2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On the Structure of State-Feedback LQG Controllers for Distributed Systems With Communication Delays

Andrew Lamperski and John C. Doyle

***Abstract*— This paper presents explicit solutions for a few distributed LQG problems in which players communicate their states with delays. The resulting control structure is reminiscent of a simple management hierarchy, in which a top level input is modified by newer, more localized information as it gets passed down the chain of command. It is hoped that the controller forms arising through optimization may lend insight into the control strategies of biological and social systems with communication delays.**

I. Introduction

Imagine a large event, such as a conference, is being planned. A group of people must move chairs and tables, prepare and serve food, and set up audio-visual equipment. In addition, typically, there are also people who do little or none of the physical, work, but whose main function is coordination. Someone oversees the food, while someone else might coordinate the audio-visual equipment. At the top level, there is often an individual, or group, that manages the coordinators. Management hierarchies, such as the one just described, are common, even though they are not actually necessary to perform the actions needed for the event.

This paper studies how hierarchical control structures can arise as optimal methods to deal with communication delays. While the problem studied is a simple variant of distributed linear quadratic Gaussian (LQG) control, the results have some intuitive similarities with the event planning example discussed above. In the problems studied, a group of players works together to minimize a quadratic cost. The players have access to local state information, but can only communicate their state with a delay. This paper shows that for such problems, a hierarchical control structure emerges as the optimal strategy. In particular, the optimal inputs can be decomposed into partially ordered components. The lowest level components represent the players (which do the physical work), while higher level components are used for coordination.

A. Related Work

The focus of this paper, LQG control with communication delays, is a basic problem in distributed or decentralized control. Decentralized control has a long history, [1], [2], [3], [4], [5], but computationally tractable solutions to nontrivial problems have been rare, until recently. Notably, in the past ten years, certain decentralized optimal control problems were shown to be convex [6], [7]. More recently, computationally efficient solutions to some of the convex problems have been found [8], [9], [10], [11], [12]. Of the work cited, Rantzer's paper on linear quadratic teams is the most closely related [8]. In this paper, Rantzer solves the problem studied in this chapter using semidefinite programming (SDP), but does not explore the structure of the solution. The solution techniques used in this chapter are closely related to the dynamic programming methods used in [11]. Also related are the works of [3], [4], [5], which give solutions to the problem of Section III. The work in Section III differs from these works, in that it naturally leads to generalizations.

B. Motivation and Contributions

The main motivation for studying the present problem is the analysis of biological and social systems. Thus, the primary goal is to develop intuition about the consequences of communication delay in optimized control systems.

Given the focus on intuition, the primary contribution of this paper is the explicit structure of the optimal controllers found. While the problems of this paper can be solved using the SDP method of [8], the structures of the optimal solutions are not explicit. Using a novel derivation, this paper shows that by simply assuming that communication between players is delayed, a control hierarchy arises as the optimal solution. Furthermore, the optimal solution is computed from a standard algebraic Riccati equation. The new knowledge about the control structure can give insight into hierarchies found in existing control systems.

Finally, the techniques developed in this paper are useful for other problems in decentralized control. Indeed, the dynamic programming arguments below can be modified to recover the results of [10], [11], [12].

C. Overview

The article is structured as follows. Section II defines the problems studied in the paper. Section III gives a full solution for the first problem, called the two-player problem. This section develops basic ideas on decoupling information into independent terms, and the associated dynamic programming problem. Next, a solution for a more complex example termed the three-player chain is sketched in Section IV. To solve this problem, a systematic method for decoupling information based on how it is shared by the players is introduced. Conclusions and future work are outlined in Section V.

Notation. The expected value of a random variable, x, is denoted by $\mathbb{E}[x]$. The conditional expectation of x given y is denoted by $\mathbb{E}[x|y]$. The sequence $x(0), x(1), \ldots, x(t)$ is denoted by $x(0:t)$.

A. Lamperski and J. C. Doyle are with Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125 andyl@cds.caltech.edu

978-1-61284-799-3/11/$26.00 ©2011 IEEE

II. Problem Statement

Consider dynamics of the form

$$x(t+1) = Ax(t) + Bu(t) + w(t),$$

with initial condition $x(0) = 0$. Here (A, B) is assumed to be stabilizable. The state is given by x, u is the input, and w is process noise. In the two problems studied, the goal is to minimize the steady-state cost

$$\lim_{t\to\infty} \mathbb{E}\left[x(t)^T Q x(t) + u(t)^T R u(t)\right]$$

subject to input constraints to be specified below. The matrix Q is positive semidefinite, and R is positive definite. To guarantee a stabilizing solution to the corresponding algebraic Riccati equation, $(\sqrt{Q}, A)$ is assumed to be stabilizable. Aside from that, no other assumptions are made about Q and R.

