J Syst Sci Complex (2025) 38(1): 210–237

Best Linear Controllers for Decentralized Linear Quadratic Systems*

AFSHARI Mohammad · MAHAJAN Aditya

DOI: 10.1007/s11424-025-4505-9
Received: 16 October 2024 / Revised: 12 January 2025

Abstract In this paper, the authors revisit decentralized control of linear quadratic (LQ) systems. Instead of imposing an assumption that the process and observation noises are Gaussian, the authors assume that the controllers are restricted to be linear. The authors show that the multiple decentralized control models, the form of the best linear controllers is identical to the optimal controllers obtained under the Gaussian noise assumption. The main contribution of the paper is the solution technique. Traditionally, optimal controllers for decentralized LQ systems are identified using dynamic programming, maximum principle, or spectral decomposition. The authors present an alternative approach which is based by combining elementary building blocks from linear systems, namely, completion of squares, state splitting, static reduction, orthogonal projection, (conditional) independence of state processes, and decentralized estimation.

Keywords Decentralized estimation, decentralized stochastic control, linear quadratic systems, team theory.

1 Introduction

Optimal control of linear systems with quadratic cost (henceforth referred to as LQ systems) is one of the most popular areas of Systems and Control. Such models are popular because dynamical systems arising in various application domains can be approximated to have linear dynamics; moreover, minimizing the energy used to control such systems naturally corresponds to a cost that is quadratic in the state and control. But, another reason for the appeal of such models is that the optimal controllers are easy to implement because they satisfy a separation property, highlighted below.

AFSHARI Mohammad (Corresponding author)
Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, Atlanta 30332-0530*, USA.*
Email: mafshari@gatech.edu.
MAHAJAN Aditya
*Department of Electrical and Computer Engineering, McGill University, Montreal H*3*A* 0*G*4*, Canada.*
Email: aditya.mahajan@mcgill.ca.
*This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant under Grant No. RGPIN-2021-0351.
◇ *This paper was recommended for publication by Editor HUANG Minyi.*

The optimal controller of an LQ is a linear function of the controller's estimate of the state of the system. The gain of such a feedback controller can be computed based on a solution of a backward Riccati equation. When the process and the observation noises are Gaussian, the state estimate can be recursively updated based on the solution of a forward Riccati equation, where the forward and backward Riccati equations can be solved separately. This is known as the separation between estimation and control. See [1] for an overview.

However, the situation is drastically different in decentralized control (also called team theory). Decentralized control or team problems can be classified as static or dynamic: A team problem is called *static* if the observations of an agent do not depend on the past actions of any agent (including itself); otherwise the problem is called *dynamic*. Static team problems were first analyzed by Radner[2] who showed that when all the system variables are jointly Gaussian and the the cost is quadratic, the optimal decentralized control laws are linear and the corresponding gains can be obtained by solving a linear system of equations. However, the situation is drastically different for dynamic teams.

In a seminal paper, Witsenhausen[3] showed that for a two-stage system with linear dynamics, quadratic cost, and Gaussian disturbance, non-linear strategies can outperform the best linear strategies when the agent at stage 2 does not have access to all the information that was available to the agent at stage 1. A similar counterexample for longer horizons is presented in [4]. A partial resolution was provided by Ho and Chu[5], who showed that there is no loss of optimality in restricting attention to linear strategies when the system has partially nested information structure. The result was generalized in [6] to stochastic nested information structures. However, these results do not provide a way to identify sufficient statistics for the optimal control for general models.

An alternative approach is to a priori restrict attention to linear (or affine) control strategies. There are two challenges in finding the best linear controllers. The first challenge is that the optimization problem for finding the best linear controllers may not be convex in general. It may be converted into a convex model matching problem only when the sparsity pattern of the plant and the controller have a specific structure such as funnel causality[7], quadratic invariance[8], or their variations[9]. The second challenge is that the best linear controller may not have a finite-dimensional representation, as was illustrated by Whittle and Rudge[10] for a completely decentralized controller.

In spite of these challenges, there are several positive results in decentralized control[11–21] where explicit formulas for the optimal controllers are derived. We refer the reader to [22, 23] for a detailed literature review. In almost all of this existing literature, it is assumed that the process and observation noises have a Gaussian distribution.

In an essay on the use of probability theory in Systems and Control, Willems[24] had used the example of filtering to argue that there are two interpretations of the use of probability in Systems and Control. The first is *prescriptive*: For instance, in Kalman (and Wiener) filtering it is assumed that the underlying physics of the model is such that the noise processes are Gaussian. The second is *descriptive*: For instance, in least squares filtering it is assumed that the signal processing is restricted to be linear without making any assumptions on the distribution of the noise process. Both approaches give the same filtering equations but their justifications and the guarantees provided by them are different.

Inspired by [24], we revisit decentralized LQ systems. Rather than imposing an assumption that the distribution of the noise processes is Gaussian, we assume that the controller is restricted to be linear function of the data, and seek to identify the best linear controller. Our main contribution is to present an elementary approach to identify the best linear controller as an alternative to dynamic programming, maximum principle, and spectral factorization methods commonly used in the literature.

This paper is dedicated to Prof. Peter E. Caines on his 80th birthday. Peter is an esteemed colleague, an encouraging and uplifting mentor, and a role model. The approach presented in this paper combines one of Peter's favorite results — Viewing optimal estimation through the lens of orthogonal projection in Hilbert space — With fundamental ideas of linear systems (namely, state splitting and completion of squares) and probability theory (conditional independence of dynamical systems).

Notations We use the standard notation of stochastic control where x denotes the state of a system and u denotes the control input. Moreover, w denotes the process noise and v denotes the observation noise. Usually, subscripts indicate time and superscripts indicate the index of the subsystem/agent. So x_t^i denotes the state of state of subsystem i at time t. Similar notation holds for other variables as well. Given vectors x, y, z, we use $\text{vec}(x, y, z)$ as a short-hand notation for $[x^{\mathrm{T}}, y^{\mathrm{T}}, z^{\mathrm{T}}]^{\mathrm{T}}$. Given vectors x_1, x_2, $\cdots$, x_t, we use $x_{1:t}$ as a short-hand notation for $\text{vec}(x_1, \cdots, x_t)$.

We use $\mathbb{R}$ to denote the set of real numbers and $\mathbb{E}[\cdot]$ to denote expectation of a random variable. For random variables w, x, y, z defined on a common probability space, we use the notation $x \perp\!\!\!\perp y \perp\!\!\!\perp z$ to denote that (x, y, z) are independent and $x \perp\!\!\!\perp y \perp\!\!\!\perp z \mid w$ to denote that (x, y, z) are independent given w.

Given matrices A, B, Q, R, and P of appropriate dimensions, we define the *Riccati* update operator as

$$\mathcal{R}(P, A, B, Q, R) = Q + A^{\mathrm{T}} P A - A^{\mathrm{T}} P B (R + B^{\mathrm{T}} P B) B^{\mathrm{T}} P A,$$

and the feedback gain operator as

$$\mathcal{G}(P, A, B, R) = (R + B^{\mathrm{T}} P B)^{-1} B^{\mathrm{T}} P A.$$

Moreover, given matrices C, Σ^v, Σ of appropriate dimensions, we define the filtering gain opertor as

$$\mathcal{F}(\Sigma, C, \Sigma^v) = \Sigma C^{\mathrm{T}} (\Sigma^v + C \Sigma C^{\mathrm{T}})^{-1}.$$

2 Background on Linear Filtering

2.1 Linear Estimation

Let x and y be random variables defined on a common probability space that are zero mean and have finite variance. Let $\mathcal{L}(y)$ denote the linear subspace spanned by y. We use $\mathbb{L}[x \mid y]$ to denote the best linear unbiased estimator (BLUE) of x given y, i.e.,

$$\mathbb{L}[x \mid y] = \operatorname*{arg\,min}_{\widehat{x}\in\mathcal{L}(y)} \mathbb{E}[\|x-\widehat{x}\|^2].$$

A standard result in least square filtering (see [25, Theorem 3.2.1]) is that

$$\mathbb{L}[x \mid y] = Ky, \quad \text{where } K = \operatorname{cov}(x,y)\operatorname{var}(y)^{-1}. \tag{1}$$

When x and y are jointly Gaussian then $\mathbb{L}[x \mid y] = \mathbb{E}[x \mid y]$ but, in general, they are different. Immediate implication of (1) is that the error $x - \mathbb{L}[x \mid y]$ is orthogonal to y, i.e., for any $z \in \mathcal{L}(y)$,

$$\mathbb{E}[(x-\mathbb{L}[x \mid y])z^{\mathrm{T}}] = 0 \quad \text{and} \quad \mathbb{E}[(x-\mathbb{L}[x \mid y])^{\mathrm{T}}z] = 0. \tag{2}$$

2.2 Linear Filtering

Consider an autonomous linear system with state $x \in \mathbb{R}^{d_x}$ and output $y \in \mathbb{R}^{d_y}$ which starts at a known initial state x_1 and evolves as follows:

$$x_{t+1} = Ax_t + w_t, \quad y_t = Cx_t + v_t,$$

where (A, C) are matrices of appropriate dimension and $\{w_t\}_{t\geq 1}$ and $\{v_t\}_{t\geq 1}$ are process and observation noise processes. We assume that the random variables $\{x_1, w_1, w_2, \cdots, v_1, v_2, \cdots\}$ are independent random variables that are zero mean and have finite variance. Let Σ_1^x denote the variance of x_1 and Σ_t^w and Σ_t^v, $t \geq 1$, denote the variance of w_t and v_t, respectively.

Let $\widehat{x}_t \coloneqq \mathbb{L}[x_t \mid y_{1:t}]$ denote the best linear estimator of the state x_t given the outputs $y_{1:t}$. Then, the estimate $\widehat{x}_t$ can be updated recursively as follows:

$$\widehat{x}_t = \widehat{x}_{t|t-1} + \Sigma_{t|t-1}C^{\mathrm{T}}(C\Sigma_{t|t-1}C^{\mathrm{T}} + \Sigma_t^v)^{-1}(y_t - C\widehat{x}_{t|t-1}), \tag{3}$$

$$\widehat{x}_{t|t} = A\widehat{x}_{t-1|t-1} + \Sigma_{t|t-1}C^{\mathrm{T}}(C\Sigma_{t|t-1}C^{\mathrm{T}} + \Sigma_t^v)^{-1}(y_t - CA\widehat{x}_{t-1|t-1}), \tag{4}$$

$$\widehat{x}_t = A\widehat{x}_{t-1} + L_t(y_t - CA\widehat{x}_{t-1}), \tag{5}$$

where $L_t = \mathcal{F}(\Sigma_t, C, \Sigma^v)$ and the covariance matrices Σ_t are precomputable and given by the forward Riccati equation

$$\Sigma_{t+1} = \mathcal{R}(\Sigma_t, A^{\mathrm{T}}, C^{\mathrm{T}}, \Sigma_t^w, \Sigma_t^v),$$

with $\Sigma_1 = \Sigma_1^x$.

