Stochastic Nestedness and the Belief Sharing Information Pattern

Serdar Yüksel, *Associate Member, IEEE*

Abstract—Solutions to decentralized stochastic optimization problems lead to recursions in which the state space enlarges with the time-horizon, thus leading to non-tractability of classical dynamic programming. A common joint information state supplied to each of the agents leads to a tractable recursion, as is evident in the one-step-delayed information sharing structure case or when deterministic nestedness in information holds when there is a causality relationship as in the case of partially nested information structure. However, communication requirements for such 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, 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 discuss the minimum amount of information exchange required to achieve such an information pattern for $k = 1$. The information exchange needed is generally strictly less than what is needed for deterministic nestedness and is zero whenever stochastic nestedness applies. In view of nestedness, we present a discussion on the monotone values of information channels.

Index Terms—Communication complexity, decentralized control, information-control structure (ICS), stochastic control, team decision theory.

I. INTRODUCTION

DECENTRALIZED stochastic control problems span a large venue of applications, however, these contain many issues yet to be further understood [29]. In a general decentralized system, different information sets are available to different decision makers who try to act on a common system towards a common goal as in team problems [25] or towards a variety of goals as in multi-criteria optimization problems [6]
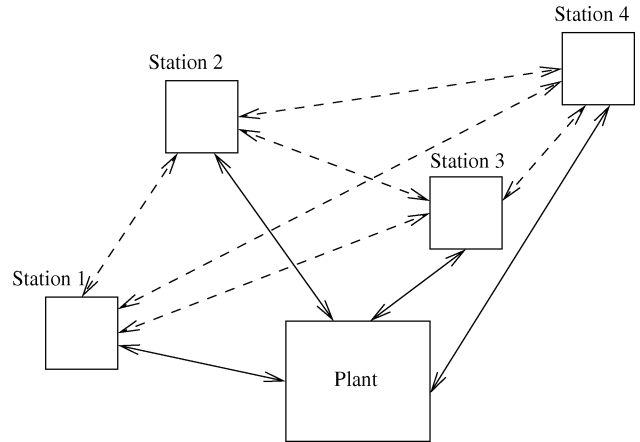

Fig. 1. Decentralized control system. Dashed lines depict communication links between the stations.

(Fig. 1). 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, as is evident in the Witsenhausen counterexample [40], and of very high complexity [23].

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

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, $\{DM^i, i = 1, 2, \ldots, L\}$. 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 $y_t^i \in \mathbb{Y}^i$ at the channel output. We refer to a policy Π^i as a sequence of control functions measurable with respect to the sigma-algebra generated by

$$I_t^i = \left\{ y_t^i, Z_t^i; y_{[0,t-1]}^i, u_{[0,t-1]}^i, Z_{[0,t-1]}^i \right\} \quad t \geq 1$$
$$I_0^i = \left\{ y_0^i, Z_0^i \right\}$$

with control actions $u_t^i \in \mathbb{U}^i$, with the notation for $t \geq 1$

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

The author is with Mathematics and Engineering Program, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6 Canada (e-mail: yuksel@mast.queensu.ca).

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 $\{u_t^i, t \geq 0\}$, $u_t^i \in \mathbb{U}^i$, and let a dynamical system and observation channels be described by the following discrete-time equations:

$$x_{t+1} = f\left(x_t, u_t^1, u_t^2, \ldots, u_t^L, w_t\right)$$
$$y_t^i = g^i\left(x_t, v_t^i\right)$$

for some measurable functions f, $\{g^i\}$, 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. The disturbance processes are uncorrelated across time.

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

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

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

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

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

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

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

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 [7], [16], [30]):

Centralized Information Structure: 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, ensuring the existence of a solution, is possible, the information structure is said to have a quasi-classical pattern; thus, under such a structure, the optimization problem is computationally feasible and the problem is said to be tractable. This structure includes the *one-step delayed observation sharing* information pattern (see [18] and [6]), 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: $Z_t^i = \{\mathbf{y}_{t-1}, \mathbf{u}_{t-1}\}$.

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; but it will be a result of our paper to show that such a characterization of information sets is too strong. It was observed by Radner [25] that a static LQG team problem with a non-nested information structure admits an optimal solution which is linear. The proof for this result follows from the observation that the team cost is convex in the joint strategies of the DMs, and it suffices to find the unique 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: $Z_t^j = \{y_t^i, \text{ if } \mathrm{DM}^i \to \mathrm{DM}^j\}$. Here the notation $\mathrm{DM}^i \to \mathrm{DM}^j$ denotes the fact that the actions of DM^i affects the information at DM^j (which is also known as signaling, see [43] for a review of signaling). 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 DMs. 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. Partially nested structures can also have a dynamic evolution [11], and as a special case, this includes the case where information propagation is faster than dynamics propagation, where in the above definition, delay is also considered [11], [26], [37].

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 [2], [6], [30].

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 [36] 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 [18]. The optimal control problems might also include different assumptions on the ordering information of control actions, see [34] and the references therein. A sequential design for decentralized optimization problems has been recently reported in [19]. In a parallel line of research related to the notion of partial nestedness, Voulgaris [37], Bamieh and Voulgaris [5] and Rotkowitz and Lall [27] have studied sufficiency conditions for tractability and convexity in optimal decentralized control problems.

An important related information pattern is the *k-step periodic information sharing pattern* studied by Yoshikawa [41] and Ooi *et al.* [21] with $Z_t^i = \{\mathbf{y}_{[t-k-(t \bmod k), t-(t \bmod k)]}, \mathbf{u}_{[t-k-(t \bmod k), t-(t \bmod k)]}\}$, 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 tractable information structure. We will discuss this pattern further in the paper, and provide an alternative derivation of the main results presented in [21] via Witsenhausen's equivalent model for discrete-stochastic control [39].

When the information structures are non-nested, controllers might choose to communicate via their control actions, that is might wish to pursue signaling. Different types of signaling can occur: signaling to learn the dynamics of the system, 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 or signaling what the agent's own future control actions will be, depending on the effects on the cost performance. 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 amounts of data noiselessly especially in a real-time fashion. In particular, it is impossible to exchange a real number noiselessly over a practical channel in finite time [10].

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*. Such a pattern minimizes the information exchanges required for tractability, without any loss of performance in comparison with its 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.

A brief summary of the rest of the paper is as follows. In the following section we review the topic of optimal control under partial observations as relevant to the topic of the paper and then we introduce in Section III the notion of stochastic nestedness with various examples. We then investigate the issue of a common information state and a sufficient statistic generated by the common information state in Section IV, where we discuss the Belief Sharing Information Pattern. We then discuss the information requirements leading to this information pattern. Finally, in Section V, we investigate various communication channels as observation channels and study their influence on nestedness, before presenting some concluding remarks in Section VI.

II. CONTROL OF A MARKOV CHAIN UNDER PARTIAL OBSERVATIONS

Many of the results in this paper will be based 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 (see for example [3]).

Consider a state process $\{x_t, y_t, t \geq 0, t \in \mathbb{Z}_+ \cup \{0\}\} \in \mathbb{X} \times \mathbb{Y}$. The time-relationship of the processes are given by the following equations:

$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = g(x_t, v_t).$$

Here, x_t is the state, $u_t \in \mathbb{U}$ is the control, $(w_t, v_t) \in \mathbb{W} \times \mathbb{V}$ are second order, zero-mean, i.i.d noise processes and w_t is independent of v_t. We also assume that the state noise w_t either has a probability mass function, or admits a probability measure which is absolutely continuous with respect to the Lebesgue measure; this will ensure that the probability measure admits a density function. Hence, the notation $p(x)$ will denote either the probability mass for discrete-valued spaces or probability density function for uncountable spaces. The controller only has causal access to the second component $\{y_t\}$ of the process. A policy $\{\Pi\}$ is measurable with respect to $\sigma(\{y_s, s \leq t\})$. We denote the observed history space as: $H_0 := \mathcal{P}, H_t =$

$H_{t-1} \times \mathbb{Y} \times \mathbb{U}$. Here $\mathcal{P}$, denotes the space of probability measures on $\mathbb{X}$. Hence, the set of (wide-sense) causal control policies are such that $P(u(h_t) \in \mathbb{U}|h_t) = 1, \forall h_t \in H_t$.

