Decentralized Linear Quadratic Systems With Major and Minor Agents and Non-Gaussian Noise

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

***Abstract*—A decentralized linear quadratic system with a major agent and a collection of minor agents is considered. The major agent affects the minor agents, but not vice versa. The state of the major agent is observed by all agents. In addition, the minor agents have a noisy observation of their local state. The noise process is *not* assumed to be Gaussian. The structures of the optimal strategy and the best linear strategy are characterized. It is shown that the major agent's optimal control action is a linear function of the major agent's minimum mean-squared error (MMSE) estimate of the system state while the minor agent's optimal control action is a linear function of the major agent's MMSE estimate of the system state and a "correction term" that depends on the difference of the minor agent's MMSE estimate of its local state and the major agent's MMSE estimate of the minor agent's local state. Since the noise is non-Gaussian, the minor agent's MMSE estimate is a nonlinear function of its observation. It is shown that replacing the minor agent's MMSE estimate with its linear least mean square estimate gives the best linear control strategy. The results are proved using a direct method based on conditional independence, common-information-based splitting of state and control actions, and simplifying the per-step cost based on conditional independence, orthogonality principle, and completion of squares.**

***Index Terms*—Decentralized linear quadratic systems, decentralized stochastic control, dynamic team theory, separation of estimation and control, non-Gaussian noise.**

I. Introduction

In many modern decentralized control systems, such as self-driving cars, robotics, unmanned aerial vehicles (UAVs), and others, the environment is sensed using vision and Lidar sensors; the raw sensor observations are filtered through a deep-neural-network-based object classifier and the classifier outputs are used as the inputs to the controllers. In such systems, the assumption that the observation noise is Gaussian breaks down. Therefore, the optimal design of such decentralized systems requires understanding the structure of optimal controllers when the observation noise is non-Gaussian.

For centralized control of linear systems with quadratic per-step cost, the classical two-way separation between estimation and control continues to hold even when the observations (and the process noises) are non-Gaussian. In particular, the optimal control action is a linear function of the minimum mean-squared error (MMSE) estimator of the state given the observations and the past actions at the controller. Moreover, the MMSE estimator does not depend on the choice of the control strategy. See [1], [2], and [3] for details.

Although the optimal control action is a linear function of the MMSE estimate, the MMSE estimate is, in general, a nonlinear function of the past observations and actions. Thus, the optimal control action is a nonlinear function of the past observations and the actions. In certain applications, it is desirable to restrict attention to linear control strategies. The best linear strategy is similar to the optimal strategy where the MMSE estimate is replaced by the linear least mean squares (LLMS) estimate.[1] Moreover, the LLMS estimate does not depend on the choice of the control strategy. See [4, Sec. 15.5.3] for details.

In summary, in centralized control of linear quadratic systems with non-Gaussian noise, there is a two-way separation of estimation and control; the optimal control action is a linear function of the MMSE estimate of the state given the data at the controller. The best linear controller has the same structure except the MMSE estimate of the state is replaced by the LLMS estimate. Both the MMSE and LLMS estimators can be computed as functions of sufficient statistics that can be recursively updated.[2] In contrast, the current state of the art in decentralized systems is significantly limited.

In the literature on optimal decentralized control of linear quadratic systems, most papers assume that the noise processes are Gaussian. Even with Gaussian noise, nonlinear policies may outperform the best linear policies [5]; linear strategies are globally optimal only for specific information structures (e.g., partially nested [6] and its variants). Even for systems

Manuscript received 30 July 2021; revised 30 June 2022; accepted 9 September 2022. Date of publication 27 September 2022; date of current version 28 July 2023. This work was supported by the Natural Sciences and Engineering Research Council of Canada Discovery under Grant RGPIN-2016-05165. Recommended by Associate Editor 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@mail.mcgill.ca; adityam@cim.mcgill.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2022.3210049.

Digital Object Identifier 10.1109/TAC.2022.3210049

[1] For linear models driven by uncorrelated noise, the LLMS estimate is the best linear unbiased estimator of the state.

[2] MMSE estimator is the mean of the conditional density, which can be recursively updated via Bayesian filtering; LLMS estimator can be recursively updated via recursive least squares filtering.

with Gaussian noise and partially nested information structures, there is no general method to identify sufficient statistics for the optimal controller; the optimal strategy is known to have a finite-dimensional sufficient statistic only for specific models (e.g., the one-step delayed sharing information structure [7], [8], [9]; asymmetric one-step delayed sharing [10]; chain structures [11]; two-agent problem [12] and its variant [13]). As far as we are aware, there are no existing results on sufficient statistics for optimal decentralized control of linear quadratic systems with output feedback and non-Gaussian noise.

If attention is restricted to linear strategies, the problem of finding the best linear control strategy for a decentralized linear quadratic system is not convex in general but can be converted to a convex problem when the controller and the plant have specific sparsity pattern (funnel causality [14], quadratic invariance [15], and their variants). Even for such models, the best linear control strategy may not have a finite-dimensional sufficient statistic [16]; the best linear strategy is known to have a finite-dimensional sufficient statistic only for specific models (e.g., poset causality [17], two-agent problem [18], [19], [20], [21], [22], [23], [24], [25] and its variants [26], [27], [28]). A general method for identifying sufficient statistics for the best linear strategy in linear quadratic systems with partial history sharing was proposed in [29], but this method did not provide an efficient algorithm to compute all the gains at the controllers.

In this article, we investigate a decentralized control system with a major agent and a collection of minor agents. The agents are coupled in their dynamics as well as cost. In particular, the dynamics are linear; the state and the control actions of the major agent affect the state evolution of all the minor agents but the state and control actions of the minor agents do not affect the state evolution of the major or other minor agents. The cost is an arbitrarily coupled quadratic cost. The information structure is partially nested with partial output feedback. In particular, the major agent perfectly observes its own state while each minor agent perfectly observes the state of the major agent and partially observes its own state. We assume that the process and the observation noises have zero mean and finite variance but do not impose any restrictions on the distribution of the noise processes. We are interested in identifying both the optimal and the best linear control strategy for this model.

There are two motivations for considering this specific model. First, such systems arise in certain applications in decentralized control of UAVs and, for that reason, there has been considerable interest in understanding special cases of such models [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. Variations of this model with weak coupling between the agents have also been considered in the literature on mean-field games [30], [31], [32], [33]. Second, the information structure may be viewed as a "star network," where the major agent is the central hub and the minor agents are on the periphery. Understanding the optimal design of such systems is an important intermediate step in understanding the optimal design of decentralized systems where agents are connected over a general graph.

Even though the information structure of our model is partially nested, we cannot use the results of [6] because the noise processes are not Gaussian. There is information that is commonly known to all agents in our model, however the information structure is not partial history sharing [34]. Hence, we cannot directly use the dynamic programming decomposition of [34] which was derived for models with finite-state and finite action spaces. In addition, the local information at the minor agents is increasing with time. So, we cannot use the method of [29] to identify sufficient statistics.

When there is only one minor agent, our model is similar to the two agent problem considered in [12], [18], [19], [20], [21], [22], [23], and [25]. However, none of these results are directly applicable: in [18], [19], [20], attention is restricted to state feedback; in [22], [23], [25], continuous time systems with output or partial output feedback are considered but attention is restricted to linear strategies; in [12], output feedback is considered but it is assumed that the noise is Gaussian. A model similar to ours has been considered in [28] and [21]. In [28], a continuous time system with major and minor agents with output feedback is considered but it is assumed that there is no cost coupling between the minor agents, the system dynamics is stable, and attention is restricted to linear strategies. In [21], a discrete-time system with a major and a single minor agent is considered but it is assumed that the system dynamics is stable and attention is restricted to linear strategies.

Our first main result is to show that the qualitative features of centralized control of linear quadratic control continue to hold for decentralized control of linear systems with major and minor agents. In particular, we show the following.

1) The optimal control action of the major agent is a linear function of the major agent's MMSE estimate of the state of the entire system. The corresponding gains are determined by the solution of a single "global" Riccati equation that depends on the dynamics and the cost of the entire system.
2) The optimal control action of the minor agent is a linear function of the minor agent's MMSE estimate of its local state and the major agent's MMSE estimate of the local state of the minor agent. The corresponding gains are determined by the solution of two Riccati equations: A "global" Riccati equation that depends on the dynamics and the cost of the entire system and a "local" Riccati equation that depends on the dynamics and the cost of the minor agent.

Moreover, there is a separation between estimation and control. The MMSE estimation strategies of both the major and the minor agents do not depend on the choice of the control strategies. In addition, the choice of the controller gains does not depend on the estimation strategies used by the agents. See Theorem 2 for a precise statement of these results. Note that the MMSE estimator of the major agent is a linear function of the data while the MMSE estimator of the minor agent is a nonlinear function of the data.

Our second main result is to show that the best linear strategy has the same structure as the optimal strategy where the MMSE estimate is replaced by the LLMS estimate. Moreover, the LLMS estimate does not depend on the choice of the control strategy.

We show that both the MMSE and the LLMS estimates can be computed as a function of sufficient statistics that can be updated recursively. In particular, we show that the MMSE estimate at the minor agent is the mean of the conditional density of the state of the minor agent given the past observations. The conditional density can be recursively updated using (nonlinear) Bayesian filtering. The LLMS estimates at the minor agent can be updated using recursive least squares filtering. Note that unlike the results of [12] and [25], the recursive update of both the MMSE and the LLMS estimates do not depend on the Riccati gains.

Finally, we believe that our proof technique might be considered a contribution in its own right. The two most commonly used techniques in decentralized control of linear systems are 1) time-domain dynamic programming decomposition, which is used to identify optimal strategies; and 2) frequency domain decomposition using Youla parameterization, which is used to identify the best linear control strategy. In this article, we present a unified approach to identify both the optimal and the best linear control strategies. Our approach is based on the following:

1) conditional independence of the states of the minor agents given the common information;
2) splitting the state and the control actions based on the common information;
3) simplifying the per-step cost based on conditional independence, orthogonality principle, and completion of squares.

Our approach sidesteps the technical difficulties related to measurability and existence of value functions in dynamic programming. At the same time, unlike the spectral factorization methods, it can be used to identify both the optimal and the best linear control strategy. Given the paucity of positive results in decentralized control, we believe that a new solution approach is of interest.

A. Notation

Given a matrix A, A_{ij} denotes its (i,j)th block element, $A^\intercal$ denotes its transpose, $\operatorname{vec}(A)$ denotes the column vector of A formed by vertically stacking the columns of A. Given a square matrix A, $\operatorname{Tr}(A)$ denotes the sum of its diagonal elements. I_n denotes an $n \times n$ identity matrix. We simply use I when the dimension is clear for context. Given any vector-valued process $\{y(t)\}_{t\geq 1}$ and any time instances t_1, t_2 such that $t_1 \leq t_2$, $y(t_1{:}t_2)$ is a short hand notation for $\operatorname{vec}(y(t_1), y(t_1+1), \ldots, y(t_2))$.

Given random vectors x, y, and z, $\mathbb{E}[x]$ denotes the mean of x, $\mathbb{E}[x|y]$ denotes the conditional mean of random variable x given random variable y, $\operatorname{cov}(x,y)$ denotes the covariance between x and y, and $x \perp\!\!\!\perp y | z$ denotes that x and y are conditionally independent given z.

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

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

$$\begin{aligned}
\mathcal{R}(P,A,B,Q,R) &= Q + A^\intercal P A \\
&\quad - A^\intercal P B (R + B^\intercal P B)^{-1} B^\intercal P A \\
\mathcal{G}(P,A,B,R) &= (R + B^\intercal P B)^{-1} B^\intercal P A \\
\mathcal{K}(P,A,C,\Sigma,\Sigma') &= (A P A^\intercal C^\intercal + \Sigma C^\intercal) \\
&\quad (C A P A^\intercal C^\intercal + C \Sigma C^\intercal + \Sigma')^{-1}
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{F}(P,A,C,\Sigma,\Sigma') &= A P A^\intercal + \Sigma \\
&\quad - K (C A P A^\intercal C^\intercal + C \Sigma C^\intercal + \Sigma') K^\intercal
\end{aligned}$$

where $K = \mathcal{K}(P,A,C,\Sigma,\Sigma')$.

II. Model and Problem Formulation

A. 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, $i \in N := \{1,\ldots,n\}$, to indicate a minor agent. We also define $N_0 := \{0,1,\ldots,n\}$ as the set of all agents. Let $x_i(t) \in \mathbb{R}^{d_x^i}$ and $u_i(t) \in \mathbb{R}^{d_u^i}$ denote the state and control input of agent $i \in N_0$.

1) System Dynamics: 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 $x_0(1)$, and for $t \geq 1$, the state of the major agent evolves according to

$$x_0(t+1) = A_{00} x_0(t) + B_{00} u_0(t) + w_0(t) \tag{1}$$

where $\{w_0(t)\}_{t\geq 1}$, $w_0(t) \in \mathbb{R}^{d_x^0}$, is a noise process.

