Optimal decentralized control of two agent linear system with partial output feedback: certainty equivalence and optimality of linear strategies

Mohammad Afshari and Aditya Mahajan

Department of Electrical and Computer Engineering, McGill University (email: mohammad.afshari2@mail.mcgill.ca, aditya.mahajan@mcgill.ca)

Abstract: We consider the optimal decentralized control of two agent linear system in which the agent 2's state is affected by the state and control actions of agent 1 but not vice versa. The state of agent 1 is perfectly observed by both agents while the state of agent 2 is observed with noise by agent 2. Thus, the information structure is partially nested. However, we *do not* assume that the process and observation noises are Gaussian. Therefore, it is not known a priori whether linear strategies are globally optimal. Without using dynamic programming, we show that the optimal strategy is linear and certainty equivalent. Our proof is based on three steps: static reduction via state splitting, orthogonal projection, and completion of squares. We believe that our solution methodology is of independent interest for decentralized control of linear systems.

Keywords: Decentralized stochastic control, team theory, non-classical information structure, certainty equivalence.

1. INTRODUCTION

In centralized stochastic control of linear systems with quadratic per-stage cost, the optimal control strategy is a linear function of the controller's estimate of the state of the system. In addition, there is a two-way separation between estimation and control: the optimal gain is the same as in the case of state feedback; the optimal state estimate is the optimal estimate of the uncontrolled system shifted by an amount that depends on the past control actions. Thus, the optimal control strategy does not depend on the estimation strategy and the estimation strategy does not depend on the control strategy. Although the separation result is typically stated under the assumption that the noise is Gaussian (such systems are typically referred to as LQG (linear quadratic and Gaussian) systems), it holds more generally as long as the noise has finite second moment (Wonham, 1968; Root, 1969; Bertsekas, 2000).

The situation in decentralized stochastic control is significantly different. Even for LQG systems, linear strategies are not globally optimal (Witsenhausen, 1968). Linear strategies are optimal if the information structure is partially nested *and the primitive random variables are Gaussian* (Ho and Chu, 1972). If attention is restricted to linear strategies, the problem of finding the best linear strategy need not be convex. It is convex if the system is quadratic invariant (Rotkowitz and Lall, 2006). Based on these characterizations, the literature on decentralized control of linear systems is broadly split into two parts: (i) models with partially nested information structure where the noise is assumed to be Gaussian and the structure of *optimal linear* strategy is established; and (ii) quadratic invariant models where no restriction is imposed on the noise process and the structure of the *best linear* [1] strategy is established. In the interest of space, we omit a detailed literature overview and refer the reader to Mahajan et al. (2012) for an overview. To the best of our knowledge, the problem of identifying the *optimal* strategy when the noise is not Gaussian has not been investigated. Note that, a priori, it is not obvious that linear strategies are optimal in such a setup.

Two main proof techniques are used to establish the structure of optimal strategies: (i) time domain decomposition using dynamic programming for partially nested systems and (ii) frequency domain Youla parametrization for quadratic invariant systems. Both of these techniques are difficult to use to identify *optimal* control strategies when the noise is not Gaussian. Since it is not known that linear strategies are optimal, frequency domain techniques cannot be used. Dynamic programming techniques only work for partial history sharing information structures (Nayyar et al., 2013) and even in those cases, the sufficient statistics are distribution valued and at each step of the dynamic program, one needs to search over all non-linear prescriptions.

In this paper, we revisit the so called two player problem with partial output feedback but do not assume that the noise is Gaussian. Variants of this model have been considered in Swigart and Lall (2010, 2011); Kim and Lall (2011, 2012); Lessard and Lall (2012, 2015); Lessard and

[1] The distinction between "optimal linear" and the "best linear" strategies is that in the former case, restricting attention to linear strategies is without loss of optimality while in the latter case the restriction to linear strategies is arbitrary and may lead to loss of optimality.

Nayyar (2013); Nayyar et al. (2018). The model is partially nested and quadratic invariant; all of the previous papers assume that the noise is Gaussian (and hence, the optimal strategy is known to be linear). We do not impose any restriction on the noise or on the class of control strategies.

The main result of this paper is to show that even when the noise is not Gaussian, the optimal control strategy retains the main feature of the Gaussian noise case. In particular, there is a two way separation between estimation and control. On the one hand, the optimal control strategy is a linear function of the state estimates given the common and the local information; the corresponding gains are obtained by solving two standard Riccati equations. On the other hand, the optimal state estimates are given by the estimates of an uncontrolled system shifted by amounts that depend on the past control actions.

It is worth highlighting that since the noise is not Gaussian, the state estimate need not be a linear function of the observations. Thus, although the optimal control action is a linear function of the conditional estimate, it need not be a linear function of the observations. Furthermore, it may not be possible to recursively compute the state estimates and we need to recursively keep track of the conditional distributions. Thus, although we have shown that the state estimate is a sufficient statistics for control, it is not an information state for dynamic programming.

Our proof is based on a new solution methodology that consists of three steps: static reduction via state splitting, completion of squares, and orthogonal projection. We believe that our solution methodology is of independent interest for decentralized control of linear systems.

1.1 Notations

Given a matrix A, A_{ij} denotes its (i,j)-th block element, $A^\intercal$ denotes its transpose, $\mathrm{vec}(A)$ denotes the column vector of A formed by vertically stacking the columns of A. Given a square matrix A, $\mathrm{Tr}(A)$ denotes the sum of its diagonal elements. I_n denotes an $n \times n$ identity matrix. We simply use I when the dimension is clear for context.