The problems considered in this paper are referred to as the *two-player problem* and the *three-player chain*. The ideas required to solve these problems extend naturally to a more general class of decentralized control problems [13].

For the two-player problem, the state, input, and process noise are partitioned as

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix},$$

where $x_i \in \mathbb{R}^{k_i}$, $u_i \in \mathbb{R}^{p_i}$, and $w_i \in \mathbb{R}^{k_i}$. The state matrices have sparsity structure (conforming to the partitions of x and u)

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix},$$

and inputs are restricted to the form

$$\begin{aligned} u_1(t) &= \gamma_{1,t}(x_1(0:t), x_2(0:t-1)) \\ u_2(t) &= \gamma_{2,t}(x_1(0:t-1), x_2(0:t)). \end{aligned} \tag{1}$$

Here $\gamma_{i,t}$ are Borel-measurable functions to be chosen in the optimization procedure. The process noise $w(t)$ is Gaussian white noise with covariance given by

$$\mathbb{E}[ww^T] = \mathbb{E}\left[\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \begin{bmatrix} w_1^T & w_2^T \end{bmatrix}\right] = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}.$$

For the three-player chain, the state, input, and process noise are partitioned as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix},$$

where $x_i \in \mathbb{R}^{k_i}$, $u_i \in \mathbb{R}^{p_i}$, and $w_i \in \mathbb{R}^{k_i}$. The state matrices have the sparsity

$$A = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & B_3 \end{bmatrix},$$

and the inputs have the delay structure

$$\begin{aligned} u_1(t) &= \gamma_{1,t}(x_1(0:t), x_2(0:t-1), x_3(0:t-2)) \\ u_2(t) &= \gamma_{2,t}(x_1(0:t-1), x_2(0:t), x_3(0:t-1)) \\ u_3(t) &= \gamma_{3,t}(x_1(0:t-2), x_2(0:t-1), x_3(0:t)). \end{aligned} \tag{2}$$

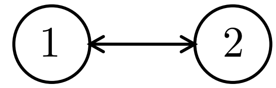

Fig. 1. Two-player graph.

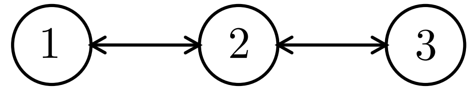

Fig. 2. Three-player chain.

In this case the process noise is Gaussian white noise with covariance given by

$$\mathbb{E}[ww^T] = \mathbb{E}\left[\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \begin{bmatrix} w_1^T & w_2^T & w_3^T \end{bmatrix}\right] = \begin{bmatrix} W_1 & 0 & 0 \\ 0 & W_2 & 0 \\ 0 & 0 & W_3 \end{bmatrix}.$$

Throughout the paper, $u_i(t)$ will be interpreted as the "input chosen by player i" at time t. Note that for the two-player problem, Equation (1) shows that at time $t+1$ player 2 has access to all of the information available to player 1 at time t, and vice versa. This can be interpreted as players 1 and 2 communicating their information after a one-step delay. Similarly, Equation (2) specifies that, in the three-player chain, players 1 and 2 communicate their information after a one-step delay and players 2 and 3 communicate their information after a one-step delay. The delay patterns from the two-player problem and the three-player chain can be captured by the graphs shown in Figures 1 and 2, respectively.

Note that the assumptions about the structure of the input and the sparsity of the dynamics guarantee that communication between the players choosing u_i occurs as least as fast as information travels through the plant. This assumption implies that the information structure (the set of input constraints) is *partially nested*, which in turn implies that the optimal inputs are linear in the associated information [1].

In the deriving the optimal controller, the following finite-horizon variant of the control problem is studied. Minimize

$$\mathbb{E}\left[\sum_{t=0}^{N-1} \left(x(t)^T Q x(t) + u(t)^T R u(t)\right) + x(N)^T \Lambda x(N)\right] \tag{3}$$

with inputs of the form of Equation (1) or Equation (2). Here Λ is a positive semidefinite matrix of appropriate dimensions, corresponding to a terminal cost. If Λ is positive definite, then as $N \to \infty$, the optimal controller for this finite-horizon problem approaches the steady-state controller.

III. Two-Player Problem

This section presents a solution to the two-player problem which extends to other delay patterns. The delay pattern of the two-player problem is often referred to in the literature as the "one-step delay information pattern." Dynamic programming solutions for the this problem have been known since the 1970s [3], [4], [5]. The approaches in the cited work differ from the work of this paper, and it is not immediately clear how to generalize them to other delay structures, such as the three-player chain. The method of this paper is to decompose the information into independent components, a priori, and use this decomposition to decouple the dynamic programming problem into independent subproblems.