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

$$x_i(t+1) = A_{ii} x_i(t) + A_{i0} x_0(t) + B_{ii} u_i(t) + B_{i0} u_0(t) + w_i(t) \tag{2}$$

where $\{w_i(t)\}_{t\geq 1}$, $w_i(t) \in \mathbb{R}^{d_x^i}$, is a noise process. Furthermore, the minor agent $i \in N$ generates an output $y_i(t) \in \mathbb{R}^{d_y^i}$ given by

$$y_i(t) = C_{ii} x_i(t) + v_i(t) \quad i \in N \tag{3}$$

where $\{v_i(t)\}_{t\geq 1}$, $v_i(t) \in \mathbb{R}^{d_y^i}$, is a noise process.

Assumption 1: We assume that all primitive random variables—the initial states $\{x_0(1), x_1(1), \ldots, x_n(1)\}$, the process noises $\{w_i(1), \ldots, w_i(T)\}_{i \in N_0}$, and the observation noises $\{v_i(1), \ldots, v_i(T)\}_{i\in N}$ are defined on a common probability space, are independent and have zero mean and finite variance. We use Σ_i^x to denote the variance of the initial state $x_i(1)$, Σ_i^w to denote the variance of the process noise $w_i(t)$ and Σ_i^v to denote the variance of the observation noise $v_i(t)$.

Note that we do not assume that the primitive random variables have a Gaussian distribution. For some of the results, we impose an additional assumption that the primitive random variables have a density.

Assumption 2: All primitive random variables (which are defined on a common probability space) have a joint density. We denote the marginal density of $x_i(1), i \in N_0, w_i(t), i \in N_0$, and $v_i(t), i \in N$, by $\pi_{x_i(1)}$, $\varphi_{i,t}$, and $\nu_{i,t}$, respectively.

Let $x(t) = \operatorname{vec}(x_0(t), \ldots, x_n(t))$ denote the state of the system, $u(t) = \operatorname{vec}(u_0(t), \ldots, u_n(t))$ denote the control actions of all controllers, and $w(t) = \operatorname{vec}(w_0(t), \ldots, w_n(t))$ denote the system disturbance. Then, the dynamics (1) and (2) can be written in vector form as

$$x(t+1) = A x(t) + B u(t) + w(t) \tag{4}$$

where

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

and

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

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

2) Information Structure: The system has partial output feedback: the major agent observes its own state while minor agent i, $i \in N$, observes the state of the major agent and its own output. Thus, the information $I_0(t)$ available to the major agent is given by

$$I_0(t) := \{x_0(1{:}t), u_0(1{:}t-1)\} \tag{5}$$

while the information $I_i(t)$ available to minor agent i, $i \in N$, is given by

$$I_i(t) := \{x_0(1{:}t), y_i(1{:}t), u_0(1{:}t-1), u_i(1{:}t-1)\}. \tag{6}$$

3) Admissible Control Strategies: At time t, controller $i \in N_0$ chooses control action $u_i(t)$ as a function of the information $I_i(t)$ available to it, i.e.,

$$u_i(t) = g_{i,t}(I_i(t)), \quad i \in N_0.$$

The function $g_{i,t}$ is called the control law of controller i, $i \in N_0$, at time t. The collection $g_i := (g_{i,1}, \ldots, g_{i,T})$ is called the control strategy of controller i and $(g_0, \ldots, g_n)$ is called the control strategy of the system.

Let $\mathcal{L}_2(\mathbb{R}^n)$ denote the family of all square integrable random variables, i.e., random variables $Z \in \mathbb{R}^n$ such that $\mathbb{E}[|Z|^2] < \infty$. We consider two classes of control strategies. The first, which we call *general control strategies* and denote by $\mathscr{G}$, is where $g_{i,t}$ is a measurable function that maps $I_i(t)$ to $u_i(t)$ that satisfies the property that for any $I_i(t) \in \mathcal{L}_2(\mathbb{R}^{d_I^i})$, where $d_I^i = t \times (d_x^0 + d_y^i) + (t-1) \times (d_u^0 + d_u^i), i \in N,$ we have $\mathbb{E}[|g_{i,t}(I_i(t))|^2] < \infty$.

The second, which we call *affine control strategies* and denote by $\mathscr{G}_A$, is where $g_{i,t}$ is an affine function that maps $I_i(t)$ to $u_i(t)$.

4) System Performance and Control Objective: At time $t \in \{1, \ldots, T-1\}$, the system incurs a per-step cost of

$$c(x(t), u(t)) = x(t)^\intercal Q x(t) + u(t)^\intercal R u(t) \tag{7}$$

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

$$C(x(T)) = x^\intercal(T) Q_T x(T). \tag{8}$$

It is assumed that Q and Q_T are positive semidefinite and R is positive definite.

The performance of any strategy $(g_0, \ldots, g_n)$ is given by

$$J(g_0, \ldots, g_n) = \mathbb{E}\left[\sum_{t=1}^{T-1} c(x(t), u(t)) + C(x(T))\right] \tag{9}$$

where the expectation is with respect to the joint measure on all the system variables induced by the choice of the strategy $(g_0, \ldots, g_n) \in \mathcal{G}$.

We are interested in the following optimization problems.

Problem 1: In the system described above, choose a general control strategy $(g_0, \ldots, g_n) \in \mathscr{G}$ to minimize the total expected cost given by (9).

The information structure of the model is partially nested [6], but the noise is not Gaussian. So, we cannot assert that there is no loss of optimality in restricting attention to linear strategies. In fact, our main result shows that the optimal policy of Problem 1 is nonlinear. In certain applications, it is desirable to restrict attention to linear strategies. For that reason, we also consider the following optimization problem.

Problem 2: In the system described above, choose an affine strategy $(g_0, \ldots, g_n) \in \mathscr{G}_A$ to minimize the total expected cost given by (9).

B. Roadmap of the Solution Approach

The rest of the article is organized as follows. In Section III, we present several preliminary results to simplify the analysis. These include a common-information-based splitting of state and control actions, a static reduction of the information structure, and establishing conditional independence of the various components of the state. We combine these results to split the per-step cost and then use completion of squares to rewrite the total cost as the sum of three terms: The first depends on the common component of the state and control action, the second depends on the local component of the state and control action, and the third depends on the stochastic component of the state. A key feature of this decomposition is that the third term does not depend on the choice of the control strategy. So, we can focus on the first two terms to find the optimal or the best linear strategy.

Our next step is to use orthogonal projection to simplify the first two terms. In Section IV, we simplify these terms using orthogonality properties of the MMSE estimate and the estimation error; in Section V, we simplify these terms using orthogonality properties of the LLMS estimate and the estimation error. The final expression of the total cost in both cases is such that the optimal and best linear strategies can be identified by inspection.

III. Preliminary Results

A. Common Information Based State and Control Splitting

Following [34], we split the information at each agent into common and local information. The common information is defined as

$$I^c(t) := \bigcap_{i\in N_0} I_i(t) = \{x_0(1{:}t), u_0(1{:}t-1)\} = I_0(t). \quad (10)$$

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

$$I_0^\ell(t) := I_0(t) \setminus I^c(t) = \emptyset \quad (11a)$$

$$I_i^\ell(t) := I_i(t) \setminus I^c(t) = \{y_i(1{:}t), u_i(1{:}t-1)\}. \quad (11b)$$

Thus, although there is common information among the agents, the system *does not* have a partially history sharing information structure [34] because the local information at agent $i \in N$ is increasing with time. Hence, the approach in [29] and [34] cannot be used directly.

Instead, we combine the idea of common information with a standard idea in linear systems and split the state and the control actions into different components based on the common information. First, we split the control action into two components: $u(t) = u^c(t) + u^\ell(t)$, where

$$u^c(t) = \mathbb{E}[u(t)|I^c(t)], \qquad u^\ell(t) = u(t) - u^c(t). \quad (12)$$

We refer to $u^c(t)$ and $u^\ell(t)$ as the common control and the local control, respectively.

Based on the above splitting of control actions, we split the state into three components: $x(t) = x^c(t) + x^\ell(t) + x^s(t)$, where

$$x^c(1) = 0, \qquad x^c(t+1) = Ax^c(t) + Bu^c(t) \quad (13a)$$

$$x^\ell(1) = 0, \qquad x^\ell(t+1) = Ax^\ell(t) + Bu^\ell(t) \quad (13b)$$

$$x^s(1) = x(1), \qquad x^s(t+1) = Ax^s(t) + w(t). \quad (13c)$$

We refer to $x^c(t)$, $x^\ell(t)$, $x^s(t)$ as the common, local, and stochastic components of the state, respectively. Note that the stochastic component is control free (i.e., does not depend on the control actions).

Based on the above splitting of state, we split the observations of agent $i \in N$ into three components as well: $y_i(t) = y_i^c(t) + y_i^\ell(t) + y_i^s(t)$, where

$$y_i^c(t) = C_{ii}x_i^c(t) \quad (14a)$$

$$y_i^\ell(t) = C_{ii}x_i^\ell(t) \quad (14b)$$

$$y_i^s(t) = C_{ii}x_i^s(t) + v_i(t). \quad (14c)$$

We refer to $y_i^c(t)$, $y_i^\ell(t)$, and $y_i^s(t)$ as the common, local, and stochastic components of the observation, respectively. Note that since $x_i^s(t)$ is control free, so is $y_i^s(t)$.

Lemma 1: For any strategy $g \in \mathscr{G}$, the split components of the state and the control actions satisfy the following properties:

P1) $u_0^\ell(t) = 0$.
P2) $x_0^\ell(t) = 0$.
P3) $\mathbb{E}[u_i^\ell(t)|I^c(t)] = 0, i \in \{1,\dots,n\}$.
P4) $\mathbb{E}[u^c(t)^\intercal M u^\ell(t)] = 0$, where M is any matrix of compatible dimensions.
P5) $\mathbb{E}[u_i^\ell(t)] = 0, i \in \{1,\dots,n\}$.
P6) $\mathbb{E}[x^c(t)|I^c(t)] = x^c(t)$.

The proof is presented in Appendix A.

B. Static Reduction

We define the following information structure that does not depend on the control strategy:

$$I_0^s(t) = \{x_0^s(1{:}t)\} \quad (15a)$$

$$I_i^s(t) = \{x_0^s(1{:}t), y_i^s(1{:}t)\}, \quad i \in N. \quad (15b)$$

We now show that the above information structure may be viewed as the *static reduction* of the original information structure [6] and [35].

Lemma 2: For any arbitrary but fixed strategy $g \in \mathscr{G}$,

$$I_i(t) \equiv I_i^s(t), \quad i \in N_0$$

i.e., both sets generate the same sigma-algebra or, equivalently, they are functions of each other. Moreover, if $g \in \mathscr{G}_A$ then $I_i(t)$ and $I_i^s(t)$, $i \in N_0$, are linear functions of each other.

The proof is presented in Appendix B. In the sequel, we use Lemma 2 to replace conditioning on $I_i(t)$ by conditioning on $I_i^s(t)$ and to replace a linear function of $I_i(t)$ by a linear function of $I_i^s(t)$. As a first implication, we derive the following additional properties of the split components of the state.

Lemma 3: For any strategy $g \in \mathscr{G}$, the split components of the state and the control action satisfy the following additional properties: for any $i \in N$,

P7) For any $\tau \le t$, $\mathbb{E}[u_i^\ell(\tau)|I^c(t)] = 0$.
P8) For any $\tau \le t$, $\mathbb{E}[x_i^\ell(\tau)|I^c(t)] = 0$.
For any matrix M of appropriate dimensions:
P9) $\mathbb{E}[x_i^\ell(t)^\intercal M x_0^s(t)] = 0$.
P10) $\mathbb{E}[x_i^\ell(t)^\intercal M x^c(t)] = 0$.
P11) $\mathbb{E}[u_i^\ell(t)^\intercal M x_0^s(t)] = 0$.

The proof is presented in Appendix C.

C. Conditional Independence and Split of Per-Step Cost

Lemma 4: For any strategy $g \in \mathscr{G}$ and any $i, j \in N$, $i \ne j$, we have the following:

1) $(x_i(1{:}t), u_i(1{:}t)) \perp\!\!\!\perp (x_j(1{:}t), u_j(1{:}t)) \mid I^c(t)$.
2) $x_i^s(1{:}t) \perp\!\!\!\perp x_j^s(1{:}t) \mid I_0^s(t)$.
3) $(x_i^\ell(1{:}t), u_i^\ell(1{:}t)) \perp\!\!\!\perp (x_j^\ell(1{:}t), u_j^\ell(1{:}t)) \mid I^c(t)$.

The proof is presented in Appendix D.

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

$$z^c(t) = x^c(t) + x^s(t), \qquad z_i^\ell(t) = x_i^\ell(t) + x_i^s(t). \quad (16)$$

Due to the conditional independence of Lemma 4, the per-step cost simplifies as follows.