When the noise processes $\{w_t\}_{t\geq 1}$ and $\{v_t\}_{t\geq 1}$ are jointly Gaussian, then the least squares estimate is optimal over all (possibly non-linear) estimators and the update equation above coincides with Kalman filtering equation.

3 Centralized Linear Quadratic Regulator Under Output Feedback

In this section, we revisit optimal centralized linear quadratic regulation by a single agent with output feedback. This is a classical result[1, 26]. Our motivation for presenting a self-contained proof is two-fold. First, instead of assuming that the process and observation noise processes are Gaussian, we do not impose any assumption on the distribution of the noise process; rather, we assume that the agent is restricted to linear processing. Second, instead of proving the result using the standard dynamic programming argument, we present an alternative proof which introduces four of the critical blocks that we use in our solution framework: Completion of squares, state splitting, static reduction, and orthogonal projection.

3.1 System Model and Problem Formulation

Consider a discrete-time stochastic dynamical system that runs for a finite horizon T. Let $x_t \in \mathbb{R}^{d_x}$ denote the state, $u_t \in \mathbb{R}^{d_u}$ denote the control input, and $y_t \in \mathbb{R}^{d_y}$ denote the output. We assume that the system starts from an initial state x_1 and for $t \geq 1$ evolves as

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad y_t = Cx_t + v_t, \tag{6}$$

where (A, B, C) are matrices of appropriate dimension. We assume that the primitive random variables $\{x_1, w_1, \cdots, w_{T-1}, v_1, \cdots, v_{T-1}\}$ are independent random variables that are zero mean and have finite variance. We use Σ_1^x to denote the variance of x_1 and Σ_t^w and Σ_t^v to denote the variance of w_t and v_t, respectively.

Remark 3.1 In the above model, the assumption that the matrices A, B, and C are time-invariant is made for notational simplicity. The results generalize to time-varying A, B, and C in a natural manner.

Information structure: We assume that an agent observes the output of the system and chooses the control input. The information I_t available to the agent at time t is given by

$$I_t = \{y_{1:t}, u_{1:t-1}\}.$$

This information structure is typically called output feedback in the literature.

Admissible control strategies: The controller chooses the control input as a *linear*[†] function of its information. In particular, we assume that the control input is chosen as

$$u_t = g_t(I_t), \tag{7}$$

where g_t is a linear function and is called the *control law* at time t. The collection $g := (g_1, \cdots, g_{T-1})$ is called the (linear) *control strategy.*

System performance and control objective: For time $t \in \{1, \cdots, T-1\}$, the system incurs a per-step cost

$$c_t(x_t, u_t) = x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t \tag{8}$$

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

$$c_T(x_T) = x_T^{\mathrm{T}} Q_T x_T. \tag{9}$$

[†]In principle, we should consider affine (i.e., linear plus a constant term) controllers. However, since the process and observation noises are zero mean, we can show that the constant term in an optimal affine strategy will always be zero.

We assume that the matrices $\{Q_1, \cdots, Q_T\}$ are symmetric and positive semi-definite and the matrices $\{R_1, \cdots, R_{T-1}\}$ are symmetric and positive definite.

The performance of any control strategy g is given by

$$J(g) = \mathbb{E}^g\left[\sum_{t=1}^{T-1} c_t(x_t, u_t) + c_T(x_T)\right], \tag{10}$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of the strategy g.

We are interested in the following optimization problem:

Problem 3.2 For the system described above, given the horizon T, system dynamics (A, B, C), the cost matrices $(Q_{1:T}, R_{1:T-1})$, and the noise statistics $\Sigma^w_{1:T-1}$ and $\Sigma^v_{1:T-1}$, choose a *linear* control strategy g to minimize the total expected cost given by (10).

3.2 Building Blocks of the Optimal Solution

In this section, we present four building blocks that form the basis of our approach to solve Problem 3.2.

Block 1: Completion of Squares

By a standard completion of squares argument, we can show the following:

Lemma 3.3 *The performance* (10) *of any control strategy g can be written as*

$$J(g) = \underbrace{\mathbb{E}\left[\sum_{t=1}^{T-1}(u_t + K_t x_t)^{\mathrm{T}}\Delta_t(u_t + K_t x_t)\right]}_{\text{Term I}} + \underbrace{\sum_{t=1}^{T-1}\mathbf{Tr}(\Sigma^w_t P_{t+1}) + x_1^{\mathrm{T}} S_1 x_1}_{\text{Term II}}, \tag{11}$$

where $\Delta_t = R_t + B^{\mathrm{T}} P_{t+1} B$, $K_t = \mathcal{G}(P_{t+1}, A, B, R_t)$, and the matrices $\{P_t\}_{t=1}^T$ are computed backward in time using the following recursion:

$$P_T = Q_T \text{ and for } t \in \{T-1, \cdots, 1\}, \quad P_t = \mathcal{R}(P_{t+1}, A, B, Q_t, R_t).$$

Proof This is a standard result. See, for example, [26, Lemma 6.1]. ∎

When the agent can observe the system state (the so called state feedback setting), then Lemma 3.3 can be used to infer the optimal controller. Observe that Term II in (11) is *control free* (i.e., does not depend on the choice of the control actions). Therefore, minimizing $J(g)$ is equivalent to minimizing Term I of (11). Since R_t is positive definite and we can recursively show that P_t is positive semi-definite, we have that $\Delta_t = R_t + B^{\mathrm{T}} P_{t+1} B$ is positive definite. Therefore, Term I of (11) is a sum of squares. Choosing

$$u_t = -K_t x_t \tag{12}$$

sets Term I to its minimum value of 0. Hence, in the case of state feedback, the strategy (12) is optimal. However, the above argument does not work for output feedback.

In the rest of this section, we describe three additional blocks which help in generalizing the solution approach for state feedback described above to the output feedback setting of Problem 3.2.

Block 2: State Splitting

The dynamical system of (6) consists of two inputs: The control input u_t and the stochastic inputs (w_t, v_t). We exploit the fact that the system dynamics are linear and split the system into two components: a controlled component with initial state $x_1^g = 0$ which is driven by the controlled input as follows:

$$x_{t+1}^g = Ax_t^g + Bu_t, \quad y_t^g = Cx_t^g,$$

and a stochastic component with initial state $x_1^s = x_1$ which is driven by the stochastic inputs as follows:

$$x_{t+1}^s = Ax_t^s + w_t, \quad y_t^s = Cx_t^s + v_t.$$

Due to linearity of the system dynamics, we have $x_t = x_t^g + x_t^s$ and $y_t = y_t^g + y_t^s$. Moreover, the state and output (x_t^s, y_t^s) of the stochastic component are control free (i.e., they do not depend on the control actions).

Block 3: Static Reduction

The term static reduction of an information structure is due to Witsenhausen[27], but the idea has been used earlier in the literature as well (e.g., [5, 28]). Static reduction means cancelling out the impact of the past control actions on the information available to the agent. For the output feedback model being considered here, static reduction implies the following.

Lemma 3.4 *The information structure* $I_t = \{y_{1:t}, u_{1:t-1}\}$ *is equivalent to the information structure*

$$I_t^s = \{y_{1:t}^s\},$$

i.e., for a fixed control strategy both information sets generate the same sigma algebra or, equivalently, they are functions of each other. Furthermore, when the control strategy is affine, the two information sets are linear functions of each other, i.e., $\mathcal{L}(I_t) = \mathcal{L}(I_t^s)$.

Proof A proof of the first part of the result is given in [28]. Since we are interested in showing that both information sets are *linear* functions of each other, we present the complete proof. A similar proof argument also appears in [5] for a decentralized control problem.

Both results are immediate implications of state splitting. In particular, to show that I_t^s is a linear function of I_t observe that state splitting implies that $y_{1:t}^g$ is a linear function of $u_{1:t}$. Therefore, $y_t^s = y_t - y_t^g$ is a linear function of $(y_{1:t}, u_{1:t-1})$. To show the other direction, we use induction to show that $(y_{1:t}, u_{1:t-1})$ is a linear function of $y_{1:t}^s$. At $t = 1$, $y_t = y_t^s$, so u_1 is a linear function of y_1^s. This forms the basis of induction. Now assume that the result is true for $t-1$, i.e., $I_{t-1} = (y_{1:t-1}, u_{1:t-2})$ is a linear function of $y_{1:t-1}^s$. By assumption, u_t is a linear function I_{t-1} and, by the induction hypothesis of $y_{1:t-1}^s$. Hence $(y_{1:t-1}, u_{1:t-1})$ is a linear function of $y_{1:t-1}^s$. Since x_t^g is a linear function of $u_{1:t-1}$, it is a linear function of $y_{1:t-1}^s$. Consequently, y_t^g is a linear function of $y_{1:t-1}^s$, which implies that $y_t = y_t^g + y_t^s$ is a linear function of $y_{1:t}^s$. This proves the induction step. ∎

Block 4: Orthogonal Projection

Define $\widehat{x}_t = \mathbb{L}[x_t \mid I_t]$ and $\widetilde{x}_t = x_t - \widehat{x}_t$. A consequence of orthogonality of the estimation error (2) is the following.

Lemma 3.5 *For any fixed control strategy g, we have*

$$\mathbb{E}[(u_t + K_t x_t)^{\mathrm{T}} \Delta_t (u_t + K_t x_t)] = \mathbb{E}[(u_t + K_t \widehat{x}_t)^{\mathrm{T}} \Delta_t (u_t + K_t \widehat{x}_t)] + \mathbb{E}[(K_t \widetilde{x}_t)^{\mathrm{T}} \Delta_t (K_t \widetilde{x}_t)].$$

Proof Since $x_t = \widehat{x}_t + \widetilde{x}_t$, we have that

$$\begin{aligned}\mathbb{E}[(u_t + K_t x_t)^{\mathrm{T}} \Delta_t (u_t + K_t x_t)] &= \mathbb{E}[(u_t + K_t \widehat{x}_t)^{\mathrm{T}} \Delta_t (u_t + K_t \widehat{x}_t)] + \mathbb{E}[(K_t \widetilde{x}_t)^{\mathrm{T}} \Delta_t (K_t \widetilde{x}_t)] \\ &\quad + 2\mathbb{E}[(u_t + K_t \widehat{x}_t)^{\mathrm{T}} \Delta_t (K_t \widetilde{x}_t)].\end{aligned}$$

Since both $u_t, \widehat{x}_t \in \mathcal{L}(I_t)$, we have $u_t + K_t \widehat{x}_t \in \mathcal{L}(I_t)$. Therefore, the third term is zero because the error $\widetilde{x}_t$ is orthogonal to the linear subspace $\mathcal{L}(I_t)$. ∎

A key result which ties state splitting, static reduction, and orthogonal projection together is the following.