The section is organized as follows. First the optimal controller is presented in Subsection III-A. Subsection III-B derives the optimal solution to a finite-horizon version of the optimal control problem. Finally, the optimal controller is derived in Subsection III-C by applying a limiting argument.

A. Two-Player Problem: Optimal Solution

In order to find a structure for the optimal controller, decompose $x(t)$ into three independent terms

$$x(t) = \begin{bmatrix}\zeta_1(t)\\ \zeta_2(t)\end{bmatrix} + \hat{x}(t), \tag{4}$$

where $\hat{x}(t) = \mathbb{E}[x(t)|x(0:t-1)]$. Since the input $u(t-1)$ depends on $x(0:t-1)$, it follows that

$$\begin{aligned}\hat{x}(t) &= Ax(t-1) + Bu(t-1)\\ \zeta_1(t) &= w_1(t-1)\\ \zeta_2(t) &= w_2(t-1).\end{aligned} \tag{5}$$

Thus $\hat{x}(t)$, $\zeta_1(t)$, and $\zeta_2(t)$ are, indeed, pairwise independent. The term $\hat{x}(t)$ denotes the expected value of $x(t)$ given the information shared by both player 1 and player 2. The term $\zeta_1(t)$ depends on the information available only to player 1, and similarly $\zeta_2(t)$ depends on the information available only to player 2.

Theorem 1: *There exist matrices K, H_1, H_2, X_1, and X_2, such that the optimal controller for the two-player problem is given by*

$$u(t) = -\begin{bmatrix}H_1\zeta_1(t)\\ H_2\zeta_2(t)\end{bmatrix} - K\hat{x}(t),$$

and the optimal cost is given by

$$\mathrm{Tr}(W_1X_1) + \mathrm{Tr}(W_2X_2). \tag{6}$$

Remark 1: The input $-K\hat{x}(t)$ could be interpreted as a command sent by a "manager" based information $x(0:t-1)$. Player 1 then applies a correction term $-H_1\zeta_1(t)$, based on newer information unavailable to the "manager." Similarly, $-H_2\zeta_2(t)$ represents player 2's correction term.

The gains, as well as the costs are specified by the stabilizing solution to the algebraic Riccati equation, S:

$$S = Q + A^TSA - A^TSB(R+B^TSB)^{-1}B^TSA.$$

For more compact notation, define the block columns of A and B as

$$\begin{aligned}\left[\begin{array}{c|c} A_1 & A_2\end{array}\right] &= \left[\begin{array}{c|c} A_{11} & A_{12}\\ A_{21} & A_{22}\end{array}\right],\\ \left[\begin{array}{c|c} B_1 & B_2\end{array}\right] &= \left[\begin{array}{c|c} B_{11} & 0\\ 0 & B_{22}\end{array}\right].\end{aligned} \tag{7}$$

The gains are then given by

$$\begin{aligned}K &= (R+B^TSB)^{-1}B^TSA\\ H_1 &= (R_{11}+B_1^TSB_1)^{-1}B_1^TSA_1\\ H_2 &= (R_{22}+B_2^TSB_2)^{-1}B_2^TSA_2,\end{aligned}$$

and the cost matrices, X_1 and X_2, are given by

$$\begin{aligned}X_1 &= \begin{array}{c}Q_{11} + A_1^TSA_1\\ -A_1^TSB_1(R_{11}+B_1^TSB_1)^{-1}B_1^TSA_1\end{array}\\ X_2 &= \begin{array}{c}Q_{22} + A_2^TSA_2\\ -A_2^TSB_2(R_{22}+B_2^TSB_2)^{-1}B_2^TSA_2.\end{array}\end{aligned}$$

B. Two-Player Problem: Finite-Horizon Derivation

In order to derive the optimal controller, the finite-horizon version, with cost given by Equation (3), will be solved, and the infinite-horizon version follows by taking limits.

The following lemma shows how an input structure based on the distribution of information between the players can be assumed. The proof is omitted for space.

Lemma 1: *The optimal input can be decomposed as*

$$u(t) = \begin{bmatrix}\varphi_1(t)\\ \varphi_2(t)\end{bmatrix} + \hat{u}(t), \tag{8}$$

where $\varphi_1(t)$, $\varphi_2(t)$, and $\hat{u}(t)$ are independent random variables which are linear functions of $\zeta_1(t)$, $\zeta_2(t)$, and $x(0:t-1)$, respectively.