Lemma 5: The per-step cost simplifies as follows:

$$\mathbb{E}\big[x(t)^\intercal Q x(t)\big] = \mathbb{E}\Big[z^c(t)^\intercal Q z^c(t) + \sum_{i=1}^n z_i^\ell(t)^\intercal Q_{ii} z_i^\ell(t) - \sum_{i=1}^n x_i^s(t) Q_{ii} x_i^s(t)\Big] \quad (17)$$

and

$$\mathbb{E}\big[u(t)^\intercal Ru(t)\big] = \mathbb{E}\Big[u^c(t)^\intercal Ru^c(t) + \sum_{i\in N} u_i^\ell(t)^\intercal R_{ii}u_i^\ell(t)\Big]. \tag{18}$$

The proof is presented in Appendix E.

D. Completion of Squares

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

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

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

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

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

Now, we can show

$$\begin{aligned}&u^\intercal Ru + (Ax + Bu)^\intercal S(Ax + Bu)\\ &\quad = (u + Lx)^\intercal\Delta(u + Lx) + x^\intercal\tilde{S}x\end{aligned}$$

by expanding both sides and combining the coefficients. The proof follows by combining both the equations. ■

Let $S^c(1{:}T)$ and $S_i^\ell(1{:}T)$ denote the solution to the following Riccati equations: Initialize $S^c(T) = Q_T$ and $S_i^\ell(T) = [Q_T]_{ii}$, $i \in N$. Then, for $t \in \{T-1, \ldots, 1\}$, recursively define

$$S^c(t) = \mathcal{R}(S^c(t+1), A, B, Q, R) \tag{19}$$

$$S_i^\ell(t) = \mathcal{R}(S_i^\ell(t+1), A_{ii}, B_{ii}, Q_{ii}, R_{ii}), \quad i \in N. \tag{20}$$

Define the gains

$$L^c(t) = \mathcal{G}(S^c(t+1), A, B, R) \tag{21}$$

$$L_i^\ell(t) = \mathcal{G}(S_i^\ell(t+1), A_{ii}, B_{ii}, R_{ii}), \quad i \in N \tag{22}$$

and the matrices

$$\Delta^c(t) = [R + B^\intercal S^c(t+1)B]$$

$$\Delta_i^\ell(t) = [R_{ii} + B_{ii}^\intercal S_i^\ell(t+1)B_{ii}].$$

Lemma 7: For any strategy $g \in \mathscr{G}$, the total cost may be split as

$$J(g) = J^c(g) + \sum_{i\in N} J_i^\ell(g) + J^s \tag{23}$$

where $J^c(g)$ is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(u^c(t) + L^c(t)z^c(t))^\intercal\Delta^c(t)(u^c(t) + L^c(t)z^c(t))\right]$$

and $J_i^\ell(g)$, $i \in N$, is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(u_i^\ell(t) + L_i^\ell(t)z_i^\ell(t))^\intercal\Delta_i^\ell(t)(u_i^\ell(t) + L_i^\ell(t)z_i^\ell(t))\right]$$

and J^s is given by

$$\begin{aligned}\mathbb{E}\Bigg[&x(1)^\intercal S^c(1)x(1) + \sum_{i=1}^{n} x_i(1)^\intercal S_i^\ell(1)x_i(1)\\ &+ \sum_{t=1}^{T-1}\Big[w(t)^\intercal S^c(t+1)w(t) + \sum_{i=1}^{n} w_i(t)^\intercal S_i^\ell(t+1)w_i(t)\Big]\\ &+ \sum_{t=1}^{T-1}\sum_{i=1}^{n}\big[(A_{i0}x_0^s(t))^\intercal S_i^\ell(t+1)(A_{i0}x_0^s(t) + 2A_{ii}x_i^s(t))\big]\\ &- \sum_{t=1}^{T-1}\sum_{i=1}^{n} x_i^s(t)Q_{ii}x_i^s(t) - \sum_{i=1}^{n} x_i^s(T)[Q_T]_{ii}x_i^s(T)\Bigg].\end{aligned}$$

Proof: We start by rewriting the total cost using the result of Lemma 5. In particular, $J(g)$ can be written as

$$\begin{aligned}&\mathbb{E}\left[\sum_{t=1}^{T-1} z^c(t)^\intercal Qz^c(t) + u^c(t)^\intercal Ru^c(t) + z^c(T)^\intercal Q_T z^c(T)\right]\\ &+ \mathbb{E}\Bigg[\sum_{t=1}^{T-1}\sum_{i=1}^{n} z_i^\ell(t)^\intercal Q_{ii}z_i^\ell(t) + u_i^\ell(t)^\intercal R_{ii}u_i^\ell(t)\\ &+ \sum_{i=1}^{n} z_i^\ell(T)^\intercal[Q_T]_{ii}z_i^\ell(T)\Bigg]\\ &- \mathbb{E}\left[\sum_{t=1}^{T-1}\sum_{i=1}^{n} x_i^s(t)^\intercal Q_{ii}x_i^s(t) - \sum_{i=1}^{n} x_i^s(T)^\intercal[Q_T]_{ii}x_i^s(T)\right].\end{aligned}$$

The dynamics of $z^c(t)$ and $z^\ell(t)$ may be written as

$$\begin{aligned}z^c(t+1) &= Az^c(t) + Bu^c(t) + w(t)\\ z_i^\ell(t+1) &= A_{ii}z_i^\ell(t) + A_{i0}x_0^s(t) + B_{ii}u_i^\ell(t) + w_i(t).\end{aligned}$$

Note that $w(t)$ is zero mean and independent of $(z^c(t), u^c(t))$ (because both $z^c(t)$ and $u^c(t)$ depend on $w(1{:}t-1)$, which is independent of $w(t)$). Similarly, $w(t)$ is zero mean and independent of $(\mathrm{vec}(x_0^s(t), z_i^\ell(t)), u_i^\ell(t))$. The result then follows from recursively applying Lemma 6, (P9), and (P11). ■

Remark 1: The term J^s is control-free and depends on only the primitive random variables. Hence minimizing $J(g)$ is equivalent to minimizing $J^c(g) + \sum_{i\in N} J_i^\ell(g)$.

In the next two sections, we simplify $J^c(g) + \sum_{i\in N} J_i^\ell(g)$ using orthogonality properties of MMSE/ LLMS estimates and the corresponding estimation error.

IV. Main Results for Problem 1

A. Orthogonal Projection

As explained in Remark 1, minimizing $J(g)$ is equivalent to minimizing $J^c(g) + \sum_{i\in N} J_i^\ell(g)$ defined in Lemma 7. To simplify $J^c(g) + \sum_{i\in N} J_i^\ell(g)$, define

$$\hat{z}(t|c) := \mathbb{E}[z^c(t)|I^c(t)] \tag{24a}$$

$$\breve{z}_i^\ell(t|i) := \mathbb{E}[z_i^\ell(t)|I_i(t)] - \mathbb{E}[z_i^\ell(t)|I_0(t)]. \tag{24b}$$

Define the "estimation errors"

$$\tilde{z}^c(t) = z^c(t) - \hat{z}(t|c), \quad \tilde{z}_i^\ell(t) = z_i^\ell(t) - \breve{z}_i^\ell(t|i).$$

Lemma 8: For any strategy $g \in \mathscr{G}$, the variables defined above satisfy the following properties.

C1) $\tilde{z}^c(t)$ and $\tilde{z}_i^\ell(t)$ are control-free and may be written just in terms of the primitive random variables.

C2) $\mathbb{E}[\tilde{z}^c(t)|I^c(t)] = 0$.

For any matrix M of appropriate dimensions:

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

C4) $\mathbb{E}[u^c(t)^\intercal M \tilde{z}^c(t)] = 0$.

C5) $\mathbb{E}[\tilde{z}_i^\ell(t)^\intercal M \breve{z}_i^\ell(t|i)] = 0$.

C6) $\mathbb{E}[u_i^\ell(t)^\intercal M \tilde{z}_i^\ell(t)] = 0$.

The proof is presented in Appendix F.

An implication of the above is the following.

Lemma 9: The per-step terms in $J^c(g)$ and $J_i^\ell(g)$ simplify as follows:

$$\begin{aligned}
&\mathbb{E}\left[(u^c(t) + L^c(t)z^c(t))^\intercal \Delta^c(t)(u^c(t) + L^c(t)z^c(t))\right] \\
&\quad = \mathbb{E}\left[(u^c(t) + L^c(t)\hat{z}(t|c))^\intercal \Delta^c(t)(u^c(t) + L^c(t)\hat{z}(t|c))\right] \\
&\quad + \mathbb{E}\left[\tilde{z}^c(t)^\intercal L^c(t)^\intercal \Delta^c(t) L^c(t)\tilde{z}^c(t)\right]
\end{aligned} \tag{25}$$

and

$$\begin{aligned}
&\mathbb{E}\left[(u_i^\ell(t) + L_i^\ell(t)z_i^\ell(t))^\intercal \Delta_i^\ell(t)(u_i^\ell(t) + L_i^\ell(t)z_i^\ell(t))\right] \\
&\quad = \mathbb{E}\left[(u_i^\ell(t) + L_i^\ell(t)\breve{z}_i^\ell(t|i))^\intercal \Delta_i^\ell(t)(u_i^\ell(t) + L_i^\ell(t)\breve{z}_i^\ell(t|i))\right] \\
&\quad + \mathbb{E}\left[\tilde{z}_i^\ell(t)^\intercal L_i^\ell(t)^\intercal \Delta_i^\ell(t) L_i^\ell(t)\tilde{z}_i^\ell(t)\right].
\end{aligned} \tag{26}$$

Proof: Equation (25) follows from (C2) and is equivalent to

$$\mathbb{E}[u^c(t)^\intercal \Delta^c(t) L^c(t)\tilde{z}^c(t)] = 0 \tag{27}$$

$$\mathbb{E}[\hat{z}(t|c)(t)^\intercal L^c(t)^\intercal \Delta^c(t) L^c(t)\tilde{z}^c(t)] = 0 \tag{28}$$

which is the direct result of (C3) and (C4).

Equation (26) is equivalent to

$$\mathbb{E}[u_i^\ell(t)^\intercal \Delta_i^\ell(t) L_i^\ell(t)\tilde{z}^\ell(t)] = 0 \tag{29}$$

$$\mathbb{E}[\tilde{z}_i^\ell(t)^\intercal L_i^\ell(t)^\intercal \Delta_i^\ell(t) L_i^\ell(t)\breve{z}_i^\ell(t|i)] = 0 \tag{30}$$

which is a direct result of (C5) and (C6). ∎

An immediate implication of Lemma 9 is the following.

Lemma 10: For any strategy $g \in \mathscr{G}$, the cost $J^c(t)$ and $J_i^\ell(t)$ defined in Lemma 7 may be further split as

$$J^c(g) = \hat{J}^c(g) + \tilde{J}^c, \quad J_i^\ell(g) = \breve{J}_i^\ell(g) + \tilde{J}_i^\ell$$

where $\hat{J}^c(g)$ is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(u^c(t) + L^c(t)\hat{z}(t|c))^\intercal \Delta^c(t)(u^c(t) + L^c(t)\hat{z}(t|c))\right]$$

and $\tilde{J}^c$ is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(L^c(t)\tilde{z}^c(t))^\intercal \Delta^c(t) L^c(t)\tilde{z}^c(t)\right]$$

and $\breve{J}_i^\ell(g)$, $i \in N$, is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(u_i^\ell(t) + L_i^\ell(t)\breve{z}_i^\ell(t|i))^\intercal \Delta_i^\ell(t)(u_i^\ell(t) + L_i^\ell(t)\breve{z}_i^\ell(t|i))\right]$$

and $\tilde{J}_i^\ell$, $i \in N$, is given by

$$\mathbb{E}\left[\sum_{t=1}^{T-1}(L_i^\ell(t)\tilde{z}_i^\ell(t))^\intercal \Delta_i^\ell(t) L_i^\ell(t)\tilde{z}_i^\ell(t)\right].$$

Remark 2: Property (C1) implies that the terms $\tilde{J}^c$ and $\tilde{J}_i^\ell$ are control-free and depend only on the primitive random variables. Combined with Remark 1, this implies that minimizing $J(g)$ is equivalent to minimizing $\hat{J}^c(g) + \sum_{i \in N} \breve{J}^i(g)$.

Theorem 1: The optimal control strategy of Problem 1 is unique and is given by

$$u^c(t) = -L^c(t)\hat{z}(t|c) \tag{31a}$$

$$u_i^\ell(t) = -L_i^\ell(t)\breve{z}_i^\ell(t|i). \tag{31b}$$

Furthermore, the optimal performance is given by

$$J^* := \inf_{g \in \mathscr{G}} J(g) = \tilde{J}^c + \sum_{i \in N} \tilde{J}_i^\ell$$

where $\tilde{J}^c$ and $\tilde{J}_i^\ell$ are defined in Lemma 10.