Given any vector valued process $\{y(t)\}_{t\geq 1}$ and any time instances t_1, t_2 such that $t_1 \leq t_2$, $y(t_1{:}t_2)$ is a short hand notation for $\mathrm{vec}(y(t_1), y(t_1+1), \dots, y(t_2))$.

Given random vectors x and y, $\mathbb{E}[x]$ denotes the mean of x and $\mathbb{E}[x|y]$ denotes the conditional mean of random variable x given random variable y.

Superscript index agents and local, common, and stochastic components of state and control. Subscripts denote components of vectors and matrices. The notation $\hat{x}(t|i)$ denotes the estimate of variable x at time t conditioned on the info at agent i at time t.

2. PROBLEM FORMULATION AND MAIN RESULT

2.1 Problem formulation

Consider a decentralized control system with two coupled agents. For $i \in \{1,2\}$, let $x_i(t) \in \mathbb{R}^{d_x^i}$ and $u_i(t) \in \mathbb{R}^{d_u^i}$ denote the state and control action of agent i. The initial state of the system is $(x_1(1), x_2(1))$ and for $t \geq 1$, the dynamics are given by

$$x(t+1) = Ax(t) + Bu(t) + w(t), \tag{1}$$

where $x(t) = \mathrm{vec}(x_1(t), x_2(t))$ denotes the state of the system, $u(t) = \mathrm{vec}(u_1(t), u_2(t))$ denotes the control action of both controllers, $w(t) = \mathrm{vec}(w_1(t), w_2(t))$ is the system disturbance, and the system matrices are given by

$$A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix}.$$

The A and B matrices are block lower triangular. Thus, agent 1 affects the dynamics of agent 2 but not vice-versa. We may think of agent 1 as a major agent and agent 2 as a minor agent.

The system has partial output feedback. In particular, the state of agent 1 is perfectly observed by both agents while the state of agent 2 is observed by agent 2 with noise (but not observed by agent 1). In particular, the observation of agent 2 is

$$y_2(t) = C_{21}x_1(t) + C_{22}x_2(t) + v_2(t). \tag{2}$$

Thus, the information $I^i(t)$ available to agent i, $i \in \{1,2\}$, is given by

$$I^1(t) \coloneqq \{x_1(1{:}t), u_1(1{:}t-1)\}, \tag{3}$$

$$I^2(t) \coloneqq \{x_1(1{:}t), y_2(1{:}t), u_1(1{:}t-1), u_2(1{:}t-1)\}. \tag{4}$$

Furthermore, agent i chooses its control action according to

$$u_i(t) = g_{i,t}(I_i(t)),$$

where $g_i \coloneqq (g_{i,1}, \dots g_{i,T})$ is called the control strategy of agent i.

The primitive random variables $\{x_1(1), x_2(1), w_1(1:T), w_2(1:T), v_2(1:T)\}$ are independent and have zero mean and finite variance. Note that we do not assume that the primitive random variables are Gaussian.

The system incurs a per-step cost of

$$c(x(t), u(t)) = x(t)^\intercal Q x(t) + u(t)^\intercal R u(t), \tag{5}$$

at time $t \in \{1, \dots, T-1\}$ and a terminal cost of

$$C(x(T)) = x^\intercal(T) Q_T x(T). \tag{6}$$

at time T. We assume that Q and Q_T are positive semidefinite and R is positive definite.

The performance of any strategy (g_1, g_2) is given by

$$J(g_1, g_2) = \mathbb{E}\Bigg[\sum_{t=1}^{T} c(x(t), u(t)) + C(x(T))\Bigg]. \tag{7}$$

We are interested in the following optimization problem.

Problem 1. Given horizon T, the system dynamic matrices A and B, the cost matrices Q, Q_T, and R, and the covariance matrices of the primitive random variables, choose a strategy (g_1, g_2) to minimize the total expected cost given by (7).

Remark 1. The system described above is partially nested (Ho and Chu, 1972). Thus, if the primitive random variables are jointly Gaussian, then linear strategies are optimal. The system described above is also quadratic invariant (Rotkowitz and Lall, 2006). Thus, if we arbitrarily restrict attention to linear strategies, then the problem of finding the best linear strategy is convex. However, there is no general methodology to identify sufficient statistics for partially nested or quadratic invariant problems. In addition, it is not known if linear strategies are globally optimal when the primitive random variables are not Gaussian.

2.2 Main result

Following Nayyar et al. (2013), we split the information at each agent into common information and local information. The common information is defined as the information commonly known to both agents, i.e.,

$$I^{\mathrm{c}}(t) \coloneqq I^1(t) \cap I^2(t) = \{x_1(1:t), u_1(1:t-1)\} = I^1(t).$$

The local information is the remaining information at each agent. Thus,

$$I^{1,\ell}(t) \coloneqq I^1(t) \setminus I^{\mathrm{c}}(t) = \emptyset,$$
$$I^{2,\ell}(t) \coloneqq I^2(t) \setminus I^{\mathrm{c}}(t) = \{y_2(1:t), u_2(1:t-1)\}.$$

Now, we split the control action into two parts: $u(t) = u^{\mathrm{c}}(t) + u^{\ell}(t)$, where

$$u^{\mathrm{c}}(t) = \mathbb{E}[u(t)|I^{\mathrm{c}}(t)],$$
$$u^{\ell}(t) = u(t) - u^{\mathrm{c}}(t).$$

We refer to $u^{\mathrm{c}}(t)$ and $u^{\ell}(t)$ as the common control and the local control, respectively. We have the following properties.

(H1) $u_1^{\ell}(t) = 0$.