The lemma combined with the decomposition of the state in terms of $\hat{x}$, ζ_1, and ζ_2 implies that the summand of the cost function can be decomposed as

$$\begin{aligned}&\mathbb{E}\left[x(t)^TQx(t) + u(t)^TRu(t)\right]\\ &\quad= \mathbb{E}\left[\hat{x}(t)^TQ\hat{x}(t) + \hat{u}(t)^TR\hat{u}(t)\right]\\ &\qquad+\mathbb{E}\left[\zeta_1(t)^TQ_{11}\zeta_1(t) + \varphi_1(t)^TR_{11}\varphi_1(t)\right]\\ &\qquad+\mathbb{E}\left[\zeta_2(t)^TQ_{22}\zeta_2(t) + \varphi_2(t)^TR_{22}\varphi_2(t)\right].\end{aligned} \tag{9}$$

The solution will proceed via a dynamic programming argument. Let $\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,t)]$ denote the optimal expected cost-to-go function, when the state is decomposed as $\hat{x}$, ζ_1, and ζ_2 at time t. By independence, $\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,N)]$ can be decoupled as

$$\begin{aligned}&\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,N)]\\ &\quad= \mathbb{E}\left[\hat{x}^T\Lambda\hat{x}\right] + \mathbb{E}\left[\zeta_1^T\Lambda_{11}\zeta_1\right] + \mathbb{E}\left[\zeta_2^T\Lambda_{22}\zeta_2\right].\end{aligned}$$

Let $S(N) = \Lambda$, $X_1(N) = \Lambda_{11}$, and $X_2(N) = \Lambda_{22}$. For $t \leq N$, it will be shown that $J(\hat{x},\zeta_1,\zeta_2,t)$ has the form

$$\begin{aligned}J(\hat{x},\zeta_1,\zeta_2,t) &= \hat{x}^TS(t)\hat{x} + \zeta_1^TX_1(t)\zeta_1 + \zeta_2^TX_2(t)\zeta_2\\ &\quad+ \sum_{j=t+1}^{N}\left(\mathrm{Tr}(W_1X(j)) + \mathrm{Tr}(W_2X_2(j))\right),\end{aligned} \tag{10}$$

for some matrices $S(t)$, $X_1(t)$, and $X_2(t)$ to be specified.

Inductively assume that $J(\hat{x},\zeta_1,\zeta_2,t+1)$ has the form given in Equation (10). Then $\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,t)]$ is given by the Bellman equation:

$$\begin{aligned}&\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,t)] =\\ &\quad\min_{\hat{u},\zeta_1,\zeta_2}\mathbb{E}\left[x^TQx + u^TRu + J(Ax+Bu,w_1,w_2,t+1)\right].\end{aligned}$$

Note that $J(Ax+Bu, w_1, w_2, t+1)$ can be expanded as

$$\begin{aligned} J(Ax+Bu,w_1,w_2,t+1) = \\ &(Ax+Bu)^T S(t+1)(Ax+Bu) \\ &+w_1^T X_1(t+1)w_1 + w_2^T X_2(t+1)w_2 \\ &+ \sum_{j=t+2}^{N} (\mathrm{Tr}(W_1 X_1(j)) + \mathrm{Tr}(W_2 X_2(j))). \end{aligned} \tag{11}$$

The expected value of the third and fourth lines of Equation (11) can be grouped as

$$\sum_{j=t+1}^{N} (\mathrm{Tr}(W_1 X_1(j)) + \mathrm{Tr}(W_2 X_2(j))).$$

The independence of the terms in Equations (4) and (8) implies that the expected value of the first term on the right-hand side of Equation (11) can be expanded further as follows:

$$\begin{aligned} \mathbb{E}[(Ax+Bu)^T S(t+1)(Ax+Bu)] = \\ &\mathbb{E}[(A\hat{x}+B\hat{u})^T S(t+1)(A\hat{x}+B\hat{u})] \\ &+\mathbb{E}\left[(A_1\zeta_1+B_1\varphi_1)^T S(t+1)\,(A_1\zeta_1+B_1\varphi_1)\right] \\ &+\mathbb{E}\left[(A_2\zeta_2+B_2\varphi_2)^T S(t+1)\,(A_2\zeta_2+B_2\varphi_2)\right]. \end{aligned} \tag{12}$$

Here A_1, A_2, B_1, and B_2 are the block columns defined in Equation (7). Note that in the expansion, independent cross terms are set to zero.

