Forty-Seventh Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 30 - October 2, 2009

Sequential team form and its simplification using graphical models

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

Sekhar Tatikonda
Dept. of Electrical Engineering,
Yale University,
New Haven, CT, USA 06520.
Email: sekhar.tatikonda@yale.edu

***Abstract*—In this paper a sequential team form for dynamic teams is defined. A team form captures the properties of teams that only depend on the independence relation between the system variables but do not depend on the exact functional form of these relationships or on the state space of the variables. A notion of simplifying team forms is presented. We represent team forms using a directed acyclic factor graph and present a graphical model algorithm to simplify the team form.**

I. Introduction

A. Motivation

Qualitative properties of optimal control laws play an important role in stochastic control. Such properties allow us to restrict attention to only a subclass of control laws, thereby reducing the solution complexity. Such properties are also called "structural results". They have two distinct flavors, which we will call structural results of the first and second kind.

The first kind of structural results do not depend on the specifics of the state spaces, the underlying probability measure, and the cost function. For example, in MDPs (Markov decision processes) without loss of optimality we can choose control actions based on the current state of the process; in POMDPs (partially observable Markov decision processes) without loss of optimality we can choose control actions based on the information state (also called the belief state) of the process. These structural results hold under mild assumptions on the underlying model. The nature of the process (discrete or continuous valued), the specific form of the probability measure (uniform, binomial, Gaussian, etc.), and the specific properties of the cost function (concave, convex, monotone, etc.) are immaterial. These results only depend on the process being control Markov and the total cost being additive.

The second kind of structural results depend on the specifics of the state space, the underlying probability measure, and the cost function. For example, in LQG (linear quadratic Gaussian) systems (i.e., systems with linear dynamics, quadratic cost, Gaussian random variables and perfect recall at the controller) without loss of optimality we can choose the control action to be an affine function of the estimate of the state. If one of the assumptions of the model is relaxed, the result ceases to be true. Similar results exist in queuing theory which are only true when the holding cost is linear in the queue size.

Both kinds of structural results are well studied for centralized stochastic control systems. However, structural results of centralized systems break down in decentralized systems; For decentralized systems, structural results are known for only a few specific models [1]–[11]. These results are proved on a case-by-case basis. No general framework to study structural results of decentralized control systems exist. In this paper, we present one such framework to derive structural results of the first kind.

B. Literature Overview

In this paper, we restrict attention to *sequential teams*. *Teams* are decentralized control systems in which all agents have the same objective (in contrast to games and multi-objective systems). In *sequential* teams, the order in which the agents act does not depend on the actions of nature or other agents. We briefly overview the different models for sequential teams that have appeared in the literature.

Perhaps the most general model for decentralized control systems is Witsenhausen's intrinsic model [12] (for general non-sequential multi-objective systems) and [13] (when restricted to sequential teams). This model is extremely useful in resolving the conceptual issues related with decentralized control systems. However, this model does not include explicit observations, so it is of limited value in establishing structural results.

A second class of models are the sequential team model and the standard form considered in [14]. These models are equivalent to the intrinsic model [12] and were used to show that any sequential team problem can be sequentially decomposed into nested subproblems. However, like the intrinsic model, these models do not include explicit observations, so they are of limited value in establishing structural results.

A third class of models is the model considered in [15] which is also equivalent to the intrinsic model. This model was used to show that any dynamic team (where the actions of one agent affects the observations of others) can be reduced to a static team (where the actions of any agent do not affect the observations of others). This model includes explicit observations, and as such can be useful in establishing structural results. We will follow a slight variation of this model. Our variation allows for a partial order on the agents (the model in [15] assumed a total order), and expresses the

978-1-4244-5871-4/09/$26.00

joint probability in terms of factorization based on conditional independences (the model in [15] expressed the joint probability in terms of a product measure on primitive random variables).

We use graphical models to represent sequential teams. Various graphical models for sequential teams already exist in the literature. Witsenhausen [1] used a graphical model where each node corresponds to a system dynamics function or a control law. A directed edge from a node f to a node g indicates that the output of f is an input of g. Ho and Chu [2] used a directed tree to model partially nested teams. Each node of the tree corresponds to an agent. A directed edge from node i to node j indicates that the action of agent i affects the observation of agent j. Yoshikawa [4] used a directed multi-graph to represent a sequential team. As in [2], each node corresponds to an agent. A solid directed edge from node i to node j indicates that the action of agent i affects the observation of agent j; a dashed directed edge from node i to node j indicates that agent j knows all the data observed at agent i.

None of the above models are appropriate for out purpose. So, we use directed factor graphs to model sequential teams. Factor nodes correspond to system dynamics functions and control functions; variable nodes correspond to system variables. A directed edge from variable x to factor f indicates that x is an input to f; a directed edge from factor f to variable x indicates that x is an output of f. Such a directed factor graph allows us to easily reason about the conditional independence relations betweens the system variables.

C. Contribution

The contributions of this paper are twofold. Firstly, we present a graphical representation of sequential teams. This representation is easy to understand, and at the same time, is general enough to model any finite horizon sequential team (with unconstrained control laws). Secondly, we present a methodology to simplify the team using its graphical representation. This methodology can be automated using standard algorithms from graphical models.

D. Notation

Given two sets A and B, $A \times B$ denotes their Cartesian product. Given two measurable spaces $(\mathcal{X}, \mathfrak{X})$ and $(\mathcal{Y}, \mathfrak{Y})$, $\mathfrak{X} \otimes \mathfrak{Y}$ denotes the product sigma algebra on $\mathcal{X} \times \mathcal{Y}$. Given two measures μ_X on $(\mathcal{X}, \mathfrak{X})$ and μ_Y on $(\mathcal{Y}, \mathfrak{Y})$, $\mu_X \otimes \mu_Y$ denotes the product measure on $(\mathcal{X} \times \mathcal{Y}, \mathfrak{X} \otimes \mathfrak{Y})$.

Given any set M, X_M denotes the vector $\{X_m : m \in M\}$ while $\mathcal{X}_M$ denotes the product space $\prod_{m \in M} \mathcal{X}_m$ and $\mathfrak{X}_M$ denotes the product σ-algebra $\bigotimes_{m \in M} \mathfrak{X}_m$.

E. Organization

The rest of this paper is organized as follows. In Section! II We present a model for sequential teams, define a team form, and explain what we mean by simplification of a team form. In Section III, we present some preliminaries on partial orders and graphical models. In Section IV we show how to represent a sequential team form as a graphical model. We discuss completing a team form in Section V and simplifying a team form in Section VI. We conclude in Section VII.

II. SEQUENTIAL TEAMS AND TEAM FORMS

A. Sequential teams

A sequential team consists of the following.

1) A collection of n system variables, X_k, $k \in N$ where $N = \{1, \dots, n\}$.
2) A collection $\{(\mathcal{X}_k, \mathfrak{F}_k)\}_{k \in N}$ of measurable spaces. The system variable X_k takes values in $(\mathcal{X}_k, \mathfrak{F}_k)$, $k \in N$. The σ-algebras $\mathfrak{F}_k$, $k \in N$ contain all singletons and are either Borelian or countably generated.
3) A collection $\{I_k\}_{k \in N}$ of sets such that $I_k \subseteq \{1, \dots, k-1\}$. I_k is the *information set* for variable X_k.
4) A set $A \subset N$ of control agents. For $k \in A$, agent k chooses the system variable X_k. X_k is the control action of agent k.
5) The variables $X_{N \setminus A}$ are chosen by nature according to a collection $\{p_k\}_{k \in N \setminus A}$ of stochastic kernels where p_k is a stochastic kernel from $(\mathcal{X}_{I_k}, \mathfrak{F}_{I_k})$ to $(\mathcal{X}_k, \mathfrak{F}_k)$.
6) A set $R \subset N$. The variables X_R are the *reward variables*.

