

Optimal Decentralized Control with Asymmetric One-Step Delayed Information Sharing

Naumaan Nayyar, Dileep Kalathil and Rahul Jain

Abstract—We consider optimal control of decentralized LQG problems for plants having nested subsystems controlled by two players having asymmetric information sharing patterns between them. In the main scenario, players are assumed to have a unidirectional error-free, unlimited rate communication channel with a unit delay in one direction and no communication in the other. A secondary model presented for completeness considers a channel with no delay in one direction and a unit delay in the other. Delayed information sharing patterns in general do not admit linear optimal control laws and are thus difficult to control optimally. However, in these scenarios, we show that the problems have a partially nested information structure, and thus linear optimal control laws exist. Summary statistics are identified and analytical solutions to the optimal control laws are derived. State and output feedback cases are solved for both scenarios.

I. INTRODUCTION

Recently, many problems of decentralized ‘non-classical’ control have been looked at in the context of practical systems [1], [2], [3]. Examples include cyberphysical systems, formation flight, and other networked control systems wherein multiple agents try to achieve a common objective in a decentralized manner. Such situations arise, for example, due to controllers not having access to the same information. One possible reason is network delays, and a consequent time-lag to communicate observations to other controllers.

These problems were first formulated by Marschak in the 1950s [4] as team decision problems, and further studied by Radner [5], though in such problems communication between the controllers was usually ignored. In a celebrated paper [6], Witsenhausen showed that even for seemingly simple systems with just two players, non-linear controllers could outperform any linear controller. Witsenhausen also consolidated and conjectured results on separation of estimation and control in decentralized

problems in [7]. However, the structure of decentralized optimal controllers for LQG systems with time-delays has been hard to identify. Indeed, in [8] it was proved that the centralized-like separation principle that was conjectured for delayed systems does not hold for a system having a delay of more than one timestep. The more general delayed sharing pattern was only recently solved by Nayyar, et al. in [9]. It is an open problem to actually compute the optimal decentralized control law for the n -player setting even using such a structure result.

However, not all results for such problems have been negative. Building on results of Radner on team decision theory, Ho and Chu [10] showed that for a unit-delay information sharing pattern, the optimal controller is linear. This was used by Kurtaran and Sivan [11] and others [12], [13] to derive optimal controllers for the finite-horizon case. It is also known that for a class of problems having partially nested structures, person-by-person optimality implies global optimality [14]. These approaches do not extend to multi-unit delayed sharing patterns. This is because the former are examples of systems with *partially nested information structure*, for which linear optimal laws are known to exist [10].

A criterion for determining convexity of optimal control problems, *funnel causality*, was presented in [15]. Recently, another characterization, *quadratic invariance*, was discovered under which optimal decentralized controls laws can be solved by convex programming [1]. This has led to a resurgence of interest in decentralized optimal control problems, which since the 1970s had been assumed to be intractable. Subsequently, in a series of papers, Lall, Lessard and others have computed the optimal control laws for a suite of decentralized nested LQG systems, involving no communication delays [16], [17], [18] and others. More general networked structures that dealt with state-feedback delayed information sharing in networked control with sparsity were also looked at in [19], [20], wherein recursive solutions for optimal control laws were derived. Output feedback without delays in a network was considered in [21]. Time varying delays in 2-player state-feedback LQR problems have known explicit solutions [22].

The authors are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089. Emails: (nnayyar,manisser,rahul.jain)@usc.edu.

This research was supported by AFOSR grant FA9550-10-1-0307, the NSF CAREER Award CNS-0954116 and the ONR Young Investigator Award N000141210766. Parts of this paper have been presented in ACC 2014, without all proof details and only the $(1, \infty)$ case. Theorems 4 and 6 are new, which now fully solve the problem.

For a subclass of quadratic invariant problems known as *poset-causal* problems, Parrilo and Shah [23] showed that the computation of the optimal law becomes easier by lending itself to decomposition into subproblems. Solutions to certain cases of the state-feedback and output-feedback delayed information sharing problem have also been presented by Lamperski et al. [24] where a networked graph structure of the controllers (strongly connected) is considered, with constraints on system dynamics. In this work, we restrict our attention to two-player systems. A summary of results pertaining to these systems is given in Table I.

(d_{12}, d_{21})	Literature	Comments
(0, 0)	Classical	no plant restrictions
(1, 1)	[11],[12],[13]	no plant restrictions
(0 or 1, 0 or 1)	[24], [25]	b.d B matrix, inf. horizon
(0, ∞)	[16], [17], [18]	l.b.t matrix
(1, ∞)	[19]	l.b.t matrix, state f/b, u.c
(∞, ∞)	[26]	b.d. dynamics, state f/b
(1, 0)	here	no plant restrictions
(1, ∞)	here	l.b.t matrix

TABLE I
SUMMARY OF RESULTS FOR SOME INFORMATION SHARING PATTERNS WITH TWO-PLAYERS.

In Table I, d_{12} is the delay in information transmission from player 1 to player 2, and vice versa. 'u.c', 'b.d' and 'l.b.t' refer to uncoupled noise, block diagonal and lower block triangular matrices respectively. Unlike in previous work [19], noise vectors are not restricted to be uncorrelated in our work.

In this paper, we consider two asymmetric scenarios, (1, 0) and (1, ∞) information sharing patterns between two players in an LQG system. Both scenarios have a partially nested information structure and also satisfy quadratic invariance. They are not poset-causal and do not lend themselves to easy decompositions. We derive optimal control laws in both cases. This results in Riccati-type iterations for computing the gain matrix of the optimal control law which is linear. This paper extends our work presented in ACC 2014 [27] to derive explicit optimal controllers for the output feedback case.

II. THE (1, ∞) INFORMATION SHARING PATTERN

State-feedback case

We consider a coupled two-player discrete linear time-invariant system with nested structure. Player 1's actions affect block 1 of the plant whereas player 2's actions affect both blocks 1 and 2. The system dynamics are,

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}, \quad (1)$$

for $t \in \{0, 1, \dots, N-1\}$.

We will use the notation $x(t+1) = Ax(t) + Bu(t) + v(t)$ where definition is clear from context. Initial state $x(0)$ is assumed to be zero-mean Gaussian, independent of the system noise $v(t)$, and $v(t)$ is zero-mean Gaussian with covariance V , which is independent across time t .

In the formulation above, restricting plant dynamics to be lower block triangular ensures that optimal control laws exist that are linear in the observations. Otherwise, the system fails the only known sufficiency test for existence of linear optimal laws for the two-player class [10].

The two players have a team/common objective to find the decentralized control law $u(t)$ that minimizes,

$$\mathbb{E} \left[\sum_{t=0}^{N-1} (x(t)'Qx(t) + u(t)'Ru(t)) + x(N)'Sx(N) \right], \quad (2)$$

where Q , R and S are positive definite matrices.

At each instant, the control action $u_i(t)$ taken by each player can only depend on the *information structure*, the information available to them denoted by $\mathcal{F}(t) = (\mathcal{F}_1(t), \mathcal{F}_2(t))$. The information structure for the problem is,

$$\begin{aligned} \mathcal{F}_1(t) &= \{x_1(0:t), u_1(0:t-1)\}, \\ \mathcal{F}_2(t) &= \{x_1(0:t-1), u_1(0:t-1); \\ &\quad x_2(0:t), u_2(0:t-1)\}. \end{aligned} \quad (3)$$

The set of all control laws is characterized by $u_i(t) = f(\mathcal{F}_i(t))$ for some time-varying function f . In the next subsection, we show that the optimal law is linear.

1) *Linearity of the optimal control law:* It was shown by Varaiya and Walrand [8] that the optimal decentralized control law with general delayed information sharing between players may not be linear. However, Ho and Chu [10] had established earlier that to show the existence of a linear optimal control law, it is sufficient to prove that the LQG problem has a *partially nested information structure* [10, Theorem 2].

Formally, if player i 's actions at time t_1 affect player j 's information at time t_2 , then a partially nested information structure is defined to have $\mathcal{F}_i(t_1) \subseteq \mathcal{F}_j(t_2)$, where \mathcal{F} denotes the information available to the players when making their decision. Intuitively, this can be described as an information pattern that allows communication of all information used by a player to make a decision to all other players whose system dynamics are affected by that decision. We now show that the (1, ∞) pattern has a partially nested structure.

