Optimal Local and Remote Controllers With Unreliable Uplink Channels: An Elementary Proof

Mohammad Afshari, *Student Member, IEEE*, and Aditya Mahajan, *Senior Member, IEEE*

***Abstract*—Recently, a model of a decentralized control system with local and remote controllers connected over unreliable channels was presented in [1]. The model has a nonclassical information structure that is not partially nested. Nonetheless, it is shown in [1] that the optimal control strategies are linear functions of the state estimate (which is a nonlinear function of the observations). Their proof is based on a fairly sophisticated dynamic programming argument. In this article, we present an alternative and elementary proof of the result which uses common information-based conditional independence and completion of squares.**

***Index Terms*—Certainty equivalence, common information approach, linear systems, networked control systems, separation of estimation and control.**

I. Introduction

In a recent paper, a methodology for synthesizing optimal control laws for local and remote controllers for networked control of a linear system over unreliable uplink channel was presented [1]. Such models arise in applications such as temperature control in smart buildings, control of UAVs, vehicle to infrastructure communication, etc.

The model proposed in [1] is a decentralized control system with nonclassical information structure. Due to the unreliable nature of the uplink channels, the information structure is not partially nested. Therefore, one cannot *a priori* restrict attention to linear strategies. Nonetheless, it is shown in [1] that the optimal local and remote control laws are linear functions of the state estimate (which is a nonlinear function of the observations). See Theorem 1 for a precise statement of the result.

The proof technique employed in [1] uses ideas from the common information approach of [2] to compute the optimal control laws. Using a conditional independence argument, it is first shown that the local controllers can ignore the past realizations of their local states without any loss of optimality [1, Lemma 1]. When attention is restricted to control strategies with such a structure, the resulting information structure is partial history sharing [2]. So, in principle, the common information approach of [2] is applicable. However, there are several technical difficulties in extending the argument given in [2] for finite valued random variables to continuous random variables. The key result of [1] is to carefully resolve these technical difficulties—issues of measurability, existence of well-defined value function, and infinite dimensional strategy space—and then obtain a closed form solution of the dynamic program.

In this article, we provide an alternative and elementary proof of the result of [1]. Our proof also relies on the split of total information into common and local information as proposed in [2]. However, instead of using the dynamic program proposed in [2], we develop an alternative solution methodology which relies on: 1) the conditional independence of the local states given the common information (which was established in [1]); 2) simplifying the per-step cost-based on this conditional independence, the orthogonality principle, and the completion of squares. The key advantage of this solution approach is that it completely sidesteps the technical difficulties with measurability and existence of value functions present in a dynamic programming-based approach. Given the paucity of positive results in optimal control of decentralized systems, we believe that a new solution approach is interesting in its own right.

The model considered in [1] consists of N local controllers and one remote controller. For ease of exposition, we assume that $N = 2$. It will be clear from the proof that the steps extend to general N. For the most part, we broadly follow the notation and terminology of [1], but we occasionally deviate from it to be consistent with the standard notation used in linear systems.

A. Notations

We use superscripts to indicate subsystems/controllers and subscripts to indicate time. Thus, x^i_t denotes the state of subsystem i at time t. The superscript $^\intercal$ denotes transpose (of a vector or a matrix). $\mathbf{0}_{m\times n}$ is a $m \times n$ matrix with all elements being equal to zero. We omit the subscript from $\mathbf{0}_{m\times n}$ when the dimension is clear from context. Given column vectors x and y, the notation $\operatorname{vec}(x, y)$ is a short hand for the vector formed by stacking x on top of y. Given random variables x, y, and z, the notation $x \perp\!\!\!\perp y \mid z$ indicates that x and y are conditionally independent given z. Given matrices A and B with the same number of columns, $\operatorname{rows}(A, B)$ denotes the matrix obtained by stacking A on top of B.

Given matrices A, B, Q, M, R, and P of appropriate dimensions, we use the following operators:

$$\begin{aligned}\mathcal{R}(P, A, B, Q, M, R) &= Q + A^\intercal PA \\ &\quad - (M + A^\intercal PB)(R + B^\intercal PB)^{-1}(M + A^\intercal PB)^\intercal \\ \mathcal{G}(P, A, B, M, R) &= (R + B^\intercal PB)^{-1}(M + A^\intercal PB)^\intercal\end{aligned}$$

which denote the one step update of the discrete time Riccati equation and the gain of a linear system, respectively.

II. Model and Problem Formulation

A. System Dynamics

Consider a discrete-time linear dynamical system consisting of $N = 2$ subsystems. Let $x^i_t \in \mathbb{R}^{d^i_x}$ denotes the state of subsystem i,

Manuscript received December 9, 2018; revised July 2, 2019; accepted October 23, 2019. Date of publication November 6, 2019; date of current version July 28, 2020. This work was supported by NSERC under Discovery Accelerator Grant 493011. Recommended by Associate Editor Prof. R. Jain. *(Corresponding author: Mohammad Afshari.)*

The authors are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A-0E9, Canada (e-mail: mohammad.afshari2@mcgill.ca; aditya.mahajan@mcgill.ca).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2019.2951658

$i \in \{1, 2\}$. There is a local controller C^i colocated with subsystem i. In addition, there is a remote controller C^0. The information available to the controllers will be described later. Let $u_t^i \in \mathbb{R}^{d_u^i}$, $i \in \{1, 2\}$, denote the control action of local controller C^i and $u_t^0 \in \mathbb{R}^{d_u^0}$ denote the control action of remote controller C^0.