Lemma 3.6 *For any fixed control strategy g, we have*

$$\widehat{x}_t = x_t^g + \widehat{x}_t^s, \quad \text{where } \widehat{x}_t^s \coloneqq \mathbb{L}[x_t^s \mid I_t^s]. \tag{13}$$

The estimate $\widehat{x}_t^s$ is the standard linear estimation of an uncontrolled linear system and can be computed recursively as follows

$$\widehat{x}_t^s = A\widehat{x}_{t-1}^s + L_t(y_t^s - CA\widehat{x}_{t-1}^s),$$

where L_t is given by $L_t = \mathcal{F}(\Sigma_t, C, \Sigma_t^v)$ and the covariance matrices $\{\Sigma_t\}_{t\geq 1}$ are precomputable and are given by the forward Riccati equation

$$\Sigma_{t+1} = \mathcal{R}(\Sigma_t, A^{\mathrm{T}}, C^{\mathrm{T}}, \Sigma_t^w, \Sigma_t^v),$$

with $\Sigma_1 = \Sigma_1^x$.

An implication of (13) *is that the estimation error can be simplified as*

$$\widetilde{x}_t \coloneqq x_t - \widehat{x}_t = x_t^s - \mathbb{L}[x_t^s \mid I_t^s]$$

and is, therefore, control free.

Proof From state splitting, we have that $x_t = x_t^g + x_t^s$. Therefore,

$$\mathbb{L}[x_t \mid I_t] = x_t^g + \mathbb{L}[x_t^s \mid I_t], \tag{14}$$

where we have used the fact that x_t^g is a linear function of $u_{1:t-1}$ (and hence I_t). Now, static reduction implies that $\mathcal{L}(I_t) = \mathcal{L}(I_t^s)$. Therefore, $\mathbb{L}[x_t^s \mid I_t] = \mathbb{L}[x_t^s \mid I_t^s]$. Substituting this in (14) establishes the first result. The second result follows from the definition of $\widetilde{x}_t$ and the fact that both x_t^s and I_t^s are control free. ∎

We can combine the update of x_t^g and $\mathbb{L}[x_t^s \mid I_t^s]$ to write the update of $\widehat{x}_t$ as follows.

Lemma 3.7 *We have that*

$$\widehat{x}_{t+1} = Ax_t + Bu_t + L_t(y_t - C\widehat{x}_t),$$

where L_t *is as given in the statement of Lemma* 3.6.

Proof This is a simple consequence of the definitions and the update equations for x_t^g and $\mathbb{L}[x_t^s \mid I_t^s]$. In particular,

$$\begin{aligned}\widehat{x}_{t+1} &= \widehat{x}_{t+1}^g + \mathbb{L}[x_{t+1}^s \mid I_{t+1}^s] \\ &= Ax_t^g + Bu_t + A\mathbb{L}[x_t^s \mid I_t^s] + L_t(y_t^s - C\mathbb{L}[x_t^s \mid I_t^s]) \\ &= A(x_t^g + \mathbb{L}[x_t^s \mid I_t^s]) + Bu_t + L_t(y_t - y_t^g - C\mathbb{L}[x_t^s \mid I_t^s]) \\ &= Ax_t + Bu_t + L_t(y_t - C\widehat{x}_t).\end{aligned}$$

The proof is completed. ∎

3.3 Putting Everything Together

Substituting the result of Lemma 3.5 in Lemma 3.3 we have that the total cost $J(g)$ of a control strategy g can be written as

$$J(g) = \widehat{J}(g) + \widetilde{J},$$

where

$$\widehat{J}(g) = \mathbb{E}\Big[\sum_{t=1}^{T-1}(u_t + K_t\widehat{x}_t)^{\mathrm{T}}\Delta_t(u_t + K_t\widehat{x}_t)\Big]$$

and

$$\widetilde{J} = \mathbb{E}\Big[\sum_{t=1}^{T-1}\widetilde{x}_t^{\mathrm{T}}K_t^{\mathrm{T}}\Delta_t K_t\widetilde{x}_t\Big] + \sum_{t=1}^{T-1}\mathbf{Tr}(\Sigma_t^w P_{t+1}) + x_1^{\mathrm{T}}S_1x_1.$$

From Lemma 3.6, we get that the term $\widetilde{J}$ is control free and is, therefore, not affected by the choice of control strategy g. Thus, we can pick g to minimize the term $\widehat{J}(g)$. By an argument similar to one given after Lemma 3.3 for state feedback, we know that the term $\widehat{J}(g)$ is a sum of squares and choosing $u_t = -K_t\widehat{x}_t$ sets $\widehat{J}(g)$ to its minimum value of zero.

We summarize the main result

Proposition 3.8 *The best linear controller for Problem* 3.2 *is given by*

$$u_t = -K_t\widehat{x}_t,$$

where the gain K_t *is chosen as* $K_t = \mathcal{G}(P_{t+1}, A, B, R_t)$ *and the matrices* $\{P_t\}_{t=1}^T$ *are computed backward in time using the following recursion:*

$$P_T = Q_T \text{ and for } t \in \{T-1, \cdots, 1\}, \quad P_t = \mathcal{R}(P_{t+1}, A, B, Q_t, R_t).$$

Furthermore, the best linear estimate $\widehat{x}_t$ *is initialized as* $\widehat{x}_1 = x_1$ *and is recursively updated as*

$$\widehat{x}_t = A\widehat{x}_{t-1} + Bu_{t-1} + L_t(y_t - CA\widehat{x}_{t-1}),$$

where L_t *is given by* $L_t = \mathcal{F}(\Sigma_t, C, \Sigma_t^v)$ *and the covariance matrices* $\{\Sigma_t\}_{t\geq 1}$ *are precomputable and are given by the forward Riccati equation*

$$\Sigma_{t+1} = \mathcal{R}(\Sigma_t, A^{\mathrm{T}}, C^{\mathrm{T}}, \Sigma_t^w, \Sigma_t^v),$$

with $\Sigma_1 = \Sigma_1^x$.

So far, we have shown how the problem of optimal centralized control can be solved by a combination of four elementary blocks. These four blocks, while useful in the decentralized setting, are not sufficient. In the next sections, we present two additional building blocks that are specific to the decentralized setting. We present the simplest models where these blocks are used.

4 Additional Building Blocks for Multi-Agent Systems

4.1 Optimal Decentralized Control of Dynamically Decoupled Subsystems

4.1.1 System Model and Problem Formulation

Consider a decentralized control system consisting of n subsystems indexed by the set $\mathcal{N} := \{1, \cdots, n\}$ that runs for a finite horizon T. Let $x_t^i \in \mathbb{R}^{d_x^i}$ and $u_t^i \in \mathbb{R}^{d_u^i}$, $i \in \mathcal{N}$, denote the state and control input of subsystem i at time t. We use $x_t = \mathrm{vec}(x_t^1, \cdots, x_t^n)$ and $u_t = \mathrm{vec}(u_t^1, \cdots, u_t^n)$ to denote the global state and control inputs of the system. The subsystems are dynamically decoupled but coupled via cost. In particular, the system starts at an initial state x_1 and each subsystem i, $i \in \mathcal{N}$, evolves as follows

$$x_{t+1}^i = Ax_t^i + Bu_t^i + w_t^i, \quad y_t^i = Cx_t^i + v_t^i,$$

where $(A^i, B^i, C^i)_{i\in\mathcal{N}}$ are system matrices of the appropriate dimension. We assume that the primitive variables $\{x_1^i, w_1^i, \cdots, w_{T-1}^i, v_1^i, \cdots, v_{T-1}^i\}_{i\in\mathcal{N}}$ are independent random variables that are zero mean and have finite variance. We use $\Sigma_1^{x^i}$, $i \in \mathcal{N}$, to denote the variance of x_1^i and $\Sigma_t^{w^i}$ and $\Sigma_t^{v^i}$, $i \in \mathcal{N}$, $t \geq 1$, to denote the variance of w_t^i and v_t^i, respectively.

Information structure: The system has a completely decentralized information structure. The information available to agent i at time t is given by

$$I_t^i = \{y_{1:t}^i, u_{1:t-1}^i\}.$$

Admissible control strategies: As in the centralized setting, we assume that each agent chooses its control input as a *linear* function of the information available to it. In particular,

$$u_t^i = g_t^i(I_t^i), \tag{15}$$

where g_t^i is a linear function and is called the *control law of agent* i at time t; the collection $g^i = (g_1^i, \cdots, g_{T-1}^i)$ is called the (linear) *control strategy of agent* i, and $g = (g^1, \cdots, g^n)$ is called the (linear) *control strategy of the system*.

System performance and control objective: The subsystems are coupled via the cost. In particular, for time $t \in \{1, \cdots, T-1\}$, the system incurs a per-step cost

$$c_t(x_t, u_t) = x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t \tag{16}$$

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

$$c_T(x_T) = x_T^{\mathrm{T}} Q_T x_T. \tag{17}$$

We assume that the matrices $\{Q_1, \cdots, Q_T\}$ are symmetric and positive semi-definite and the matrices $\{R_1, \cdots, R_{T-1}\}$ are symmetric and positive definite. We will sometimes consider Q_t and R_t matrices in a block form as follows:

$$Q_t = \begin{bmatrix} Q_t^{11} & \cdots & Q_t^{1n} \\ \vdots & \ddots & \vdots \\ Q_t^{n1} & \cdots & Q_t^{nn} \end{bmatrix} \quad \text{and} \quad R_t = \begin{bmatrix} R_t^{11} & \cdots & R_t^{1n} \\ \vdots & \ddots & \vdots \\ R_t^{n1} & \cdots & R_t^{nn} \end{bmatrix}.$$

The performance of a strategy g is given by

$$J(g) = \mathbb{E}\left[\sum_{t=1}^{T-1} c_t(x_t, u_t) + c_T(x_T)\right], \tag{18}$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of the strategy g.

We are interested in the following optimization problem:

Problem 4.1 For the system described above, given the horizon T, system dynamics $(A^i, B^i, C^i)_{i\in\mathcal{N}}$, the cost matrices $(Q_{1:T}, R_{1:T-1})$, and the noise statistics $\{\Sigma^{w^i}_{1:T}, \Sigma^{v^i}_{1:T-1}\}_{i\in\mathcal{N}}$, choose a *linear* control strategy g to minimize the total expected cost given by (18).

The solution to Problem 4.1 relies on establishing an independence property of the state, which we present below in its simplest form. We will later present a generalization of this property.

Block 5: Independence of the State, Output, and Control Processes

The main idea is the following.

