

Joint Communication-Control Strategy Optimization with Partially Nested Information Structures: The Linear-Quadratic Case

Haoyi You[†]

Kaiqing Zhang[†]

Abstract

In this paper, we formalize a joint communication-control strategy optimization (JCCO) problem in multi-agent linear systems with quadratic costs, under the common-information-based (CIB) framework from decentralized stochastic control. For computational tractability, we focus on such JCCO problems with partially nested (PN) information structures (ISs). In particular, with a baseline communication protocol that leads to a PN IS, we establish a series of conditions under which the partial nestedness is preserved under the (additional) communication strategies to be optimized, while violating them may cause nonlinearity or non-existence of the optimal control strategies, with open-loop communication strategies. We then develop a dynamic-programming-based approach to compute the optimal control strategies of JCCO with open-loop communication strategies, which yields a set of closed-form Riccati Equations. As a byproduct of independent interest, such an approach also offers a new way to solve decentralized linear-quadratic control with PN ISs and *output* feedback, under the CIB framework. Finally, we extend such an approach to JCCOs with closed-loop communication strategies, yielding a more tractable dynamic program than an infinite-dimensional CIB-belief-based one.

I. Introduction

The design of communication protocols jointly with control has been extensively studied in both the control theory and multi-agent learning literature, from different perspectives and under different models. For example, [1, 2] investigated the design of communication channels with capacity constraints in networked control systems; [3, 4] studied communication architecture and control strategy co-design via norm minimization; [5, 6, 7] investigated the joint optimization of the quantization and control policies, with the quantized signals being communicated to the controller; and [8, 9] developed heuristic learning-to-communicate algorithms to jointly learn the control and communication strategies for accumulated reward maximization.

In this paper, we formalize a joint communication-control strategy optimization (JCCO) problem in multi-agent linear systems with quadratic costs. Motivated by the empirical studies [8, 9], we model communication as some *actions/strategies* to be optimized *jointly* with the control strategies, under an *optimal control* formulation to minimize some accumulated costs. The agents may follow some fixed *baseline* communication protocols for information sharing, and then share their private information through *additional sharing*, following their communication strategies. We

[†]The authors are affiliated with the University of Maryland, College Park, MD, USA, 20742. Emails: {yuri-iyou, kaiqing}@umd.edu.

formalize the problem under the common-information-based (CIB) framework from decentralized stochastic control [10, 11], which has also been adopted in several recent studies on control-communication strategy co-optimization [12, 13, 14]. In comparison, [12, 13] focused specifically on sharing instantaneous observations and on systems with *decoupled* state dynamics. Our formalism in [14] is more general and allows the sharing of private-information histories in discrete-space problems, together with computational and sample complexity analyses. We here focus on an instantiation of our formalism in [14] in the continuous-space, linear-quadratic (LQ) setting, which yields fundamentally different technical challenges as detailed below.

To facilitate tractable computation of the optimal strategies, we focus on deriving *dynamic-programming* (DP)-based approaches to solve JCCO. To this end, we concentrate on such co-optimization problems with partially nested (PN) information structures (ISs) [15]. Indeed, with a *fixed* open-loop communication strategy (and thus a fixed IS), partial nestedness is known to be a favorable IS that not only ensures the *linearity* of an optimal control strategy, but also yields a *convex reformulation* of the decentralized LQ control problem [15]. However, as pointed out in [16], the reformulated convex program can be too large to be computationally efficient, which precisely motivated the DP-based solution therein. Moreover, the DP-based approach can provide more insights into the structure of the optimal controller [16], by identifying the proper *sufficient statistics* for decision-making [16, 10, 17, 18].

Yet, the results in [16, 18] do not apply to our setting, as they focused on the special cases with *factorized* states, and with either *state-feedback* [16] or *multi-tree* coupling and communication graphs [18]. Under the CIB framework [10], [17] considered the output-feedback setting, and showed that when control strategies are *restricted to be linear*, and the private-information component of the strategy is *fixed*, the resulting problem is linear-quadratic-Gaussian (LQG) control, which can be solved via Riccati Equations. However, it remains unclear how to solve the overall decentralized LQ control problem by further optimizing over the private-information component. Since given a *fixed* communication strategy, our JCCO problem reduces to a decentralized LQ control one, we need to advance these results in order to fully address JCCO. We thus make the following contributions:

Contributions. (i) We formalize *joint communication-control strategy optimization* as an optimal control problem with linear systems and quadratic costs, under the common-information-based framework [10]. (ii) For better computational tractability, we focus on JCCO with *partially nested* baseline ISs [15], and identify structural conditions under which partial nestedness is preserved under additional information sharing, while violating them may cause nonlinearity or non-existence of the optimal control strategies, under *open-loop* communication strategies. (iii) We then develop a closed-form dynamic program, i.e., a set of Riccati Equations, to solve for the optimal control strategy with respect to fixed open-loop communication strategies in JCCO. This program is also of independent interest: it may be viewed as a new way to solve decentralized linear-quadratic control with PN ISs and *output* feedback, under the CIB framework, which thus advances the results in [17]. The key is to identify a novel connection between the (strictly) PN IS [15] and the strategy-independent common-information-based belief (SI-CIB) condition [11] in this linear-quadratic setting. (iv) We then extend such an approach to JCCO with *closed-loop* communication strategies, yielding reduced-dimensional sufficient statistics and a more tractable dynamic program than an infinite-dimensional belief-based one, when applying the CIB framework [10, 11] to general LQ settings directly.

II. Preliminaries

Notation. Random variables are denoted by bold upper case letters, and their realizations are denoted by the corresponding non-bold upper case letters. For any matrix X , we use $X^\dagger$ to denote the pseudo-inverse of X if $X \geq 0$. For any two integers $0 \leq a < b$, we denote $[a : b] := \{a, a+1, \dots, b\}$, and denote $[a] = [1 : a]$. For any vector space $\mathcal{X}$, we denote by $\mathcal{P}(\mathcal{X})$ the space of all probability distributions over $\mathcal{X}$.

II-A Problem formulation

For a team of $n > 1$ agents, a *joint communication-control strategy optimization* problem with *output feedback* can be described by the following tuple: $\mathcal{D} = \langle H, \mathcal{X}, \{\mathcal{Y}_i\}_{i \in [n]}, \{\mathcal{U}_i\}_{i \in [n]}, \{\mathcal{M}_{i,h}\}_{i \in [n], h \in [H]}, \{A_h\}_{h \in [H]}, \{B_{i,h}\}_{i \in [n], h \in [H]}, \{E_{i,h}\}_{i \in [n], h \in [H]}, \{Q_h^1\}_{h \in [H+1]}, \{Q_h^2\}_{h \in [H]}, \{\mathcal{K}_h\}_{h \in [H]} \rangle$, where H is the time horizon and $\mathbf{X}_h \in \mathcal{X} = \mathbb{R}^{d_x}$ is the state. At each timestep $h \in [H]$, each agent $i \in [n]$ receives a noisy observation $\mathbf{Y}_{i,h} \in \mathcal{Y}_i = \mathbb{R}^{d_y^i}$ of the state $\mathbf{X}_h$, and chooses a control action $\mathbf{U}_{i,h} \in \mathcal{U}_i = \mathbb{R}^{d_u^i}$. At timestep $h \in [H]$, we denote by $\mathbf{U}_h = [\mathbf{U}_{1,h}^\top \ \mathbf{U}_{2,h}^\top \ \dots \ \mathbf{U}_{n,h}^\top]^\top$ the joint control action of all the n agents, and by $\mathcal{U} = \mathbb{R}^{\sum_{i=1}^n d_u^i}$ the joint control action space; we denote by $\mathbf{Y}_h = [\mathbf{Y}_{1,h}^\top \ \mathbf{Y}_{2,h}^\top \ \dots \ \mathbf{Y}_{n,h}^\top]^\top$ the joint observation, and by $\mathcal{Y} = \mathbb{R}^{\sum_{i=1}^n d_y^i}$ the joint observation space. The system evolves as follows for each timestep $h \in [H]$:

$$\mathbf{X}_{h+1} = A_h \mathbf{X}_h + \sum_{i=1}^n B_{i,h} \mathbf{U}_{i,h} + \mathbf{W}_{0,h}, \quad \mathbf{Y}_{i,h} = E_{i,h} \mathbf{X}_h + \mathbf{W}_{i,h}, \forall i \in [n], \quad (\text{II.1})$$

where for each $i \in [0 : n], h \in [H]$, $\mathbf{W}_{i,h}$ is a Gaussian random variable with distribution $\mathcal{N}(\mathbf{0}, \Sigma_{i,h})$ for some covariance matrix $\Sigma_{i,h} \geq 0$, while $A_h, E_{i,h}$ are matrices of appropriate dimensions. We define $B_h := [B_{1,h} \ B_{2,h} \ \dots \ B_{n,h}]$ and $E_h := [E_{1,h}^\top \ E_{2,h}^\top \ \dots \ E_{n,h}^\top]^\top$. Here, we assume that the initial state $\mathbf{X}_1$ follows a Gaussian distribution $\mathcal{N}(\mathbf{0}, \Sigma_1)$ with covariance matrix $\Sigma_1 \geq 0$, and that $\mathbf{X}_1$ and $\{\mathbf{W}_{i,h}\}_{i \in [0:n], h \in [H]}$ are mutually independent.

At timestep $h \in [H]$, each agent will share part of her information with other agents. The shared information $\mathbf{Z}_h := \mathbf{Z}_h^b \cup \mathbf{Z}_h^a$ consists of two parts, the *baseline-sharing* part $\mathbf{Z}_h^b$, which originates from some existing sharing protocol, and the *additional-sharing* part $\mathbf{Z}_h^a$, which is *decided/learned* by agents, with joint additional-sharing information $\mathbf{Z}_h^a := \cup_{i=1}^n \mathbf{Z}_{i,h}^a$. The baseline-sharing part is introduced for generality (i.e., $\mathbf{Z}_h^b$ may be set as $\emptyset$), and some benign information structures to be introduced later may require a certain amount of baseline sharing; see [19, 14] for concrete examples. At timestep h , the common information among all the agents is thus defined as the union of all the shared information so far: $\mathbf{C}_{h^-} = \cup_{t=1}^{h-1} \mathbf{Z}_t \cup \mathbf{Z}_h^b$, $\mathbf{C}_{h^+} = \cup_{t=1}^h \mathbf{Z}_t$, where $\mathbf{C}_{h^-}$ and $\mathbf{C}_{h^+}$ denote the common information *before* and *after* additional sharing, respectively. The private information of agent i at timestep h before and after additional sharing is denoted by $\mathbf{P}_{i,h^-}, \mathbf{P}_{i,h^+}$, respectively, where $\mathbf{P}_{i,h^-} \subseteq \{\mathbf{Y}_{1:h}, \mathbf{U}_{1:h-1}\} \setminus \mathbf{C}_{h^-}$, $\mathbf{P}_{i,h^+} \subseteq \{\mathbf{Y}_{1:h}, \mathbf{U}_{1:h-1}\} \setminus \mathbf{C}_{h^+}$. Then, we define $\mathbf{I}_{i,h^-} := \mathbf{P}_{i,h^-} \cup \mathbf{C}_{h^-}$, $\mathbf{I}_{i,h^+} := \mathbf{P}_{i,h^+} \cup \mathbf{C}_{h^+}$ as the information available to agent i before and after additional sharing, respectively. We denote by $\mathbf{P}_{h^-} := [\mathbf{P}_{1,h^-}^\top \ \dots \ \mathbf{P}_{n,h^-}^\top]^\top$ the joint private information at timestep h before additional sharing, and similarly by $\mathbf{P}_{h^+}$ the joint private information after additional sharing. We denote by $\mathcal{C}_{h^-}, \mathcal{C}_{h^+}, \mathcal{P}_{i,h^-}, \mathcal{P}_{i,h^+}, \mathcal{P}_{h^-}, \mathcal{P}_{h^+}, \mathcal{I}_{i,h^-}, \mathcal{I}_{i,h^+}$ the sets of all possible values of the respective random variables $\mathbf{C}_{h^-}, \mathbf{C}_{h^+}, \mathbf{P}_{i,h^-}, \mathbf{P}_{i,h^+}, \mathbf{P}_{h^-}, \mathbf{P}_{h^+}, \mathbf{I}_{i,h^-}, \mathbf{I}_{i,h^+}$.

At each timestep h , in addition to the control action, each agent i needs to choose a commu-

nication action $\mathbf{M}_{i,h} \in \mathcal{M}_{i,h}$ to determine what information $\mathbf{Z}_{i,h}^a$ she will share, as specified below. We use $\mathbf{M}_h := (\mathbf{M}_{1,h}, \dots, \mathbf{M}_{n,h}) \in \mathcal{M}_h$ to denote the joint communication action at timestep h .

We assume that the evolution of common and private information in $\mathcal{D}$ follows the rules below, where we adopt the convention that any quantity at timestep 0 is empty/null.

Assumption II.1 (Information Evolution).

- (a) (Baseline sharing). The baseline sharing evolves as $\mathbf{Z}_h^b = \chi_h(\mathbf{P}_{(h-1)^+}, \mathbf{U}_{h-1}, \mathbf{Y}_h)$ for some fixed projection function χ_h .
- (b) (Additional sharing). For each $i \in [n]$, the additional sharing evolves as $\mathbf{Z}_{i,h}^a = \phi_{i,h}(\mathbf{M}_{i,h}, \mathbf{P}_{i,h^-})$ for some function $\phi_{i,h}$, where for each realization $M_{i,h} \in \mathcal{M}_{i,h}$, $\phi_{i,h}(M_{i,h}, \cdot)$ is a fixed projection function. The joint additional sharing $\mathbf{Z}_h^a$ is thus generated by $\mathbf{Z}_h^a = \phi_h(\mathbf{M}_h, \mathbf{P}_{h^-})$ for some function ϕ_h .
- (c) For each $i \in [n]$, $\mathbf{P}_{i,h^-} = \zeta_{i,h}(\mathbf{P}_{i,(h-1)^+}, \mathbf{U}_{i,h-1}, \mathbf{Y}_{i,h})$ for some fixed projection function $\zeta_{i,h}$, and the joint private information thus evolves as $\mathbf{P}_{h^-} = \zeta_h(\mathbf{P}_{(h-1)^+}, \mathbf{U}_{h-1}, \mathbf{Y}_h)$ for some projection function ζ_h .
- (d) For each $i \in [n]$, $\mathbf{P}_{i,h^+} = \mathbf{P}_{i,h^-} \setminus \mathbf{Z}_{i,h}^a$.
- (e) For each $i \in [n]$ and $h \in [H]$, $\mathbf{Y}_{i,h} \in \mathbf{I}_{i,h^-}$ and $\mathbf{I}_{i,h^-} \subseteq \mathbf{I}_{i,h^+}$; for $h \in [H-1]$, $\mathbf{I}_{i,h^+} \subseteq \mathbf{I}_{i,(h+1)^-}$.