A sequential team is denoted by the tuple $(N, A, R, \{I_k\}_{k \in N}, \{(\mathcal{X}_k, \mathfrak{F}_k)\}_{k \in N}, \{p_k\}_{k \in N \setminus A})$. Given a sequential team, the system designer has to choose *control strategy* $\{g_k\}_{k \in A}$ such that g_k is a measurable function from $(\mathcal{X}_{I_k}, \mathfrak{F}_{I_k})$ to $(\mathcal{X}_k, \mathfrak{F}_k)$. The choice of a control strategy induces a probability measure on X_N which is given by

$$P(dX_N) = \bigotimes_{k \in N \setminus A} p_k(dX_k | X_{I_k}) \bigotimes_{k \in A} \delta_{g_k(X_{I_k})}(dX_k) \quad (1)$$

The design objective is to choose a control strategy $\{g_k\}_{k \in A}$ to maximize the expectation of $\sum_{k \in R} X_k$ where the expectation is with respect to the induced probability measure given by (1). The corresponding maximum value of this expectation is the *value of the team*.

Remark: The optimal design problem described above in an unconstrained optimization problem because the control laws g_k can be any measurable function from $(\mathcal{X}_{I_k}, \mathfrak{F}_{I_k})$ to $(\mathcal{X}_k, \mathfrak{F}_k)$. Consequently, using randomized control laws does not improve performance [13]. Hence, there is no loss of generality in the above assumption that all control laws g_k are non-randomized.

B. Sequential team forms

A *sequential team form* is a tuple $(N, A, R, \{I_k\}_{k \in N})$, that satisfies the first and third condition of the definition of a team. A sequential team form along with measurable spaces $\{(\mathcal{X}_k, \mathfrak{F}_k)\}_{k \in N}$ and stochastic kernels $\{p_k\}_{k \in N}$ specifies a sequential team. In the rest of this paper, we will abbreviate "sequential team form" to "team form". In general, we can also have nonsequential team forms.

Two team forms $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ and $\mathcal{T}' = (N', A', R', \{I'_k\}_{k \in N'})$ are *equivalent* if the following conditions hold:

1) $N = N'$, $A = A'$, and $R = R'$;

2) for all $k \in N \setminus A$, we have $I_k = I'_k$;
3) for any choice of measurable spaces $\{(\mathcal{X}_k, \mathfrak{F}_k)\}_{k \in N}$ and stochastic kernels $\{p_k\}_{k \in N \setminus A}$, the values of the teams corresponding to $\mathcal{T}$ and $\mathcal{T}'$ are the same.

The first two conditions can be verified trivially. There is no easy way to check the last condition.

A team form $\mathcal{T}' = (N', A', R', \{I'_k\}_{k \in N'})$ is a *simplification* of a team form $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ if $\mathcal{T}'$ is equivalent to $\mathcal{T}$ and

$$\sum_{k \in A} |I'_k| < \sum_{k \in A} |I_k| .$$

$\mathcal{T}'$ is a *strict simplification* of $\mathcal{T}$ if $\mathcal{T}'$ is equivalent to $\mathcal{T}$, $|I'_k| \leq |I_k|$ for $k \in N$, and at least one of these inequalities is strict.

We are interested in the following question. ***Given a sequential team form, can we simplify it?*** In this paper, we present a methodology to simplify sequential team forms This methodology is based on ideas from partial orders and graphical models, so we briefly overview these areas in the next section.

III. Preliminaries

A. Partial orders

A *strict partial order* $\prec$ on a set S is a binary relation that is transitive, irreflexive, and asymmetric. A set with a partial order on it is called a *partially ordered set* or a *poset*.

The *reflexive closure* $\preceq$ of a partial order $\prec$ is given by $a \preceq b$ if and only if $a \prec b$ or $a = b$. The relation $\preceq$ is transitive, reflexive, and antisymmetric. It is also called a *non-strict partial order*. A non-strict partial order is *total* if for all a and b in S, $a \preceq b$ or $b \preceq a$.

Let A be a subset of a partially ordered set $(S, \prec)$. Then, the *lower set* of A, denoted by $\overleftarrow{A}$ is defined as $\overleftarrow{A} := \{b \in S : b \preceq a \text{ for some } a \in A\}$. By duality, the *upper set* of A, denoted by $\overrightarrow{A}$ is defined as $\overrightarrow{A} := \{b \in S : a \preceq b \text{ for some } a \in A\}$.

For singleton sets $\{a\}$, we will denote $\overleftarrow{\{a\}}$ and $\overrightarrow{\{a\}}$ by $\overleftarrow{a}$ and $\overrightarrow{a}$.

B. Directed graphs

A *directed graph* $\mathcal{G}$ is a tuple (N, E) where N is the set of nodes and $E \subseteq N \times N$ is the set of edges. An edge (u, v) in E is considered directed from u to v; u is *in-neighbor* or *parent* of v, v a *out-neighbor* or *child* of u, and u and v are *neighbors*. The set of in-neighbors of v, also called the *in-neighborhood* of v, is denoted by $N^-_{\mathcal{G}}(v)$; the set of out-neighbors of u, also called the *out-neighborhood* of u, is denoted by $N^+_{\mathcal{G}}(v)$; the set of neighbors of u, also called the *neighborhood* of u, is denoted by $N_{\mathcal{G}}(u)$.

A *path* is a sequence of nodes such that each node has a directed edge to the next node in the sequence. The first node of a path is its *start node*, the last node is its *end node*. A *cycle* is a path with the same start and end node. A *trail* is a sequence of nodes such that each node is a neighbor of the next node in the sequence.

A *directed acyclic graph* (DAG) is a directed graph with no cycles. A DAG $\mathcal{G}(N, E)$ gives rise to a partial order $\prec_{\mathcal{G}}$ on its its nodes: for u, v in N, $u \prec_{\mathcal{G}} v$ when there exists a path from u to v. Thus, the induced partial order $\prec_{\mathcal{G}}$ is the transitive closure of the edge set E. The lower set $\overleftarrow{v}$ of a node v in N is equivalent to the set of all nodes u in N such that there is a path from u to v. This set is also called the *ancestors* of v. In addition, the lower set $\overleftarrow{A}$ of a subset A of N is called the *ancestral set* of A. Similarly, the upper set $\overrightarrow{u}$ of a node u in N is equivalent to the set of all nodes v in N such that there is a path from u to v. This set is also called the *descendants* of u. In addition, the upper set $\overrightarrow{A}$ of a subset A of N is called the *descendant set* of A.

A *bipartite directed graph* is a directed graph whose nodes can be divided into two disjoint sets N_1 and N_2 such that $E \subseteq (N_1 \times N_2) \cup (N_2 \times N_1)$, that is, every edge has a node in N_1 and another in N_2.

C. Graphical model

A *graphical model* is a graph that captures conditional independence relation between random variables. In this paper, we will use a *directed acyclic factor graph* (DAFG) as a graphical model. A DAFG $\mathcal{G}(V, F, E)$ is an acyclic bipartite directed graph with two kinds of nodes, variable nodes V and factor nodes F. Each variable node v in V corresponds to a random variable X_v. The random variable X_v takes values in some measurable space $(\mathcal{X}_v, \mathfrak{F}_k)$. Each factor node f in F corresponds to a stochastic kernel p_f from $(\mathcal{X}_{N^-_{\mathcal{G}}(f)}, \mathfrak{F}_{N^-_{\mathcal{G}}(f)})$ to $(\mathcal{X}_{N^+_{\mathcal{G}}(f)}, \mathfrak{F}_{N^+_{\mathcal{G}}(f)})$. The joint probability measure on X_V is given by

$$P\left(dX_V\right) = \bigotimes_{f \in F} p_f(dX_{N^+_{\mathcal{G}}(f)} | X_{N^-_{\mathcal{G}}(f)}).$$

Such a joint measure is said to have a *recursive factorization* according to the DAFG $\mathcal{G}$. All joint measures that have a recursive factorization according to a DAFG are denoted by $\mathcal{P}_{\mathcal{G}(V,F,E)}$.