One could transform a partially observable Markov Decision Problem to a Fully Observed Markov Decision Problem via an enlargement of the sample space [31]. In particular, we obtain via the properties of total probability the following dynamical recursion:

$$\pi_t(x) := P\left(x_t = x | y_0^t, u_0^{t-1}\right)$$
$$= \frac{\sum_{\mathbb{X}} \pi_{t-1}(x_{t-1}) P(y_t|x_t) P(x_t|x_{t-1}, u_{t-1})}{\sum_{\mathbb{X}} \sum_{\mathbb{X}} \pi_{t-1}(x_{t-1}) P(y_t|x_t) P(x_t|x_{t-1}, u_{t-1})}.$$

Here, the summation needs to be exchanged with an integral in case the variables live in an uncountable space. The conditional measure process becomes a controlled Markov chain in $\mathcal{P}$. Let the cost function to be minimized be $\sum_{t=0}^{T-1} E_{x_0}^{\Pi}[c(x_t, u_t)]$, where $E_{x_0}^{\Pi}[]$ denotes the expectation over all sample paths with initial state given by x_0 under policy Π. We transform the system into a fully observed Markov model as follows. Define the new cost as $\tilde{c}(\pi, u) = \sum_{\mathbb{X}} c(x, u)\pi(x)$, $\pi \in \mathcal{P}$. The stochastic transition kernel q is given by

$$q(x, y|\pi, u) = \sum_{\mathbb{X}} P(x, y|x', u)\pi(x'), \quad \pi \in \mathcal{P}.$$

This kernel can be decomposed as $q(x, y|\pi, u) = P(y|\pi, u)P(x|\pi, u, y)$. The second term here is the filtering equation, mapping $(\pi, u, y) \in (\mathcal{P} \times \mathbb{U} \times \mathbb{Y})$ to $\mathcal{P}$. It follows that $(\mathcal{P}, \mathbb{U}, \mathcal{K}, \tilde{c})$ defines a completely observable controlled Markov process. Here, we have

$$\mathcal{K}(B|\pi, u) = \sum_{\mathbb{Y}} 1_{(P(.|\pi, u, y) \in B)} P(y|\pi, u), \quad \forall B \in \sigma(\mathcal{P})$$

with $1_{(.)}$ denoting the indicator function. As such, one can obtain the optimal solution by using the filtering equation as a sufficient statistic in a centralized setting, as Markov policies (policies that use the Markov state as their sufficient statistics) are optimal for control of Markov chains, under well-studied sufficiency conditions for the existence of optimal selectors [15].

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 the state and as the time horizon increases, the beliefs on the beliefs and so on, leading to a non-tractable optimization problem in a dynamic setting. This issue still remains even if one adopts a team-decision based approach through Witsenhausen's equivalent model.

III. 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. 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 evolving in $\mathbb{R}^n$ with the following description:

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

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

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

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 controller 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 $C^1 A^l B^2 \neq 0$ or $C^2 A^l B^1 \neq 0$ ([43], 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. 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 or partial nestedness for the convexity argument above to hold. We now proceed to define stochastic nestedness. In essence, the following show that it suffices that **relevant information on the state/optimal cost given what is already available at the receiver end** is to be exchanged.

B. 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 3.1: For measurable functions $f, g^i, i \in \{1, 2, \ldots, L\}$, consider a system described by

$$x_{t+1} = f(x_t, \mathbf{u}_t, w_t)$$
$$y_t^i = g^i(x_t, v_t^i), \quad i \in \{1, 2, \ldots, L\}.$$

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

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

forms a Markov chain

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

and

$$y_t^i = h_t(y_0^i)$$

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

Theorem 3.1: 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 $Q \geq 0, R > 0$. Consider an optimization problem with the objective functional to be minimized as

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

with the system dynamics

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

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.

Proof: First, let us consider the case with $L = 2$, that is with two Decision Makers. We refer to these as inner and outer DMs, such that the inner DM is DM^i, and outer DM is denoted by DM^o and

$$x_0 \leftrightarrow y_0^o \leftrightarrow y_0^i$$

forms a Markov chain and

$$y_t^i = h_t(y_0^i)$$

for some deterministic function h. Now, it should be noted that one can decompose the positive-definite matrix $R > 0$ as

$$R = \begin{bmatrix} R_{ii} & R_{io}/2 \\ R_{io}^T/2 & R_{oo} \end{bmatrix}$$

with $R_{ii} > 0$ and $R_{oo} > 0$. The proof for this argument follows from the fact that for some vector $v = [v^i \ v^0]$, it must be that $vRv^T > 0$, for all $v \neq 0$. Taking $v = v_1 = [v^i \ 0]$ and $v = v_2 = [0 \ v^o]$, the desired result follows. In the following, we express $\mathbf{u}_t^T R \mathbf{u}_t$ as $u_t^{iT} R_{ii} u_t^i + u_t^{iT} R_{io} u_t^o + u_t^{oT} R_{oo} u_t^o$, and adopt

a Stackelberg type approach to obtain the optimal team solution. Let the inner DM's decision rule be given by, for a sequence of functions $\gamma_t^i : \mathbb{Y}^{i(t+1)} \to \mathbb{U}^i$:

$$u_t^i = \gamma_t^i\left(y_{[0,t]}^i\right), \quad t \in \{0, 1, \ldots, T-1\}.$$

Let the outer DM's decision rule be given by, for a sequence of functions $\gamma_t^o : \mathbb{Y}^{o(t+1)} \times \mathbb{U}^{iT} \to \mathbb{U}^o$

$$u_t^o = \gamma_t^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right), \quad t \in \{0, 1, \ldots, T-1\}.$$

In the following we drop the subscripts in $\gamma^i()$ functions, as the domain of the functions will make the context clear. It follows that:

$$\inf_{\gamma^i, \gamma^o} E\left[\sum_{t=0}^{T-1} x_t^T Q x_t + \gamma^i\left(y_{[0,t]}^i\right)^T R_{ii} \gamma^i\left(y_{[0,t]}^i\right) \right.$$
$$+ \gamma^i\left(y_{[0,t]}^i\right)^T R_{io} \gamma^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right)$$
$$\left. + \gamma^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right)^T R_{oo} \gamma^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right) \right]$$

$$= \inf_{\gamma^i} E\left[\sum_{t=0}^{T-1} \gamma^i\left(y_{[0,t]}^i\right)^T R_{ii} \gamma^i\left(y_{[0,t]}^i\right) \right.$$
$$+ \inf_{\gamma^o} E\left[\sum_{t=0}^{T-1} x_t^T Q x_t \right.$$
$$+ \gamma^i\left(y_{[0,t]}^i\right)^T R_{io} \gamma^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right)$$
$$+ \gamma^o\left(y_{[0,t]}^o, u_{[0,T-1]}^i\right)^T R_{oo} \gamma^o$$
$$\left.\left. \times \left(y_{[0,t]}^o, u_{[0,T-1]}^i\right) \right] \right] \quad (2)$$

$$= \inf_{\gamma^i} E\left[\sum_{t=0}^{T-1} \gamma^i\left(y_{[0,t]}^i\right)^T R_{ii} \gamma^i\left(y_{[0,t]}^i\right) \right.$$
$$\left. + \inf_{\gamma^o} E\left[\sum_{t=0}^{T-1} c^o(P(x_t | I_t^o)) \right] \right]. \quad (3)$$

Here (2) follows from the fact that γ^o does not affect $\gamma^i(y_{[\cdot]}^i)$ since it does not affect y_t^i. Here, $c^o(P(x_t | I_t^o))$ denotes the fact that the cost is only a function of the belief of the outer DM on the state of the system. The problem reduces to a regular LQG formulation **for any** $\gamma^i(.)$ policy, as these are not affected by $\{u_t^o\}$ values, and can be treated as constants. In this case, the inner optimization problem $\inf_{\gamma^o} E[\sum_{t=0}^{T-1} c^o(P(x_t | I_t^o))]$ is a classical LQG optimization problem, and is convex in the outer DM's decision policies as $R_{oo} > 0$. Furthermore, the past applied control laws are known and hence the estimation error is independent of the control laws applied by the inner DM and the estimation law is always linear as the inner DM need not use signaling as given the information at the outer DM, the inner DM's observation is useless due to the Markov chain condition: $P(x_t | y_{[0,t]}^o, y_{[0,t]}^i, u_{[0,T-1]}^i) = P(x_t | y_{[0,t]}^o, u_{[0,T-1]}^i)$ for $t \geq 0$.