The initial state x_0^i of subsystem i, $i \in \{1, 2\}$, is random and the dynamics of subsystem i is given by

$$x_{t+1}^i = A^{ii} x_t^i + \begin{bmatrix} B^{i0} & B^{ii} \end{bmatrix} \begin{bmatrix} u_t^0 \\ u_t^i \end{bmatrix} + w_t^i \tag{1}$$

where $w_t^i \in \mathbb{R}^{d_x^i}$ is the process noise and A^{ii}, B^{i0}, and B^{ii} are matrices of appropriate dimensions. We assume that random variables $\{w_0^1, \dots, w_{T-1}^1, w_0^2, \dots, w_{T-1}^2\}$ are independent and have zero mean and finite variance. Let $x_t := \operatorname{vec}(x_t^1, x_t^2)$, $u_t := \operatorname{vec}(u_t^0, u_t^1, u_t^2)$, and $w_t := \operatorname{vec}(w_t^1, w_t^2)$ denote the state, control actions, and noise of the overall system. Then, the system dynamics can be written as

$$x_{t+1} = A x_t + B u_t + w_t \tag{2}$$

where the matrices A and B are given by

$$A = \begin{bmatrix} A^{11} & 0 \\ 0 & A^{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B^{10} & B^{11} & 0 \\ B^{20} & 0 & B^{22} \end{bmatrix}.$$

B. Information Structure

At time t, the local controller C^i, $i \in \{1, 2\}$, perfectly observes the state x_t^i of subsystem i and sends it to the remote controller C^0 over an unreliable packet drop channel. Let $\Gamma_t^i \in \{0, 1\}$ denote the state of the channel, where $\Gamma_t^i = 0$ means that the channel is in the OFF state where the transmitted packet gets dropped while $\Gamma_t^i = 1$ means that the channel is in the ON state where the transmitted packet gets delivered. Thus, Γ_t^i is a Bernoulli random variable and we denote the packet drop probability $\mathbb{P}(\Gamma_t^i = 0)$ by p^i. We use Γ_t to denote (Γ_t^1, Γ_t^2).

Let z_t^i denote the output of the channel i, $i \in \{1, 2\}$, i.e.,

$$z_t^i = f(x_t^i, \Gamma_t^i) = \begin{cases} x_t^i, & \text{if } \Gamma_t^i = 1 \\ \mathfrak{E}, & \text{if } \Gamma_t^i = 0 \end{cases} \tag{3}$$

where $\mathfrak{E}$ denotes a dropped packet. It is assumed that there are perfect channels from C^0 to C^1 and C^2. Using these channels, C^0 can share $z_t := \operatorname{vec}(z_t^1, z_t^2)$ and u_{t-1}^0 with local controllers C^1 and C^2. Note that it is possible to recover Γ_t^i from z_t^i. Hence, all controllers also have access to Γ_t. The fact that Γ_t is available at all controllers is critical to derive the main result of the model (presented in Theorem 1).

Let H_t^i, $i \in \{0, 1, 2\}$, denote the information available to controller C^i to take decisions at time t. Then

$$H_t^0 = \{z_{0:t}, \Gamma_{0:t}, u_{0:t-1}^0\} \tag{4a}$$

$$H_t^i = \{x_{0:t}^i, u_{0:t-1}^i, z_{0:t}, \Gamma_{0:t}, u_{0:t-1}^0\}, \quad i \in \{1, 2\}. \tag{4b}$$

Let $\mathcal{H}_t^i$ be the space of all possible realizations of H_t^i. Then, controller C^i chooses it's control action according to

$$u_t^i = g_t^i(H_t^i), \quad i \in \{0, 1, 2\} \tag{5}$$

where the Borel measurable function $g_t^i : \mathcal{H}_t^i \to \mathbb{R}^{d_u^i}$ is called the *control law of controller* C^i at time t. The collection $\mathbf{g}^i = (g_0^i, \dots, g_{T-1}^i)$ is called the *control strategy of controller* C^i. The collection $\mathbf{g} := (\mathbf{g}^0, \mathbf{g}^1, \mathbf{g}^2)$ is called the *strategy profile of the system*.

C. System Performance and the Optimization Problem

The system operates for a finite horizon T. At time $t < T$, the system incurs a per-step cost

$$c_t(x_t, u_t) = \begin{bmatrix} x_t \\ u_t \end{bmatrix}^\intercal \begin{bmatrix} Q_t & M_t \\ M_t^\intercal & R_t \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix}$$

and at the terminal time T, the system incurs a terminal cost

$$c_T(x_T) = x_T^\intercal Q_T x_T$$

where Q_t, M_t, and R_t are matrices of appropriate dimensions. We assume the following block-wise structure of Q_t, M_t, and R_t:

$$Q_t = \begin{bmatrix} Q_t^{11} & Q_t^{12} \\ Q_t^{21} & Q_t^{22} \end{bmatrix}, \quad M_t = \begin{bmatrix} M_t^{10} & M_t^{11} & M_t^{12} \\ M_t^{20} & M_t^{21} & M_t^{22} \end{bmatrix}$$

and

$$R_t = \begin{bmatrix} R_t^{00} & R_t^{01} & R_t^{02} \\ R_t^{10} & R_t^{11} & R_t^{12} \\ R_t^{20} & R_t^{21} & R_t^{22} \end{bmatrix}.$$

The performance of a strategy profile $\mathbf{g}$ is given by

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}} \left[\sum_{t=0}^{T-1} c_t(x_t, u_t) + c_T(x_T) \right] \tag{6}$$

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy joint profile $\mathbf{g}$.

The following assumptions are imposed on the system.

A1) The primitive random variables $\{x_0^1, x_0^2, w_0^1, \dots, w_{T-1}^1, w_0^2, \dots, w_{T-1}^2, \Gamma_0^1, \dots, \Gamma_{T-1}^1, \Gamma_0^2, \dots, \Gamma_{T-1}^2\}$ are independent.

A2) The variables $\{x_0^1, x_0^2, w_0^1, \dots, w_{T-1}^1, w_0^2, \dots, w_{T-1}^2\}$ have zero mean and finite variance. We use Σ_t^i and Σ_x^i to denote the variance of w_t^i and x_0^i, respectively.

A3) For each t, the matrix $\left[\begin{smallmatrix} Q_t & M_t \\ M_t^\intercal & R_t \end{smallmatrix}\right]$ is symmetric and positive semidefinite, and the matrix R_t is symmetric and positive definite.

We are interested in the following optimization problem.

Problem 1: In the model described above, find a strategy profile $\mathbf{g}^* = (\mathbf{g}^{*,0}, \mathbf{g}^{*,1}, \mathbf{g}^{*,2})$ that minimizes (6), i.e.,

$$J^* := J(\mathbf{g}^*) = \inf_{\mathbf{g}} J(\mathbf{g})$$

where the infimum is taken over all strategy profiles of the form (5).