Proposition 1. *The problem (1)-(2) with information structure (3) has a linear optimal control law.*

Proof. Observe that the information structure defined in (3) can be simplified as $\mathcal{F}_1(t) = \{x_1(0:t)\}$ and $\mathcal{F}_2(t) = \{x_1(0:t-1), x_2(0:t)\}$ as the observations can be used to determine inputs. Because of the nested

system structure, viz. $A_{12} = B_{12} = 0$, the propagation of inputs in the plant dynamics is as shown in Fig 1. Additionally, at any time instant t , $\mathcal{F}_i(t) \subset \mathcal{F}_i(t + \tau)$, $\tau \geq 0$, $i = 1, 2$ and $\mathcal{F}_1(t) \subset \mathcal{F}_2(t + \tau)$, $\tau \geq 1$. This implies that the information communication diagram is also the same as in Figure 1.

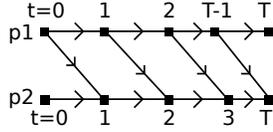


Fig. 1. Information communication in $(1, \infty)$ -delayed sharing pattern in two-player D-LTI systems with nested structures

Thus, the system described with its information structure has a partially nested structure [10, Theorem 2], and consequently, a linear optimal control law exists. \square

Denote by $\mathcal{H}_i(t) = (x_i(0:t-1), u_i(0:t-1))$, the history of observations for block i . Using linearity of the optimal law, it follows that $u(t) = [u_1(t), u_2(t)]' =$

$$\begin{bmatrix} F_{11}^*(t)x_1(t) + L_1(\mathcal{H}_1(t)) \\ F_{22}^*(t)x_2(t) + L_2(\mathcal{H}_1(t), \mathcal{H}_2(t)) \end{bmatrix}, \quad (4)$$

where $L_i(\cdot)$ are linear, possibly time-varying, functions. However, the arguments of L_i , namely the output observations, increase with time. So, obtaining a closed form expression for the optimal control law is intractable in the current form, necessitating use of summary statistics.

2) *Summary Statistic to Compute Optimal Control Law:* We now show that the estimate of the current state given the history of observations of the players is a summary statistic that characterizes the optimal control law. The notation used is summarized in Table II.

Note that $\hat{x}(t)$ and $\hat{\hat{x}}(t)$ are estimates of the current states by Players 1 and 2 respectively.

Lemma 1. *For the nested system (1), the estimator dynamics for $\hat{x}(t)$ and $\hat{\hat{x}}(t)$ are given by*

$$\hat{x}(t+1) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix} + \mathcal{R}(t)\phi(t), \quad (5)$$

$$\hat{\hat{x}}(t+1) = A\hat{\hat{x}}(t) + B \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \bar{\mathcal{R}}(t)\bar{\phi}(t), \quad (6)$$

$\mathcal{H}_i(t)$	$\{x_i(0:t-1), u_i(0:t-1)\}$
$\mathcal{F}_1(t)$	$\{x_1(t), \mathcal{H}_1(t)\}$
$\mathcal{F}_2(t)$	$\{x_2(t), \mathcal{H}_1(t), \mathcal{H}_2(t)\}$
$\hat{x}(t)$	$\mathbb{E}[x(t) \mathcal{H}_1(t)]$
$\hat{\hat{x}}(t)$	$\mathbb{E}[x(t) \mathcal{H}_1(t), \mathcal{H}_2(t)]$
$\phi(t)$	$x_1(t) - \hat{x}_1(t)$
$\bar{\phi}(t)$	$x(t) - \hat{\hat{x}}(t)$

TABLE II

A GLOSSARY OF NOTATIONS USED IN THE FOLLOWING SECTIONS

$\bar{T}(t)$	$\mathbb{E}[(x(t) - \hat{\hat{x}}(t))(x(t) - \hat{\hat{x}}(t))]$
$T(t)$	$\mathbb{E}[(x(t) - \hat{x}(t))(x_1(t) - \hat{x}_1(t))]$
$R_x(t)$	$\bar{T}(t)$
$\bar{R}_{xx}(t)$	$A\bar{T}(t)$
$R_{x_1}(t)$	$T_1(t)$
$R_{xx_1}(t)$	$AT(t)$
$\bar{\phi}(t), \phi(t)$	$x(t) - \hat{\hat{x}}(t), x_1(t) - \hat{x}_1(t)$
$\bar{\mathcal{R}}(t)$	$\bar{R}_{xx}(t)\bar{R}_x^{-1}(t)$
$\mathcal{R}(t)$	$R_{xx_1}(t)R_{x_1}^{-1}(t)$
$\hat{u}_2(t)$	$\mathbb{E}[u_2(t) \mathcal{H}_1(t)]$

TABLE III

A LIST OF NOTATIONS USED IN LEMMA 1

with $\hat{x}(0) = \hat{\hat{x}}(0) = \mathbb{E}[x(0)]$. $\bar{\mathcal{R}}(t) = A$. $\bar{\phi}(t)$ and $\bar{\phi}(t)$ are zero-mean, uncorrelated Gaussian random vectors and $\mathbb{E}[\phi(t)\phi(t)'] = T_1(t)$ and $\mathbb{E}[\bar{\phi}(t)\bar{\phi}(t)'] = \bar{T}(t)$. At each time step, $\phi(t)$ is uncorrelated with $\mathcal{H}_1(t)$, and so $\mathbb{E}[\phi(t)x_1(\tau)'] = \mathbf{0}$ for $\tau < t$. $\bar{\phi}(t)$ is uncorrelated with $\mathcal{H}_1(t)$, $\mathcal{H}_2(t)$, and so $\mathbb{E}[\bar{\phi}(t)x(\tau)'] = \mathbf{0}$ for $\tau < t$.

Proof. The proof is in the Appendix. \square

Before showing that $\hat{x}_1(t)$ and $\hat{\hat{x}}_2(t)$ are summary statistics for the history of state observations, we first obtain the optimal decentralized control law for the LQG problem (1)-(2) with a nested information structure having no communication delay - i.e., $(0, \infty)$. The information structure here is described by,

$$z_1(t) = \{x_1(0:t), u_1(0:t-1)\}, \quad (7)$$

$$z_2(t) = \{x_1(0:t), u_1(0:t-1); x_2(0:t), u_2(0:t-1)\}.$$

Proposition 2 considers the $(0, \infty)$ state-feedback problem, already solved for output-feedback in [18]:

Proposition 2. *The optimal control law for the LQG problem (1)-(2) with information structure (7) is,*

$$\begin{bmatrix} u_1^*(t) \\ u_2^*(t) \end{bmatrix} = \begin{bmatrix} K_{11}(t) & K_{12}(t) & \mathbf{0} \\ K_{21}(t) & K_{22}(t) & J(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \tilde{x}_2(t) \\ x_2(t) - \tilde{x}_2(t) \end{bmatrix},$$

$$\begin{aligned} \tilde{x}_2(t+1) &= A_{21}x_1(t) + A_{22}\tilde{x}_2(t) + B_{21}u_1(t) + B_{22}\tilde{u}_2(t), \\ \tilde{x}_2(0) &= 0, \end{aligned} \quad (8)$$

where $\tilde{u}_2(t) = [K_{21}(t) \ K_{22}(t)] \begin{bmatrix} x_1(t) \\ \tilde{x}_2(t) \end{bmatrix}$. $K(t)$ is obtained from the backward recursion,

$$\begin{aligned} P(T) &= S, \\ P(t) &= A'P(t+1)A + A'P(t+1)BK(t) + Q, \\ K(t) &= -(B'P(t+1)B + R)^{-1}B'P(t+1)A, \end{aligned}$$

and $J(t)$ is obtained from the backward recursion,

$$\begin{aligned}\tilde{P}(T) &= S_{22}, \\ \tilde{P}(t) &= A'_{22}\tilde{P}(t+1)A_{22} + A'_{22}\tilde{P}(t+1)B_{22}J(t) + Q_{22}, \\ J(t) &= -(B'_{22}\tilde{P}(t+1)B_{22} + R_{22})^{-1}B'_{22}\tilde{P}(t+1)A_{22}.\end{aligned}$$