Given a joint probability measure $P\left(dX_V\right)$ on X_V, and sets A, B, $C \subseteq V$, X_A is *independent* of X_B given X_C if

$$P\left(X_A \,|\, X_B, X_C\right) = P\left(X_A \,|\, X_C\right)$$

D. Graphical model with deterministic nodes

Traditionally, graphical models capture conditional independence relation between random variables that hold for all probability measures that have recursive factorization according to a graph. In stochastic control problems, the control actions are generated using non-randomized functions. As such, in stochastic control problems, one is interested in conditional independence relation between random variables that hold for all probability measures where certain variables are deterministic functions of their parent variables, while others are stochastic functions of their parent variables. Such conditional independence relations can be captured using DAFG with deterministic nodes.

A DAFG $\mathcal{G}$ with deterministic nodes is a tuple (V, F, E, D) where (V, F, E) is a DAFG and $D \subseteq F$ is the set of deterministic factors. When we draw a DAFG with deterministic nodes, we use circles to indicate variable nodes and squares to indicate factor nodes. Deterministic factors are indicated by solid squares whilst stochastic factors are indicated by hollow squares. In the rest of this paper, we will call a DAFG with deterministic nodes simply as a DAFG.

In a DAFG $\mathcal{G}(V, F, E, D)$, each variable v in V corresponds to a random variable X_v, each stochastic factor node f in $F \setminus D$ corresponds to a stochastic kernel p_f from $(\mathcal{X}_{N^-_{\mathcal{G}}(f)}, \mathfrak{F}_{N^-_{\mathcal{G}}(f)})$ to $(\mathcal{X}_{N^+_{\mathcal{G}}(f)}, \mathfrak{F}_{N^+_{\mathcal{G}}(f)})$, and each deterministic node d in D corresponds to a measurable function g_d from $(\mathcal{X}_{N^-_{\mathcal{G}}(f)}, \mathfrak{F}_{N^-_{\mathcal{G}}(f)})$ to $(\mathcal{X}_{N^+_{\mathcal{G}}(f)}, \mathfrak{F}_{N^+_{\mathcal{G}}(f)})$. The joint probability measure on X_V is given by

$$P(dX_V) = \bigotimes_{f \in F \setminus D} p_f(X_{N^+_{\mathcal{G}}(f)} | X_{N^-_{\mathcal{G}}(f)}) \times \bigotimes_{f \in D} \delta_{g_f(X_{N^-_{\mathcal{G}}(f)})}(dX_{N^+_{\mathcal{G}}(f)}).$$

Such a joint measure is said to have a *deterministic recursive factorization* according to the DAFG $\mathcal{G}$ (with deterministic nodes). All joint measures that have deterministic recursive factorization according to a DAFG are denoted by $\mathcal{P}_{\mathcal{G}(V,F,E,D)}$, or, when there is no ambiguity, simply by $\mathcal{P}_{\mathcal{G}}$.

Given a DAFG $\mathcal{G}(V, F, E, D)$ and sets A, B, $C \subseteq V$, X_A is *deterministically irrelevant* to X_B given X_C if X_A is independent to X_B given X_C for *all* joint measures $P(X_V)$ in $\mathcal{P}_{\mathcal{G}}$. All variable nodes that are deterministically irrelevant to X_A given X_C, denoted by $R^-_{\mathcal{G}}(X_A | X_C)$, can be determined in linear time in the size of the graph using a deterministic generalization of d-separation called D-separation [17], [18].

E. Observations and functional observations

Given a DAFG $\mathcal{G}(V, F, E, D)$, a variable X_v is *observed* at a factor node f if $v \in N^-_{\mathcal{G}}(f)$. A variable X_v is *functionally observed* at a factor node f if every variable node in $N^-_{\mathcal{G}}(N^-_{\mathcal{G}}(v))$ is either observed or functionally observed at f. The set of functional observations at a factor node is denoted by $O^+_{\mathcal{G}}(f)$ and equals to the set of variables that are deterministically irrelevant to X_V given $X_{N^-_{\mathcal{G}}(f)}$, i.e., $O^+_{\mathcal{G}}(f) = R^-_{\mathcal{G}}(X_V | X_{N^-_{\mathcal{G}}(f)})$. A variable v is *strict functional observation* at a factor node f, if v is not observed at f but is functionally observed at f. The set of strict functional observations at a factor node is denoted by $O^-_{\mathcal{G}}(f)$ and equals to the set of functional observations minus the in-neighborhood, i.e., $O^-_{\mathcal{G}}(f) = O^+_{\mathcal{G}}(f) \setminus N^-_{\mathcal{G}}(f)$.

IV. Graphical model for a team form

A team form $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ can be represented as a DAFG as follows. Let k^0 denote $(k, 0)$ and k^1 denote $(k, 1)$. Now, consider a DAFG with $V = N \times \{0\}$, $F = N \times \{1\}$, $E = \{(k^1, k^0) : k \in N\} \cup \{(i^0, k^1) : k \in N, i \in I_k\}$, and $D = A \times \{1\}$. Thus, for each system variable X_k, we have a variable node k^0 and a factor node k^1 in the graph. The variable node represents X_k and the factor node represents the "function", g_k for control actions and p_k for others, that generates the variable node. For each $k \in N$, we add an edge from the factor node k^1 to the variable node k^0. For $i, k \in N$ such that $i \in I_k$, we add an edge from the variable node i^0 to the factor node k^1. The factor nodes corresponding to the control actions are deterministic. This DAFG $\mathcal{G} = (V, F, E, D)$ along with the reward nodes $R^0 = R \times \{0\}$ is a graphical model for the team form $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$.

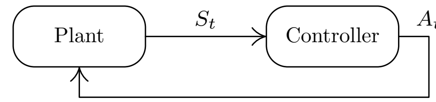

Fig. 1. A Markov decision process

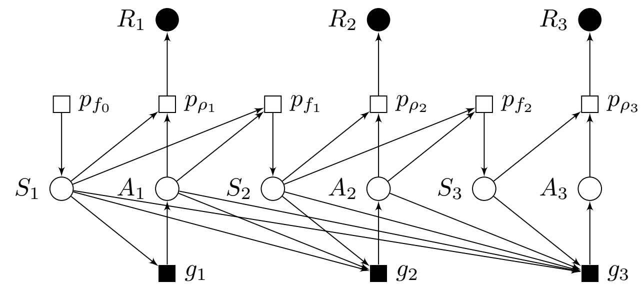

Fig. 2. The team form of a Markov decision process.

We assign a label to each node in the DAFG $\mathcal{G}$. For an abstract sequential team form, X_k is the label of k^0, p_k is the label of k^1, $k \in N \setminus A$, and g_k is the label of k^1, $k \in A$. For a specific sequential team form, we will use labels that are natural to the setup. In the remainder of this paper, we will the index and the label of nodes interchangeably.

Since $\mathcal{G}$ is a DAFG, it induces a partial order $\prec_{\mathcal{G}}$ on its nodes. As explained earlier, this partial order $\prec_{\mathcal{G}}$ is the transitive closure on the edge set E. When restricted to $V \times V$, this partial order $\prec_{\mathcal{G}}$ induces a partial order $\prec_N$ on X_N. Likewise, when restricted to $D \times D$, the partial order $\prec_{\mathcal{G}}$ induces a partial order $\prec_A$ on $\{g_k\}_{k \in A}$. Thus, for any sequential team form, there is a partial order on the decision makers. This is consistent with the result in [13] that a team problem is sequential if and only if there is a partial order on the decision makers.

To fix ideas, we consider some examples of sequential teams and their corresponding graphical model.