D. Some Remarks

The per-step cost function defined above differs slightly from the per-step cost function considered in [1] in the following ways.

1) In [1], the matrix $\left[\begin{smallmatrix} Q_t & M_t \\ M_t^\intercal & R_t \end{smallmatrix}\right]$ was denoted by R_t. We follow the standard notation here.
2) In [1], it was assumed that the performance of a strategy profile is

$$\mathbb{E}^{\mathbf{g}} \left[\sum_{t=0}^{T} c_t(x_t, u_t) \right].$$

This is effectively the same as assuming that there is no terminal cost (i.e., $Q_{T+1} = 0$) and therefore the terminal control actions u_T^i are 0 for both local and remote controllers. To avoid such triviality, we assume a performance function of the form (6).

III. Main Result

In this section, we restate the main results of [1] but we present them in a slightly different manner.

A. Common Information-Based Estimates

Following [2], we define the *common information* H_t^{com} between agents as

$$H_t^{\text{com}} = H_t^0 \cap H_t^1 \cap H_t^2.$$

The information structure of the model (4) implies that $H_t^{\text{com}} = H_t^0 = \{z_{0:t}, \Gamma_{0:t}, u^0_{0:t-1}\}$.

Now we define the common information-based "estimates" of the state and control actions and the corresponding "estimation errors" as follows:

$$\hat{x}_t = \mathbb{E}[x_t \mid H_t^{\text{com}}], \quad \tilde{x}_t = x_t - \hat{x} \tag{7}$$

$$\hat{u}_t = \mathbb{E}[u_t \mid H_t^{\text{com}}], \quad \tilde{u}_t = u_t - \hat{u}_t. \tag{8}$$

For ease of notation, we use $\hat{x}_t^i$ to denote the i_{th} component of $\hat{x}_t$, i.e., $\hat{x}_t = \text{vec}(\hat{x}_t^1, \hat{x}_t^2)$. Similar interpretation holds for $\tilde{x}_t^i$, $\hat{u}_t^i$, and $\tilde{u}_t^i$.

It can be shown that the state estimates and the estimation error satisfy the following property.

Lemma 1: The state estimates and estimation errors evolve as follows: for $i \in \{1, 2\}$

$$\hat{x}_0^i = \begin{cases} 0, & \text{if } \Gamma_0^i = 0 \\ x_0^i, & \text{if } \Gamma_0^i = 1 \end{cases}$$

and for $t > 0$,

$$\hat{x}_{t+1}^i = \begin{cases} A^{ii}\hat{x}_t^i + B^{i0}u_t^0 + B^{ii}\hat{u}_t^i, & \text{if } \Gamma_{t+1}^i = 0 \\ x_{t+1}^i, & \text{if } \Gamma_{t+1}^i = 1. \end{cases}$$

Therefore,

$$\tilde{x}_0^i = \begin{cases} x_0^i, & \text{if } \Gamma_0^i = 0 \\ 0, & \text{if } \Gamma_0^i = 1 \end{cases}$$

and for $t > 0$

$$\tilde{x}_{t+1}^i = \begin{cases} A^{ii}\tilde{x}_t^i + B^{ii}\tilde{u}_t^i + w_t^i, & \text{if } \Gamma_{t+1}^i = 0 \\ 0, & \text{if } \Gamma_{t+1}^i = 1. \end{cases} \qquad \square$$

A proof is presented in Section IV-D.

Remark 1: Lemma 1 along with the definition of the state and control estimates (7) and (8) and the information structure (4) imply that all controllers know the value of $\text{vec}(\hat{x}_t^1, \hat{x}_t^2)$ at time t. Consequently, controller C^i knows the value of $\tilde{x}_t^i$ at time t. The main result of the model, explained in the next section, is that the optimal control action at controller C^i is linear in $(\hat{x}_t, \tilde{x}_t^i)$. $\square$

B. Structure of Optimal Control Laws

In order to present the main result of [1], we recursively define matrices $\{P_t\}_{t=1}^T$ as follows: $P_T = Q_T$ and for $t \in \{T-1, \dots, 1\}$

$$P_t = \mathcal{R}(P_{t+1}, A, B, Q_t, M_t, R_t). \tag{9}$$

Furthermore, let P_t^{ii} denote the (i, i)th block of P_t. Then for $i \in \{1, 2\}$, recursively define the matrices $\{\Pi_t^i\}_{t=1}^T$ and $\{\tilde{P}_t^i\}_{t=1}^T$ as follows: $\Pi_T^i = Q_T^{ii}$ and $\tilde{P}_T^i = Q_T^{ii}$ and for $t \in \{T-1, \dots, 1\}$, let

$$\tilde{P}_t^i = \mathcal{R}(\Pi_{t+1}^i, A^{ii}, B^{ii}, Q_t^{ii}, M_t^{ii}, R_t^{ii}) \tag{10}$$

and

$$\Pi_{t+1}^i = (1 - p^i)P_{t+1}^{ii} + p^i\tilde{P}_{t+1}^i. \tag{11}$$

The main result of [1] is the following.

Theorem 1: The optimal control strategy for Problem 1 is given by

$$\begin{bmatrix} u_t^0 \\ \hat{u}_t^1 \\ \hat{u}_t^2 \end{bmatrix} = -K_t\hat{x}_t \tag{12}$$

and

$$\tilde{u}_t^i = -\tilde{K}_t^i\tilde{x}_t^i, \quad i \in \{1, 2\}, \tag{13}$$

where the time evolution of $\hat{x}_t$ and $\tilde{x}_t$ are given by Lemma 1.

The gains $\{K_t\}_{t=0}^{T-1}$ and $\{\tilde{K}_t\}_{t=0}^{T-1}$ are given by

$$K_t = \mathcal{G}(P_{t+1}, A, B, M_t, R_t)$$

$$\tilde{K}_t^i = \mathcal{G}(\Pi_{t+1}^i, A^{ii}, B^{ii}, M_t^{ii}, R_t^{ii}), \quad i \in \{1, 2\}$$

where the matrices $\{P_t\}_{t=1}^T$, $\{\Pi_t^i\}_{t=1}^T$, and $\{\tilde{P}^i\}_{t=1}^T$ are given by (9), (10), and (11). $\square$

Remark 2: Let $K_t = \text{rows}(K_t^0, K_t^1, K_t^2)$. Then, Theorem 1 implies that the optimal control actions are given by

$$u_t^0 = -K_t^0\hat{x}_t, \tag{14}$$

$$u_t^i = -K_t^i\hat{x}_t - \tilde{K}_t^i(x_t^i - \hat{x}_t^i), \quad i \in \{1, 2\}. \tag{15}$$

Such a control law is feasible because, $\tilde{x}_t^i$ is available at controller C^i as explained in Remark 1.