Note that (a) and (c) on baseline sharing follow from those in [11, 19, 14]. (b) and (d) on additional sharing dictate how the communication action affects the additional sharing. For example, a common choice of $(\mathcal{M}_{i,h}, \phi_{i,h})$ is that $\mathcal{M}_{i,h} = \{0, 1\}^{\max_{P_{i,h^-} \in \mathcal{P}_{i,h^-}} |P_{i,h^-}|}$, where $|P_{i,h^-}|$ denotes the number of elements in P_{i,h^-} , and for any $P_{i,h^-} \in \mathcal{P}_{i,h^-}$, $M_{i,h} \in \mathcal{M}_{i,h}$, $\phi_{i,h}(\mathbf{P}_{i,h^-} = P_{i,h^-}, \mathbf{M}_{i,h})$ consists of the k -th element ($k \in [|P_{i,h^-}|]$) of P_{i,h^-} if and only if the k -th element of $M_{i,h}$ is 1. Condition (e) means that each agent has full memory of the available information, and has a *closed-loop* (baseline) information structure with $\mathbf{Y}_{i,h} \in \mathbf{I}_{i,h^-}$, which are standard assumptions also made in [14, 20].

II-B Objective and strategies

At each timestep $h \in [H]$, agents will incur two types of costs, communication cost κ_h and control cost c_h . The control cost inherits the cost of a decentralized LQG problem without communication, see e.g., [17, 16], and is defined as $c_h = \mathbf{X}_h^\top \mathbf{Q}_h^1 \mathbf{X}_h + \mathbf{U}_h^\top \mathbf{Q}_h^2 \mathbf{U}_h$, with a terminal cost of $c_{H+1} = \mathbf{X}_{H+1}^\top \mathbf{Q}_{H+1}^1 \mathbf{X}_{H+1}$. Here, for every $h \in [H]$, $\mathbf{Q}_h^1, \mathbf{Q}_h^2 \geq 0$ and $\mathbf{Q}_{H+1}^1 \geq 0$ are symmetric matrices of appropriate dimensions. The communication cost is determined by the additional sharing and is defined as $\kappa_h = \mathcal{K}_h(\mathbf{M}_h)$ for some function $\mathcal{K}_h : \mathcal{M}_h \rightarrow \mathbb{R}_+$, depending on the communication action $\mathbf{M}_h$ chosen from a finite set $\mathcal{M}_h$.

At each timestep h , each agent i needs to choose a communication action $\mathbf{M}_{i,h}$ and a control action $\mathbf{U}_{i,h}$, based on her communication strategy $g_{i,h}^m$ and control strategy $g_{i,h}^a$, respectively. The class of *open-loop* communication strategies is defined as $\mathcal{G}_{i,h}^m := \{g_{i,h}^m : \{*\} \rightarrow \mathcal{M}_{i,h}\}$. We identify each such strategy with its single value $g_{i,h}^m(*)$, so $\mathbf{M}_{i,h}$ is pre-selected. In this case, we consider the control strategies $g_{i,h}^a \in \mathcal{G}_{i,h}^a := \{g_{i,h}^a : \mathcal{I}_{i,h^+} \rightarrow \mathcal{U}_i\}$. Also, we may extend the setting to the class of *closed-loop* communication strategies, $\mathcal{G}_{i,h}^m := \{g_{i,h}^m : \mathcal{I}_{i,h^-} \times \mathcal{M}_{1:h-1} \rightarrow \mathcal{M}_{i,h}\}$ where $\mathbf{M}_{i,h}$ is thus chosen based on $\mathbf{I}_{i,h^-}$ and $\mathbf{M}_{1:h-1}$. Following [14, 12, 13], in this closed-loop setting, the communication action $\mathbf{M}_h$ is also assumed to be (implicitly) shared, since it can be inferred from *what has been*

shared. Hence, in this case, we consider the control strategies $g_{i,h}^a \in \mathcal{G}_{i,h}^a := \{g_{i,h}^a : \mathcal{I}_{i,h^+} \times \mathcal{M}_{1:h} \rightarrow \mathcal{U}_i\}$ for generality. We will focus on the open-loop communication strategies throughout and extend our approach to the closed-loop ones in §V. We use $g_h^m := (g_{1,h}^m, \dots, g_{n,h}^m)$ and $g_h^a := (g_{1,h}^a, \dots, g_{n,h}^a)$ to denote the joint communication and control strategies, respectively. We denote by $\mathcal{G}_h^a, \mathcal{G}_h^m$ the spaces of these joint strategies.

The total cost (objective) of all the agents is defined as

$$J_{\mathcal{D}}(g_{1:H}^m, g_{1:H}^a) := \mathbb{E} \left[\sum_{h=1}^H (c_h + \kappa_h) + c_{H+1} \mid g_{1:H}^m, g_{1:H}^a \right],$$

which is the expected accumulated sum of the control and communication costs over H timesteps. With this objective, we define the solution concept of *team optimality* below.

Definition II.2. Given a JCCO problem $\mathcal{D}$, a strategy $(g_{1:H}^{m,*}, g_{1:H}^{a,*})$ is *team-optimal* if $\forall g_{1:H}^m \in \mathcal{G}_{1:H}^m, g_{1:H}^a \in \mathcal{G}_{1:H}^a$,

$$J_{\mathcal{D}}(g_{1:H}^m, g_{1:H}^a) \geq J_{\mathcal{D}}(g_{1:H}^{m,*}, g_{1:H}^{a,*}),$$

with $g_{1:H}^{m,*}, g_{1:H}^{a,*}$ being referred to as the *optimal communication* and *control* strategies, respectively.

Note that the team-optimal strategies depend on the (communication) strategy space being used (i.e., open-loop versus closed-loop), which will be clear later from the context.

II-C Information structures

Information structure is a well-studied notion in decentralized stochastic control that captures *who knows what and when* [21, 22]. For any JCCO problem $\mathcal{D}$, as the additional sharing via communication will also affect the IS and is not determined beforehand, when discussing the IS of $\mathcal{D}$, we mean that of the problem *without additional sharing*, which is essentially a decentralized LQG control problem, as formally defined below.

Definition II.3 (Decentralized LQG Induced by JCCO). For a JCCO problem $\mathcal{D}$, we call $\check{\mathcal{D}}$ the *decentralized LQG control problem induced by $\mathcal{D}$* . The information structure (IS) of $\mathcal{D}$ is defined to be that of $\check{\mathcal{D}}$.

Definition II.4 ((Strictly) Partially Nested JCCO). A JCCO problem $\mathcal{D}$ is called *partially nested* (PN) if its induced decentralized LQG control problem $\check{\mathcal{D}}$ is PN [15]. Namely, for any $i_1, i_2 \in [n]$ and $h_1, h_2 \in [H]$ with $h_1 < h_2$, if $\mathbf{U}_{i_1, h_1}$ influences $\mathbf{I}_{i_2, h_2^-}$ in $\check{\mathcal{D}}$, then $\mathbf{I}_{i_1, h_1^-} \subseteq \mathbf{I}_{i_2, h_2^-}$. It is called *strictly PN* if, under the same condition, $\mathbf{I}_{i_1, h_1^-} \cup \{\mathbf{U}_{i_1, h_1}\} \subseteq \mathbf{I}_{i_2, h_2^-}$.

III. Linearity and Structural Conditions for JCCO with Open-Loop Communication Strategies

We now focus on the setting with *open-loop* communication strategies. Given any such fixed $g_{1:H}^m$ (and thus pre-selected $M_{1:H}$), the problem becomes a decentralized LQG control one. However, it is known that the optimal control strategy of a decentralized LQG control problem may be *non-linear* in general [23], making it computationally challenging. Fortunately, with partially nested

ISs, the optimal control strategies can be linear [15]¹. Hence, it is natural to ask: when it comes to JCCO (with open-loop $g_{1:H}^m$), when would the optimal control strategy $g_{i,h}^a$ still be *linear*, i.e., a linear function of the available information $\mathbf{I}_{i,h^+}$, for all $i \in [n], h \in [H]$?

To answer this question, first, if the IS of JCCO (i.e., the IS of baseline sharing) is not even PN, then in general it is hard for the IS of the decentralized LQG to be PN *after* the optimal additional sharing, especially with (high) communication costs. This may thus further break the linearity of the optimal control strategies. To formalize this intuition, we introduce the following result, whose proof, together with other omitted proofs for this section, can be found in §B.

Lemma III.1. There exists a JCCO problem $\mathcal{D}$ with open-loop communication strategies that is not PN but satisfies Assumptions II.1, III.2, and III.4, such that the optimal control strategy is *nonlinear*.

Due to Lemma III.1, we focus on PN JCCO problems, in order to obtain linearity of the optimal control strategies and thus computational tractability. However, even when the baseline sharing is PN, the additional sharing may *break* the PN IS, and thus cause computational hardness again. Specifically, suppose agent i_1 's control action $\mathbf{U}_{i_1,h_1}$ does not influence another agent i_2 's information $\mathbf{I}_{i_2,h_2}$ in baseline sharing, but her control action $\mathbf{U}_{i_1,h_1}$ is shared with others via additional sharing. Then, $\mathbf{U}_{i_1,h_1}$ starts to influence $\mathbf{I}_{i_2,h_2}$ and the PN IS may thus break. Here, the *non-influence* of $\mathbf{U}_{i_1,h_1}$ in baseline sharing may come from two cases: 1) $\mathbf{U}_{i_1,h_1}$ does not influence the underlying state $\mathbf{X}_{h_1+1}$; or 2) $\mathbf{U}_{i_1,h_1}$ does influence the underlying state $\mathbf{X}_{h_1+1}$, but the effect is *not observed* by other agents. To avoid the impact of such non-influential actions, we make the following two assumptions, followed by justifications of their necessity regarding the *existence/linearity* of the optimal control strategies.

For case 1) above, we make the following assumption that such useless actions will *not* be used or shared later. In fact, for the linear systems considered in (II.1), a useless control action will not even appear as a random variable in the system evolution, and it is natural to disregard it in the available information of later agents. Moreover, we highlight that such an assumption has been implicitly made in the common examples in decentralized stochastic control [16, 19].

Assumption III.2. For any $i \in [n], h \in [H]$, if $B_{i,h} = \mathbf{0}$, then $\mathbf{U}_{i,h} \notin \mathbf{I}_{j,(h')^-}, \mathbf{U}_{i,h} \notin \mathbf{I}_{j,(h')^+}, \forall h' \in [H]$ with $h' > h, j \in [n]$.

The significance of Assumption III.2 in JCCO is reflected in two respects. First, it helps to preserve the PN IS (as to be shown formally in Theorem III.6). Second, without it, the optimal control strategy may either not exist or be nonlinear, as shown in the following lemma.

Lemma III.3. There exist PN JCCO problems $\mathcal{D}_1, \mathcal{D}_2$ with open-loop communication strategies that satisfy Assumptions II.1, III.4 but not Assumption III.2, such that the optimal control strategy of $\mathcal{D}_1$ is *nonlinear*, and the team-optimal strategy of $\mathcal{D}_2$ *does not exist*.

For case 2) above, the PN IS may break due to the *non-informativeness* of agents' observations. We thus make the following assumption.

Assumption III.4. For any $h \in [H-1], i \in [n]$, if $B_{i,h} \neq \mathbf{0}$, then $E_{-i,h+1}B_{i,h} \neq \mathbf{0}$, where $E_{-i,h+1} := \begin{bmatrix} E_{1,h+1}^\top & \cdots & E_{i-1,h+1}^\top & E_{i+1,h+1}^\top & \cdots & E_{n,h+1}^\top \end{bmatrix}^\top$.

¹For completeness, we provide Corollary A.2 to establish the extension needed here for possibly singular Gaussian covariances and positive-semidefinite state and control cost matrices $\{Q_h^1\}_{h \in [H+1]}$ and $\{Q_h^2\}_{h \in [H]}$.

Assumption III.4 means that, for any $i \in [n]$ and $h \in [H-1]$, if $B_{i,h} \neq \mathbf{0}$, then $E_{j,h+1}B_{i,h} \neq \mathbf{0}$ for at least one $j \neq i$. Thus, it does not require $E_{-i,h+1}$ to have full rank or the other agents to jointly observe the whole state $\mathbf{X}_{h+1}$. For instance, if $B_{i,h} = \mathbf{0}$, then no condition is imposed on $E_{-i,h+1}$. As shown below, Assumption III.4 is necessary to ensure the linearity of the optimal control strategy of $\mathcal{D}$.

Lemma III.5. There exists a PN JCCO problem $\mathcal{D}$ with open-loop communication strategies that satisfies Assumptions II.1, III.2 but not Assumption III.4, such that the optimal control strategy of $\mathcal{D}$ is *nonlinear*.

We have thus justified the importance of the above assumptions and the partial nestedness of baseline sharing for JCCO problems, by showing that missing any one of them (while keeping others) may cause the nonlinearity (or even non-existence) of the optimal control strategy.

On the other hand, for any fixed open-loop communication strategy $g_{1:H}^m \in \mathcal{G}_{1:H}^m$, we denote by $\mathcal{D}(g_{1:H}^m)$ the subproblem of finding the optimal *control* strategy with respect to $g_{1:H}^m$, where we add $-$ to the notation in $\mathcal{D}(g_{1:H}^m)$ to distinguish it from the notation in the original JCCO problem. We remove $+$ in the timestep index for all the notation in $\mathcal{D}(g_{1:H}^m)$ for convenience, as the IS is now *fixed* and only the control problem will be analyzed. Then, we can show that if $\mathcal{D}$ satisfies all the assumptions above, the optimal control strategy of $\mathcal{D}$ is linear. Indeed, under any fixed open-loop $g_{1:H}^m \in \mathcal{G}_{1:H}^m$, the subproblem $\mathcal{D}(g_{1:H}^m)$ can be shown to be a *decentralized LQG* problem with a partially nested information structure (see §A for a formal definition).

