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 controlled by two players having asymmetric information sharing patterns between them. In one scenario, players are assumed to have a bidirectional error-free, unlimited rate communication channel with no delay in one direction and a unit delay in the other. In another scenario, the communication channel is assumed to be unidirectional with a unit delay. 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 problem has a partially nested information structure, and thus linear optimal control laws exist. Summary statistics to characterize these laws are developed and deterministic convex optimization problems are formulated whose solutions yield the optimal control laws. The state feedback case is solved for both scenarios and extended to output and partial output feedback in case of bidirectional and unidirectional channels respectively.

I. Introduction

Recently, many problems of decentralized control have arisen in practical systems. Examples include cyberphysical systems, formation flight, and several other networked control systems.
wherein multiple agents try to achieve a common objective in a decentralized manner. Such situations arise, for instance, due to controllers not having access to the same information at the same time. One possible scenario is because of network delays, and a consequent time-lag to communicate observations to other controllers.

Such problems were first formulated by Marschak in the 1950s [11] as team decision problems, and further studied by Radner [2], though in such problems communication between the controllers was usually ignored. Problems were formulated as a ‘one-shot’ decision problem where each controller acted once. A few cases were identified that guaranteed linear optimal control laws, making their computation tractable. In [3], Witsenhausen - in a much celebrated paper - showed that even for seemingly simple systems with communication between controllers but with multi-unit delays, the optimal linear controllers were sub-optimal. This was established via a counterexample for which a non-linear controller was given that outperformed any linear controller. Witsenhausen also consolidated and conjectured results on separation of estimation and control in decentralized problems in [4]. 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 conjectured separation principle 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 [5]. It remains an open problem to actually compute the optimal decentralized control law 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 established in [6] that for a unit-delay information sharing pattern, the optimal controller is linear. This was used by Kurtaran and Sivan [7] to derive an optimal controller using dynamic programming for the finite-horizon case. Unfortunately, the results 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 control laws are known to exist [6].

Recently, another sufficient characterization, called quadratic invariance, was discovered under which optimal decentralized controls laws can be solved by convex programming [9]. This
was later shown to be equivalent to ‘partially nested information structure’. This has led to a resurgence of interest in decentralized optimal control problems, which since the 1970s had been assumed to be intractable. In a series of papers, Lall and his co-workers have computed the optimal (linear) control laws for a suite of decentralized nested \textit{LQG systems}, involving no communication delays \cite{10}, \cite{11}. They computed the optimal law for both state-feedback and output-feedback cases. For a subclass of quadratic invariant problems known as \textit{poset-causal} problems, Parrilo and Shah \cite{12} showed that the computation of the optimal law becomes easier by lending itself to decomposition into subproblems. Solutions to restricted cases of the delayed information sharing problem have been presented by Lamperski et al. \cite{13} wherein a specific networked graph structure of the controllers is considered with constraints on the system dynamics. It is shown that this problem can be decomposed into a hierarchy of subproblems that can be solved independently by dynamic programming for the finite-horizon case. We note that in these problems, no rate constraints are imposed on the communication channels. They are assumed error-free and to have no delay. A summary of results pertaining to two-player decentralized controller systems is given in Table \ref{tab:summary}

<table>
<thead>
<tr>
<th>(d_{12})</th>
<th>(d_{21})</th>
<th>Literature</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Classical LQG</td>
<td>no restrictions on plant</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Kurtaran \cite{7}, Sandell \cite{14}, Yoshikawa \cite{15}</td>
<td>no restrictions on plant</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Lamperski \cite{13}</td>
<td>b.d B matrix, state f/b only</td>
</tr>
<tr>
<td>0</td>
<td>(\infty)</td>
<td>Lall et al. \cite{10}, \cite{11}</td>
<td>l.b.t matrix</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>This paper</td>
<td>no restrictions on plant</td>
</tr>
<tr>
<td>1</td>
<td>(\infty)</td>
<td>This paper</td>
<td>l.b.t matrix, state + partial output f/b</td>
</tr>
</tbody>
</table>

\begin{table}[h]
\centering
\caption{Summary of results pertaining to two-player unit-delayed information sharing patterns}
\begin{tabular}{|c|c|c|}
\hline
\(d_{12}\) & \(d_{21}\) & Literature & Comments \\
\hline
0 & 0 & Classical LQG & no restrictions on plant \\
1 & 1 & Kurtaran \cite{7}, Sandell \cite{14}, Yoshikawa \cite{15} & no restrictions on plant \\
1 & 1 & Lamperski \cite{13} & b.d B matrix, state f/b only \\
0 & \(\infty\) & Lall et al. \cite{10}, \cite{11} & l.b.t matrix \\
0 & 1 & This paper & no restrictions on plant \\
1 & \(\infty\) & This paper & l.b.t matrix, state + partial output f/b \\
\hline
\end{tabular}
\end{table}

In Table \ref{tab:summary}, \(d_{12}\) is the delay in information transmission from player 1 to player 2, and vice versa. 'b.d' and 'l.b.t' refer to block diagonal and lower block triangular matrices respectively. Standard conditions for stability and controllability of the system are assumed to hold in all cases.

In this paper, we consider two scenarios of asymmetric delayed information sharing for nested
LQG systems. In both cases, we restrict our attention to two players. In the first scenario, there is a bidirectional communication channel between the two players. There is unit delay in information transmission from player 1 to player 2 and no delay from player 2 to player 1. We denote this system as having a \((1, 0)\) information sharing pattern. This system has a partially nested information structure and also satisfies quadratic invariance. It, however, is not poset-causal and does not lend itself to an easy decomposition. It will be shown that the problem has an optimal decentralized control law that is linear. Our goal is to compute this control law. This is accomplished by identifying a suitable information statistic, which then is used in formulating a deterministic convex optimization problem that yields the optimal control law. Results are proved for both state and output feedback cases.

In the second scenario, we extend the approach to the problem of one-way communication: a unit-delay exists in information transmission from player 1 to player 2 and no communication occurs from player 2 to player 1. That is, a unidirectional one-step delayed communication channel exists between the players, referred to as a \((1, \infty)\) information sharing pattern. State and partial-output feedback results are proved for this model. We note that such literature on decentralized control problems with asymmetric delayed information sharing patterns is very sparse and the area is only very newly emerging.

The paper is organized in the following manner. Sections II and III deal with \((1, 0)\) and \((1, \infty)\) information sharing patterns respectively. Each section comprises of four subsections. In the first subsection, linearity of the optimal law is proven. In the second, summary statistics are developed and their properties are studied. The third subsection proves the main result for the state-feedback case, and the fourth extends the approach to output and partial-output feedback in the respective scenarios.