Example 1 (Markov decision process): A Markov decision process, shown in FigurE 1, consists of a plant and a controller. The state S_t of the plant evolves according to

$$S_{t+1} = f_t(S_t, A_t, W_t), \quad t = 1, \dots, T-1 \tag{2}$$

where A_t is the control action and W_t is plant disturbance. The process $\{W_t,\ t = 1, \dots, T\}$ is an independent process and is also independent of S_1. The controller observes the state of the plant. It has perfect memory, so it remembers everything that it has seens and done in the past, and choose a control action A_t according to

$$A_t = g_t(S_1, \dots, S_t, A_1, \dots, A_{t-1})$$

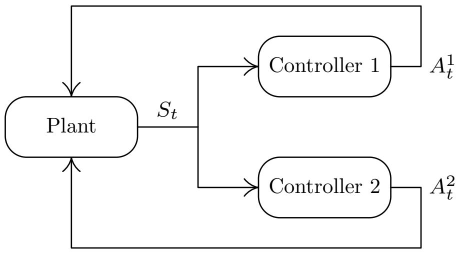

Fig. 3. A simple two agent team

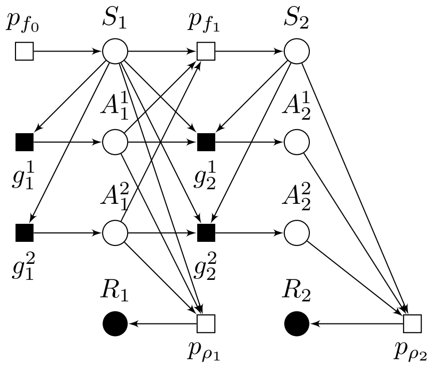

Fig. 4. The team form of a simple two agent team.

At each time a reward $R_t = \rho_t(S_t, A_t)$ is obtained. The objective is to choose a control strategy $(g_1, \dots, g_T)$ so as to maximize the expected value of $\sum_{t=1}^{T} R_t$.

The above model is a sequential team with $n = 3T$ variables—T state variables, T action variables, and T reward varaibles. Instead of denoting these by an index $N = \{1, \dots, n\}$, for ease of understanding we will denote them by their labels. Thus, the team form corresponding to a Markov decision process is given as follows.

1) $\mathcal{X}_N = \{S_1, \dots, S_T, A_1, \dots, A_T, , R_1, \dots, R_T\}$. The control variables are $\mathcal{X}_A = \{A_1, \dots, A_T\}$ and the reward variables are $\mathcal{X}_R = \{R_1, \dots, R_T\}$.
2) For $X_k = A_t \in \mathcal{X}_A$, $X_{I_k} = \{\mathcal{S}_1, \dots, \mathcal{S}_t, A_1, \dots, A_{t-1}\}$; for $X_k = R_t \in \mathcal{X}_R$, $X_{I_k} = \{X_t, A_t\}$; and for $X_k \in S_t \in \mathcal{X}_{N\setminus(A\cup R)}$, $X_{I_k} = \emptyset$ when $t = 1$, and (X_{t-1}, A_{t-1}) otherwise.

This team form can be represented as a factor graph. This factor graph is shown in Figure 2 for $T = 3$. p_{f_t} and p_{ρ_t} are the stochastic kernels corresponding to the functions f_t and ρ_t.

Example 2 (A simple two agent team): Consider a simple two agent team, shown in Figure 3, that consists of a plant and two controller. $\{S_t,\ t = 1, \dots, T\}$ denotes the state of the plant and $\{A_t^i,\ t = 1, \dots, T\}$ denotes the control actions of the controller i, $i = 1, 2$. The plant is updated according to a plant function f_t,

$$S_{t+1} = f_t(S_t, A_t^1, A_t^2, W_t), \quad t = 1, \dots, T-1 \tag{3}$$

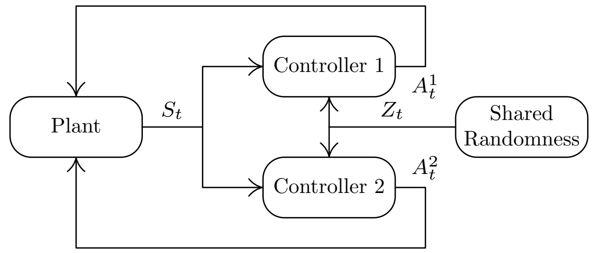

Fig. 5. A simple two agent team with shared randomness

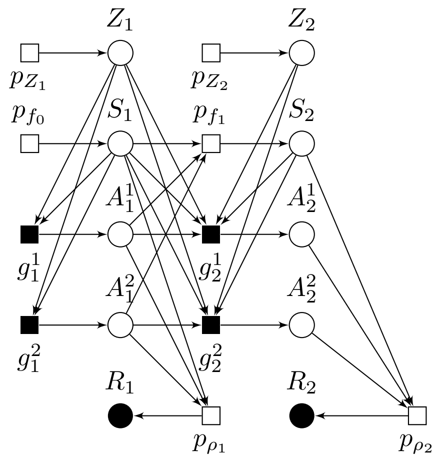

Fig. 6. The team form of a simple two agent team with common randomness.

where $\{W_t,\ t = 1, \dots, T\}$ is a sequence of independent random variables that are also independent of S_1. Each controller observes the state of the plant and has perfect memory, so it remembers everything that it has seens and done in the past. Controller i choose a control action A_t^i according to

$$A_t^i = g_t^i(S_1, \dots, S_t, A_1^i, \dots, A_{t-1}^i)$$

At each time a reward $R_t = \rho_t(S_t, A_t^1, A_t^2)$ is obtained. The objective is to choose a control strategy $(g_1^1, \dots, g_T^1;\ g_1^2, \dots, g_T^2)$ so as to maximize the expected value of $\sum_{t=1}^{T} R_t$.

The above model is a sequential team with $n = 4T$ variables—T state variables, $2T$ action variables, and T reward varaibles. As in the expample on Markov decision process, instead of denoting the variables by an index, we denote them by their labels. Thus, the team form corresponding to this model is given as follows.

1) $X_N = \{S_1, \dots, S_T, A_1^1, \dots, A_T^1, A_1^2, \dots, A_T^2, R_1, \dots, R_T\}$. The control variables are $X_A = \{A_1^1, \dots, A_T^1, A_1^2, \dots, A_T^2\}$ and the reward variables are $X_R = \{R_1, \dots, R_T\}$.
2) For $X_k = A_t^i \in \mathcal{X}_A$, $X_{I_k} = \{\mathcal{S}_1, \dots, \mathcal{S}_t, A_1^i, \dots, A_{t-1}^i\}$; for $X_k = R_t \in \mathcal{X}_R$, $X_{I_k} = \{X_t, A_t^1, A_t^2\}$; and for $X_k \in S_t \in \mathcal{X}_{N\setminus(A\cup R)}$, $X_{I_k} = \emptyset$ when $t = 1$, and $(X_{t-1}, A_{t-1}^1, A_{t-1}^2)$ otherwise.

This team form can be represented as a factor graph. This

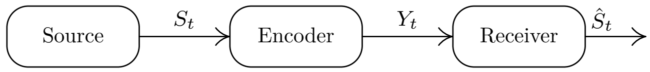

Fig. 7. A real time communication system

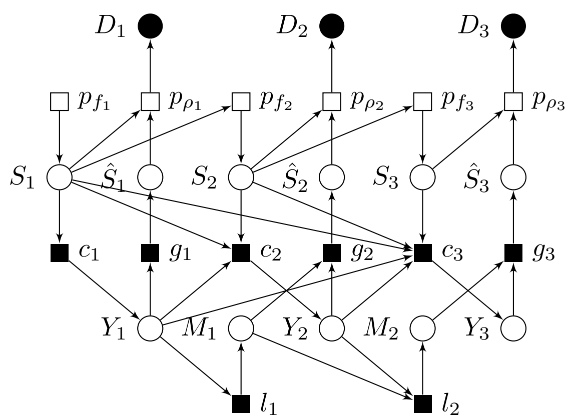

Fig. 8. The team form of a real time communication system.

factor graph is shown in Figure 4 for $T = 2$. p_{f_t} and p_{ρ_t} are the stochastic kernels corresponding to the functions f_t and ρ_t.

Example 3 (Team with shared randomness): Consider a two agent team, shown in Figure 5, which is similar to Example 2 except that both agents observe a i.i.d. process $\{Z_t,\ t = 1, \dots, T\}$ that is independent of $\{W_t,\ t = 1, \dots, T\}$ and S_1. We call this process *shared randomness*. Thus, controller i chooses a control action A^i_t according to

$$A^i_t = g^i_t(S_1, \dots, S_t, A^i_1, \dots, A^i_{t-1}, Z_1, \dots, Z_t) \qquad (4)$$

The plant dynamics, the reward function, and the objective are the same as in Example 2.