Theorem III.6. Consider a PN JCCO problem $\mathcal{D}$ that satisfies Assumptions II.1, III.2, and III.4. For any open-loop communication strategy $g_{1:H}^m \in \mathcal{G}_{1:H}^m$, the subproblem $\mathcal{D}(g_{1:H}^m)$ is a decentralized LQG control problem with a PN IS, and its information evolution rules satisfy: for each $i \in [n]$ and $h \in [H]$,

$$\begin{aligned}\bar{\mathbf{C}}_h &= \bar{\mathbf{C}}_{h-1} \cup \bar{\mathbf{Z}}_h, & \bar{\mathbf{Z}}_h &= \bar{\chi}_h(\bar{\mathbf{P}}_{h-1}, \bar{\mathbf{U}}_{h-1}, \bar{\mathbf{Y}}_h), & \bar{\mathbf{Y}}_{i,h} &\in \bar{\mathbf{I}}_{i,h}, \\ \bar{\mathbf{P}}_{i,h} &= \bar{\zeta}_{i,h}(\bar{\mathbf{P}}_{i,h-1}, \bar{\mathbf{U}}_{i,h-1}, \bar{\mathbf{Y}}_{i,h}), & \bar{\mathbf{I}}_{i,h-1} &\subseteq \bar{\mathbf{I}}_{i,h},\end{aligned}$$

for some projection functions $\bar{\chi}_h$ and $\{\bar{\zeta}_{i,h}\}_{i \in [n]}$. Furthermore, there exists an optimal joint strategy $(g_{1:H}^{m,*}, g_{1:H}^{a,*})$ of $\mathcal{D}$ with open-loop communication strategies such that $g_{1:H}^{a,*}$ is a collection of *linear* functions.

Theorem III.6 provides a way to solve $\mathcal{D}$ by finding the optimal control strategy $g_{1:H}^{a,*}$ with respect to any fixed open-loop communication strategy $g_{1:H}^m$, i.e., solving $\mathcal{D}(g_{1:H}^m)$, and then optimizing over the open-loop communication strategies $g_{1:H}^m$. We will thus investigate how to solve the problem $\mathcal{D}(g_{1:H}^m)$ next.

IV. Dynamic Programming for Decentralized LQG with Partially Nested Information Structures

Since $\mathcal{D}(g_{1:H}^m)$ is PN under the assumptions in Theorem III.6, it now suffices to develop an algorithm that can solve PN decentralized LQG control problems. For notational convenience, in this section, we will omit the superscript a for the control strategies $\bar{g}_{1:H}^a$ for $\mathcal{D}(g_{1:H}^m)$.

The algorithm to be introduced can be viewed as a dynamic-programming approach for solving a class of PN decentralized LQG problems with *output* feedback under the common-information-based framework [10, 24], which thus advances the results in [17] and may be of independent interest. Detailed proofs of results in this section are deferred to §C. Specifically, under PN IS, the known dynamic-programming-based approaches and sufficient statistics in [16, 18] do not apply here, as they focused on the special settings with *factorized* states, and with either *state-feedback* [16] or *multi-tree* coupling and communication graphs [18].

On the other hand, under the common-information-based framework [10], [17] showed that if agents are *restricted* to using *linear* control strategies, then the decentralized control problem can be reformulated as a *centralized* LQG one when *fixing* the private-information component of the strategies. However, it is not clear how to further *optimize* the private-information component, which induces a non-convex optimization problem in general [17]. Furthermore, [11] showed that, if additionally the CIB beliefs from [10] are *strategy independent* [11, Section 2.4], then the CIB beliefs can be further simplified to a finite-dimensional conditional mean of the state and private information in this linear-quadratic setting².

Inspired by [11], and to address the non-convexity issue in [17], we exploit the PN IS to obtain a finite-dimensional, conditional-mean-based sufficient statistic, by establishing a new connection between the PN IS and the SI-CIB condition.

To this end, we adapt our technique in [14] to this linear-quadratic setting, while addressing several fundamentally different challenges. First, we expand the PN IS into a *strictly* PN one, where the *actions* of agents who affect some other agent will also be known by the affected agent (in addition to their *information*). More formally, we construct a new problem $\widetilde{\mathcal{D}}(g_{1:H}^m)$, whose elements are almost identical to those in $\mathcal{D}(g_{1:H}^m)$, except that the agents' information is now expanded as follows: for each $h \in [H]$

$$\widetilde{\mathcal{C}}_h = \bar{\mathcal{C}}_h \cup \{\bar{\mathbf{U}}_{i,t} | i \in [n], t < h, \bar{\mathbf{I}}_{i,t} \subseteq \bar{\mathcal{C}}_h, \bar{\mathbf{B}}_{i,t} \neq \mathbf{0}\}, \quad \forall j \in [n], \widetilde{\mathcal{P}}_{j,h} = \bar{\mathcal{P}}_{j,h} \setminus \{\bar{\mathbf{U}}_{j,t} | t < h, \bar{\mathbf{I}}_{j,t} \subseteq \bar{\mathcal{C}}_h\}, \quad (\text{IV.1})$$

where we add $\sim$ to the notation in $\widetilde{\mathcal{D}}(g_{1:H}^m)$, and the incremental common information is defined by $\widetilde{\mathcal{Z}}_h := \widetilde{\mathcal{C}}_h \setminus \widetilde{\mathcal{C}}_{h-1}, \forall h \in [H]$. Through such an expansion, we can leverage the new problem $\widetilde{\mathcal{D}}(g_{1:H}^m)$ to solve the original $\mathcal{D}(g_{1:H}^m)$, as formalized in the following lemma.

Lemma IV.1. Let $\mathcal{D}$ be PN and satisfy Assumptions II.1, III.2, and III.4, and for any fixed open-loop $g_{1:H}^m \in \mathcal{G}_{1:H}^m$, let $\mathcal{D}(g_{1:H}^m)$ and $\widetilde{\mathcal{D}}(g_{1:H}^m)$ be the decentralized LQG control problems constructed above. Then, there exists a function φ such that for any linear optimal control strategy $\bar{g}_{1:H}^*$ of the problem $\widetilde{\mathcal{D}}(g_{1:H}^m)$, $\bar{g}_{1:H}^* = \varphi(\bar{g}_{1:H}^*, \mathcal{D}(g_{1:H}^m))$ is a linear optimal control strategy of $\mathcal{D}(g_{1:H}^m)$, and $J_{\mathcal{D}(g_{1:H}^m)}(\bar{g}_{1:H}^*) = J_{\widetilde{\mathcal{D}}(g_{1:H}^m)}(\bar{g}_{1:H}^*)$, where $J_{\mathcal{D}(g_{1:H}^m)}(\bar{g}_{1:H}^*)$ is the expected accumulated cost of $\mathcal{D}(g_{1:H}^m)$ under the strategy $\bar{g}_{1:H}^*$, and $J_{\widetilde{\mathcal{D}}(g_{1:H}^m)}(\bar{g}_{1:H}^*)$ is defined similarly (cf. §A for the formal definitions).

Before proceeding further, we introduce some additional notation in $\widetilde{\mathcal{D}}(g_{1:H}^m)$. For any $h \in [H]$, we define $\widetilde{\mathcal{S}}_h := [\widetilde{\mathbf{X}}_h^\top \quad \widetilde{\mathbf{P}}_{1,h}^\top \quad \cdots \quad \widetilde{\mathbf{P}}_{n,h}^\top]^\top$ and use $\widetilde{\mathcal{S}}_h := \widetilde{\mathcal{X}} \times \widetilde{\mathcal{P}}_{1,h} \times \cdots \times \widetilde{\mathcal{P}}_{n,h}$ to denote the space of $\widetilde{\mathcal{S}}_h$. Then, we use $\widetilde{\mathbf{B}}_h \in \mathcal{P}(\widetilde{\mathcal{S}}_h)$ with $\widetilde{\mathbf{B}}_h(d\widetilde{\mathcal{S}}_h) := \mathbb{P}^{\widetilde{\mathcal{D}}(g_{1:H}^m)}(d\widetilde{\mathcal{S}}_h | \widetilde{\mathcal{C}}_h, \widetilde{g}_{1:h-1})$ to denote the conditional probability measure of $\widetilde{\mathcal{S}}_h$ given the common information $\widetilde{\mathcal{C}}_h$ and the past strategies $\widetilde{g}_{1:h-1}$. Hence, we have $\widetilde{\mathbf{B}}_h = \Pi_h(\widetilde{\mathcal{C}}_h, \widetilde{g}_{1:h-1})$ for some functional Π_h . The benefit of solving $\widetilde{\mathcal{D}}(g_{1:H}^m)$ instead of

²Note that [11] focused on a *game* setting with strategic agents, instead of the *team* setting in [17, 16] and the present paper.

$\mathcal{D}(g_{1:H}^m)$ is that the former satisfies the strategy-independence condition for CIB beliefs under the stated assumptions, as shown below.

Theorem IV.2. If $\mathcal{D}$ is PN and satisfies Assumptions II.1, III.2, and III.4, then for any fixed open-loop $g_{1:H}^m \in \mathcal{G}_{1:H}^m$, $\widetilde{\mathcal{D}}(g_{1:H}^m)$ constructed as above is (strictly) PN and satisfies the *SI-CIB condition*: for any $h \in [H]$, any realization $\widetilde{\mathcal{C}}_h \in \widetilde{\mathcal{C}}_h$, and any two control strategies $\widetilde{g}_{1:h-1}, \widetilde{g}'_{1:h-1}$ that can reach $\widetilde{\mathcal{C}}_h$, it holds that $\Pi_h(\widetilde{\mathcal{C}}_h, \widetilde{g}_{1:h-1}) = \Pi_h(\widetilde{\mathcal{C}}_h, \widetilde{g}'_{1:h-1})$. Meanwhile, $\widetilde{\mathcal{D}}(g_{1:H}^m)$ has the following information evolution rules: for each $i \in [n]$ and $h \in [H]$,

$$\begin{aligned}\widetilde{\mathcal{C}}_h &= \widetilde{\mathcal{C}}_{h-1} \cup \widetilde{\mathcal{Z}}_h, \widetilde{\mathcal{Z}}_h = \widetilde{\chi}_h(\widetilde{\mathcal{P}}_{h-1}, \widetilde{\mathcal{U}}_{h-1}, \widetilde{\mathcal{Y}}_h), \\ \widetilde{\mathcal{P}}_{i,h} &= \widetilde{\zeta}_{i,h}(\widetilde{\mathcal{P}}_{i,h-1}, \widetilde{\mathcal{U}}_{i,h-1}, \widetilde{\mathcal{Y}}_{i,h}), \widetilde{\mathcal{Y}}_{i,h} \in \widetilde{\mathcal{I}}_{i,h}, \widetilde{\mathcal{I}}_{i,h-1} \subseteq \widetilde{\mathcal{I}}_{i,h},\end{aligned}\tag{IV.2}$$

for some projection functions $\widetilde{\chi}_h$ and $\{\widetilde{\zeta}_{i,h}\}_{i \in [n]}$.

Furthermore, we define the *prescription* [10] of each agent i at timestep h as $\widetilde{\gamma}_{i,h} : \widetilde{\mathcal{P}}_{i,h} \rightarrow \widetilde{\mathcal{U}}_i$. We denote by $\widetilde{\gamma}_h = (\widetilde{\gamma}_{1,h}, \dots, \widetilde{\gamma}_{n,h})$ the joint prescription of all the agents, and by $\widetilde{\Gamma}_{i,h}$ and $\widetilde{\Gamma}_h$ the spaces of agent i 's and the joint prescriptions at timestep h , respectively. For notational convenience, we write $\widetilde{\gamma}_h(\widetilde{\mathcal{P}}_h) = [\widetilde{\gamma}_{1,h}(\widetilde{\mathcal{P}}_{1,h})^\top \dots \widetilde{\gamma}_{n,h}(\widetilde{\mathcal{P}}_{n,h})^\top]^\top$.

Following [11], under the SI-CIB condition, the belief $\widetilde{\mathcal{B}}_h$ in $\widetilde{\mathcal{D}}(g_{1:H}^m)$ is a Gaussian distribution with mean $\widetilde{\Theta}_h = [\mathbb{E}[\widetilde{\mathcal{X}}_h | \widetilde{\mathcal{C}}_h]^\top \mathbb{E}[\widetilde{\mathcal{P}}_{1,h} | \widetilde{\mathcal{C}}_h]^\top \dots \mathbb{E}[\widetilde{\mathcal{P}}_{n,h} | \widetilde{\mathcal{C}}_h]^\top]^\top$ and covariance $\widetilde{\Sigma}_h$ that evolve over time, as formalized below.

Lemma IV.3 (Adapted from Lemmas 2.2 & 2.3 in [11]). Consider a decentralized LQG problem $\widetilde{\mathcal{D}}(g_{1:H}^m)$ constructed by the strict expansion in Theorem IV.2, that has information evolution rules in Equation (IV.2) and satisfies the SI-CIB condition. Then the CIB belief $\widetilde{\mathcal{B}}_h$ admits a Gaussian distribution $\mathcal{N}(\widetilde{\Theta}_h, \widetilde{\Sigma}_h)$, and the conditional mean $\widetilde{\Theta}_h$ and the conditional covariance $\widetilde{\Sigma}_h$ of $\widetilde{\mathcal{B}}_h$ evolve as follows: for any $h \in [2 : H]$,

$$\widetilde{\Theta}_h = \widetilde{\Psi}_h^1(\widetilde{\Theta}_{h-1}, \widetilde{\mathcal{Z}}_h), \quad \widetilde{\Sigma}_h = \widetilde{\Psi}_h^2(\widetilde{\Sigma}_{h-1}),$$

where $\widetilde{\Psi}_h^1$ is a fixed linear function, $\widetilde{\Psi}_h^2$ is a fixed deterministic function, and neither depends on the strategies $\widetilde{g}_{1:h-1}$. In particular, $\widetilde{\Sigma}_h$ is deterministic and strategy-independent.

According to Lemma IV.3, we know that $\widetilde{\Theta}_h$ depends only on $\{\widetilde{\mathcal{Z}}_t\}_{t=1}^h$. Hence, there exists a fixed linear transformation $\widetilde{\Psi}_h^3$ such that $\widetilde{\Theta}_h = \widetilde{\Psi}_h^3(\widetilde{\mathcal{C}}_h)$. Then, following [11], we can construct a new problem $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$: At timestep $h \in [H]$, the state of $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is defined as $\widetilde{\Theta}_h \in \widetilde{\mathcal{S}}_h$ and the action is defined as $\widetilde{\gamma}_h \in \widetilde{\Gamma}_h$. The stage cost is defined as $c_h^\dagger(\widetilde{\Theta}_h, \widetilde{\gamma}_h) := \mathbb{E}[\widetilde{\mathcal{X}}_h^\top \widetilde{\mathcal{Q}}_h^{-1} \widetilde{\mathcal{X}}_h + \widetilde{\gamma}_h(\widetilde{\mathcal{P}}_h)^\top \widetilde{\mathcal{Q}}_h^{-2} \widetilde{\gamma}_h(\widetilde{\mathcal{P}}_h) + \mathbb{1}[h = H] \widetilde{\mathcal{X}}_{H+1}^\top \widetilde{\mathcal{Q}}_{H+1}^{-1} \widetilde{\mathcal{X}}_{H+1}]$, where the expectation is taken over $[\widetilde{\mathcal{X}}_h^\top \widetilde{\mathcal{P}}_h^\top]^\top \sim \mathcal{N}(\widetilde{\Theta}_h, \widetilde{\Sigma}_h)$, and when $h = H$, also over the final state $\widetilde{\mathcal{X}}_{H+1}$ generated by the system dynamics. The admissible strategy at timestep h in $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is denoted by $g_h^\dagger : \widetilde{\mathcal{C}}_h \rightarrow \widetilde{\Gamma}_h$, and $\mathcal{G}_h^\dagger$ is the set of all admissible strategies. The objective of $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is then defined by

$$J_{\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)}(g_{1:H}^\dagger) := \mathbb{E} \left[\sum_{h=1}^H c_h^\dagger(\widetilde{\Theta}_h, \widetilde{\gamma}_h) \right].$$

As shown in the following theorem, the constructed problem $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is a Markov decision problem, and the optimal strategy of $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ can be used to compute the team-optimal strategy of $\widetilde{\mathcal{D}}(g_{1:H}^m)$, and vice versa.