Hence, the problem is now reduced to an LQG problem given the outcomes under the strategy of $\gamma^i(.)$. Thus, one can obtain

the solutions for $\gamma^o(.)$ given $u^i_{[0,T-1]}$. It turns out that the solution is an affine controller following the classical LQG machinery, that is it is of the form:

$$u^o_t = K^o_t E\left[x_t | I^o_t\right] + L^o_{t,t} u^i_t + L^o_{t,t+1} u^i_{t+1} + \ldots L^o_{t,T-1} u^i_{T-1}$$

for all time stages $t \in \{0, 1, \ldots, T-1\}$ and for matrices K_t and $L_{(.,.)}$ of appropriate dimensions. For details of such a derivation the reader is referred to Theorem 2.1 of [42].

Now, one needs to optimize over $u^i_{[0,T-1]}$. With the linear controls of the outer DM, there is an equivalent optimization problem that the inner DM needs to solve. Let us denote the cost as $J(\gamma^i, \gamma^{o*}(\gamma^i))$, with $\gamma^{o*}(.)$ denoting the best response of the outer DM to the inner DM's policy. We now show that this cost is strictly convex in γ^i and hence in the control actions $u^i_{[0,T-1]}$. With the above solution for the outer DM, it follows that the state dynamics evolve as follows:

$$\begin{aligned} x_{t+1} = {} & Ax_t + B^i u^i_t + B^o \left(K^o_t x_t + L^o_{t,t} u^i_t + L^o_{t,t+1} u^i_{t+1} + \ldots \right. \\ & \left. + L^o_{t,T-1} u^i_{T-1} - K^o_t \left(x_t - E\left[x_t | I^o_t\right]\right)\right) + w_t. \quad (4) \end{aligned}$$

Here, $(x_t - E[x_t | I^o_t])$ is independent of the information available at DM^i. To see this, one could append the information at DM^i to I^o_t and realize that the estimate would not change. Hence, by the Projection Theorem, the estimation error would be orthogonal to the information available at DM^o and hence at DM^i. Furthermore $(x_t - E[x_t | I^o_t])$ is Gaussian. Hence, the inner DM solves another LQG optimization problem. Since $R_{ii} > 0$ is positive definite and the system described by (4) is linear, it follows that the cost function:

$$E\left[\sum_{t=0}^{T-1} x^T_t Q x_t + \mathbf{u}^T_t R \mathbf{u}_t | y^i_0\right]$$

is strictly convex in $u^i_{[0,T-1]}$ and thus there exists a unique solution.

Now, for the general case, we can regard the set of all DM's which can signal to DM^j as a single inner DM and apply the above analysis, for every decision maker, leading to a unique solution. Under such a nestedness condition, the optimal solutions will be linear.

In case there is no signaling, the problem reduces to the setups of Ho and Chu ([17], Section III) and Radner [25], which in turn is a static, convex team problem. In this case, the optimal solutions 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 actions. ◇

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 [2], Sandell and Athans [30] and Bismut [8]. In those works, ϵ-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 ob-

servation, 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 $\mathbb{R}$, for any ϵ, there exists an $n \in \mathbb{Z}$ such that an n-decimal representation which is at most at an ϵ distance from any real number in a compact set is possible. Therefore, if one is to represent a finite dimensional control variable in a compact set $\mathbb{U} \subset \mathbb{R}^r$, and a finite dimensional observation variable in a compact set $\mathbb{Y} \subset \mathbb{R}^m$; these signals can be represented with an arbitrarily small error by a single rational number. Thus, one may embed in this number, the ϵ-approximate decimal expansions of the numbers to be represented; thus, leading to a total of $n(m + r)$ decimal letters, by allocating the most significant nr letters for the control variable.

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 [30].

In our setup, the resulting policy is optimal (and not only ϵ-optimal), and unlike the setups of [30] and [8], is applicable to cases where (i) the control policy is discontinuous, or (ii) the state space has finite cardinality (hence arbitrarily precise representation of two signals is not possible via encoding these into one-signal since there is only finite information that can be transmitted in one signal), or (iii) the observation and control action sets are not compact, eliminating the possibility of approximation with a uniform error as discussed earlier. These are some conditions under which the assumptions of [30] and [8] are not applicable.

The applicability of the above scenarios under the stochastically 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, we introduce a pattern, where further assumptions on the cost function will be imposed to weaken the stochastic nestedness condition.

D. Stochastically Decoupled Information Structure

We define a stochastically decoupled information structure as follows.

Definition 3.2: Let a state be explicitly represented by its individual components $\mathbf{x} = \begin{bmatrix} x^1 & x^2 & \ldots & x^L \end{bmatrix}^T$, and evolve under the following dynamics:

$$\begin{aligned} x^i_{t+1} &= f^i\left(x^i_t, u^i_t, w^i_t\right), \quad i \in \{1, 2, \ldots, L\} \\ z_{t+1} &= f\left(z_t, u^1_t, u^2_t, \ldots, u^L_t\right) \\ y^i_t &= g^i\left(x^i_t, v^i_t\right), \quad i \in \{1, 2, \ldots, L\} \end{aligned}$$

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 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

$$\tilde{y}_t^i = \left[y_t^i, z_t \right], \quad i \in \{1, 2, \ldots, L\}.$$

If the information available at each controller is such that

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

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

Theorem 3.2: Let there be an optimization problem with the objective to be minimized as

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

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.

Proof: Toward the proof, we first provide a lower bound and then show that the optimal policies under centralized structure leading to the lower bound is attainable with decentralized controllers under the nestedness condition. Let $\mathbf{x} = \{x^1, x^2, \ldots, x^L\}$. Suppose each of the DMs has access to the other DMs' total information, that is the information structure is centralized, that is all have access to $\mathbf{I}_t$. In this case, the sufficient statistic for the centralized controller is equal to

$$P(\mathbf{x}_t | \tilde{\mathbf{y}}_t, \mathbf{I}_{t-1})$$
$$= \frac{P(\mathbf{x}_t, \tilde{\mathbf{y}}_t | \mathbf{I}_{t-1})}{\sum_{\mathbb{X}} P(\mathbf{x}_t, \tilde{\mathbf{y}}_t | \mathbf{I}_{t-1})}$$
$$= \frac{P(\mathbf{x}_t, \mathbf{y}_t, z_t | \mathbf{I}_{t-1})}{\sum_{\mathbb{X}} P(\mathbf{x}_t, \mathbf{y}_t, z_t | \mathbf{I}_{t-1})}$$
$$= \frac{P(z_t | \mathbf{I}_{t-1}) P(\mathbf{x}_t, \mathbf{y}_t | \mathbf{I}_{t-1})}{\sum_{\mathbb{X}} P(z_t | \mathbf{I}_{t-1}) P(\mathbf{x}_t, \mathbf{y}_t | \mathbf{I}_{t-1})}$$
$$= \frac{P(z_t | \mathbf{I}_{t-1}) P(\mathbf{x}_t, \mathbf{y}_t | \mathbf{I}_{t-1})}{P(z_t | \mathbf{I}_{t-1}) \sum_{\mathbb{X}} P(\mathbf{x}_t, \mathbf{y}_t | \mathbf{I}_{t-1})}$$
$$= P(\mathbf{x}_t | \mathbf{y}_t, \mathbf{I}_{t-1})$$
$$= \prod_{i=1}^{L} P \left(x_t^i | \mathbf{y}_t, x_t^1, x_t^2, \ldots, x_t^{i-1}, \mathbf{I}_{t-1} \right)$$
$$= \prod_{i=1}^{L} P \left(x_t^i | \mathbf{y}_t, \left(F_t^1 \left(x_0^1, w_{[0,t-1]}^1, \mathbf{I}_{t-1} \right) \right), \ldots, \right.$$
$$\left. \left(F_t^{i-1} \left(x_0^{i-1}, w_{[0,t-1]}^{i-1}, \mathbf{I}_{t-1} \right) \right), \mathbf{I}_{t-1} \right)$$
$$= \prod_{i=1}^{L} P \left(x_t^i | y_t^i, I_{t-1}^i \right). \tag{5}$$

Here, (5) follows from the Markov chain hypothesis of the theorem and the rest of the arguments uses the properties of total probability. In the above, we write $x_t^i = F_t^i(x_0^i, w_{[0,t-1]}^i, \mathbf{I}_{t-1}))$, to express the explicit dependence on the variables. Hence, the optimal centralized control policy will act on each of the states x^i, $i \in \{1, 2, \ldots, L\}$ using some local information.