### II. The \((1, 0)\) Information Sharing Pattern

We consider a two-player discrete linear time-invariant system. The system dynamics are specified as

\[
x(t + 1) = Ax(t) + Bu(t) + v(t)
\]  

(1)
We will assume that the initial state $x(0)$ is a zero-mean Gaussian random vector, independent of the system noise $v(t)$, and that $v(t)$ are zero-mean Gaussian random vectors with covariance $V$, that are independent across time.

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

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

(2)

where $Q$, $R$ and $S$ are positive semi-definite matrices.

At each instant, the control action $u_i(t)$ taken by each player can only depend on the information available to them. This is called the information structure, denoted by $\mathcal{F}(t) = (\mathcal{F}_1(t), \mathcal{F}_2(t))$. We consider the following information structure. At each time $t$, player $i$ observes its state $x_i(t)$. And then, player 1 shares its information (state $x_1(t)$ and control action $u_1(t - 1)$) with player 2. There is a unit delay in this communication. However, when player 2 shares its information (state $x_2(t)$ and control action $u_2(t)$) with player 1, there is no delay. The information available to the players at time $t$ is given by

$$\mathcal{F}_1(t) = \{x_1(0 : t), u_1(0 : t - 1), x_2(0 : t), u_2(0 : t - 1)\},$$

$$\mathcal{F}_2(t) = \{x_1(0 : t - 1), u_1(0 : t - 1), x_2(0 : t), u_2(0 : t - 1)\},$$

(3)

where the notation $x(0 : t)$ denotes the vector $(x(0), \cdots, x(t))$.

Given this information structure, we refer to the optimal control problem as having a $(1, 0)$ information sharing pattern. Thus, the optimal law is characterized by $u_i^*(t) = f(\mathcal{F}_i(t))$ for some function $f$. This implies a huge function space in which to search for the optimal control law. In the next section, we show that the optimal decentralized control law for the LQG system with asymmetric unit-delay information structure is linear.

A. 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, earlier Ho
and Chu [6] had established that for a symmetric one-step delayed information structure, an optimal decentralized control law that is linear exists. They also showed that to establish linearity in general, it is sufficient to show that the LQG problem formulation has a *partially nested information structure*.

Intuitively, a partially nested structure can be described as an information structure that allows communication of all information used by a player to make a decision to all players whose system dynamics are affected by that decision. That is, if a certain player’s system dynamics are affected by actions taken by other players, then that player has all that same information (that others used to make decisions) to make its decision, and possibly some more. The following proposition shows that \((1, 0)\)-delayed sharing problem has a partially nested information structure.

**Proposition 1.** There exists a linear optimal decentralized control law for the LQG problem \((1)\)\-(2) with the information structure \((3)\).

**Proof:** Observe that the information structure defined in \((3)\) can be simplified as \(\mathcal{F}_1(t) = \{x_1(0 : t), x_2(0 : t)\}\) and \(\mathcal{F}_2(t) = \{x_1(0 : t - 1), x_2(0 : t)\}\). This is because the control actions \(u_i(\cdot)\) can be derived from the state information available, given a fixed control law. From the simplified expressions, it can be seen that, at any time instant \(t\), \(\mathcal{F}_i(t) \subset \mathcal{F}_i(t + \tau), \tau \geq 0, i = 1, 2\) and \(\mathcal{F}_1(t) \subset \mathcal{F}_2(t + \tau), \tau \geq 0\). This is illustrated in the information communication diagram shown in Fig 1.

![Information communication diagram for one-step delayed information-sharing in two-player D-LTI systems with nested structures](image)

Fig. 1. Information communication diagram for one-step delayed information-sharing in two-player D-LTI systems with nested structures.

The dependence of state variables on the input is also the same as in Fig 1. The \((1, 0)\)-delayed sharing system formulation thus has a partially nested structure [6] and by Theorem 1 therein the optimal control law is linear. 

Denote by \(\mathcal{H}(t) = (x(0 : t - 1), u(0 : t - 1))\), the history of observations (common information)
at time $t$. Using the linearity of the optimal control law, it follows that the optimal control law can be written as $u(t) = [u_1(t), u_2(t)]' = \begin{bmatrix} F_{11}^*(t) & F_{12}^*(t) \\ 0 & F_{22}^*(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} L_1(H(t)) \\ L_2(H(t)) \end{bmatrix}$, \hspace{1cm} (4)

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

However, since the arguments of $L_i$, namely the state observations, increase with time, obtaining a closed form expression for the optimal control law is not possible in the current form. Instead, we identify summary statistics that compress the information structure without losing necessary information, but keeps the dimension of the argument of the functions $L_i$ fixed.

**B. A Summary Statistic to Compute The Optimal Control Law**

We will now show that the estimate of the current state given the history of observations available to the players is a summary statistic that characterizes the optimal control law. Recall that the information available to player 1 at time $t$ is $\mathcal{F}_1(t) = \{x_1(t), x_2(t), H(t)\}$, and that available to Player 2 is $\mathcal{F}_2(t) = \{x_2(t), H(t)\}$.

Define the estimator $\hat{x}(t) := \mathbb{E}[x(t)|H(t)]$ to be the estimate of the current states by the players, based on the history of information available to them. The dynamics of these estimators are known through the following lemma.

**Lemma 1. \[16\] For the system (1), the estimator dynamics for $\hat{x}(t)$ are given by

\[ \hat{x}(t + 1) = A\hat{x}(t) + Bu(t) + R(t)\phi(t) \hspace{1cm} (5) \]

with $\hat{x}(0) = \mathbb{E}[x(0)]$, where $R(t) := R_{xx}(t)R_x^{-1}(t)$, $\phi(t) := x(t) - \hat{x}(t)$, $R_{xx}(t) := AT(t)$, $R_x(t) = T(t)$, $T(t) := \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))]$.

Also, $\phi(t)$ is zero-mean, uncorrelated Gaussian and $\mathbb{E}[\phi(t)\phi(t)'] = T(t)$. Additionally, $\mathbb{E}[\phi(t)x(\tau)'] = 0$ for $\tau < t$.

**Proof:** The proof is available in [16]. \[ \blacksquare \]

The sufficiency of the estimator $\hat{x}(t)$ for optimal control is established below.
Theorem 1. For the system (1) with the information structure (3) and objective function (2), the optimal control law is given by

$$u(t) = F^*(t)x(t) + G^*(t)\hat{x}(t),$$

(6)

where $F^*(t) = \begin{bmatrix} F_{11}^*(t) & F_{12}^*(t) \\ 0 & F_{22}^*(t) \end{bmatrix}$ is referred to as the optimal gain matrix (to be derived later) and $G^*(t) = -(F^*(t) + H^*(t))$ where $H^*(t)$ is the optimal gain matrix for the classical LQR problem with plant dynamics described by (32).

Proof: The proof follows from a transformation of the original problem to which existing LQR results can be applied. It is given in Appendix A.