Theorem IV.4. Let $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ be the problem constructed from $\widetilde{\mathcal{D}}(g_{1:H}^m)$ satisfying the SI-CIB condition. Then, $\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is a Markov decision problem with horizon H , state $\{\widetilde{\Theta}_h\}_{h \in [H]}$, control action $\{\widetilde{\gamma}_h\}_{h \in [H]}$, and cost function $\{c_h^\dagger\}_{h \in [H]}$. Moreover, there exist bijections $\varsigma_h : \widetilde{\mathcal{G}}_h \rightarrow \mathcal{G}_h^\dagger$ for all $h \in [H]$ that can be defined as follows: for any strategy $\widetilde{g}_h \in \widetilde{\mathcal{G}}_h$ and realization of common information $\widetilde{C}_h \in \widetilde{\mathcal{C}}_h$, $\varsigma_h(\widetilde{g}_h)(\widetilde{C}_h)(\cdot) = \widetilde{g}_h(\widetilde{C}_h, \cdot)$. These bijections also satisfy the following equality for every strategy $\widetilde{g}_{1:H} \in \widetilde{\mathcal{G}}_{1:H}$: $J_{\widetilde{\mathcal{D}}(g_{1:H}^m)}(\widetilde{g}_{1:H}) = J_{\widetilde{\mathcal{D}}^\dagger(g_{1:H}^m)}(g_{1:H}^\dagger)$, where $g_h^\dagger = \varsigma_h(\widetilde{g}_h), \forall h \in [H]$.

By Theorem IV.4 and Lemma IV.5 below, an optimal strategy exists and may be chosen *Markovian*, with $g_h^{\dagger,*}$ depending on the common record only through $\widetilde{\Theta}_h$. In particular, for any $h \in [H]$, we can write $g_h^{\dagger,*}$ as $g_h^{\dagger,*}(\widetilde{C}_h) = \mathcal{G}_h^*(\widetilde{\Psi}_h^3(\widetilde{C}_h))$ for some function $\mathcal{G}_h^*$. Furthermore, if we define the strategy $\widetilde{g}_h^* = \varsigma_h^{-1}(g_h^{\dagger,*}), \forall h \in [H]$, then $\widetilde{g}_{1:H}^*$ is an optimal strategy of $\widetilde{\mathcal{D}}(g_{1:H}^m)$. Therefore, there is no loss of optimality in restricting attention to $\widetilde{g}_{i,h}$ that chooses action $\widetilde{U}_{i,h}$ depending only on $\widetilde{\Theta}_h$ and $\widetilde{P}_{i,h}$ for any $h \in [H], i \in [n]$ in solving the problem $\widetilde{\mathcal{D}}(g_{1:H}^m)$. We categorize this type of strategy as *common-information-based Markovian (CIB-Markovian) strategy* [24, 11]. Moreover, we can define the associated value functions for every $h \in [H]$ as

$$\widetilde{V}_h(\widetilde{\Theta}_h) := \min_{\widetilde{g}_{h:H}} \mathbb{E} \left[\sum_{t=h}^H (\widetilde{\mathbf{X}}_t^\top \widetilde{\mathbf{Q}}_t^\dagger \widetilde{\mathbf{X}}_t + \widetilde{\mathbf{U}}_t^\top \widetilde{\mathbf{Q}}_t^\dagger \widetilde{\mathbf{U}}_t) + \widetilde{\mathbf{X}}_{H+1}^\top \widetilde{\mathbf{Q}}_{H+1}^\dagger \widetilde{\mathbf{X}}_{H+1} \mid \widetilde{\Theta}_h, \widetilde{g}_{h:H} \right],$$

which satisfy the following Bellman Equations:

$$\widetilde{V}_h(\widetilde{\Theta}_h) = \begin{cases} \min_{\widetilde{\gamma}_h \in \widetilde{\Gamma}_h} \left\{ c_h^\dagger(\widetilde{\Theta}_h, \widetilde{\gamma}_h) + \mathbb{E}[\widetilde{V}_{h+1}(\widetilde{\Theta}_{h+1}) \mid \widetilde{\Theta}_h, \widetilde{\gamma}_h] \right\}, & h \in [H-1], \\ \min_{\widetilde{\gamma}_H \in \widetilde{\Gamma}_H} c_H^\dagger(\widetilde{\Theta}_H, \widetilde{\gamma}_H), & h = H. \end{cases} \quad (\text{IV.3})$$

Then, we can derive an approach to obtain the optimal strategy $\widetilde{g}_{1:H}^*$ by dynamic programming: for all $h \in [H]$ and $\widetilde{C}_h \in \widetilde{\mathcal{C}}_h$, $\widetilde{g}_h^*(\widetilde{C}_h, \cdot)$ minimizes the right-hand side of Equation (IV.3) at $\widetilde{\Theta}_h = \widetilde{\Psi}_h^3(\widetilde{C}_h)$. However, there are uncountably many possible $\widetilde{C}_h$ (or $\widetilde{\Theta}_h$), making this approach computationally intractable without a closed-form solution.

For computational tractability, we leverage the *linearity* property of the optimal strategy. Although it is known that there exists a *linear* optimal strategy under PN IS (cf. Corollary A.2), and if it exists, the optimal strategy can be *CIB-Markovian* under the SI-CIB condition (cf. Theorem IV.4), the existence of an optimal strategy with both properties remains unclear. We show the existence of such an optimal strategy below.

Lemma IV.5. Suppose $\widetilde{\mathcal{D}}(g_{1:H}^m)$ is a decentralized LQG control problem constructed from a PN JCCO problem $\mathcal{D}$ that satisfies Assumptions II.1, III.2, and III.4. Then, there exists an optimal linear strategy $\widetilde{g}_{1:H}^*$, and matrices $\{\widetilde{\mathcal{E}}_{i,h}^*\}_{i \in [n], h \in [H]}, \{\widetilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]}$ such that the $\widetilde{\gamma}_{i,h}^*(\cdot)$, defined as $\widetilde{\gamma}_{i,h}^*(\cdot) := \widetilde{\mathcal{E}}_{i,h}^* \widetilde{\Theta}_h + \widetilde{\mathcal{F}}_{i,h}^* \cdot$, minimizes the right-hand side of Equation (IV.3) at every CIB mean $\widetilde{\Theta}_h$ reachable under a preceding strategy, and $\widetilde{g}_{i,h}^*(\widetilde{C}_h, \widetilde{P}_{i,h}) = \widetilde{\mathcal{E}}_{i,h}^* \widetilde{\Psi}_h^3(\widetilde{C}_h) + \widetilde{\mathcal{F}}_{i,h}^* \widetilde{P}_{i,h}$ for any $i \in [n], h \in [H], \widetilde{C}_h \in \widetilde{\mathcal{C}}_h, \widetilde{P}_{i,h} \in \widetilde{\mathcal{P}}_{i,h}$.

Lemma IV.5 shows that for each agent i at timestep h , it suffices to compute the optimal strategy that is linear in $\tilde{\Theta}_h$ and $\tilde{\mathbf{P}}_{i,h}$. Then, for any $i \in [n], h \in [H]$, we can parameterize the strategy $\tilde{g}_{i,h}$ by two matrices $\tilde{\mathcal{E}}_{i,h}, \tilde{\mathcal{F}}_{i,h}$, where $\tilde{g}_{i,h}(\tilde{\mathbf{C}}_h, \tilde{\mathbf{P}}_{i,h}) = \tilde{\mathcal{E}}_{i,h} \tilde{\Psi}_h^3(\tilde{\mathbf{C}}_h) + \tilde{\mathcal{F}}_{i,h} \tilde{\mathbf{P}}_{i,h}$. Define $\tilde{\mathcal{E}}_h^* := [(\tilde{\mathcal{E}}_{1,h}^*)^\top \cdots (\tilde{\mathcal{E}}_{n,h}^*)^\top]^\top$ and $\tilde{\mathcal{F}}_h^* := \text{diag}(\tilde{\mathcal{F}}_{1,h}^*, \dots, \tilde{\mathcal{F}}_{n,h}^*)$. With these definitions, we can find the optimal control strategy $\tilde{g}_h^*$ by solving for the optimal matrices:

$$(\tilde{\mathcal{E}}_h^*, \tilde{\mathcal{F}}_h^*) \in \underset{\tilde{\mathcal{E}}_h, \tilde{\mathcal{F}}_h}{\text{argmin}} \begin{cases} c_h^\dagger(\tilde{\Theta}_h, \tilde{\gamma}_h) + \mathbb{E}[\tilde{V}_{h+1}(\tilde{\Theta}_{h+1}) | \tilde{\Theta}_h = \tilde{\Theta}_h, \tilde{\mathcal{E}}_h, \tilde{\mathcal{F}}_h], & h \in [H-1], \\ c_H^\dagger(\tilde{\Theta}_H, \tilde{\gamma}_H), & h = H. \end{cases} \quad (\text{IV.4})$$

As shown in the following theorem, such optimal matrices can be computed through a set of Riccati Equations with better computational tractability.

Theorem IV.6. Suppose $\tilde{\mathcal{D}}(g_{1:H}^m)$ is a decentralized LQG control problem constructed from a PN JCCO problem $\mathcal{D}$ that satisfies Assumptions II.1, III.2, and III.4. For $h \in [H-1]$, the conditional mean evolves as $\tilde{\Theta}_{h+1} = K_{h+1}^1 \tilde{\Theta}_h + K_{h+1}^2 \tilde{\mathbf{P}}_h + K_{h+1}^3 \tilde{\mathbf{U}}_h + K_{h+1}^4 \tilde{\mathbf{Y}}_{h+1}$, where these matrices are computed by a centralized Kalman filter. For $h \in [H-1]$, define

$$\begin{aligned} \tilde{L}_h^1 &= \tilde{L}_h^3 := K_{h+1}^3 + K_{h+1}^4 \tilde{E}_{h+1} \tilde{B}_h, \\ \tilde{L}_h^2 &:= K_{h+1}^1 + K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^4 \tilde{E}_{h+1} \tilde{A}_h \mathbb{I}_{x,h}, \\ \tilde{L}_h^4 &:= K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^4 \tilde{E}_{h+1} \tilde{A}_h \mathbb{I}_{x,h}, \end{aligned}$$

and set $\tilde{L}_H^1 = \tilde{L}_H^3 := \tilde{B}_H$ and $\tilde{L}_H^2 = \tilde{L}_H^4 := \tilde{A}_H \mathbb{I}_{x,H}$. Then, the value function $\tilde{V}_h(\tilde{\Theta}_h)$ has the quadratic form $\tilde{V}_h(\tilde{\Theta}_h) = \tilde{\Theta}_h^\top \tilde{R}_h \tilde{\Theta}_h + \tilde{c}_h$ for some matrix $\tilde{R}_h$ and constant $\tilde{c}_h$ that do not depend on $\tilde{\Theta}_h$, which can be computed through the following Riccati Equations:

$$\begin{aligned} \tilde{R}_h &= (\tilde{\mathcal{E}}_h^*)^\top \tilde{Q}_h^2 \tilde{\mathcal{E}}_h^* + (\tilde{L}_h^1 \tilde{\mathcal{E}}_h^* + \tilde{L}_h^2)^\top \tilde{R}_{h+1} (\tilde{L}_h^1 \tilde{\mathcal{E}}_h^* + \tilde{L}_h^2) + \mathbb{I}_{x,h}^\top \tilde{Q}_h^1 \mathbb{I}_{x,h}, & \tilde{R}_{H+1} &= \tilde{Q}_{H+1}^1, \\ \tilde{\mathcal{E}}_h^* &= -(\tilde{Q}_h^2 + (\tilde{L}_h^1)^\top \tilde{R}_{h+1} \tilde{L}_h^1)^\dagger (\tilde{L}_h^1)^\top \tilde{R}_{h+1} \tilde{L}_h^2, \end{aligned} \quad (\text{IV.5})$$

where $\mathbb{I}_{x,h}, \mathbb{I}_{p,h}$ are the projection matrices that project vectors in $\tilde{\mathcal{S}}_h = \tilde{\mathcal{X}} \times \tilde{\mathcal{P}}_h$ to the corresponding vectors in $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{P}}_h$, respectively. For $h \in [H]$, define

$$\Sigma_t^y := \text{diag}(\tilde{\Sigma}_{1,t}, \dots, \tilde{\Sigma}_{n,t}), \quad \tilde{\Upsilon}_h := \begin{cases} K_{h+1}^4 (\tilde{E}_{h+1} \Sigma_{0,h} \tilde{E}_{h+1}^\top + \Sigma_{h+1}^y) (K_{h+1}^4)^\top, & h \in [H-1], \\ \Sigma_{0,H}, & h = H. \end{cases}$$

Also, $\{\tilde{\mathcal{F}}_{i,h}^*\}_{i \in [n]}, \tilde{c}_h$ can be computed as follows:

$$\begin{aligned} \tilde{\mathcal{F}}_h^* &\in \underset{\tilde{\mathcal{F}}_h = \text{diag}(\tilde{\mathcal{F}}_{1,h}, \dots, \tilde{\mathcal{F}}_{n,h})}{\text{argmin}} \quad \text{Tr}(\tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\Sigma}_h \mathbb{I}_{p,h}^\top \tilde{\mathcal{F}}_h^\top \tilde{Q}_h^2) + \text{Tr}((\tilde{L}_h^3 \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \tilde{L}_h^4) \tilde{\Sigma}_h (\tilde{L}_h^3 \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \tilde{L}_h^4)^\top \tilde{R}_{h+1}), \\ \tilde{\mathcal{F}}_h^* &:= \text{diag}(\tilde{\mathcal{F}}_{1,h}^*, \dots, \tilde{\mathcal{F}}_{n,h}^*), \\ \tilde{c}_h &= \text{Tr}(\tilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} \tilde{\Sigma}_h \mathbb{I}_{p,h}^\top (\tilde{\mathcal{F}}_h^*)^\top \tilde{Q}_h^2) + \text{Tr}(\mathbb{I}_{x,h} \tilde{\Sigma}_h \mathbb{I}_{x,h}^\top \tilde{Q}_h^1) + \text{Tr}((\tilde{L}_h^3 \tilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} + \tilde{L}_h^4) \tilde{\Sigma}_h (\tilde{L}_h^3 \tilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} + \tilde{L}_h^4)^\top \tilde{R}_{h+1}) \\ &\quad + \text{Tr}(\tilde{\Upsilon}_h \tilde{R}_{h+1}) + \tilde{c}_{h+1}, \\ \tilde{c}_{H+1} &= 0. \end{aligned} \quad (\text{IV.6})$$