Proof: As argued in Remark 2, minimizing $J(g)$ is equivalent to minimizing $\hat{J}^c(g) + \sum_{i \in N} \breve{J}^i(g)$. By assumption, R is symmetric and positive definite and, therefore, so is R_{ii}. It can be shown recursively that $S^c(t)$ and $S_i^\ell(t)$ are symmetric and positive-semidefinite. Hence, both $\Delta^c(t)$ and $\Delta_i^\ell(t)$ are symmetric and positive definite. Therefore

$$\hat{J}^c(g) + \sum_{i \in N} \breve{J}_i^\ell(g) \geq 0$$

with equality if and only if the strategy g is given by (31). ∎

The optimal control strategy in Theorem 1 is described in terms of the common and local components of the control. We can write it in terms of the control actions of the agents as follows. Let

$$\hat{x}(t|c) = \mathbb{E}[x(t) \mid I^c(t)] \quad \text{and} \quad \hat{x}(t|i) = \mathbb{E}[x(t) \mid I_i(t)]$$

denote the major and ith minor agent's MMSE estimate of the state. Equations (16) and (24) imply the following.

Lemma 11: The common and local information based estimates $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ are related to the major and minor agents' MMSE estimates as follows:

$$\hat{z}(t|c) = \hat{x}(t|c) \quad \text{and} \quad \breve{z}_i^\ell(t|i) = \hat{x}_i(t|i) - \hat{x}_i(t|c).$$

Proof: (P8) implies that $\hat{x}(t|c) = \hat{z}(t|c)$. Moreover, since $x_i^c(t)$ is a function of $I^c(t)$ (and, therefore, a function of $I_i(t)$), we have

$$\begin{aligned}
\hat{x}_i(t|i) - \hat{x}_i(t|c) &= x_i^c(t) + \mathbb{E}[x_i^\ell(t) + x_i^s(t) \mid I_i(t)] \\
&\quad - x_i^c(t) - \mathbb{E}[x_i^\ell(t) + x_i^s(t) \mid I_0(t)] \\
&= \breve{z}_i^\ell(t|i)(t).
\end{aligned}$$

∎

Let $\hat{x}_i(t|c)$ and $\hat{x}_i(t|i)$ denote the ith element of $\hat{x}(t|c)$ and $\hat{x}(t|i)$, respectively. Moreover, let $f_{i,t}$ denote the conditional density of $x_i(t)$ given $I_i(t)$. Note that $\hat{x}_i(t|i)$ is the mean of $f_{i,t}$.

Theorem 2: The optimal control strategy of Problem 1 is unique and is given by

$$u_0(t) = -L_0^c(t)\hat{x}(t|c) \tag{32a}$$

and for all $i \in N$,

$$u_i(t) = -L_i^c(t)\hat{x}(t|c) - L_i^\ell(t)(\hat{x}_i(t|i) - \hat{x}_i(t|c)) \tag{32b}$$

where $L_i^c(t)$ denote the ith row of $L^c(t)$. The major agent's MMSE estimate can be recursively updated as follows: $\hat{x}(1|c) = \text{vec}(x_1(1), 0, \dots, 0)$ and

$$\hat{x}(t+1|c) = A \begin{bmatrix} x_0(t) \\ \hat{x}_1(t|c) \\ \vdots \\ \hat{x}_n(t|c) \end{bmatrix} + B \begin{bmatrix} u_0(t) \\ u_1^c(t|c) \\ \vdots \\ u_n^c(t|c) \end{bmatrix} + \begin{bmatrix} w_0(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{33}$$

where

$$w_0(t) = x_0(t+1) - A_{00}x_0(t) - B_{00}u_0(t)$$

and $u_i^c(t|c) = -L_i^c(t)\hat{x}(t|c)$. Furthermore, under Assumption 2, the ith minor agent's MMSE estimate is given by

$$\hat{x}_i(t|i) = x_i^c(t) + x_i^\ell(t) + \int x_i^s(t) f_{i,t}(x_{i,t}^s) dx_i^s(t) \tag{34}$$

where the conditional density $f_{i,t}$ may be updated using the following Bayesian filter: for any $x_i^s(t)$,

$$f_{i,t}(x_i^s(t)) = \frac{\beta_i(t)\int \gamma_i(t)\gamma_0(t) f_{i,t-1}(x_i^s(t-1))dx_i^s(t-1)}{\int \beta_i(t)\int \gamma_i(t)\gamma_0(t) f_{i,t-1}(x_i^s(t-1))dx_i^s(t-1)dx_i^s(t)} \tag{35}$$

where

$$\beta_i(t) = \nu_{i,t}(y_i^s(t) - C_{ii}x_i^s(t))$$
$$\gamma_0(t) = \varphi_{0,t}\big(x_0^s(t) - A_{00}x_0^s(t-1)\big)$$
$$\gamma_i(t) = \varphi_{i,t}\big(x_i^s(t) - A_{ii}x_i^s(t-1) - A_{i0}x_0^s(t-1)\big)$$

and $\varphi_{i,t}$ and $\nu_{i,t}$ are the distributions of the noise variables $w_i(t)$ and $v_i(t)$, respectively.

Proof: The structure of optimal policies follows from Lemma 11 and Theorem 1.

We establish the update of the major agent's MMSE estimate in two steps. First note that

$$\hat{x}_0(t+1|c) = \mathbb{E}[x_0(t+1)|I^c(t+1)] = x_0(t+1) \tag{36}$$

because $x_0(t+1)$ is part of $I^c(t+1)$. This proves the zeroth component of (33). Next, for any $i \in N$,

$$\begin{aligned} \hat{x}_i(t+1|c) &= \mathbb{E}[x_i(t+1)|I^c(t+1)] \\ &\overset{(a)}{=} \mathbb{E}[A_{i0}x_0(t) + B_{i0}u_0(t) + A_{ii}x_i(t) + B_{ii}u_i(t)|I^c(t+1)] \\ &\overset{(b)}{=} A_{i0}x_0(t) + B_{i0}u_0(t) + \mathbb{E}[A_{ii}x_i(t) + B_{ii}u_i(t)|I^c(t)] \\ &= A_{i0}x_0(t) + A_{ii}\hat{x}_i(t|c) + B_{i0}u_0(t) + B_{ii}u_i^c(t) \end{aligned} \tag{37}$$

where (a) is because $w_i(t)$ is zero mean and independent of $I^c(t+1)$ and (b) follows from the following:

1) $x_0(t)$ and $u_0(t)$ are part of $I^c(t+1)$ so can be taken out of the expectation;
2) $I^c(t+1)$ is equivalent to $(I^c(t), u_0(t), x_0(t+1))$ which, in turn, is equivalent to $(I^c(t), u_0(t), w_0(t))$. Now,

$$\begin{aligned} &\mathbb{E}[A_{ii}x_i(t) + B_{ii}u_i(t)|I^c(t), u_0(t), w_0(t)] \\ &= \mathbb{E}[A_{ii}x_i(t) + B_{ii}u_i(t)|I^c(t)] \end{aligned}$$

because $u_0(t)$ can be removed from the conditioning since it is a function of $I^c(t)$ and $w_0(t)$ can be removed from the conditioning because it is independent of $x_i(t)$ and $u_i(t)$.

This proves the ith component of (33).

Finally, to compute $\hat{x}_i(t|i)$, we use the state split in (13b). We have

$$\begin{aligned} \hat{x}_i(t|i) &= \mathbb{E}[x_i(t)|I_i(t)] \\ &= \mathbb{E}[x_i^c(t) + x_i^\ell(t) + x_i^s(t)|I_i(t)] \\ &\overset{(a)}{=} x_i^c(t) + x_i^\ell(t) + \mathbb{E}[x_i^s(t)|I_i(t)] \\ &\overset{(b)}{=} x_i^c(t) + x_i^\ell(t) + \mathbb{E}[x_i^s(t)|I_i^s(t)], \end{aligned}$$

where in (a), we use the fact that $x_i^c(t)$ and $x_i^\ell(t)$ are measurable functions of $I_i(t)$, and in (b), we use Lemma 2. Now, we consider the update of the conditional density. With a slight abuse of notation, we use $\mathbb{P}(y_i^s(t)|x_i^s(t))$ to denote the conditional density of $y_i^s(t)$ given $x_i^s(t)$ and similar interpretations hold for other terms. Consider

$$\begin{aligned} f_{i,t}(x_i^s(t)) &= \mathbb{P}(x_i^s(t)|I_i^s(t)) \\ &= \int \mathbb{P}(x_i^s(t), x_i^s(t-1)|I_i^s(t))dx_i^s(t-1). \end{aligned} \tag{38}$$

Substituting $I_i^s(t) = (I_i^s(t-1), y_i^s(t), x_0^s(t))$ in (38) and using Bayes rule, we get that $f_{i,t}(x_i^s(t))$ is equal to

$$\frac{\int \mathbb{P}(y_i^s(t), x_i^s(t), x_0^s(t)|I_i^s(t))dx_i^s(t-1)}{\iint \mathbb{P}(y_i^s(t), x_i^s(t), x_0^s(t)|I_i^s(t))dx_i^s(t-1)dx_i^s(t)}. \tag{39}$$

Now consider

$$\begin{aligned} &\mathbb{P}(y_i^s(t), x_i^s(t), x_0^s(t)|I_i^s(t)) \\ &= \mathbb{P}(y_i^s(t)|x_i^s(t)) \\ &\times \mathbb{P}(x_i^s(t)|x_0^s(t-1), x_i^s(t-1)) \\ &\times \mathbb{P}(x_0^s(t)|x_0^s(t-1)) \times \mathbb{P}(x_i^s(t-1)|I_i^s(t-1)). \end{aligned} \tag{40}$$

Substituting (40) into (39) gives the updated equation (35). ■

B. Implementation of the Optimal Control Strategy

Based on Theorem 2, the optimal control strategy can be implemented as follows.

1) Computation of the Gains:

Before the system starts running, the agents perform the following computations.

1) All agents solve the Riccati equation (19) and compute the gains $L^c(t)$ using (21). The major agent stores the row

$L_0^c(t)$ while minor agent i stores the row $L_i^c(t)$. For ease of reference, we repeat the equations here:

$$S^c(t) = \mathcal{R}(S^c(t+1), A, B, Q, R)$$
$$L^c(t) = \mathcal{G}(S^c(t+1), A, B, R).$$

Note that these are global equations that depend on the dynamics and the cost of the complete system.

2) Minor agent i solves the Riccati equation (20) and computes and stores the gains $L_i^\ell(t)$ using (22). For ease of reference, we repeat them here

$$S_i^\ell(t) = \mathcal{R}(S_i^\ell(t+1), A_{ii}, B_{ii}, Q_{ii}, R_{ii})$$
$$L_i^\ell(t) = \mathcal{G}(S_i^\ell(t+1), A_{ii}, B_{ii}, R_{ii}).$$

Note that these are local equations that depend on the local dynamics and the cost of the minor agent i.

2) Filtering and Tracking of Different Components of the State: Once the system is running, the agents keep track of the following components of the state and their estimates:

1) All agents keep track of the major agent's MMSE estimate using (33), which we repeat here: $\hat{x}(1|c) = \text{vec}(x_1(0), 0, \ldots, 0)$ and

$$\hat{x}(t+1|c) = A \begin{bmatrix} x_0(t) \\ \hat{x}_1(t|c) \\ \vdots \\ \hat{x}_n(t|c) \end{bmatrix} + B \begin{bmatrix} u_0(t) \\ u_1^c(t|c) \\ \vdots \\ u_n^c(t|c) \end{bmatrix} + \begin{bmatrix} w_0(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

2) Agent i keeps track of the density $f_{i,t}$ of $x_i(t)$ given $I_i^s(t)$ using the Bayesian filter (35) and computes the mean $\hat{x}_i(t|i)$ of this density. Note that the Bayesian filter (35) does not depend on the control strategy.

3) Implementation of the Control Strategies: Finally, the agents choose the control actions as follows.

1) The major agent chooses $u_0(t)$ using (32a), which we repeat as follows:

$$u_0(t) = u_0^c(t) = -L_0^c(t)\hat{x}(t|c).$$

2) The minor agent chooses $u_i(t)$ using (32b), which we repeat as follows:

$$u_i(t) = u_i^c(t) + u_i^\ell(t)$$
$$= -L_i^c(t)\hat{x}(t|c) - L_i^\ell(t)(\hat{x}_i(t|i) - \hat{x}_i(t|c)).$$

C. Special Case of State Feedback

Consider the special case of the model when each minor agent observes its state perfectly. This corresponds to $C_{ii} = I$ and $v_i(t) = 0$. The information structure remains the same as before. In this case, the result of Theorem 2 simplifies as follows. The optimal control action of the major agent is

$$u_0(t) = L_0^c(t)\hat{x}(t|c) \tag{41}$$

and that of the ith minor agent, $i \in N$, is

$$u_i(t) = L_i^c(t)\hat{x}(t|c) + L_i^\ell(t)(x_i(t) - \hat{x}_i(t|c)) \tag{42}$$

where $\hat{x}(t|c) = \mathbb{E}[x(t)|I_0(t)]$. A similar result for only one minor agent was derived in [19].

The following remarks are in order.