Combining Equations (9), (11), and (12) shows that the right-hand side of the Bellman equation can be decomposed into three independent minimizations, plus a constant term:

$$\begin{aligned} &\min_{\hat{u},\zeta_1,\zeta_2} \mathbb{E}\left[x^T Qx + u^T Ru + J(Ax+Bu,w_1,w_2,t+1)\right] = \\ &\min_{\hat{u}} \mathbb{E}\left[\Psi(\hat{x},\hat{u})\right] + \min_{\varphi_1} \mathbb{E}\left[\Theta_1(\zeta_1,\varphi_1)\right] + \min_{\varphi_2} \mathbb{E}\left[\Theta_2(\zeta_2,\varphi_2)\right] \\ &\quad + \sum_{j=t+1}^{N} (\mathrm{Tr}(W_1 X_1(j)) + \mathrm{Tr}(W_2 X_2(j))), \end{aligned}$$

with functions given by

$$\begin{aligned} \Psi(\hat{x},\hat{u}) &= \hat{x}^T Q\hat{x} + \hat{u}^T R\hat{u} + \\ &\quad (A\hat{x}+B\hat{u})^T S(t+1)(A\hat{x}+B\hat{u}) \\ \Theta_1(\zeta_1,\varphi_1) &= \zeta_1^T Q_{11}\zeta_1 + \varphi_1^T R_{11}\varphi_1 + \\ &\quad (A_1\zeta_1+B_1\varphi_1)^T S(t+1)(A_1\zeta_1+B_1\varphi_1) \\ \Theta_2(\zeta_2,\varphi_2) &= \zeta_2^T Q_{22}\zeta_2 + \varphi_2^T R_{22}\varphi_2 + \\ &\quad (A_2\zeta_2+B_2\varphi_2)^T S(t+1)(A_2\zeta_2+B_2\varphi_2). \end{aligned}$$

Quadratic minimization shows that the optimal inputs are given by

$$\begin{aligned} \hat{u}(t) &= -K(t)\hat{x}(t) \\ \varphi_1(t) &= -H_1(t)\zeta_1(t) \\ \varphi_2(t) &= -H_2(t)\zeta_2(t), \end{aligned}$$

where the gains are given by

$$\begin{aligned} K(t) &= (R+B^T S(t+1)B)^{-1}B^T S(t+1)A \\ H_1(t) &= (R_{11}+B_1^T S(t+1)B_1)^{-1}B_1^T S(t+1)A_1 \\ H_2(t) &= (R_{22}+B_2^T S(t+1)B_2)^{-1}B_2^T S(t+1)A_2. \end{aligned}$$

Finally, the matrices $S(t)$, $X_1(t)$, and $X_2(t)$ are computed recursively as follows:

$$\begin{gathered} S(t) = Q + A^T S(t+1)A \\ -A^T S(t+1)B(R+B^T S(t+1)B)^{-1}B^T S(t+1)A \\ X_1(t) = Q_{11} + A_1^T S(t+1)A_1 \\ -A_1^T S(t+1)B_1(R_{11}+B_1^T S(t+1)B_1)^{-1}B_1^T S(t+1)A_1 \\ X_2(t) = Q_{22} + A_2^T S(t+1)A_2 \\ -A_2^T S(t+1)B_2(R_{22}+B_2^T S(t+1)B_2)^{-1}B_2^T S(t+1)A_2. \end{gathered}$$

By construction, $J(\hat{x},\zeta_1,\zeta_2,t)$ satisfies the Bellman equation for all $t \le N$. Thus, since $\mathbb{E}[J(\hat{x},\zeta_1,\zeta_2,N)]$ is the optimal expected cost-to-go at time N, it follows inductively that $\mathbb{E}[J(\hat{x},\zeta_1,\zeta_t,t)]$ is the optimal expected cost-to-go for all $t \le N$, and the optimal control has been found. Noting that $x(0) = 0$, the optimal expected cost is given by

$$\sum_{t=1}^{N} (\mathrm{Tr}(W_1 X_1(t)) + \mathrm{Tr}(W_2 X_2(t))). \tag{13}$$

C. Two-Player Problem: Steady State

To derive the steady state regulator from the finite-horizon regulator, assume that as N approaches ∞, $S(t)$ approaches the stabilizing solution of the corresponding algebraic Riccati equation. Then $K(t)$, $H_1(t)$, $H_2(t)$, $X_1(t)$, and $X_2(t)$ will approach the values of K, H_1, H_2, X_1, and X_2 specified by the theorem and the derivation of the controller is complete.

To compute the steady state cost, note that the average cost approaches the steady state cost as $N \to \infty$:

$$\begin{aligned} &\lim_{t\to\infty} \mathbb{E}\left[x(t)^T Qx(t) + u(t)^T Ru(t)\right] = \\ &\quad \lim_{N\to\infty} \tfrac{1}{N}\textstyle\sum_{t=1}^{N} (\mathrm{Tr}(W_1 X_1(t)) + \mathrm{Tr}(W_2 X_2(t))), \end{aligned}$$

where the optimal finite-horizon cost has been taken from Equation (13). Furthermore, since $X_1(t) \to X_1$ and $X_2(t) \to X_2$, Equation (6) follows.