The above model is a sequential team with $n = 5T$ variables—T state variables, $2T$ action variables, T reward variabls, and T shared randomness variables. As before, we denote the variables by their label. The team form corresponding to this model is given as follows.

1) $X_N = \{S_1, \dots, S_T, A^1_1, \dots, A^1_T, A^2_1, \dots, A^2_T, R_1, \dots, R_T, Z_1, \dots, Z_T\}$. The control and the reward variables are the same as in Example 2.
2) For $X_k = A^i_t \in \mathcal{X}_A$, $X_{I_k} = \{\mathcal{S}_1, \dots, \mathcal{S}_t, A^i_1, \dots, A^i_{t-1}, Z_1, \dots, Z_t\}$; for $X_k = Z_t$, $X_{I_k} = \emptyset$; for $X_k = R_t$ and $X_k = S_t$, X_{I_k} is the same as in Example 2.

This team form can be represented as a factor graph. This factor graph is shown in Figure 6 for $T = 2$. p_{f_t} and p_{ρ_t} are the stochastic kernels corresponding to the functions f_t and ρ_t; p_{Z_t} the probability density function for Z_t.

Example 4 (Real-time source coding): Consider a simple real-time communication system, shown in Figure 7, that consists of a source, an encoder and a decoder. This system was originally studied in [19]. $\{S_t,\ t = 1, \dots, T\}$ denotes the source output; $\{Y_t,\ t = 1, \dots, T\}$ denotes the encoded symbols; and $\{\hat{S}_t,\ t = 1, \dots, T\}$ denotes the decoder outputs. The source is a first-order Markov source, so

$$P(S_{t+1} \mid S_1, \dots, S_t) = P(S_{t+1} \mid S_t) = Q_t(S_{t+1} | S_t).$$

The encoded symbols are generated according to

$$Y_t = c_t(S_1, \dots, S_t, Y_1, \dots, Y_{t-1}).$$

The decoder is a finite memory decoder. $\{M_t,\ t = 1, \dots, T\}$ denotes the memory of the decoder. The decoder generates a decoded symbol according to

$$\hat{S}_t = g_t(Y_t, M_{t-1}),$$

and then updates its memory according to

$$M_t = l_t(Y_t, M_{t-1}).$$

At each time a distortion $D_t = \rho_t(S_t, \hat{S}_t)$. The objective is to choose a communication strategy $(c_1, \dots, c_T; g_1, \dots, g_T; l_1, \dots, l_T)$ so as to minimize the expected value of $\sum_{t=1}^{T} D_t$.

The above model is a sequential team with $n = 5T$ variables—T source outputs, T encoded symbols, T memory contents, T decoded symbols, T distortion variables. As in previous examples, we denote the variables by their index. The team form corresponding to this model is given as follows.

1) $X_N = \{S_1, \dots, S_T, Y_1, \dots, Y_T, M_1, \dots, M_T, \hat{S}_1, \dots, \hat{S}_T, D_1, \dots, D_T\}$. The control variables are $X_A = \{Y_1, \dots, Y_T, M_1, \dots, M_T, \hat{S}_1, \dots, \hat{S}_T\}$ and the reward variables are $X_R = \{D_1, \dots, D_T\}$.
2) For $X_k = Y_t$, $X_{I_k} = \{S_1, \dots, S_t, Y_1, \dots, Y_{t-1}\}$; for $X_k = M_t$, $X_{I_k} = \{Y_1\}$ when $t = 1$, and $\{Y_t, M_{t-1}\}$ otherwise; for $X_k = \hat{S}_t$, $X_{I_k} = \{Y_1\}$ when $t = 1$, and $\{Y_t, M_{t-1}\}$ otherwise; for $X_k = D_t$, $X_{I_k} = \{S_t, \hat{S}_t\}$; and for $X_k = S_t$, $X_{I_k} = \emptyset$ when $t = 1$, and $\{S_{t-1}\}$ otherwise.

This team form can be represented as a factor graph. This factor graph is shown in Figure 8 for $T = 3$. p_{ρ_t} is the stochastic kernels corresponding to the function ρ_t.

V. Completion of a team form

A team form $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ is *complete* if for $k, l \in \mathcal{A}$, $k \neq l$, such that $I_k \subset I_l$ we have $X_k \in I_l$. Equivalently, a team form is complete if its graphical model has no deterministic factor node with strict functional observations.

Any team form $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ can be completed as follows. Let $\mathcal{G} = (V, F, E, D)$ be its corresponding graphical model. If for all f in $\mathcal{D}$, $O^-_{\mathcal{G}}(f) = \emptyset$ then $\mathcal{T}$ is complete. Otherwise, pick an $f \in \mathcal{D}$ such that $O^-_{\mathcal{G}}(f) \neq \emptyset$. Create a new graphical model $\mathcal{G}' = (V', F', E', D')$ with $\mathcal{V}' = \mathcal{V}$, $\mathcal{F}' = \mathcal{F}$, $\mathcal{E}' = \mathcal{E} \cup \{(v, f) : v \in O^-_{\mathcal{G}}(f)\}$, and $\mathcal{D}' = \mathcal{D}$. If $\mathcal{G}'$ is complete then stop, otherwise repeat the above procedure with $\mathcal{G}'$. Since the number of vertices (and hence the maximum number of possible edges) in the graphical model are finite, the above process, which adds edges to the graph at each step, will always terminate. The team form $\mathcal{T}'$ corresponding to the terminal graphical model is called the *completion*[1] of $\mathcal{T}$.

[1] This notion of completion of a team form is similar to the notion of strict expansion of an information structure used in [20].

The competition of a team is not unique. Depending on the order in which we complete the graphical model, we may end up with different completions. Nevertheless, all of them are equivalent to the original team form.

Proposition 1: All the completions of a team form are equivalent to it.

Proof: Let $\mathcal{T} = (N, A, R, \{I_k\}_{k\in N})$ be a team form and $\mathcal{T}' = (N', A', R', \{I'_k\}_{k\in N'})$ be (one of) its completion. The procedure for completing a team form adds edges into the deterministic factors of graphical model corresponding to $\mathcal{T}$. So, the first two properties of equivalence—(1) $\mathcal{N}' = \mathcal{N}$, $\mathcal{A}' = \mathcal{A}$, $\mathcal{R}' = \mathcal{R}$; and (2) for all $k \in \mathcal{N} \setminus A$, $I'_k = I_k$—are satisfied by construction. Thus, we only need to check the third property. This means that we need to show that for any choice of measurable spaces $\{(\mathcal{X}_k, \mathfrak{F}_k)\}_{k\in\mathcal{N}}$ and stochastic kernels $\{p_k\}_{k\in\mathcal{N}\setminus\mathcal{A}}$, the values of the teams corresponding to $\mathcal{T}$ and $\mathcal{T}'$ are the same. To prove this, we show that each step of the completion process does not change the value of the team. Such a result will imply that the teams corresponding to the starting and the terminal teams forms have the same value.