Proof. The proof follows by use of Theorems 10 and 12 of [18], and replacing the information structure with the state-feedback structure given in (7). The conditional expectations reduce to $\mathbf{z}(t) := \mathbb{E}[x(t)|z_2(t)] = x(t)$ and $\hat{\mathbf{z}}(t) := \mathbb{E}[x(t)|z_1(t)] = \begin{bmatrix} x_1(t) \\ \tilde{x}_2(t) \end{bmatrix}$, where $\tilde{x}_2(t) := \mathbb{E}[x_2(t)|z_1(t)]$. The optimal control law is,

$$u(t) = K(t)\hat{\mathbf{z}}(t) + \hat{K}(t)(\mathbf{z}(t) - \hat{\mathbf{z}}(t)),$$

$$\text{where } \hat{K}(t) = \begin{bmatrix} 0 & 0 \\ \hat{K}_{21} & J(t) \end{bmatrix}$$

Substituting $\mathbf{z}(t)$ and $\hat{\mathbf{z}}(t)$, the control law reduces to,

$$\begin{aligned}u(t) &= K(t) \begin{bmatrix} x_1(t) \\ \tilde{x}_2(t) \end{bmatrix} + \hat{K}(t) \begin{bmatrix} 0 \\ x_2(t) - \tilde{x}_2(t) \end{bmatrix} \\ &= \begin{bmatrix} K_{11}(t) & K_{12}(t) & \mathbf{0} \\ K_{21}(t) & K_{22}(t) & J(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \tilde{x}_2(t) \\ x_2(t) - \tilde{x}_2(t) \end{bmatrix}.\end{aligned}$$

Consequently, we also have, $\tilde{u}_2(t) := \mathbb{E}[u_2(t)|z_1(t)] = \begin{bmatrix} K_{21}(t) & K_{22}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$, which when substituted in (1), gives (8). \square

We now establish the sufficiency of $\hat{x}_1(t)$ and $\hat{x}_2(t)$.

Theorem 1. *For the nested system (1) with the information structure (3) and objective function (2), the optimal control law is given by*

$$u(t) = F^*(t) \begin{bmatrix} x_1(t) - \hat{x}_1(t) \\ x_2(t) - \hat{x}_2(t) \end{bmatrix} + \tilde{K}(t), \quad (9)$$

where $F^*(t) = \begin{bmatrix} F_{11}^*(t) & 0 \\ 0 & F_{22}^*(t) \end{bmatrix}$ and $\tilde{K}(t)$ is given by

$$\begin{bmatrix} K_{11}(t) & K_{12}(t) & \mathbf{0} \\ K_{21}(t) & K_{22}(t) & J(t) \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_2(t) - \hat{x}_2(t) \end{bmatrix},$$

where $K(t)$ and $J(t)$ are as in Proposition 2.

The proof is in the Appendix. Note that the classical separation principle does not hold as the statistics $\hat{x}(t)$ and $\hat{\hat{x}}(t)$ cannot be used to compute the optimal law.

3) *Deriving the Optimal Gain Matrix:* To characterize the optimal gain matrix, $F^*(t)$, the stochastic optimization problem is converted into a deterministic matrix optimization, and solved analytically.

Theorem 2. *The matrix $F^*(t)$ is characterized by the linear equations,*

$$\begin{aligned}& \begin{bmatrix} R + \tilde{B}'M(t+1)\tilde{B} \end{bmatrix}_{11} F_{11}^*(t)\bar{T}_{11}(t) \\ & + \begin{bmatrix} R + \tilde{B}'M(t+1)\tilde{B} \end{bmatrix}_{12} F_{22}^*(t)\bar{T}_{21}(t) \\ & = - \begin{bmatrix} \tilde{B}'M(t+1)V(t)' \end{bmatrix}_{11} \quad (10)\end{aligned}$$

$$\begin{aligned}& \begin{bmatrix} R + \tilde{B}'M(t+1)\tilde{B} \end{bmatrix}_{21} F_{11}^*(t)\bar{T}_{12}(t) \\ & + \begin{bmatrix} R + \tilde{B}'M(t+1)\tilde{B} \end{bmatrix}_{22} F_{22}^*(t)\bar{T}_{22}(t) \\ & = - \begin{bmatrix} \tilde{B}'M(t+1)V(t)' \end{bmatrix}_{22}, \quad (11)\end{aligned}$$

where $[A]_{ij} = E'_i A E_j$, E_i and E_j being suitable partitions of the identity matrix $I = [E_i \ E_j]$.

$$\begin{aligned}\tilde{A} &= \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ 0 & 0 & A_{22} \end{bmatrix}, \quad G(t) = \\ & \begin{bmatrix} BH(t) \\ [0 \ 0 \ B_{22}J(t)] \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \\ 0 & B_{22} \end{bmatrix}, \quad H(t) = \\ & \begin{bmatrix} K_{11}(t) & K_{12}(t) & -K_{12}(t) \\ K_{21}(t) & K_{22}(t) & J(t) - K_{22}(t) \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}, \\ \tilde{S} &= \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}, \quad V(t) = \mathbf{E}[\bar{\phi}(t)n_1(t)'], \quad \text{and} \\ n_1(t) &:= \begin{bmatrix} (\mathcal{R}(t)\bar{\phi}_1(t))_1 \\ B_{21}F_{11}(t)\bar{\phi}_1(t) + (\mathcal{R}(t)\bar{\phi}(t))_2 \\ + (\mathcal{R}(t)\bar{\phi}(t))_2 - (\mathcal{R}(t)\bar{\phi}_1(t))_2 \end{bmatrix}.\end{aligned}$$

The proof is in the Appendix.

Output-feedback case

The approach extends to the output-feedback case where players' observations are corrupted by noise. The system dynamics are described by the same lower block triangular matrices as,

$$\begin{aligned}\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \\ &+ \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}, \\ \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} &= \begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}. \quad (12)\end{aligned}$$

Here, $v(t) := \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$ and $w(t) := \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$ are zero-mean, Gaussian random vectors with covariance matrices V and W respectively that are independent across time and independent of initial system state $x(0)$ and of each other. $x(0)$ is zero-mean, Gaussian with known mean and covariance.

The information structure for the problem is,

$$\begin{aligned}\mathcal{F}_1(t) &= \{y_1(0:t), u_1(0:t-1)\}, \\ \mathcal{F}_2(t) &= \{y_1(0:t-1), u_1(0:t-1); \\ & \quad y_2(0:t), u_2(0:t-1)\}.\end{aligned} \quad (13)$$

Redefining the history of observations as, $\mathcal{H}_i(t) = (y_i(t-1), \dots, y_i(0), u_i(t-1), \dots, u_i(0))$ for block i , $i = 1, 2$, the information available to Player 1 at time t can also be expressed as $\mathcal{F}_1(t) = \{y_1(t), \mathcal{H}_1(t)\}$, and that available to Player 2 as $\mathcal{F}_2(t) = \{y_2(t), \mathcal{H}_1(t), \mathcal{H}_2(t)\}$.

Given this information structure, the players' objective is to find the control law $u_i^*(t)$ as a function of $\mathcal{F}_i(t)$ ($i = 1, 2$) that minimizes the finite-time quadratic cost criterion (2). It can be seen from Proposition 1 that the optimal control law for the formulated output-feedback problem is linear.

Further, defining estimators $\hat{x}(t) := \mathbb{E}[x(t)|\mathcal{H}_1(t)]$ and $\hat{\hat{x}}(t) := \mathbb{E}[x(t)|\mathcal{H}_2(t)]$, the following result can be reconstructed from the state-feedback version.

Lemma 2. *For the nested system (12), the estimator dynamics for the redefined $\hat{x}(t)$ and $\hat{\hat{x}}(t)$ are as follows.*

$$\hat{x}(t+1) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix} + \mathcal{R}(t)\phi(t), \quad (14)$$

$$\hat{\hat{x}}(t+1) = A\hat{\hat{x}}(t) + B \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \bar{\mathcal{R}}(t)\bar{\phi}(t) \quad (15)$$

and $\hat{x}(0) = \hat{\hat{x}}(0) = \mathbb{E}[x(0)]$, where $\hat{u}_2(t) = \mathbb{E}[u_2(t)|\mathcal{H}_1(t)]$. $\mathcal{R}(t) = R_{xy_1}(t)R_{y_1}^{-1}(t)$, $\bar{\mathcal{R}}(t) = \bar{R}_{xy}\bar{R}_y^{-1}$, and $\phi(t) = y_1(t) - C_{11}\hat{x}_1(t)$, $\bar{\phi}(t) = y(t) - C\hat{\hat{x}}(t)$. $R_{xy_1} = AT(t)C'_{11}$, $R_{y_1} = C_{11}T_1(t)C'_{11} + W_{11}$, $\bar{R}_{xy} = \bar{A}\bar{T}(t)C'$, $\bar{R}_y = \bar{C}\bar{T}(t)C' + W$. $T(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x_1(t) - \hat{x}_1(t))]$, $\bar{T}(t) = \mathbb{E}[(x(t) - \hat{\hat{x}}(t))(x(t) - \hat{\hat{x}}(t))']$.