Lemma 4.2 *For any fixed control strategy g, we have*

$$(x^1_{1:T}, y^1_{1:T}, u^1_{1:T}) \perp\!\!\!\perp (x^2_{1:T}, y^2_{1:T}, u^2_{1:T}) \perp\!\!\!\perp \cdots \perp\!\!\!\perp (x^n_{1:T}, y^n_{1:T}, u^n_{1:T}).$$

Proof The proof follows from induction. For $t = 1$, the components of $(x^i_1)_{i\in\mathcal{N}}$ and $(y^i_1)_{i\in\mathcal{N}}$ are independent by assumption. Since u^i_1 is a function of y^i_1, the components of $(u^i_1)_{i\in\mathcal{N}}$ are independent. This forms the basis of induction. We now assume that the result is true for $T = t$. Now consider $T = t+1$. Since we assumed that the components of $(x^i_t)_{i\in\mathcal{N}}$ are independent, the form of the dynamics and the assumptions on the noise imply that the components of $(y^i_t)_{i\in\mathcal{N}}$ are independent. Combined with the induction hypothesis, this implies that the components of $(I^i_t)_{i\in N}$ are independent. Therefore, the components of $(u^i_t)_{i\in\mathcal{N}}$, which are functions of I^i_t are also independent. This completes the induction step. ∎

A key implication of Lemma 4.2 is the following.

Lemma 4.3 *For $t \in \{1, \cdots, T-1\}$, the expected per-step cost can be written as*

$$\mathbb{E}[c_t(x_t, u_t)] = \mathbb{E}\bigg[\sum_{i \in \mathcal{N}} c_t^i(x_t^i, u_t^i)\bigg], \quad \textit{where } c_t^i(x_t^i, u_t^i) = (x_t^i)^{\mathrm{T}} Q_t^{ii} x_t^i + (u_t^i)^{\mathrm{T}} R_t^{ii} u_t^i, \quad i \in \mathcal{N}$$

and, for the terminal time T, the expected terminal cost can be written as

$$\mathbb{E}[c_T(x_T)] = \mathbb{E}\bigg[\sum_{i \in \mathcal{N}} c_T^i(x_T^i)\bigg], \quad \textit{where } c_T^i(x_T^i) = (x_T^i)^{\mathrm{T}} Q_T^{ii} x_T^i, \quad i \in \mathcal{N}.$$

Proof Note that

$$\mathbb{E}[x_t^{\mathrm{T}} Q_t x_t] = \mathbb{E}\bigg[\sum_{i \in \mathcal{N}} (x_t^i)^{\mathrm{T}} Q_t^{ii} x_t^i\bigg] + 2\mathbb{E}\bigg[\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N} \setminus \{i\}} (x_t^i)^{\mathrm{T}} Q_t^{ij} x_t^j\bigg]$$

and

$$\mathbb{E}[u_t^{\mathrm{T}} R_t u_t] = \mathbb{E}\bigg[\sum_{i \in \mathcal{N}} (u_t^i)^{\mathrm{T}} R_t^{ii} u_t^i\bigg] + 2\mathbb{E}\bigg[\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N} \setminus \{i\}} (u_t^i)^{\mathrm{T}} R_t^{ij} u_t^j\bigg].$$

The result then follows from observing that in both these expressions, the cross terms are zero because the state and the control inputs are zero mean and independent across subsystems (due to Lemma 4.2). ∎

4.1.2 Solution of Problem 4.1

Lemma 4.3 implies that the total expected cost under strategy g can be written as

$$J(g) = \sum_{i \in \mathcal{N}} J^i(g^i),$$

where

$$J^i(g^i) = \mathbb{E}\bigg[\sum_{t=1}^{T-1} c_t^i(x_t^i, u_t^i) + c_T^i(x_T^i)\bigg].$$

Thus, the decentralized control problem with decoupled dynamics and independent noise is effectively equivalent to a decentralized control problem with decoupled dynamics and decoupled cost. Therefore, $J(g)$ can be minimized by separately choosing g^i to minimize $J^i(g^i)$ for each $i \in \mathcal{N}$. Each of these optimization problems is a centralized optimal control problem and can be solved in the same manner as Problem 3.2. Thus, the optimal control strategy is given as follows.

Proposition 4.4 *The best linear controller for Problem* 4.1 *is given by*

$$u_t^i = -K_t^i \widehat{x}_t^i, \quad i \in \mathcal{N},$$

where the gain K_t^i for $i \in N$ is chosen as $K_t^i = \mathcal{G}(P_{t+1}^i, A^i, B^i, R_t^{ii})$ and the matrices $\{P_t^i\}_{t=1}^T$ are computed backward in time using the following recursion: For each $i \in \mathcal{N}$,

$$P_T^i = Q_T^{ii} \textit{ and for } t \in \{T-1, \cdots, 1\}, \quad P_t^i = \mathcal{R}(P_{t+1}^i, A^i, B^i, Q_t^{ii}, R_t^{ii}).$$

Furthermore, the best linear estimate $\widehat{x}_t^i$ *is initialized as* $\widehat{x}_1^i = x_1^i$ *and is recursively updated as*

$$\widehat{x}_t^i = A^i \widehat{x}_{t-1}^i + B^i u_{t-1}^i + L_t^i (y_t^i - C^i A^i \widehat{x}_{t-1}^i),$$

where L_t^i *is given by* $L_t^i = \mathcal{F}(\Sigma_t^i, C^i, \Sigma_t^{v^i})$ *and the covariance matrices* $\{\Sigma_t^i\}_{t \geq 1}$ *are precomputable and are given by the forward Riccati equation*

$$\Sigma_{t+1}^i = \mathcal{R}(\Sigma_t^i, (A^i)^{\mathrm{T}}, (C^i)^{\mathrm{T}}, \Sigma_t^{w^i}, \Sigma_t^{v^i}),$$

with $\Sigma_1^i = \Sigma_1^{x^i}$.

4.2 Best Decentralized Linear Estimation to Minimize Team Mean Squared Error

Consider a system with n agents that are indexed by a set $\mathcal{N} = \{1, \cdots, n\}$. The agents are interested in estimating a state $x \in \mathbb{R}^{d_x}$. Each agent i, $i \in \mathcal{N}$, observe a local measurement $y^i \in \mathbb{R}^{d_y^i}$; in addition, all agents observe a common measurement $y^0 \in \mathbb{R}^{d_y^0}$. We use $\mathcal{N}_0$ to denote the set $\{0, 1, \cdots, n\}$.

The variables $(x, y^0, y^1, \cdots, y^n)$ are random variables defined on a common probability space, are zero mean, and have finite variance. For $i, j \in \mathcal{N}_0$, we use Θ^i to denote $\mathrm{cov}(x, y^i)$ and Σ^{ij} to denote $\mathrm{cov}(y^i, y^j)$.

Agent i, $i \in \mathcal{N}$, generates an estimate $\widehat{z}^i \in \mathbb{R}^{d_z^i}$ according to a *linear* estimation rule $\widehat{z}^i = g^i(y^0, y^i)$. The collection $g = (g^1, \cdots, g^n)$ is called the estimation strategy of the system. Let $\widehat{z} = \mathrm{vec}(z^1, \cdots, \widehat{z}^n)$ denote the estimates generated by all agents.

The performance of an estimation strategy g is given by an expected cost of estimation error given by

$$J(g) = \mathbb{E}[c(x, \widehat{z})], \quad \text{where } c(x, \widehat{z}) = \sum_{i \in \mathcal{N}} \sum_{i \in \mathcal{N}} (L^i x - \widehat{z}^i) S^{ij} (L^j x - \widehat{z}^j), \tag{19}$$

where $\{L^i\}_{i \in \mathcal{N}}$ and $\{S^{ij}\}_{i,j \in \mathcal{N}}$ are weight matrices of appropriate dimension. We assume that the matrix S given by

$$S = \begin{bmatrix} S^{11} & \cdots & S^{1n} \\ \vdots & \ddots & \vdots \\ S^{n1} & \cdots & S^{nn} \end{bmatrix}$$

is positive definite.

The above model of decentralized estimation was considered in [29] under the assumption that $(x, y^0, \cdots, y^n)$ are jointly Gaussian. The model considered above is also effectively the same as the static team problem considered in [2], again under the assumption that the random variables are jointly Gaussian.

In the model above, we do not assume that the random variables are jointly Gaussian. Instead of imposing assumptions on the joint distribution of the random variables, we assume that the estimates are linear function of the measurements. Our main result is to show that this distinction does not change the nature of the solution.

As in [29], we define three auxiliary variables:

- The *common (linear) estimate of state* x given the common measurement y^0 of all agents. We denote this estimate by $\widehat{x}^0$ and it is equal to $\mathbb{L}[x \mid y^0]$.
- All agent's *common (linear) estimate of agent i's measurement* y^i given the common measurement y^0. We denote this estimate by $\widehat{y}^i$ and it is equal to $\mathbb{L}[y^i \mid y^0]$.
- The *innovation in local measurement of agent* i with respect to the common measurement. We denote this innovation by $\widetilde{y}^i$ and it is equal to $y^i - \widehat{y}^i$.

We also define $\widehat{\Theta}^i$ to denote the covariance $\operatorname{cov}(x, \widetilde{y}^i)$ and $\widehat{\Sigma}^{ij}$ to denote the covariance $\operatorname{cov}(\widetilde{y}^i, \widetilde{y}^j)$. Following a proof argument very similar to that given in [29], we can show the following.

Proposition 4.5 *The best linear strategy that minimizes the team mean-squared error defined in* (19) *is given by*

$$\widehat{z}^i = L^i \widehat{x}^0 + F^i \widetilde{y}^i, \quad \forall i \in \mathcal{N}, \tag{20}$$

where the gains $\{F^i\}_{i \in \mathcal{N}}$ *satisfy the following system of matrix equations:*

$$\sum_{j \in \mathcal{N}} \Big[S^{ij} F^j \widehat{\Sigma}^{ji} - S^{ij} L^j \widehat{\Theta}^i \Big] = 0, \quad \forall i \in \mathcal{N},$$

which has a unique solution when $\widehat{\Sigma}^{ii}$ *is positive definite.*

The above result forms our final building block for decentralized control. We will refer to it as decentralized estimation block.

5 Solution of Some Multi-Agent Decentralized Control Problems

In this section, we show how the building blocks described earlier can be combined to provide simple solutions to some multi-agent decentralized control problems.

5.1 Best Linear Decentralized Control of Multi-Agent Systems with One-Step Delayed Sharing

5.1.1 System Model and Problem Formulation

Consider a decentralized multi-agent system with n systems, indexed by the set $\mathcal{N} = \{1, \cdots, n\}$, that runs for a finite horizon T. Let $x_t \in \mathbb{R}^{d^x}$ denote the state of the system, $y_t^i \in \mathbb{R}^{d_i^y}$ denote the observation of agent i, and $u_t^i \in \mathbb{R}^{d_i^u}$ denote the control action of agent t. We will use $y_t = \operatorname{vec}(y_t^1, \cdots, y_t^n)$ and $u_t = \operatorname{vec}(u_t^1, \cdots, u_t^n)$ to denote the set of all observations and all control actions, respectively, of all agents.

