal as in team problems [12] or towards a variety of goals as in multi-criteria optimization problems [13]. Such team and decentralized multi-criteria optimization problems are challenging since the information patterns determining which agent has access to what information and the influence of her actions, can fall into the categories such that the generation of the optimal control laws can be very difficult, and of very high complexity.

We now proceed to make the decentralized system considered in this paper precise.

[1]Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Research Supported By the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Queen's Advisory Research Committee. Email: yuksel@mast.queensu.ca

A. Decentralized System Model

Let $\mathbb{X}$ be a space in which elements of a random sequence, $\{x_t, t \in \mathbb{Z}_+ \cup \{0\}\}$ live in. Let an observation channel $\mathcal{C}^i$ be defined as a stochastic kernel on $\mathbb{X} \times \mathbb{Y}^i$, such that for every $x \in \mathbb{X}$, $p(.|x)$ is a probability distribution on $\sigma(\mathbb{Y}^i)$ and for every $A \in \sigma(\mathbb{Y}^i)$, $p(A|.)$ is a function of x. Let there be L decision makers, $\{DM^i, i = 1, 2, \ldots, L\}$. Let a Decision Maker (DM) DM^i be located at one end of an observation channel $\mathcal{C}^i$, with inputs x_t generated as y_t^i at the channel output. We refer to a policy Π^i as a sequence of control functions which are causal such that the action of DM^i at time t, u_t^i, under Π^i is a causal function of its local information, that is, it is a measurable mapping with respect to the sigma-algebra generated by

$$I_t^i = \{y_t^i, Z_t^i; y_{[0,t-1]}^i, u_{[0,t-1]}^i, Z_{[0,t-1]}^i\} \quad t \geq 1,$$

with $I_0^i = \{y_0^i, Z_0^i\}$, to $\mathbb{U}^i$, with the notation for $t \geq 1$

$$y_{[0,t-1]}^i = \{y_s^i, 0 \leq s \leq t-1\}$$

Here Z_t^i denotes the additional information that can be supplied to DM^i at time t. Let DM^i have a policy Π^i and under this policy generate control actions $\{u_t^i, u_t^i \in \mathbb{U}^i, t \geq 0\}$, and let a dynamical system and observation channels be described by the following discrete-time equations:

$$x_{t+1} = f(x_t, u_t^1, u_t^2, \ldots, u_t^L, w_t),$$

$$y_t^i = g^i(x_t, v_t^i),$$

with $\{w_t\}$ independent, identical, white system noise process and $\{v_t^i, i = 1, 2, \ldots, L\}$ be disturbance processes. The disturbance processes might be correlated, but are independent of the system noise process.

Let $\mathbb{X}^T = \prod_{t=0}^{T-1} \mathbb{X}$ be the $T-$product space of $\mathbb{X}$. For the above setup, under a sequence of control policies $\{\Pi^1, \Pi^2, \ldots, \Pi^L\}$, we define an *Information-Control Structure* (ICS) as a probability space

$$(\Psi, \sigma(\Psi), P)$$

with

$$\Psi = \mathbb{X}^T \times \prod_{t=0}^{T-1} \prod_{k=1}^{L} \mathbb{Y}^k \times \prod_{t=0}^{T-1} \prod_{k=1}^{L} \mathbb{U}^k$$

Here, P is the probability measure on the sigma-algebra $\sigma(\mathbb{X}^T \times \prod_{i=0}^{T-1} \prod_{k=1}^{L} \mathbb{Y}^k \times \prod_{i=0}^{T-1} \prod_{k=1}^{L} \mathbb{U}^k)$.

*Information Pattern*s determine the sub-fields for all decision makers and time stages $\sigma(I_t^i) \subset \sigma(\mathbb{X}^T \times \prod_{i=0}^{T-1} \prod_{k=1}^{L} \mathbb{Y}^k \times \prod_{i=0}^{T-1} \prod_{k=1}^{L} \mathbb{U}^k)$. Hence, the control actions are measurable on the sub-fields, which are characterized by I_t^i for all DMs, through the term Z_t^i. In other words, an Information Pattern determines what the control action can depend on, inducing an information-control structure.

With the above formulation, let the objective of the decision makers be the minimization of

$$E_{x_0}^{\Pi^1, \Pi^2, \ldots, \Pi^L} \left[\sum_{i=0}^{T-1} c(x_t, u_t^1, u_t^2, \ldots, u_t^L) \right],$$

over all policies $\Pi^1, \Pi^2, \ldots, \Pi^L$, with initial condition x_0.