Also, $\phi(t)$ and $\bar{\phi}(t)$ are zero-mean, uncorrelated Gaussian random vectors. $\mathbb{E}[\phi(t)\phi(t)'] = C_{11}T_1(t)C'_{11} + W_{11}$ and $\mathbb{E}[\bar{\phi}(t)\bar{\phi}(t)'] = \bar{C}\bar{T}(t)C' + W$. Additionally, at each time step, $\phi(t)$ is uncorrelated with $\mathcal{H}_1(t)$, and so $\mathbb{E}[\phi(t)y_1(\tau)'] = \mathbf{0}$ for $\tau < t$. $\bar{\phi}(t)$ is uncorrelated with $\mathcal{H}_1(t)$, $\mathcal{H}_2(t)$, and so $\mathbb{E}[\bar{\phi}(t)y(\tau)'] = \mathbf{0}$ for $\tau < t$.

Proof. Proof is similar to Lemma 1 and is omitted. \square

Theorem 3. *For the nested system (12) with the information structure (13) and objective function (2), the optimal decentralized control law is given by,*

$$u(t) = F^*(t) \begin{bmatrix} y_1(t) - \hat{x}_1(t) \\ y_2(t) - C_{21}\hat{x}_1(t) - C_{22}\hat{x}_2(t) \end{bmatrix} + \tilde{K}(t), \quad (16)$$

where $F^*(t) = \begin{bmatrix} F_{11}^*(t) & 0 \\ 0 & F_{22}^*(t) \end{bmatrix}$ is the optimal gain matrix, and,

$$\tilde{K}(t) = \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ [\hat{K}_{21}(t) \quad J(t)] [\hat{\hat{x}}(t) - \hat{x}(t)] \end{bmatrix},$$

where $K(t)$ and $J(t)$ are as in Proposition 2, \hat{K}_{21} is the optimal control gain for the second player's input given the optimal centralized controller only based on common information. [18, Theorem 12].

The proof is in the appendix.

Remark 1. *Observe that the classical separation principle does not hold in the $(1, \infty)$ -delayed sharing pattern [28], i.e., $u_2(t)$ cannot be computed using just $\mathbb{E}[x(t)|\mathcal{F}_2(t)]$.*

Theorem 4. *The optimal gain matrix $F^*(t)$ is given by the recursion,*

$$\begin{aligned} F^*(t) &= -\left(R + \tilde{B}'M(t+1)\tilde{B}\right)^{-1} \times \\ &\quad \left(\Psi^*(t) + \tilde{B}'M(t+1)V(t)'\right) \times \\ &\quad \left(C\bar{T}(t)C' + W\right)^{-1}, \\ M^*(t) &= \tilde{Q} + H(t)'RH(t) \\ &\quad + (\tilde{A} + G(t))'M^*(t+1)(\tilde{A} + G(t)), \\ M^*(N) &= \tilde{S}, \end{aligned} \quad (17)$$

where, $\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$, $G(t) = \begin{bmatrix} BK(t) & 0 \\ 0 & B \begin{bmatrix} \hat{K}_{21}(t) & J(t) \end{bmatrix} \end{bmatrix}$, $\tilde{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}$, $H(t) = \begin{bmatrix} K & \begin{bmatrix} 0 \\ \hat{K}_{21}(t) & J(t) \end{bmatrix} \end{bmatrix}$, $\tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}$, $\tilde{S} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$, $V(t) = \mathbb{E}[\phi_o(t)n_1(t)']$, and $n_1(t) := \begin{bmatrix} \mathcal{R}(t)\phi(t) \\ \bar{\mathcal{R}}(t)\bar{\phi}(t) - \mathcal{R}(t)\phi(t) \end{bmatrix}$.

Remark 2. *As in Theorem 2, it is not difficult to obtain complete linear expressions for $F^*(t)$ from it using the sparsity constraints on $\Psi^*(t)$ and $F^*(t)$.*

III. THE $(1, 0)$ INFORMATION SHARING PATTERN

Another application of the above approach, which is also an extension to [11], is the $(1, 0)$ -pattern. A two-player discrete linear time-invariant system with output feedback is considered. The system dynamics are,

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + v(t), \\ y(t) &= Cx(t) + w(t) \end{aligned} \quad (18)$$

The two players have a (team/common) objective to find the decentralized control law $(u_1^*(\cdot), u_2^*(\cdot))$ that minimizes a finite-time quadratic cost criterion

$$\mathbb{E}\left[\sum_{t=0}^{N-1} (x(t)'Qx(t) + u(t)'Ru(t)) + x(N)'Sx(N)\right], \quad (19)$$

We consider the following information structure:

$$\begin{aligned}
\mathcal{F}_1(t) &= \{y_1(0:t), u_1(0:t-1), y_2(0:t), u_2(0:t-1)\}, & F^*(t) &= -(R + B'M(t+1)B)^{-1} \times \\
\mathcal{F}_2(t) &= \{y_1(0:t-1), u_1(0:t-1), & & \left(\Psi^*(t)(CT(t)C' + W(t))^{-1} + B'M(t+1)\mathcal{R}(t) \right), \\
& y_2(0:t), u_2(0:t-1)\}. & & M^*(t) = Q + H(t)'RH(t) \\
\end{aligned} \tag{20}$$

where $y(0:t)$ denotes the vector $(y(0), \dots, y(t))$.

We refer to this as the $(1, 0)$ information sharing pattern. Finding the optimal decentralized controller under this information pattern is similar to that of the $(1, 1)$ pattern [11], with the difference being the asymmetry in observation history. We outline the application of the same approach we used in $(1, \infty)$ to solve this case.

4) Linearity of the optimal control law:

Proposition 3. *The LQG problem (18)-(19) with the $(1, 0)$ -delayed sharing information pattern (20) has a partially nested structure, and a linear optimal control law exists.*

See, for instance, [10, Theorem 2] for a proof.

Denote by $\mathcal{H}(t) = (y(0:t-1), u(0:t-1))$, the common information at time t [7]. Using the linearity of the optimal control law, it follows that the optimal control law can be written as $u(t) = [u_1(t), u_2(t)]' =$

$$\begin{bmatrix} F_{11}^*(t) & F_{12}^*(t) \\ 0 & F_{22}^*(t) \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} L_1(\mathcal{H}(t)) \\ L_2(\mathcal{H}(t)) \end{bmatrix}, \tag{21}$$

where F^* is the optimal gain matrix and $L_i(\cdot)$ are linear, possibly time-varying, functions.

5) *Derivation of the Optimal Control Law:* Define the estimator $\hat{x}(t) := \mathbb{E}[x(t)|\mathcal{H}(t)]$.

Lemma 3. [29, Section 6.5] *For the nested system (18), the estimator dynamics for $\hat{x}(t)$ are as follows.*

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + \mathcal{K}(t)\phi(t), \tag{22}$$

and $\hat{x}(0) = \mathbb{E}[x(0)]$, where, $\mathcal{K}(t) = K_{xy}(t)K_y(t)^{-1}$ and $\phi(t) = y(t) - C\hat{x}(t)$. $K_{xy}(t) = AT(t)C'$, $K_y(t) = CT(t)C' + W$. $T(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))']$. $\phi(t)$ is zero-mean, uncorrelated Gaussian and $\mathbb{E}[\phi(t)\phi(t)'] = CT(t)C' + W$. $\mathbb{E}[\phi(t)y(\tau)'] = 0$, $\tau < t$.

Theorem 5. *For the system (18) with the information structure (20) and objective function (19),*

$$\begin{bmatrix} u_1^*(t) \\ u_2^*(t) \end{bmatrix} = F^*(t) \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + G^*(t)\hat{x}(t), \tag{23}$$

where $F^*(t) = \begin{bmatrix} F_{11}^*(t) & F_{12}^*(t) \\ \mathbf{0} & F_{22}^*(t) \end{bmatrix}$ is the optimal gain and $G^*(t) = -(F^*(t)C + \bar{H}^*(t))$ where $H^*(t)$ is the gain for classical LQR with dynamics described by (18).