The structure of the control laws (14) and (15) implies that the optimal action is a linear function of the state estimate $\hat{x}_t$. Note that the evolution of the state estimate, given by Lemma 1, is a nonlinear function of the data available at controller C^i, $i \in \{1, 2\}$. $\square$

Remark 3: The result does not depend on the distribution of the noise processes $\{w_t^i\}_{t\geq 0}$, $i \in \{1, 2\}$, as long as the random variables $\{w_0^1, \dots, w_{T-1}^1, w_0^2, \dots, w_{T-1}^2\}$ are independent and have finite second moment. For convenience we have presented the result under the additional assumption that the noise is zero-mean but that assumption can be relaxed using a change of variables. $\square$

IV. Proof of the Main Result

A. Roadmap of the Proof

Our proof is based on the following fact which is typically referred to as the *completion of squares* in the literature.

Fact Given a linear system $x_{t+1} = Ax_t + Bu_t + w_t$, the quadratic cost

$$\sum_{t=0}^{T-1}[x_t^\intercal Q_t x_t + u_t^\intercal R_t u_t] + x_T^\intercal Q_T x_T$$

may be rewritten as

$$x_0^\intercal P_0 x_0 + \sum_{t=0}^{T-1}(u_t + L_t x_t)^\intercal \Delta_t(u_t + L_t x_t) + \sum_{t=0}^{T-1} w_t^\intercal P_{t+1} w_t$$

where $P_T = Q_T$ and for $t \in \{T-1, \dots, 0\}$, $P_t = \mathcal{R}(P_{t+1}, A, B, Q_t, \mathbf{0}, R_t)$, $L_t = \mathcal{G}(P_{t+1}, A, B, \mathbf{0}, R_t)$, and $\Delta_t = (R_t + B^\intercal P_{t+1}B)$.

Using this fact, one can prove the structure of optimal strategy for the centralized control of stochastic linear systems for both complete and partial state observation. See, for example, [3, Ch. 8]. However, the completion of squares argument does not work directly for decentralized control systems.

In our proof, we exploit a fundamental property of the model, which was established in [1, Claim 2] and is formally stated as Lemma 2 below: $x_t^1 \perp\!\!\!\perp x_t^2 | H_t^{\text{com}}$. As a consequence of this conditional independence, the past realizations $(x_{0:t-1}^i, u_{0:t-1}^i)$ are irrelevant at controller C^i and may be shed without loss of optimality. This follows from Blackwell's principle of irrelevant information [4] as generalized to decentralized control systems in [5]. The simplified structure of the optimal controller was established in [1, Lemma 1] and is formally stated as Lemma 3 below.

Using these two results and basic properties of conditional expectations, we prove the structure of the dynamics of the state estimates and the estimation error (Lemma 1). This structure was also established in [1, Th. 3] as part of the result that establishes the structure of the optimal controller. However, as we show below, one only needs the conditional independence property of Lemma 2 and its consequences to establish Lemma 1.

As a next step, we use orthogonal projections and the specific form of the information structure to simplify the per-step cost (Lemma 5). We combine this simplified form of the cost with the dynamics of the state estimates and estimation error (established in Lemma 1) to prove completion of squares result for the cost (Theorem 2) tailored to the specific model of the system.

Subsequently, we follow the standard steps of the "completion of squares" argument to establish the structure of the optimal strategy.

B. Conditional Independence of Local States and Its Implications

A key property of the model established in [1, Claim 2] is the following.

Lemma 2: For any control strategy profile $\mathbf{g}$ of the form (5)

$$x_t^1 \perp\!\!\!\perp x_t^2 \mid H_t^{\text{com}}. \qquad \square$$

Furthermore, it is shown in [1, Lemma 1] that the above conditional independence implies the following.

Lemma 3: In Problem 1, there is no loss of optimality to restrict attention to local controllers of the form

$$u_t^i = g_t^i(x_t^i, H_t^{\text{com}}), \quad i \in \{1, 2\}. \tag{16}$$

$\square$

An immediate consequence of the above lemma is the following.

Corollary 1: For any control strategy profile $\mathbf{g}$ of the form (16), we have the following:

1) $(x_t^1, u_t^1) \perp\!\!\!\perp (x_t^2, u_t^2) \mid H_t^{\text{com}}$;
2) $(\tilde{x}_t^1, \tilde{u}_t^1) \perp\!\!\!\perp (\tilde{x}_t^2, \tilde{u}_t^2) \mid H_t^{\text{com}}$. $\square$

Proof: Property 1 follows from the Lemma 2 and the structure of the control strategy. Property 2 follows from Property 1, (7) and (8) and the fact that $\hat{x}_t^i$ and $\hat{u}_t^i$ are functions of H_t^{com}. ■

C. Some Preliminary Properties

Lemma 4: For any control strategy profile $\mathbf{g}$ of the form (5), we have the following.

H1) $\hat{u}_t^0 = u_t^0$ and $\tilde{u}_t^0 = \mathbf{0}$. Thus, $\hat{u}_t = \text{vec}(u_t^0, \hat{u}_t^1, \hat{u}_t^2)$ and $\tilde{u}_t = \text{vec}(\mathbf{0}, \tilde{u}_t^1, \tilde{u}_t^2)$.
H2) $\mathbb{E}[\tilde{x}_t \mid H_t^{\text{com}}] = 0$ and $\mathbb{E}[\tilde{u}_t \mid H_t^{\text{com}}] = 0$.
H3) For any matrix W of appropriate dimensions, $\mathbb{E}[\hat{s}_t^\intercal W \tilde{s}_t] = 0$, where $\hat{s}_t = \text{vec}(\hat{x}_t, \hat{u}_t)$ and $\tilde{s}_t = \text{vec}(\tilde{x}_t, \tilde{u}_t)$.