Let for a general vector q, $\mathbf{q}$ denote $\{q^1, q^2 \ldots, q^L\}$. Let $\boldsymbol{\Pi} = \{\Pi^1, \Pi^2, \ldots, \Pi^L\}$ denote the ensemble of policies. Under an ensemble of policies $\boldsymbol{\Pi}$ and a given information pattern, with an initial condition x_0, the attained performance index is

$$J_{x_0}(\boldsymbol{\Pi}) = E_{x_0}^{\boldsymbol{\Pi}} \left[\sum_{i=0}^{T-1} c(x_t, \mathbf{u}_t) \right]$$

In the above problem, information patterns affect the difficulty of control design, especially as the horizon T or the cardinality of $\mathbb{X}$ and $\mathbb{U}^i, i \in \{1, 2, \ldots, n\}$ increases.

B. Relevant Literature and Information Patterns

Various information structures have received particular interest. It has been almost customary to categorize such information structures as follows (see [1], [14], [3]):

Centralized Information Structure: All agents have the same information regarding the current value of the state. Here $Z_t^i = \{\mathbf{y}_t\}$ for all decision makers and time stages.

Quasi-Classical Information Structure: Whenever a dynamic programming recursion with a fixed complexity per time stage is possible, the information structure is said to have a quasi-classical pattern. This structure includes the *one-step delayed observation sharing* information pattern (see [6] and [13]), which allows the Decision Makers to share all their observations with a unit delay: $Z_t^i = \{\mathbf{y}_{t-1}\}$. If the agents also share their decisions, then the information pattern is called *one-step delayed information sharing* pattern: $Z_t^i = \{\mathbf{y}_{t-1}, \mathbf{u}_{t-1}\}$.

It was observed by Radner [12] that a static LQG team problem with non-nested information structures admit a linear solution. This argument follows from the observation that the team cost is convex in the joint strategies of the DM's, and it suffices to find the uique fixed point. This, in turn, is satisfied by a linear set of solutions for each DM. However, the extension of this result to a dynamic setup is not always possible. The following information structure present cases where, this result still holds in a dynamic setting.

Partially Nested Information Structure: An information structure is partially nested, if whenever the control actions of a DM^i affects the observations of another decision maker DM^j, the information available at DM^i is known noiselessly by the affected decision maker, that is: $Z_t^j = \{y_t^i, \quad \text{if} \quad \text{DM}^i \rightarrow \text{DM}^j\}$. Here the notation $DM^i \rightarrow DM^j$ denotes the fact that the actions of DM^i affects the information at DM^j. The partially nested structure effectively reduces the dynamic LQG team problem to a static optimization problem in the sense that the signaling (inner) agent (whose information sigma algebra is a subset of the signaled (outer) agent's information sigma algebra) makes all her decisions statically and the outer agent can generate such *pure strategy* decisions and the joint decisions can be regarded as one single-DM's decision, effectively making the problem static among such single DM's. Due to the static nature of the problem, one shows that the optimization is jointly convex in the decision variables and there is a person-by-person optimal solution, which turns out to be a linear set of solutions. As a special case, partially-nested structure includes the case where information propagation is faster than dynamics propagation [19], [11].

Non-classical Information Structures: If a decision maker's, DM^j, information is dependent on the actions of another, say DM^k, and DM^j does not have access to the information available to DM^k, this information structure is said to be *non-classical*. Hence, an information pattern which is not partially nested is a non-classical information pattern. The *one-step delayed control sharing pattern* $Z_t^i = \mathbf{u}_{t-1}$ is one such example [5], [3], [13].

Other information structures include the ones induced by the *n-step delayed information pattern* with $Z_t^i = \{\mathbf{y}_{t-n}, \mathbf{u}_{t-n}\}$. Such a pattern does not lead to a separation property [15] for $n \geq 2$. Here, by separation we mean that the conditional measure on a sufficient time in the past and the received observations thereafter are

sufficient statistics for the generation of optimal control laws.

