2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Stochastic Nestedness and the Belief Sharing Information Pattern in Decentralized Control

Serdar Yüksel[1]

***Abstract*— In a dynamic decentralized control problem, a common information state supplied to each of the Decision Makers leads to a tractable dynamic programming recursion. However, communication requirements for such conditions require exchange of very large data noiselessly, hence these assumptions are generally impractical. We present a weaker notion of nestedness, which we term as stochastic nestedness, which is characterized by a sequence of Markov chain conditions. It is shown that if the information structure is stochastically nested, then an optimization problem is tractable, and in particular for LQG problems, the team optimal solution is linear, despite the lack of deterministic nestedness or partial nestedness. One other contribution of this paper is that, by regarding the multiple decision makers as a single decision maker and using Witsenhausen's equivalent model for discrete-stochastic control, it is shown that the common state required need not consist of observations and it suffices to share beliefs on the state and control actions; a pattern we refer to as *k-stage belief sharing pattern*. We evaluate a precise expression for the minimum amount of information required to achieve such an information pattern for $k = 1$. The information exchange needed is generally strictly less than the information exchange needed for deterministic nestedness and is zero whenever stochastic nestedness applies.**

I. Introduction

In a decentralized system, different information is available to different decision makers who act on a system towards a common goal as in team problems or towards a variety of goals as in multi-criteria optimization problems. Such 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]Mathematics and Engineering Program and the 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 $\mathbb{Y}^i$, be another space for $i = 1, 2, \ldots, L$ and 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 the (Borel) sigma-algebra $\sigma(\mathbb{Y}^i)$ and for every $A \in \sigma(\mathbb{Y}^i)$, $p(A|.)$ is a function of x. We will mostly be concerned with cases when $\mathbb{X}$ and $\mathbb{Y}^i$ are either finite sets or are finite-dimensional real vector spaces. Let there be L decision makers, $\{\mathrm{DM}^i, i = 1, 2, \ldots, L\}$. 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 $y^i_t \in \mathbb{Y}^i$ 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^i_t, 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

$$I^i_t = \{y^i_t, Z^i_t; y^i_{[0,t-1]}, u^i_{[0,t-1]}, Z^i_{[0,t-1]}\} \quad t \geq 1,$$
$$I^i_0 = \{y^i_0, Z^i_0\},$$

to $\mathbb{U}^i$, with the notation for $t \geq 1$, $y^i_{[0,t-1]} = \{y^i_s, 0 \leq s \leq t-1\}$. Here Z^i_t 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 $\{u^i_t, t \geq 0\}$, $u^i_t \in \mathbb{U}^i$, and let a dynamical system and observation channels be described by the following discrete-time equations:

$$x_{t+1} = f(x_t, u^1_t, u^2_t, \ldots, u^L_t, w_t),$$
$$y^i_t = g^i(x_t, v^i_t),$$

for some measurable functions $f, \{g^i\}$, with $\{w_t\}$ independent, identical, white system noise process and $\{v^i_t, i = 1, 2, \ldots, L\}$ 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

$$(\mathbb{X}^T \times \prod_{t=0}^{T-1}\prod_{k=1}^{L} \mathbb{Y}^k \times \prod_{t=0}^{T-1}\prod_{k=1}^{L} \mathbb{U}^k, \sigma(.), P)$$

Here, P is the probability measure on the (Borel) 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

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

over all policies $\Pi^1, \Pi^2, \ldots, \Pi^L$, with initial condition x_0. Here, $E_{x_0}^{\Pi^1,\Pi^2,,\ldots,\Pi^L}[]$ denotes the expectation over all sample paths with initial state given by x_0 under policies $\{\Pi^1, \Pi^2, \ldots, \Pi^L\}$. For a general vector q, let $\mathbf{q}$ denote $\{q^1, q^2 \ldots, q^L\}$. Let $\mathbf{\Pi} = \{\Pi^1, \Pi^2, \ldots, \Pi^L\}$ denote the ensemble of policies. Under an ensemble of policies $\mathbf{\Pi}$ and a given information pattern, with an initial condition x_0, the attained performance index is

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

B. Relevant Literature and Information Patterns

It has been almost customary to categorize information structures as follows (see [3], [6], [10]). Please see [15] for a detailed review of the literature:

Centralized Information Structure: Here, all DMs 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.