IV. Three-Player Chain

This section studies the three-player chain. While the problem description is not much more complex than that of the two-player problem, the controller structure is noticeably more involved. The ideas presented in this section generalize naturally to other delay structures [13], but the generalization is outside the scope of this paper.

This section is organized as follows. Subsection IV-A describes how to decompose the state and input into independent components. Next, Subsection IV-B presents the optimal controller. Finally, Subsection IV-C sketches a derivation of the controller.

Notation. For a matrix partitioned into blocks

$$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$$

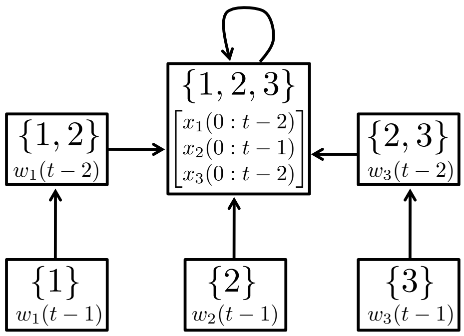

Fig. 3. Information Hierarchy Graph. The nodes correspond to subsets of $\{1,2,3\}$ and are labeled information vectors. If a node $v \subset \{1,2,3\}$ is labeled by a vector $\mathscr{L}_v(t)$, then $\mathscr{L}_v(t)$ is available to player i if $i \in v$ and not available to player i if $i \notin v$.

and $s, v \subset \{1,2,3\}$, let $M^{s,v} = (M_{i,j})_{i \in s, j \in v}$. For instance

$$M^{\{1,2,3\},\{1,2\}} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \\ M_{31} & M_{32} \end{bmatrix}.$$

A. Information Decomposition

This subsection describes how to decouple the state and input into independent random variables, similar to what was done for the two-player problem. First, the information available to the various players is decomposed. Next the state and input are decomposed as functions as of the independent information components.

The information available to the players is decomposed according to a labeled graph, called the *information hierarchy graph*, $\mathscr{I} = (\mathscr{V}, \mathscr{E})$ (Figure 3). The nodes,

$$\mathscr{V} = \{\{1\},\{2\},\{3\},\{1,2\},\{2,3\},\{1,2,3\}\},$$

correspond to sets of players, or alternatively sets of nodes in the delay graph (Figure 2). Each node v is labeled by a vector $\mathscr{L}_v(t)$ which corresponds to information available only to the players in the set v. For instance, at time t, players 1 and 2 share $w_1(t-2)$ but it is unavailable to player 3, and thus $\mathscr{L}_{\{1,2\}}(t) = w_1(t-2)$. Likewise, node $\{1,2,3\}$ is labeled by the common information $(x_1(0:t-2), x_2(0:t-1), x_3(0:t-2))$. Note that the labels, $\mathscr{L}_v(t)$ are pairwise independent. There is an edge (v,s) if and only if s is the set of nodes of the delay graph reachable from v within one step. Thus if $(v,s) \in \mathscr{E}$, then $v \subset s$ and so the edges induce a partial order on the nodes.

Lemma 2: *The state and optimal input can be decomposed as*

$$x(t) = \begin{bmatrix} \zeta_{\{1\}}(t) \\ \zeta_{\{2\}}(t) \\ \zeta_{\{3\}}(t) \end{bmatrix} + \begin{bmatrix} \zeta_{\{1,2\}}(t) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta_{\{2,3\}}(t) \end{bmatrix} + \zeta_{\{1,2,3\}}(t), \tag{14}$$

$$u(t) = \begin{bmatrix} \varphi_{\{1\}}(t) \\ \varphi_{\{2\}}(t) \\ \varphi_{\{3\}}(t) \end{bmatrix} + \begin{bmatrix} \varphi_{\{1,2\}}(t) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \varphi_{\{2,3\}}(t) \end{bmatrix} + \varphi_{\{1,2,3\}}(t), \tag{15}$$

where $\varphi_v(t)$ and $\zeta_v(t)$ are linear functions of the label $\mathscr{L}_v(t)$. Furthermore, the $\zeta_v(t)$ terms are computed by the following dynamic equations:

$$\begin{aligned} \zeta_{\{i\}}(t+1) &= w_i(t),\ i = 1,2,3 \\ \zeta_{\{1,2\}}(t+1) &= A^{\{1,2\},\{1\}}\zeta_{\{1\}}(t) + B^{\{1,2\},\{1\}}\varphi_{\{1\}}(t) \\ \zeta_{\{2,3\}}(t+1) &= A^{\{2,3\},\{3\}}\zeta_{\{3\}}(t) + B^{\{2,3\},\{3\}}\varphi_{\{3\}}(t) \end{aligned} \tag{16}$$