A very important related information pattern is the *n-step periodic information sharing* pattern of Ooi et al [8], with $Z_t^i = \{\mathbf{y}_{[t-k-(t \mod k),t-(t \mod k)]}, \mathbf{u}_{[t-k-(t \mod k),t-(t \mod k)]}\}$, where $k \in \mathbb{Z}_+$ denotes the period of information sharing. This pattern does admit a separation structure for the generation of optimal control laws, and hence this leads to a quasi-classical information structure. We will discuss this pattern further in the paper, and provide an alternative derivation of the main results presented in [8] via Witsenhausen's equivalent model for discrete-stochastic control [10].

When the information structures are non-nested, controllers might choose to communicate via their control actions, that is might wish to pursue signaling. Three types of signaling can occur: signaling what the belief (that is, the conditional probability measure) on the state of the system is, signaling what the belief on the other agents controls are and signaling what the agent's own future control actions will be. These are all distinct issues and affect the classes of problems that we will discuss in the remainder of the paper.

C. Contributions of the paper: Stochastic Nestedness and the Belief Sharing Information Pattern

The information structures leading to tractable solutions require large *noiseless* data transfer between decision makers. In practice, it is not possible to send large amount data noiselessly especially in a real-time fashion. In particular, it is impossible to exchange a real number noiselessly over a practical channel.

In this paper, we present a weaker notion of nestedness, which we term as *Stochastic Nestedness*. We show that under stochastic nestedness, there is a rich class of problems which lead to tractable solutions despite the lack of deterministic nestedness or partial nestedness. The stochastically nested information structure entails a Markov chain condition between the observations at different sites and a state that is being controlled and in essence requires the information to be stochastically more informative at a decision maker than at another one together with the availability of actions.

In this paper, we also introduce another information pattern, which we call *Belief Sharing Information Pattern*. We note however, that, a similar pattern was discussed in [8], although we adopt an approach based on stochastic control, Witsenhausen's equivalent model and information theory. Such a pattern minimizes the information exchanges required for tractability, without any loss of performance in comparison with its de-

terministic counterpart in the observation sharing information pattern. The belief-sharing pattern allows us to formulate an optimization problem with minimum information exchange leading to a finite-complexity dynamic programming recursion to be applicable. We also address the communication rate minimization problem.

Finally, we investigate the effects of various communication channels on stochastic nestedness, when the channels are used in a decentralized control system.

D. Control of a Markov Chain under Partial Observations

Many of the results in this paper will base on the fact that while controlling a partially observed Markov chain, one could study the optimal control problem by enlarging the state space, via replacing the state with the belief on the state in a centralized setting (or a decentralized setting, in which case the notion of state becomes more complicated) and applying the control machinery on the belief process [2].

One could transform a partially observable Markov Decision Problem to a Fully Observed Markov Decision Problem via an enlargement of the event space [7], if the cost can be written as a summation of additive, per-stage costs.

In a general decentralized setting, the above discussion also applies, however, the notion of state x_t and the Markov recursion now involves a much larger space since the effective state includes the actions of the other DM's, and the beliefs of all DM's on every other DM's actions and as the time horizon increases, the beliefs on the beliefs and so on, leading to a non-tractable optimization problem.

II. STOCHASTICALLY NESTED INFORMATION STRUCTURE

In this section we will present three types of information patterns, all of which are non-classical, yet admit tractable recursions and when applied to LQG problems lead to the optimality of linear policies. We will later build on the findings of this section to present a new information sharing pattern. First, however, we discuss why nestedness is important for team decision problems.

A. A Linear Quadratic Gaussian Example and Importance of Nestedness

Consider a two-controller system:

$$x_{t+1} = Ax_t + B^1 u_t^1 + B^2 u_t^2 + w_t$$
$$y_t^1 = C^1 x_t + v_t^1$$
$$y_t^2 = C^2 x_t + v_t^2,$$

with w, v^1, v^2 zero-mean, i.i.d. disturbances. For $\rho_1, \rho_2 > 0$, let the goal be the minimization of

$$J = E\left[\left(\sum_{t=0}^{1} \|x_t\|_2^2 + \rho_1\|u_t^1\|_2^2 + \rho_2\|u_t^2\|_2^2\right) + \|x_2\|_2^2\right]$$

over the control policies of the form:

$$u_t^i = \mu_t^i(y_{[0,t]}^i), \quad i = 1, 2, t = 0, 1$$

For a two-stage problem, the cost is in general no-longer quadratic in the action of the controllers acting in the first stage $t = 0$: This is because these actions might affect the estimation quality of the other controllers in the second stage, if one DM can signal information at the other DM in one stage. We note that this condition is equivalent to $C^1 A^l B^2 \neq 0$ or $C^2 A^l B^1 \neq 0$ ([17], Lemma 3.1), with l denoting the delay in signaling. Hence, it is not immediate whether the cost function is jointly convex in the control policies, and as such finding a fixed point in the optimal policies does not necessarily lead to the conclusion that such policies are optimal.