Proposition 4. *The optimal gain matrix $F^*(t)$ is given by the recursion,*

$$\begin{aligned}
F^*(t) &= -(R + B'M(t+1)B)^{-1} \times \\
& \left(\Psi^*(t)(CT(t)C' + W(t))^{-1} + B'M(t+1)\mathcal{R}(t) \right), \\
M^*(t) &= Q + H(t)'RH(t) \\
& + (A - BH(t))'M^*(t+1)(A - BH(t)), \\
M^*(N) &= \tilde{S}. \tag{24}
\end{aligned}$$

From Theorems 4 and 5, we see that the form of the optimal controller for the $(1, \infty)$ case is a mixture of those for $(1, 0)$ and $(0, \infty)$ cases. Specifically, the first component that includes the innovation sequence resembles the $(1, 0)$ setting, and the second component resembles the $(0, \infty)$ solution. In fact, it can be seen from Theorem 4 that removing the innovation sequence term (as is the case when there are no delays) results in the same solution as the $(0, \infty)$ case.

IV. FUTURE WORK

We believe that the approach of combining linearity and summary statistics in the manner of this paper can be used to compute optimal control laws for more general asymmetric information sharing patterns. Specifically, the n -player network extension to the output feedback scenario would be a novel extension to [20].

V. ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of the editor and our reviewers to this paper, in particular to the reviewer who pointed out an approach to use sparsity patterns to solve the gain matrix equations.

APPENDIX A

PROOFS OF THEOREMS AND LEMMAS

Proof of lemma 1. Let us first derive the dynamics of block 1. Denote $\phi(t) := x_1(t) - \hat{x}_1(t)$. Then, $\phi(t)$ and $\mathcal{H}_1(t)$ are independent by the projection theorem for Gaussian random variables [29, Section 6.5] and,

$$\begin{aligned}
\hat{x}(t+1) &= \mathbb{E}[x(t+1)|\mathcal{H}_1(t+1)] \\
&= \mathbb{E}[x(t+1)|\mathcal{H}_1(t), x_1(t), u_1(t)] \\
&= \mathbb{E}[x(t+1)|\mathcal{H}_1(t), u_1(t)] + \mathbb{E}[x(t+1)|\phi(t)] \\
&\quad - \mathbb{E}[x(t+1)],
\end{aligned}$$

where the last equality follows from the independence of $\phi(t)$ and $\mathcal{H}_1(t)$ [29, Section 6.5]. The first term on the RHS above, $\mathbb{E}[x(t+1)|\mathcal{H}_1(t), u_1(t)]$

$$\begin{aligned}
&= \mathbb{E}[Ax(t) + Bu(t)|\mathcal{H}_1(t), u_1(t)] \\
&= A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix},
\end{aligned}$$

where $\hat{u}_2(t) = \mathbb{E}[u_2(t)|\mathcal{H}_1(t)]$.

The second and third terms can be related through the conditional estimation of Gaussian random vectors as,

$$\mathbb{E}[x(t+1)|\phi(t)] - \mathbb{E}[x(t+1)] = R_{xx_1}R_{x_1}^{-1}\phi(t),$$

where defining $T(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x_1(t) - \hat{x}_1(t))]$ and partitioning $T(t)$ as $T(t) = [T_1(t)|T_2(t)]$, we have $R_{xx_1} = AT(t)$, $R_{x_1} = T_1(t)$ [29, Section 6.5].

Putting these three terms together, we have

$$\hat{x}(t+1) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix} + \mathcal{R}(t)\phi(t).$$

Similarly, by defining $\bar{\phi}(t) = x(t) - \hat{x}(t)$ and proceeding as above, we have

$$\hat{\bar{x}}(t+1) = A\hat{\bar{x}}(t) + B \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \bar{\mathcal{R}}(t)\bar{\phi}(t).$$

To derive the properties of $\phi(t)$ and $\bar{\phi}(t)$, note that the mean and variance follow immediately by observing that $\mathbb{E}[\hat{x}_1(t)] = \mathbb{E}[x_1(t)]$ and $\mathbb{E}[\hat{\bar{x}}(t)] = \mathbb{E}[x(t)]$.

The projection theorem for Gaussian RVs again implies that $\phi(t)$ and $x_1(\tau)$ are uncorrelated for $\tau < t$, and that $\bar{\phi}(t)$ is uncorrelated with $x(\tau)$, $\tau < t$. \square

Proof of Theorem 1. We begin by outlining our proof technique. The approach, while superficially similar to [11], is far more challenging due to the asymmetry in observation history. The optimal control law for the $(1, \infty)$ state-feedback pattern, already shown to be linear through the problem being partially nested, is proven to be the same as the optimal control law for the dynamical system (1) with the objective function $\arg \min_{\bar{u}(t)} \mathbb{E}[\sum_{t=0}^{N-1} (\bar{x}(t)'Q\bar{x}(t) + \bar{u}(t)'R\bar{u}(t)) + \bar{x}(N)'S\bar{x}(N)]$, where $\bar{x}(t) = [\hat{x}_1(t) \ \hat{\bar{x}}_2(t)]$ and $\bar{u}(t)$ is an input variable that is an invertible transformation of the original input $u(t)$. The crucial point behind this transformation of the objective function and the input variable is that $\bar{u}(t)$ will be shown to be a function of just the summary statistics $\hat{x}(t)$ and $\hat{\bar{x}}(t)$, hence allowing efficient computation of the control law.

Since the new objective depends on $\bar{x}(t)$ and $\bar{u}(t)$, the dynamics of $x(t)$ can be rewritten in terms of these. It can then be shown that the modified problem is of the form as in Proposition 2. Thus, the optimal control law can be obtained for the modified problem, and inverted to get the optimal law for the $(1, \infty)$ state-feedback pattern.

As a preliminary result, the following lemma derives a required uncorrelatedness property and also proves the equivalence of two of the estimator variables.

Lemma 4. *For the nested system (1), $\hat{x}_1(t) = \hat{\bar{x}}_1(t)$. Additionally, $\phi(t) = \bar{\phi}_1(t)$, which implies from Lemma 1 that $\phi(t)$ is uncorrelated with $\mathcal{H}_2(t)$, and so $\mathbb{E}[\phi(t)x(\tau)] = \mathbf{0}$, $\tau < t$.*

Proof. Intuitively, the claim is true because Player 2 has no additional information about block 1's dynamics as compared to Player 1. Formally,

$$\begin{aligned} \hat{x}_1(t) &= \mathbb{E}[x_1(t)|\mathcal{H}_1(t), \mathcal{H}_2(t)], \\ &= A_{11}x_1(t-1) + B_{11}u_1(t-1) = \hat{\bar{x}}_1(t), \end{aligned}$$

where we have used the independence of the zero-mean noise $v_1(t)$ from both $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$, and the definitions of $\phi(t)$ and $\bar{\phi}(t)$. \square

Note that the estimator dynamics of $[\hat{x}_1(t), \hat{\bar{x}}_2(t)]$ can be written in the following way,

$$\begin{aligned} \begin{bmatrix} \hat{x}_1(t+1) \\ \hat{\bar{x}}_2(t+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{\bar{x}}_2(t) \end{bmatrix} \\ &+ \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\ &+ \begin{bmatrix} (\mathcal{R}(t)\phi(t))_1 \\ (\bar{\mathcal{R}}(t)\bar{\phi}(t))_2 \end{bmatrix}. \end{aligned} \quad (25)$$

This shall be denoted in shorthand notation as,

$$\bar{x}(t+1) = A\bar{x}(t) + Bu(t) + \begin{bmatrix} (\mathcal{R}(t)\phi(t))_1 \\ (\bar{\mathcal{R}}(t)\bar{\phi}(t))_2 \end{bmatrix}.$$

Let us denote estimation error by $e(t) = [e_1(t) \ e_2(t)] := [x_1(t) - \hat{x}_1(t) \ x_2(t) - \hat{\bar{x}}_2(t)]$ respectively. Then, $e_1(t) = \phi_1(t)$ and $e_2(t) = \bar{\phi}_2(t)$. By the projection theorem and some manipulations, it can be verified that $\mathbb{E}[e_i(t)\bar{x}_j(t)'] = 0$ for $i, j = 1, 2$.