Furthermore, if the strategy profile is of the form (16), we have:

H4) For any matrix W of appropriate dimensions, $\mathbb{E}[(\tilde{s}_t^1)^\intercal W \tilde{s}_t^2] = 0$, where $\tilde{s}_t^i = \text{vec}(\tilde{x}_t^i, \tilde{u}_t^i)$. $\square$

Proof: Property (H1) follows from the fact that u_t^0 is a measurable function of H_t^0 (which is the same as H_t^{com}).

Property (H2) is a standard property of error estimates and can be shown as follows:

$$\mathbb{E}[\tilde{x}_t \mid H_t^{\text{com}}] = \mathbb{E}[x_t - \mathbb{E}[x_t \mid H_t^{\text{com}}] \mid H_t^{\text{com}}] = 0.$$

Property (H3) follows from the generalized orthogonality principle and can be shown as follows:

$$\mathbb{E}[\hat{s}_t^\intercal W \tilde{s}_t] = \mathbb{E}\big[\mathbb{E}[\hat{s}_t^\intercal W \tilde{s}_t \mid H_t^{\text{com}}]\big] = \mathbb{E}\big[\hat{s}_t^\intercal W \underbrace{\mathbb{E}[\tilde{s}_t \mid H_t^{\text{com}}]}_{=0 \text{ (by (H2))}}\big].$$

Property (H4) follows from Corollary 1 and (H2). ■

D. Proof of Lemma 1

By definition, $H_{t+1}^{\text{com}} = H_t^{\text{com}} \cup \{z_{t+1}^1, z_{t+1}^2, \Gamma_{t+1}^1, \Gamma_{t+1}^2, u_t^0\}$. Thus

$$\begin{aligned}\hat{x}_{t+1}^i &= \mathbb{E}[x_{t+1}^i \mid H_{t+1}^{\text{com}}] \\ &= \mathbb{E}[x_{t+1}^i \mid H_t^{\text{com}}, z_{t+1}^1, z_{t+1}^2, \Gamma_{t+1}^1, \Gamma_{t+1}^2, u_t^0] \\ &= \mathbb{E}[x_{t+1}^i \mid H_t^{\text{com}}, z_{t+1}^i, \Gamma_{t+1}^i],\end{aligned} \tag{17}$$

where we can remove z_{t+1}^{-i} (where $-i$ means the controller other than i), Γ_{t+1}^{-i}, and u_t^0 due to the following reasons.

1) By (3), $z_{t+1}^{-i} = f(x_{t+1}^{-i}, \Gamma_{t+1}^{-i})$ and hence conditionally independent of x_{t+1}^i given H_t^{com} due to (A1), Lemmas 2 and 3.
2) Γ_{t+1}^{-i} is conditionally independent of x_{t+1}^i given $\{H_t^{\text{com}}, u_t^0\}$ due to (1) and (A1).
3) $u_t^0 = g_t^0(H_t^0)$ and hence may be removed from the conditioning (since $H_t^{\text{com}} = H_t^0$).

Now, we consider the two cases $\Gamma_{t+1}^i = 0$ and $\Gamma_{t+1}^i = 1$ separately. When $\Gamma_{t+1}^i = 0$, $z_{t+1}^i = \mathfrak{E}$ and from (17) we have

$$\begin{aligned}\hat{x}_{t+1}^i &= \mathbb{E}[x_{t+1}^i \mid H_t^{\text{com}}, z_{t+1}^i = \mathfrak{E}, \Gamma_{t+1}^i = 0] \\ &\overset{(a)}{=} \mathbb{E}[x_{t+1}^i \mid H_t^{\text{com}}] \\ &\overset{(b)}{=} A^{ii}\hat{x}_t^i + B^{i0}u_t^0 + B^{ii}\hat{u}_t^i\end{aligned}$$

where (a) follows from (A1) and (b) follows from (1), (7), (8), (H1), (A1), and (A2). Consequently

$$\tilde{x}_{t+1}^i = x_{t+1}^i - \hat{x}_{t+1}^i = A^{ii}\tilde{x}_t^i + B^{ii}\tilde{u}_t^i + w_t^i.$$

Now consider the case when $\Gamma_{t+1}^i = 1$, i.e., $z_{t+1}^i = x_{t+1}^i$. Therefore,

$$\hat{x}_{t+1}^i = \mathbb{E}[x_{t+1}^i \mid H_t^{\text{com}}, z_{t+1}^i = x_{t+1}^i, \Gamma_{t+1}^i = 1] = x_{t+1}^i.$$

Consequently, $\tilde{x}_{t+1}^i = x_{t+1}^i - \hat{x}_{t+1}^i = 0$.

E. Orthogonal Projection for Per-Step Cost

Lemma 5: For any strategy profile of the form (16), we have

$$\mathbb{E}[x_t^\intercal Q_t x_t] = \mathbb{E}\left[\hat{x}_t^\intercal Q_t \hat{x}_t + \sum_{i\in\{1,2\}} (\tilde{x}_t^i)^\intercal Q_t^{ii} \tilde{x}_t^i\right] \tag{18}$$

$$\mathbb{E}[u_t^\intercal R_t u_t] = \mathbb{E}\left[\hat{u}_t^\intercal R_t \hat{u}_t + \sum_{i\in\{1,2\}} (\tilde{u}_t^i)^\intercal R_t^{ii} \tilde{u}_t^i\right] \tag{19}$$

$$\mathbb{E}[x_t^\intercal M_t u_t] = \mathbb{E}\left[\hat{x}_t^\intercal M_t \hat{u}_t + \sum_{i\in\{1,2\}} (\tilde{x}_t^i)^\intercal M_t^{ii} \tilde{u}_t^i\right]. \tag{20}$$