Finally, $\tilde{\mathcal{E}}_h^* := [(\tilde{\mathcal{E}}_{1,h}^*)^\top \ \cdots \ (\tilde{\mathcal{E}}_{n,h}^*)^\top]^\top$ is computed as

$$\tilde{\mathcal{E}}_h^* = \tilde{\mathcal{E}}_h^* - \text{diag}(\tilde{\mathcal{F}}_{1,h}^*, \tilde{\mathcal{F}}_{2,h}^*, \dots, \tilde{\mathcal{F}}_{n,h}^*) \mathbb{I}_{p,h}. \quad (\text{IV.7})$$

Solving JCCO with open-loop $g_{1:H}^m$: Now we are ready to solve the problem $\mathcal{D}(g_{1:H}^m)$. First, from Theorem IV.6, we can solve for $\tilde{g}_{1:H}^*$ as $\tilde{g}_{i,h}^*(\tilde{\mathcal{C}}_h, \tilde{\mathcal{P}}_{i,h}) = \tilde{\mathcal{E}}_{i,h}^* \tilde{\Psi}_h^3(\tilde{\mathcal{C}}_h) + \tilde{\mathcal{F}}_{i,h}^* \tilde{\mathcal{P}}_{i,h}$ for any $i \in [n], h \in [H]$, and obtain the optimal value of $\min_{\tilde{g}_{1:H}} J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}) = \mathbb{E}[\tilde{V}_1(\tilde{\Theta}_1)] := \tilde{c}_0$. Note that $\tilde{\Theta}_1 = \tilde{\Psi}_1^3(\tilde{\mathcal{C}}_1) = K_1^1 \tilde{\mathcal{C}}_1, \tilde{\mathcal{C}}_1 = \tilde{\mathcal{Z}}_1 = \tilde{\chi}_1(\tilde{\mathcal{Y}}_1) = K_1^2 \tilde{\mathcal{Y}}_1, \tilde{\mathcal{Y}}_1 = \tilde{E}_1 \tilde{\mathcal{X}}_1 + \tilde{\mathcal{W}}_{1:n,1}$, where $\tilde{\mathcal{X}}_1 \sim \mathcal{N}(\mathbf{0}, \Sigma_1)$ for some matrices K_1^1, K_1^2 . Therefore, we can write $\tilde{c}_0 = \text{Tr}(\tilde{E}_1^\top (K_1^2)^\top (K_1^1)^\top \tilde{R}_1 K_1^1 K_1^2 \tilde{E}_1 \Sigma_1) + \text{Tr}((K_1^2)^\top (K_1^1)^\top \tilde{R}_1 K_1^1 K_1^2 \text{diag}(\Sigma_{1,1}, \dots, \Sigma_{n,1})) + \tilde{c}_1$. Then, by Lemma IV.1, we can obtain an optimal control strategy of $\mathcal{D}(g_{1:H}^m)$ as $\tilde{g}_{1:H}^* = \varphi(\tilde{g}_{1:H}^*, \mathcal{D}(g_{1:H}^m))$ such that $J_{\mathcal{D}(g_{1:H}^m)}(\tilde{g}_{1:H}^*) = J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}^*)$. If we assign $g_{1:H}^* = \tilde{g}_{1:H}^*$, then $g_{1:H}^*$ is an optimal control strategy with respect to the communication strategy $g_{1:H}^m$ in $\mathcal{D}$. Let $M_{1:H}$ be the communication actions chosen by $g_{1:H}^m$ such that $M_h = g_h^m$. Then, $J_{\mathcal{D}(g_{1:H}^m)}(g_{1:H}^*) = J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}^*) + \sum_{h=1}^H \mathcal{K}_h(g_h^m)$. Therefore, we can finally solve the JCCO problem by solving $g_{1:H}^{m,*} \in \text{argmin}_{g_{1:H}^m} J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}^*) + \sum_{h=1}^H \mathcal{K}_h(g_h^m)$.

V. Extension to JCCO with Closed-loop Communication Strategies

In this section, we extend our study to the JCCO problem with *closed-loop* communication strategies, i.e., for any $h \in [H], i \in [n], g_{i,h}^m : \mathcal{I}_{i,h^-} \times \mathcal{M}_{1:h-1} \rightarrow \mathcal{M}_{i,h}$, with associated (closed-loop) control strategies $g_{i,h}^a : \mathcal{I}_{i,h^+} \times \mathcal{M}_{1:h} \rightarrow \mathcal{U}_i$. Since each agent may influence others through the additional sharing determined by the communication strategy $g_{i,h}^m$, while others may not access the input of $g_{i,h}^m$, PN IS may break. Therefore, the SI-CIB condition does not hold in general and the common-information-based beliefs need not be Gaussian distributions that can be characterized by the finite-dimensional conditional mean and covariance. For the recursion at $h = 1$, identify $(\mathbf{X}_1, \mathbf{P}_{0^+}) = (\mathbf{X}_1, \emptyset)$ with $\mathbf{X}_1$ and set $\mathbf{B}_{0^+} := \mathcal{N}(\mathbf{0}, \Sigma_1)$, $\Theta_{0^+} := \mathbf{0}$, and $\Sigma_{0^+} := \Sigma_1$. Since the strict expansion has the same initial law, set $\tilde{\mathbf{B}}_{0^+} := \mathbf{B}_{0^+}$, $\tilde{\Theta}_{0^+} := \Theta_{0^+}$, $\tilde{\Sigma}_{0^+} := \Sigma_{0^+}$, and $\gamma_0^a = \tilde{\gamma}_0^a := \emptyset$. Fortunately, we can still apply the approach designed in [14], which expanded the problem into a new one that satisfies the SI-CIB condition, and then solved it via dynamic programming. To guarantee the SI-CIB condition, [14] introduced the following assumption.

Assumption V.1. The communication strategies take common information and past communication actions as input, i.e., $\forall i \in [n], h \in [H], g_{i,h}^m : \mathcal{C}_{h^-} \times \mathcal{M}_{1:h-1} \rightarrow \mathcal{M}_{i,h}$.

To this end, we first introduce some additional notation. We denote by $\mathcal{P}_{h^-}(M_{1:h-1}) \subseteq \mathcal{P}_{h^-}$ the set of all possible P_{h^-} given $M_{1:h-1}$. Similarly, we denote by $\mathcal{P}_{h^+}(M_{1:h}) \subseteq \mathcal{P}_{h^+}$ the set of all possible P_{h^+} given $M_{1:h}$. Then, given any $M_{1:h-1}$, we can define $\mathbf{B}_{h^-} \in \mathcal{P}(\mathcal{X} \times \mathcal{P}_{h^-}(M_{1:h-1}))$ as the conditional probability measure of state $\mathbf{X}_h$ and private information $\mathbf{P}_{h^-}$, given the past strategies $g_{1:h-1}^m, g_{1:h-1}^a$, communication actions $M_{1:h-1}$, and common information $\mathcal{C}_{h^-}$ before additional sharing. Similarly, we define $\mathbf{B}_{h^+} \in \mathcal{P}(\mathcal{X} \times \mathcal{P}_{h^+}(M_{1:h}))$ as the conditional probability measure of $\mathbf{X}_h$ and $\mathbf{P}_{h^+}$ given $g_{1:h}^m, g_{1:h}^a, \mathbf{M}_{1:h}$, and $\mathcal{C}_{h^+}$. Hence, we can write $\mathbf{B}_{h^-} = \Pi_{h^-}(\mathcal{C}_{h^-}, \mathbf{M}_{1:h-1}, g_{1:h-1}^m, g_{1:h-1}^a)$ and $\mathbf{B}_{h^+} = \Pi_{h^+}(\mathcal{C}_{h^+}, \mathbf{M}_{1:h}, g_{1:h}^m, g_{1:h}^a)$ for some functionals Π_{h^-} and Π_{h^+} , respectively. Then, we can extend the SI-CIB condition as follows.

Definition V.2 (SI-CIB Condition of $\mathcal{D}$). We say that a JCCO problem $\mathcal{D}$ satisfies the *strategy-independent common-information-based belief* condition if

- For any $h \in [H]$, any realization $C_{h^-} \in \mathcal{C}_{h^-}, M_{1:h-1} \in \mathcal{M}_{1:h-1}$ and any two strategies $(g_{1:h-1}^m, g_{1:h-1}^a)$ and $(g_{1:h-1}^{m'}, g_{1:h-1}^{a'})$ that can reach C_{h^-} and $M_{1:h-1}$, it holds that $\Pi_{h^-}(C_{h^-}, M_{1:h-1}, g_{1:h-1}^m, g_{1:h-1}^a) = \Pi_{h^-}(C_{h^-}, M_{1:h-1}, g_{1:h-1}^{m'}, g_{1:h-1}^{a'})$.
- For any $h \in [H]$, any realization $C_{h^+} \in \mathcal{C}_{h^+}, M_{1:h} \in \mathcal{M}_{1:h}$ and any two strategies $(g_{1:h}^m, g_{1:h}^a)$ and $(g_{1:h}^{m'}, g_{1:h}^{a'})$ that can reach C_{h^+} and $M_{1:h}$, it holds that $\Pi_{h^+}(C_{h^+}, M_{1:h}, g_{1:h}^m, g_{1:h}^a) = \Pi_{h^+}(C_{h^+}, M_{1:h}, g_{1:h}^{m'}, g_{1:h}^{a'})$.

Under Assumption V.1, for both timesteps h^- and h^+ , we can expand the problem $\mathcal{D}$ into a new one by adding some actions that affect other agents into the common information, similarly to Equation (IV.1). After the strict expansion, the new problem satisfies the SI-CIB condition, and the team-optimal strategy of the expanded JCCO problem can be used to obtain that of the original one. More importantly, satisfying the SI-CIB condition ensures that $\tilde{\mathbf{B}}_{h^-}$ and $\tilde{\mathbf{B}}_{h^+}$ are Gaussian distributions, as shown below.

Theorem V.3. Let $\tilde{\mathcal{D}}$ be the problem obtained by the above strict expansion from a JCCO problem $\mathcal{D}$ with PN IS that satisfies Assumptions II.1, III.2, III.4, and V.1. Then, $\tilde{\mathcal{D}}$ is a JCCO problem satisfying the SI-CIB condition. Moreover, for any $h \in [H]$, $\tilde{\mathbf{B}}_{h^-}, \tilde{\mathbf{B}}_{h^+}$ are Gaussian distributions. Formally, $\tilde{\mathbf{B}}_{h^-} = \mathcal{N}(\tilde{\boldsymbol{\Theta}}_{h^-}, \tilde{\Sigma}_{h^-}), \tilde{\mathbf{B}}_{h^+} = \mathcal{N}(\tilde{\boldsymbol{\Theta}}_{h^+}, \tilde{\Sigma}_{h^+})$, and

$$\begin{aligned} \tilde{\boldsymbol{\Theta}}_{h^-} &= \tilde{\Psi}_{h^-}^1(\tilde{\boldsymbol{\Theta}}_{(h-1)^+}, \tilde{\mathbf{M}}_{1:h-1}, \tilde{\mathbf{Z}}_h^b), & \tilde{\Sigma}_{h^-} &= \tilde{\Psi}_{h^-}^2(\tilde{\mathbf{M}}_{1:h-1}), \\ \tilde{\boldsymbol{\Theta}}_{h^+} &= \tilde{\Psi}_{h^+}^1(\tilde{\boldsymbol{\Theta}}_{h^-}, \tilde{\mathbf{M}}_{1:h}, \tilde{\mathbf{Z}}_h^a), & \tilde{\Sigma}_{h^+} &= \tilde{\Psi}_{h^+}^2(\tilde{\mathbf{M}}_{1:h}), \end{aligned} \quad (\text{V.1})$$

for some functions $\tilde{\Psi}_{h^-}^1, \tilde{\Psi}_{h^+}^1, \tilde{\Psi}_{h^-}^2, \tilde{\Psi}_{h^+}^2$ that do not depend on the strategies $\tilde{g}_{1:h}^m, \tilde{g}_{1:h}^a$.

Theorem V.3 provides a way to solve JCCO with closed-loop communication strategies via dynamic programming. Specifically, we can conduct dynamic programming over the spaces of $(\tilde{\mathbf{M}}_{1:h}, \tilde{\boldsymbol{\Theta}}_{h^+}) \in \tilde{\mathcal{M}}_{1:h} \times (\tilde{\mathcal{X}} \times \tilde{\mathcal{P}}_{h^+}(\tilde{\mathcal{M}}_{1:h}))$ and the spaces of $(\tilde{\mathbf{M}}_{1:h-1}, \tilde{\boldsymbol{\Theta}}_{h^-}) \in \tilde{\mathcal{M}}_{1:h-1} \times (\tilde{\mathcal{X}} \times \tilde{\mathcal{P}}_{h^-}(\tilde{\mathcal{M}}_{1:h-1}))$, to solve for the optimal strategies $\tilde{g}_h^{a,*}$ and $\tilde{g}_h^{m,*}$, respectively. The overall algorithm, under its stated attainment assumption, is tabulated as Algorithm 1 in §D.

VI. Numerical Example and Experimental Results

In this section, we demonstrate the proposed method through concrete JCCO examples. We first present a numerical example that illustrates the implementation of our method step by step. We then evaluate the approach experimentally and analyze its performance.