C. Deriving the Optimal Gain Matrix

We now characterize the optimal gain matrix, $F^*(t)$. This problem is solved in a manner similar to the symmetric one-step delayed information sharing scenario [7]: The stochastic optimization problem is converted to a deterministic matrix optimization problem, after which standard matrix optimization tools can be applied to get the optimal law.

Theorem 2. The optimal gain matrix $F^*(t)$ is obtained by solving the deterministic optimization problem,

$$J_F = \sum_{t=0}^{N-1} \text{trace}\left((Q + H^*(t)'RH^*(t))\Sigma(t)\right) + \text{trace}(S\Sigma(N)) + \sum_{t=0}^{N-1} \text{trace}\left(F(t)'RF(t)T(t)\right),$$

subject to the deterministic system dynamics

$$\Sigma(t + 1) = (A - BH^*(t))\Sigma(t)(A - BH^*(t))' + (BF(t) + R(t))T(t)(BF(t) + R(t))', \quad (7)$$

$$\Sigma(0) = 0. \quad (8)$$

Proof: The proof is in Appendix B.
The optimal gain matrix $F^*(t)$ may now be obtained using matrix optimization tools such as the discrete matrix minimum principle through the Hamiltonian. This gives the same expression as in [7] with the structure of $F^*(t)$ defined above.

D. Optimal Control law for the Output-feedback problem

We now turn our attention to the output-feedback problem. It will be shown that the results from the state-feedback problem can be extended to the output-feedback case directly with a change of notation. Before proceeding with the problem, it is useful to formulate it precisely.

The system dynamics, given that the output variables $y(t) := \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$ are observed, are governed by the following equations.

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

$$
y(t) = Cx(t) + w(t),
$$

(9)

where $v(t)$ and $w(t)$ are zero-mean, Gaussian random vectors with covariance matrices $V$ and $W$ respectively that are independent across time $t$ and independent of initial system state $x(0)$ and of each other. $x(0)$ is a zero-mean, Gaussian random vector with known mean and covariance.

The information structure for the problem is,

$$
\mathcal{F}_1(t) = \{y_1(0 : t), u_1(0 : t - 1), y_2(0 : t), u_2(0 : t - 1)\},
$$

$$
\mathcal{F}_2(t) = \{y_1(0 : t - 1), u_1(0 : t - 1), y_2(0 : t), u_2(0 : t - 1)\}.
$$

(10)

We redefine the history of observations as, $\mathcal{H}(t) = (y(0 : t - 1), u(0 : t - 1))$. 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 easily seen from Proposition [1] that the optimal control law for the output-feedback problem is linear. Further, defining estimator $\hat{x}(t) := \mathbb{E}[x(t)|\mathcal{H}(t)]$ as before, the following lemma can be reconstructed from Lemma [1] Note that the only significant differences are in the definition of $\phi(t)$ and $R(t).$
Lemma 2. For the nested system (9), the estimator dynamics for the redefined \( \hat{x}(t) \) are as follows.

\[
\hat{x}(t + 1) = A\hat{x}(t) + Bu(t) + K(t)\phi(t),
\]

and \( \hat{x}(0) = \hat{x}(0) = \mathbb{E}[x(0)] \), where, \( K(t) = K_{xy}(t)K_y(t)^{-1} \) and \( \phi(t) = y(t) - C\hat{x}(t) \). \( K_{xy}(t) = AT(t)C' \), \( K_y(t) = CT(t)C' + W \). \( T(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))] \).

Also, \( \phi(t) \) is zero-mean, uncorrelated Gaussian and \( \mathbb{E}[\phi(t)\phi(t)'] = CT(t)C' + W \). Additionally, \( \mathbb{E}[\phi(t)y(\tau)'] = 0 \) for \( \tau < t \).

**Proof:** The proof is available in [16].

The following theorem without proof (following the same steps as in the proof of Theorem 1), describes the structure of the optimal control law for the output-feedback problem.

**Theorem 3.** For the nested system (9) with the information structure (10) and objective function (2),

\[
\begin{bmatrix}
  u^*_1(t) \\
  u^*_2(t)
\end{bmatrix} = F^*(t) \begin{bmatrix}
  y_1(t) \\
  y_2(t)
\end{bmatrix} + G^*(t)\hat{x}(t),
\]

(12)

where \( F^*(t) = \begin{bmatrix}
  F^*_{11}(t) & F^*_{12}(t) \\
  0 & F^*_{22}(t)
\end{bmatrix} \) is the optimal gain matrix and \( G^*(t) = -(F^*(t)C + H^*(t)) \).

**Proof:** The proof is the same as that for the state-feedback case in Theorem [1].

Finally, to complete the output-feedback problem, the optimal gain matrix \( F^*(t) \) is characterized by the following theorem.

**Theorem 4.** The optimal gain matrix \( F^*(t) \) is obtained by solving the convex optimization problem,

\[
J_F = \sum_{t=0}^{N-1} \text{trace} \left( (Q + H^*(t)'RH^*(t))\Sigma(t) \right) + \text{trace} \left( S\Sigma(N) \right) + \sum_{t=0}^{N-1} \text{trace} \left( F(t)'RF(t)(CT(t)C' + W) \right),
\]
subject to the deterministic system dynamics,

\[ \Sigma(t + 1) = (A - BH^*(t))\Sigma(t)(A - BH^*(t))' + (BF(t) + K(t))(CT(t)C' + W)(BF(t) + K(t))', \]

\[ \Sigma(0) = 0 \]

with the terms defined in Lemma 2.

Proof: The proof is the same as that for the state-feedback case in Theorem 2.

\( F^*(t) \) can be computed analytically through the above deterministic dynamic optimization problem as in the state-feedback case by using the Hamiltonian through the discrete maximum principle [17]. This gives the same expression as in [7] with the structure of \( F^*(t) \) as defined above. This completes the output-feedback case for the \((1, 0)\) information sharing pattern.

III. The \((1, \infty)\) Information Sharing Pattern

In this scenario, we consider another coupled two-player discrete linear time-invariant (LTI) system with a nested structure. However, we assume additional conditions on the coupling structure. Player 1’s actions affect only block 1 of the plant whereas player 2’s actions affect both blocks 1 and 2. Formally, the system dynamics are specified as

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

for \( t \in \{0, 1, \cdots, N - 1\} \).

Here, \( x_i(t) \) is block \( i \)'s state output at time \( t \) and \( u_i(t) \) is player \( i \)'s control input at time \( t \).

Henceforth, we will often use the notation \( x(t+1) = Ax(t) + Bu(t) + v(t) \) where the definition of \( A, B, x(t), u(t) \) and \( v(t) \) should be clear from the above equation. We will assume that the initial state \( x(0) \) is a zero-mean Gaussian random vector, independent of the system noise \( v(t) \), and that \( v(t) \) is zero-mean Gaussian with covariance \( V \), which is independent across time \( t \).
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

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

(14)

where \(Q, R\) and \(S\) are positive semi-definite matrices.