Since the cost function is also decoupled, it turns out that one can write the dynamic programming recursions as L decoupled optimization problems. The cost can be written as a summation of decoupled systems. The optimal centralized policy is implementable under a decentralized structure, such that the dynamic programming can be carried out locally.   ◇

An example with relaxed conditions which do obey Markov chain conditions is the following. Consider a dynamical system described by

$$\begin{aligned} x_{t+1}^1 &= a_1 x_t^1 + u_t^1 + w_t^1 \\ x_{t+1}^2 &= a_2 x_t^2 + u_t^2 + w_t^2 \\ x_{t+1}^3 &= a_3 x_t^3 + u_t^1 + u_t^2 + w_t^3 \\ y_t^1 &= \left(x_t^1 + v_t^1, x_t^2 + v_t^2 + v_t^{21}, x_t^3 + v_t^{31} \right) \\ y_t^2 &= \left(x_t^1 + v_t^1 + v_t^{12}, x_t^2 + v_t^2, x_t^3 + v_t^{32} \right) \end{aligned}$$

where the goal is the minimization of

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

with $\rho_1, \rho_2 > 0$ constants. The control actions are measurable with the sigma-algebra generated by their causal observations and past controls: $I_t^i = \{y_t^i, I_{t-1}^i\}$, with $I_0^i = y_0^i$. This system has a non-classical information structure, as controller 1 affects the observation at controller 2, but controller 2 cannot recover the information at controller 1. In the above, the third state acts as a communications medium between the controllers. However, the communication channel is not needed by the Markov chain condition. The optimal policy is linear.

In the following section we investigate the cases where there is no apriori nestedness and we evaluate the information requirements needed to obtain a quasi-classical information structure, basically by making the information structure stochastically nested.

IV. 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 ([34], [39]) 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: $I^c_{ks-1} = \{\mathbf{y}_{[0,ks-1]}, \mathbf{u}_{[0,ks-1]}\}$ and for $I^c_0 = \{P(x_0)\}$, 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 $\{u^i_t = \bar{u}^i_s(y^i_{[ks,t]}, I^c_{ks-1})\}$ 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 $y^i_{[ks+1,t]}$, under $\bar{u}^i_s(y^i_{[ks,t]}, I^c_{ks-1})$, 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 $\bar{\mathbf{u}}_s(., I^c_{ks-1})$ is the joint team decision rule mapping I^c_{ks-1} into a space of action vectors: $\{u^i(I^c_{ks-1}, y^i_{[ks,t]}), i \in \{1, 2 \ldots, L\}, t \in \{ks, ks+1, \ldots, k(s+1) - 1\}\}$. In this case, the cost function is also modified as

$$J_{x_0}(\mathbf{\Pi}) = E^{\mathbf{\Pi}}_{x_0}\left[\sum_{s=0}^{\frac{T}{k}-1} \bar{c}\left(\bar{\mathbf{u}}_s\left(., I^c_{ks-1}\right), \bar{x}_s\right)\right]$$

with

$$\bar{c}\left(\bar{\mathbf{u}}_s\left(., I^c_{ks-1}\right), \bar{x}_s\right) = E^{\mathbf{\Pi}}_{\bar{x}_s}\left[\sum_{t=ks}^{k(s+1)-1} c(x_t, \mathbf{u}_t)\right].$$

Lemma 4.1: Consider the decentralized system setup in Section I-A, with the observation noise processes being independent. Let I^c_t 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, $\{\bar{x}_s := x_{ks}, \bar{\mathbf{u}}_s(., I^c_{ks-1}), s \geq 0\}$ forms a Controlled Markov chain.

Proof: Let $\bar{\mathbf{u}}_{[0,s]} = \{\bar{\mathbf{u}}_m(., I^c_{km-1}), 0 \leq m \leq s\}$. It follows that:

$$P\left(\bar{x}_{s+1}|\bar{\mathbf{u}}_{[0,s]}, \bar{x}_{[0,s]}\right)$$
$$= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}}$$
$$P\left(\bar{x}_{s+1}, x_{[ks,k(s+1)-1]}, \mathbf{y}_{[ks,k(s+1)-1]}|\bar{\mathbf{u}}_{[0,s]}, \bar{x}_{[0,s]}\right)$$
$$= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}}\left(\prod_{m=ks}^{k(s+1)-1}\right.$$
$$P\left(x_{m+1}|x_m, \bar{\mathbf{u}}_s\left(\mathbf{y}_{[ks,m]}, I^c_{ks-1}\right),\right.$$
$$\left.\bar{x}_s, \bar{\mathbf{u}}_{[0,s-1]}, \bar{x}_{[0,s-1]}\right)$$
$$\left. P\left(y^1_m|x_m\right) P\left(y^2_m|x_m\right) \ldots P\left(y^L_m|x_m\right)\right)$$
$$= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}}\left(\prod_{m=ks}^{k(s+1)-1}\right.$$
$$P\left(x_{m+1}|x_m, \bar{\mathbf{u}}_s\left(\mathbf{y}_{[ks,m]}, I^c_{ks-1}\right)\right)$$
$$\left. P\left(y^1_m|x_m\right) P\left(y^2_m|x_m\right) \ldots P\left(y^L_m|x_m\right)\right)$$
$$= P\left(\bar{x}_{s+1}|\bar{\mathbf{u}}_s\left(., I^c_{ks-1}\right), \bar{x}_s\right). \qquad (6)$$
$$\diamond$$

In view of the above, we now present a result on a separation property. We note that the following also has been studied in [21]. We present a shorter proof, using the result above directly.

Lemma 4.2: Let I^c_t be a common information vector supplied to the DMs regularly at every k time steps. There is no loss in performance if I^c_{ks-1} is replaced by $P(\bar{x}_s|I^c_{ks-1})$.

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

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

with

$$\tilde{c}(\bar{\mathbf{u}}_s, \bar{x}_s) = E^{\mathbf{\Pi}}_{\bar{x}_s}\left[\sum_{t=ks}^{k(s+1)-1} c(x_t, \mathbf{u}_t)\right].$$

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 [31] as discussed in Section II. Hence $P(\bar{x}_s|I^c_{ks-1})$ acts as a sufficient statistic (See also [3]) [32]. $\diamond$

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 4.1: 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:

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

writes as

$$\frac{P\left(\bar{x}_{s+1}, (\mathbf{y},\mathbf{u})_{[ks,k(s+1)-1]}|\pi_s\right)}{\sum_{\bar{x}_{s+1}} P\left(\bar{x}_{s+1}, (\mathbf{y},\mathbf{u})_{[ks,k(s+1)-1]}|\pi_s\right)}$$
$$= \frac{\sum_{x_{[ks,k(s+1)-1]}} P\left(\bar{x}_{s+1}, (x,\mathbf{y},\mathbf{u})_{[ks,k(s+1)-1]}|\pi_s\right)}{\sum_{x_{[ks,k(s+1)-1]}, \bar{x}_{s+1}} P\left(\bar{x}_{s+1}, (x,\mathbf{y},\mathbf{u})_{[ks,k(s+1)-1]}|\pi_s\right)}. \qquad (7)$$

We now express the numerator in (7) more explicitly as

$$\sum_{x_{k(s+1)-1}} \left(P\left(x_{k(s+1)}|x_{k(s+1)-1}, \mathbf{u}_{k(s+1)-1}\right) \right.$$

$$\left(\prod_{l=1}^{L} P\left(y^l_{k(s+1)-1}|x_{k(s+1)-1}\right)\right)$$

$$\sum_{x_{k(s+1)-2}} \left(P\left(x_{k(s+1)-1}|x_{k(s+1)-2}, \mathbf{u}_{k(s+1)-2}\right) \right.$$

$$\left(\prod_{l=1}^{L} P(y^l_{k(s+1)-2}|x_{k(s+1)-2})\right)$$

$$\cdots \quad \cdots \quad \cdots$$

$$\sum_{x_{ks+1}} \left(P(x_{ks+2}|x_{ks+1}, \mathbf{u}_{ks+1}) \right.$$

$$\left(\prod_{l=1}^{L} P\left(y^l_{ks+1}|x_{ks+1}\right)\right)$$

$$\sum_{x_{ks}} \left(P(x_{ks+1}|x_{ks}, \mathbf{u}_{ks}) \right.$$

$$\left(\prod_{l=1}^{L} P\left(y^l_{ks}|x_{ks}\right)\right)$$

$$\left. \left. \left. P\left(x_{ks}|I^c_{ks-1}\right) \right) \right) \cdots \right) \right). \tag{8}$$

As such, if $k > 1$, 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. We will discuss this further, while considering the belief propagation algorithm in the following.