Further, we define a transformed system input $\bar{u}(t)$ by

$$\bar{u}(t) = \begin{bmatrix} \bar{u}_1(t) \\ \bar{u}_2(t) \end{bmatrix} := u(t) - F^*(t)e(t), \quad (26)$$

where $F^*(t)$ is the optimal gain matrix in (4). Note that the transformation is invertible. From (4), it can be observed that two components of $\bar{u}(t)$ are linear functions of $\mathcal{H}_1(t)$ and $(\mathcal{H}_1(t), \mathcal{H}_2(t))$ respectively, since,

$$\begin{aligned} \bar{u}(t) &= u(t) - F^*(t) \begin{bmatrix} x_1(t) - \hat{x}_1(t) \\ x_2(t) - \hat{\bar{x}}_2(t) \end{bmatrix} \\ &= \begin{bmatrix} F_{11}^*(t)\hat{x}_1(t) + L_1(\mathcal{H}_1(t)) \\ F_{22}^*(t)\hat{\bar{x}}_2(t) + L_2(\mathcal{H}_1(t), \mathcal{H}_2(t)) \end{bmatrix} \\ &= \begin{bmatrix} L'_1(\mathcal{H}_1(t)) \\ L'_2(\mathcal{H}_1(t), \mathcal{H}_2(t)) \end{bmatrix}, \end{aligned} \quad (27)$$

since $\hat{x}(t)$ and $\hat{\bar{x}}(t)$ are linear functions of $\mathcal{H}_1(t)$ and $(\mathcal{H}_1(t), \mathcal{H}_2(t))$, respectively.

Using transformed input $\bar{u}(t)$, (25) can be rewritten as

$$\bar{x}(t+1) = A\bar{x}(t) + B\bar{u}(t) + \tilde{v}(t), \quad (28)$$

where $\tilde{v}(t)$ is equal to

$$\begin{bmatrix} B_{11}F_{11}(t)\phi(t) + (\mathcal{R}(t)\phi(t))_1 \\ B_{21}F_{11}(t)\phi(t) + B_{22}F_{22}(t)\bar{\phi}_2(t) + (\bar{\mathcal{R}}(t)\bar{\phi}(t))_2 \end{bmatrix}.$$

It can be verified from Lemmas 1 and 4 that $\tilde{v}(t)$ is Gaussian, zero-mean and independent across time. With the estimator system well-characterized, we now show that the original objective function (2) is equivalently expressed in terms of $\bar{x}(t)$ and $\bar{u}(t)$.

Rewriting (2) in terms of the new state variable

$\bar{x}(t)$ and the transformed input $\bar{u}(t)$, after noting that the terms involving $e_1(t)$ and $e_2(t)$ are either zero or independent of the input $\bar{u}(t)$, we have,

$$\begin{aligned} & \min_{u(t), 0 \leq t \leq N-1} \mathbb{E} \left[\sum_{t=0}^{N-1} (x(t)' Q x(t) + u(t)' R u(t)) \right. \\ & \quad \left. + x(N)' S x(N) \right] \\ &= \min_{\bar{u}(t), 0 \leq t \leq N-1} \mathbb{E} \left[\sum_{t=0}^{N-1} (\bar{x}(t)' Q \bar{x}(t) + \bar{u}(t)' R \bar{u}(t))' \right. \\ & \quad \left. + e(t)' F^*(t) R F^*(t) e(t) \right. \\ & \quad \left. + 2e(t)' F^*(t) R \bar{u}(t) + \bar{x}(N)' S \bar{x}(N) \right]. \end{aligned} \quad (29)$$

The term $e(t)' F^*(t) R F^*(t) e(t)$ is an estimation error and is independent of the control input $\bar{u}(t)$. This can be seen from the fact that $e_1(t) = \phi_1(t)$ and $e_2(t) = \bar{\phi}_2(t)$ are independent of $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$, by Lemmas 1 and 4. $\bar{u}(t)$ is a linear function of $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$ and so is independent of $e(t)$. In a sense, we have ‘‘subtracted’’ out the dependence on $e(t)$ in (26).

$\bar{u}_1(t)$ and $\bar{u}_2(t)$ are linear functions of $\mathcal{H}_1(t)$ and $(\mathcal{H}_1(t), \mathcal{H}_2(t))$ respectively. So the fourth term has zero expected value since $\mathbb{E}[\phi(t)x(\tau)'] = \mathbf{0} = \mathbb{E}[\bar{\phi}(t)x(\tau)']$, $\tau < t$, from Lemmas 1 and 4. Thus, the following is an equivalent function,

$$\min_{\bar{u}(t)} \mathbb{E} \left[\sum_{t=0}^{N-1} (\bar{x}(t)' Q \bar{x}(t) + \bar{u}(t)' R \bar{u}(t)) + \bar{x}(N)' S \bar{x}(N) \right]. \quad (30)$$

Observe that, with this transformation, solving the $(1, \infty)$ problem is equivalent to solving the cost (30) for the system (28) with no communication delay ($(0, \infty)$ -delayed sharing pattern) in the new state variables. A similar observation has been remarked in [25, IV].

Letting $[\bar{u}_1(t) \ \bar{u}_2(t)]'$ take up the role of players’ inputs, (28) has a nested system structure with no communication delays. By Proposition 2, the optimal controller for the system (28)-(30) is given by,

$$\begin{bmatrix} \bar{u}_1^*(t) \\ \bar{u}_2^*(t) \end{bmatrix} = \begin{bmatrix} K_{11}(t) & K_{12}(t) & \mathbf{0} \\ K_{21}(t) & K_{22}(t) & J(t) \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{\bar{x}}_2(t) - \hat{x}_2(t) \end{bmatrix}, \quad (31)$$

where $\hat{\bar{x}}_2(t) = \mathbb{E}[\hat{x}_2(t) | \bar{x}_1(0:t), \bar{u}_1(0:t-1)]$. As shown below, $\hat{\bar{x}}_2(t)$ is just $\hat{x}_2(t)$.

$$\begin{aligned} \hat{x}_2(t) &= \mathbb{E}[x_2(t) | \mathcal{H}_1(t)], \\ &= \mathbb{E}[\mathbb{E}[x_2(t) | \mathcal{H}_1(t)] | \bar{x}_1(0:t), \bar{u}_1(0:t-1)], \\ &= \mathbb{E}[x_2(t) | \bar{x}_1(0:t), \bar{u}_1(0:t-1)], \\ &= \mathbb{E}[\mathbb{E}[x_2(t) | \mathcal{H}_1(t), \mathcal{H}_2(t)] | \bar{x}_1(0:t), \bar{u}_1(0:t-1)], \\ &= \hat{x}_2(t), \end{aligned}$$

where, in the second equality, the conditional expectation is taken w.r.t. itself and other random variables. The 3rd and 4th equalities follow from tower rule.

Recall that $\bar{u}(t)$ was allowed to be any linear function

of the histories $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$. However, $\bar{u}(t)$ finally depends on just the summary statistics $\hat{x}(t)$ and $\hat{\bar{x}}(t)$, thereby proving that these information statistics are indeed sufficient to compute the optimal law.

From (31) and (26), $u(t)$ is obtained as,

$$u(t) = F^*(t) \begin{bmatrix} x_1(t) - \hat{x}_1(t) \\ x_2(t) - \hat{x}_2(t) \end{bmatrix} + \tilde{K}(t). \quad (32)$$

Proof of Theorem 2. Assuming that optimal gain matrix $F^*(t)$ is unknown, we replace it by arbitrary $F(t)$ and solve for it by deterministic matrix optimization.