(H2) $\mathbb{E}[u_2^{\ell}(t)|I^{\mathrm{c}}(t)] = 0$.

(H3) $\mathbb{E}[u^{\mathrm{c}}(t)^{\intercal} u^{\ell}(t)] = 0$.

Both (H1) and (H2) follow from the definition of $u^{\ell}(t)$ and (H3) follows because mean square error is orthogonal to the estimate.

Our main result is the following.

Theorem 1. Given the split of control action $u(t)$ as $u^{\mathrm{c}}(t)$ and $u^{\ell}(t) = \mathrm{vec}(0, u_2^{\ell}(t))$, the optimal control actions are

$$u^{\mathrm{c}}(t) = -L^{\mathrm{c}}(t)\hat{x}(t|\mathrm{c}),$$
$$u_2^{\ell}(t) = -L^{\ell}(t)(\hat{x}_2(t|2) - \hat{x}_2(t|\mathrm{c})),$$

where

$$\hat{x}(t|\mathrm{c}) = \mathbb{E}[x(t)|I^{\mathrm{c}}(t)],$$
$$\hat{x}(t|2) = \mathbb{E}[x(t)|I^2(t)].$$

Furthermore, the optimal gains are given by

$$L^{\mathrm{c}}(t) = [R + B^{\intercal} S^{\mathrm{c}}(t+1)B]^{-1} B^{\intercal} S^{\mathrm{c}}(t+1)A,$$
$$L^{\ell}(t) = [R_{22} + B_{22}^{\intercal} S^{\ell}(t+1)B_{22}]^{-1} B_{22}^{\intercal} S^{\ell}(t+1)A_{22},$$

where $S^{\mathrm{c}}(1{:}T)$ and $S^{\ell}(1{:}T)$ are the solution of the following standard Riccati equations: $S^{\mathrm{c}}(T) = Q_T$, $S^{\ell}(T) = [Q_T]_{22}$, and for $t \in \{T-1, \dots, 1\}$,

$$S^{\mathrm{c}}(t) = Q + A^{\intercal} S^{\mathrm{c}}(t+1)A - L^{\mathrm{c}}(t)^{\intercal}[R + B^{\intercal} S(t+1)B]L^{\mathrm{c}}(t),$$

and

$$S^{\ell}(t) = Q_{22} + A_{22}^{\intercal} S^{\ell}(t+1)A_{22} - L^{\ell}(t)^{\intercal}[R_{22} + B_{22}^{\intercal} S^{\ell}(t+1)B_{22}]L^{\ell}(t).$$

3. PROOF OF THE MAIN RESULT

The proof consists of three steps. In Step 1, we split the state into controlled and uncontrolled components and use that to provide a static reduction of the information structure. In Step 2, we use completion of squares to write an alternative expression for the total cost. In Step 3, we further simplify the expression of total cost using orthogonal projections. The final expression is such that the form of the optimal controller can be identified by inspection.

Step 1: Static reduction via state splitting

We split the state into three components: $x^{\mathrm{c}}(t)$, $x^{\ell}(t)$, $x^{\mathrm{s}}(t)$ (called the common, local, and stochastic components, respectively) as follows.

$$x^{\mathrm{c}}(1) = 0, \qquad x^{\mathrm{c}}(t+1) = Ax^{\mathrm{c}}(t) + Bu^{\mathrm{c}}(t),$$
$$x^{\ell}(1) = 0, \qquad x^{\ell}(t+1) = Ax^{\ell}(t) + Bu^{\ell}(t),$$
$$x^{\mathrm{s}}(1) = x(1), \qquad x^{\mathrm{s}}(t+1) = Ax^{\mathrm{s}}(t) + w(t).$$

Due to the linearity of dynamics, we have

$$x(t) = x^{\mathrm{c}}(t) + x^{\ell}(t) + x^{\mathrm{s}}(t).$$

Now, define

$$y_2^{\mathrm{s}}(t) = C_{21}x_1^{\mathrm{s}}(t) + C_{22}x_2^{\mathrm{s}}(t) + v_2(t)$$

and consider the information structure

$$I^{1,s}(t) = \{x_1^{\mathrm{s}}(1:t)\},$$
$$I^{2,s}(t) = \{y_2^{\mathrm{s}}(1:t)\}.$$

Lemma 2. For any arbitrary but fixed control strategy (g_1, g_2), $I^1(t) \equiv I^{1,s}(t)$ and $I^2(t) \equiv I^{2,s}(t)$ (i.e., both sets generate the same σ-field or, equivalently, they are functions of each other.

Proof. The proof is omitted due to space limitations but is similar in spirit to similar results for centralized stochastic control. For example, see Kumar and Varaiya (1986, Chapter 7.3). ☐

Thus, we can assume that agents choose their control actions based on $I^{i,\mathrm{s}}(t)$ rather than $I^i(t)$. Note that $I^{i,\mathrm{s}}(t)$ does not depend on any of the previous control actions. Therefore, following Witsenhausen (1988), we call $\{I^{1,s}(t), I^{2,s}(t)\}$ to be the *static reduction* of $\{I^1(t), I^2(t)\}$. We exploit the static reduction to simplify the expressions for conditional expectations.

Some preliminary properties

Now we establish some preliminary properties of the different components of the state and the control that follow from state splitting and static reduction.

Lemma 3. The following properties hold.

(P1) $x_1^{\ell}(t) = 0$.

(P2) $\mathbb{E}[u_2^{\ell}(t)] = 0$.

(P3) For any $\tau \le t$, $\mathbb{E}[u_2^{\ell}(\tau)|I^{\mathrm{c}}(t)] = 0$.