1) The major agent observes its local state and the minor agents observe their local state and the state of the major agent. Nonetheless, the optimal control strategy involves the major agent's MMSE estimate of the global state.
2) As argued before, the major agent's MMSE estimate of the state of the system evolves according to a linear filter. Therefore, the optimal control action is a linear function of the data.
3) In light of the above result, we may view the optimal solution for partial output feedback as a certainty equivalence solution. In particular, the optimal strategy (32b) of the minor agent in partial output feedback is the same as the optimal strategy in state feedback where the state $x_i(t)$ is replaced by the MMSE estimate of the state.

V. Main Results for Problem 2

The main idea of this section is the same as that of Section IV; however, instead of defining $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ in terms of expectation (which can be nonlinear), we define them in terms of Hilbert space projections that are linear. We first start with an overview of basic results for Hilbert space projections.

A. Preliminaries of Hilbert Space Projections

Given zero-mean random variables x and y defined on a common probability space, the least linear mean square estimate (LLMS) $\mathbb{L}[x|\text{span}\,(y)]$ is the projection of x on to $Y = \text{span}\,(y)$ and satisfies the orthogonal projection property: for any $z \in Y$,

$$\mathbb{E}[(x - \mathbb{L}[x|Y])z^\intercal] = 0 \text{ and } \mathbb{E}[(x - \mathbb{L}[x|Y])^\intercal z] = 0. \tag{43}$$

For any arbitrary but fixed strategy $g \in \mathscr{G}_A$ and any agent $i \in N_0$, define $H_i(t) = \text{span}\,\{I_i(t)\}$ and $H_i^s(t) = \text{span}\,\{I_i^s(t)\}$. We can split $H_i(t)$ and $H_i^s(t)$ into orthogonal subspaces

$$H_i(t) = H_0(t) \oplus \tilde{H}_i(t) \quad \text{and} \quad H_i^s(t) = H_0^s(t) \oplus \tilde{H}_i^s(t)$$

where $\tilde{H}_i(t)$ is the orthogonal complement of $H_0(t)$ with respect to $H_i(t)$ and a similar interpretation holds for $\tilde{H}_i^s(t)$. Thus, for any random variable v

$$\mathbb{L}[v \mid H_i(t)] = \mathbb{L}[v \mid H_0(t)] + \mathbb{L}[v \mid \tilde{H}_i(t)] \tag{44}$$

and similar interpretations holds for projections on $H_i^s(t)$.

Now, define $W_0(t) = \text{span}\{x_0(1), w_0(1{:}t-1)\}$, and, for any minor agent $i \in N$, $W_i(t) = \text{span}\{x_i(1), w_i(1{:}t-1), v_i(1{:}t)\}$. An immediate implication of Lemma 2 is the following.

Lemma 12: For any $g \in \mathscr{G}_A$ and $i \in N_0$, $H_i(t) = H_i^s(t)$; therefore, $\tilde{H}_i(t) = \tilde{H}_i^s(t)$. Furthermore, for all t and $i \in N$,

1) $H_0(t) = H_0^s(t) = W_0(t)$.
2) $H_i(t) = H_i^s(t) \subseteq W_0(t) \oplus W_i(t)$.
3) $\tilde{H}_i(t) = \tilde{H}_i^s(t) \subseteq W_i(t)$.

Proof: By construction, $x_0^s(t) \in W_0(t)$ and, it is easy to show that $w_0(t-1) \in H^s(t)$. Hence, $H_0^s(t) = W_0(t)$. Similarly, by construction, $y_i^s(t) \in W_0(t) \oplus W_i(t)$. Hence, $H_i^s(t) \subseteq W_0(t) \oplus W_i(t)$. Finally, consider any vector $b_i \in \tilde{H}_i^s(t)$. Then,

$b_i \in W_i^s(t)$ as each element of $\tilde{H}_i^s$ is a specific linear function of $W_i(t)$ due to linear dynamics of the system. ■

Lemma 13: For any strategy $g \in \mathscr{G}_A$,

$$u^c(t) = \mathbb{E}[u(t) \mid I^c(t)] \in H_0^s(t)$$
$$u_i^\ell(t) = u_i(t) - u^c(t) \in \tilde{H}_i^s(t).$$

Proof: For any strategy $g \in \mathscr{G}_A$, $u_i(t) \in H_i(t) = H_i^s(t) = H_0^s(t) \oplus \tilde{H}_i^s(t)$. Thus, by Lemma 12, $u_i(t) \in W_0(t) \oplus W_i(t)$, which are independent subspaces. Therefore, the result follows from orthogonal projection (43) and independence of $W_0(t)$ and $W_i(t)$. ■

Proof: For any strategy $g \in \mathscr{G}_A$, $u_i(t) \in H_i(t) = H_i^s(t) = H_0^s(t) \oplus \tilde{H}_i^s(t)$. Hence, there exist unique vectors $a_i(t) \in H_0^s(t)$ and $b_i(t) \in \tilde{H}_i^s(t)$, such that $u_i(t) = a_i(t) + b_i(t)$.

We have

$$\mathbb{E}[u_i(t) \mid I^c(t)] \stackrel{(a)}{=} \mathbb{E}[a_i(t) + b_i(t) \mid I^c(t)]$$
$$\stackrel{(b)}{=} \mathbb{E}[a_i(t) \mid I^c(t)] \stackrel{(c)}{=} a_i(t)$$

where (a) uses the unique orthogonal decomposition $u_i(t) = a_i(t) + b_i(t)$, (b) uses $\mathbb{E}[b_i(t) \mid I^c(t)] = 0$ from Lemma 12, Part 3, and (c) uses $\mathbb{E}[a_i(t) \mid I^c(t)] = a_i(t)$ from Lemma 12, Part 2. Hence, $u^c(t) = a_i(t) \in H_0^s(t)$. Moreover, $u_i^\ell(t) = u(t) - u^c(t) = u(t) - a_i(t) = b_i(t) \in \tilde{H}_i^s(t)$. ■

Lemma 14: For any $g \in \mathscr{G}_A$, we have the following.

S1) For any $\tau < t$, $u^c(\tau) \in H_0(\tau) \subset H_0(t)$.

S2) For any $\tau \leq t$, $x^c(\tau) \in H_0(t)$.

S3) For any $\tau \leq t$, $\mathbb{L}[x_i^\ell(\tau) | H_0(t)] = 0$.

Proof: Using (13), we have the following.

S1) From the results of Lemma 13, for any $\tau < t$, $u^c(\tau) \in H_0(\tau)$ where $H_0(\tau) \subset H_0(t)$.

S2) For any $\tau \leq t$, by construction, $x^c(\tau)$ is a linear function of $u^c(1{:}\tau - 1)$. Hence, by (S1), $x^c(\tau) \in H_0(\tau - 1) \subset H_0(t)$.

S3) For any $\tau \leq t$, by construction, $x_i^\ell(\tau)$ is a linear function of $u_i^\ell(1{:}\tau - 1)$. Hence, it belongs to $\tilde{H}_i(t)$ by Lemma 13. ■

B. Orthogonal Projection

We use the same notation as in Section IV with the understanding that the terms are defined differently. We do not use any result from Section IV here, so the overlap of notation should not cause any confusion.

As explained in Remark 1, minimizing $J(g)$ is equivalent to minimizing $J^c(g) + \sum_{i\in N} J_i^\ell(g)$ defined in Lemma 7. To simplify $J^c(g) + \sum_{i\in N} J_i^\ell(g)$, define

$$\hat{z}(t|c) := \mathbb{L}[z^c(t)|H_0(t)] \tag{45}$$

$$\breve{z}_i^\ell(t|i) := \mathbb{L}[z_i^\ell(t)|H_i(t)] - \mathbb{L}[z_i^\ell(t)|H_0(t)]. \tag{46}$$

Equations (44) and (46) imply that

$$\breve{z}_i^\ell(t|i) = \mathbb{L}[z_i^\ell(t)|\tilde{H}_i(t)]. \tag{47}$$

Define the estimation errors

$$\tilde{z}^c(t) = z^c(t) - \hat{z}(t|c), \quad \tilde{z}_i^\ell(t) = z_i^\ell(t) - \breve{z}_i^\ell(t|i).$$

Lemma 15: For any strategy $g \in \mathcal{G}_A$ the properties (C1) and (C3)–(C6) hold for $\hat{z}(t|c)$, $\breve{z}_i^\ell(t|i)$, $\tilde{z}^c(t)$, and $\tilde{z}_i^\ell(t)$ defined above.

The proof is presented in Appendix G. An implication of the above is the following.

Lemma 16: For any strategy $g \in \mathscr{G}_A$, the results of Lemma 9, hold with $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ defined by (45) and (46).

Proof: As mentioned in the proof of Lemma 9, (25) follows from (C3) and (C4) and is equivalent to (27) and (28).

Equation (26) follows from (C5) and (C6) and is equivalent to (29) and (30). ■

An immediate implication of Lemma 16 is the following.

Lemma 17: For any strategy $g \in \mathscr{G}_A$, the results of Lemma 10 hold with $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ defined by (45) and (46).

Remark 3: The terms $\tilde{J}^c$ and $\tilde{J}_i^\ell$ are control-free and depend only on the primitive random variables. Combined with Remark 1, this implies that minimizing $J(g)$ is equivalent to minimizing $\hat{J}^c(g) + \sum_{i\in N} \breve{J}^i(g)$.

C. Main Results

Theorem 3: The optimal control strategy of Problem 2 is unique and is given by

$$u^c(t) = -L^c(t)\hat{z}(t|c) \tag{48a}$$
$$u_i^\ell(t) = -L_i^\ell(t)\breve{z}_i^\ell(t|i). \tag{48b}$$

Furthermore, the optimal performance is given by

$$J_A^* := \inf_{g\in\mathscr{G}_A} J(g) = \tilde{J}^c + \sum_{i\in N} \tilde{J}_i^\ell$$

where $\tilde{J}^c$ and $\tilde{J}_i^\ell$ are defined in Lemma 10 with $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ defined by (45) and (46).

Proof: The proof relies on symmetric property and positive definiteness of both $\Delta^c(t)$ and $\Delta_i^\ell(t)$ and is the same as that of Theorem 1. ■

Now let

$$\hat{x}(t|c) = \mathbb{L}[x(t) \mid I^c(t)] \quad \text{and} \quad \hat{x}(t|i) = \mathbb{L}[x(t) \mid I_i(t)]$$

denote the major and the ith minor agent's LLMS estimate of the state. Let $\hat{x}_i(t|c)$ and $\hat{x}_i(t|i)$ denote the ith element of $\hat{x}(t|c)$ and $\hat{x}(t|i)$, respectively. Equations (16), (45), and (46) imply the following.

Lemma 18: The common and local information based estimates $\hat{z}(t|c)$ and $\breve{z}_i^\ell(t|i)$ are related to the major and minor agents' LLMS estimates as follows:

$$\hat{z}(t|c) = \hat{x}(t|c) \quad \text{and} \quad \breve{z}_i^\ell(t|i) = \hat{x}_i(t|i) - \hat{x}_i(t|c).$$

Proof: First observe that (P8) implies $\hat{x}(t|c) = \hat{z}(t|c) \in H_0(t)$. Now consider that

$$\hat{x}(t|i) - \hat{x}(t|c) \stackrel{(a)}{=} x_i^c(t) + \mathbb{L}[x_i^\ell(t) + x_i^s(t) \mid H_i(t)]$$
$$- x_i^c(t) - \mathbb{L}[x_i^\ell(t) + x_i^s(t) \mid H_0(t)]$$
$$\stackrel{(b)}{=} \mathbb{L}[x_i^\ell(t) + x_i^s(t) \mid \tilde{H}_i(t)] + \mathbb{L}[x_i^\ell(t) + x_i^s(t) \mid H_0(t)]$$
$$- \mathbb{L}[x_i^\ell(t) + x_i^s(t) \mid H_0(t)]$$
$$= \breve{z}_i^\ell(t|i)$$

where (a) follows from (S2) and (b) uses (44). ■

Theorem 4: The optimal control strategy of Problem 2 is unique and is given by

$$u_0(t) = -L_0^c(t)\hat{x}(t|c) \tag{49a}$$

and for all $i \in N$,

$$u_i(t) = -L_i^c(t)\hat{x}(t|c) - L_i^\ell(t)(\hat{x}_i(t|i) - \hat{x}_i(t|c)) \tag{49b}$$

where $L_i^c(t)$ denote the ith row of $L^c(t)$. The major agent's LLMS estimate follows the same recursive update rule (33) as the major agent's MMSE estimate. Furthermore, the ith minor agent's LLMS estimate is given as follows: $\hat{x}_i(t|0) = 0$ and for $t > 1$:

$$\begin{aligned}\hat{x}_i(t|i) &= A_{ii}\hat{x}_i(t-1|i) + A_{i0}x_0(t-1)\\ &\quad + B_{ii}u_i(t-1) + B_{i0}u_0(t-1) + K_i(t)\tilde{y}_i(t)\end{aligned} \tag{50}$$

where

$$\begin{aligned}\tilde{y}_i(t) = y_i(t) - C_{ii}\big(&A_{i0}x_0(t-1) + A_{ii}\hat{x}_i(t-1|i)\\ &+ B_{i0}u_0(t-1) + B_{ii}u_i(t-1)\big)\end{aligned}$$

and $K_i(t)$ is computed by the following standard recursive least square equations: $K_i(1) = 0$, and for $t > 1$,

$$K_i(t) = \mathcal{K}(P_i(t-1), A_{ii}, C_{ii}, \Sigma_i^w, \Sigma_i^v). \tag{51}$$

Finally in the above equation, $P_i(t) = \mathrm{var}(x_i(t) - \hat{x}_i(t|i))$ and can be recursively updated as follows. $P_i(1) = \Sigma_i^x$, and for $t > 1$,

$$P_i(t) = \mathcal{F}(P_i(t-1), A_{ii}, C_{ii}, \Sigma_i^w, \Sigma_i^v)$$

Proof: The structure of optimal policies for the major agent follows from Lemma 18 and Theorem 3.