Thus, we have that

$$\mathbb{E}\left[\begin{bmatrix}x_t\\u_t\end{bmatrix}^\intercal\begin{bmatrix}Q_t & M_t\\M_t^\intercal & R_t\end{bmatrix}\begin{bmatrix}x_t\\u_t\end{bmatrix}\right]=\mathbb{E}\left[\begin{bmatrix}\hat{x}_t\\\hat{u}_t\end{bmatrix}^\intercal\begin{bmatrix}Q_t & M_t\\M_t^\intercal & R_t\end{bmatrix}\begin{bmatrix}\hat{x}_t\\\hat{u}_t\end{bmatrix}\right]$$
$$+\sum_{i\in\{1,2\}}\mathbb{E}\left[\begin{bmatrix}\tilde{x}_t^i\\\tilde{u}_t^i\end{bmatrix}^\intercal\begin{bmatrix}Q_t^{ii} & M_t^{ii}\\(M_t^{ii})^\intercal & R_t^{ii}\end{bmatrix}\begin{bmatrix}\tilde{x}_t^i\\\tilde{u}_t^i\end{bmatrix}\right].\quad\square$$

Proof: To show (18), we recall that $x_t=\hat{x}_t+\tilde{x}_t$. Thus

$$\mathbb{E}[x_t^\intercal Q_t x_t]=\mathbb{E}[\hat{x}_t^\intercal Q_t\hat{x}_t+\tilde{x}_t^\intercal Q_t\tilde{x}_t+2\hat{x}_t^\intercal Q_t\tilde{x}_t].\quad(21)$$

Consider the second term of (21)

$$\mathbb{E}[\tilde{x}_t^\intercal Q_t\tilde{x}_t]=\sum_{i\in\{1,2\}}\mathbb{E}[(\tilde{x}_t^i)^\intercal Q_t^{ii}\tilde{x}_t^i]+2\underbrace{\mathbb{E}[(\tilde{x}_t^1)^\intercal Q_t^{12}\tilde{x}_t^2]}_{=0\text{ (by (H4))}}.\quad(22)$$

Substituting (22) in (21) and observing that the third term of (21) is 0 due to (H3), we get (18).

Equations (19) and (20) can be proved in a similar manner. ■

F. Change of Variables

For ease of notation, we define

$$\hat{x}_{t+1}^{i,\mathrm{OFF}}=A^{ii}\hat{x}_t^i+B^{i0}u_t^0+B^{ii}\hat{u}_t^i\quad(23)$$
$$\tilde{x}_{t+1}^{i,\mathrm{OFF}}=A^{ii}\tilde{x}_t^i+B^{ii}\tilde{u}_t^i+w_t^i.\quad(24)$$

Thus, we can write

$$\hat{x}_{t+1}^i=\begin{cases}\hat{x}_{t+1}^{i,\mathrm{OFF}}, & \text{if }\Gamma_{t+1}^i=0\\ x_{t+1}^i, & \text{if }\Gamma_{t+1}^i=1,\end{cases}$$

and

$$\tilde{x}_{t+1}^i=\begin{cases}\tilde{x}_{t+1}^{i,\mathrm{OFF}}, & \text{if }\Gamma_{t+1}^i=0\\ 0, & \text{if }\Gamma_{t+1}^i=1.\end{cases}$$

Let $\hat{x}_t^{\mathrm{OFF}}=\mathrm{vec}(\hat{x}_t^{1,\mathrm{OFF}},\hat{x}_t^{2,\mathrm{OFF}})$ and $\tilde{x}_t^{\mathrm{OFF}}=\mathrm{vec}(\tilde{x}_t^{1,\mathrm{OFF}},\tilde{x}_t^{2,\mathrm{OFF}})$. It follows that $\hat{x}_{t+1}^{i,\mathrm{OFF}}+\tilde{x}_{t+1}^{i,\mathrm{OFF}}=x_{t+1}^i$ and

$$\hat{x}_{t+1}^{\mathrm{OFF}}=A\hat{x}_t+B\hat{u}_t.\quad(25)$$

Lemma 6: For any strategy profile of the form (5), we have the following:

H5) For any matrix W of appropriate dimensions, $\mathbb{E}[(\hat{x}_t^{\mathrm{OFF}})^\top W\tilde{x}_t^{\mathrm{OFF}}]=0$.

Furthermore, if the strategy profile is of the form (16), we have:

H6) For any matrix W of appropriate dimensions, $\mathbb{E}[(\tilde{x}_t^{1,\mathrm{OFF}})^\intercal W\tilde{x}_t^{2,\mathrm{OFF}}]=0$. □

Proof: Property (H5) follows immediately from (H3). Property (H6) follows immediately from (H4) and (A1). ■

Lemma 7: For any strategy profile of the form (16), we have the following:

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^i)^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^i\right]$$
$$=\mathbb{E}\left[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^{i,\mathrm{OFF}})^\intercal\Pi_{t+1}^i\tilde{x}_{t+1}^{i,\mathrm{OFF}}\right].\quad(26)$$

□

Proof: We compute the conditional value of the left hand side given the realization of $\Gamma_{t+1}=(\Gamma_{t+1}^1,\Gamma_{t+1}^2)$ and using Lemma 1. We have four cases

1) $\Gamma_{t+1}=(0,0)$: In this case $\hat{x}_{t+1}=\hat{x}_{t+1}^{\mathrm{OFF}}$ and $\tilde{x}_{t+1}=\tilde{x}_{t+1}^{\mathrm{OFF}}$. Thus

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^i)^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^i\middle|\Gamma_{t+1}=(0,0)\right]$$
$$=\mathbb{E}\left[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^{i,\mathrm{OFF}})^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^{i,\mathrm{OFF}}\right].$$

2) $\Gamma_{t+1}=(1,0)$: In this case $\hat{x}_{t+1}=\mathrm{vec}(x_{t+1}^1,\hat{x}_{t+1}^{2,\mathrm{OFF}})=\hat{x}_{t+1}^{\mathrm{OFF}}+\mathrm{vec}(\tilde{x}_{t+1}^{1,\mathrm{OFF}},\mathbf{0})$ and $\tilde{x}_{t+1}=\mathrm{vec}(\mathbf{0},\tilde{x}_{t+1}^{2,\mathrm{OFF}})$. Thus,