Under the one-step delayed information structure case, or the partially nested case, this ceases to be true; there is no need for signaling, since all of the information that can be signaled is already available at the DMs that can be signaled. Thus, the cost is convex in both the second stage controls and the first stage ones; in particular, under any policy for the controls in the first stage, the second stage controls are linear and independent of an estimation error or improvement caused by control actions applied at the first stage. The optimization problem is still convex, and linear policies are person-by-person-optimal, leading to a globally optimal solution.

We will see that, one may not need nestedness for the convexity argument above to hold. We now proceed to define stochastic nestedness.

B. Stochastic Nestedness

Definition 2.1: Let for some measurable functions $f, g_i, i \in \{1, 2, \ldots, L\}$, a system be described by

$$x_{t+1} = f(x_t, u_t, w_t),$$
$$y_t^i = g_i(x_t, v_t^i), \quad i \in \{1, 2, \ldots, L\}$$

Under the decentralized model description of Section I-A: If whenever $DM^i \to DM^j$, it follows that:

$$x_0 \leftrightarrow y_0^j \leftrightarrow y_0^i$$

forms a Markov chain,

$$I_t^j = \{y_{[0,t]}^j, u_{[0,T-1]}^i\}$$

and

$$y_t^i = s_t(y_0^i),$$

where s_t is a deterministic function for $t \in \{0, 1, \ldots, T - 1\}$, then the information structure is stochastically nested.

Theorem 2.1: Under the decentralized system description of Section I-A, let $\mathbf{u}_t = [u_t^1 u_t^2 \ldots u_t^L]^T$ and $Q \geq 0, R > 0$ and there be an optimization problem with the objective to be minimized as:

$$J^i = E[\sum_{t=0}^{T-1} x_t^T Q x_t + \mathbf{u}_t^T R \mathbf{u}_t]$$

with the system dynamics:

$$\begin{aligned} x_{t+1} &= A x_t + \sum_{j=1}^{L} B^j u_t^j + w_t, \\ y_t^i &= C^i x_t + v_t^i, \quad 1 \leq i \leq L, \end{aligned} \quad (1)$$

where w_t, v_t^i are Gaussian and the disturbances and the noise processes are such that the information structure is stochastically nested. In this case, the optimal control laws are linear.

Remark: Note that, if we relax the Markov chain condition there will be an incentive for signaling from the inner DM to the outer DM on what the inner DM thinks regarding the initial state. The availability of the control actions is also essential, for otherwise, there will be an incentive for the inner DM to signal information on its future control signals. $\diamond$

C. Comparison with the Control Sharing Information Pattern

The stochastically nested information structure discussed above brings to mind the *Control Sharing Information Pattern* of Aoki [5], Sandell and Athans [3] and Bismut [4]. In those works, $\epsilon-$optimal policies were obtained for the control sharing pattern. The ϵ term arises due to the fact that the control policy is to encode information on both the control action and the observation, with as minimum damage as possible to the control action; and this is possible due to the fact that a real number carries infinite amount of information (when information is measured in Shannon information theoretic bits). One way to achieve this is as follows: Since rational numbers are dense in reals, for any ϵ, there exists an n such that an n-decimal representation which is at most at an ϵ distance (in the sup norm) from any real number in a compact set is possible. Hence, if one is to represent a finite dimensional control $\mathbb{U} = \mathbb{R}^r$, and a finite dimensional observation variable $\mathbb{Y} = \mathbb{R}^m$ taking

values in some compact set, all of these signals can be represented uniformly by an arbitrarily small error with a real number by transmitting all the $\epsilon-$ approximate decimal expansion of the numbers leading to a total of $n(m + r)$ decimal letters, by allocating the most significant nr letters for the control signal.

If the control and observation variables take values in a non-compact set, then, by separability, a countable representation is possible but the mapping in the transformation needs to be infinite, and a uniform number of decimal letters will not be sufficient, hence, the coding design becomes further impractical. In practical applications, there cannot exist a noiseless exchange of arbitrary real numbers, as this amounts to infinite amount of information exchange. Also, note that, such a setup is extremely sensitive to even an arbitrarily small noise [3].