The update of the major agent's MMSE estimate in Theorem 2 is linear. Hence, the major agent's LLMS estimate is the same as the MMSE estimate and follows the same recursive equations.

To prove the update of the ith agent's LLMS estimate, we split the state of agent i into two components: $x_i(t) = x_i^g(t) + x_i^w(t)$, where

$$\begin{aligned}x_i^g(t+1) &= A_{ii}x_i^g(t) + A_{i0}x_0(t) + B_{ii}u_i(t) + B_{i0}u_0(t)\\ x_i^w(t+1) &= A_{ii}x_i^w(t) + w_i(t).\end{aligned}$$

Based on this splitting of state, we split the observation of agent $i \in N$ into two components as follows: $y_i(t) = y_i^g(t) + y_i^w(t)$, where

$$y_i^g(t) = C_{ii}x_i^g(t), \quad \text{and} \quad y_i^w(t) = C_{ii}x_i^w(t) + v_i(t).$$

Observe that $x_i^w(t)$ and $y_i^w(t)$ do not depend on the control actions at agent $i \in N$. Now, we have

$$\begin{aligned}\hat{x}_i(t|i) = \mathbb{L}[x_i(t)|I_i(t)] &\overset{(a)}{=} x_i^g(t) + \mathbb{L}[x_i^w(t)|I_i(t)]\\ &\overset{(b)}{=} x_i^g(t) + \mathbb{L}[x_i^w(t)|x_0^w(1{:}t), y_i^w(1{:}t)]\\ &\overset{(c)}{=} x_i^g(t) + \mathbb{L}[x_i^w(t)|y_i^w(1{:}t)]\end{aligned} \tag{52}$$

where (a) follows from the state split to $x_i^g(t)$ and $x_i^w(t)$, (b) follows from static reduction argument similar to the one presented in Lemma 2, and (c) follows from Assumption 1.

Let us define $\hat{x}_i^w(t|i) = \mathbb{L}[x_i^w(t)|y_i^w(1{:}t)]$. Observe that $\hat{x}_i^w(t|i)$ can be recursively updated using the standard LLMS updates [4] as follows:

$$\hat{x}_i^w(t|i) = A_{ii}\hat{x}_i^w(t-1|i) + K_i(t)\tilde{y}_i^w(t) \tag{53}$$

where

$$\tilde{y}_i^w(t) = y_i^w(t) - C_{ii}A_{ii}\hat{x}_i^w(t-1|i)$$

and $K_i(t)$ is given by (51) where $P_i(t) = \mathrm{var}(x_i^w(t) - \hat{x}_i^w(t|i)) = \mathrm{var}(x_i(t) - \hat{x}_i(t|i))$, which follows from (52). Note that (52) also implies that

$$\begin{aligned}\tilde{y}_i^w(t) &= y_i(t) - y_i^g(t) - C_{ii}A_{ii}\hat{x}_i^w(t-1|i)\\ &= y_i(t) - C_{ii}(x_i^g(t) + A_{ii}\hat{x}_i^w(t-1|i))\\ &= \tilde{y}_i(t)\end{aligned} \tag{54}$$

where we use the dynamics of $x_i^g(t)$ and (52) to simplify the last step.

Finally, to show the recursive form of $\hat{x}_i(t|i)$, substitute (53) in (52), to get

$$\begin{aligned}\hat{x}_i(t|i) &= x_i^g(t) + \hat{x}_i^w(t|i)\\ &= A_{ii}x_i^g(t-1) + A_{i0}x_0(t-1) + B_{ii}u_i(t-1)\\ &\quad + B_{i0}u_0(t-1) + A_{ii}\hat{x}_i^w(t-1|i) + K_i(t)\tilde{y}_i^w(t)\\ &= A_{ii}\hat{x}_i(t-1|i) + A_{i0}x_0(t-1) + B_{ii}u_i(t-1)\\ &\quad + B_{i0}u_0(t-1) + K_i(t)\tilde{y}_i^w(t).\end{aligned}$$

The result then follows from substituting (54) in the above equation. ■

Remark 4: The best linear strategies derived in Theorem 4 have a similar structure to the best linear strategies derived in [21] using spectral factorization techniques for a model with only one minor agent and stable A.

Remark 5: Due to the separation of estimation and control, the difference in performance J^* of the optimal policy derived in Theorem 2 and the performance J_A^* of the best linear policy derived in Theorem 4 depends on the difference in error covariance between MMSE and LLMS filters. This error covariance depends on the exact distribution of the non-Gaussian noise. There is evidence to suggest that MMSE filters can perform significantly better than LLMS filters in some settings (low signal-to-noise ratio with a noise that differs significantly from Gaussian) [36].

D. Implementation of the Optimal Control Strategy

Remarkably, the implementation of the best linear control strategy is exactly the same as that of the optimal strategy with one difference: The minor agents use a recursive least squares filter instead of a Bayesian filter to update the estimate $\hat{x}_i(t|i)$. The rest of the implementation is the same as described in Section IV-B.

VI. Discussion and Conclusion

We consider a decentralized linear quadratic system with a major agent and a collection of minor agents with a partially nested information structure and partial output feedback. The key feature of our model is that we do not assume that the noise has a Gaussian distribution. Therefore, the optimal strategy is not necessarily linear. Nonetheless, we show that the optimal strategy has an elegant structure and the following salient features.

1) The common component $u^c(t)$ of the control actions is a linear function of the major agent's MMSE estimate $\hat{x}(t|c)$ of the system state. The MMSE estimate $\hat{x}(t|c)$ can be updated using a linear filter and the corresponding gains $L^c(t)$ are computed from the solution of a "global" Riccati equation.
2) The local component $u_i^\ell(t)$ of the control action at minor agent i is a linear function of offset between the minor agent's MMSE estimate $\hat{x}_i(t|i)$ of the minor agent's state and the major agent's estimate $\hat{x}_i(t|c)$ of the minor agent's state. The corresponding gains $L_i^\ell(t)$ are computed from the solution of a "local" Riccati equation.
3) The minor agent's MMSE estimate $\hat{x}_i(t|i)$ is, in general, a nonlinear function of the data $I_i(t)$. Thus, the optimal strategy of the minor agent is a nonlinear function of its data. Nonetheless, the update (35) of the conditional density does not depend on the control strategy. Thus, there is a separation between estimation and control.

Interestingly, the optimal strategy is closely related to the best linear strategy. The best linear strategy has the following salient features.

1) Since the major agents' MMSE estimate $\hat{x}(t|c)$ is a linear function of the data, the major agent's LLMS estimate is the same as the MMSE estimate. Therefore, the common component $u^c(t)$ of the control actions remains the same as the optimal controller.
2) The minor agent's LLMS estimate $\hat{x}_i(t|i)$ is updated according to the recursive least squares filter rather than the Bayesian filter used for updating MMSE estimates.
3) Therefore, the structure of the best linear controller is the same as the structure of the optimal control with the exception that the minor agent's MMSE estimate of its local state are replaced by its LLMS estimates!

In light of the results presented in this article, a natural question is whether these salient features are specific to the model presented in this article or they hold for more general models with delayed sharing of information and coupling between minor agents as well. We hope to be able to address these questions in the future.

Appendix A
Proof of Lemma 1

We prove each property separately.

P1) $u_0(t)$ is a function of $I_0(t)$ which, by (10), equals $I^c(t)$. Thus, $u_0^c(t) = u(t)$, and hence, $u_0^\ell(t) = 0$.

P2) This follows from (P1) and the fact that A and B matrices are block lower triangular.

P3) This follows from the definition of $u_i^\ell(t)$.

P4) This follows from the following:

$$\mathbb{E}[u^c(t)^\intercal M u^\ell(t)] \overset{(a)}{=} \mathbb{E}[\mathbb{E}[u^c(t)^\intercal M u^\ell(t)|I^c(t)]] \overset{(b)}{=} \mathbb{E}[u^c(t)^\intercal M \mathbb{E}[u^\ell(t)|I^c(t)]] = 0$$

where (a) uses the towering property and (b) uses (P3).

P5) This follows from (P4) and the smoothing property of conditional expectation.

P6) By construction, $x^c(t)$ is a function of $u^c(1{:}t-1)$, which, by definition, is a function of $I^c(t)$.

Appendix B
Proof of Lemma 2

For notational convenience, we use $S_A \leftsquigarrow S_B$ to denote that set S_A is a function of set S_B. Note that the relation $\leftsquigarrow$ is transitive.

We consider the cases $i = 0$ and $i \neq 0$ separately. For both cases, we will show that $I_i(t) \leftsquigarrow I_i^s(t)$ and $I_i^s(t) \leftsquigarrow I_i(t)$.

For $i = 0$, first note that (P2) implies

$$x_0(t) = x_0^c(t) + x_0^s(t). \tag{55}$$

By construction $u_0^c(t) \leftsquigarrow u_0(1{:}t-1) \subset I_0(t)$. Thus, $x_0^s(t) = x_0(t) - x_0^c(t)$, both of which are functions of $I_0(t)$. Hence, $I_0^s(t) \leftsquigarrow I_0(t)$.

We prove the reverse implication by induction. Note that $x_0(1) = x_0^s(1)$. Thus, $I_0(1) \leftsquigarrow I_0^s(1)$. This forms the basis of induction. Now assume that $I_0(t) \leftsquigarrow I_0^s(t)$ and consider $I_0(t+1) = \{I_0(t), x_0(t+1), u_0(t)\}$. Since $u_0(t) \leftsquigarrow I_0(t)$ and, by the induction hypothesis, $I_0(t) \leftsquigarrow I_0^s(t)$, we have $u_0(t) \leftsquigarrow I_0^s(t)$. Moreover, by (55), $x^c(t) = x_0(t) - x_0^s(t)$ and, therefore, by the induction hypothesis, $x^c(t) \leftsquigarrow I_0^s(t)$. Since both $u_0(t) \leftsquigarrow I_0^s(t)$ and $x^c(t) \leftsquigarrow I_0^s(t)$, we have $x_0^c(t+1) \leftsquigarrow I_0^s(t)$, and hence, $x_0^c(t+1) \leftsquigarrow I_0(s)$. By (55), $x_0(t+1) = x_0^c(t+1) + x_0^s(t+1)$. Hence, $x_0(t+1) \leftsquigarrow I_0^s(t+1)$. Thus, we have shown that each components of $I_0(t+1) = \{I_0(t), x_0(t+1), u_0(t)\} \leftsquigarrow I_0^s(t+1)$. Thus, by induction, $I_0(t) \leftsquigarrow I_0^s(t)$.

We have thus shown that $I_0^s(t) \leftsquigarrow I_0(t)$ and $I_0(t) \leftsquigarrow I_0^s(t)$. This proves that $I_0(s) \equiv I_0^s(t)$.

Now consider $i \neq 0$. By construction, $x_i^c(t) + x_i^\ell(t) \leftsquigarrow \{u_0(1{:}t-1), u_i(1{:}t-1)\} \subset I_i(t)$. Thus, $y_i^c(t) + y_i^\ell(t) \leftsquigarrow I_i(t)$ and, hence $y_i^s(t) = y_i(t) - y_i^c(t) - y_i^\ell(t)$ is a function of $I_i(t)$. We have already shown that $x_0^s(1{:}t) \leftsquigarrow x_0(1{:}t)$. Thus, $I_i^s(t) \leftsquigarrow I_i(t)$.

We prove the reverse implication by induction. Note that $y_i^c(1) = y_i^\ell(1) = 0$. Thus, $y_i(1) = y_i^s(1)$ and, as shown before $x_0(1) = x_0^s(1)$. Thus, $I_i(1) \leftsquigarrow I_i^s(1)$. This forms the basis of induction. Now assume that $I_i(t) \leftsquigarrow I_i^s(t)$ and consider $I_i(t+1) = \{I_i(t), x_0(t+1), u_0(t), y_i(t+1), u_i(t)\}$. We have already shown that $x_0(t+1)$ and $u_0(t)$ are functions of $I_0^s(t+1) \subset I_i^s(t+1)$. For $u_i(t)$, observe that $u_i(t) \leftsquigarrow I_i(t)$ and therefore, by the induction hypothesis, $u_i(t) \leftsquigarrow I_i^s(t)$. As was the case for $i = 0$, we can argue that $x_i^c(t+1) + x_i^\ell(t+1) \leftsquigarrow I_i^s(t)$, and therefore, $y_i^c(t+1) + y_i^\ell(t+1) \leftsquigarrow I_i^s(t)$. Thus, from (14), $y_i(t+1) \leftsquigarrow I_i^s(t+1)$. Thus, by induction $I_i(t) \leftsquigarrow I_i^s(t)$.