A. Belief Sharing Through an Iterative Belief Propagation Algorithm

In this section, we present an iterative approach for the belief sharing pattern. This will be an extension of the Belief Propagation algorithm, which is a local message exchange algorithm among several remote DMs/sensors located on vertices in a graph through the edges [14]. In a belief propagation algorithm, each DM has an apriori belief about the state of the system. With local observations, the DMs generate an aposteriori information, and then exchange these with the other DMs. Belief propagation reaches to the correct measure (one that would be achieved under a centralized information structure) if there are no cycles, and hence if the topology of the communication graph between the DMs forms a tree.

We briefly summarize now, how belief exchanges can be performed to achieve the belief sharing pattern in view of (7) and (8). We first consider a cycle-free network of DMs. Consider Fig. 2. For $m \geq 0$, suppose the DMs have an agreement on x_{ks+m}, before the DMs
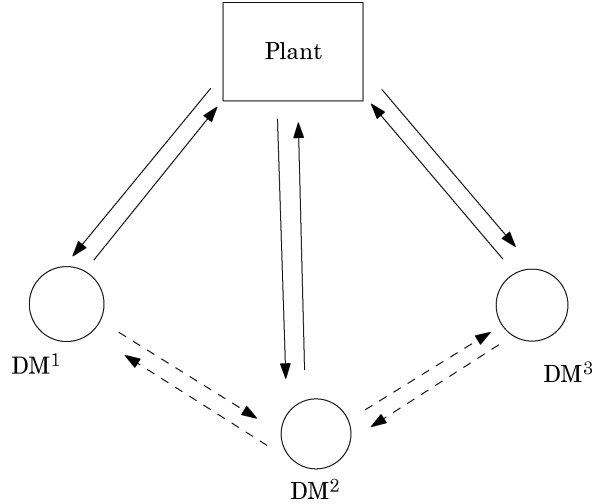

Fig. 2. Belief propagation converges to the true conditional measure in a finite number of iterations in a cycle-free network.

receive local observations, that is, they all have access to $P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|I^c_{ks+m-1}))$ for all x_{ks+m} values. Once DMs observe local measurements y^i_{ks+m}, first, DM^1 sends to DM^2 its belief on x_{ks+m}: $P(x_{ks+m}|y^1_{ks+m}, P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|I^c_{ks+m-1})))$. Thus, DM^1 sends its belief about the state of the system at time $ks + m$ to DM^2, for all possible x_{ks+m} values. DM^2 then for all x_{ks+m} values, generates its own belief on x_{ks+m}, using its local information, to obtain (only using the belief sent by DM^1 and its own information, together with the prior belief)

$$P\left(x_{ks+m}|(y^1, y^2)_{ks+m}, P(x_{ks+m}|.)\right)$$
$$= \frac{P(y^2|x_{ks+m})P\left(x_{ks+m}|y^1_{ks+m}, P(x_{ks+m}|.)\right)}{\sum_{x_{ks+m}} P(y^2|x_{ks+m})P\left(x_{ks+m}|y^1_{ks+m}, P(x_{ks+m}|.)\right)}$$

where $P(x_{ks+m}|.)$ stands for $P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|I^c_{ks+m-1}))$ (and is the prior belief on x_{ks+m} using the information from the previous time stage). DM^2 then sends this information to DM^3, who upon receiving the information from DM^2 generates the final conditional measure

$$P\left(x_{ks+m}|(y^1, y^2, y^3)_{ks+m}, P(x_{ks+m}|.)\right).$$

In the next iteration DM^3 sends this belief information back to DM^2, and finally to DM^1. Upon such a forward and backward sweeping, all the DMs have access to the correct joint belief on the state of the system, which we denote by $P(x_{ks+m}|I^c_{ks+m})$.

Now that all the DMs have the same belief on the state, they share all their control actions that they applied at time $ks + m$. With this information, all of the DMs have access to $P(x_{ks+m+1}|\mathbf{u}_{ks+m}, P(x_{ks+m}|I^l_{ks+m}))$. This now acts as a common prior for the next iteration in this algorithm. DM^1 once again sends its belief on the state x_{ks+m+1} and the iteration continues. As such, upon such a double forward and backward sweeping (first for the beliefs, then for the control actions), all DMs have access to the correct joint belief on the state of the system. Once this convergence occurs for all time stages $ks, ks + 1, ks + 2\ldots, ks + s - 1$, the DMs start the iteration for the next time stage, eventually agreeing on the conditional
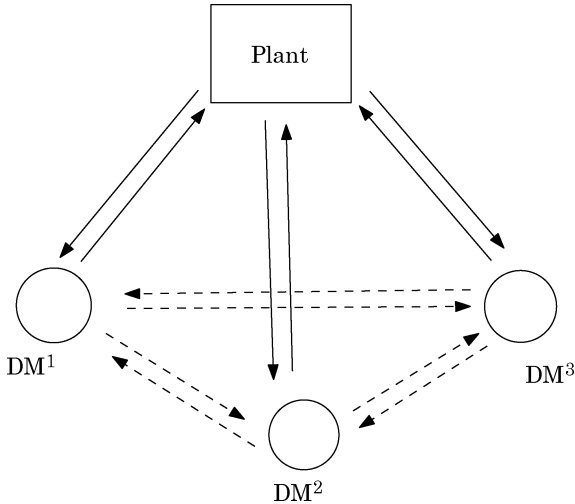
Fig. 3. When there are cycles in the communication topology, in our setting, one needs to pick an ordering apriori and avoid cycles in the ordering until convergence is reached.

measure for the time stage $k(s + 1) - 1$ and subsequently having a common prior on $k(s + 1)$.

Consider Fig. 3. In this case, as there is a cycle (loop) in the communication topology, it is essential that the DMs have a predefined route for information exchange for convergence to the actual joint belief: If all DMs talk to their neighbors, then one DM might have more than one paths to send a message to another DM and hence her opinions might be incorrectly emphasized by a virtue of the cycle. For a large system, however, for scalability considerations, it is more reasonable to apply the belief propagation algorithm, at the expense of reaching an incorrect agreement. In such a case of incorrect agreement, one needs to apply the dynamic programming algorithm by considering the probability of such agreements and generating policies accordingly. This requires further future research. For a discussion of the convergence of loopy belief propagation algorithms, with no control and in a static setting, we refer the reader to [20], [33] and the references therein.

If the sharing occurs at every stage (that is, $k = 1$), then control actions are not shared, since every DM can generate every DM's control action given the joint belief, under deterministic policies. In this case, we only need to apply the original belief sharing algorithm, which, in the cycle-free case only requires one forward and backward message passing.

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

B. 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 4.1 and 4.2, 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 (that is the joint belief on the state) and a dynamic programming recursion is possible?

The information will be measured by the number of bits; when the coding is variable-rate, information is measured by the average number of bits needed to be exchanged among the decision makers; whereas when the coding scheme is fixed-rate, information is measured by the actual number of bits that is exchanged for any given time stage $t \geq 0$.