(P4) $\mathbb{E}[x_2^{\ell}(t)|I^{\mathrm{c}}(t)] = 0$.

(P5) $\mathbb{E}[x^{\mathrm{c}}(t)|I^{\mathrm{c}}(t)] = x^{\mathrm{c}}(t)$.

(P6) For any matrix M of compatible dimensions, $\mathbb{E}[(x_2^{\ell}(t))^{\intercal} M x_1^{\mathrm{s}}(t)] = 0$.

Proof. We prove each property separately.

(P1) This is an immediate consequence of the fact that A and B are block lower triangular, $x^{\ell}(1) = 0$, and $u_1^{\ell}(t) = 0$.

(P2) By the smoothing property of conditional expectation, we have

$$\mathbb{E}[u_2^\ell(t)] = \mathbb{E}\big[\mathbb{E}[u_2^\ell(t)|I^c(t)]\big] = 0$$

where the last equality follows from (H2).

(P3) By (H2), we have that $\mathbb{E}[u^\ell(t)|I^c(t)] = 0$. Now consider $\tau < t$. By Lemma 2,

$$\mathbb{E}[u_2^\ell(t)|I^c(t)] = \mathbb{E}[u_2^\ell(t)|I^{1,s}(t)].$$

Now observe that,

$$\begin{aligned} I^{1,\mathrm{s}}(t) = \{x_1^\mathrm{s}(1{:}t)\} &\equiv \{x_1^\mathrm{s}(1{:}\tau), w_1(\tau{:}t-1)\} \\ &= \{I^{1,\mathrm{s}}(\tau), w_1(\tau{:}t-1)\} \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}[u_2^\ell(t)|I^{1,s}(t)] &= \mathbb{E}[u_2^\ell(\tau)|I^{1,\mathrm{s}}(\tau), w_1(\tau{:}t-1)] \\ &\overset{(a)}{=} \mathbb{E}[u_2^\ell(\tau)|I^{1,\mathrm{s}}(\tau)] \\ &\overset{(b)}{=} \mathbb{E}[u_2^\ell(\tau)|I^1(\tau)] \\ &\overset{(c)}{=} 0 \end{aligned}$$

where (a) holds because $u_2^\ell(\tau)$ is independent of future noise $w_1(\tau{:}t-1)$, (b) uses Lemma 2, and (c) uses (H2).

(P4) Since $x^\ell(1) = 0$, from (H1) and (P1) we have that

$$x_2^\ell(t) = \sum_{\tau=1}^{t-1} A_{22}^{\tau-1} B_{22} u_2^\ell(t-\tau).$$

Hence, the result follows from (P3).

(P5) By definition, $u^\mathrm{c}(1{:}t)$ is a function of $I^\mathrm{c}(t)$, and therefore, $x^\mathrm{c}(t)$ is a function of $I^\mathrm{c}(t)$.

(P6) By the smoothing property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}[(x_2^\ell(t))^\intercal Q_{21} x_1^\mathrm{s}(t)] &= \mathbb{E}\big[\mathbb{E}[(x_2^\ell(t))^\intercal Q_{21} x_1^\mathrm{s}(t)|I^\mathrm{c}(t)]\big] \\ &\overset{(a)}{=} \mathbb{E}\big[\mathbb{E}[(x_2^\ell(t))^\intercal|I^\mathrm{c}(t)]\, Q_{21} x_1^\mathrm{s}(t)\big] \\ &\overset{(b)}{=} 0, \end{aligned}$$

where (a) follows because $x_1^\mathrm{s}(t)$ is part of $I^\mathrm{c}(t)$ and (b) follows from (P4). ☐

For ease of notation, we consider the following combinations of different components of the state:

$$\begin{aligned} z^\mathrm{c}(t) &= x^\mathrm{c}(t) + x^\mathrm{s}(t), \\ z_2^\ell(t) &= x_2^\ell(t) + x_2^\mathrm{s}(t). \end{aligned}$$

Based on the properties of Lemma 3, we show that the per-step cost can be split as follows.

Lemma 4. The following properties hold:

(1) $\mathbb{E}\big[u(t)^\intercal R u(t)\big] = \mathbb{E}\big[u^\mathrm{c}(t)^\intercal R u^\mathrm{c}(t) + u_2^\ell(t)^\intercal R_{22} u_2^\ell(t)\big]$.

(2) $\mathbb{E}\big[x(t)^\intercal Q x(t)\big] = \mathbb{E}\big[(z^\mathrm{c}(t))^\intercal Q (z^\mathrm{c}(t)) + (z^\ell(t))^\intercal Q_{22} (z^\ell(t)) - x_2^\mathrm{s}(t) Q_{22} x_2^\mathrm{s}(t)\big]$.

Proof. We prove the two parts separately.

(1) From smoothing property of conditional expectation and (H3), we get

$$\mathbb{E}\big[u(t)^\intercal R u(t)\big] = \mathbb{E}\big[u^\mathrm{c}(t)^\intercal R u^\mathrm{c}(t) + u^\ell(t)^\intercal R u^\ell(t)\big].$$

The result then follows from (H1).