The information structure, denoted by \(\mathcal{F}(t) = (\mathcal{F}_1(t), \mathcal{F}_2(t))\), for this problem is as follows. At each time \(t\), player \(i\) observes its state \(x_i(t)\). And then, player 1 shares its information (state \(x_1(t)\) and control action \(u_1(t)\)) with player 2. However, this is received after a unit delay. The information available to the players at time \(t\) is given by

\[
\mathcal{F}_1(t) = \{x_1(0 : t), u_1(0 : t – 1)\},
\]

\[
\mathcal{F}_2(t) = \{x_1(0 : t – 1), u_1(0 : t – 1); x_2(0 : t), u_2(0 : t – 1)\},
\]  

(15)

where the notation \(x(0 : t)\) denotes the vector \((x(0), \cdots, x(t))\).

Given this information structure, the set of all control laws is characterized by \(u_i(t) = f(\mathcal{F}_i(t))\) for some time-varying, possibly non-linear function \(f\). In the next subsection, we show that the optimal decentralized control law for the problem formulated above is linear.

A. Linearity of the optimal control law

As in the \((1, 0)\) information sharing pattern, we show that the \((1, \infty)\) pattern also has a partially nested structure.

**Proposition 2.** There exists a linear optimal decentralized control law for the LQG problem \((13)-(14)\) with the information structure \((15)\).

**Proof:** Observe that the information structure defined in \((15)\) can be simplified as \(\mathcal{F}_1(t) = \{x_1(0 : t)\}\) and \(\mathcal{F}_2(t) = \{x_1(0 : t – 1), x_2(0 : t)\}\). Because of the nested system structure, \(x_1(t)\) does not depend on \(u_2(\cdot)\). 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 is illustrated in the information communication diagram shown in Fig 2 that shows how observations are communicated between players.
Observe that $A_{12} = B_{12} = 0$ implies that the dependence of state variables on the input is also the same as in Fig 2. The nested system formulation thus has a partially nested structure [6] and by Theorem 2 therein the optimal control law is linear.

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

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

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

B. A Summary Statistic to Compute The Optimal Control Law

We now show that the estimate of the current state given the history of observations available to the players is a summary statistic that characterizes the optimal control law. Denoting by $\mathcal{H}_i(t) = \{x_i(0 : t - 1), u_i(0 : t - 1)\}$, the history of observations for block $i$, the information available to player 1 at time $t$ is $\mathcal{F}_1(t) = \{x_1(t), \mathcal{H}_1(t)\}$, and that available to Player 2 is $\mathcal{F}_2(t) = \{x_2(t), \mathcal{H}_1(t), \mathcal{H}_2(t)\}$.

Define the estimators $\hat{x}(t) := E[x(t)|\mathcal{H}_1(t)]$ and $\hat{x}(t) := E[x(t)|\mathcal{H}_1(t), \mathcal{H}_2(t)]$ to be the estimates of the current states by Players 1 and 2 respectively, based on the history of information available to them. The dynamics of these estimators are derived in the following lemma.
Lemma 3. For the nested system (13), the estimator dynamics for \( \hat{x}(t) \) and \( \hat{x}(t) \) are given by

\[
\hat{x}(t + 1) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix} + R(t)\phi(t) \tag{17}
\]

and

\[
\hat{x}(t + 1) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \overline{R}(t)\overline{\phi}(t), \tag{18}
\]

with \( \hat{x}(0) = \hat{x}(0) = \mathbb{E}[x(0)] \), where \( \hat{u}_2(t) := \mathbb{E}[u_2(t)|\mathcal{H}_1(t)] \), \( R(t) := R_{xx}(t)R^{-1}_{x1}(t) \), \( \overline{R}(t) := \overline{R}_{xx}(t)\overline{R}^{-1}_{x1}(t) \), \( \phi(t) := x_1(t) - \hat{x}_1(t) \), \( \overline{\phi}(t) := x(t) - \hat{x}(t) \), \( R_{xx}(t) := AT(t) \), \( R_{x1}(t) = T_1(t) \), \( \overline{R}_{xx}(t) := \overline{A}\overline{T}(t) \), \( \overline{R}_{x1}(t) := \overline{T}(t) \), \( T(t) := \mathbb{E}[(x(t) - \hat{x}(t))(x_1(t) - \hat{x}_1(t))] \), and \( \overline{T}(t) := \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))] \).

Also, \( \phi(t) \) and \( \overline{\phi}(t) \) are zero-mean, uncorrelated Gaussian random vectors and \( \mathbb{E}[\phi(t)\phi(t)'] = T_1(t) \) and \( \mathbb{E}[\overline{\phi}(t)\overline{\phi}(t)'] = \overline{T}(t) \). Additionally, \( \mathbb{E}[\phi(t)x_1(\tau)'] = 0 \) and \( \mathbb{E}[\overline{\phi}(t)x(\tau)'] = 0 \) for \( \tau < t \).

Proof: The proof is in Appendix C.

We claim that \( \hat{x}_1(t) \) and \( \hat{x}_2(t) \) are indeed summary statistics for the history of state observations to determine the optimal control law. To do this, we first recall the optimal decentralized control law for the LQG problem (13)-(14) with no communication delay. For such a system, the
information structure is described by,

\[ z_1(t) = \{x_1(0 : t), u_1(0 : t - 1)\}, \]
\[ z_2(t) = \{x_1(0 : t), u_1(0 : t - 1); x_2(0 : t), u_2(0 : t - 1)\}, \quad (19) \]

Recall the following result from [10] for instantaneous communication.

**Proposition 3.** [10] Define estimator \( \hat{x}_2(t) = \mathbb{E}[x_2(t) | z_1(t)] \). The optimal decentralized control law for LQG problem \((13)-(14)\) with no communication delay is,

\[
\begin{bmatrix}
  u_1^*(t) \\
  u_2^*(t)
\end{bmatrix} = - \begin{bmatrix}
  K_{11}(t) & K_{12}(t) & 0 \\
  K_{21}(t) & K_{22}(t) & J(t)
\end{bmatrix} \begin{bmatrix}
  x_1(t) \\
  \hat{x}_2(t) \\
  x_2(t) - \hat{x}_2(t)
\end{bmatrix} \quad (20)
\]

and

\[
\hat{x}_2(t + 1) = A_{21}x_1(t) + A_{22}\hat{x}_2(t) + B_{21}u_1(t) + B_{22}\tilde{u}_2(t),
\]

where \( \tilde{u}_2(t) = \mathbb{E}[u_2(t) | z_1(t)] \). (The exact expressions for \( K_{ij}(t) \) and \( J(t) \) are given in [10].)

We can now establish the sufficiency of the estimators \( \hat{x}_1(t) \) and \( \hat{x}_2(t) \) for optimal control.