VI-A Numerical example

We illustrate the method using a linear-quadratic JCCO problem $\mathcal{D}$ with $n = 2, H = 2$, and open-loop communication strategies. The baseline sharing of $\mathcal{D}$ is *one-step measurement delay*; that is, for any $i \in [2]$ and $h \in [2]$, $C_{h^-} = \{\mathbf{Y}_{1:h-1}\}$ and $\mathbf{P}_{i,h^-} = \{\mathbf{Y}_{i,h}\}$. The communication action spaces are $\mathcal{M}_{i,h} = \{0, 1\}$ for all $i \in [2]$ and $h \in [2]$, where $\mathbf{M}_{i,h} = 1$ indicates $\mathbf{Z}_{i,h}^a = \{\mathbf{Y}_{i,h}\}$ and $\mathbf{M}_{i,h} = 0$ indicates $\mathbf{Z}_{i,h}^a = \emptyset$. The communication cost is $\mathcal{K}_h(\mathbf{M}_h) = 0.2(\mathbf{M}_{1,h} + \mathbf{M}_{2,h})$. The system is defined by $\mathcal{X} = \mathcal{Y}_i = \mathcal{U}_i = \mathbb{R}$ for all $i \in [2]$, with the following data:

$A_1 = 1.25, A_2 = 1.15, B_{1,1} = 1, B_{2,1} = 0.9, B_{1,2} = 0.85, B_{2,2} = 1.1, E_{1,1} = 1.1, E_{2,1} = 0.8, E_{1,2} = 0.9, E_{2,2} = 1.2,$
 $\Sigma_{0,1} = 1.5, \Sigma_{0,2} = 2, \Sigma_{1,1} = 1, \Sigma_{2,1} = 1.5, \Sigma_{1,2} = 2, \Sigma_{2,2} = 0.8, \Sigma_1 = 2, Q_1^1 = 5, Q_2^1 = 6, Q_3^1 = 8, Q_1^2 = Q_2^2 = \mathbb{I}_2.$

The problem $\mathcal{D}$ has PN IS and satisfies Assumptions II.1, III.2, and III.4. We solve it as follows.

Step 1: There are 16 possible communication strategies with $g_{i,h}^m \in \{0, 1\}$ for all $h \in [2]$ and $i \in [2]$. For each $g_{1:2}^m$, we construct the associated decentralized LQG problem $\mathcal{D}(g_{1:2}^m)$.

Step 2: We expand $\mathcal{D}(g_{1:2}^m)$ to $\tilde{\mathcal{D}}(g_{1:2}^m)$ and compute the corresponding filter matrices and covariance matrices from the information structure of $\tilde{\mathcal{D}}(g_{1:2}^m)$. To illustrate the computation, consider $g_{1,1}^m = g_{2,2}^m = 1$ and $g_{2,1}^m = g_{1,2}^m = 0$. This choice yields $\tilde{\mathcal{C}}_1 = \{\tilde{\mathbf{Y}}_{1,1}\}, \tilde{\mathcal{C}}_2 = \{\tilde{\mathbf{Y}}_1, \tilde{\mathbf{U}}_1, \tilde{\mathbf{Y}}_{2,2}\}, \tilde{\mathbf{P}}_{2,1} = \{\tilde{\mathbf{Y}}_{2,1}\}, \tilde{\mathbf{P}}_{1,2} = \{\tilde{\mathbf{Y}}_{1,2}\},$ and $\tilde{\mathbf{P}}_{1,1} = \tilde{\mathbf{P}}_{2,2} = \emptyset$. For timestep $h = 2$, $\{K_2^j\}_{j \in [4]}$ and $\tilde{\Sigma}_2$ are given by

$$\begin{aligned} K_2^1 &= E_{1,2}^{\text{ext}} \left(1 - \frac{\dot{P}_2 E_{2,2}^2}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} \right) A_1 \dot{P}_1 \frac{E_{1,1}^2 \Sigma_1 + \Sigma_{1,1}}{\Sigma_1 \Sigma_{1,1}} \begin{bmatrix} 1 & 0 \end{bmatrix}, K_2^2 = E_{1,2}^{\text{ext}} \left(1 - \frac{\dot{P}_2 E_{2,2}^2}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} \right) A_1 \dot{P}_1 \frac{E_{2,1}}{\Sigma_{2,1}}, \\ K_2^3 &= E_{1,2}^{\text{ext}} \left(1 - \frac{\dot{P}_2 E_{2,2}^2}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} \right) B_1, K_2^4 = E_{1,2}^{\text{ext}} \frac{\dot{P}_2 E_{2,2} \begin{bmatrix} 0 & 1 \end{bmatrix}}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}}, \tilde{\Sigma}_2 = \begin{bmatrix} \frac{\dot{P}_2 \Sigma_{2,2}}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} & \frac{E_{1,2} \dot{P}_2 \Sigma_{2,2}}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} \\ \frac{E_{1,2} \dot{P}_2 \Sigma_{2,2}}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} & \Sigma_{1,2} + \frac{E_{1,2}^2 \dot{P}_2 \Sigma_{2,2}}{E_{2,2}^2 \dot{P}_2 + \Sigma_{2,2}} \end{bmatrix}, \end{aligned} \quad (\text{VI.1})$$

where $\dot{P}_1 = (\frac{1}{\Sigma_1} + \frac{E_{1,1}^2}{\Sigma_{1,1}} + \frac{E_{2,1}^2}{\Sigma_{2,1}})^{-1}$, $\dot{P}_2 = A_1^2 \dot{P}_1 + \Sigma_{0,1}$ and $E_{1,2}^{\text{ext}} = \begin{bmatrix} 1 & E_{1,2} \end{bmatrix}^\top$. The initial quantities K_1^1, K_1^2 , and $\tilde{\Sigma}_1$ are computed similarly.

Step 3: Apply the backward recursion and finite-dimensional convex quadratic minimizations in Theorem IV.6 to compute $\tilde{c}_0$ and $\tilde{g}_{1:2}^*$ for the selected $g_{1:2}^m$, and define $f(g_{1:2}^m) := \min_{g_{1:2}^a} J_{\mathcal{D}}(g_{1:2}^m, g_{1:2}^a) = \tilde{c}_0 + \mathcal{K}_1(g_{1:2}^m) + \mathcal{K}_2(g_{1:2}^m)$.

Step 4: Compute $g_{1:2}^{m,*} \in \arg\min_{g_{1:2}^m \in \mathcal{G}_{1:2}^m} f(g_{1:2}^m)$. The optimizer is $g_{1,1}^{m,*} = g_{2,2}^{m,*} = 1$ and $g_{2,1}^{m,*} = g_{1,2}^{m,*} = 0$. The optimal strategy $\tilde{g}_{1:2}^{m,*}$ for the problem $\tilde{\mathcal{D}}(g_{1:2}^{m,*})$ is³

$$\begin{aligned} \tilde{g}_{1,1}^*(\tilde{\mathbf{I}}_{1,1}) &= -0.410 \tilde{\mathbf{Y}}_{1,1}, \tilde{g}_{2,1}^*(\tilde{\mathbf{I}}_{2,1}) = -0.219 \tilde{\mathbf{Y}}_{1,1} - 0.292 \tilde{\mathbf{Y}}_{2,1}, \\ \tilde{g}_{1,2}^*(\tilde{\mathbf{I}}_{1,2}) &= -0.038 \tilde{\mathbf{Y}}_{1,1} - 0.019 \tilde{\mathbf{Y}}_{2,1} - 0.060 \tilde{\mathbf{U}}_{1,1} - 0.054 \tilde{\mathbf{U}}_{2,1} - 0.196 \tilde{\mathbf{Y}}_{1,2} - 0.200 \tilde{\mathbf{Y}}_{2,2}, \\ \tilde{g}_{2,2}^*(\tilde{\mathbf{I}}_{2,2}) &= -0.079 \tilde{\mathbf{Y}}_{1,1} - 0.038 \tilde{\mathbf{Y}}_{2,1} - 0.123 \tilde{\mathbf{U}}_{1,1} - 0.110 \tilde{\mathbf{U}}_{2,1} - 0.410 \tilde{\mathbf{Y}}_{2,2}. \end{aligned}$$

The optimal objective of $\tilde{\mathcal{D}}(g_{1:2}^{m,*})$ is $\tilde{c}_0 = 45.947$. We then compute $g_{1:2}^{a,*} = \varphi(\tilde{g}_{1:2}^*, \mathcal{D}(g_{1:2}^{m,*}))$, where $g_1^{a,*} = \tilde{g}_1^*$, and

$$\begin{aligned} g_{1,2}^{a,*}(\mathbf{I}_{1,2^+}) &= -0.002 \mathbf{Y}_{1,1} - 0.003 \mathbf{Y}_{2,1} - 0.196 \mathbf{Y}_{1,2} - 0.200 \mathbf{Y}_{2,2}, \\ g_{2,2}^{a,*}(\mathbf{I}_{2,2^+}) &= -0.004 \mathbf{Y}_{1,1} - 0.006 \mathbf{Y}_{2,1} - 0.410 \mathbf{Y}_{2,2}, \end{aligned}$$

The optimal objective of $\mathcal{D}$ is $J_{\mathcal{D}}(g_{1:2}^{m,*}, g_{1:2}^{a,*}) = 46.347$.

VI-B Experimental results

We evaluate our approach on JCCO problems with the following four information structures. In the descriptions below, labeled observations and actions with nonpositive time indices are omit-

³All numerical values are rounded to three decimal places.

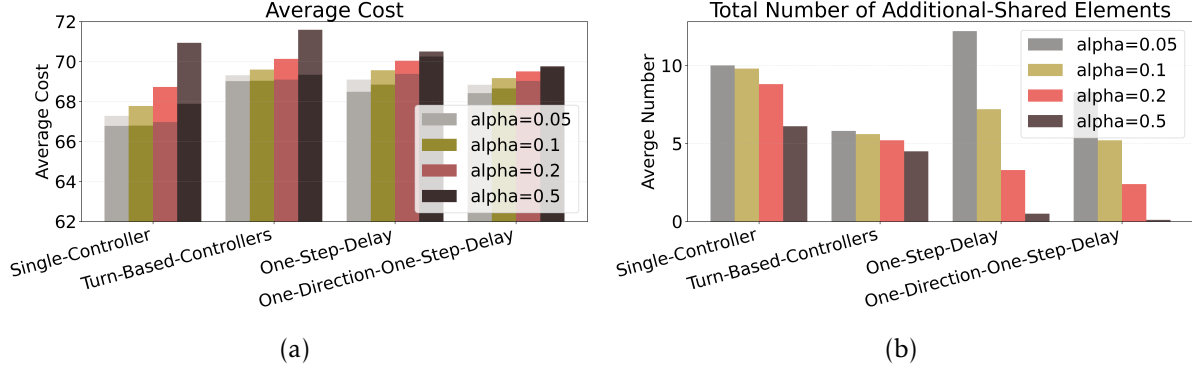


Figure 1: Average values over 10 random seeds under different baseline sharing protocols and values of α . Figure (a): For each bar, the dark and light portions correspond to the control cost and communication cost, respectively. Figure (b): Each value represents the total number of additionally shared elements.

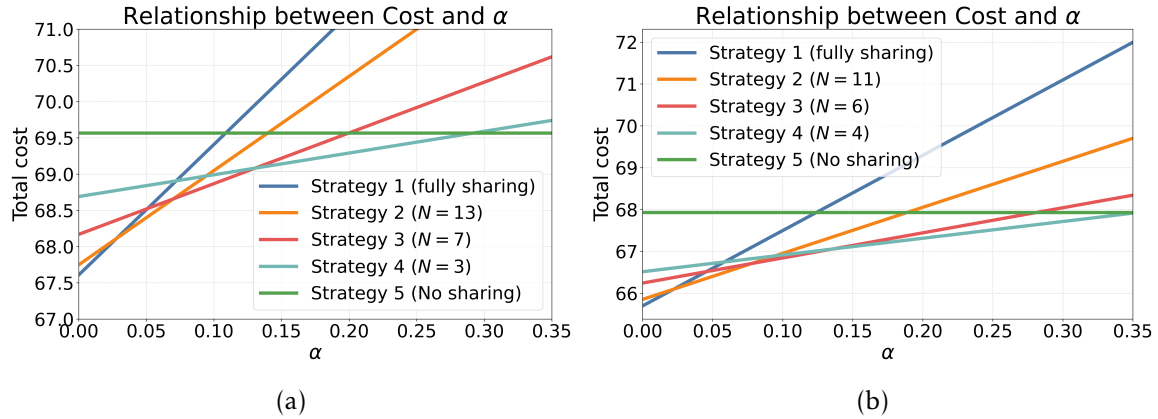


Figure 2: Total cost under different values of α and communication strategies. Each line corresponds to a different communication strategy, with N representing the number of elements shared through additional sharing. Figure (a): One-Step-Delay baseline sharing. Figure (b): One-Direction-One-Step-Delay baseline sharing.

ted, and $C_{0^+} = U_0 = \emptyset$.

- **Single-Controller:** Only agent 1 can control the underlying state X_h , i.e., $B_{i,h} = \mathbf{0}$ for all $h \in [H]$ and $i \in [2 : n]$. Agent 1 shares all of its information with the other agents, while the remaining agents share their information with a d -step delay for some $d > 1$; that is, $C_{h^-} = C_{(h-1)^+} \cup \{Y_{1,h}, U_{1,h-1}, Y_{2:n,h-d}\}$, $P_{1,h^-} = \emptyset$, and $P_{i,h^-} = P_{i,(h-1)^+} \cup \{Y_{i,h}\} \setminus \{Y_{i,h-d}\}$ for all $h \in [H]$ and $i \in [2 : n]$. In the experiments, we set $n = 3, H = 6$.
- **Turn-based-Controllers:** For each $h \in [H]$ with the unique representation $h = kn + i$, where $k \in \mathbb{Z}_{\geq 0}$ and $i \in [n]$, only agent i can control the state at timestep h , i.e., $B_{j,h} = \mathbf{0}$ for all $j \neq i$. For $h < H$, at timestep $h + 1$, agent i shares its past observations and actions with the other agents; that is, $C_{(h+1)^-} = C_{h^+} \cup \{Y_{i,h-n+1:h}, U_{i,h}\}$. In the experiments, we set $n = 3, H = 6$.
- **One-Step-Delay:** Through baseline sharing, each agent shares its past observations and actions with the other agents, i.e., $C_{h^-} = C_{(h-1)^+} \cup \{Y_{h-1}, U_{h-1}\}$ and $P_{i,h^-} = \{Y_{i,h}\}$ for all $h \in [H]$

and $i \in [n]$. In the experiments, we set $n = 3, H = 7$.

- **One-Direction-One-Step-Delay:** Agent 1 shares all of its information with the other agents, while the remaining agents share their past observations and actions through baseline sharing; that is, $\mathbf{C}_{h^-} = \mathbf{C}_{(h-1)^+} \cup \{\mathbf{Y}_{1,h}, \mathbf{Y}_{2:n,h-1}, \mathbf{U}_{h-1}\}, \mathbf{P}_{1,h^-} = \emptyset, \mathbf{P}_{i,h^-} = \{\mathbf{Y}_{i,h}\}$ for all $h \in [H]$ and $i \in [2 : n]$. In the experiments, we set $n = 3, H = 7$.