In our setup, the resulting policy is optimal (and not only ϵ-optimal), and unlike the setups of [3] and [4], is applicable to cases where (i) the control policy is discontinuous, or (ii) the state space has finite cardinality (hence arbitrarily small precision of two signals is not possible via encoding into one-signal since there is only finite information that can be transmitted in one signal), (iii) the observation and control sets are not compact, eliminating the possibility of uniform approximation, (iv) the observation space of DM^i, $\mathbb{Y}^i$, is an inseparable space leading to an absence of a countable dense subset in turn leading to a uniformly close finite truncated representation, or (v) the time-horizon is not finite. These are some conditions under which the assumptions of [3] and [4] are not applicable.

The applicability of the above scenarios under the stochastically partially nested structure follows because in the information structure presented here, the signaling DM does not need to encode any information on her observations, as what she can encode is useless for the other decision makers which have more informative observations. The exchange of the control signals, however, is essential.

In the following section we investigate the case where there is no nestedness. We evaluate the information requirements to obtain stochastic nestedness.

III. BELIEF SHARING INFORMATION PATTERN

The computationally attractive aspects of a partially nested, or nested information structure comes with a price of exchanging *all* of the information available by the preceding controllers noiselessly. This is, however, impractical. In the analysis heretofore, we have weakened the information requirements for tractability

in a class of decentralized optimization problems. We now investigate the quantitative minimization of the information requirements needed for tractability in a large class of decentralized optimal control problems.

Before proceeding further, let us recall Witsenhausen's equivalent model ([10], [16]) for dynamic team problems in terms of an extensive form static team problem. Let there be a common information vector I_t^c at some time t, which is available at all of the decision makers. Let at times ks, $k \in \mathbb{Z}_+ \cup \{0\}$ and T divisible by k, $s \in \mathbb{Z}_+$, the decision makers share all their past information: $I_{ks}^c = \{\mathbf{y}_{[0,ks-1]}, \mathbf{u}_{[0,ks-1]}\}$. In this case, until the next observation instant $t = (k+1)s$ we can regard the individual decision functions specific to DM^i as $\{u_t^i = \bar{u}_s^i(y_{[ks,t]}^i, I_{ks}^c)\}$ and we let $\bar{\mathbf{u}}$ denote the ensemble of such decision functions. In essence, it suffices to generate $\bar{\mathbf{u}}_s$ for all $s \geq 0$, as the decision outputs conditioned on $y_{[ks,t]}^i$, under $\bar{u}_s^i(y_{[ks,t]}^i, I_{ks}^c)$, can be generated. Witsenhausen achieved this by transforming the effects of the control action into the costs and formulating an equivalent control problem. In such a case, we have that $\bar{\mathbf{u}}_s(., I_{ks}^c)$ is the joint team decision rule mapping I_{ks}^c into a space of action vectors, $\{u^i(I_{ks}^c, y_{[ks,t]}^i), i \in \{1, 2 \ldots, L\}, t \in \{ks, ks+1, \ldots, k(s+1) - 1\}\}$.

In this case, the cost function is also adjusted as:

$$J_{x_0}(\Pi) = E_{x_0}^{\mathbf{\Pi}}\Big[\sum_{s=0}^{\frac{T}{k}-1} \bar{c}(\bar{\mathbf{u}}_s(., I_{ks}^c), \bar{x}_s)\Big]$$

with

$$\bar{c}(\bar{\mathbf{u}}_s(., I_{ks}^c), \bar{x}_s) = \sum_{t=ks}^{(k+1)s-1} c(x_t, \mathbf{u}_t)$$

Lemma 3.1: Consider the decentralized system setup in Section I-A, with the observation noise processes being independent. Let I_t^c be a common information vector supplied to the DMs regularly at every k time steps, so that the DMs have common memory with a control policy generated as described above. Then, $\{\bar{x}_s := x_{ks}, \bar{\mathbf{u}}_s(., I_{ks}^c), s \geq 0\}$ form a Controlled Markov chain

In view of the above, we now present a result on a separation property. We note that the following has been proven in [8]. We present a shorter proof, using the result above directly.

Lemma 3.2: Let I_t^c be a common information vector supplied to the DMs regularly at every k time steps. There is no loss in performance if I_{ks}^c is replaced by $P(\bar{x}_s|I_{ks}^c)$.