**Theorem 5.** For the nested system \((13)\) with the information structure \((15)\) and objective function \((14)\), the optimal control law is given by

\[
u(t) = F^*(t) \begin{bmatrix}
  x_1(t) - \hat{x}_1(t) \\
  x_2(t) - \hat{x}_2(t)
\end{bmatrix} - \tilde{K}(t), \quad (21)
\]

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

\[
\begin{bmatrix}
  K_{11}(t) & K_{12}(t) & 0 \\
  K_{21}(t) & K_{22}(t) & J(t)
\end{bmatrix} \begin{bmatrix}
  \hat{x}_1(t) \\
  \hat{x}_2(t) \\
  \hat{x}_2(t) - \hat{x}_2(t)
\end{bmatrix},
\]
where $K(t)$ and $J(t)$ are as in Proposition 3.

Remark: $K(t)$ and $J(t)$ are explicitly computed from (40)-(42) in the same way as in Proposition 3.

Proof: The proof is in Appendix D.

C. Deriving the Optimal Gain Matrix

It remains to characterize the optimal gain matrix, $F^*(t)$. As earlier, the stochastic optimization problem is converted to a deterministic convex optimization problem, after which standard matrix optimization tools can be applied to compute the optimal law analytically.

Theorem 6. The optimal gain matrix $F^*(t)$ is obtained by solving the deterministic convex optimization problem, $F^*(t) \in$

$$
\arg\min_{F(t)} \left[ \sum_{t=0}^{N-1} \text{trace}\left((\bar{Q} + H(t)'RH(t))\Sigma(t)\right) + \text{trace}\left(\bar{S}\Sigma(N)\right) + \sum_{t=0}^{N-1} \text{trace}\left(F(t)'RF(t)T(t)\right) \right],
$$

subject to the deterministic system dynamics

$$
\Sigma(t+1) = (\tilde{A} - G(t))\Sigma(t)(\tilde{A} - G(t))' + \sigma_F,
$$

$$
\Sigma(0) = 0.
$$

Here, $\bar{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}$, $\bar{S} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$, and $G(t) := \begin{bmatrix} BH(t) \\ 0 \\ 0 \\ B_{22}J(t) \end{bmatrix}$, where $H(t) = \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}$, $\sigma_F$ is the covariance matrix of $n(t)$, given by,

$$
n(t) = \begin{bmatrix} B_{11}F_{11}(t)\phi(t) + (R(t)\phi(t))_1 \\ B_{21}F_{11}(t)\phi(t) + B_{22}F_{22}(t)\phi_2(t) + (R(t)\phi(t))_2 \\ B_{21}F_{11}(t)\phi(t) + B_{22}F_{22}(t)\phi_2(t) + (R(t)\phi(t))_2 - (R(t)\phi(t))_2 \end{bmatrix}.
$$
Proof: The proof is in Appendix E.

The optimal gain matrix $F^\ast(t)$ may now be obtained using matrix optimization tools such as the discrete matrix minimum principle \cite{17} through the Hamiltonian to minimize (29). This completes the solution to the state-feedback problem for the $(1, \infty)$ information sharing pattern.

D. Optimal Control law for the Partial Output-feedback problem

The approach is shown to extend to a partial output-feedback problem where one of the blocks’ states is perfectly observed and the other gives a noisy observation. Before proceeding, it is useful to formulate the problem precisely.

The system dynamics, given that the output variables $y(t) := \begin{bmatrix} x_1(t) \\ y_2(t) \end{bmatrix}$ are observed, are governed by the following equations.

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

\[
\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} I & 0 \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ w(t) \end{bmatrix}.
\tag{24}
\]

Here, $v(t) := \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$ and $w(t)$ 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 a zero-mean, Gaussian random vector with known mean and covariance. For ease of notation, we will denote $y_1(t) := x_1(t)$.

The information structure for the problem is,

\[
\mathcal{F}_1(t) = \{y_1(0 : t), u_1(0 : t - 1)\},
\]

\[
\mathcal{F}_2(t) = \{y_1(0 : t - 1), u_1(0 : t - 1); y_2(0 : t), u_2(0 : t - 1)\},
\tag{25}
\]

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

Redefining the history of observations as, $\mathcal{H}_i(t) = (y_i(t - 1), \ldots, y_i(0), u_i(t - 1), \ldots, u_i(0))$
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 (14).

It can easily be seen from Proposition 2 that the optimal control law for the formulated partial output-feedback problem is linear.

Further, defining estimators $\hat{x}(t) := \mathbb{E}[x(t)|\mathcal{H}_1(t)]$ and $\hat{\hat{x}}(t) := \mathbb{E}[x(t)|\mathcal{H}_2(t)]$ as before, the following lemma can be reconstructed from Lemma 3. Note that the only significant differences are in the definition of $\phi(t), \overline{\phi}(t), R(t)$ and $\overline{R}(t)$.

**Lemma 4.** For the nested system (24), the estimator dynamics for the redefined $\hat{x}(t)$ and $\hat{\hat{x}}(t)$ are as follows.

\[
\begin{align*}
\hat{x}(t + 1) &= A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix} + R(t)\phi(t) \\
\hat{\hat{x}}(t + 1) &= A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \overline{R}(t)\overline{\phi}(t)
\end{align*}
\tag{26}
\]

and $\hat{x}(0) = \hat{\hat{x}}(0) = \mathbb{E}[x(0)]$, where $\hat{u}_2(t) = \mathbb{E}[u_2(t)|\mathcal{H}_1(t)]$. $R = R_{xy}R_{y1}^{-1}$, $\overline{R} = \overline{R}_{xy}\overline{R}_{y1}^{-1}$, and $\phi(t) = y_1(t) - \hat{x}_1(t), \overline{\phi}(t) = y(t) - C\hat{x}(t)$. $R_{xy} = AT(t)$, $R_{y1} = T_1(t) + W_{11}$, $\overline{R}_{xy} = AT(t)C'$, $\overline{R}_{y1} = CT(t)C' + W$. $T(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))]$, $\overline{T}(t) = \mathbb{E}[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))']$.

Also, $\phi(t)$ and $\overline{\phi}(t)$ are zero-mean, uncorrelated Gaussian random vectors. $\mathbb{E}[\phi(t)\phi(t)'] = T_1(t)$ and $\mathbb{E}[\overline{\phi}(t)\overline{\phi}(t)'] = C\overline{T}(t)C' + W$. Also, $\mathbb{E}[\phi(t)y_1(\tau)'] = 0$ and $\mathbb{E}[\overline{\phi}(t)y(\tau)'] = 0$ for $\tau < t$.

**Proof:*** The proof is same as that for Lemma 3.

Using the above tools, the following theorem without proof describes the structure of the optimal control law for the partial output-feedback problem.

**Theorem 7.** For the nested system (24) with the information structure (25) and objective
function (14), the optimal control law is given by

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

(28)

where \( F^*(t) = \begin{bmatrix} F^*_{11}(t) & 0 \\ 0 & F^*_{22}(t) \end{bmatrix} \) is the optimal gain matrix, and \( \tilde{K}(t) \) is given by

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

where \( K(t) \) and \( J(t) \) are as in Proposition 3.