We say the information available at DM^i is nested in that of DM^j at time t, if $\sigma(I_t^i) \subset \sigma(I_t^j)$. Nestedness, as we will observe in the development of the paper, has very important implications. It was observed by Radner [9] that a static LQG team problem with a non-nested information structure admits an optimal solution which is linear.

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: $Z_t^j = \{y_t^i, \quad \text{if} \quad \mathrm{DM^i} \to \mathrm{DM^j}\}$. Here the notation DM^i $\to$ DM^j denotes the fact that the actions of DM^i affects the information at DM^j.

Non-classical Information Structures: An information pattern which is not nested or 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. 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 for $n \geq 2$. Here, by separation we mean that the conditional probability measure on a sufficient time in the past and the received observations thereafter are sufficient statistics for the generation of optimal control laws. Studies of this information pattern with separation results are reported in [1] and [7]. A related information pattern is the *n-step periodic information sharing pattern* studied by Yoshikawa [14] and Ooi et al [8] with Z_t^i consisting of $\mathbf{y}_{[t-k-(t \mod k), t-(t \mod k)]}$, and $\mathbf{u}_{[t-k-(t \mod k), t-(t \mod k)]}$, where $k \in \mathbb{Z}_+$ denotes the period of information sharing. This pattern admits a separation structure for the generation of optimal control laws. 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 [13].

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.

II. Stochastically Nested Information Structure

In this section we will present a class of information patterns, which is non-classical, yet its related optimization problems admit tractable recursions and when applied to LQG problems, leads to the optimality of linear policies. First, we discuss why nestedness is important for team decision problems. Consider a two-controller system evolving in $\mathbb{R}^n$ with the following description:

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

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 to the other DM in one stage. We note that this condition is equivalent to $C^1A^lB^2 \neq 0$ or $C^2A^lB^1 \neq 0$ ([**?**], Lemma 3.1), with $l+1$ denoting the delay in signaling, with $l = 0$ in the problem considered. 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.

A. Stochastic Nestedness

Before proceeding further, we introduce a related notion: Consider three random variables A, B, C in some common probability space. If A and C are conditionally independent given B, we say that $A \leftrightarrow B \leftrightarrow C$ form a Markov chain, and it follows that $P(A|B,C) = P(A|B)$.

Definition 2.1: For measurable functions $f, g^i, i \in \{1, 2, \dots, L\}$, consider a system described by

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

Under the decentralized model description of Section I-A: If whenever $\mathrm{DM}^i \to \mathrm{DM}^j$, 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 $y_t^i = h_t(y_0^i)$, where h_t is a deterministic function for $t \in \{0, 1, \dots, T-1\}$, then the information structure is stochastically nested.

Theorem 2.1: Under the decentralized system description of Section I-A, let $\mathbf{u}_t = \begin{bmatrix} u_t^1 & u_t^2 & \dots & u_t^L \end{bmatrix}^T$ and $Q \geq 0, R > 0$. Consider an optimization problem with the objective to be minimized as: $J = E[\sum_{t=0}^{T-1} x_t^T Q x_t + \mathbf{u}_t^T R \mathbf{u}_t]$, with the system dynamics:

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

where x_0, 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: It should be noted 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. $\diamond$

The stochastically nested information structure discussed above brings to mind the *Control Sharing Information Pattern* of Aoki [2], Sandell and Athans [10] 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.

In our setup, the resulting policy is optimal (and not only ϵ-optimal), and unlike the setups of [10] 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), or (iii) the observation and control sets are not compact.