(2) Since $x(t) = x^\mathrm{c}(t) + x^\ell(t) + x^\mathrm{s}(t)$, we can write

$$\begin{aligned} \mathbb{E}\big[x(t)^\intercal Q x(t)\big] &= \mathbb{E}\big[(x^\mathrm{c}(t) + x^\mathrm{s}(t))^\intercal Q (x^\mathrm{c}(t) + x^\mathrm{s}(t))\big] \\ &+ \mathbb{E}[x^\ell(t)^\intercal Q x^\ell(t) + 2x^\ell(t)^\intercal Q (x^\mathrm{s}(t) + x^\mathrm{c}(t))]. \end{aligned} \tag{8}$$

Now from (P1) we have

$$\mathbb{E}[x^\ell(t)^\intercal Q x^\ell(t)] = \mathbb{E}[x_2^\ell(t)^\intercal Q_{22} x_2^\ell(t)]. \tag{9}$$

From (P1) and (P6) we have

$$\mathbb{E}[x^\ell(t)^\intercal Q x^\mathrm{s}(t)] = \mathbb{E}[x_2^\ell(t)^\intercal Q_{22} x_2^\mathrm{s}(t)]. \tag{10}$$

Finally, from smoothing property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}[x^\ell(t)^\intercal Q x^\mathrm{c}(t)] &= \mathbb{E}[x^\ell(t)^\intercal Q x^\mathrm{c}(t)|I^\mathrm{c}(t)]] \\ &\overset{(a)}{=} \mathbb{E}[x^\ell(t)^\intercal|I^\mathrm{c}(t)] Q x^\mathrm{c}(t)] \\ &\overset{(b)}{=} 0, \end{aligned} \tag{11}$$

where (a) follows from (P5), and (b) follows from (P1) and (P4). Substituting (9), (10), and (11) in (8) and competing the squares, we get the results. ☐

Step 2. Completion of squares

Lemma 5. For variables x and u and matrices A, B, R and S of appropriate dimensions, we have

$$u^\intercal R u + (Ax+Bu)^\intercal S (Ax+Bu) = (u+Lx)^\intercal \Delta (u+Lx) + x^\intercal \tilde{S} x,$$

where

$$\Delta = [R + B^\intercal S B], \quad L = \Delta^{-1} B^\intercal S A, \quad \tilde{S} = A^\intercal S A - L^\intercal \Delta L.$$

Proof. The proof follows easily from multiplying the second term in the left hand side and adding and subtracting the term $(Lx)^\intercal \Delta L x$. ☐

Lemma 6. For random variables (x, u, w) such that w is zero-mean and independent of (x, u), and given matrices A, B, R, and S of appropriate dimensions, we have

$$\begin{aligned} &\mathbb{E}[u^\intercal R u + (Ax + Bu + w)^\intercal S (Ax + Bu + w)] \\ &\quad = \mathbb{E}[(u + Lx)^\intercal \Delta (u + Lx)] + \mathbb{E}[x^\intercal \tilde{S} x] + \mathbb{E}[w^\intercal S w], \end{aligned}$$

where Δ, L, and $\tilde{S}$ are as in Lemma 5.

Proof. Since w is zero mean and independent of (x, u), we have

$$\begin{aligned} &\mathbb{E}[(Ax + Bu + w)^\intercal S (Ax + Bu + w)] \\ &\quad = \mathbb{E}[(Ax + Bu)^\intercal S (Ax + Bu) + w^\intercal S w]. \end{aligned}$$

Simplifying the first term using Lemma 6, we get the result. ☐

Lemma 7. The total cost $J(g_1, g_2)$ may be written as

$$\begin{aligned} \mathbb{E}\Bigg[& x(1)^\intercal S^\mathrm{c}(1) x(1) + x_2(1)^\intercal S^\ell(1) x_2(1) \\ &+ \sum_{t=1}^{T-1} \Big[w(t)^\intercal S^\mathrm{c}(t+1) w(t) + w_2(t)^\intercal S^\ell(t+1) w_2(t) \Big] \\ &+ \sum_{t=1}^{T-1} \Big[(u^\mathrm{c}(t) + L^\mathrm{c}(t) z^\mathrm{c}(t))^\intercal \Delta^\mathrm{c}(t) (u^\mathrm{c}(t) + L^\mathrm{c}(t) z^\mathrm{c}(t)) \Big] \\ &+ \sum_{t=1}^{T-1} \Big[(u_2^\ell(t) + L^\ell(t) z_2^\ell(t))^\intercal \Delta^\ell(t) (u_2^\ell(t) + L^\ell(t) z_2^\ell(t)) \Big] \\ &+ \sum_{t=1}^{T-1} \Big[(A_{21} x_1^\mathrm{s}(t))^\intercal S^\ell(t+1) (A_{21} x_1^\mathrm{s}(t) + 2A_{22} x_2^\mathrm{s}(t)) \\ &- \sum_{t=1}^{T} x_2^s(t) Q_{22} x_2^s(t) \Big] \Bigg] \end{aligned}$$

where

$$\Delta^{c}(t) = [R + B^{\intercal} S^{c}(t+1) B],$$
$$\Delta^{\ell}(t) = [R_{22} + B_{22}^{\intercal} S^{\ell}(t+1) B_{22}].$$

Proof. We start rewriting the total cost using the result of Lemma 4. Now, the dynamics of $z^{c}(t)$ and $z^{\ell}(t)$ may be written as

$$z^{c}(t+1) = A z^{c}(t) + B u^{c}(t) + w(t),$$
$$z_2^{\ell}(t+1) = A_{22} z_2^{\ell}(t) + A_{21} x_1^{s}(t) + B_{22} u_2^{\ell}(t) + w_2(t),$$

where we can write the full dynamics of second equation as follows:

$$\begin{bmatrix} x_1^{s}(t+1) \\ z_2^{\ell}(t+1) \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1^{s}(t) \\ z_2^{\ell}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ B_{22} \end{bmatrix} u_2^{\ell}(t) + \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}.$$

Note that $w(t)$ is zero mean and independent of $(z^{c}(t), u^{c}(t))$ (because both $z^{c}(t)$ and $u^{c}(t)$ depend on $w(1:t-1)$ which is independent of $w(t)$). Similarly, $w(t)$ is zero mean and independent of $(\text{vec}(x_1^{2}(t), z_2^{\ell}(t)), u_2^{\ell}(t))$. The result then follows from recursively applying Lemma 6. □

An immediate consequence of Lemma 7 is the following.