**Proof:** The proof follows in the same manner as that for Theorem 5.

Finally, to complete the partial output-feedback problem, the optimal gain matrix \( F^*(t) \) is characterized by the following theorem.

**Theorem 8.** The optimal gain matrix \( F^*(t) \) is obtained by solving the deterministic convex optimization problem,

\[
    F^*(t) \in \arg \min_{F(t)} \left[ \sum_{t=0}^{N-1} \text{trace} \left( (\tilde{Q} + H(t)^T R H(t)) \Sigma(t) \right) + \text{trace} \left( \tilde{S} \Sigma(N) \right) + \sum_{t=0}^{N-1} \text{trace} \left( F(t)^T R F(t) T(t) \right) \right],
\]

(29)

subject to the deterministic system dynamics

\[
    \Sigma(t+1) = (\tilde{A} - G(t)) \Sigma(t) (\tilde{A} - G(t))^T + \sigma_F,
\]

\[
    \Sigma(0) = 0.
\]

Here, \( \tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \), \( \tilde{S} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \), and \( G(t) := \begin{bmatrix} BH(t) \\ 0 & 0 & B_{22} J(t) \end{bmatrix} \), where \( H(t) = \).
\[
\begin{bmatrix}
K_{11}(t) & K_{12}(t) & -K_{12}(t) \\
K_{21}(t) & K_{22}(t) & J(t) - K_{22}(t)
\end{bmatrix},
\]
\(\sigma_F\) is the covariance matrix of \(n(t)\), given by,
\[
n(t) = \begin{bmatrix}
B_{11}F_{11}(t)\phi(t) + (R(t)\phi(t))_1 \\
B_{21}F_{11}(t)\phi(t) + B_{22}F_{22}(t)\phi_2(t) + (R(t)\phi(t))_2 \\
B_{21}F_{11}(t)\phi(t) + B_{22}F_{22}(t)\phi_2(t) + (R(t)\phi(t))_2 - (R(t)\phi(t))_2
\end{bmatrix}.
\]

**Proof:** The proof is the same as that for Theorem 6.

As stated earlier, \(F^*(t)\) can be computed as in the state-feedback case using the Hamiltonian through the discrete maximum principle [17]. This concludes the partial output-feedback case for the \((1, \infty)\) information sharing pattern.

**IV. Conclusion**

In this paper, optimal decentralized control laws for two classes of delayed asymmetric information sharing patterns with finite time horizons were derived. An approach was developed to make the computation of the optimal law feasible through linearity. Subsequently, summary statistics to make the problem tractable were identified and a deterministic matrix optimization problem to compute the optimal control law was derived.

This approach was applied in two scenarios of two-player systems. The first involved an undelayed channel in one direction and a one unit timestep-delayed channel in the other. Optimal control laws were derived for both state and output-feedback cases. In the second scenario, the approach was applied to a one-way communication channel with a one unit delay. Optimal control laws were derived for state and partial output-feedback cases. The result could not be extended to the full output feedback case because the approach required player 2 to have the same history as player 1. If block 1’s observations are noisy, this assumption does not hold.

We believe that the approach of combining linearity and summary statistics in this manner can be used to compute optimal control laws for more general asymmetric information sharing scenarios and will be an interesting research direction for the future.
APPENDIX A

PROOF OF THEOREM 1 - FORM OF OPTIMAL CONTROL LAW FOR (1, 0) INFORMATION SHARING PATTERN

Proof: We begin by outlining our proof technique. The desired optimal control law is first shown to be the same as that for system (1) with the objective function $E\left[\sum_{t=0}^{N-1}(\dot{x}(t)'Q\dot{x}(t) + u(t)'Ru(t)) + \dot{x}(N)'S\dot{x}(N)\right]$. Thus, finding the optimal control law for the original problem is equivalent to finding the optimal law for the estimator dynamics (5) and the new objective.

Let us denote the estimation errors made by Players 1 and 2 by $e(t) := x(t) - \hat{x}(t)$. Then, $\phi(t) = \dot{x}(t)$ and by the projection theorem [16], $E[e(t)\dot{x}(t)'] = 0$.

Define a modified system input, $\overline{u}(t)$ by,

$$\overline{u}(t) := u(t) - F^*(t)\phi(t), \quad (31)$$

where $F^*(t)$ is the optimal gain matrix in (4). From (4), it can also be observed that the components of $\overline{u}(t)$ are functions of $H(t)$.

Using the modified input, (5) can be rewritten as,

$$\dot{x}(t+1) = A\dot{x}(t) + B\overline{u}(t) + \overline{v}(t), \quad (32)$$

where $\overline{v}(t) = (BF^*(t) + R(t))\phi(t)$.

It can be verified using Lemma 1 that $\overline{v}(t)$ is Gaussian, zero-mean and independent across time $t$. With the estimator system well-characterized, let us now turn our attention to the objective function. Rewriting the objective function (2) in terms of the estimator variables after noting that the terms involving $e(t)$ are either zero or independent of the input,

$$J = E\left[\sum_{t=0}^{N-1}(\dot{x}(t)'Q\dot{x}(t) + \overline{u}(t)'R\overline{u}(t)') + \dot{x}(N)'S\dot{x}(N) + \right.$$  
$$
\left.\sum_{t=0}^{N-1}(\phi(t)'F^*(t)'RF(t)\phi(t)) + 2\sum_{t=0}^{N-1}(\phi(t)'F^*(t)'R\overline{u}(t))\right].
$$

In the above equation, the term $\phi(t)'F^*(t)'RF(t)\phi(t)$ is independent of the input $\overline{u}(t)$. Also, $E[(\phi(t)'F^*(t)'R\overline{u}(t))] = 0 \forall t$, because $E[\overline{u}(t)] = 0$. 

21
Thus, when optimizing over $\tilde{u}(t)$, the following is an equivalent cost function,

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

(S3)

Solving our original optimal control problem is equivalent to solving the optimal control problem for the system (32) with objective function (33). Observe that this is just the classical LQR problem as there is no delay in the information communication, hence making the system ‘centralized’. Thus, the optimal controller is given by,

$$
\tilde{u}^*(t) = -H^*(t)\hat{x}(t),
$$

(34)

where $H^*(t)$ is the optimal gain matrix for the classical LQG problem [16].

Therefore, from (34) and (31), the optimal decentralized control law $u(t)$ is obtained as,

$$
u(t) = F^*(t)x(t) + G^*(t)\hat{x}(t),
$$

(35)

where $F^*(t) = \begin{bmatrix} F^*_{11}(t) & F^*_{12}(t) \\ 0 & F^*_{22}(t) \end{bmatrix}$ and $G^*(t) = -(F^*(t) + H^*(t)).$}

**APPENDIX B**

**Proof of Theorem 2 - Optimal Gain Matrix for (1, 0) Information Sharing Pattern**

*Proof:* Recall that we need to solve for the optimal control law to the modified system (32) with the cost function defined by (33). Assuming that the optimal gain matrix $F^*(t)$ is unknown, let us replace it by an arbitrary matrix $F(t)$ and find the deterministic optimization problem that has $F^*(t)$ as its solution.