Let us now introduce two standard information-theoretic notions, namely entropy and mutual information. *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-alphabet valued 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 [13]. 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 function of I_t^i with its range in $\mathcal{W}^{i,j} = \{W^{i,j}(0), W^{i,j}(2), \ldots, W^{i,j}(T-1)\}$, the set of entropy-coder variable-rate or fixed-rate codewords for communication from DM^i to DM^j. Hence, at each time stage t, DM^i sends $R^{i,j}(t)$ bits over an external channel to DM^j. Let $\mathcal{R}(t) = \{R^{i,j}(t), i \neq j, i, j \in 1, 2, \ldots, L\}$ such that belief sharing is possible. Define $R(t) := \inf_{\mathcal{R}(t)}\{\sum_{i=1}^{L}\sum_{j=1, j\neq i}^{L} R^{i,j}(t)\}$, such that belief sharing is realized for a given time stage t. We wish to obtain such $R(t)$ values.

The multi-terminal source coding theorems [13], although insightful, are not always applicable for a real-time setting, as the asymptotic partitioning arguments in classical information theory [13] do not apply. In a control context, however, one method to achieve the information theoretic bounds is via binning; see [43], [44] for discussions on binning in a decentralized control context and [24] for a discussion on binning in a general communications context.

1) One-Stage Belief Sharing Pattern: 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 $p(\bar{x}_t | I_t^1, I_t^2)$, 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 $k = 1$.

We note that, when control policies are deterministic, the actions can uniquely be identified by both DMs. As such, control signals need not be exchanged.

Theorem 4.1: To achieve the one-stage belief sharing information pattern, the following rate region is achievable using fixed-rate codes:

$$
\begin{aligned}
\mathcal{R}(t) = \Big\{ &(R^{i,j}, R^{j,i}) : R^{i,j} = \lceil \log_2\left(|\mathcal{S}_t|\right)\rceil, \\
&R^{j,i} = \Big\lceil \log_2\left(\sup_{\pi^i} |\mathcal{S}_{\pi^i, t}|\right)\Big\rceil \\
&\mathcal{S}_t = \big\{ \pi^i = P\left(x_t | y_t^i = y^i, P(.|.)\right) : \\
&\quad P\left(y_t^i = y^i | P(.|.)\right) > 0, y^i \in \mathbb{Y}^i \big\} \\
&\mathcal{S}_{\pi^i, t} = \Big\{ P\left(x_t | y_t^j = y^j, \pi^i, P(.|.)\right) : \\
&\quad P\left(y_t^j = y^j | \pi^i, P(.|.)\right) > 0, y^j \in \mathbb{Y}^j \big\} \Big\}
\end{aligned}
$$

where $P(.|.)$ denotes $P(x_t | I_{t-1}^c)$.

Proof: Proof follows from a graph based combinatorial extension of the binning arguments used in [44] in the context of decentralized communication for control systems (see [24] for an asymptotic context). First DM^2 learns the conditional belief of DM^1; and then computes the joint belief using the algorithm discussed in the previous section. DM^2 then sends the set of all distinct possible aposteriori joint beliefs to DM^1 consistent with the belief at DM^1. ◇

A discussion is also available when the communicate rate is measured by the average number of bits.

Theorem 4.2: 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^i from DM^j, $i, j \in \{1, 2\}$, $i \neq j$ and in the opposite direction are

$$R^{j,i} \geq H\left(P\left(x_t|I_{t-1}^c, y_t^i, y_t^j\right) | P\left(x_t|I_{t-1}^c\right), y_t^i\right)$$
$$R^{i,j} \geq H\left(P\left(x_t|I_{t-1}^c, y_t^i, y_t^j\right) | P\left(x_t|I_{t-1}^c\right), y_t^j, Z_t^i\right)$$

where Z_t^i is the variable sent to DM^1 from DM^2.

Proof: Let Z_t^1 be the random variable that is transmitted from DM^2 to DM^1. In the following, let $\|.\|$ denote the total variation norm. It follows that:

$$
\begin{aligned}
R^{2,1} \geq & \inf\left\{H\left(Z_t^1\right) : \left\|P\left(x_t|I_{t-1}^c, Z_t^1, y_t^1\right)\right.\right.\\
& \left.\left.- P\left(x_t|I_{t-1}^c, y_t^1, y_t^2\right)\right\| = 0\right\}\\
\geq & \inf\left\{H\left(Z_t^1|P\left(x_t|I_{t-1}^c\right), y_t^1\right) :\right.\\
& \left.\left\|P\left(x_t|I_{t-1}^c, Z_t^1, y_t^1\right) - P\left(x_t|I_{t-1}^c, y_t^1, y_t^2\right)\right\| = 0\right\}\\
\geq & \inf\left\{H\left(Z_t^1|P\left(x_t|I_{t-1}^c\right), y_t^1\right)\right.\\
& - H\left(Z_t^1|P\left(x_t|I_{t-1}^c, y_t^2, y_t^1\right), P\left(x_t|I_{t-1}^c\right), y_t^1\right) :\\
& \left.\left\|P\left(x_t|I_{t-1}^c, Z_t^1, y_t^1\right) - P\left(x_t|I_{t-1}^c, y_t^1, y_t^2\right)\right\| = 0\right\}\\
= & \inf\left\{I\left(Z_t^1; P\left(x_t|I_{t-1}^c, y_t^2, y_t^1\right) | P\left(x_t|I_{t-1}^c\right), y_t^1\right) :\right.\\
& \left.\left\|P\left(x_t|I_{t-1}^c, Z_t^1, y_t^1\right) - P\left(x_t|I_{t-1}^c, y_t^1, y_t^2\right)\right\| = 0\right\}\\
= & \inf\left\{H\left(P\left(x_t|I_{t-1}^c, y_t^2, y_t^1\right) | P\left(x_t|I_{t-1}^c\right), y_t^1\right)\right.\\
& - H\left(P\left(x_t|I_{t-1}^c, y_t^2, y_t^1\right) | P\left(x_t|I_{t-1}^c\right), Z_t^1, y_t^1\right) :\\
& \left.\left\|P\left(x_t|I_{t-1}^c, Z_t^1, y_t^1\right) - P\left(x_t|I_{t-1}^c, y_t^1, y_t^2\right)\right\| = 0\right\}\\
= & H\left(P\left(x_t|I_{t-1}^c, y_t^2, y_t^1\right) | P\left(x_t|I_{t-1}^c\right), y_t^1\right).
\end{aligned}
$$

Here the last two steps follows from the observation that given Z^1 the conditional measures are identical, and the constraint is independent of $H(P(x_t|I_{t-1}^c, y_t^1, y_t^1)|P(x_t|I_{t-1}^c), y_t^1)$. A parallel discussion applies to the reverse direction. ◇

We note that the information rate needed is less than one needed for achieving the centralized information pattern. By the above argument, one would need $R^{i,j} \geq H(y_t^i|y_t^j, I_{t-1}^c)$ for the centralized information pattern as a lower bound. The entropy of the conditional measure is at most as much as the entropy of the observed variable. This is because, different outputs may lead to the same values for $P(y_t^2 = y|x_t, I_{t-1}^c)$. Hence, we have the following corollary to Theorem 4.2.

Corollary 4.1: 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.

Proof: We only provide a discussion for the lower bounds. Following the proof of Theorem 4.2, the required rate for the one-step delay sharing pattern would be lower bounded by $H(y_t^2|y_t^1, I_{t-1}^c)$. It follows that:

$$
\begin{aligned}
& H\left(y_t^2|y_t^1, I_{t-1}^c\right)\\
& = H\left(y_t^2|y_t^1, P\left(x_t|I_{t-1}^c\right)\right)\\
& \geq H\left(P\left(x_t|P\left(x_t|I_{t-1}^c\right), y_t^1, y_t^2\right) | y_t^1, P\left(x_t|I_{t-1}^c\right)\right).
\end{aligned}
$$

The last argument follows from the following. Let A, B be defined on a common probability space and take values in a finite set and let κ be a deterministic function of A and B. Then, $H(A|B) = H(\kappa(A, B)|A, B) + H(A|B) = H(\kappa(A, B), A|B) = H(A|\kappa(A, B), B) + H(\kappa(A, B)|B) \geq H(\kappa(A, B)|B)$. Here, we use the fact that $H(\kappa(A, B)|A, B) = 0$, and use the properties of joint entropy. Now, since $P(x_t|P(x_t|I_{t-1}^c), y_t^1, y_t^2)$ is a deterministic function of y_t^2, given $p(x_t|I_{t-1}^c), y_t^1$, as a result of the filtering equation, the desired conclusion follows. ◇

We may also obtain an upper bound on the communication rates for a two decision maker setting for variable-rate schemes. The result is intuitive and its proof is omitted. In the following, we assume that entropy bound is tight for variable-rate coding.