We have thus shown that $I_i^s(t) \leftsquigarrow I_i(t)$ and $I_i(t) \leftsquigarrow I_i^s(t)$. This proves that $I_i(s) \equiv I_i^s(t)$.

Finally, if $g \in \mathscr{G}_A$, all the relationships $\leftsquigarrow$ in the above argument are linear functions. Thus, $I_i(t)$ and $I_i^s(t)$ are linear functions of each other.

Appendix C
Proof of Lemma 3

We prove each property separately.

P7) For $\tau = t$, the result is the same as (P4). Now consider $\tau < t$. Recall that $I^c(t) = I_0(t)$. Thus, by Lemma 2

$$\mathbb{E}[u_i^\ell(t)|I^c(t)] = \mathbb{E}[u_i^\ell(t)|I_0^s(t)].$$

Now observe that

$$\begin{aligned} I_0^s(t) &= \{x_0^s(1{:}t)\} \equiv \{x_0^s(1{:}\tau), w_0(\tau{:}t-1)\} \\ &= \{I_0^s(\tau), w_0(\tau{:}t-1)\}. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E}[u_i^\ell(\tau)|I_0^s(t)] &= \mathbb{E}[u_i^\ell(\tau)|I_0^s(\tau), w_0(\tau{:}t-1)] \\ &\overset{(a)}{=} \mathbb{E}[u_i^\ell(\tau)|I_0^s(\tau)] \overset{(b)}{=} \mathbb{E}[u_i^\ell(\tau)|I_0(\tau)] \overset{(c)}{=} 0 \end{aligned}$$

where *(a)* holds because $u_0^\ell(\tau)$ is independent of future noise $w_0(\tau{:}t-1)$, *(b)* uses Lemma 2, and *(c)* follows from (P4).

P8) Combining (13b) and (P1), we get

$$x_i^\ell(\tau) = \sum_{\sigma=1}^{\tau-1} A_{ii}^{\sigma-1} B_{ii} u_i^\ell(\tau-\sigma).$$

Hence, the result follows from (P7).

P9) By the smoothing property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}[(x_i^\ell(t))^\intercal M x_0^s(t)] &= \mathbb{E}\big[\mathbb{E}[(x_i^\ell(t))^\intercal M x_0^s(t)|I_0^s(t)]\big] \\ &\overset{(a)}{=} \mathbb{E}\big[\mathbb{E}[(x_i^\ell(t))^\intercal|I_0^s(t)]\, M x_0^s(t)\big] \\ &\overset{(b)}{=} 0 \end{aligned}$$

where (a) follows because $x_0^s(t)$ is part of $I_0^s(t)$ and (b) follows from Lemma 2 and (P8).

P10) By the smoothing property of conditional expectation, we have

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

where (a) follows because $x^c(t)$ is a function of $I^c(t)$ and (b) follows from (P8).

P11) By the smoothing property of conditional expectation, we have

$$\begin{aligned} \mathbb{E}[(u_i^\ell(t))^\intercal M x_0^s(t)] &= \mathbb{E}\big[\mathbb{E}[(u_i^\ell(t))^\intercal M x_0^s(t)|I^c(t)]\big] \\ &\overset{(a)}{=} \mathbb{E}\big[\mathbb{E}[(u_i^\ell(t))^\intercal|I^c(t)]\, M x_0^s(t)\big] \\ &\overset{(b)}{=} 0 \end{aligned}$$

where (a) follows because $x_0^s(t)$ is in $I_0^s(t)$, and therefore, a function of $I^c(t)$ and (b) follows from (P4).

Appendix D
Proof of Lemma 4

We prove each part separately.

1) Arbitrarily fix a strategy $g \in \mathcal{G}$ and define the following σ-algebras:

$$\begin{aligned} &\mathcal{F}_0(t) = \sigma(x_0(1), w_0(1{:}t-1)) \\ &\mathcal{F}_i(t) = \sigma(x_0(1), x_i(1), w_0(1{:}t-1), w_i(1{:}t-1)), i \in N. \end{aligned}$$

It follows from Assumption 1 that $\{\mathcal{F}_i(t)\}_{i\in N}$ are conditionally independent given $\mathcal{F}_0(t)$. From an argument similar to the one used in the proof of Lemma 2, we can show that $x_i(t)$ is function (which may depend on the strategy g) of $(x_0(1), x_i(1), w_0(1:t-1), w_i(1:t-1))$. Thus, for any Borel measurable subset $D_i(t)$ of $\mathbb{R}^{t(d_x^i+d_u^i)}$, the event $E_i(t) = \{(x_i(1{:}t), u_i(1{:}t)) \in D_i(t)\}$ is $\mathcal{F}_i(t)$ measurable.

Similarly, from an argument similar to Lemma 2, we can show that $\sigma(I^c(t)) = \sigma(I_0^s(t)) = \mathcal{F}_0(t)$. Thus,

$$\begin{aligned} &\mathbb{P}(\{(x_i(1{:}t), u_i(1{:}t)) \in D_i(t)\}_{i\in N}|I^c(t)) \\ &\quad = \mathbb{P}(\{E_i(t)\}_{i\in N}|\mathcal{F}_0(t)) = \prod_{i=1}^n \mathbb{P}(E_i(t)|\mathcal{F}_0(t)) \end{aligned}$$

where the last equality follows from the fact that $\{\mathcal{F}_i(t)\}_{i\in N}$ are conditionally independent given $\mathcal{F}_0(t)$.

2) We prove this by induction. For $t = 1$, $x_i^s(1) = x_i(1)$ and $I_0^s(1) = \{x_0^s(1)\} = \{x_0(1)\}$. By Assumption 1, $x_i(1) \perp\!\!\!\perp x_j(1) \mid x_0(1)$. Thus, $x_i^s(1) \perp\!\!\!\perp x_j^s(1) \mid x_0^s(1)$. This forms the basis of induction. Now assume that $x_i^s(1{:}t) \perp\!\!\!\perp x_j^s(1{:}t) \mid I_0^s(t)$. From the dynamics (13c), we have

$$\begin{aligned} &x_0^s(t+1) = A_{00} x_0^s(t) + w_0(t) \\ &x_i^s(t+1) = A_{ii} x_i^s(t) + A_{i0} x_0^s(t) + w_i(t), \quad i \in N. \end{aligned}$$

By Assumption 1, $w_0(t) \perp\!\!\!\perp w_i(t) \perp\!\!\!\perp w_j(t)$. This, combined with the induction hypothesis implies that $x_i^s(1{:}t+1) \perp\!\!\!\perp x_j^s(1{:}t+1) \mid I_0^s(t+1)$. Hence, the result holds by induction.

3) Recall that $x_i^\ell(t) = x_i(t) - x_i^c(t) - x_i^s(t)$ and $u_i^\ell(t) = u_i(t) - u_i^c(t)$. Since $x_i^c(t)$ and $u_i^c(t)$ are functions of $I^c(t)$, the result follows from the result of the previous two parts.

Appendix E
Proof of Lemma 5

First consider (17). Since $x(t) = z^c(t) + x^\ell(t)$, we have

$$\begin{aligned} \mathbb{E}\big[x(t)^\intercal Q x(t)\big] = \mathbb{E}\Big[&z^c(t)^\intercal Q z^c(t) + x^\ell(t)^\intercal Q x^\ell(t) \\ &+ 2x^\ell(t)^\intercal Q z^c(t)\Big]. \end{aligned} \tag{56}$$

Now from (P2) and Lemma 4, we have

$$\mathbb{E}[x^\ell(t)^\intercal Q x^\ell(t)] = \sum_{i\in N} \mathbb{E}[x_i^\ell(t)^\intercal Q_{ii} x_i^\ell(t)]. \tag{57}$$

From (P10), we have

$$\begin{aligned}\mathbb{E}[x^\ell(t)^\intercal Q z^c(t)] &= \mathbb{E}[x^\ell(t)^\intercal Q x^s(t)]\\ &= \sum_{i\in N} \mathbb{E}[x_i^\ell(t)^\intercal Q_{ii} x_i^s(t)]\end{aligned} \tag{58}$$

where the last equality follows from (P2), (P9), and Lemma 4.

Substituting (57) and (58) into (56) and completing the squares, we get (17).

Now consider (18). From (P4), we get

$$\mathbb{E}\big[u(t)^\intercal R u(t)\big] = \mathbb{E}\left[u^c(t)^\intercal R u^c(t) + u^\ell(t)^\intercal R u^\ell(t)\right]. \tag{59}$$

From (P1) and Lemma 4, we get

$$\mathbb{E}[u^\ell(t)^\intercal R u^\ell(t)] = \sum_{i\in N} \mathbb{E}[u_i^\ell(t)^\intercal R_{ii} u_i^\ell(t)]. \tag{60}$$

Substituting (60) into (59), we get (18).

Appendix F
Proof of Lemma 8

We prove each property separately.

C1) For $\tilde z^c(t)$, observe that

$$\hat z(t|c) = \mathbb{E}[x^c(t) + x^s(t)|I^c(t)] = x^c(t) + \mathbb{E}[x^s(t)|I_0^s(t)]$$

where the second equality uses (P6) and Lemma 2. Thus

$$\tilde z^c(t) := z^c(t) - \hat z(t|c) = x^s(t) - \mathbb{E}[x^s(t)|I_0^s(t)]$$

which is control-free and depends only on the primitive random variables.

For $\tilde z_i^\ell(t)$, observe that

$$\begin{aligned}\breve z_i^\ell(t|i) &= \mathbb{E}[z_i^\ell(t)|I_i(t)] - \mathbb{E}[z_i^\ell(t)|I_0(t)]\\ &= x_i^\ell(t) + \mathbb{E}[x_i^s(t)|I_i(t)]\\ &\quad - \mathbb{E}[x_i^\ell(t)|I_0(t)] - \mathbb{E}[x_i^s(t)|I_0(t)]\\ &\overset{(a)}{=} x_i^\ell(t) + \mathbb{E}[x_i^s(t)|I_i^s(t)] - \mathbb{E}[x_2^s(t)|I_0^s(t)]\end{aligned}$$

where (a) uses Lemma 2 and (P8). Thus

$$\begin{aligned}\tilde z_i^\ell(t) &= z_i^\ell(t) - \breve z_i^\ell(t|i)\\ &= x_i^s(t) - \mathbb{E}[x_i^s(t)|I_i^s(t)] + \mathbb{E}[x_i^s(t)|I_0^s(t)]\end{aligned}$$

which is control-free and depends only on the primitive random variables.

C2) Observe that

$$\mathbb{E}[\tilde z^c(t)|I^c(t)] = \mathbb{E}[z^c(t) - \hat z(t|c)|I^c(t)] = 0.$$

C3) This follows immediately from the fact that the error of a mean-squared estimator is orthogonal to the estimate.

C4) Using the smoothing property we have,

$$\begin{aligned}\mathbb{E}[u^c(t)M\tilde z^c(t)] &= \mathbb{E}[\mathbb{E}[u^c(t)M\tilde z^c(t)|I^c(t)]]\\ &\overset{(a)}{=} \mathbb{E}[u^c(t)M\mathbb{E}[\tilde z^c(t)|I^c(t)]] \overset{(b)}{=} 0\end{aligned}$$

where *(a)* uses the fact that $u^c(t)$ is measurable with respect to the common information and *(b)* uses (C2).