From Theorem 1, using the modified input $\bar{u}(t) := u(t) - \begin{bmatrix} F_{11}(t)\phi(t) \\ F_{22}(t)\bar{\phi}_2(t) \end{bmatrix}$, we have,

$$\bar{u}(t) = H(t)\tilde{x}(t), \quad (33)$$

where $H(t) = \begin{bmatrix} K_{11}(t) & K_{12}(t) & -K_{12}(t) \\ K_{21}(t) & K_{22}(t) & J(t) - K_{22}(t) \end{bmatrix}$, and

$$\tilde{x}(t) := \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{\bar{x}}_2(t) - \hat{x}_2(t) \end{bmatrix}.$$

The dynamics of $\tilde{x}(t)$ are obtained from (6), (28) as,

$$\tilde{x}(t+1) = (\tilde{A} + G(t))\tilde{x}(t) + noise(t), \quad (34)$$

where $\tilde{A} := \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A_{22} \end{bmatrix}$, $G(t) := \begin{bmatrix} BH(t) & & \\ 0 & 0 & B_{22}J(t) \\ & \tilde{v}(t) & \\ \tilde{v}_2(t) - (\mathcal{R}(t)\phi(t))_2 & & \end{bmatrix}$, and $noise(t) :=$

Let $\sigma_{F(t)}$ denote the variance of $noise(t)$. Since $F(t)$ minimizes (29), keeping only terms that depend on $F(t)$,

$$\begin{aligned} J_F = \min_{F(t)} \mathbb{E} \left[\sum_{t=0}^{N-1} \left(\tilde{x}(t)' \tilde{Q} \tilde{x}(t) + \tilde{x}(t)' H(t)' R H(t) \tilde{x}(t) \right. \right. \\ \left. \left. + \bar{\phi}(t)' F(t)' R F(t) \bar{\phi}(t) \right) + \tilde{x}(N)' \tilde{S} \tilde{x}(N) \right], \end{aligned} \quad (35)$$

where $\tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}$ and $\tilde{S} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$. Using matrix algebra, 35 simplifies to,

$$\begin{aligned} J_F = \min_{F(t)} \left[\sum_{t=0}^{N-1} \text{trace} \left((\tilde{Q} + H(t)' R H(t)) \Sigma(t) \right) \right. \\ \left. + \text{trace} \left(\tilde{S} \Sigma(N) \right) + \sum_{t=0}^{N-1} \text{trace} \left(F(t)' R F(t) \bar{T}(t) \right) \right], \end{aligned}$$

where $\Sigma(t) := \mathbb{E}[\bar{x}(t)\bar{x}(t)']$, with dynamics from (34),

$$\begin{aligned} \Sigma(t+1) &= (\tilde{A} + G(t))\Sigma(t)(\tilde{A} + G(t))' + \sigma_{F(t)}, \\ \Sigma(0) &= \mathbb{E}[\bar{x}(0)\bar{x}(0)'] = \mathbf{0}. \end{aligned} \quad (36)$$

where $\sigma_{F(t)}$ is the covariance of $n(t)$. $F^*(t)$ may now be obtained using the discrete matrix minimum princi-

ple [30] through the Hamiltonian to minimize (35).

$$\begin{aligned}\mathcal{H} &= \text{trace}[(\tilde{Q} + H(t)'RH(t))\Sigma(t)] \\ &+ \text{trace}[F(t)'RF(t)\bar{T}(t)] \\ &+ \text{trace}[(\tilde{A} + G(t))\Sigma(t)(\tilde{A} + G(t))'M(t+1)'] \\ &+ \text{trace}[\sigma_{F(t)}M(t+1)'] + \text{trace}[2F(t)\Psi(t)'] \quad (37)\end{aligned}$$

where $\Psi(t) = \begin{bmatrix} 0 & \Psi_{12}(t) \\ \Psi_{21}(t) & 0 \end{bmatrix}$ is a suitably partitioned Lagrange multiplier matrix and $M(t)$ is the costate matrix. The optimality criteria are,

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial F(t)} &= 0, \quad \frac{\partial \mathcal{H}}{\partial \Psi(t)} = 0, \quad \frac{\partial \mathcal{H}}{\partial M(t)} = \Sigma^*(t), \quad \Sigma^*(0) = \Sigma(0) \\ \frac{\partial \mathcal{H}}{\partial \Sigma(t)} &= M^*(t), \quad M^*(N) = \frac{\partial \text{trace}[S\Sigma(N)]}{\partial \Sigma(N)}\end{aligned}$$

Solving the above equations, we get,

$$2RF^*(t)\bar{T}(t) + \frac{\partial \text{trace}[\sigma_{F(t)}M(t+1)']}{\partial F(t)} + 2\Psi^*(t) = 0$$

$$F_{12}^*(t) = F_{21}^*(t) = 0$$

$$\Sigma^*(t+1) = (\tilde{A} + G(t))\Sigma(t)(\tilde{A} + G(t))' + \sigma_{F(t)},$$

$$\Sigma^*(0) = 0$$

$$M^*(t) = \tilde{Q} + H(t)'RH(t)$$

$$+ (\tilde{A} + G(t))'M^*(t+1)(\tilde{A} + G(t)), \quad M^*(N) = S. \quad (38)$$

The noise term in (34) can be rewritten as, $\text{noise}(t) =$

$$\begin{bmatrix} B_{11}F_{11}(t)\bar{\phi}_1(t) + (\mathcal{R}(t)\bar{\phi}_1(t))_1 \\ B_{21}F_{11}(t)\bar{\phi}_1(t) + B_{22}F_{22}(t)\bar{\phi}_2(t) + (\mathcal{R}(t)\bar{\phi}(t))_2 \\ B_{22}F_{22}(t)\bar{\phi}_2(t) + (\mathcal{R}(t)\bar{\phi}(t))_2 - (\mathcal{R}(t)\bar{\phi}_1(t))_2 \end{bmatrix} = \tilde{B}F(t)\bar{\phi}(t) + n_1(t),$$

$$\text{where } \tilde{B} := \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \\ 0 & B_{22} \end{bmatrix}, \quad \text{and } n_1(t) :=$$

$$\begin{bmatrix} (\mathcal{R}(t)\bar{\phi}_1(t))_1 \\ (\mathcal{R}(t)\bar{\phi}(t))_2 \\ (\mathcal{R}(t)\bar{\phi}(t))_2 - (\mathcal{R}(t)\bar{\phi}_1(t))_2 \end{bmatrix}. \quad \text{Then,}$$

$$\begin{aligned}\sigma_{F(t)} &= \mathbb{E}[(\tilde{B}F\bar{\phi}(t) + n_1(t))(\tilde{B}F\bar{\phi}(t) + n_1(t))'] \\ &\equiv \mathbb{E}[\tilde{B}F(t)\bar{\phi}(t)n_1(t)' + \tilde{B}F(t)\bar{\phi}(t)(\tilde{B}F(t)\bar{\phi}(t))' \\ &\quad + n_1(t)n_1(t)' + n_1(t)(\tilde{B}F(t)\bar{\phi}(t))'] \\ &= \tilde{B}F(t)\mathbb{E}[\bar{\phi}(t)n_1(t)'] + \tilde{B}F(t)\mathbb{E}[\bar{\phi}(t)\bar{\phi}(t)']F(t)'\tilde{B}' \\ &\quad + \mathbb{E}[n_1(t)\bar{\phi}(t)']F(t)'\tilde{B}' \\ &= \tilde{B}F(t)V(t) + \tilde{B}F(t)\bar{T}(t)F(t)'\tilde{B}' + V(t)'F(t)'\tilde{B}'\end{aligned}$$

where $V(t) = \mathbb{E}[\bar{\phi}(t)n_1(t)']$ and equivalence in the 2^{nd} line is from taking partial derivative of $\sigma_{F(t)}$ w.r.t. $F(t)$.

Taking the partial derivative of $\text{trace}[\sigma_{F(t)}M(t+1)']$,

$$\begin{aligned}&\frac{\partial \text{trace}[\sigma_{F(t)}M(t+1)']}{\partial F(t)} \\ &= \frac{\partial}{\partial F(t)} \left(\text{trace}[(\tilde{B}F(t)V(t) + \tilde{B}F(t)\bar{T}(t)F(t)'\tilde{B}' \right. \\ &\quad \left. + V(t)'F(t)'\tilde{B}')M(t+1)'] \right) \\ &= \tilde{B}'M(t+1)V(t)' + \tilde{B}'M(t+1)\tilde{B}F(t)\bar{T}(t) \\ &\quad + \tilde{B}'M(t+1)'\tilde{B}F(t)\bar{T}(t) + \tilde{B}'M(t+1)'V(t)'\end{aligned}$$

Substituting into the optimality condition (38) we get,

$$\begin{aligned}F^*(t) &= - \left(2R + \tilde{B}'M(t+1)\tilde{B} + \tilde{B}'M(t+1)'\tilde{B} \right)^{-1} \times \\ &\quad \left(2\Psi^*(t) + \tilde{B}'M(t+1)V(t)' \right. \\ &\quad \left. + \tilde{B}'M(t+1)'V(t) \right) \bar{T}^{-1}(t) \\ &= - \left(R + \tilde{B}'M(t+1)\tilde{B} \right)^{-1} \times \\ &\quad \left(\Psi^*(t) + \tilde{B}'M(t+1)V(t)' \right) \bar{T}^{-1}(t) \quad (39)\end{aligned}$$

where the last equality uses $M(t)$ being symmetric.