The system starts at an initial state x_1 and the state evolves as

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \quad y_t^i = C_t^i x_t + v_t^i, \quad i \in \mathcal{N}.$$

We assume that the variables $(x_1, w_1, \cdots, w_{T-1}, \{v_1^i, \cdots, v_{T-1}^i\}_{i \in \mathcal{N}})$ are independent random variables that are zero mean and have finite variance. We use Σ_1^x to denote the variance of x_1 and Σ_t^w and $\Sigma_t^{v^i}$, $t \geq 1$, to denote the variance of w_t and v_t^i, $i \in \mathcal{N}$, respectively.

As usual, we assume that for $t \in \{1, \cdots, T-1\}$ the system incurs a per-step cost given by

$$c_t(x_t, u_t) = x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t$$

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

$$c_T(x_T) = x_T^{\mathrm{T}} Q_T x_T.$$

We assume that the matrices $\{Q_1, \cdots, Q_T\}$ are symmetric and positive semi-definite and the matrices $\{R_1, \cdots, R_{T-1}\}$ are symmetric and positive definite.

Information structure: We assume that each agent observes its local observations and control inputs as well as one-step delayed observations and controls of all other agents. Thus, the information I_t^i available to agent i at time t is given by

$$I_t^i = \{y_{1:t-1}, u_{1:t-1}, y_t^i\}.$$

This information structure is called one-step delayed sharing. It was proposed by Witsenhausen[30]. Under the assumption that the primitive random variables are jointly Gaussian, the optimal solution has been proposed by various authors including[20, 31, 32]. In contrast, we do not impose any assumption on the distribution of the primitive random variables; rather we restrict attention to linear control strategies.

Admissible control strategies: Each agent chooses its control input as a *linear* function of its information, i.e.,

$$u_t^i = g_t^i(I_t^i), \quad i \in \mathcal{N}, \tag{21}$$

where g_t^i is a linear function and called the *control law of agent i at time t*. The collection $g^i := (g_1^i, \cdots, g_{T-1}^i)$ is called the (linear) *control strategy of agent i* and $g = (g^1, \cdots, g^n)$ is called the (linear) *control strategy of all agents*.

System performance and control objective: The performance of any control strategy g is given by

$$J(g) = \mathbb{E}^g\Bigg[\sum_{t=1}^{T-1} c_t(x_t, u_t) + c_T(x_T)\Bigg], \tag{22}$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of the strategy g.

We are interested in the following optimization problem:

Problem 5.1 For the system described above, given the horizon T, system dynamics $(A, B, C^1, \cdots, C^n)$, the cost matrices $(Q_{1:T}, R_{1:T-1})$, and the noise statistics $\Sigma^w_{1:T-1}$ and $\{\Sigma^{v^i}_{1:T-1}\}_{i\in\mathcal{N}}$, choose a *linear* control strategy g to minimize the total expected cost given by (22).

5.1.2 Solution of Problem 5.1

We now show how to solve Problem 5.1 using the different building blocks that we have presented earlier.

- **Completion of squares.** By completion of squares, we can argue that minimizing $J(g)$ is equivalent to minimizing

$$J^{\text{equiv}}(g) \coloneqq \mathbb{E}\Bigg[\sum_{t=1}^{T-1} (u_t + K_t x_t)^{\mathrm{T}} \Delta_t (u_t + K_t x_t)\Bigg],$$

where K_t and Δ_t are as described in Lemma 3.3.

- **State splitting.** Following the idea of state splitting for centralized control, we split the state and output processes into two parts: A controlled part (x_t^g, y_t^g) driven by the control input u_t and a stochastic part (x_t^s, y_t^s) driven by the stochastic inputs (w_t, u_t). In particular we have $x_1^g = 0$, $x_1^s = x_1$ and

$$x_{t+1}^g = A x_t^g + B u_t, \quad y_t^g = C x_t^g,$$
$$x_{t+1}^s = A x_t^s + w_t, \quad y_t^s = C x_t^s + v_t.$$

We also split the control actions similar to what is done for decentralized estimation. For that matter, we first define common information as

$$I_t^c = \bigcap_{i \in \mathcal{N}} I_t^i = \{y_{1:t-1}, u_{1:t-1}\}. \tag{23}$$

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

$$I_t^{i,\ell} = I_t^c \setminus I_t^i = \{y_t^i\}. \tag{24}$$

Now, as we did for decentralized estimation, we split the control action into two components: A common control u_t^c defined as $\mathbb{L}[u_t \mid I_t^c]$ and local control $u_t^\ell = u_t - u_t^c$.

- **Static reduction.** Following arguments similar to static reduction for the single agent setting, we can show that the original information structure is equivalent to

$$I_t^{i,s} = \{y_t^{i,s}, y_{1:t-1}^s\}.$$

In particular, $\mathcal{L}(I_t^i) = \mathcal{L}(I_t^{i,s})$, $i \in \mathcal{N}$.

A direct result of the above equation is that the common information in the original information structure is equivalent to the common information in the static reduction, which is given by

$$I_t^{c,s} = \{y_{1:t-1}^s\}.$$

The implication of static reduction is that in both conditional expectations and linear estimation we can replace conditioning on I_t^c by $I_t^{c,s}$.

- **Orthogonal projection.** Define $\widehat{x}_t = \mathbb{L}[x_t \mid I_t^c]$ and $\widetilde{x}_t = x_t - \widehat{x}_t$. By construction, we have that $\mathbb{L}[u_t^\ell \mid I_t^c] = 0$. Since x_t^g is a linear function of $u_{1:t-1}$, which is part of I_t^c, we have

$$\widehat{x}_t = x_t^g + \widehat{x}_t^s, \quad \text{where } \widehat{x}_t^s \coloneqq \mathbb{L}[x_t^s \mid I_t^{c,s}],$$

and consequently,

$$\widetilde{x}_t = x_t - \widehat{x}_t = x^s_t - \widehat{x}^s_t.$$

Recall that x^s_t is a control free processes; therefore $\widetilde{x}_t$ is control free. We can show that the estimate $\widehat{x}^s_t$ equals $A\widehat{x}^s_{t-1|t-1}$, where $\widehat{x}^s_{t-1|t-1}$ can be recursively updated as follows:

$$x^s_{t|t} = Ax^s_{t-1|t-1} + L_t(y_t - CA\widehat{x}^s_{t-1|t-1}),$$

where $L_t = \mathcal{F}(\Sigma_t, C, \Sigma^v_t)$ and the covariance matrices $\{\Sigma_t\}_{t=1}^{T-1}$ can be precomputed as follows:

$$\Sigma_1 = \Sigma^x_1 \quad \text{and for} \quad t \in \{1, \cdots, T-1\} \;\; \Sigma_{t+1} = \mathcal{R}(\Sigma_t, A^{\mathrm{T}}, C^{\mathrm{T}}, \Sigma^w_t, \Sigma^v_t).$$

Then, a consequence of orthogonality of the estimation error is the following: For any fixed control strategy g, we have

$$\mathbb{E}[(u_t+K_tx_t)^{\mathrm{T}}\Delta_t(u_t+K_tx_t)] = \mathbb{E}[(u^c_t+K_t\widehat{x}_t)^{\mathrm{T}}\Delta_t(u^c_t+K_t\widehat{x}_t)]+\mathbb{E}[(u^\ell_t+K_t\widetilde{x}_t)^{\mathrm{T}}\Delta_t(u^\ell_t+K_t\widetilde{x}_t)].$$

Consequently, the total cost $J^{\text{equiv}}(g)$ can be written as

$$J^{\text{equiv}}(g) = J^c(g) + \sum_{t=1}^{T-1} J^\ell_t(g),$$

where

$$J^c(g) = \mathbb{E}\biggl[\sum_{t=1}^{T-1}(u^c_t + K_t\widehat{x}_t)^{\mathrm{T}}\Delta_t(u^c_t + K_t\widehat{x}_t)\biggr]$$

and

$$J^\ell_t(g) = \mathbb{E}[(u^\ell_t + K_t\widetilde{x}_t)^{\mathrm{T}}\Delta_t(u^\ell_t + K_t\widetilde{x}_t)].$$

Putting everything together. Minimizing $J^{\text{equiv}}(g)$ is equivalent to minimizing the sum of $J^c(g)$ and $\sum_{t=1}^{T-1} J^\ell_t(g)$. Observe that both these terms are sum of squares. The first term $J^c(g)$ takes its minimum value of 0 when u^c_t is chosen to be $-K_t\widehat{x}_t$. The second term $\sum_{t=1}^{T-1} J^\ell_t(g)$ is a sum of decentralized estimation problems. For each t, minimizing $J^\ell_t(g)$ is a decentralized estimation problem, where the state is $\widetilde{x}_t = \widetilde{x}^s_t$, each agent has a common observation of $y^s_{1:t-1}$ and agent i has a local observation of $y^{i,s}_t$. Therefore, by Proposition 4.5, the optimal controller is given by

$$u^{i,\ell}_t = F^i_t\widetilde{y}^{i,s}_t, \quad \text{where} \;\; \widetilde{y}^{i,s}_t = y^{i,s}_t - \mathbb{L}[y^{i,s}_t \mid y^s_{1:t-1}],$$

where the gains F^i_t are found by solving the following system of linear matrix equation:

$$\sum_{j\in\mathcal{N}}((B^i)^{\mathrm{T}}P_{t+1}B^j + R^{ij})F^j_t((C^j)^{\mathrm{T}}\Sigma_tC^i + \Sigma^{ij,v}_t) = (B^i)^{\mathrm{T}}P_{t+1}A\Sigma_t(C^i)^{\mathrm{T}}, \quad \forall i \in \mathcal{N},$$

where $\Sigma^{ji,v}_t = 0$ for $i \neq j$ is the ji-element of Σ^v_t. To summarize, let K^i_t denote the i-th row of K_t. Then, the optimal control action at each agent is given by

$$u^i_t = -K^i_t\widehat{x}_t + F^i_t\widetilde{y}^{i,s}_t, \quad i \in \mathcal{N}.$$

The form of the optimal controller derived above is identical to the form of the optimal controller derived in [20, 31, 32] under the assumption that the process and observation noises are Gaussian.

5.2 Optimal Decentralized Control of System with Major and Minor Agents and State Sharing

5.2.1 System Model and Problem Formulation

Consider a decentralized control system with one major and n minor agents that evolve in discrete time over a finite horizon T. We use index 0 to indicate the major agent and use index i, $i \in \mathcal{N} := \{1, \cdots, n\}$, to index the minor agents. We also define $\mathcal{N}_0 := \{0, 1, \cdots, n\}$ as the set of all agents. Let $x_t^i \in \mathbb{R}^{d_x^i}$ and $u_t^i \in \mathbb{R}^{d_u^i}$ denote the state and control input of agent $i \in \mathcal{N}_0$.

All agents have linear dynamics. The dynamics of the major agent is not affected by the minor agents. In particular, the initial state of the major agent is given by x_1^0, and for $t \geq 1$, the state of the major agent evolves according to

$$x_{t+1}^0 = A^{00} x_t^0 + B^{00} u_t^0 + w_t^0, \tag{25}$$

where $\{w_t^0\}_{t \geq 1}$, $w_t^0 \in \mathbb{R}^{d_x^0}$, is a noise process.