C5) For ease of notation, define

$$\begin{aligned}\hat d_1(t) &= \mathbb{E}[z_i^\ell(t)|I_i(t)], \quad \tilde d_1(t) = z_i^\ell(t) - \hat d_1(t)\\ \hat d_2(t) &= \mathbb{E}[z_i^\ell(t)|I_0(t)], \quad \tilde d_2(t) = z_i^\ell(t) - \hat d_2(t).\end{aligned}$$

So, we can write

$$\begin{aligned}z_i^\ell(t) &= \hat d_1(t) + \tilde d_1(t) = \hat d_2(t) + \tilde d_2(t)\\ \breve z_i^\ell(t|i) &= \hat d_1(t) - \hat d_2(t)\\ \tilde z_i^\ell(t) &= z_i^\ell(t) - \hat d_1(t) + \hat d_2(t) = \tilde d_1(t) + \hat d_2(t).\end{aligned}$$

From the orthogonality principle, $\tilde d_1(t) \perp \hat d_1(t)$ and $\tilde d_2(t) \perp \hat d_2(t)$. Since $I_0(t)$ is a subset of $I_i(t)$, $\tilde d_1(t) \perp \hat d_2(t)$. Then, we have

$$\begin{aligned}\mathbb{E}[(\tilde z_i^\ell(t))^\intercal \breve z_i^\ell(t|i)] &= \mathbb{E}[(\tilde d_1(t) + \hat d_2(t))^\intercal(\hat d_1(t) - \hat d_2(t))]\\ &= \mathbb{E}[\hat d_2(t)^\intercal(\hat d_1(t) - \hat d_2(t))]\\ &= \mathbb{E}[\hat d_2(t)^\intercal(\tilde d_2(t) - \tilde d_1(t))]\\ &= 0.\end{aligned} \tag{61}$$

C6) Recall the definitions of $\hat d_1(t)$ and $\hat d_2(t)$ from the proof of (C5). Since $\tilde z_i^\ell(t) = \tilde d_1(t) + \hat d_2(t)$, we have

$$\begin{aligned}\mathbb{E}[u_i^\ell(t)^\intercal M \tilde z_i^\ell(t)] &= \mathbb{E}[u_i^\ell(t)^\intercal M \tilde d_1(t)]\\ &\quad + \mathbb{E}[u_i^\ell(t)^\intercal M \hat d_2(t)].\end{aligned}$$

Now, we show that both terms are zero. Consider

$$\begin{aligned}\mathbb{E}[u_i^\ell(t)^\intercal M \tilde d_1(t)] &= \mathbb{E}[\mathbb{E}[u_i^\ell(t)^\intercal M \tilde d_1(t) \mid I_i(t)]]\\ &\overset{(a)}{=} \mathbb{E}[u_i^\ell(t)^\intercal M \mathbb{E}[\tilde d_1(t) \mid I_i(t)]]\\ &\overset{(b)}{=} 0\end{aligned}$$

where (a) follows because $u_i^\ell(t)$ is a function of $I_i(t)$ and (b) follows from the definition of $\tilde d_1(t)$. Now consider

$$\begin{aligned}\mathbb{E}[u_i^\ell(t)^\intercal M \hat d_2(t)] &= \mathbb{E}[\mathbb{E}[u_i^\ell(t)^\intercal M \hat d_2(t) \mid I_0(t)]]\\ &\overset{(c)}{=} \mathbb{E}[\mathbb{E}[u_i^\ell(t)^\intercal \mid I_0(t)] M \hat d_2(t)]\\ &\overset{(d)}{=} 0,\end{aligned}$$

where (c) follows from the definition of $\hat d_2(t)$ and (d) follows from (P4).

Appendix G
Proof of Lemma 15

We prove each property separately.

C1) For $\tilde z^c(t)$, observe that

$$\hat z(t|c) = \mathbb{L}[x^c(t) + x^s(t)|H_0(t)] = x^c(t) + \mathbb{L}[x^s(t)|H_0^s(t)]$$

where the second equality uses (S2) and Remark 12. Thus

$$\tilde{z}^c(t) := z^c(t) - \hat{z}(t|c) = x^s(t) - \mathbb{L}[x^s(t)|H_0^s(t)]$$

which is control-free and depends only on the primitive random variables.

For $\tilde{z}_i^\ell(t)$, observe that

$$\begin{aligned}\tilde{z}_i^\ell &= z_i^\ell(t) - \mathbb{L}[z_i^\ell(t)|\tilde{H}_i(t)]\\ &= x_i^\ell(t) + x_i^s(t) - \mathbb{L}[x_i^\ell(t) + x_i^s(t)|\tilde{H}_i(t)]\\ &\overset{(a)}{=} x_i^s(t) - \mathbb{L}[x_i^s(t)|\tilde{H}_i(t)]\\ &\overset{(b)}{=} x_i^s(t) - \mathbb{L}[x_i^s(t)|\tilde{H}_i^s(t)]\end{aligned}$$

where (a) uses (S3) and (b) uses Remark 12. Thus, $\tilde{z}_i^\ell(t)$ is control-free and depends only on the primitive random variables.

C3) By definition, $M\hat{z}(t|c)$ is a linear function of $I^c(t)$. Hence, $\mathbb{E}[\tilde{z}^c(t)^\intercal M\hat{z}(t|c)] = 0$ by (43).

C4) $M^\intercal u^c(t)$ is a linear function of $u^c(t)$ and, hence, by (S1) belongs to $H_0(t)$. Hence, $\mathbb{E}[\tilde{z}^c(t)^\intercal M^\intercal u^c(t)] = 0$ by (43). Therefore, $\mathbb{E}[u^c(t)^\intercal M\tilde{z}^c(t)] = 0$.

C5) Again by definition, $M\breve{z}_i^\ell(t|i)$ is a linear function of $\tilde{I}_i(t)$. Hence, $\mathbb{E}[\tilde{z}_i^\ell(t)^\intercal M\breve{z}_i^\ell(t|i)] = 0$ by (43).

C6) $M^\intercal u_i^\ell(t)$ is a linear function of $u_i^\ell(t)$ which belongs to $\tilde{H}_i(t)$ by Lemma 13, and, hence, is a linear function of $\tilde{I}_i(t)$. Therefore $\mathbb{E}[\tilde{z}_i^\ell(t)^\intercal M^\intercal u_i^\ell(t)] = 0$ by (43), which results in $\mathbb{E}[u_i^\ell(t)^\intercal M\tilde{z}_i^\ell(t)] = 0$.

References

[1] W. M. Wonham, "On the separation theorem of stochastic control," *SIAM J. Control*, vol. 6, no. 2, pp. 312–326, 1968.

[2] J. G. Root, "Optimum control of non-Gaussian linear stochastic systems with inaccessible state variables," *SIAM J. Control*, vol. 7, no. 2, pp. 317–323, 1969.

[3] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Belmont, MA, USA: Athena Scientific, 2000.

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.

[5] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM J. Control*, vol. 6, no. 1, pp. 131–147, 1968.

[6] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems—Part I," *IEEE Trans. Autom. Control*, vol. AC-17, no. 1, pp. 15–22, Feb. 1972.

[7] T. Yoshikawa, "Dynamic programming approach to decentralized stochastic control problems," *IEEE Trans. Autom. Control*, vol. AC-20, no. 6, pp. 796–797, Dec. 1975.

[8] P. Varaiya and J. Walrand, "On delayed sharing patterns," *IEEE Trans. Autom. Control*, vol. AC-23, no. 3, pp. 443–445, Jun. 1978.

[9] M. Afshari and A. Mahajan, "Multi-agent estimation and filtering for minimizing team mean-squared error," *IEEE Trans. Signal Process.*, vol. 69, pp. 5206–5221, Aug. 2021.

[10] N. Nayyar, D. Kalathil, and R. Jain, "Optimal decentralized control with asymmetric one-step delayed information sharing," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 1, pp. 653–663, Mar. 2018.

[11] H. R. Feyzmahdavian, A. Alam, and A. Gattami, "Optimal distributed controller design with communication delays: Application to vehicle formations," in *Proc. IEEE Conf. Decis. Control*, 2012, pp. 2232–2237.

[12] L. Lessard and A. Nayyar, "Structural results and explicit solution for two-player LQG systems on a finite time horizon," in *Proc. IEEE Conf. Decis. Control*, 2013, pp. 6542–6549.

[13] M. Afshari and A. Mahajan, "Optimal local and remote controllers with unreliable uplink channels: An elementary proof," *IEEE Trans. Autom. Control*, vol. 65, no. 8, pp. 3616–3622, Aug. 2020.

[14] B. Bamieh and P. G. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Syst. Control Lett.*, vol. 54, no. 6, pp. 575–583, 2005.

[15] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 274–286, Feb. 2006.

[16] P. Whittle and J. Rudge, "The optimal linear solution of a symmetric team control problem," *J. Appl. Probability*, vol. 11, no. 2, pp. 377–381, 1974.

[17] P. Shah and P. A. Parrilo, "H_2-optimal decentralized control over posets: A. state-space solution for state-feedback," *IEEE Trans. Autom. Control*, vol. 58, no. 12, pp. 3084–3096, Dec. 2013.

[18] J. Swigart and S. Lall, "An explicit state-space solution for a decentralized two-player optimal linear-quadratic regulator," in *Proc. Amer. Control Conf.*, 2010, pp. 6385–6390.

[19] J. Swigart and S. Lall, "An explicit dynamic programming solution for a decentralized two-player optimal linear-quadratic regulator," in *Proc. Math. Theory Netw. Syst.*, 2010, pp. 1443–1447.

[20] J. Swigart and S. Lall, "Optimal synthesis and explicit state-space solution for a decentralized two-player linear-quadratic regulator," in *Proc. IEEE Conf. Decis. Control*, 2010, pp. 132–137.

[21] J. Swigart and S. Lall, "Optimal controller synthesis for a decentralized two-player system with partial output feedback," in *Proc. Amer. Control Conf.*, 2011, pp. 317–323.

[22] L. Lessard and S. Lall, "A state-space solution to the two-player decentralized optimal control problem," in *Proc. Annu. Allerton Conf. Commun., Control, Comput.*, 2011, pp. 1559–1564.

[23] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," in *Proc. Amer. Control Conf.*, 2012, pp. 6314–6321.

[24] L. Lessard, "Optimal control of a fully decentralized quadratic regulator," in *Proc. Annu. Allerton Conf. Commun., Control, Comput.*, 2012, pp. 48–54.

[25] L. Lessard and S. Lall, "Optimal control of two-player systems with output feedback," *IEEE Trans. Autom. Control*, vol. 60, no. 8, pp. 2129–2144, Aug. 2015.

[26] J. H. Kim and S. Lall, "A unifying condition for separable two player optimal control problems," in *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 3818–3823.

[27] J. H. Kim and S. Lall, "Separable optimal cooperative control problems," in *Proc. Amer. Control Conf.*, 2012, pp. 5868–5873.

[28] L. Lessard, "Decentralized LQG control of systems with a broadcast architecture," in *Proc. IEEE Conf. Decis. Control*, 2012, pp. 6241–6246.

[29] A. Mahajan and A. Nayyar, "Sufficient statistics for linear control strategies in decentralized systems with partial history sharing," *IEEE Trans. Autom. Control*, vol. 60, no. 8, pp. 2046–2056, Aug. 2015.

[30] M. Huang, "Large-population LQG games involving a major player: The Nash certainty equivalence principle," *SIAM J. Control Optim.*, vol. 48, no. 5, pp. 3318–3353, Jan. 2010.

[31] P. E. Caines and A. C. Kizilkale, "ϵ-Nash equilibria for partially observed LQG mean field games with a major player," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3225–3234, Jul. 2017.

[32] D. Firoozi and P. E. Caines, "ϵ-Nash equilibria for major minor LQG mean field games with partial observations of all agents," *IEEE Trans. Automat. Control*, vol. 66, no. 6, pp. 2778–2786, Jun. 2021, doi: 10.1109/TAC.2020.3010129.

[33] J.-M. Lasry and P.-L. Lions, "Mean-field games with a major player," *Comptes Rendus Mathematique*, vol. 356, no. 8, pp. 886–890, Aug. 2018.

[34] 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.

[35] H. S. Witsenhausen, "Equivalent stochastic control problems," *Math. Control Signals Syst.*, vol. 1, pp. 3–11, 1988.

[36] B. S. Rao and H. F. Durrant-Whyte, "Fully decentralised algorithm for multisensor Kalman filtering," *Proc. Inst. Elect. Eng. D.—Control Theory Appl.*, vol. 138, no. 5, pp. 413–420, Sep. 1991.

Mohammad Afshari (Member, IEEE) received the B.S. and the M.S. degrees in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 2010 and 2012, respectively, and the Ph.D. degree in electrical and computer engineering from McGill University, Montreal, QC, Canada, in 2021. From 2021 to 2022, he was a Postdoctoral Fellow with the University of Alberta, Alberta, Canada. Currently, he is a Postdoctoral Fellow with the Georgia Institute of Technology, Atlanta, GA, USA.

His current research interests include decentralized stochastic control, multi-agent systems, and reinforcement learning.

Mr Afshari is a member of the Institute for Robotics and Intelligent Machines (IRIM) and the Dynamics and Control Systems Laboratory (DCSL) at Georgia Institute of Technology.

Aditya Mahajan (Senior Member, IEEE) received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 2003, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Michigan, Ann Arbor, MI, USA, in 2006 and 2008, respectively.

From 2008 to 2010, he was a Postdoctoral Researcher with Yale University, New Haven, CT, USA. Since 2010, he has been with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, where he is currently an Associate Professor. His principal research interests include learning and control of decentralized multiagent systems, multiarmed bandits, and reinforcement learning.

Dr. Mahajan is currently an Associate Editor for the IEEE TRANSACTIONS OF AUTOMATIC CONTROL, IEEE CONTROL SYSTEM LETTERS, and *Mathematics of Control, Signal, and Systems*. He was an Associate Editor for the IEEE Control Systems Society Conference Editorial Board from 2014 to 2017. He is the recipient of the 2015 George Axelby Outstanding Paper Award, the 2014 CDC Best Student Paper Award (as a supervisor), and the 2016 NecSys Best Student Paper Award (as a supervisor).