Let $\mathcal{T}_1 = (N, A, R, \{I^1_k\}_{k\in N})$ and $\mathcal{T}_2 = (N, A, R, \{I^2_k\}_{k\in N})$ be the team forms corresponding to two consecutive steps of the completion of $\mathcal{T}$ such that by adding incoming edges to a node $(k^*, 1)$ in the graphical model of $\mathcal{T}_1$ we get the graphical model of $\mathcal{T}_2$. By construction, $I^1_{k^*} \subset I^2_{k^*}$ while for all other $k \in \mathcal{A}$, $k \neq k^*$, $I^1_k = I^2_k$. Consider any strategy $\{g^1_k\}_{k\in\mathcal{A}}$ for $\mathcal{T}_1$. The control law g_k is a measurable function from $(\mathcal{X}_{I^1_k}, \mathfrak{F}_{I^1_k})$ to $(\mathcal{X}_k, \mathfrak{F}_k)$. Since $I^1_{k^*} \subset I^2_{k^*}$, the measurable space $(\mathcal{X}_{I^1_{k^*}}, \mathfrak{F}_{I^1_{k^*}})$ is a natural projection of $(\mathcal{X}_{I^2_{k^*}}, \mathfrak{F}_{I^2_{k^*}})$; for all other $k \in \mathcal{A}$, $k \neq k^*$, the measurable spaces $(\mathcal{X}_{I^1_k}, \mathfrak{F}_{I^1_k})$ and $(\mathcal{X}_{I^2_k}, \mathfrak{F}_{I^2_k})$ are the same. Therefore, for all $k \in \mathcal{A}$, g^1_k is also a measurable function from $(\mathcal{X}_{I^2_k}, \mathfrak{F}_{I^2_k})$ to $(\mathcal{X}_k, \mathfrak{F}_k)$. Hence, any strategy $\{g^1_k\}_{k\in\mathcal{A}}$ for $\mathcal{T}_1$ is also a valid strategy for $\mathcal{T}_2$ and consequently, the value of $\mathcal{T}_1$ is less than or equal to the value of $\mathcal{T}_2$.

Next consider a strategy $\{g^2_k\}_{k\in\mathcal{A}}$ for $\mathcal{T}_2$. By construction, there exists a function F such that $X_{I^2_{k^*}} = F(X_{I^1_{k^*}}, \{g^2_k\}_{k\in\mathcal{A}})$; for all $k \in \mathcal{A}$, $k \neq k^*$, $I^2_k = I^1_k$. Now consider a strategy $\{g^1_k\}_{k\in\mathcal{A}}$ such that $g^1_{k^*}(X_{I^1_{k^*}}) = g^2_{k^*}(F(X_{I^1_{k^*}}, \{g^2_k\}_{k\in\mathcal{A}}))$ and for all other $k \in \mathcal{A}$, $k \neq k^*$, $g^1_k(\cdot) = g^2_k(\cdot)$. Then, the strategies $\{g^2_k\}_{k\in\mathcal{A}}$ in $\mathcal{T}_2$ and $\{g^1_k\}_{k\in\mathcal{A}}$ in $\mathcal{T}_2$ induce the same joint measure on $X_{\mathcal{N}}$ which is given by (1), and therefore, have the same performance. Thus, for any strategy of $\mathcal{T}_2$, we can find a strategy in $\mathcal{T}_1$ that has the same performance. Consequently, the value of $\mathcal{T}_2$ is less than or equal to the value of $\mathcal{T}_1$.

We have shown that the value of $\mathcal{T}_1$ should be less than or equal to as well as greater than or equal to the value of $\mathcal{T}_2$. Hence, the value of $\mathcal{T}_1$ and $\mathcal{T}_2$ must be the same. This completes the proof of the third property of equivalence of team forms. ■

We now reconsider the examples of Section IV and complete their team forms.

Example 1 (continued): The team form of a Markov decision process, which is shown in Figure 2, is complete.

Example 2 (continued): The team form of the simple two agent team, which is shown in Figure 4, is not complete. All factor nodes have strict functional opbservations. Depending on the order in which we proceed, we end up with different completions. Let $\mathcal{G}_1$ denote the team form shown in Figure 6. Suppose we start with factor node g^2_2. For this node, $O^-_{\mathcal{G}_1}(g^2_2) = \{A^1_1\}$. So, we add an edge (A^2_1, g^2_2) in $\mathcal{G}_1$ and get a new graph $\mathcal{G}_2$. Now, suppose we pick g^2_2 again in $\mathcal{G}_2$. Since $O^-_{\mathcal{G}_2}(g^2_2) = \{A^1_2\}$, we add the edge (A^1_2, g^2_2) in $\mathcal{G}_2$ to get $\mathcal{G}_3$. Next, lets pick g^2_1 in $\mathcal{G}_3$. For this node, $O^-_{\mathcal{G}_3}(g^2_1) = \{A^1_1\}$. So, we add an edge (A^2_1, g^2_1) in $\mathcal{G}_3$ and get a new graph $\mathcal{G}_4$, which is complete. $\mathcal{G}_4$ is shown in Figure 9(a). The added edges are shown by thick lines. Had we proceeded by picking g^2_2, g^1_2, and g^1_1, we would get Figure 9(b) as the completion. Other orderings can result in different completions.

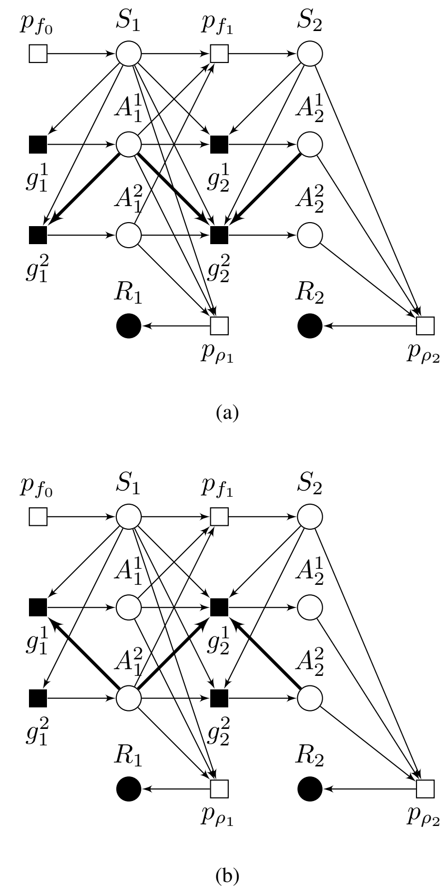

Fig. 9. Two of the completions of the team form of Figure 4

Example 3 (continued): The team form of the simple two agent team with common randomness, shown in Figure 6, is not complete. As in Example 2, All control factor nodes have strict functional observations. Depending on the order in which we proceed, we end up with different completions. Since the process $\{Z_t\}$ is observed by both the control stations, the completions of this team form are similar to the completions of the team form of Example 2. One of these completions is shown in Figure 10.

Example 4 (continued): The team form of the real-time

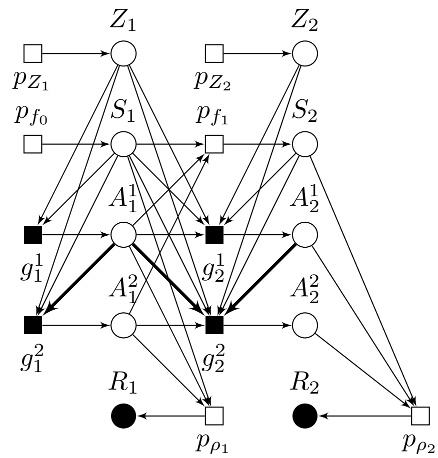

Fig. 10. One of the completions of the team form of Figure 6

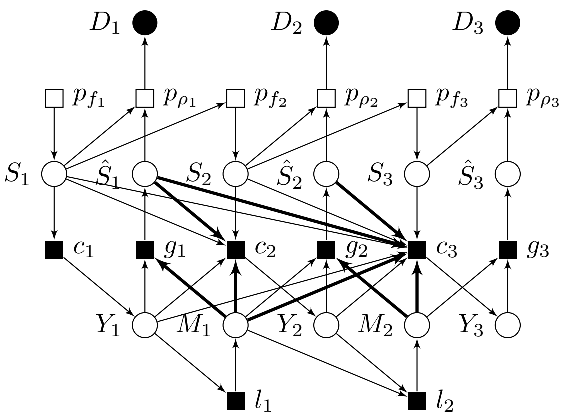

Fig. 11. One of the completions of the team form of Figure 8

communication system, which is shown in Figure 8, is not complete. To complete it we can proceed as follows. Since M_1 is a function of Y_1, M_1 is a strict observation at g_1, c_2 and c_3. So, we add edges (M_1, g_1), (M_1, c_2), and (M_1, c_3) to obtain a new graph $\mathcal{G}_2$. Now, in $\mathcal{G}_2$, M_2 is a strict observation at g_2 and c_3, so we can add the edges (M_2, g_2) and (M_2, c_3) to obtain a new graph $\mathcal{G}_3$. The graph $\mathcal{G}_3$ is complete and is shown in Figure 11. (The thick lines show the new edges). Had we proceeded in a different order, we could have ended with a complete graph with edges $(\hat{S}_1, l_1)$, $(\hat{S}_2, l_2)$, $(\hat{S}_3, l_3)$, instead of the edges (M_1, g_1), (M_2, g_2) in $\mathcal{G}_3$.