Using the modified input relation $\tilde{u}(t) := -H^*(t)\hat{x}(t)$, we can write (32) as,

$$
\bar{x}(t+1) = (A - BH^*(t))\hat{x}(t) + (R(t) + BF(t))\phi(t).
$$

(36)

Observe that the variance of ‘noise’ $(R(t)+BF(t))\phi(t)$ is $(BF(t)+R(t))T(t)(BF(t)+R(t))^\dagger$.

Considering the cost function (33) and retaining only those terms in it that depend on $F(t)$,
the relevant part for our optimization problem is,

\[ J_F = \mathbb{E} \left[ \sum_{t=0}^{N-1} \left( \hat{x}(t)'Q\hat{x}(t) + \hat{x}(t)'H^*(t)'RH^*(t)\hat{x}(t) \\
+ \phi(t)'F(t)'RF(t)\phi(t) \right) + \hat{x}(N)'S\hat{x}(N) \right]. \] (37)

Defining \( \Sigma(t) := \mathbb{E}[\hat{x}(t)\hat{x}(t)'] \), the dynamics of \( \Sigma(t) \) are obtained from (36) as,

\[ \Sigma(t+1) = (A - BH^*(t))\Sigma(t)(A - BH^*(t))' + (BF(t) + R(t))T(t)(BF(t) + R(t))', \]

\[ \Sigma(0) = 0. \]

Using matrix algebra, (37) simplifies to

\[ J_F = \sum_{t=0}^{N-1} \text{trace}\left( (Q + H^*(t)'RH^*(t))\Sigma(t) \right) \\
+ \text{trace}\left( S\Sigma(N) \right) + \sum_{t=0}^{N-1} \text{trace}\left( F(t)'RF(t)T(t) \right). \]

\[ \blacksquare \]

APPENDIX C

PROOF OF LEMMA - DYNAMICS OF ESTIMATORS FOR \( (1, \infty) \) INFORMATION SHARING PATTERN

Proof: Let us first derive the dynamics of block 1. Denote \( \phi(t) := x_1(t) - \hat{x}_1(t) \). Then, \( \phi(t) \) and \( H_1(t) \) are independent by the projection theorem for Gaussian random variables [16] and,

\[ \hat{x}(t+1) = \mathbb{E}[x(t+1)|H_1(t+1)] \\
= \mathbb{E}[x(t+1)|H_1(t), x_1(t), u_1(t)] \\
= \mathbb{E}[x(t+1)|H_1(t), u_1(t)] + \mathbb{E}[x(t+1)|\phi(t)] - \mathbb{E}[x(t+1)], \]

23
where the last equality follows from the independence of $\phi(t)$ and $H_1(t)$ \[16\]. The first term on the RHS above, $E[x(t + 1)|H_1(t), u_1(t)]$

$$= E[Ax(t) + Bu(t)|H_1(t), u_1(t)]$$

$$= A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \hat{u}_2(t) \end{bmatrix},$$

where $\hat{u}_2(t) = E[u_2(t)|H_1(t)].$

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

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

where defining $T(t) = E[(x(t) - \hat{x}(t))(x_1(t) - \hat{x}_1(t))]$ and partitioning $T(t)$ as $T(t) = [T_1(t)|T_2(t)],$ we have $R_{xx_1} = AT(t), R_{x_1} = T_1(t)$ \[16\].

Putting these three terms together, we have

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

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

$$\phi(t) = x(t) - \hat{x}(t) = A\hat{x}(t) + B \begin{bmatrix} u_1(t) \\ \phi_2(t) \end{bmatrix} + R(t)\phi(t).$$

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

The projection theorem for Gaussian random variables \[16\] implies the uncorrelatedness of $\phi(t)$ and $x_1(\tau)$ for $\tau < t,$ and also implies that $\overline{\phi}(t)$ is uncorrelated with $x(\tau)$ for $\tau < t.$ □
APPENDIX D

PROOF OF THEOREM[5] - FORM OF OPTIMAL CONTROL LAW FOR \((1, \infty)\) INFORMATION SHARING PATTERN

Proof: We begin by outlining our proof technique. The desired optimal control law is first shown to be the same as that for system \((13)\) with the objective function \(E[\sum_{t=0}^{N-1} (\bar{x}(t)'Q\bar{x}(t) + \bar{u}(t)'R\bar{u}(t)) + \bar{x}(N)'S\bar{x}(N)]\), where \(\bar{x}(t) = [\hat{x}_1(t) \hat{x}_2(t)]\) and \(\bar{u}(t)\) is a suitably defined input variable. Thus, finding the optimal control law for the original problem is equivalent to finding the optimal control law for the estimator dynamics system \((17)-(18)\) and the new objective function. Then, the proof argument follows that of Proposition 3.

The following lemma derives a required uncorrelatedness property and also proves the equivalence of two of the estimator variables.

Lemma 5. For the nested system \((13)\), \(\hat{x}_1(t) = \hat{x}_1(t)\). Additionally, \(\phi(t) = \bar{\phi}_1(t)\), which implies from Lemma 3 that \(E[\phi(t)x(\tau)'] = 0, \tau < t\).

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

\[
\hat{x}_1(t) = E[x_1(t)|\mathcal{H}_1(t), \mathcal{H}_2(t)],
\]
\[
= A_{11}x_1(t - 1) + B_{11}u_1(t - 1),
\]
\[
= \hat{x}_1(t),
\]

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

\[
\begin{bmatrix}
\hat{x}_1(t+1) \\
\hat{x}_2(t+1)
\end{bmatrix} =
\begin{bmatrix}
A_{11} & 0 \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
\hat{x}_1(t) \\
\hat{x}_2(t)
\end{bmatrix} +
\begin{bmatrix}
B_{11} & 0 \\
B_{21} & B_{22}
\end{bmatrix}
\begin{bmatrix}
u_1(t) \\
u_2(t)
\end{bmatrix} +
\begin{bmatrix}
(R(t)\phi(t))_1 \\
(T(t)\bar{\phi}(t))_2
\end{bmatrix}.
\] (38)

This shall be referred to in shorthand notation as

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

Let us denote estimation error by \(e(t) = [e_1(t) e_2(t)] := [x_1(t) - \hat{x}_1(t) \ x_2(t) - \hat{x}_2(t)]\) respectively. Then, \(e_1(t) = \phi_1(t)\) and \(e_2(t) = \bar{\phi}_2(t)\). By the projection theorem \[16\] and some manipulations, it can be shown that \(\mathbb{E}[e_i(t)\bar{x}_j(t)'] = 0\) for \(i, j = 1, 2\).