In contrast, the dynamics of the minor agents are affected by the state of the major agent. For agent $i \in \mathcal{N}$, the initial state is given by x_1^i, and for $t \geq 1$, the state evolves according to

$$x_{t+1}^i = A^{ii} x_t^i + A^{i0} x_t^0 + B^{ii} u_t^i + B^{i0} u_t^0 + w_t^i, \tag{26}$$

where $\{w_t^i\}_{t \geq 1}$, $w_t^i \in \mathbb{R}^{d_x^i}$, is a noise process.

We assume that all primitive random variables-the initial states $\{x_1^0, x_1^1, \cdots, x_1^n\}$, and the process noises $\{w_1^i, \cdots, w_T^i\}_{i \in \mathcal{N}_0}$, are defined on a common probability space, are independent and have zero mean and finite variance. We use $\Sigma_1^{x^i}$ to denote the variance of the initial state x_1^i and use $\Sigma_t^{w^i}$ and $\Sigma_t^{v^i}$ to denote the variance of w_t^i and v_t^i, respectively, $i \in \mathcal{N}$.

Let $x_t = \text{vec}(x_t^0, \cdots, x_t^n)$ denote the state of the system, $u_t = \text{vec}(u_t^0, \cdots, u_t^n)$ denote the control actions of all controllers, and $w_t = \text{vec}(w_t^0, \cdots, w_t^n)$ denote the system disturbance. Then the dynamics (25) and (26) can be written in vector form as

$$x_{t+1} = A x_t + B u_t + w_t, \tag{27}$$

where

$$A = \begin{bmatrix} A_{00} & 0 & 0 & \cdots & 0 \\ A_{10} & A_{11} & 0 & \cdots & 0 \\ A_{20} & 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_{n0} & 0 & \cdots & 0 & A_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} B_{00} & 0 & 0 & \cdots & 0 \\ B_{10} & B_{11} & 0 & \cdots & 0 \\ B_{20} & 0 & B_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B_{n0} & 0 & \cdots & 0 & B_{nn} \end{bmatrix}.$$

Note that A and B are sparse block lower triangular matrices.

Information structure: We assume that the major agent observes its own state, while minor agent i, $i \in N$, observes the state of both the major agent and its own state. Thus, the information I_t^0 available to the major agent is given by

$$I_t^0 \coloneqq \{x_{1:t}^0, u_{1:t-1}^0\}, \tag{28}$$

while the information I_t^i available to minor agent i, $i \in \mathcal{N}$, is given by

$$I_t^i \coloneqq \{x_{1:t}^0, x_{1:t}^i, u_{1:t-1}^0, u_{1:t-1}^i\}. \tag{29}$$

The information structure addressed in this section is studied in the literature and is commonly referred to as the two-agent problem. See [16] and [33, 34] and the references therein.

Admissible control strategies: At time t, controller $i \in \mathcal{N}_0$ chooses control action u_t^i as a *linear* function of the information I_t^i available to it, i.e.,

$$u_t^i = g_t^i(I_t^i), \quad i \in \mathcal{N}_0,$$

where g_t^i is a linear function and is called the *control law of controller* i, $i \in \mathcal{N}_0$, at time t. The collection $g^i \coloneqq (g_1^i, \cdots, g_T^i)$ is called the *control strategy of controller* i and $g = (g^0, \cdots, g^n)$ is called the *control strategy of the system.*

System performance and control objective: At time $t \in \{1, \cdots, T-1\}$, the system incurs a per-step cost of

$$c_t(x_t, u_t) = x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t \tag{30}$$

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

$$c_T(x_T) = x_T^{\mathrm{T}} Q_T x_T. \tag{31}$$

It is assumed that Q and Q_T are symmetric and positive semi-definite and R is symmetric and positive definite.

The performance of any strategy g is given by

$$J(g) = \mathbb{E}\Big[\sum_{t=1}^{T-1} c_t(x_t, u_t) + c_T(x_T)\Big], \tag{32}$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of the strategy g.

We are interested in the following optimization problem.

Problem 5.2 For the system described above, given the horizon T, system dynamics (A, B), the cost matrices $(Q_{1:T}, R_{1:T-1})$, and the noise statistics $\{\Sigma_{1:T-1}^{w^i}, \Sigma_{1:T-1}^{v^i}\}_{i\in\mathcal{N}}$, choose a control strategy g to minimize the total expected cost given by (32).

5.2.2 Solution of Problem 5.2

We now show how to solve Problem 5.2 using the different building blocks that we have presented earlier.

- **Common information based state splitting.** We combine the idea of using common information based decomposition of the state presented in decentralized estimation with that of state splitting. We first define common information as

$$I_t^c \coloneqq \bigcap_{i \in \mathcal{N}_0} I_t^i = \{x_{1:t}^0, u_{1:t-1}^0\} = I_t^0. \tag{33}$$

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

$$I_t^{0,\ell} \coloneqq I_t^0 \setminus I_t^c = \emptyset, \tag{34a}$$

$$I_t^{i,\ell} \coloneqq I_t^i \setminus I_t^c = \{x_{1:t}^i, u_{1:t-1}^i\}, \quad i \in \mathcal{N}. \tag{34b}$$

Now, as we did for decentralized estimation, we split the control action into two components: A common control u_t^c defined as $\mathbb{L}[u_t \mid I_t^c]$ and local control u_t^ℓ defined as $u_t - u_t^c$.

Finally, based on the above splitting of control actions, we split the state into three components: Common component of the state x_t^c which is driven by common control u_t^c, local component of the state x_t^ℓ which is driven by the local control u_t^ℓ, and the stochastic component of the state x_t^s which is driven by the stochastic input w_t. In particular, we have $x_1^c = 0$, $x_1^\ell = 0$, $x_1^s = x_1$, and

$$x_{t+1}^c = Ax_t^c + Bu_t^c, \quad x_{t+1}^\ell = Ax_t^\ell + Bu_t^\ell, \quad x_{t+1}^s = Ax_t^s + w_t.$$

By construction, the stochastic component is control free (i.e., does not depend on the control actions). By linearity of the dynamics, we have $x_t = x_t^c + x_t^\ell + x_t^s$. Moreover, we define $x_t^{c,i}$, etc. such that

$$x_t^c = \text{vec}(x_t^{0,c}, \cdots, x_t^{n,c}), \quad x_t^\ell = \text{vec}(x_t^{0,\ell}, \cdots, x_t^{n,\ell}), \quad x_t^s = \text{vec}(x_t^{0,s}, \cdots, x_t^{n,s}).$$

Note that by construction $(x_t^{0,c}, u_t^{0,c}) = (x_t^0, u_t^0)$; therefore, $x_t^{0,\ell} = x_t^{0,s} = 0$ and $u_t^{0,\ell} = 0$.

- **Static reduction.** Following arguments similar to static reduction for the single agent setting, we can show that the original information structure is equivalent to

$$I_t^{s,0} = \{x_{1:t}^{0,s}\}, \quad I_t^{s,i} = \{x_{1:t}^{0,s}, x_{1:t}^{i,s}\}.$$

In particular, $\mathcal{L}(I_t^i) = \mathcal{L}(I_t^{i,s})$, $i \in \mathcal{N}_0$.

The implication of static reduction is that in both conditional expectations and linear estimation we can replace conditioning on I_t^0 by $I_t^{0,s}$.

- **Conditional independence of state and control processes.** We generalize the idea of independence of state and control processes to establish conditional independence of state and control processes given the common information $I_t^c = I_t^0$. In particular, for any control strategy g, we have

$$(x_{1:t}^1, u_{1:t}^1) \perp\!\!\!\perp (x_{1:t}^2, u_{1:t}^2) \perp\!\!\!\perp \cdots \perp\!\!\!\perp (x_{1:t}^n, u_{1:t}^n) \mid I_t^0.$$

Moreover, since $\mathcal{L}(I_t^0) = \mathcal{L}(I_t^{0,s})$ (thus, both I_t^0 and $I_t^{0,s}$ are linear functions of each other), we can replace I_t^0 in the conditioning with $I_t^{0,s}$.

- **Orthogonal projection.** Define $\widehat{x}_t = \mathbb{L}[x_t \mid I_t^0]$ and $\widetilde{x}_t = x_t - \widehat{x}_t$. In order to simplify $\widehat{x}_t$, we observe that by construction, we have that $\mathbb{L}[u_t^\ell \mid I_t^c] = 0$. Moreover, since $I_t^{0,s}$ is equivalent to $w_{1:t-1}^0$, for any $i \in \mathcal{N}$ and $\tau < t$, we have

$$\mathbb{L}[u_\tau^\ell \mid I_t^0] = \mathbb{L}[u_\tau^\ell \mid I_t^s] = \mathbb{L}[u_\tau^\ell \mid w_{1:t-1}^0] = \mathbb{L}[u_\tau^\ell \mid w_{1:\tau-1}^0] = \mathbb{L}[u_\tau^\ell \mid I_\tau^{0,s}] = 0.$$

Consequently, since x_t^ℓ is a linear function of $u_{1:t-1}^\ell$, we have

$$\mathbb{L}[x_t^\ell \mid I_t^0] = \mathbb{L}[x_t^\ell \mid I_t^{0,s}] = 0.$$

Therefore,

$$\mathbb{L}[x_t \mid I_t^0] = \mathbb{L}[x_t^c + x_t^\ell + x_t^s \mid I_t^0] = x_t^c + \mathbb{L}[x_t^s \mid I_t^0] = x_t^c + \mathbb{L}[x_t^s \mid I_t^{0,s}].$$

We write this as

$$\widehat{x}_t = x_t^c + \widehat{x}_t^s, \quad \text{where } \widehat{x}_t^s \coloneqq \mathbb{L}[x_t^s \mid I_t^{0,s}].$$

Recall that x_t^s is a control free processes. We can show that the estimate $\widehat{x}_t^s$ can be recursively updated as

$$\widehat{x}_{t+1}^s = A\widehat{x}_t^s + \operatorname{vec}(w_t^0, 0, \cdots, 0).$$

Therefore, the update of $\widehat{x}_t$ simplifies to

$$\widehat{x}_t = A\widehat{x}_t + Bu_t^c + \operatorname{vec}(w^0, 0, \cdots, 0) \tag{35}$$

and, consequently,

$$\widetilde{x}_t = x_t - \widehat{x}_t = \begin{bmatrix} 0 \\ A^{11}\widetilde{x}_t^1 + B^1 u_t^{1,\ell} \\ \vdots \\ A^{nn}\widetilde{x}_t^n + B^n u_t^{n,\ell} \end{bmatrix} + \begin{bmatrix} 0 \\ w_t^1 \\ \vdots \\ w_t^n \end{bmatrix}. \tag{36}$$