$$\begin{aligned} \zeta_{\{1,2,3\}}(t+1) = \\ & A^{\{1,2,3\},\{1,2,3\}}\zeta_{\{1,2,3\}} + B^{\{1,2,3\},\{1,2,3\}}\varphi_{\{1,2,3\}} + \\ & A^{\{1,2,3\},\{1,2\}}\zeta_{\{1,2\}} + B^{\{1,2,3\},\{1,2\}}\varphi_{\{1,2\}} + \\ & A^{\{1,2,3\},\{2,3\}}\zeta_{\{2,3\}} + B^{\{1,2,3\},\{2,3\}}\varphi_{\{2,3\}} \end{aligned} \tag{17}$$

with initial conditions $\zeta_v(0) = 0$.

Lemma 2 can be proved by noting that the input can be assumed to be linear and inductively applying Equations (16) and (17). Note that the partial order induced by $\mathscr{E}$ can be applied to the input decomposition to distinguish between "high-level" and "low-level" inputs. Furthermore, note that $\zeta_s(t+1)$ depends on $\zeta_v(t)$ if and only if $(v,s) \in \mathscr{E}$. Thus Equations (16) and (17) can be viewed as specifying how information "travels up the information hierarchy graph".

B. Three-Player Chain: Optimal Solution

This section gives the optimal solution for the three-player chain. The controller is structurally similar to the controller for the two-player problem.

Theorem 2: *There are matrices K_v and X_v such that the optimal control has the form of Equation (15) with*

$$\varphi_v(t) = -K_v \zeta_v(t),$$

and the optimal cost is given by

$$\mathrm{Tr}(W_1 X_{\{1\}}) + \mathrm{Tr}(W_2 X_{\{2\}}) + \mathrm{Tr}(W_3 X_{\{3\}}).$$

The gains are computed as

$$K_v = \left(R^{v,v} + B^{s,v\,T} X_s B^{s,v}\right)^{-1} B^{s,v\,T} X_s A^{s,v}, \tag{18}$$

where s is the unique node such that (v,s) is an edge in the information hierarchy graph, $\mathscr{I}$.

The matrices X_v are computed as follows. Let $X_{\{1,2,3\}}$ be the stabilizing solution to the algebraic Riccati equation

$$S = Q + A^T S A - A^T S B (R + B^T S B)^{-1} B^T S A.$$

Then assuming that X_s is defined and (v,s) is an edge, define X_v by

$$\begin{aligned} X_v = Q^{v,v} + A^{s,v\,T} X_s A^{s,v} - \\ A^{s,v\,T} X_s B^{s,v} \left(R^{v,v} + B^{s,v\,T} X_s B^{s,v}\right)^{-1} B^{s,v\,T} X_s A^{s,v}. \end{aligned} \tag{19}$$

Note that the gain corresponding to node $\{1,2,3\}$ is exactly the standard LQR gain. The other gains are found by propagating the solution to the LQR Riccati equation through the information hierarchy graph, based on the equations for X_v.

Remark 2: In terms of a management hierarchy, $\varphi_{\{1,2,3\}}$ could be thought of as the command sent by the "boss," based on delayed global information. Next $\varphi_{\{1,2\}}$ and $\varphi_{\{2,3\}}$ correspond to modifications made by "middle managers" at nodes $\{1,2\}$ and $\{2,3\}$. Finally $\varphi_{\{1\}}$, $\varphi_{\{2\}}$, and $\varphi_{\{3\}}$ are the corrections made at the site of the physical activity, based on the most recent local information.

C. Three-Player Chain: Controller Derivation

This subsection sketches a derivation of the optimal controller. A more complete derivation based on the limits of an associated finite-horizon problem is possible, but omitted for space reasons.

Lemma 2 shows that the cost function can be decomposed as

$$\mathbb{E}\left[x(t)^T Q x(t) + u(t)^T R u(t)\right] = \sum_{v\in\mathscr{V}} \mathbb{E}\left[\zeta_v(t)^T Q^{v,v}\zeta_v(t) + \varphi_v(t)^T R^{v,v}\varphi_v(t)\right]. \tag{20}$$

Using the decomposition of the state and input into ζ_v and φ_v, respectively, a function J and a constant c will be found satisfying the steady-state Bellman equation:

$$\mathbb{E}[J(\zeta)] + c = \min_{\varphi} \mathbb{E}\left[\sum_{v\in\mathscr{V}}\left(\zeta_v^T Q^{v,v}\zeta_v + \varphi_v^T R^{v,v}\varphi_v\right) + J(\zeta')\right], \tag{21}$$

where ζ' corresponds to the updated variables. Note that c will specify the optimal steady-state cost.