VI. Simplification of sequential team form

In this section, we assume that the team form that we want to simplify is complete. If not, we start with any one of its completions. The simplification of team form is based on two ideas.

1) For any control law g_k, there is no loss of optimality in ignoring the subset of available data that is independent of future rewards (rewards that depend on X_k) given X_{I_k} and X_k.
2) For any $B \subset A$, we consider a problem with a planner b that knows the information known to all agents in B and uses this information to compute the control laws of all agents in B. This problem with a planner is equivalent to the original problem. Furthermore, if some information is irrelevant to the planner then all agents in B can ignore that information without any loss of optimality.

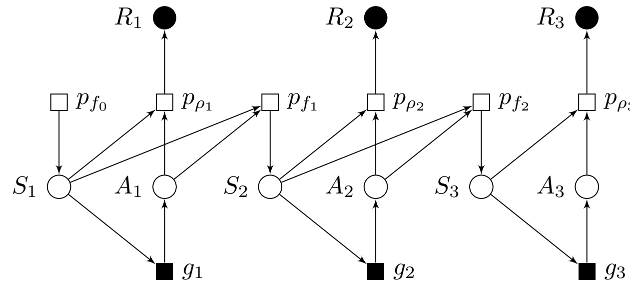

Fig. 12. Simplification of the team form of Figure 2

Due to lack of space, we only present the first simplification here. The second simplification is presented in the extended version of this paper [21].

Theorem 1: Let $\mathcal{T} = (N, A, R, \{I_k\}_{k\in N})$ be a *complete* team form and $\mathcal{G} = (V, F, E, D)$ its graphical model. Then, without loss of optimality, we can choose X_k according to $X_k = g_k(X_M)$ where $X_M = X_{I_k} \setminus \big(R^-_{\mathcal{G}}(X_R \cap \overrightarrow{X}_k \mid X_{I_k}, X_k) \cup \{X_k\}\big)$.

Proof: Fix the policy g_{-k} of all controllers $k' \in A \setminus k$. Since the σ-fields $\{\mathfrak{F}_i\}_{i\in N}$ are either Borelian or countably generated, the total cost conditioned on the information and control action of controller k does not depend on g_k [22]. So, we can write

$$\mathbb{E}^{(g_k, g_{-k})}\Big\{\sum_{i\in R} X_i | X_{I_k}, X_k\Big\} = \mathbb{E}^{g_{-k}}\Big\{\sum_{i\in R} X_i | X_{I_k}, X_k\Big\}.$$

Partition the set of reward variables X_R into two groups: $X_C = X_R \cap \overrightarrow{X}_k$ and $X_{R\setminus C}$. Then

$$\mathbb{E}^{g_{-k}}\Big\{\sum_{i\in R} X_i | X_{I_k}, X_k\Big\} = \mathbb{E}^{g_{-k}}\Big\{\sum_{i\in C} X_i | X_{I_k}, X_k\Big\} + \mathbb{E}^{g_{-k}}\Big\{\sum_{i\in R\setminus C} X_i | X_{I_k}, X_k\Big\}. \quad (5)$$

For any $i \in R \setminus C$, $X_k \not\rightsquigarrow X_i$, thus $P^{g_{-k}}(dX_i \mid X_{I_k}, X_k) = P^{g_{-k}}(dX_i \mid X_{I_k})$. Hence, the second term of (5) does not depend on the choice of X_k. Thus, in order to choose an optimal g_k, we only need to consider the first term of (5). By definition of deterministic irrelevant nodes, we can write

$$\mathbb{E}^{g_{-k}}\left\{\sum_{i\in C} X_i \,\middle|\, X_{I_k}, X_k\right\} = \mathbb{E}^{g_{-k}}\left\{\sum_{i\in C} X_i \,\middle|\, X_M, X_k\right\} =: F_k(X_M, X_k; g_{-k})$$

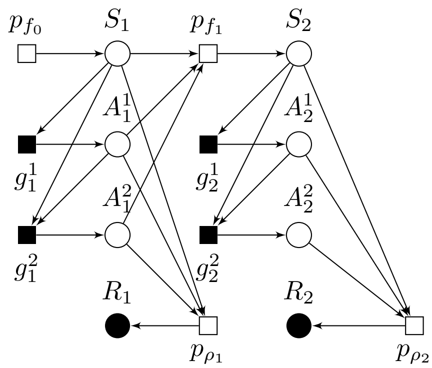

Fig. 13. Simplification of the team form of Figure 9(a)

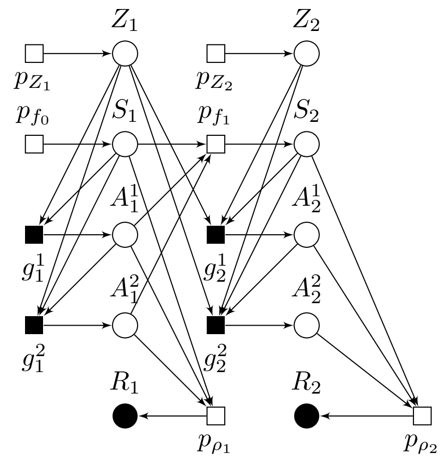

Fig. 14. Simplification of the team form of Figure 10

where

$$X_M = X_{I_k} \setminus \big(R^-_{\mathcal{G}}(X_C \mid X_{I_k}, X_k) \cup \{X_k\}\big).$$

For a fixed g_{-k}, the optimization problem at controller k is

$$\max_{g_k} \mathbb{E}^{(g_k, g_{-k})} \left\{ \sum_{i \in R} X_i \,\middle|\, X_{I_k}, X_k \right\}$$
$$= \max_{g_k} \{F_k(X_M, X_k; g_{-k})\} + \mathbb{E}^{g_{-k}} \left\{ \sum_{i \in R \setminus C} X_i \,\middle|\, X_{I_k} \right\}.$$

Thus, controller k has to solve

$$\max_{g_k} F_k(X_M, g_k(X_M, X_{I_k \setminus M}; g_{-k}).$$

By the method described in [19, Appendix], for any g_{-k} we can choose a function $\hat{g}_k(X_M)$ that is $\mathfrak{F}_k/\mathfrak{F}_M$ measurable such that for any g_k

$$F_k(X_M, g_k(X_M, X_{I_k \setminus M}; g_{-k}) \leq F_k(X_M, \hat{g}_k(X_M)).$$

Thus, without loss of optimality, we can choose a control law of the form

$$X_k = \hat{g}_k(X_M)$$

which proces the result of the theorem. ■

Theorem 1 implies that in the graphical model, we can drop the edges

$$\{(i^0, k^1) : i \in R^-_{\mathcal{G}}(X_R \cap \overrightarrow{X}_k \mid X_{I_k}, X_k)\}.$$

As stated in Section III-D, deterministic irrelevant nodes can be determinied using D-separation. Thus, a team form can be simplified as follows.