Lemma 8. We can minimize $J(g_1, g_2)$ by minimizing $\tilde{J}(g_1, g_2)$ defined as

$$\mathbb{E}\Bigg[\sum_{t=1}^{T-1} \Big[(u^{c}(t) + L^{c}(t) z^{c}(t))^{\intercal} \Delta^{c}(t) (u^{c}(t) + L^{c}(t) z^{c}(t))\Big] + \sum_{t=1}^{T-1} \Big[(u_2^{\ell}(t) + L^{\ell}(t) z_2^{\ell}(t))^{\intercal} \Delta^{\ell}(t) (u_2^{\ell}(t) + L^{\ell}(t) z_2^{\ell}(t))\Big]\Bigg]. \tag{12}$$

Proof. The result follows from Lemma 7 and observing that the remaining terms in the expression for $J(g_1, g_2)$ in Lemma 7 are a function of primitive random variables and hence, do not depend on the choice of control actions.

Step 3. Orthogonal projection

In order to minimize $\tilde{J}(g_1, g_2)$ defined in (12), define

$$\hat{z}^{c}(t) := \mathbb{E}[z^{c}(t) | I^{c}(t)],$$
$$\hat{z}_2^{\ell}(t) := \mathbb{E}[z_2^{\ell}(t) | I^{2}(t)] - \mathbb{E}[z_2^{\ell}(t) | I^{1}(t)].$$

Lemma 9. Let $\tilde{z}^{c}(t) = z^{c}(t) - \hat{z}^{c}(t)$ and $\tilde{z}_2^{\ell}(t) = z_2^{\ell}(t) - \hat{z}_2^{\ell}(t)$. Then, we have the following:

(C1) $\tilde{z}^{c}(t)$ and $\tilde{z}_2^{\ell}(t)$ may be written in terms of the primitive random variables. Hence, they do not depend on the control strategies.

(C2) $\mathbb{E}[\tilde{z}^{c}(t)^{\intercal} \hat{z}^{c}(t)] = 0$

(C3) $\mathbb{E}[\tilde{z}_2^{\ell}(t)^{\intercal} \hat{z}_2^{\ell}(t)] = 0.$

(C4) For any matrix M of appropriate dimensions, we have $\mathbb{E}[u_2^{\ell}(t)^{\intercal} M \tilde{z}_2^{\ell}(t)] = 0.$

Proof.

(C1) Note that

$$\hat{z}^{c}(t) = \mathbb{E}[x^{c}(t) + x^{s}(t) | I^{c}(t)] = x^{c}(t) + \mathbb{E}[x^{s}(t) | I^{1,s}(t)].$$

where the second equality uses (P5) and Lemma 2. Thus,

$$\tilde{z}^{c}(t) := z^{c}(t) - \hat{z}^{c}(t) = x^{s}(t) - \mathbb{E}[x^{s}(t) | I^{1,s}(t)],$$

which only depends on the primitive random variables.

For the second part, observe that

$$\begin{aligned}\hat{z}_2^{\ell}(t) &= \mathbb{E}[z_2^{\ell}(t) | I^{2}(t)] - \mathbb{E}[z_2^{\ell}(t) | I^{1}(t)] \\ &= x_2^{\ell}(t) + \mathbb{E}[x_2^{s}(t) | I^{2}(t)] \\ &\qquad - \mathbb{E}[x_2^{\ell}(t) | I^{1}(t)] - \mathbb{E}[x_2^{s}(t) | I^{1}(t)] \\ &\overset{(a)}{=} x_2^{\ell}(t) + \mathbb{L}[x_2^{s}(t) | I^{2,s}(t)] - \mathbb{L}[x_2^{s}(t) | I^{1,s}(t)],\end{aligned} \tag{13}$$

where (a) uses Lemma 2 and (P4). Thus,

$$\begin{aligned}\tilde{z}_2^{\ell}(t) &= z_2^{\ell}(t) - \hat{z}_2^{\ell}(t) \\ &= x_2^{s}(t) - \mathbb{E}[x_2^{s}(t) | I^{2,s}(t)] + \mathbb{E}[x_2^{s}(t) | I^{1,s}(t)],\end{aligned} \tag{14}$$

which only depends on the primitive random variables.

(C2) This follows immediately from the fact that error of a mean-squared estimator is orthogonal to the estimate.