B. Stochastically Decoupled Information Structure

Definition 2.2: Let a state be explicitly represented by its individual components $\mathbf{x} = \begin{bmatrix} x^1 & x^2 & \dots & x^L \end{bmatrix}$, and evolve under the following dynamics

$$x_{t+1}^i = f_1(x_t^i, u_t^i, w_t^i), \quad i \in \{1, 2, \dots, L\}$$
$$z_{t+1} = f(z_t, u_t^1, u_t^2, \dots, u_t^L),$$

$$y_t^i = g^i(\mathbf{x}_t, v_t^i), \quad i \in \{1, 2, \ldots, L\},$$

for some measurable functions f_i, g^i, f and $\{w_t^i\}$ independent state disturbance processes and $\{v_t^i\}$, observation noise processes for $i \in \{1, 2, \ldots, L\}$. In the above $(x_t^i, w_t^i, v_t^i, i \in \{1, 2, \ldots, L\}, z_0)$ are independent second-order processes. Under the decentralized system description of Section I-A, suppose each of the DMs has access to the additional $\{z_t\}$ process: $\tilde{y}_t^i = [y_t^i, z_t], \quad i \in \{1, 2, \ldots, L\}$. If the information available at each controller is such that

$$x_t^i \leftrightarrow (y_{[0,t]}^i, u_{[0,t-1]}^i) \leftrightarrow \{x_0^j, z_0, w_{[0,t-1]}^j, y_{[0,t]}^j, \quad j \neq i\}$$

form Markov chains, for all t and i, then such an information structure is said to be stochastically decoupled.

Theorem 2.2: Let there be an optimization problem with the objective to be minimized as:

$$E[\sum_{t=0}^{T-1} c_1(x_t^1, u_t^1) + c_2(x_t^2, u_t^2) + \ldots c_L(x_t^L, u_t^L)].$$

If the controllers have stochastically decoupled information structures, then the optimal control problem is tractable and a finite dimensional dynamic programming recursion can be used to obtain the solutions.

An example is the following dynamical system described by

$$x_{t+1}^1 = a_1 x_t^1 + u_t^1 + w_t^1$$
$$x_{t+1}^2 = a_2 x_t^2 + u_t^2 + w_t^2$$
$$x_{t+1}^3 = a_3 x_t^3 + u_t^1 + u_t^2 + w_t^3$$
$$y_t^1 = (x_t^1 + v_t^1, x_t^2 + v_t^2 + v_t^{21}, x_t^3 + v_t^{31})$$
$$y_t^2 = (x_t^1 + v_t^1 + v_t^{12}, x_t^2 + v_t^2, x_t^3 + v_t^{32}),$$

with the goal of the minimization of

$$J = E\Bigg[\sum_{t=0}^{T-1} \bigg((x_t^1)^2 + (x_t^2)^2 + \rho_1 (u_t^1)^2 + \rho_2 (u_t^2)^2 \bigg)\Bigg],$$

with $\rho_1, \rho_2 > 0$ constants.

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 at 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 [13] 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 $ks-1$, $k \in \mathbb{Z}_+ \cup \{0\}$ and T divisible by k, $s \in \mathbb{Z}_+$, the decision makers share all their information: $I_{ks-1}^c = \{\mathbf{y}_{[0,ks-1]}, \mathbf{u}_{[0,ks-1]}\}$ and for $I_0^c = \{P(x_0)\}$, that is at time 0 the DMs have the same apriori belief on the initial state. Until the next observation instant $t = k(s+1)-1$ we can regard the individual decision functions specific to DM^i as $\{u_t^i = \bar{u}_s^i(y_{[ks,t]}^i, I_{ks-1}^c)\}$ 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 $y_{[ks+1,t]}^i$, under $\bar{u}_s^i(y_{[ks,t]}^i, I_{ks-1}^c)$, 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 $\bar{\mathbf{u}}_s(., I_{ks-1}^c)$ is the joint team decision rule mapping I_{ks-1}^c into a space of action vectors: $\{u^i(I_{ks-1}^c, y_{[ks,t]}^i), i \in \{1, 2 \ldots, L\}, t \in \{ks, ks+1, \ldots, k(s+1)-1\}\}$. In this case, the cost function is also modified as:

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

with

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

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 stages, so that the DMs have common memory with a control policy generated as described above. Then, $\{\bar{x}_s := x_{ks}, \bar{\mathbf{u}}_s(., I_{ks-1}^c), s \geq 0\}$ forms a Controlled Markov chain

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-1}^c is replaced by $P(\bar{x}_s | I_{ks-1}^c)$.