The following is a method to obtain the gain matrix $F^*(t)$ explicitly from the above equation. Rearranging (39) results in the following equation:

$$\begin{aligned}\left(R + \tilde{B}'M(t+1)\tilde{B} \right) F^*(t)\bar{T}(t) + \Psi^*(t) \\ + \tilde{B}'M(t+1)V(t)'\end{aligned} \quad (40)$$

Using $\Psi_{11}^*(t) = \Psi_{22}^*(t) = 0$, we get,

$$\begin{aligned}E_1' \left(R + \tilde{B}'M(t+1)\tilde{B} \right) F^*(t)\bar{T}(t)E_1 \\ = -E_1'\tilde{B}'M(t+1)V(t)'\end{aligned} \quad (41)$$

$$\begin{aligned}E_2' \left(R + \tilde{B}'M(t+1)\tilde{B} \right) F^*(t)\bar{T}(t)E_2 \\ = -E_2'\tilde{B}'M(t+1)V(t)'\end{aligned} \quad (42)$$

where E_1 and E_2 are suitable partitions of the identity matrix $I = [E_1 \ E_2]$. We also have, $F_{12}^*(t) = F_{21}^*(t) = 0$, giving us $F^*(t) = E_1F_{11}^*(t)E_1' + E_2F_{22}^*(t)E_2'$. Substituting into the above equation, we get,

$$\begin{aligned}\left[R + \tilde{B}'M(t+1)\tilde{B} \right]_{11} F_{11}^*(t)\bar{T}_{11}(t) \\ + \left[R + \tilde{B}'M(t+1)\tilde{B} \right]_{12} F_{22}^*(t)\bar{T}_{21}(t) \\ = - \left[\tilde{B}'M(t+1)V(t)' \right]_{11}\end{aligned} \quad (43)$$

$$\begin{aligned}\left[R + \tilde{B}'M(t+1)\tilde{B} \right]_{21} F_{11}^*(t)\bar{T}_{12}(t) \\ + \left[R + \tilde{B}'M(t+1)\tilde{B} \right]_{22} F_{22}^*(t)\bar{T}_{22}(t) \\ = - \left[\tilde{B}'M(t+1)V(t)' \right]_{22}, \quad (44)\end{aligned}$$

where $[A]_{ij} = E_i' A E_j$. This is a complete set of linear equations that can be solved directly, for example, by vectorization using the Kronecker product $AXB = C \Leftrightarrow (B' \otimes A)x = c$, where \otimes is the Kronecker product and x, c are vectors made by assembling columns of X, C . \square

Proof of Theorem 3. Only the outline of the proof is given here for brevity. As in the proof of Theorem 1, we define a modified input,

$$\bar{u}(t) := u(t) - F^*(t)\phi_o(t),$$

$$\text{where } \phi_o(t) := \begin{bmatrix} y_1(t) - C_{11}\hat{x}_1(t) \\ y_2(t) - C_{21}\hat{x}_1(t) - C_{22}\hat{x}_2(t) \end{bmatrix}.$$

It can be shown that $\bar{u}(t)$ can be decomposed into a coordinator's action, based on common history $\mathcal{H}_1(t)$, and the remaining part based on the entire information available to player 2 [18, Theorem 10]. That is,

$$\bar{u}(t) = \tilde{u}(t) + \tilde{\tilde{u}}(t),$$

where $\tilde{u}(t)$ is obtained by fixing $\bar{u}(t) := \tilde{u}(t) + \hat{K}(t)\hat{x}(t)$ and solving the optimal control problem. Subsequently, $\tilde{\tilde{u}}(t)$ is obtained by fixing $\bar{u}(t) := K(t)\hat{x}(t) + [0 \ \tilde{\tilde{u}}]'$ and solving the LQR problem [18, Theorem 10]. Person-by-person optimality implies global optimality since the problem is partially nested [14].

It can be verified through the above approach that the additional $F^*(t)\phi_o(t)$ term does not affect the computation of $\tilde{u}(t)$ and $\tilde{\tilde{u}}(t)$ and,

$$\tilde{u}(t) = (K(t) - \hat{K}(t))\hat{x}(t), \tilde{\tilde{u}}(t) = \begin{bmatrix} 0 \\ \hat{K}(t)(\hat{x}(t) - \hat{x}(t)) \end{bmatrix}$$

\square

REFERENCES

- [1] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 274–286, 2006.
- [2] C. Langbort, R. S. Chandra, and R. D'Andrea, "Distributed control design for systems interconnected over an arbitrary graph," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1502–1519, Sept 2004.
- [3] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *IEEE Transactions on Automatic Control*, vol. 53, pp. 1616–1629, Aug 2008.
- [4] J. Marschak and R. Radner, "Economic theory of teams," Cowles Foundation Discussion Papers 59e, Cowles Foundation for Research in Economics, Yale University, 1958.
- [5] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 857–881, 1962.
- [6] H. S. Witsenhausen, "A Counterexample in Stochastic Optimum Control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.
- [7] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.
- [8] P. Varaiya and J. Walrand, "On delayed sharing patterns," *IEEE Transactions on Automatic Control*, vol. 23, no. 3, pp. 443–445, 1978.
- [9] A. Nayyar, A. Mahajan, and D. Teneketzis, "Optimal control strategies in delayed sharing information structures," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1606–1620, 2011.
- [10] 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, pp. 15–22, 1972.
- [11] 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.
- [12] 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.
- [13] T. Yoshikawa, "Dynamic programming approach to decentralized stochastic control problems," *IEEE Transactions on Automatic Control*, vol. 20, no. 6, pp. 796–797, 1975.
- [14] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yuksel, "Information structures in optimal decentralized control," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec 2012.
- [15] B. Bamieh and P. G. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Systems and Control Letters*, vol. 54, no. 6, pp. 575 – 583, 2005.
- [16] J. Swigart and S. Lall, "An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator," in *American Control Conference (ACC), 2010*, pp. 6385–6390, 2010.
- [17] J. Swigart and S. Lall, "Optimal controller synthesis for a decentralized two-player system with partial output feedback," in *American Control Conference (ACC), 2011*, pp. 317–323, 2011.
- [18] L. Lessard and A. Nayyar, "Structural results and explicit solution for two-player lqg systems on a finite time horizon," in *IEEE Conference on Decision and Control (CDC)*, pp. 6542–6549, 2013.
- [19] A. Lamperski and L. Lessard, "Optimal state-feedback control under sparsity and delay constraints," in *IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pp. 204–209, 2012.
- [20] A. Lamperski and L. Lessard, "Optimal decentralized state-feedback control with sparsity and delays," *Automatica*, vol. 58, pp. 143 – 151, 2015.
- [21] A. Nayyar and L. Lessard, "Structural results for partially nested lqg systems over graphs," in *American Control Conference (ACC), 2015*, pp. 5457–5464, July 2015.
- [22] N. Matni, A. Lamperski, and J. C. Doyle, "Optimal two player lqr state feedback with varying delay," in *IFAC WC 2014*, 2014.
- [23] P. Shah and P. A. Parrilo, "H2-optimal decentralized control over posets: A state space solution for state-feedback," in *IEEE Conference on Decision and Control (CDC)*, pp. 6722–6727, 2010.
- [24] A. Lamperski and J. C. Doyle, "Dynamic programming solutions for decentralized state-feedback lqg problems with communication delays," in *American Control Conference (ACC), 2012*, 2012.
- [25] A. Lamperski and J. C. Doyle, "Output feedback H2 model matching for decentralized systems with delays," in *American Control Conference (ACC), 2013*, pp. 5778–5783, 2013.
- [26] L. Lessard, "Optimal control of a fully decentralized quadratic regulator," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pp. 48–54, 2012.
- [27] N. Nayyar, D. Kalathil, and R. Jain, "Optimal decentralized control in unidirectional one-step delayed sharing pattern with partial output feedback," in *American Control Conference (ACC), 2014*, pp. 1906–1911, June 2014.
- [28] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, Inc., 1986.
- [29] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. John Wiley & Sons, 1972.
- [30] D. L. Kleinmann and M. Athans, "The discrete minimum principle with application to the linear regulator problem," *Electronic Systems Laboratory, MIT, Rep 260*, 1966.