(C3) Using the expressions for $\hat{z}_2^{\ell}(t)$ and $\tilde{z}_2^{\ell}(t)$ from (13) and (14), we get

$$\begin{aligned}&\mathbb{E}[\tilde{z}_2^{\ell}(t)^{\intercal} \hat{z}_2^{\ell}(t)] \\ &= \mathbb{E}\big[x_2^{s}(t)^{\intercal} x_2^{\ell}(t) - \mathbb{E}[x_2^{s}(t) | I^{2}(t)]^{\intercal} x_2^{\ell}(t)\big] \\ &+ \mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{1}(t)]^{\intercal} x_2^{\ell}(t)\big] \\ &+ \mathbb{E}\big[x_2^{s}(t)^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{2}(t)]\big] \\ &\quad - \mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{2}(t)]^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{2}(t)]\big] \\ &\left.\begin{aligned}&+ 2\mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{2}(t)]^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{1}(t)]\big] \\ &\quad - \mathbb{E}\big[x_2^{s}(t)^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{1}(t)]\big] \\ &\quad - \mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{1}(t)]^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{1}(t)]\big].\end{aligned}\right\} \text{(Term I)}\end{aligned} \tag{15}$$

Now, we consider each of the terms separately.

$$\begin{aligned}&\mathbb{E}\big[x_2^{s}(t)^{\intercal} x_2^{\ell}(t) - \mathbb{E}[x_2^{s}(t) | I^{2}(t)]^{\intercal} x_2^{\ell}(t)\big] \\ &\overset{(a)}{=} \mathbb{E}\big[x_2^{s}(t)^{\intercal} x_2^{\ell}(t) - \mathbb{E}[x_2^{s}(t)^{\intercal} x_2^{\ell}(t) | I^{2}(t)]\big] = 0,\end{aligned} \tag{16}$$

where (a) follows because $x_2^{\ell}(t)$ is a function of $I^{2}(t)$ and the last equation follows from smoothing property of conditional expectation. For the remaining terms, we will simplify them by first using the smoothing property of conditional expectation (and conditioning on $I^{1}(t)$ or $I^{2}(t)$ as appropriate), moving one term out of the inner expectation, and simplifying using Lemma 3. Thus,

$$\begin{aligned}&\mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{1}(t)]^{\intercal} x_2^{\ell}(t)\big] \\ &\quad = \mathbb{E}\big[\mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{1}(t)]^{\intercal} x_2^{\ell}(t) \big| I^{1}(t)\big]\big] \\ &\quad = \mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{1}(t)]^{\intercal} \mathbb{E}[x_2^{\ell}(t) | I^{1}(t)]\big] = 0.\end{aligned} \tag{17}$$

Next, consider

$$\begin{aligned}&\mathbb{E}[x_2^{s}(t)^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{2}(t)]] \\ &\quad = \mathbb{E}\big[\mathbb{E}[x_2^{s}(t)^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{2}(t)] | I^{2}(t)]\big] \\ &\quad = \mathbb{E}\big[\mathbb{E}[x_2^{s}(t) | I^{2}(t)]^{\intercal} \mathbb{E}[x_2^{s}(t) | I^{2}(t)]\big].\end{aligned} \tag{18}$$

Finally, by a similar argument we can show that (Term I) is 0. Substituting (16)–(18) in (15), we get the result.

(C4) Using the expression for $\tilde{z}_2^{\ell}(t)$ from (14), we get

$$\begin{aligned}&\mathbb{E}[u_2^{\ell}(t)^{\intercal} M \tilde{z}_2^{\ell}(t)] \\ &\quad = \mathbb{E}[u_2^{\ell}(t)^{\intercal} M \big(x_2^{s}(t) - \mathbb{E}[x_2^{s}(t) | I^{2}(t)]\big)] \\ &\qquad + \mathbb{E}[u_2^{\ell}(t)^{\intercal} M \mathbb{E}[x_2^{s}(t) | I^{1}(t)]]\end{aligned}$$

$$\stackrel{(a)}{=} \mathbb{E}\big[\mathbb{E}[u_2^\ell(t)^\intercal M\big(x_2^{\mathrm{s}}(t) - \mathbb{E}[x_2^{\mathrm{s}}(t)|I^2(t)]\big)|I^2(t)]\big] + \mathbb{E}\big[\mathbb{E}[u_2^\ell(t)^\intercal M \mathbb{E}[x_2^{\mathrm{s}}(t)|I^1(t)]|I^1(t)]\big]$$

$$\stackrel{(b)}{=} \mathbb{E}\big[u_2^\ell(t)^\intercal M \underbrace{\mathbb{E}[x_2^{\mathrm{s}}(t) - \mathbb{E}[x_2^{\mathrm{s}}(t)|I^2(t)]|I^2(t)]}_{=0}\big] + \mathbb{E}\big[\underbrace{\mathbb{E}[u_2^\ell(t)^\intercal|I^1(t)]}_{=0 \text{ by (H2)}} M \mathbb{E}[x_2^{\mathrm{s}}(t)|I^1(t)]\big]$$

$$= 0,$$

where (a) follows from the smoothing property of conditional expectation; the first term in (b) uses the fact that $u_2^\ell(t)$ is a function of $I^2(t)$ and the second term in (b) uses the fact that $\mathbb{E}[x_2^{\mathrm{s}}(t)|I^1(t)]$ is a function of $I^1(t)$.

Proof of Theorem 1

Using (C2) and the fact that $u^{\mathrm{c}}(t)$ is a function of $I^{\mathrm{c}}(t)$, we get

$$\mathbb{E}\big[(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)z^{\mathrm{c}}(t))^\intercal \Delta^{\mathrm{c}}(t)(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)z^{\mathrm{c}}(t))\big] = \mathbb{E}\big[(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)\hat{z}^{\mathrm{c}}(t))^\intercal \Delta^{\mathrm{c}}(t)(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)\hat{z}^{\mathrm{c}}(t))\big] + \mathbb{E}\big[\tilde{z}^{\mathrm{c}}(t)^\intercal L^{\mathrm{c}}(t)^\intercal \Delta^{\mathrm{c}}(t) L^{\mathrm{c}}(t)\tilde{z}^{\mathrm{c}}(t)\big].$$