The function J is proposed to have the form

$$J(\zeta) = \sum_{v\in\mathscr{V}} \zeta_v^T X_v \zeta_v,$$

where X_v are the matrices defined above.

Applying Equations (16) and (17) and dropping independent cross terms shows that

$$\mathbb{E}[J(\zeta')] = \sum_{i=1}^{3}\mathrm{Tr}(W_i X_{\{i\}}) + \sum_{v\in\mathscr{V}} \mathbb{E}\left[(A^{s,v}\zeta_v + B^{s,v}\varphi_v)^T X_s (A^{s,v}\zeta_v + B^{s,v}\varphi_v)\right], \tag{22}$$

where s is the unique node such that (v,s) is an edge in $\mathscr{I}$.

Combining Equations (21) and (22) shows that the right-hand side of the Bellman equation decomposes into independent minimizations plus a constant term:

$$\min_{\varphi} \mathbb{E}\left[\sum_{v\in\mathscr{V}}\left(\zeta_v^T Q^{v,v}\zeta_v + \varphi_v^T R^{v,v}\varphi_v\right) + J(\zeta')\right] = \sum_{v\in\mathscr{V}} \min_{\varphi_v}\mathbb{E}\left[\Theta_v(\zeta_v,\varphi_v)\right] + \sum_{i=1}^{3}\mathrm{Tr}(W_i X_{\{i\}}),$$

with functions given by

$$\Theta_v(\zeta_v,\varphi_v) = \zeta_v^T Q^{v,v}\zeta_v + \varphi_v^T R^{v,v}\varphi_v + (A^{s,v}\zeta_v + B^{s,v}\varphi_v)^T X_s (A^{s,v}\zeta_v + B^{s,v}\varphi_v).$$

Standard quadratic minimization arguments show that $\mathbb{E}[\Theta_v(\zeta_v,\varphi_v)]$ is minimized by $\varphi_v = -K_v\zeta_v$, where K_v is computed from Equation (18). Plugging in the optimal inputs and applying Equation (19) shows that the Bellman equation is satisfied with $c = \sum_{i=1}^{3}\mathrm{Tr}(W_i X_{\{i\}})$.

V. Conclusion

This paper presents Riccati-based solutions to two distributed control problems with communication delays. The controllers can be interpreted as simple management schemes. In these schemes, a top level "boss" generates an input, based on delayed global information. The input is modified using newer, more localized information as it gets passed down the chain of command.

Future work will fall into two main categories. First, output feedback will be studied. Next, links between the hierarchy resulting from the optimization process and control structures from biology and social sciences will be studied in greater depth.

References

[1] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems—part i," *IEEE Transactions on Automatic Control*, vol. 17, no. 1, 1972.

[2] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal of Control*, vol. 6, no. 1, 1968.

[3] N. R. Sandell and M. Athans, "Solution of some nonclassical lqg stochastic decision problems," *IEEE Transactions on Automatic Control*, vol. 19, no. 2, pp. 108–116, 1974.

[4] B.-Z. Kurtaran and R. Sivan, "Linear-quadratic-gaussian control with one-step-delay sharing pattern," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 571–574, 1974.

[5] T. Yoshikawa, "Dynamic programming approach to decentralized stochastic control problems," *IEEE Transactions on Automatic Control*, vol. 20, no. 6, pp. 796–797, 1975.

[6] X. Qi, M. V. Salapaka, P. G. Voulgaris, and M. Khammash, "Structured optimal and robust control with multiple criteria: A convex solution," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1623–1640, 2004.

[7] M. Rotkowitz and S. Lall, "A characterization of covex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 1984–1996, 2006.

[8] A. Rantzer, "Linear quadratic team theory revisited," in *American Control Conference*, 2006.

[9] ——, "A separation principle for distributed control," in *IEEE Conference on Decision and Control*, 2006.

[10] J. Swigart and S. Lall, "An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator," in *American Control Conference*, 2010, pp. 6385–6390.

[11] ——, "An explicit dynamic programming solution for a decentralized two-player optimal linear-quadratic regulator," in *Symposium on the Mathematical Theory of Networks*, 2010, pp. 1443 – 1447.

[12] P. Shah and P. Parrilo, "$\mathcal{H}_2$-optimal decentralized control over posets: A state space solution for state-feedback," in *IEEE Conference on Decision and Control*, 2010.

[13] A. Lamperski, "Hierarchies, spikes, and hybrid systems: Physiologically inspired control problems," Ph.D. dissertation, California Institute of Technology, 2011.