Proof:

The cost can be written as a function of additive costs:

$$J_{x_0}(\mathbf{\Pi}, IS) = E_{x_0}^{\mathbf{\Pi}}\left[\sum_{s=0}^{\frac{T}{k}-1} \tilde{c}(\bar{\mathbf{u}}_s, \bar{x}_s)\right]$$

with

$$\tilde{c}(\bar{\mathbf{u}}_s, \bar{x}_s) = \sum_{t=ks}^{(k+1)s-1} c(x_t, \mathbf{u}_t)$$

For the minimization of an additive cost in Partially Observed Markov Chains, it suffices to transform the state to an equivalent state of conditional densities [7] as discussed in Section I-D. Hence $P(\bar{x}_s|I_s^c)$ acts as a sufficient statistic. See also [2].

$\diamond$

Clearly, the larger k is, the larger the complexity in the design of the computation of the control law $\bar{\mathbf{u}}_{ks}$. This grows at least exponentially in k.

The essential issue for a tractable solution is to ensure a common information vector which will act as a sufficient statistic for future control policies. This can be done via a one-step delayed structure, or some structure possibly requiring larger but finite delay.

Definition 3.1: Belief Sharing Information Pattern: An information pattern in which the DMs share their beliefs about the system state is called the *belief sharing information pattern*. If the belief sharing occurs periodically at every k-stages, and the DMs also share the control actions they applied in the last k-1 stages, we call this the *k-stage belief sharing information pattern*.

We now discuss how the beliefs are shared sequentially. We proceed by induction. Suppose at time ks, the DMs have an agreement on $P(\bar{x}_s|I_{ks}^c)$ and know the policies used by each of the DM's, hence know the ICS and the probability measure P. It follows that,

$$P(\bar{x}_{s+1}|\mathbf{y}_{[ks,k(s+1)-1]}, P(\bar{x}_s|I_{ks}^c))$$
$$= \frac{P(\bar{x}_{s+1}, \mathbf{y}_{[ks,k(s+1)-1]}, P(\bar{x}_s|I_{ks}^c))}{\sum_{\mathbb{X}} P(\bar{x}_{s+1}, \mathbf{y}_{[ks,k(s+1)-1]}, P(\bar{x}_s|I_{ks}^c))}$$
$$= \frac{\sum_{\mathbb{X}} P(\bar{x}_{s+1}, \bar{x}_s, \mathbf{y}_{[ks,k(s+1)-1]}, P(\bar{x}_s|I_{ks}^c))}{\sum_{\mathbb{X}}\sum_{\mathbb{X}} P(\bar{x}_{s+1}, \bar{x}_s, \mathbf{y}_{[ks,k(s+1)-1]}, P(\bar{x}_s|I_{ks}^c))}$$

Hence, to recover the joint belief, it suffices that the DMs share: $P(y_{[ks,k(s+1)-1]}^1, y_{[ks,k(s+1)-1]}^2, \ldots, y_{[ks,k(s+1)-1]}^L|\bar{x}_s)$, which is a $|\mathbb{X}|-$dimensional vector with $\mathbb{R}$-valued components for every $\bar{x}_s \in \mathbb{X}$ given $\mathbf{y}_{[ks,k(s+1)-1]}$ and the control strategy $\bar{\mathbf{u}}_s(., I_{ks}^c)$, and actions. In case the state space is uncountable, a probability measure on $\sigma(\mathbb{X})$ is to be exchanged. With this description, the recursions can be obtained.

A. Minimum Communication Rate Needed for the Belief Sharing Pattern

The exchange of the common information states under deterministic nestedness might lead to a large information exchange *noiselessly*. This is impractical for many scenarios. However, as a result of Lemma 3.1 and 3.2, what needs to be exchanged is a sufficient amount of information such that the DMs have a common $P(\bar{x}_s|I_s^c)$, so that their recursions can be based on this information. The question that we are interested in this section is the following: How much information exchange is needed between the decision makers so that the decision makers have an agreement on the state of the system and a dynamic programming recursion is tractable? The information is measured by information bits, that is the average number of bits needed to be exchanged among the decision makers.