Proposition 4.1: To achieve the one-stage belief sharing information pattern, an upper bound on the minimum average amount of bits to be transmitted to DM^i from DM^j, $i, j \in \{1, 2\}$, $i \neq j$, is given by

$$R^{j,i} \leq \min\left\{H\left(\zeta\left(y_t^j, P(.|.)\right) | P(.|.)\right) : P\left(x_t|P(.|.), y_t^1, y_t^2\right)\right.$$
$$\left. = P\left(x_t|P(.|.), y_t^i, \zeta\left(y_t^j, P(.|.)\right)\right)\right\}$$

where $P(.|.)$ denotes $P(x_t|I_{t-1}^c)$.

Remark: In the setup considered, the goal is that each DM can compute the joint belief. We note here the interesting discussion between decentralized computation and communication provided in Csiszar and Körner ([12], Thm. 4.6) and Orlitsky and Roche [22]. However, the setting presented in these works assumes an infinite copy of messages to be encoded and functions to be computed, which is not applicable in a real-time setting. ◇

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.

Proposition 4.2: To achieve the belief sharing information structure, a lower bound on the minimum average amount of bits that needs to be supplied to some DM^i, $\{i = 1, 2, \ldots, L\}$ is given by

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

C. Stochastic Nestedness in a Dynamic Setting

We now can extend the notion of stochastic nestedness to a dynamic context. Let, as described earlier, at times ks, $k \in \mathbb{Z}_+ \cup \{0\}$ and T divisible by k, $s \in \mathbb{Z}_+$, the decision makers share information so that they agree on a common prior with I_{ks-1}^c. In this case, until the next observation instant $t = k(s+1) - 1$

we regard the individual decision functions specific to DM^i as $\{u_t^i = \bar{u}_s^i(y_{[ks,t]}^i, I_{ks-1}^c)\}$.

Definition 4.2: For measurable functions f, g^i, $i \in \{1, 2, \ldots, L\}$, consider a system described by

$$x_{t+1} = f(x_t, \mathbf{u}_t, w_t)$$
$$y_t^i = g^i(x_t, v_t^i), \quad i \in \{1, 2, \ldots, L\}.$$

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

$$\bar{x}_s \leftrightarrow \left(y_{ks}^j, I_{ks-1}^c\right) \leftrightarrow \left(y_{ks}^i, I_{ks-1}^c\right)$$

forms a Markov chain

$$I_t^j = \left\{y_{[ks,t]}^j, u_{[ks,k(s+1)-1]}^i, I_{ks-1}^c\right\}$$

and

$$y_t^i = h_t\left(y_{ks}^i, I_{ks-1}^c\right)$$

where h_t is a deterministic function for $t \in \{ks, ks+1, \ldots, k(s+1)-1\}$, then the information structure is stochastically nested with a k-stage belief sharing pattern.

Theorem 4.3: 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 $Q \geq 0, R > 0$. Consider an optimization problem with the objective to be minimized as

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

with the system dynamics

$$x_{t+1} = A x_t + \sum_{j=1}^{L} B^j u_t^j + w_t$$
$$y_t^i = C^i x_t + v_t^i, \quad 1 \leq i \leq L \quad (9)$$

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 with a k-stage belief sharing pattern. In this case, the optimal control laws are linear.

Proof: Proof follows from dynamic programming. We can view all of the time-slots in $[ks, k(s+1)-1]$ as a single stage, and for each stage invoke Theorem 3.1, as the newly constructed process will be Markov, even when the observation noises might be correlated. Iteratively applying the algorithm backwards, we observe that the quadratic nature of the cost functions and linearity of the control policies are preserved. ◇

D. 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 4.3: Consider the case in which the observation channels for each of the DMs have zero capacity. In this case, as

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

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

Proof: By Theorem 4.2, a lower bound on the rate is

$$H\left(P\left(x_t | I_{t-1}^c, y_t^1, y_t^2\right) | P\left(x_t | I_{t-1}^c\right), y_t^1\right) = 0.$$

This rate bound is tight, since by Theorem 4.1, the rate achievable by a fixed-rate scheme is zero as well. Same discussion applies to the more than two DM cases, by Proposition 4.2. ◇

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:

Proposition 4.4: Consider the case in $x \leftrightarrow y^1 \leftrightarrow y^2$. There is no further information exchange from DM^1 to DM^2 that is needed for the belief-sharing pattern.

Proof: This follows from:

$$H\left(P\left(x_t | I_{t-1}^c, y_t^1, y_t^2\right) | P\left(x_t | I_{t-1}^c\right), y_t^1\right) = 0$$

and that this rate bound is tight. ◇

V. MONOTONE VALUE OF INFORMATION CHANNELS AND APPLICATIONS OF THE INFORMATION REQUIREMENTS

In this section, we discuss the value of information channels in decentralized control, in view of stochastic nestedness. Recall the definition of the Information Structures and the observation channels as stochastic kernels in Section I-A. Let us consider two Decision Makers individually, let each of the decision makers apply their controls under information structures ISi, $i = 1, 2$. We make the following definitions.

Definition 5.1: An information structure IS2 is deterministically sequentially degraded with respect to another one, IS1, if the observation variables are such that, for every $t \in \{0, 1, \ldots, T-1\}$, there exists a deterministic function $f_t : \mathbb{Y}^{t+1} \rightarrow \mathbb{Y}^{t+1}$ such that $y_{[0,t]}^2 = f(y_{[0,t]}^1), 0 \leq t \leq T-1$. ◇

Note that, in the above case, the information at DM^2 is deterministically nested in DM^1, that is the sigma-algebra generated

by the variables are at DM^2 is a subset of that at DM^1. This is the requirement needed in (deterministic) nestedness or (deterministic) partial nestedness [16].

Definition 5.2: An information structure IS2 is stochastically sequentially (or physically) degraded with respect to another one, IS1, if the observation variables are such that

$$x_t \leftrightarrow y_t^1 \leftrightarrow y_t^2, \quad 0 \leq t \leq T-1$$

forms a Markov chain.   ◇

In this case, note that the information at DM^2 is stochastically nested in DM^1, provided the additional requirements on the availability of control signals discussed in Section III. This is the requirement needed for stochastic nestedness presented in this paper. The name *physically degraded* comes from the information theory literature [13].

Definition 5.3: An information structure IS2 is weakly stochastically degraded with respect to another one, IS1, if the observation variables are such that, there exists an auxiliary random variable y'

$$P\left(y_t^2 = y | x_t\right) = \sum_{\mathbb{Y}} P\left(y_t' | x_t\right) P\left(y_t^2 = y | y_t'\right)$$

and

$$P\left(y_t' = y | x_t\right) = P\left(y_t^1 = y | x_t\right), \quad \forall y \in \mathbb{Y}$$

that is, the observation with auxiliary variable is equivalent to observing y^1 in distribution for $0 \leq t \leq T-1$.   ◇

The above is related to Blackwell's characterization of more informativeness [4], [9] in the context of comparison of experiments. Even though this is important in centralized stochastic control, it does not have as much value in a decentralized context as the observations at DM^2 are useful to DM^1. In information theory literature, channels satisfying equivalence to a physically degraded channel are termed as stochastically degraded channels [13]. This requirement is weaker than the nestedness condition.

Definition 5.4: An information structure IS2 is information-theoretically degraded with respect to another one, IS1 if $I(x_t; y_t^1) \geq I(x_t; y_t^2)\ 0 \leq t \leq T-1$.   ◇

Mutual information, however, is not an appropriate quantitative measure for the value of information in control systems, as has been observed in [38] and [28]. In particular, different channels with same mutual information characteristics might lead to significantly different performances; as an example consider the estimation problem of a source over two channels, mutual information and estimation error do not have a direct relationship. One can deduce the following:

Theorem 5.1: Deterministic nestedness implies stochastic sequential nestedness, which implies weakly stochastic nestedness, which in turn implies information-theoretic nestedness. ◇

From an optimal control view, however, we have only presented results for the first two structures above. Further work needs to be pursued in this direction in connection with decentralized control systems.