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^i)^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^i\middle|\Gamma_{t+1}=(1,0)\right]$$
$$=\mathbb{E}\Big[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+(\tilde{x}_{t+1}^{1,\mathrm{OFF}})^\intercal P_{t+1}^{11}\tilde{x}_{t+1}^{1,\mathrm{OFF}}$$
$$+(\tilde{x}_{t+1}^{2,\mathrm{OFF}})^\intercal\tilde{P}_{t+1}^2\tilde{x}_{t+1}^{2,\mathrm{OFF}}\Big].$$

3) $\Gamma_{t+1}=(0,1)$: Similar to case 2), we can show that

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^i)^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^i\middle|\Gamma_{t+1}=(0,1)\right]$$
$$=\mathbb{E}\Big[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+(\tilde{x}_{t+1}^{1,\mathrm{OFF}})^\intercal\tilde{P}_{t+1}^1\tilde{x}_{t+1}^{1,\mathrm{OFF}}$$
$$+(\tilde{x}_{t+1}^{2,\mathrm{OFF}})^\intercal P_{t+1}^{22}\tilde{x}_{t+1}^{2,\mathrm{OFF}}\Big].$$

4) $\Gamma_{t+1}=(1,1)$: In this case, $\hat{x}_{t+1}=x_{t+1}=\hat{x}_{t+1}^{\mathrm{OFF}}+\tilde{x}_{t+1}^{\mathrm{OFF}}$ and $\tilde{x}_{t+1}=\mathbf{0}$. Thus

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^i)^\intercal\tilde{P}_{t+1}^i\tilde{x}_{t+1}^i\middle|\Gamma_{t+1}=(1,1)\right]$$
$$=\mathbb{E}\big[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+(\tilde{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\tilde{x}_{t+1}^{\mathrm{OFF}}$$
$$+2(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\tilde{x}_{t+1}^{\mathrm{OFF}}\big]$$
$$=\mathbb{E}\left[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}}+\sum_{i\in\{1,2\}}(\tilde{x}_{t+1}^{i,\mathrm{OFF}})^\intercal P_{t+1}^{ii}\tilde{x}_{t+1}^{i,\mathrm{OFF}}\right]$$

where the last equality follows from (H5) and (H6).

Combining these four cases and using the law of total probability, we get (26). ■

G. Completion of Squares

Lemma 8: Let $x\in\mathbb{R}^{d_x}$, $u\in\mathbb{R}^{d_u}$, and $w\in\mathbb{R}^{d_x}$ be random variables defined on a common probability space. Suppose w is zero mean with finite covariance and independent of (x,u). Let $x_+=Ax+Bu+w$, where A and B are matrices of appropriate dimensions. Then given matrices P, Q, M, and R of appropriate dimensions

$$\mathbb{E}\left[\begin{bmatrix}x\\u\end{bmatrix}^\intercal\begin{bmatrix}Q & M\\M^\intercal & R\end{bmatrix}\begin{bmatrix}x\\u\end{bmatrix}+x_+^\intercal Px_+\right]$$
$$=\mathbb{E}\left[x^\intercal P_+x+(u+Kx)^\intercal\Delta(u+Kx)+w^\intercal Pw\right]$$

where

$$\Delta=R+B^\intercal PB,$$
$$K=\Delta^{-1}[M^\intercal+B^\intercal PA]$$
$$P_+=Q+A^\intercal PA-K^\intercal\Delta K.\quad\square$$

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

$$\mathbb{E}[x_+^\intercal P x_+] = \mathbb{E}\big[(Ax+Bu)^\intercal P(Ax+Bu) + w^\intercal P w\big].$$

The result follows by expanding both sides and comparing coefficients. ■

By combining Lemmas 5, 7, and 8, we get the following.

Lemma 9: For any strategy profile of the form (16)

$$\mathbb{E}\left[c_t(x_t,u_t) + \hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1} + \sum_{i\in\{1,2\}} (\tilde{x}_{t+1}^i)^\intercal \tilde{P}_{t+1}^i \tilde{x}_{t+1}^i\right]$$
$$= \mathbb{E}\Big[\hat{x}_t^\intercal P_t \hat{x}_t + (\hat{u}_t + K_t\hat{x}_t)^\intercal \Delta_t(\hat{u}_t + K_t\hat{x}_t) + \sum_{i\in\{1,2\}}\Big[(\tilde{x}_t^i)^\intercal \tilde{P}_t^i \tilde{x}_t^i + (\tilde{u}_t + \tilde{K}_t\tilde{x}_t)^\intercal \tilde{\Delta}_t(\tilde{u}_t + \tilde{K}_t\tilde{x}_t) + (w_t^i)^\intercal \tilde{\Pi}_{t+1}^i w_t^i\Big]\Big]. \quad \square$$

Theorem 2: For any strategy profile $\mathbf{g}$ of the form (16)

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\Bigg[\hat{x}_0^\intercal P_t \hat{x}_0 + \sum_{i\in\{1,2\}}(\tilde{x}_0^i)^\intercal \tilde{P}_0^i \tilde{x}_0^i + \sum_{s=0}^{T-1}(\hat{u}_s + K_s\hat{x}_s)^\intercal \Delta_s(\hat{u}_s + K_s\hat{x}_s) + \sum_{s=0}^{T-1}\sum_{i\in\{1,2\}}(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i)^\intercal \tilde{\Delta}_s^i(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i) + \sum_{s=0}^{T-1}\sum_{i\in\{1,2\}}(w_s^i)^\intercal \Pi_{t+1}^i w_s^i\Bigg] \tag{27}$$

where $\Delta_s = R_s + B^\intercal P_{s+1} B$ and $\tilde{\Delta}_s^i = R_s^{ii} + (B^{ii})^\intercal \Pi_{s+1}^i B^{ii}$, $i \in \{1,2\}$. □

Proof: For any strategy profile $\mathbf{g}$, define the expected cost to go from time t onward as

$$V_t(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\left[\sum_{s=t}^{T-1} c_s(x_s,u_s) + c_T(x_T)\right]. \tag{28}$$