Proof: The cost can be written as a function of additive costs: $J_{x_0}(\mathbf{\Pi}, IS) = E_{x_0}^{\mathbf{\Pi}}[\sum_{s=0}^{\frac{T}{k}-1} \tilde{c}(\bar{\mathbf{u}}_s, \bar{x}_s)]$, with $\tilde{c}(\bar{\mathbf{u}}_s, \bar{x}_s) = \sum_{t=ks}^{k(s+1)-1} c(x_t, \mathbf{u}_t)$. For the minimization of an additive cost in Partially Observed Markov Chains, it suffices to transform the state to an equivalent state of conditional distributions [11]. Hence $P(\bar{x}_s | I_{ks-1}^c)$ acts as a sufficient statistic. $\diamond$

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 sharing information at every stage, 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 ($k > 1$), the DMs also share the control actions they applied in the last k-1 stages, together with intermediate belief information. In this case, the information pattern is called the *k-stage belief sharing information pattern*. $\diamond$

Remark: It should be noted that, the exchange of the control actions is essential, as was discussed in view of stochastic nestedness. The DMs also need to exchange information for intermediate beliefs. The following algorithmic discussion will make this clear. $\diamond$

We now discuss how the beliefs are shared sequentially. We proceed by induction. Suppose at time $ks-1$, the DMs have an agreement on $P(\bar{x}_s|I^c_{ks-1})$ and know the policies used by each of the DMs, hence know the ICS and the probability measure P. It follows that,

$$\pi_{s+1} := P(\bar{x}_{s+1}|\mathbf{y}_{[ks,k(s+1)-1]}, \mathbf{u}_{[ks,k(s+1)-1]}, \pi_s)$$

writes as

$$\frac{P(\bar{x}_{s+1}, (\mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}{\sum_{\bar{x}_{s+1}} P(\bar{x}_{s+1}, (\mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}$$

$$= \frac{\sum_{x_{[ks,k(s+1)-1]}} P(\bar{x}_{s+1}, (x, \mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}{\sum_{x_{[ks,k(s+1)-1]}, \bar{x}_{s+1}} P(\bar{x}_{s+1}, (x, \mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}$$

We now express the numerator above more explicitly as

$$\sum_{x_{k(s+1)-1}} \Bigg(P(x_{k(s+1)}|x_{k(s+1)-1}, \mathbf{u}_{k(s+1)-1}) \\ (\prod_{l=1}^{L} P(y^l_{k(s+1)-1}|x_{k(s+1)-1})) \\ \ldots \\ \ldots \quad \ldots \quad \ldots \\ \ldots \\ \sum_{x_{ks}} \Bigg(P(x_{ks+1}|x_{ks}, \mathbf{u}_{ks}) \\ (\prod_{l=1}^{L} P(y^l_{ks}|x_{ks}))P(x_{ks}|I^c_{ks-1})\Bigg)\Bigg) \ldots \Bigg)\Bigg) \tag{2}$$

Iff $k > 1$, then the DMs also need to share the control actions applied in the previous $k-1$ time stages, as well as beliefs on individual states.

When the belief-sharing occurs at every stage, then controls can be generated by each of the DMs, hence the control actions need not be shared.

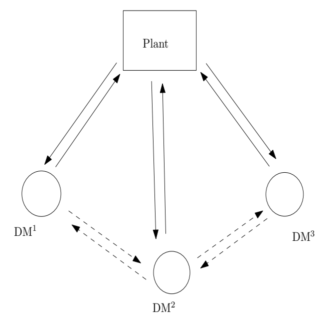

Fig. 1: A version of the belief propagation algorithm can be used for the belief sharing pattern. For a cycle-free network the analysis is simpler.