Let $\mathcal{T} = (N, A, R, \{I_k\}_{k \in N})$ be a team form and $\mathcal{G} = (V, F, E, D)$ be its graphical model. If for all f in D, $R^-_{\mathcal{G}}(X_R \cap \overrightarrow{X}_{N^+_{\mathcal{G}}(f)} \mid X_{N_{\mathcal{G}}(f)}) = X_{N_{\mathcal{G}}(f)}$ then $\mathcal{T}$ is simplified and we are done. Otherwise, pick a f in $\mathcal{D}$ such that $X_M = R^-_{\mathcal{G}}(X_R \cap \overrightarrow{X}_{N^+_{\mathcal{G}}(f)} \mid X_{N_{\mathcal{G}}(f)}) \subset X_{N_{\mathcal{G}}(f)}$. Create a new graphical model $\mathcal{G}' = (V', F', E', D')$ with $\mathcal{V}' = \mathcal{V}, \mathcal{F}' = \mathcal{F}$, $\mathcal{E}' = \mathcal{E} \setminus \{(x, f) : x \in X_M\}$. and $\mathcal{D}' = \mathcal{D}$. If $\mathcal{G}'$ is simplified, then stop; otherwise, repeat the above procedure with $\mathcal{G}'$. Since the number of edges in the graphical model are finite, the above process, which removes edges from the graph at each step, will always terminate. Theorem 1 implies that the team form $\mathcal{T}'$ corresponding to the terminal graphical model is equivalent to $\mathcal{T}$. Thus, $\mathcal{T}'$ is a simplification of $\mathcal{T}$.

We now reconsider the examples of Section IV and simplify their team forms.

Example 1 (continued): Let $\mathcal{G}_1$ be the team form shown in Figure 2. Start with factor node g_3. The irrelevant nodes for $\{R_3\}$ given $X_{N_{\mathcal{G}_1}(g_3)}$ are $\{S_1, A_1, S_2, A_2\}$. So, remove the edges (S_1, g_3), (A_1, g_3), (S_2, g_3), (A_2, g_3) and label the new graph $\mathcal{G}_2$. Now, pick factor node g_2 in $\mathcal{G}_2$. The irrelevant nodes for $\{R_2, R_3\}$ given $X_{N_{\mathcal{G}_2}(g_2)}$ are $\{S_1, A_1\}$. So, remove the edges (S_1, g_2), (A_1, g_2) and label the new graph $\mathcal{G}_3$. $\mathcal{G}_3$ is shown in Figure 12. No control factor nodes in $\mathcal{G}_3$ have any irrelevant observations. This graphical model corresponds to a team form where control actions are chosen according to

$$A_t = g_t(S_t).$$

Thus, we have obtained the structural result for MDP [23] using the above process.

Example 2 (continued): We start with the completion shown in Figure 9(a). Label the corresponding graphical model be $\mathcal{G}_1$. Start with factor node g_2^2. The irrelevant nodes for R_2 given $X_{N_{\mathcal{G}_1}(g_2^2)}$ are $\{S_1, A_1^1, A_1^2\}$. So, remove the edges (S_1, g_2^2), (A_1^1, g_2^2), and (A_1^2, g_2^2) from $\mathcal{G}_1$ and label the resulting graph $\mathcal{G}_2$. Now, pick factor node g_2^1 in $\mathcal{G}_2$. Remove the edges (S_1, g_2^1), (A_1^1, g_2^1), (A_1^2, g_2^1), which are the edges from the irrelevant nodes, from $\mathcal{G}_2$ and label the resultant graph $\mathcal{G}_3$, which is shown in Figure 13. In the corresponding team form, control actions are chosen according to

$$A_t^1 = g_t^1(S_t), \quad A_t^2 = g_t^2(S_t, A_t^1).$$

Example 3 (continued): Proceeding in a manner similar to Example 2, we obtain a simplified graph shown in Figure 14. In the corresponding team form, control actions are chosen

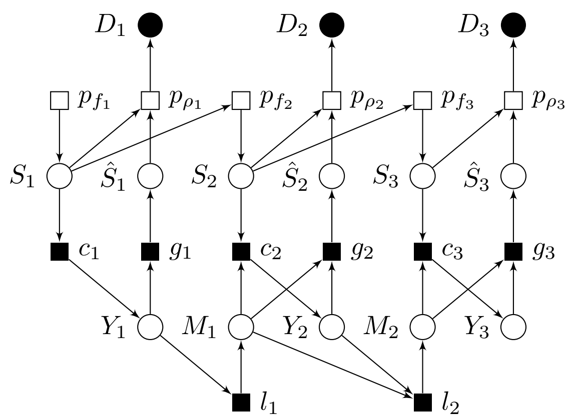

Fig. 15. Simplification of the team form of Figure 11

according to

$$A_t^1 = g_t^1(S_t, Z_1, \dots, Z_t), \quad A_t^2 = g_t^2(S_t, A_t^1, Z_1, \dots, Z_t).$$

The above example shows the limitation of Theorem 1. It can be shown that both control stations can ignore $(Z_1, \dots, Z_t)$; but the simplification according to Theorem 1 cannot capture this. To remove the corresponding edges, we need to look at a coordinator for a subset of controllers. See [21] for details.

Example 4 (continued): Start with the completion shown in Figure 11. Pick factor node g_2. Edge (M_2, g_2) can be removed. Then, pick node g_1. Edge (M_1, g_1) can be removed. Observe that these two edges were added during the completion step. Next pick factor node c_3. Edges (Y_1, c_3), (M_1, c_3), $(\hat{S}_1, c_3)$, (Y_2, c_3), $(\hat{S}_2, c_3)$ can be removed. Pick factor c_2. Edges (Y_1, c_2) and $(\hat{S}_1, c_2)$ can be removed. The resultant graphical model is shown in Figure 15. Thus, in the corresponding team form, the quantized symbols are generated according to

$$Y_t = c_t(S_t, M_{t-1}).$$

We have obtained the structural results of real-time communication [19] using the above process.

VII. Conclusion

In this paper, we formalized a method for obtaining structural results for team problems. We defined a team form as a team where the measurable spaces and stochastic kernels are unspecified. In this setup, structural results mean that we can remove some elements from the information sets of controllers without affecting the value of the team. We modeled a team form a DAFG (directed acyclic factor graph). In a DAFG, structural results mean that we can remove some incoming edge to control factor nodes, without affecting the value of the corresponding team. Such edges can be identified from the conditional independence properties between appropriate set of nodes in the graph. This model can be used as a pedagogical tool for understanding structural results for sequential teams as well as a computational tool for deriving such structural results in an automated manner.

Acknowledgement

The authors are grateful to Achilleas Anastasopoulos, Ashutosh Nayyar, and Demosthenis Teneketzis for extensive discussions and useful feedback.

References

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

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

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

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

[5] G. Casalino, F. Davoli, R. Minciardi, P. Puliafito, and R. Zoppoli, "Partially nested information structures with a common past," *IEEE Trans. Autom. Control*, vol. 29, no. 9, pp. 846–850, Sep. 1984.

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

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

[8] J. M. Ooi, S. M. Verbout, J. T. Ludwig, and G. W. 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.

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

[10] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Autom. Control*, Dec. 2009, to appear.

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

[12] H. S. Witsenhausen, "On information structures, feedback and causality," *SIAM Journal of Control*, vol. 9, no. 2, pp. 149–160, May 1971.

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

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

[15] ——, "Equivalent stochastic control problems," *Mathematics of Control Signals and Systems*, vol. 1, pp. 3–11, 1988.

[16] D. Geiger, T. Verma, and J. Pearl, "d-separation: from theorem to algorithms," in *Proceedings on the fifth Workshop on Uncertainity in Artificial Intelligence*, Windsor, Ontario, 1989, pp. 118–125.

[17] ——, "Identifying independence in Bayesian networks," *Networks*, vol. 20, no. 5, pp. 507–534, 1990.

[18] R. D. Shachter, "Bayes-ball: The rational pastime (for determining irrelevance and requisite information in belief networks and influence diagrams)," in *proceedings of the fourteenth conference in Uncertainty in Artificial Intelligence*, 1998, pp. 480–487.

[19] H. S. Witsenhausen, "On the structure of real-time source coders," *Bell System Technical Journal*, vol. 58, no. 6, pp. 1437–1451, July-August 1979.

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

[21] A. Mahajan and S. Tatikonda, "Sequential team form and its simplification using graphical models," *submitted toIEEE Trans. Autom. Control*, 2009.

[22] H. S. Witsenhausen, "On policy independence of conditional expectation," *Information and Control*, vol. 28, pp. 65–75, 1975.

[23] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Prentice Hall, 1986.