Now, by an argument similar to Lemma 3.5, we have

$$\mathbb{E}[x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t] = \mathbb{E}[\widehat{x}_t^{\mathrm{T}} Q_t \widehat{x}_t + (u_t^c)^{\mathrm{T}} R_t u_t^c] + \mathbb{E}[\widetilde{x}_t^{\mathrm{T}} Q_t \widetilde{x}_t + (u_t^\ell)^{\mathrm{T}} R_t u_t^\ell]. \tag{37}$$

Furthermore, conditional independence of state and control processes implies that

$$(\widetilde{x}_t^1, u_t^{1,\ell}) \perp\!\!\!\perp (\widetilde{x}_t^2, u_t^{2,\ell}) \perp\!\!\!\perp \cdots \perp\!\!\!\perp (\widetilde{x}_t^n, u_t^{n,\ell}) \mid I_t^0.$$

Therefore, similar to Lemma 4.3, we have that the second term of (37) can be written as

$$\mathbb{E}[\widetilde{x}_t^{\mathrm{T}} Q_t \widetilde{x}_t + (u_t^\ell)^{\mathrm{T}} R_t u_t^\ell] = \mathbb{E}\bigg[\sum_{i \in \mathcal{N}} \Big[(\widetilde{x}_t^i)^{\mathrm{T}} Q_t^{ii} \widetilde{x}_t^i + (u_t^{i,\ell})^{\mathrm{T}} R_t^{ii} u_t^{i,\ell}\Big]\bigg].$$

Consequently, the total cost $J(g)$ can be written as

$$J(g) = J^c(g) + \sum_{i \in \mathcal{N}} J^{i,\ell}(g),$$

where

$$J^c(g) = \sum_{t=1}^{T-1} \mathbb{E}[\widetilde{x}_t^{\mathrm{T}} Q_t \widetilde{x}_t + (u_t^\ell)^{\mathrm{T}} R_t u_t^\ell]$$

and

$$J^{i,\ell}(g) = \sum_{t=1}^{T-1} \mathbb{E}[(\widetilde{x}_t^i)^{\mathrm{T}} Q_t^{ii} \widetilde{x}_t^i + (u_t^{i,\ell})^{\mathrm{T}} R_t^{ii} u_t^{i,\ell}].$$

- **Completion of squares.** Now, we separately perform completion of squares of $J^c(g)$ and $J^i(g)$ and show that minimizing $J^c(g)$ is equivalent to minimizing

  $$\widetilde{J}^c(g) = \mathbb{E}\Big[\sum_{t=1}^{T-1}(u_t^c + K_t \widehat{x}_t)^{\mathrm{T}} \Delta_t (u_t^c + K_t \widehat{x}_t)\Big],$$

  where $\Delta_t = R_t + B^{\mathrm{T}} P_{t+1}^c B$, $K_t = \mathcal{G}(P_{t+1}^c, A, B, R_t)$, and the matrices $\{P_t^c\}_{t=1}^T$ are computed backwards in time using the following recursion:

  $$P_T^c = Q_T \text{ and for } t \in \{T-1, \cdots, 1\}, \quad P_t^c = \mathcal{R}(P_t^c, A, B, Q_t, R_t).$$

  Moreover, minimizing $J^{i,\ell}(g)$ is equivalent to minimizing

  $$\widetilde{J}^{i,\ell}(g) = \mathbb{E}\Big[\sum_{t=1}^{T-1}(u_t^{i,\ell} + K_t^{i,\ell} \widetilde{x}_t^i)^{\mathrm{T}} \Delta_t^{i,\ell} (u_t^{i,\ell} + K_t^i \widetilde{x}_t^i)\Big],$$

  where $\Delta_t^{i,\ell} = R_t^{ii} + (B^i i)^{\mathrm{T}} P_{t+1}^{i,\ell} B^{ii}$, $K_t^{i,\ell} = \mathcal{G}(P_{t+1}^{i,\ell}, A^{ii}, B^{ii}, R_t^{ii})$, and the matrices $\{P_t^{i,\ell}\}_{t=1}^T$ are computed backwards in time using the following recursion: For each $i \in \mathcal{N}$,

  $$P_T^{i,\ell} = Q_T^{ii} \text{ and for } t \in \{T-1, \cdots, 1\}, \quad P_t^{i,\ell} = \mathcal{R}(P_t^{i,\ell}, A^{ii}, B^{ii}, Q_t^{ii}, R_t^{ii}).$$

Putting everything together. Now, to find the optimal solution, observe that both $J^c(g)$ and $J^{i,\ell}(g)$ are sum of squares and they take their minimum value of 0 if we choose

$$u_t^c = -K_t \widehat{x}_t, \quad u_t^{i,\ell} = -K_t^{i,\ell} \widetilde{x}_t.$$

Let K_t^i, $i \in \mathcal{N}_0$, denote the i-th row of K_t. Then, the best linear controllers can be written as follows: The best linear control strategy of the major agent is given by

$$u_t^0 = -K^0 \widehat{x}_t,$$

and at the minor agent i, $i \in \mathcal{N}$, the best linear control strategy is given by

$$u_t^i = -K_t^i \widehat{x}_t - K_t^{i,\ell}(x_t^i - \widehat{x}_t^i).$$

The form of the optimal controller derived above is identical to form of the controller derived in [16] under the assumption that the process and observation noise processes are Gaussian. Generalization of [16] to output feedback are presented in [13, 14, 33, 35].

5.3 Optimal Decentralized Control of System with Remote and Local Controllers and Packet Drop State Sharing

5.3.1 System Model and Problem Formulation

Consider a discrete-time linear dynamical system consisting of 2 controllers: A remote controller and a local controller co-located with the system. The information available to the controllers will be described later. Let $x_t \in \mathbb{R}^{d_x}$ denote the state of the system, $u_t^0 \in \mathbb{R}^{d_u^0}$, denote the control action of remote controller and $u_t^1 \in \mathbb{R}^{d_u^1}$ denote the control action of local controller.

The initial state x_1 of the system is random and the dynamics of the system is given by

$$x_{t+1} = Ax_t + \begin{bmatrix} B^0 & B^1 \end{bmatrix} \begin{bmatrix} u_t^0 \\ u_t^1 \end{bmatrix} + w_t, \tag{38}$$

where $w_t \in \mathbb{R}^{d_x}$ is the process noise and A, B^0, and B^1 are matrices of appropriate dimensions. We assume that random variables $\{x_1, w_0, \cdots, w_{T-1}\}$ are independent and have zero mean and finite variance. We use Σ_1^x to denote the variance of x_1 and Σ_t^w, $t \geq 1$, to denote the variance of w_t.

Let $u_t \coloneqq \text{vec}(u_t^0, u_t^1)$ denote the control actions of the overall system. Then, the system dynamics can be written as

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{39}$$

where B is given by $B = [B^0 \quad B^1]$.

At time t, the local controller perfectly observes the state x_t of the system and sends it to the remote controller over an unreliable packet drop channel. Let $\Gamma_t \in \{0, 1\}$ denote the state of the channel, where $\Gamma_t = 0$ means that the channel is in the OFF state where the transmitted packet gets dropped while $\Gamma_t = 1$ means that the channel is in the ON state where the transmitted packet gets delivered. Thus, Γ_t is a Bernoulli random variable and we denote the packet drop probability $\mathbb{P}(\Gamma_t = 0)$ by p. We assume that the primitive random variables $\{x_0, w_0, \cdots, w_{T-1}, \Gamma_0, \cdots, \Gamma_{T-1}\}$ are independent.

Let z_t denote the output of the channel, i.e.,

$$z_t = f(x_t, \Gamma_t) = \begin{cases} x_t, & \text{if } \Gamma_t = 1, \\ \mathfrak{E}, & \text{if } \Gamma_t = 0, \end{cases} \tag{40}$$

where $\mathfrak{E}$ denotes a dropped packet. It is assumed that there is a perfect channels from the remote controller to the local controller. Using this channel, the remote controller can share z_t and u_{t-1}^r with the local controller. Note that it is possible to recover Γ_t from z_t. Hence, all controllers also have access to Γ_t.

Information structure: Let I^0 and I_t^1 denote the information available to the remote and local controllers, respectively, at time t. We have

$$I_t^0 = \{z_{0:t}, \Gamma_{0:t}, u_{0:t-1}^0\}, \tag{41a}$$
$$I_t^1 = \{x_{0:t}, u_{0:t-1}^1, z_{0:t}, \Gamma_{0:t}, u_{0:t-1}^0\}. \tag{41b}$$

Admissible control strategies: In this section we do not explicitly restrict attention to linear control strategies. We assume that the controllers choose their control action as a measurable function of their observations, i.e.,

$$u_t^0 = g_t^0(I_t^0), \quad u_t^1 = g_t^1(I_t^1), \tag{42}$$

where g_t^0 and g_t^1 are called the *control laws of the remote and local controllers*, respectively. The collections $g^0 = (g_1^0, \cdots, g_{T-1}^0)$ and $g^1 = (g_1^1, \cdots, g_{T-1}^1)$ are called the *control strategies of the remote and local controllers*, respectively and $g = (g^0, g^1)$ is called the *control strategy profile of the system.*

System performance and control objective: The system operates for a finite horizon T. For time $t \in \{1, \cdots, T-1\}$, the system incurs a per-step cost

$$c_t(x_t, u_t) = x_t^{\mathrm{T}} Q_t x_t + u_t^{\mathrm{T}} R_t u_t,$$

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

$$c_T(x_T) = x_T^{\mathrm{T}} Q_T x_T,$$

where Q_t, and R_t are matrices of appropriate dimensions. It is assumed that Q and Q_T are symmetric and positive semi-definite and R is symmetric and positive definite. We also assume that R_t has a block-wise structure given by

$$R_t = \begin{bmatrix} R_t^{00} & R_t^{01} \\ R_t^{10} & R_t^{11} \end{bmatrix}.$$

The performance of a strategy profile g is given by

$$J(g) = \mathbb{E}^g \Big[\sum_{t=0}^{T-1} c_t(x_t, u_t) + c_T(x_T) \Big], \tag{43}$$

where the expectation is with respect to the measure induced on all the system variables by the choice of strategy profile g.

Problem 5.3 For the system described above, given the horizon T, system dynamics (A, B), the cost matrices $(Q_{1:T}, R_{1:T-1})$, and the noise statistics $\Sigma_{1:T-1}^w$ and the packet drop probability p, choose a control strategy g to minimize the total expected cost given by (43).

The above model was considered in [36] where a dynamic programming solution was presented. The solution presented below is adapted from [37].

5.3.2 Solution of Problem 5.3

We now show how to solve Problem 5.3 using the different building blocks that we have presented earlier.