In the following, we study communication requirements such that such belief-sharing can be achieved for the case when $k = 1$.

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 $P(\bar{x}_s|I^c_s)$, so that their recursions can be based on this information.

Let us consider the one-stage belief sharing 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 $p(\bar{x}_t|I^1_t, I^2_t)$, where I^i_t 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 $k = 1$.

Theorem 3.1: Suppose the observation variables are discrete valued, 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 given by:

$$R^{2,1} \geq H\bigg(P(x_t|I^c_{t-1}, y^1_t, y^2_t)\bigg|P(x_{t-1}|I^c_{t-1}), y^1_t\bigg)$$

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

$$R^{1,2} \geq H\bigg(P(x_t|I^c_{t-1}, y^1_t, y^2_t)\bigg|P(x_{t-1}|I^c_{t-1}), y^2_t\bigg)$$

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

B. Case Studies

In the following, we provide a few explicit examples, which exhibit the weaker conditions required by stochastic nestedness.

1) Zero-Capacity Channels:

Proposition 3.1: Consider the case in which the channels have zero capacity. In this case, as

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

for all η, β values that the observation can take, there is no further information that is needed for the belief-sharing pattern.

Proof: This follows from $H\left(P(x_t|I_{t-1}^c, y_t^1, y_t^2)\middle|P(x_{t-1}|I_{t-1}^c), y_t^1\right) = 0$. This rate bound is tight. $\diamond$

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. We note that, in such a case, the controls do not need to be exchanged either, as there is already an agreement on the beliefs, based on the apriori belief; and the optimal team decisions can be generated decentrally. 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 performance benefit of this is zero. This example exhibits the efficiency difference between the two information patterns.

2) Stochastically Decoupled Structure: For this information pattern, belief sharing is not needed, as the problem is tractable under the mentioned structures and the problem is already partitioned into L independent, centralized optimal control under partial observation problems.

3) Stochastically Nested Structure: In case there is stochastic nestedness, the outer DM need not receive any information, except for the control actions of the inner DM.

IV. Conclusion

This paper presented a new information structure, the *stochastically nested* information structure, which is weaker than the previously known information structures (such as the nested, or partially nested information structures, or the observation sharing information patterns) which lead to tractable decentralized optimization solutions, which also lead to the optimality of linear solutions for LQG team optimization problems. We also presented a new information sharing pattern, the *belief sharing information pattern*. Under this pattern, the communication exchange requirements is shown to be strictly less than the requirements for deterministic nestedness, or the deterministic observation sharing patterns. We provided quantitative examples exhibiting the benefit of the new information structure.

References

[1] M. Aicardi, F. Davoli, R. Minciardi, "Decentralized optimal control of Markov chains with a common past information set", *IEEE Trans. Automatic Cotrol*, vol. 32, pp. 1028-1031, Nov. 1987.

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

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

[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] G. Casalino, F. Davoli, R. Minciardi, P.P. Puliafito and R. Zoppoli, "Partially nested information structures with a common past", *IEEE Trans. Automatic Control*, vol. 29, pp. 846-850, Sept. 1984.

[6] Y. C. Ho, K. C. Chu "Team decision theory and information structures in optimal control problems - Part I" *IEEE Trans. Automatic Control*, 17, pp. 15 - 22, Feb. 1972.

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

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

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

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

[12] P. Varaiya and J. Walrand, "On delayed-sharing patterns", *IEEE Trans Automatic Cotrol*, vol. 23. pp. 443-445, June 1978.

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

[14] T. Yoshikawa, "Decomposition of dynamic team decision problems", *IEEE Trans. Automatic Control*, vol. 23, pp. 627-632, August 1978.

[15] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern", *IEEE Trans. Automatic Control*, to appear.