For additional sharing, each agent decides whether to share each element of its private information with the other agents. The communication cost is α times the number of elements shared additionally, where $\alpha \in \{0.05, 0.1, 0.2, 0.5\}$. We set $\mathcal{X} = \mathcal{U}_i = \mathcal{Y}_i = \mathbb{R}$ for all $i \in [n]$. The system matrices are sampled using a random seed β . Each nonzero block of the dynamics matrices $A_{1:H}, B_{1:H}, E_{1:H}$ is sampled uniformly from $[0.8, 1.3]$, $[0.5, 1.2]$, and $[0.7, 1.3]$, respectively. The noise covariance matrices $\{\Sigma_{i,h}\}_{i \in [0:n], h \in [H]}$ are set to be block diagonal, and each diagonal block is sampled uniformly from $[0.5, 2.0]$. The covariance matrix Σ_1 is also set to be block diagonal, with each diagonal block sampled uniformly from $[1.0, 3.0]$. We run experiments for all four information structures, the four values of α in $\{0.05, 0.1, 0.2, 0.5\}$, and 10 random seeds $\beta = 1, \dots, 10$.

The results are presented in Figure 1, which shows that lower communication costs encourage agents to share more information, thereby reducing both the control cost and the total objective value. We also evaluate the total cost under different values of α and different communication strategies, each paired with its optimal control strategy, as presented in Figure 2. The results show that when α is small, strategies with more additional sharing can attain lower total costs. Conversely, when α is large, the communication penalty makes strategies with less additional sharing preferable.

VII. Concluding Remarks

In this paper, we formalized joint communication-control strategy optimization in decentralized multi-agent LQG systems. We identified structural conditions under which additional information sharing preserves partial nestedness and showed that excluding them can yield either nonlinearity or non-existence of the optimal control strategies. For each fixed open-loop communication strategy, we derived Riccati-type dynamic-programming recursions for the optimal controller, which, of independent interest, can also be applied to decentralized LQG problems with partially nested information structures and output feedback, thus advancing the results in [17]. We then further extended the approach to JCCO with closed-loop communication strategies. Our work opens several important directions for future research, including extensions to non-Gaussian noise and non-cooperative settings, as well as relaxations of the assumptions for specific system structures such as those with factorized states.

Acknowledgement

The authors acknowledge the valuable feedback from the anonymous reviewers of IEEE CDC 2026, and acknowledge the support from the Army Research Office grant W911NF-24-1-0085, the NSF CAREER Award 2443704, the AFOSR YIP Award FA 9550-25-1-0258, a JP Morgan Faculty Research Award, and an AI Safety Research Award from Coefficient Giving.

References

- [1] L. Zhang and D. Hristu-Varakelis, “Communication and control co-design for networked control systems,” *Automatica*, vol. 42, no. 6, pp. 953–958, 2006.
- [2] C. Peng and T. C. Yang, “Event-triggered communication and h_∞ control co-design for networked control systems,” *Automatica*, vol. 49, no. 5, pp. 1326–1332, 2013.
- [3] N. Matni, “Communication delay co-design in $\mathcal{H}_2$ -distributed control using atomic norm minimization,” *IEEE Trans. Control Netw. Syst.*, vol. 4, no. 2, pp. 267–278, 2015.
- [4] N. Matni and V. Chandrasekaran, “Regularization for design,” *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 3991–4006, 2016.
- [5] S. Yüksel, “Jointly optimal LQG quantization and control policies for multi-dimensional systems,” *IEEE Trans. Autom. Control*, vol. 59, pp. 1612–1617, 2013.
- [6] M. Fu, “Lack of separation principle for quantized linear quadratic Gaussian control,” *IEEE Trans. Autom. Control*, vol. 57, no. 9, pp. 2385–2390, 2012.
- [7] D. Maity and P. Tsiotras, “Optimal controller synthesis and dynamic quantizer switching for linear-quadratic-gaussian systems,” *IEEE Trans. Autom. Control*, vol. 67, no. 1, pp. 382–389, 2021.
- [8] J. Foerster, I. A. Assael, N. De Freitas, and S. Whiteson, “Learning to communicate with deep multi-agent reinforcement learning,” in *NeurIPS*, 2016.
- [9] S. Sukhbaatar, A. Szlam, and R. Fergus, “Learning multiagent communication with back-propagation,” in *NeurIPS*, 2016.
- [10] 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, 2013.
- [11] A. Gupta, A. Nayyar, C. Langbort, and T. Basar, “Common information based Markov perfect equilibria for linear-gaussian games with asymmetric information,” *SIAM J. Control Optim.*, vol. 52, no. 5, pp. 3228–3260, 2014.
- [12] S. Sudhakara, D. Kartik, R. Jain, and A. Nayyar, “Optimal communication and control strategies in a cooperative multiagent MDP problem,” *IEEE Trans. Autom. Control*, vol. 69, no. 10, pp. 6959–6966, 2024.
- [13] D. Kartik, S. Sudhakara, R. Jain, and A. Nayyar, “Optimal communication and control strategies for a multi-agent system in the presence of an adversary,” in *IEEE Conf. Decis. Control (CDC)*, 2022.
- [14] X. Liu, H. You, and K. Zhang, “Principled learning-to-communicate with quasi-classical information structures,” in *IEEE Conf. Decis. Control (CDC)*, 2025.
- [15] Y.-C. Ho and K.-C. Chu, “Team decision theory and information structures in optimal control problems – part I,” *IEEE Trans. Autom. Control*, vol. 17, pp. 15–22, 1972.

- [16] A. Lamperski and L. Lessard, “Optimal decentralized state-feedback control with sparsity and delays,” *Automatica*, pp. 143–151, 2015.
- [17] 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, 2015.
- [18] A. Nayyar and L. Lessard, “Structural results for partially nested LQG systems over graphs,” in *American Control Conf. (ACC)*, 2015, pp. 5457–5464.
- [19] X. Liu and K. Zhang, “Partially observable multiagent reinforcement learning with information sharing,” *SIAM J. Control Optim.*, vol. 64, no. 2, pp. 673–697, 2026.
- [20] S. Yüksel and T. Başar, *Stochastic Teams, Games, and Control under Information Constraints*. Springer Nature, 2023.
- [21] H. S. Witsenhausen, “On information structures, feedback and causality,” *SIAM J. Control*, vol. 9, no. 2, pp. 149–160, 1971.
- [22] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, “Information structures in optimal decentralized control,” in *IEEE Conf. Decis. Control (CDC)*, 2012.
- [23] H. S. Witsenhausen, “A counterexample in stochastic optimum control,” *SIAM J. Control*, vol. 6, no. 1, pp. 131–147, 1968.
- [24] A. Nayyar, A. Gupta, C. Langbort, and T. Başar, “Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games,” *IEEE Trans. Autom. Control*, vol. 59, pp. 555–570, 2013.
- [25] S. Axler, *Linear Algebra Done Right*, 4th ed. Springer, 2024.
- [26] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank,” *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
- [27] H.-F. Chen, P. R. Kumar, and J. H. van Schuppen, “On Kalman filtering for conditionally gaussian systems with random matrices,” *Syst. Control Lett.*, vol. 13, no. 5, pp. 397–404, 1989.
- [28] D. Rappaport and L. M. Silverman, “Structure and stability of discrete-time optimal systems,” *IEEE Trans. Autom. Control*, vol. 16, no. 3, pp. 227–233, June 1971.
- [29] A. Ferrante and L. Ntogramatzidis, “The generalised discrete algebraic Riccati equation arising in LQ optimal control problems: Part I,” in *IEEE Conf. Decis. Control (CDC)*, 2012, pp. 6394–6399.

A. Decentralized Linear-Quadratic-Gaussian Control Problem

For $n > 1$ agents, a (cooperative) decentralized linear-quadratic-Gaussian (decentralized LQG) control problem $\check{D}$ is described by a tuple $\langle H, \mathcal{X}, \{\mathcal{Y}_i\}_{i \in [n]}, \{\mathcal{U}_i\}_{i \in [n]}, \{A_h\}_{h \in [H]}, \{B_{i,h}\}_{i \in [n], h \in [H]}, \{E_{i,h}\}_{i \in [n], h \in [H]}, \{Q_h^1\}_{h \in [H+1]}, \{Q_h^2\}_{h \in [H]} \rangle$, where H is the time horizon and $\mathbf{X}_h \in \mathcal{X} = \mathbb{R}^{d_x}$ is the state. At each timestep $h \in [H]$, each agent $i \in [n]$ receives a noisy observation $\mathbf{Y}_{i,h} \in \mathcal{Y}_i = \mathbb{R}^{d_y^i}$ of the state $\mathbf{X}_h$, and chooses a control action $\mathbf{U}_{i,h} \in \mathcal{U}_i = \mathbb{R}^{d_u^i}$. At timestep $h \in [H]$, we denote by $\mathbf{U}_h = [\mathbf{U}_{1,h}^\top \ \mathbf{U}_{2,h}^\top \ \cdots \ \mathbf{U}_{n,h}^\top]^\top$ the joint control action of all the n agents, and by $\mathcal{U} = \mathbb{R}^{\sum_{i=1}^n d_u^i}$ the joint control action space; we denote by $\mathbf{Y}_h = [\mathbf{Y}_{1,h}^\top \ \mathbf{Y}_{2,h}^\top \ \cdots \ \mathbf{Y}_{n,h}^\top]^\top$ the joint observation, and by $\mathcal{Y} = \mathbb{R}^{\sum_{i=1}^n d_y^i}$ the joint observation space. The system evolves as follows: for each timestep $h \in [H]$

$$\begin{aligned} \mathbf{X}_{h+1} &= A_h \mathbf{X}_h + B_h \mathbf{U}_h + \mathbf{W}_{0,h} = A_h \mathbf{X}_h + \sum_{i=1}^n B_{i,h} \mathbf{U}_{i,h} + \mathbf{W}_{0,h}, \\ \mathbf{Y}_{i,h} &= E_{i,h} \mathbf{X}_h + \mathbf{W}_{i,h}, \forall i \in [n], \end{aligned} \quad (\text{A.1})$$

where $\mathbf{W}_{i,h}$ is a random variable taking values in some real-vector space $\mathcal{W}_{i,h}$ for all $h \in [H], i = 0, 1, \dots, n$. $A_h, B_h, E_{i,h}$ are matrices of appropriate dimensions, with $B_h = [B_{1,h} \ B_{2,h} \ \cdots \ B_{n,h}]$. Here, $\mathbf{W}_{i,h} \sim \mathcal{N}(\mathbf{0}, \Sigma_{i,h})$ for some covariance matrices $\Sigma_{i,h} \geq 0$ for all $i \in [0 : n], h \in [H]$ and $\mathbf{X}_1 \sim \mathcal{N}(\mathbf{0}, \Sigma_1)$ for some covariance matrix $\Sigma_1 \geq 0$. The random variables $\mathbf{X}_1$ and $\{\mathbf{W}_{i,h}\}_{i \in [0:n], h \in [H]}$ are mutually independent.

At timestep $h \in [H]$, each agent can access some information $\mathbf{I}_{i,h} \subseteq \{\mathbf{Y}_{1:h}, \mathbf{U}_{1:h-1}\}$, and the collection of all possible such available information is denoted by $\mathcal{I}_{i,h}$. We use $\mathbf{I}_h$ to denote the joint available information at timestep h . Meanwhile, agents may share part of their information with each other, and the shared information is denoted by $\mathbf{Z}_h$. At timestep h , the common information among all agents is defined as the union of all the shared information so far: $\mathbf{C}_h = \bigcup_{t=1}^h \mathbf{Z}_t$. The private information of agent i at timestep h is thus defined as $\mathbf{P}_{i,h} = \mathbf{I}_{i,h} \setminus \mathbf{C}_h$, and the joint private information is denoted by $\mathbf{P}_h := [\mathbf{P}_{1,h}^\top \ \cdots \ \mathbf{P}_{n,h}^\top]^\top$. We denote by $\mathcal{C}_h, \mathcal{P}_h, \mathcal{P}_{i,h}, \mathcal{I}_{i,h}, \mathcal{I}_h$ the sets of values of the random variables $\mathbf{C}_h, \mathbf{P}_h, \mathbf{P}_{i,h}, \mathbf{I}_{i,h}, \mathbf{I}_h$ for each $h \in [H], i \in [n]$.

Each agent i at timestep h chooses the control action $\mathbf{U}_{i,h}$ based on some strategy $g_{i,h} : \mathcal{I}_{i,h} \rightarrow \mathcal{U}_i$. We denote by $g_h := (g_{1,h}, g_{2,h}, \dots, g_{n,h})$ the joint control strategy of all the agents, and by $g_{1:h} := (g_1, g_2, \dots, g_h), \forall h \in [H]$ the sequence of joint strategies from timesteps 1 to h . We use $\mathcal{G}_{i,h}$ to denote the strategy space of $g_{i,h}$, and use $\mathcal{G}_h, \mathcal{G}_{1:h}$ to denote the corresponding joint strategy spaces. At each timestep $h \in [H]$, the cost is defined as $c_h = \mathbf{X}_h^\top Q_h^1 \mathbf{X}_h + \mathbf{U}_h^\top Q_h^2 \mathbf{U}_h$, and the cost of final timestep is $c_{H+1} = \mathbf{X}_{H+1}^\top Q_{H+1}^1 \mathbf{X}_{H+1}$. Here, $\forall h \in [H], Q_h^1, Q_h^2 \geq 0, Q_{H+1}^1 \geq 0$ are symmetric matrices with appropriate dimensions. The objective among all the agents is thus defined as

$$J_{\check{D}}(g_{1:H}) := \mathbb{E} \left[\sum_{h=1}^H c_h + c_{H+1} \mid g_{1:H} \right].$$

With this objective, we define the notion of team optimality for the problem $\check{D}$ as follows.

Definition A.1 (Team optimality). We call a joint strategy $g_{1:H}^* \in \mathcal{G}_{1:H}$ a *team-optimal strategy* of

the decentralized LQG control problem $\check{\mathcal{D}}$ if

$$\forall g_{1:H} \in \mathcal{G}_{1:H}, \quad J_{\check{\mathcal{D}}}(g_{1:H}) \geq J_{\check{\mathcal{D}}}(g_{1:H}^*). \quad (\text{A.2})$$

Corollary A.2 ([15] with positive semidefinite matrices). Every finite-horizon decentralized LQG problem above with a partially nested information structure admits a team-optimal linear strategy, even when the primitive Gaussian covariance matrices and the state and control cost matrices $\{Q_h^1\}_{h \in [H+1]}$ and $\{Q_h^2\}_{h \in [H]}$ are positive semidefinite and possibly singular.

Proof. We follow Theorems 1 and 2 in [15]. Regarding each *agent-time* pair (i, h) as a *decision maker*, partial nestedness allows the effects of preceding actions to be recursively subtracted from each information vector, exactly as in Eqs. (28) and (29) of [15]. Denote the resulting static information by $\widehat{\mathbf{I}}_{i,h}$. This transformation and its reverse preserve the actions and the value of $J_{\check{\mathcal{D}}}$. It therefore remains to establish linear optimality for the resulting static team.