VI. CONCLUSION

This paper attempted to obtain further insight into the problem of *the value of information channels in decentralized stochastic control*. In this direction, we 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 are shown to be strictly less than those for deterministic nestedness, or the deterministic observation sharing patterns. We provided quantitative examples exhibiting the benefit of the new information structure.

At the expense of communications, the controllers benefit from a reduced computational complexity. Further work is needed to obtain a tradeoff between computational complexity and communications, although, we believe we have been able to provide some discussion towards a unifying theory in this direction. It is perhaps counterintuitive first to think that more information means less complexity, because there is more data to work on (or attach a parameter to, if one is to design a suboptimal policy), but stochastic control theory shows that all one needs is to construct a Markov Decision Problem (with an appropriate Markovian state, and a cost function), and work with the chain. The intractability for some decentralized control problems (but not all) stems from not being able to construct such a non-exploding chain in a general setting. Belief sharing is an attempt to provide a systematic, rate-efficient way to construct a tractable setting, via an appropriate Markov chain. This pattern also provides a partial answer on what needs to be exchanged in real-time to obtain optimal performance subject to communication constraints when the communication allows an agreement on beliefs.

One question requiring further study is the following: What is the weakest set of assumptions that leads to the optimality of linear policies for a distributed LQG problem. Are stochastically nested and decoupled information structures the weakest structures one could find? Here, the common property is that the relevant information affecting the cost of the future stages is what need to be exchanged. Stochastic nestedness attempts to express this intuitive property in a formal setting.

One related problem is how to optimally encode the control actions; such an analysis will explicitly depend on the cost functions [35]. The discussion on monotone values of information channels requires further work, in particular with regard to stochastically degraded systems presented in the previous section, in view of Blackwell's notion of more informativeness. Finally, the analysis of belief propagation algorithms for the belief sharing information pattern seems to be an interesting and a practically important problem. Further directions will include the analysis of such problems.

REFERENCES

[1] M. Aicardi, F. Davoli, and R. Minciardi, "Decentralized optimal control of Markov chains with a common past information set," *IEEE Trans. Autom. Control*, vol. AC-32, no. 11, pp. 1028–1031, Nov. 1987.

[2] M. Aoki, "On decentralized linear stochastic control problems with quadratic cost," *IEEE Trans. Autom. Control*, vol. AC-18, no. 3, pp. 243–250, Jun. 1973.

[3] 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 Optim.*, vol. 31, pp. 282–344, 1993.

[4] Y. Azrieli and E. Lehrer, "The value of a stochastic information structure," *Game Econ. Behav.*, pp. 679–693, 2006.

[5] B. Bamieh and P. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Syst. Control Lett.*, vol. 54, pp. 575–583, 2005.

[6] T. Başar, "Decentralized multicriteria optimization of linear stochastic systems," *IEEE Trans. Autom. Control*, vol. AC-23, no. 2, pp. 233–243, Apr. 1978.

[7] 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)*. Amsterdam, The Netherlands: North Holland, 1982, ch. 11, pp. 351–387.

[8] 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, Jun. 1972.

[9] D. Blackwell, "Equivalent comparison of experiments," *Annals Math. Stat.*, vol. 24, pp. 265–272, 1953.

[10] V. Borkar and P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 650–655, Jun. 1982.

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

[12] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*. Budapest, Hungary: Akademiai Kiado, 1981.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[14] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication*. Cambridge, MA: MIT Press, 1998.

[15] O. Hernandez-Lerma, J. Gonzales-Hernandez, and R. R. Lopez-Martinez, "Constrained average cost Markov control processes in Borel spaces," *SIAM J. Control Optim.*, vol. 42, pp. 442–468, 2003.

[16] Y. C. Ho, "Team decision theory and information structures," *Proc. IEEE*, vol. 68, no. 6, pp. 644–654, 1980.

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

[18] B. Kurtaran, "Corrections and extensions to decentralized control with delayed sharing information pattern," *IEEE Trans. Autom. Control*, vol. AC-24, no. 4, pp. 656–657, Aug. 1979.

[19] A. Mahajan, A. Nayyar, and D. Teneketzis, "Identifying tractable decentralized control problems on the basis of information structure," in *Proc. Annual Allerton Conf. Commun., Control Comp.*, Allerton, IL, Sep. 2008, pp. 1440–1449.

[20] M. Mooij and H. J. Kappen, "Sufficient conditions for convergence of the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4422–4437, Dec. 2007.

[21] J. Ooi, S. Verbout, J. Ludwig, and G. Wornell, "A separation theorem for periodic sharing information patterns in decentralized control," *IEEE Trans. Autom. Control*, vol. 42, no. 11, pp. 1546–1550, Nov. 1997.

[22] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[23] C. H. Papadimitriou and J. Tsitsiklis, "On the complexity of designing distributed protocols," *Inform. Control*, vol. 53, pp. 211–218, 1982.

[24] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[25] R. Radner, "Team decision problems," *Annals Math. Stat.*, vol. 33, pp. 857–881, 1962.

[26] A. Rantzer, "Linear quadratic team theory revisited," in *Proc. IEEE Amer. Control Conf.*, Minneapolis, MN, Jun. 2006, pp. 1637–1641.

[27] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized Control," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 274–286, Feb. 2006.

[28] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link part I: Scalar systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[29] N. Sandell, P. Varaiya, M. Athans, and M. Safonov, "Survey of decentralized control methods for large scale systems," *IEEE Trans. Autom. Control*, vol. AC-23, no. 2, pp. 108–128, Apr. 1978.

[30] N. Sandell and M. Athans, "Solution of some nonclassical LQG stochastic decision problems," *IEEE Trans. Autom. Control*, vol. AC-19, no. 1, pp. 108–116, Feb.. 1974.

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

[32] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[33] S. Tatikonda and M. I. Jordan, "Loopy belief propagation and Gibbs measures," in *Proc. Uncertainty Artif. Intell.*, Aug. 2002, vol. 18, pp. 493–500.

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

[35] J. N. Tsitsiklis and M. Athans, "Convergence and asymptotic agreement in distributed decision problems," *IEEE Trans Autom. Control*, vol. AC-29, no. 1, pp. 42–50, Jan. 1984.

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

[37] P. G. Voulgaris, "A convex characterization of classes of problems in control with specific interaction and communication structures," in *Proc. IEEE Amer. Control Conf.*, Arlington, VA, Jun. 2001, pp. 3128–3133.

[38] J. C. Walrand and P. Varaiya, "Optimal causal coding-decoding problems," *IEEE Trans. Inform. Theory*, vol. IT-19, no. 6, pp. 814–820, Nov. 1983.

[39] H. S. Witsenhausen, "Equivalent stochastic control problems," *Math. Control, Signals, Syst.*, vol. 1, pp. 3–11, 1988.

[40] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control Optim.*, vol. 6, pp. 131–147, 1968.

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

[42] S. Yüksel, H. Hindi, and L. S. Crawford, "Optimal tracking with a feedback-feedforward control separation over a network," in *Proc. IEEE Amer. Control Conf.*, Minneapolis, MN, Jun. 2006, pp. 3500–3506.

[43] S. Yüksel and T. Başar, "Optimal signaling policies for decentralized multi-controller stabilizability over communication channels," *IEEE Trans. Autom. Control*, vol. 52, no. 10, pp. 1969–1974, Oct. 2007.

[44] S. Yüksel and T. Başar, "Communication constraints for decentralized stabilizability with time-invariant policies," *IEEE Trans. Autom. Control*, vol. 52, no. 6, pp. 1060–1066, Jun. 2007.

Serdar Yüksel (S'02–A'07) was born in Lice, Turkey in 1979. He received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001; and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana-Champaign, in 2003 and 2006, respectively.

He was a Post-doctoral Researcher at Yale University, New Haven, CT, before joining Queen's University, Kingston, ON, Canada, as an Assistant Professor of mathematics and engineering at the Department of Mathematics and Statistics in 2007. His research interests are on stochastic control, decentralized control, information theory, source coding theory, and multi-terminal control and communication systems.