Now, let us introduce two standard information-theoretic notions, namely mutual information and rate distortion function. *Mutual information* between an input random variable, X, and another one, Y, is $I(X;Y) = H(X) - H(X|Y)$, where $H(X)$ is the entropy of X (differential entropy if X is a continuous random variable), and $H(X|Y)$ is the conditional entropy of X given Y. The entropy of a variable is an important quantity since the entropy provides an almost tight lower bound for the expected number of bits for noiseless transmission of data. We assume variable-rate, time-invariant encoding, that is, the rate is defined as the expected number of bits to be transmitted: The coding process of the controller at DM^i is a mapping measurable with respect to the sigma-algebra generated by I_t^i. The DM^i's coding policy to DM^j is a mapping from I_t^i to $\mathcal{W}^{i,j} = \{W^{i,j}(1), W^{i,j}(2), \ldots, W^{i,j}(N)\}$, the codebook for communication from DM^i to DM^j. Hence, at each time stage, t, DM^i sends $R^{i,j} = -\sum_{n=1}^{N} P(W^{i,j}(n), t)\log_2(P(W^{i,j}(n), t))$ bits on average over an external channel to DM^j at time t. Let $\mathcal{R} = \{R^{i,j}, i \neq j \in 1, 2, \ldots, L\}$ such that belief sharing is possible. Define $R := \inf_{\mathcal{R}}\{\sum_{i,j=1, j\neq i}^{L} R^{i,j}\}$, such that belief sharing is realized. We wish to obtain the infimum of such R values.

In the following, we consider the discrete-alphabet valued observation case, as the analysis is simpler to pursue. Note that we do not require the state-space to be discrete, the process $\{x_t\}$ might still live in an uncountable space. However, if the observation process takes a countable number of values the following analysis is applicable. As mentioned above, with natural extensions, one could study the most general case of uncountable $\mathbb{X}, \mathbb{Y}^i$ cases for all DM^i.

1) One-Step Belief Sharing Pattern: Let us consider the one-step delayed information pattern, first for a two DM setup. In this case, the information needed at both the controllers is such that they all need to exchange the relevant information on the state, and need to agree on $p(\bar{x}_t|I_t^1, I_t^2)$, where I_t^i denotes the information available at DM^i. In the one-step Belief-Sharing Pattern, $\bar{x}_t = x_t$, since the period for information exchange $k = 1$.

In the one-step belief sharing pattern, control actions need to be exchanged, since they can be obtained given the agreement on the beliefs.

Theorem 3.1: Suppose the observation variables are discrete values, that is $\mathbb{Y}^i, i = 1, 2$ is a countable space. To achieve the belief sharing information pattern, a lower bound on the minimum average amount of bits to be transmitted to DM^2 is lower bounded by:

$$R^{2,1} \geq H\left(P(x_t|I_{t-1}^c, y_t^1, y_t^2) \Big| P(x_{t-1}|I_{t-1}^c), y_t^1 \right)$$

A lower bound on the minimum amount of information needed to be transmitted to DM^1 from DM^2 is:

$$R^{1,2} \geq H\left(P(x_t|I_{t-1}^c, y_t^1, y_t^2) \Big| P(x_{t-1}|I_{t-1}^c), y_t^2 \right)$$

We note that the information needed is less than one needed for achieving the one-step delayed information pattern. By the above argument, one would need $R^{i,j} \geq H(y_t^i|y_t^j, I_t^c)$ for the one-step delayed observation sharing pattern. The entropy of the output is at most as much as the entropy of the observed variable. This is because, different outputs may lead to the same values for $P(y_t^2 = y|x_t, I_t^c)$. Hence, we have the following corollary to Theorem 3.1.

Corollary 3.1: When the observation space is discrete, the one-step belief sharing information pattern requires less or equal amount of information exchange between the controllers than the one-step delayed observation sharing pattern.

For the multiple-decision maker case, one has a distributed coding with side information scenario: In this case the decision makers will send correlated information to another decision maker. This leads us to the following lower bound.

Theorem 3.2: To achieve the belief sharing information structure, a lower bound on the minimum average amount of bits that needs to be supplied to any DM^i, $\{i = 1, 2, \ldots, L\}$ is given by:

$$R^i \geq H\left(P(x_t|I_{t-1}^c, y_t^1, y_t^2, \ldots, y_t^L) \Big| P(x_{t-1}|I_{t-1}^c), y_t^i \right)$$

One could extend the above result for continuous-alphabet valued for the state and observation spaces. To lead to meaningful results however, one needs to assume a given distortion level as the rate required typically becomes infinite, as the conditional measure process typically converges to $-\infty$, leading to $I(X;Y) = h(X) - h(X|Y)$ to grow unbounded. Here $h(X)$ denotes the differential entropy of the random variable X.