- **Common information based estimates.** Following [38], we define the *common information* I_t^c between agents as
$$I_t^c = I_t^0 \cap I_t^1.$$

The information structure of the model (41) implies that $I_t^c = I_t^0 = \{z_{0:t}, \Gamma_{0:t}, u^0_{0:t-1}\}$. Now we define the common information based "estimates" of the state and control actions and the corresponding "estimation errors" as follows:

$$\widehat{x}_t = \mathbb{E}[x_t \mid I_t^c], \quad \widetilde{x}_t = x_t - \widehat{x}_t, \tag{44}$$

$$\widehat{u}_t = \mathbb{E}[u_t \mid I_t^c], \quad \widetilde{u}_t = u_t - \widehat{u}_t. \tag{45}$$

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

Lemma 5.4 *The state estimates and estimation errors evolve as follows:*

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

and for $t > 0$,

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

Therefore,

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

and for $t > 0$,

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

A proof is presented in [37].

- **Orthogonal projection for per-step cost.** A direct result of the common information based state estimates is the following.

$$\mathbb{E}[x_t^{\mathrm{T}} Q_t x_t] = \mathbb{E}\Big[\widehat{x}_t^{\mathrm{T}} Q_t \widehat{x}_t + (\widetilde{x}_t)^{\mathrm{T}} Q_t \widetilde{x}_t\Big], \tag{46}$$

$$\mathbb{E}[u_t^{\mathrm{T}} R_t u_t] = \mathbb{E}\Big[\widehat{u}_t^{\mathrm{T}} R_t \widehat{u}_t + (\widetilde{u}_t^{\ell})^{\mathrm{T}} R_t \widetilde{u}_t^1\Big]. \tag{47}$$

- **Completion of squares.** Now, we utilize the result of the orthogonal projection for the completion of squares. However, the exact details are slightly different: The completion of squares must take the packet drop nature of the channel and Lemma 5.4 into account. Using such a completion of squares, we obtain the following:

$$\begin{aligned} J(g) =& \mathbb{E}^g \Bigg[\widehat{x}_0^{\mathrm{T}} P_t \widehat{x}_0 + (\widetilde{x}_0)^{\mathrm{T}} \widetilde{P}_0 \widetilde{x}_0 + \sum_{s=0}^{T-1} (\widehat{u}_s + K_s \widehat{x}_s)^{\mathrm{T}} \Delta_s (\widehat{u}_s + K_s \widehat{x}_s) \\ &+ \sum_{s=0}^{T-1} (\widetilde{u}_s + \widetilde{K}_s \widetilde{x}_s)^{\mathrm{T}} \widetilde{\Delta}_s (\widetilde{u}_s + \widetilde{K}_s \widetilde{x}_s) + \sum_{s=0}^{T-1} (w_s)^{\mathrm{T}} \Pi_{t+1} w_s \Bigg], \end{aligned} \tag{48}$$

where $\Delta_s = R_s + B^{\mathrm{T}} P_{s+1} B$, $\widetilde{\Delta}_s = R_s^{11} + (B^1)^{\mathrm{T}} \varPi_{s+1} B^1$, the gains $\{K_t\}_{t=0}^{T-1}$ and $\{\widetilde{K}_t\}_{t=0}^{T-1}$ are given by

$$K_t = \mathcal{G}(P_{t+1}, A, B, R_t), \quad \widetilde{K}_t = \mathcal{G}(\varPi_{t+1}, A, B^1, R_t^{11}),$$

where the matrices $\{P_t\}_{t=1}^{T}$, $\{\varPi_t\}_{t=1}^{T}$, and $\{\widetilde{P}_t\}_{t=1}^{T}$ are given by as follows:

$$P_T = Q_T \text{ and for } t \in \{1, \cdots, T-1\}, \text{ we have } P_t = \mathcal{R}(P_{t+1}, A, B, Q_t, R_t),$$
$$\widetilde{P}_T = Q_T \text{ and for } t \in \{1, \cdots, T-1\}, \text{ we have } \widetilde{P}_t = \mathcal{R}(\varPi_{t+1}, A, B^1, Q_t, R_t^{11}),$$
$$\varPi_t = (1-p) P_t + p \widetilde{P}_t.$$

See [37] for details.

- **Putting everything together.** The optimal control strategy for Problem 5.3 is given by

$$\operatorname{vec}(u_t^0, \widehat{u}_t^1) = -K_t \widehat{x}_t \tag{49}$$

and

$$\widetilde{u}_t^1 = -\widetilde{K}_t \widetilde{x}_t, \tag{50}$$

where the time evolution of $\widehat{x}_t$ and $\widetilde{x}_t$ are given above.

Let K_t^0 and K_t^1 denote the rows of K_t. Then, we have

$$u_t^0 = -K_t^0 \widehat{x}_t, \quad u_t^1 = -K_t^1 \widehat{x}_t - \widetilde{K}_t (x_t - \widetilde{x}_t), \tag{51}$$

which is the same as the optimal controllers derived in [36, 37].

6 Conclusion

In this paper, we revisit decentralized control of multi-agent systems. Instead of identifying the optimal decentralized controllers under the assumption that the process and observation noises are Gaussian, we identify the best linear controller without any restriction on the noise distribution. We present an elementary approach to identify the best linear controller: The fundamental ideas of our approach are completion of squares, state splitting, static reduction of information structure, and orthogonal projection. The approach presented here is not a panacea for all the conceptual challenges in decentralized control. All the models considered in the paper have a partially nested information structure[5], so we know that if the noise processes were Gaussian, then there is no loss of optimality in restricting attention to linear control strategies. Effectively, we derive the same control laws but by providing a *descriptive* justification (that we are limited to use linear controllers) rather than by assuming a *prescriptive* justification (that the underlying physics of the system being modelled is such that the noise processes are Gaussian). Verifying whether the proposed approach works for more general information structures remains an interesting future direction.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Caines P E, *Linear Stochastic Systems*, SIAM, Philadelphia, PA, 2018.

[2] Radner R, Team decision problems, *The Annals of Mathematical Statistics*, 1962, **33**(3): 857–881.

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

[4] Lipsa G M and Martins N C, Optimal memoryless control in Gaussian noise: A simple counterexample, *Automatica*, 2011, **47**(3): 552–558.

[5] Ho Y-C and Chu K-C, Team decision theory and information structures in optimal control problems-Part I, *IEEE Transactions on Automatic Control*, 1972, **17**(1): 15–22.

[6] Yuksel S, Stochastic nestedness and the belief sharing information pattern, *IEEE Transactions on Automatic Control*, 2009, **54**(12): 2773–2786.

[7] Bamieh B and Voulgaris P G, A convex characterization of distributed control problems in spatially invariant systems with communication constraints, *Systems & Control Letters*, 2005, **54**(6): 575–583.

[8] Rotkowitz M and Lall S, A characterization of convex problems in decentralized control, *IEEE Transactions on Automatic Control*, 2005, **50**(12): 1984–1996.

[9] Qi X, Salapaka M V, Voulgaris P G, et al., Structured optimal and robust control with multiple criteria: A convex solution, *IEEE Transactions on Automatic Control*, 2004, **49**(10): 1623–1640.

[10] Whittle P and Rudge J, The optimal linear solution of a symmetric team control problem, *Journal of Applied Probability*, 1974, **11**(2): 377–381.

[11] Gattami A, Generalized linear quadratic control, *IEEE Transactions on Automatic Control*, 2009, **55**(1): 131–136.

[12] Lamperski A and Lessard L, Optimal decentralized state-feedback control with sparsity and delays, *Automatica*, 2015, **58**: 143–151.

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

[14] Lessard L and Nayyar A, Structural results and explicit solution for two-player LQG systems on a finite time horizon, 52*nd IEEE Conference on Decision and Control*, 2013, 6542–6549.

[15] Liang X, Xu J, Wang H, et al., Decentralized output-feedback control with asymmetric one-step delayed information, *IEEE Transactions on Automatic Control*, 2023, **68**(12): 7871–7878.

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

[17] Shah P and Parrilo P A, $\mathcal{H}_2$-optimal decentralized control over posets: A state-space solution for state-feedback, *IEEE Transactions on Automatic Control*, 2013, **58**(12): 3084–3096.

[18] Varaiya P and Walrand J, On delayed sharing patterns, *IEEE Transactions on Automatic Control*, 1978, **23**(3): 443–445.

[19] Wang Y, Xiong J, and Ho D W, Globally optimal state-feedback LQG control for large-scale systems with communication delays and correlated subsystem process noises, *IEEE Transactions on Automatic Control*, 2019, **64**(10): 4196–4201.

[20] Yoshikawa T, Decomposition of dynamic team decision problems, *IEEE Transactions on Automatic Control*, 1978, **23**(4): 627–632.

[21] Yoshikawa T, Dynamic programming approach to decentralized stochastic control problems, *IEEE Transactions on Automatic Control*, 1975, **20**(6): 796–797.

[22] Mahajan A, Martins N C, Rotkowitz M C, et al., Information structures in optimal decentralized control, *IEEE Conf. Decision and Control* (*CDC*), Maui, Hawaii, 2012, 1291–1306.

[23] Yüksel S and Başar T, *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints*, Springer, New York, 2013.

[24] Willems J C, Probability in control?, *European Journal of Control*, 2010, **16**: 436–440.

[25] Kailath T, Sayed A H, and Hassibi B, *Linear Estimation*, Prentice Hall, Upper Saddle River, 2000.

[26] Åström K J, *Introduction to Stochastic Control Theory*, Courier Corporation, Chelmsford, Massachusetts, 2012.

[27] Witsenhausen H S, Equivalent stochastic control problems, *Mathematics of Control Signals and Systems*, 1988, **1**: 3–11.

[28] Kumar P R and Varaiya P, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, SIAM, Philadelphia, PA, 2015.

[29] Afshari M and Mahajan A, Multi-agent estimation and filtering for minimizing team mean-squared error, *IEEE Transactions on Signal Processing*, 2021, **69**: 5206–5221.

[30] Witsenhausen H S, Separation of estimation and control for discrete time systems, *Proceedings of the IEEE*, 1971, **59**(11): 1557–1566.

[31] Kurtaran B and Sivan R, Linear quadratic Gaussian control with one step delay sharing pattern, *IEEE Transactions on Automatic Control*, 1974, **19**: 571–574.

[32] Sandell N and Athans M, Solution of some nonclassical LQG stochastic decision problems, *IEEE Transactions on Automatic Control*, 1974, **19**: 108–116.

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

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

[35] Afshari M and Mahajan A, Decentralized linear quadratic systems with major and minor agents and non-Gaussian noise, *IEEE Transactions on Automatic Control*, 2022, **68**(8): 4666–4681.

[36] Asghari S M, Ouyang Y, and Nayyar A, Optimal local and remote controllers with unreliable uplink channels, *IEEE Transactions on Automatic Control*, 2019, **64**(5): 1816–1831.

[37] Afshari M and Mahajan A, Optimal local and remote controllers with unreliable uplink channels: An elementary proof, *IEEE Transactions on Automatic Control*, 2020, **65**(8): 3616–3622.

[38] Nayyar A, Mahajan A, and Teneketzis D, Decentralized stochastic control with partial history sharing: A common information approach, *IEEE Transactions on Automatic Control*, 2013, **58**(7): 1644–1658.