We claim that

$$V_t(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\Bigg[\hat{x}_t^\intercal P_t \hat{x}_t + \sum_{i\in\{1,2\}}(\tilde{x}_t^i)^\intercal \tilde{P}_t^i \tilde{x}_t^i + \sum_{s=t}^{T-1}(\hat{u}_s + K_s\hat{x}_s)^\intercal \Delta_s(\hat{u}_s + K_s\hat{x}_s) + \sum_{s=t}^{T-1}\sum_{i\in\{1,2\}}(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i)^\intercal \tilde{\Delta}_s^i(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i) + \sum_{s=t}^{T-1}\sum_{i\in\{1,2\}}(w_s^i)^\intercal \Pi_{t+1}^i w_s^i\Bigg]. \tag{29}$$

We prove the claim by backward induction. For $t = T$, Lemma 5 implies that

$$V_T(\mathbf{g}) = \mathbb{E}\left[\hat{x}_T^\intercal Q_T \hat{x}_T + \sum_{i\in\{1,2\}}(\tilde{x}_T^i)^\intercal Q_T^{ii}\tilde{x}_T^i\right].$$

Equation (29) follows from the definition of P_T and $\tilde{P}_T^i$. This forms the basis of induction. Now assume that (29) is true for $t+1$ and consider V_t. By definition, we have

$$V_t(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}[c_t(x_t,u_t)] + V_{t+1}(\mathbf{g}).$$

Using the expression for V_{t+1} and Lemma 9, we get the expression for V_t. This completes the induction step and proves the claim (29).

The result of the Theorem follows from observing that $J(\mathbf{g}) = V_0(\mathbf{g})$. ■

H. Proof of Theorem 1

By Lemma 3, there is no loss of optimality in restricting attention to control strategy profile of the form (16). By Theorem 2, the performance of a strategy of the form (16) is given by (27). Note that the first two and the last terms of (27) are control free (i.e., they depend on only primitive random variables). Thus, minimizing $J(\mathbf{g})$ is equivalent to minimizing

$$\tilde{J}(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\left[\sum_{s=0}^{T-1}\left[(\hat{u}_s + K_s\hat{x}_s)^\intercal \Delta_s(\hat{u}_s + K_s\hat{x}_s) + \sum_{i\in\{1,2\}}(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i)^\intercal \tilde{\Delta}_s^i(\tilde{u}_s^i + \tilde{K}_s^i\tilde{x}_s^i)\right]\right].$$

By (A3), R_t is symmetric and positive definite and therefore so is R_t^{ii}. It can be shown recursively that P_t and $\tilde{P}_t$ are symmetric and positive semidefinite. Hence both Δ_t and $\tilde{\Delta}_t^i$ are symmetric and positive definite. Therefore, $\tilde{J}(\mathbf{g}) \geq 0$ with equality if and only if the strategy profile $\mathbf{g}$ is given by Theorem 1.

V. Discussion

The model in [1] consisted of N local controllers and one remote controller. We restricted our discussion to $N = 2$. All steps of our proof apart from Lemma 7 extend trivially to the case of general N. To extend Lemma 7 to the case of general N, one can establish the following result.

Lemma 10: For the system with general N, for any $\gamma = (\gamma^1,\ldots,\gamma^n)$, $\gamma^i \in \{0,1\}$, we have

$$\mathbb{E}\left[\hat{x}_{t+1}^\intercal P_{t+1}\hat{x}_{t+1} + \sum_{i=1}^N(\tilde{x}_{t+1}^i)^\intercal \tilde{P}_{t+1}^i \tilde{x}_{t+1}^i \;\middle|\; \Gamma_{t+1} = \gamma\right]$$
$$= \mathbb{E}\left[(\hat{x}_{t+1}^{\mathrm{OFF}})^\intercal P_{t+1}\hat{x}_{t+1}^{\mathrm{OFF}} + \sum_{i=1}^N(\tilde{x}_{t+1}^{i,\mathrm{OFF}})^\intercal \Lambda_{t+1}^i(\gamma^i)\tilde{x}_{t+1}^{i,\mathrm{OFF}} \;\middle|\; \Gamma_{t+1} = \gamma\right]$$

where

$$\Lambda_{t+1}^i(\gamma^i) = \begin{cases}\tilde{P}_{t+1}^i, & \text{if } \gamma^i = 0\\ P_{t+1}^{ii} & \text{if } \gamma^i = 1.\end{cases} \quad \square$$

Lemma 7 then follows from observing that

$$\sum_{(\gamma^1,\ldots,\gamma^N)\in\{0,1\}^N} \mathbb{P}(\Gamma_t^1 = \gamma^1)\cdots\mathbb{P}(\Gamma_t^N = \gamma^N)\Lambda_{t+1}^i(\gamma^i)$$
$$= \sum_{\gamma^i\in\{0,1\}} \mathbb{P}(\Gamma_t^i = \gamma^i)\Lambda_{t+1}^i(\gamma^i)$$
$$= p^i\tilde{P}_{t+1}^i + (1-p^i)P_{t+1}^{ii}$$
$$= \Pi_{t+1}^i.$$

The proof of Theorem 2 is similar in spirit to the proof of centralized linear quadratic control presented in [3]. However, due to decentralized information and the presence of unreliable communication channels, the specific details are different. As far as we are aware, this is the first article which presents a methodology to synthesize optimal controllers for dynamic *decentralized* control systems without using a dynamic programming or a spectral decomposition argument. In contrast to dynamic programming-based approaches, we sidestep the subtle measurability issues that arise in common information-based dynamic program for continuous state and action spaces. In contrast to spectral decomposition-based arguments, we do not *a priori* restrict attention to linear strategies. We believe that the solution approach presented in this article is interesting in its own right and may be applicable to other decentralized control problems as well.

Acknowledgment

The authors are grateful to Ashutosh Nayyar and Yi Ouyang for helpful feedback.

References

[1] S. M. Asghari, Y. Ouyang, and A. Nayyar, "Optimal local and remote controllers with unreliable uplink channels," *IEEE Trans. Autom. Control*, vol. 64, no. 5, pp. 1816–1831, May 2019.

[2] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.

[3] K. J. Astrom, *Introduction to Stochastic Control Theory*. New York, NY, USA: Academic Press, 1970.

[4] D. Blackwell, "Memoryless strategies in finite-stage dynamic programming," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 863–865, Jun. 1964.

[5] A. Mahajan and S. Tatikonda, "An algorithmic approach to identify irrelevant information in sequential teams," *Automatica*, vol. 61, pp. 178–191, 2015.