B. A Case Study: Stochastically Nested Structure with Zero-Capacity Channels

An example is for the case in which the channels has zero capacity. In this case, as

$$P(\bar{y}_s^1 = \eta|\bar{x}_s) = P(\bar{y}_s^1 = \beta|\bar{x}_s)$$

for all η, β values that the observation can take, there is no further information that is needed for the belief-sharing pattern. As such, there is no need for information exchange, since there is no information generated by the observation for the controller with regard to the state and no transmitted information will be useful. Hence, the communication required for stochastic nestedness is zero if all of the information channels are channels with zero-capacity.

It should be noted that, when the channels are zero-capacity channels, the deterministic nestedness conditions would require all the information to be exchanged, although the benefit of this is zero. This example indeed clearly exhibits the efficiency difference between the two information patterns.

REFERENCES

[1] T. Başar and J. Cruz, "Concepts and methods in multiperson coordination and control", in em S. G. Tzafestas, editor, *Optimization and Control of Dynamic Operational Research Models*, Chapter 11, pp. 351 - 387. North Holland, 1982.

[2] A. Arapostathis, V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh and S. I. Marcus, "Discrete-Time Controlled Markov Processes with Average Cost Criterion: A Survey," *SIAM J. Control and Optimization*, vol. 31, pp. 282-344, 1993.

[3] N. Sandell and M. Athans, "Solution of some nonclassical LQG stochastic decision problems", *IEEE Trans. Automatic Control*, vol 19, pp. 108-116, 1974.

[4] J. M. Bismut, "An example of interaction between information and control: The transparency of a game", presented at the *Joint Harvard University - Imperial College Conf. Information Structures and Extensive Games*, June 1972.

[5] M. Aoki, "On decentralized linear stochastic control problems with quadratic cost *IEEE Trans. Automatic Control*, vol 18, pp. 243 - 250, Jun 1973.

[6] B. Kurtaran, "Corrections and extensions to 'Decentralized control with delayed sharing information pattern *IEEE Trans. Automatic Control*, vol 24, pp. 656-657, Aug. 1979.

[7] A. N. Shiryaev, "On Markov sufficient statistics in nonadditive Bayes problems of sequential analysis", *Theory Probab. Appl.*, vol. 9, pp. 604-618, 1964.

[8] J. Ooi, S. Verbout, J. Ludwig and G. Wornell, "A separation theorem for periodic sharing information patterns in decentralized control", *IEEE Trans. Automatic Control*, vol. 42, pp. 1546 1550, Nov 1997.

[9] J.D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Information Theory*, 19:471–480, July 1973

[10] H. S. Witsenhausen, "Equivalent stochastic control problems, *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 3-11, Springer-Verlag, 1988.

[11] A. Rantzer, "Linear Quadratic Team Theory Revisited," *Proc. IEEE American Control Conference*, Minneapolis, June 2006.

[12] R. Radner, "Team decision problems," *Annals of Mathematical Statistics*, vol. 33, pp. 857 - 881, 1962

[13] T. Başar, "Decentralized multicriteria optimization of linear stochastic systems", *IEEE Transactions on Automatic Control*, vol. 23, pp. 233 - 243, April 1978.

[14] Y. C. Ho, "Team decision theory and information structures, *Proc. IEEE*, vol. 68, no. 6, pp. 644654, 1980.

[15] P. Varaiya and J. Walrand, "On delayed-sharing patternrs", *IEEE Trans Automatic Cotrol*, vol. 23. pp. 44345, June 1978.

[16] D. Teneketzis, "On information structures and nonsequential stochastic control, *CWI Quarterly*, vol. 9, pp. 241-260, 1996.

[17] S. Yüksel and T. Başar, "Optimal signaling policies for decentralized multi-controller stabilizability over communication channels", *IEEE Trans. Automatic Control*, vol. 52, pp. 1969-1974, October 2007.

[18] S. Yüksel and T. Başar, "Communication constraints for decentralized stabilizability with time-Invariant policies", *IEEE Trans. Automatic Control*, vol. 52, pp. 1060-1066, June 2007.

[19] P. G. Voulgaris, "A convex characterization of classes of problems in control with specific interaction and communication structures, in *Proc. IEEE American Control Conference*, pp. 3128-3133, Arlington, VA, June 2001.