Using (C3) and (C4), we get

$$\mathbb{E}\big[(u_2^\ell(t) + L^\ell(t)z_2^\ell(t))^\intercal \Delta^\ell(t)(u_2^\ell(t) + L^\ell(t)z_2^\ell(t))\big] = \mathbb{E}\big[(u_2^\ell(t) + L^\ell(t)\hat{z}_2^\ell(t))^\intercal \Delta^\ell(t)(u_2^\ell(t) + L^\ell(t)\hat{z}_2^\ell(t))\big] + \mathbb{E}\big[\tilde{z}_2^\ell(t)^\intercal L^\ell(t)^\intercal \Delta^\ell(t) L^\ell(t)\tilde{z}_2^\ell(t)\big].$$

Substituting these in (12) we get that $\tilde{J}(g_1, g_2)$ is given by

$$\mathbb{E}\Bigg[\sum_{t=1}^{T-1}\Big[(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)\hat{z}^{\mathrm{c}}(t))^\intercal \Delta^{\mathrm{c}}(t)(u^{\mathrm{c}}(t) + L^{\mathrm{c}}(t)\hat{z}^{\mathrm{c}}(t))\Big] + \sum_{t=1}^{T-1}\Big[\tilde{z}^{\mathrm{c}}(t)^\intercal L^{\mathrm{c}}(t)^\intercal \Delta^{\mathrm{c}}(t) L^{\mathrm{c}}(t)\tilde{z}^{\mathrm{c}}(t)\Big] + \sum_{t=1}^{T-1}\Big[(u_2^\ell(t) + L^\ell(t)\hat{z}_2^\ell(t))^\intercal \Delta^\ell(t)(u_2^\ell(t) + L^\ell(t)\hat{z}_2^\ell(t))\Big] + \sum_{t=1}^{T-1}\Big[\tilde{z}_2^\ell(t)^\intercal L^\ell(t)^\intercal \Delta^\ell(t) L^\ell(t)\tilde{z}_2^\ell(t)\Big]\Bigg] \tag{19}$$

Each term in (19) is quadratic and hence positive. By (C1), the second and the forth term do not depend on the control actions. Choosing

$$u^{\mathrm{c}}(t) = -L^{\mathrm{c}}(t)\hat{z}^{\mathrm{c}}(t),$$
$$u^\ell(t) = -L^\ell(t)\hat{z}_2^\ell(t),$$

make the first and third term of (19) equal to zero (and hence minimizes them). Thus, that is an optimal strategy. The result follows from observing that $\hat{x}_2(t|2) - \hat{x}_2(t|c) = \hat{z}_2^\ell(t)$.

4. CONCLUSION

In this paper, we consider the optimal decentralized control of two agent linear system with partial output feedback. We do not assume that the primitive random variables are Gaussian. Using state splitting, static reduction, orthogonal projection, and completion of squares, we show that the optimal strategy is linear and certainty equivalent. The structure of the optimal strategy also shows that there is a two-way separation between estimation and control.

REFERENCES

Bertsekas, D.P. (2000). *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition.

Ho, Y.C. and Chu, K.C. (1972). Team decision theory and information structures in optimal control problems–part I. *IEEE Trans. Autom. Control*, 17(1), 15–22.

Kim, J.H. and Lall, S. (2011). A unifying condition for separable two player optimal control problems. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 3818–3823.

Kim, J.H. and Lall, S. (2012). Separable optimal cooperative control problems. In *2012 American Control Conference (ACC)*, 5868–5873.

Kumar, P.R. and Varaiya, P. (1986). *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Lessard, L. and Lall, S. (2012). Optimal controller synthesis for the decentralized two-player problem with output feedback. In *2012 American Control Conference (ACC)*, 6314–6321.

Lessard, L. and Lall, S. (2015). Optimal control of two-player systems with output feedback. *IEEE Transactions on Automatic Control*, 60(8), 2129–2144.

Lessard, L. and Nayyar, A. (2013). Structural results and explicit solution for two-player LQG systems on a finite time horizon. In *IEEE Conference on Decision and Control (CDC)*, 6542–6549.

Mahajan, A., Martins, N.C., Rotkowitz, M.C., and Yüksel, S. (2012). Information structures in optimal decentralized control. In *IEEE Conference on Decision and Control (CDC)*, 1291–1306.

Nayyar, A., Mahajan, A., and Teneketzis, D. (2013). Decentralized stochastic control with partial history sharing: A common information approach. *IEEE Trans. Autom. Control*, 58(7), 1644–1658.

Nayyar, N., Kalathil, D., and Jain, R. (2018). Optimal decentralized control with asymmetric one-step delayed information sharing. *IEEE Transactions on Control of Network Systems*, 5(1), 653–663.

Root, J.G. (1969). Optimum control of non-gaussian linear stochastic systems with inaccessible state variables. *SIAM Journal on Control*, 7(2), 317–323.

Rotkowitz, M. and Lall, S. (2006). A characterization of convex problems in decentralized control. *IEEE Trans. Autom. Control*, 51(2), 274–286.

Swigart, J. and Lall, S. (2010). An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator. In *Proceedings of the 2010 American Control Conference*, 6385–6390.

Swigart, J. and Lall, S. (2011). Optimal controller synthesis for a decentralized two-player system with partial output feedback. In *Proceedings of the 2011 American Control Conference*, 317–323.

Witsenhausen, H.S. (1968). A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1), 131–147.

Witsenhausen, H.S. (1988). Equivalent stochastic control problems. *Mathematics of Control Signals and Systems*, 1, 3–11.

Wonham, W.M. (1968). On the separation theorem of stochastic control. *SIAM Journal on Control*, 6(2), 312–326.