Further, define a modified system input \(\bar{u}(t)\) by

\[
\bar{u}(t) =
\begin{bmatrix}
\bar{u}_1(t) \\
\bar{u}_2(t)
\end{bmatrix} := u(t) - F^*(t)e(t),
\] (39)

where \(F^*(t)\) is the optimal gain matrix in \[16\]. From \[16\], it can also be observed that two components of \(\bar{u}(t)\) are functions of \(H_1(t)\) and \((H_1(t), H_2(t))\) respectively.

Using the modified input, (38) can be rewritten as

\[
\bar{x}(t+1) = A\bar{x}(t) + B\bar{u}(t) + \bar{v}(t),
\] (40)

where \(\bar{v}(t)\) is equal to

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

It can be verified using Lemmas \[3\] and \[5\] that \(\bar{v}(t)\) is Gaussian, zero-mean and independent
across time $t$. With the estimator system well-characterized, let us now turn our attention to the objective function. Rewriting the objective function (14) in terms of the estimator variables $[\hat{x}_1(t) \ \hat{x}_2(t)]$ and the modified 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,

$$J = \mathbb{E}\left[ \sum_{t=0}^{N-1} (\bar{x}(t)'Q\bar{x}(t) + \bar{u}(t)'R\bar{u}(t) + e(t)'F^*(t)RF^*(t)e(t) + 2e(t)'F^*(t)R\bar{u}(t)) + \bar{x}(N)'S\bar{x}(N) \right].$$

(41)

In the above equation, the term $e(t)'F^*(t)RF^*(t)e(t)$ is independent of the control input $u(t)$. Additionally, $\bar{u}_1(t)$ and $\bar{u}_2(t)$ are linear functions of $\mathcal{H}_1(t)$ and $(\mathcal{H}_1(t), \mathcal{H}_2(t))$ respectively. The fourth term thus has zero expected value. Hence, when optimizing over $\bar{u}(t)$, the following is an equivalent cost function,

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

(42)

Solving our original optimal control problem is equivalent to solving the cost function (42) for the system (40) with no communication delay in the new state variables.

Letting $[\bar{u}_1(t) \ \bar{u}_2(t)]'$ take up the role of players’ inputs, observe that (40) has a nested system structure with no communication delays. Thus, by Proposition 3, the optimal controller for the system (40)-(42) is given by,

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

(43)

where $\hat{x}_2(t) = \mathbb{E}[\hat{x}_2(t)|\bar{x}_1(0 : t), \bar{u}_1(0 : t - 1)]$ and the exact expressions for $K(t)$ and $J(t)$ can
be found in [10]. As shown below, \( \hat{x}_2(t) \) is just \( \hat{x}_2(t) \).

\[
\hat{x}_2(t) = \mathbb{E}[x_2(t)|H_1(t)],
\]
\[
= \mathbb{E}[\mathbb{E}[x_2(t)|H_1(t)]|x_1(0:t), \bar{u}_1(0:t-1)],
\]
\[
= \mathbb{E}[x_2(t)|x_1(0:t), \bar{u}_1(0:t-1)],
\]
\[
= \mathbb{E}[\mathbb{E}[x_2(t)|H_1(t), H_2(t)]|x_1(0:t), \bar{u}_1(0:t-1)],
\]
\[
= \hat{x}_2(t),
\]

where the second equality just takes the conditional expectation of \( \hat{x}_2(t) \) with itself and other random variables. The third and fourth equalities follow from the tower rule.

Therefore, from (43) and (39), the optimal decentralized control law \( u(t) \) is obtained as,

\[
u(t) = F^*(t) \begin{bmatrix} x_1(t) - \hat{x}_1(t) \\ x_2(t) - \hat{x}_2(t) \end{bmatrix} - K(t).
\] (44)

\[\]

**APPENDIX E**

**PROOF OF THEOREM 6 - OPTIMAL GAIN MATRIX FOR \((1, \infty)\) INFORMATION SHARING PATTERN**

**Proof:** Recall that we need to solve for the optimal control law to the modified system (40) with the cost function defined by (42). Assuming that the optimal gain matrix \( F^*(t) \) is unknown, let us replace it by an arbitrary matrix \( F(t) \) and find out the deterministic optimization problem that has \( F^*(t) \) as its solution.

From Theorem 5, using the modified input \( \bar{u}(t) := u(t) - \begin{bmatrix} F_{11}(t) \phi(t) \\ F_{22}(t) \phi_2(t) \end{bmatrix} \), we have,

\[
\bar{u}(t) = -\begin{bmatrix} K_{11}(t) & K_{12}(t) & 0 \\ K_{21}(t) & K_{22}(t) & J(t) \end{bmatrix} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_2(t) - \hat{x}_2(t) \end{bmatrix},
\] (45)
which can be rewritten as,
\[ \pi(t) = -H(t)\hat{x}(t), \]  
(46)

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

The dynamics of \( \hat{x}(t) \) are obtained from \((18)\) and \((40)\) as,
\[ \hat{x}(t + 1) = (\tilde{A} - G(t))\hat{x}(t) + n(t), \]  
(47)

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

For brevity, let \( \sigma_F \) denote the variance of \( n(t) \). Since \( F(t) \) minimizes \((41)\), keeping only the terms that depend on \( F(t) \), the relevant part of the optimization problem is,
\[ J_F = \mathbb{E}\left[ N - 1 \sum_{t=0}^{N-1} \begin{bmatrix} \tilde{x}(t)'\tilde{Q}\tilde{x}(t) + \tilde{x}(t)'H(t)'RH(t)\tilde{x}(t) + \\
\bar{\phi}(t)'F(t)'RF(t)\bar{\phi}(t) + \tilde{x}(N)\tilde{S}\tilde{x}(N) \end{bmatrix} \right], \]  
(48)

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, \((48)\) simplifies to
\[ J_F = \sum_{t=0}^{N-1} \text{trace}\left( (\tilde{Q} + H(t)'RH(t))\Sigma(t) \right) + \text{trace}\left( \tilde{S}\Sigma(N) \right) + \sum_{t=0}^{N-1} \text{trace}\left( F(t)'RF(t)\overline{T}(t) \right), \]  
(49)

where we have defined \( \Sigma(t) := \mathbb{E}[\pi(t)\pi(t)'] \).

Finally, the dynamics of \( \Sigma(t) \) are obtained from \((47)\) as,
\[ \Sigma(t + 1) = (\tilde{A} - G(t))\Sigma(t)(\tilde{A} - G(t))' + \sigma_F, \]  
\[ \Sigma(0) = \mathbb{E}[\pi(0)\pi(0)'] = 0. \]  
(50)
where $\sigma_F$ is the covariance matrix of $n(t)$.

REFERENCES