Let $\mathbf{W}$ stack $\mathbf{X}_1$ and all the process and observation noises, and let $\mathbf{X}_{1:H+1}$ and $\mathbf{U}_{1:H}$ denote the stacked state and action vectors. Recursively substituting the dynamics gives fixed matrices $\mathcal{A}$ and $\mathcal{B}$ such that

$$\mathbf{X}_{1:H+1} = \mathcal{A}\mathbf{W} + \mathcal{B}\mathbf{U}_{1:H}.$$

Consequently,

$$J_{\check{\mathcal{D}}} = \mathbb{E}[\mathbf{U}_{1:H}^\top \bar{\mathbf{Q}} \mathbf{U}_{1:H} + 2\mathbf{U}_{1:H}^\top \bar{\mathbf{S}} \mathbf{W}] + c_0,$$

where

$$\begin{aligned} \bar{\mathbf{Q}} &:= \mathcal{B}^\top \text{diag}(Q_1^1, \dots, Q_{H+1}^1) \mathcal{B} + \text{diag}(Q_1^2, \dots, Q_H^2) \geq 0, \\ \bar{\mathbf{S}} &:= \mathcal{B}^\top \text{diag}(Q_1^1, \dots, Q_{H+1}^1) \mathcal{A}, \end{aligned}$$

and c_0 is independent of the strategy.

Write

$$\mathbb{E}[\mathbf{W}\mathbf{W}^\top] = \mathbf{L}\mathbf{L}^\top, \quad \mathbf{W} = \mathbf{L}\boldsymbol{\xi}, \quad \boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}),$$

with $\mathbf{L}$ having full column rank. Since $\widehat{\mathbf{I}}_{i,h}$ is linear in $\mathbf{W}$, remove its linearly dependent coordinates and denote the resulting information by $\bar{\mathbf{I}}_{i,h} = \bar{\mathbf{C}}_{i,h} \boldsymbol{\xi}$. The vectors $\widehat{\mathbf{I}}_{i,h}$ and $\bar{\mathbf{I}}_{i,h}$ determine one another through fixed linear maps, while every nonempty $\bar{\mathbf{I}}_{i,h}$ has a positive-definite covariance matrix.

Consider static linear strategies $\mathbf{U}_{i,h} = K_{i,h} \bar{\mathbf{I}}_{i,h}$. As a function of the entries of $\{K_{i,h}\}_{i,h}$, $J_{\check{\mathcal{D}}}$ is a finite-dimensional quadratic function with a positive-semidefinite quadratic part. It is bounded below because $J_{\check{\mathcal{D}}} \geq 0$. Hence, its linear part vanishes on the null space of its quadratic part, since otherwise, the objective would be unbounded below along a null direction. Its normal equations are therefore consistent, and a minimizing collection $\{K_{i,h}^*\}_{i,h}$ exists. Let $\mathbf{U}_{i,h}^* := K_{i,h}^* \bar{\mathbf{I}}_{i,h}$ and define

$$\mathbf{G}^* := \bar{\mathbf{Q}} \mathbf{U}_{1:H}^* + \bar{\mathbf{S}} \mathbf{L} \boldsymbol{\xi}.$$

If $\mathbf{G}_{i,h}^*$ denotes the block corresponding to $\mathbf{U}_{i,h}^*$, the normal equation for $K_{i,h}$ gives $\mathbb{E}[\mathbf{G}_{i,h}^* \bar{\mathbf{I}}_{i,h}^\top] = 0$. Since $(\mathbf{G}_{i,h}^*, \bar{\mathbf{I}}_{i,h})$ is jointly centered Gaussian, this implies $\mathbb{E}[\mathbf{G}_{i,h}^* \mid \bar{\mathbf{I}}_{i,h}] = 0$. For empty effective information, the same conclusion follows from centeredness.

Now let $g_{1:H}$ be any admissible static strategy and set

$$\Delta \mathbf{U}_{i,h} := g_{i,h}(\bar{\mathbf{I}}_{i,h}) - \mathbf{U}_{i,h}^*.$$

Expanding the quadratic objective and conditioning on $\bar{\mathbf{I}}_{i,h}$ gives

$$J_{\bar{\mathcal{D}}}(g_{1:H}) - J_{\bar{\mathcal{D}}}(g_{1:H}^*) = 2 \sum_{h=1}^H \sum_{i=1}^n \mathbb{E}[\Delta \mathbf{U}_{i,h}^\top \mathbf{G}_{i,h}^*] + \mathbb{E}[\Delta \mathbf{U}_{1:H}^\top \bar{\mathbf{Q}} \Delta \mathbf{U}_{1:H}] = \mathbb{E}[\Delta \mathbf{U}_{1:H}^\top \bar{\mathbf{Q}} \Delta \mathbf{U}_{1:H}] \geq 0.$$

Thus, the linear strategy is team-optimal for the static problem.

Finally, reverse the reduction in [15] over the precedence groups. For the first group, the static and original information vectors coincide, so the corresponding actions are linear in the original information. Suppose the actions in all preceding groups have been reconstructed as linear functions of their original information. Whenever a preceding action appears in the subtraction defining the current static information, partial nestedness makes the information determining that action available to the current decision maker. Hence the preceding action, and therefore the current static information, is linear in the current original information. The current static linear action is consequently linear in the original information. Induction gives a linear dynamic strategy. The forward and reverse reductions generate the same actions and costs for every realization, so the reconstructed strategy is team-optimal for the original dynamic problem. $\square$

B. Deferred Details of §III

We start with the following auxiliary lemma.

Lemma B.1. For any $n_1, n_2 \in \mathbb{N}$, let $\mathbf{Q}_1 \in \mathbb{R}^{n_1 \times n_1}$ and $\mathbf{Q}_2 \in \mathbb{R}^{n_2 \times n_2}$ be symmetric positive-semidefinite matrices, and let $\mathbf{A}, \mathbf{B}$ be matrices of compatible dimensions. Then, $\ker(\mathbf{Q}_1 + \mathbf{B}^\top \mathbf{Q}_2 \mathbf{B}) \subseteq \ker(\mathbf{A}^\top \mathbf{Q}_2 \mathbf{B})$.

Proof. For any $\mathbf{v} \in \ker(\mathbf{Q}_1 + \mathbf{B}^\top \mathbf{Q}_2 \mathbf{B})$, we have $\mathbf{v}^\top \mathbf{Q}_1 \mathbf{v} + \mathbf{v}^\top \mathbf{B}^\top \mathbf{Q}_2 \mathbf{B} \mathbf{v} = 0$. Since $\mathbf{Q}_1 \geq 0, \mathbf{Q}_2 \geq 0$, we know $\mathbf{v}^\top \mathbf{Q}_1 \mathbf{v} = \mathbf{v}^\top \mathbf{B}^\top \mathbf{Q}_2 \mathbf{B} \mathbf{v} = 0$. Then, from $\mathbf{v}^\top \mathbf{B}^\top \mathbf{Q}_2 \mathbf{B} \mathbf{v} = 0$, we have $\mathbf{Q}_2^{\frac{1}{2}} \mathbf{B} \mathbf{v} = 0$ and thus $\mathbf{Q}_2 \mathbf{B} \mathbf{v} = 0$, which further implies that $\mathbf{A}^\top \mathbf{Q}_2 \mathbf{B} \mathbf{v} = 0$ and thus $\mathbf{v} \in \ker(\mathbf{A}^\top \mathbf{Q}_2 \mathbf{B})$. This completes the proof. $\square$

B-1 Proof of Lemma III.1

Proof. To prove this lemma, we leverage the following proposition.

Proposition B.2 (Adapted from Theorems 1 and 2 of [23]). There exists a decentralized LQG control problem with $n = 2, H = 2$ and parameters k and σ_0 such that no linear control strategy is team-optimal:

$$\begin{aligned} &\text{All random variables are scalar, } \mathbf{X}_1 \sim \mathcal{N}(0, \sigma_0^2), \mathbf{X}_2 = \mathbf{X}_1 + \mathbf{U}_{1,1}, \mathbf{X}_3 = \mathbf{X}_2 - \mathbf{U}_{2,2}, \\ &\mathbf{Y}_{1,1} = \mathbf{X}_1, \mathbf{Y}_{2,1} = 0, \mathbf{Y}_{1,2} = 0, \mathbf{Y}_{2,2} = \mathbf{X}_2 + \mathbf{W}_{2,2}, \mathbf{W}_{2,2} \sim \mathcal{N}(0, 1), \\ &c_1 = k^2 \mathbf{U}_{1,1}^2, c_2 = 0, c_3 = \mathbf{X}_3^2, \text{ no information sharing between agents.} \end{aligned}$$

Then, using the parameters k and σ_0 specified in Proposition B.2, we construct a JCCO problem $\mathcal{D}$ with $n = 2, H = 2$, with the system dynamics being defined as

$$\begin{aligned} &\mathbf{X}_1 \sim \mathcal{N}(0, \sigma_0^2), \mathbf{X}_2 = \mathbf{X}_1 + \mathbf{U}_{1,1}, \mathbf{X}_3 = \mathbf{X}_2 - \mathbf{U}_{2,2}, \\ &\mathbf{Y}_{1,1} = \mathbf{X}_1, \mathbf{Y}_{2,1} = 0, \mathbf{Y}_{1,2} = 0, \mathbf{Y}_{2,2} = \mathbf{X}_2 + \mathbf{W}_{2,2}, \mathbf{W}_{2,2} \sim \mathcal{N}(0, 1). \end{aligned}$$

For the communication part, suppose the baseline sharing is null and $\mathcal{M}_{i,1} = \{0,1\}, \mathcal{M}_{1,2} = \{0,1\}^3, \mathcal{M}_{2,2} = \{0,1\}^2$, such that $\mathbf{M}_{i,1}$ represents whether agent i shares $\mathbf{Y}_{i,1}$, each digit of $\mathbf{M}_{1,2}$ represents whether agent 1 shares $\mathbf{Y}_{1,1}, \mathbf{U}_{1,1}, \mathbf{Y}_{1,2}$, and each digit of $\mathbf{M}_{2,2}$ represents whether agent 2 shares $\mathbf{Y}_{2,1}, \mathbf{Y}_{2,2}$. For control costs, we set $c_1 = k^2 \mathbf{U}_{1,1}^2, c_2 = 0, c_3 = \mathbf{X}_3^2$, where k is specified as in Proposition B.2; for communication costs, we set $\kappa_h = (\sigma_0^2 + 1) \mathbb{1}[\mathbf{M}_h \text{ is not all } 0], \forall h \in [2]$. The information available to each agent at each timestep is defined as: $\forall i \in [2], \mathbf{I}_{i,1^-} = \{\mathbf{Y}_{i,1}\}, \mathbf{I}_{i,1^+} = \mathbf{I}_{i,1^-} \cup \mathbf{Z}_1^a$, and $\mathbf{I}_{1,2^-} = \mathbf{I}_{1,1^+} \cup \{\mathbf{U}_{1,1}, \mathbf{Y}_{1,2}\}, \mathbf{I}_{2,2^-} = \mathbf{I}_{2,1^+} \cup \{\mathbf{Y}_{2,2}\}, \mathbf{I}_{i,2^+} = \mathbf{I}_{i,2^-} \cup \mathbf{Z}_2^a$. Then, we can verify that:

- $\mathcal{D}$ satisfies Assumption II.1.
- $\mathcal{D}$ satisfies Assumption III.2: The only zero input coefficient relevant to a later decision is the scalar $B_{2,1} = 0$, and $\mathbf{U}_{2,1}$ does not appear in any later information set.
- $\mathcal{D}$ satisfies Assumption III.4: It holds that $E_{1,2}B_{2,1} = B_{2,1} = 0, E_{2,2}B_{1,1} \neq 0$ and $B_{1,1} \neq 0$.
- $\mathcal{D}$ does not have PN IS: If there is no additional sharing, agent $(1,1)$ influences agent $(2,2)$, but $\mathbf{I}_{1,1^-} \not\subseteq \mathbf{I}_{2,2^-}$.

For any communication strategy $g_{1:2}^m$, if it chooses any $\mathbf{M}_{i,h}$ with any digit non-zero (namely, sharing *some* information via additional sharing), then it cannot achieve the optimum, since it will suffer from a communication cost of $\kappa_h \geq \sigma_0^2 + 1$, which is larger than the total expected cost when the agents choose $\mathbf{U}_{1,1} = \mathbf{U}_{2,2} = 0$, and they do not have any additional sharing. Therefore, the optimal communication strategy $g_{1:2}^{m,*}$ must yield no additional sharing. The remaining control subproblem is exactly Proposition B.2; hence every team-optimal control strategy is nonlinear. This completes the proof. $\square$

B-2 Proof of Lemma III.3

Proof. This proof consists of two **Parts**. **Part 1:** we construct a JCCO problem $\mathcal{D}_1$ whose optimal control strategy is nonlinear; **Part 2:** we construct another JCCO problem $\mathcal{D}_2$ whose team-optimal strategy does not exist.

Part 1: Consider an $H = 2, n = 2$ JCCO problem $\mathcal{D}_1$ with the system dynamics being defined as

$$\begin{aligned} \forall h \in [2], \mathbf{X}_h \in \mathbb{R}^2, \mathbf{U}_{1,1}, \mathbf{U}_{1,2} \in \mathbb{R}, \mathbf{U}_{2,1}, \mathbf{U}_{2,2} \in \mathbb{R}^2, \mathbf{X}_2 = \mathbf{X}_1, \mathbf{X}_3 = \mathbf{X}_2 - \mathbf{U}_{2,2}, \\ \mathbf{X}_1 \sim \mathcal{N}(\mathbf{0}, \mathbb{I}), \mathbf{Y}_{1,1} = \mathbf{X}_1, \mathbf{Y}_{1,2} = \mathbf{Y}_{2,1} = \mathbf{Y}_{2,2} = 0, c_1 = c_2 = 0, c_3 = \mathbf{X}_3^\top \mathbf{X}_3. \end{aligned}$$

For the communication part, suppose the baseline sharing is null and $\mathcal{M}_{i,1} = \{0,1\}, \mathcal{M}_{i,2} = \{0,1\}^3, \forall i \in [2]$, where $\mathbf{M}_{i,1}$ represents whether agent i shares $\mathbf{Y}_{i,1}$ and each digit of $\mathbf{M}_{i,2}$ represents whether agent i shares $\mathbf{Y}_{i,1}, \mathbf{U}_{i,1}$, and $\mathbf{Y}_{i,2}$, respectively. Set $\kappa_1 = \mathbb{1}[\mathbf{M}_{1,1} = 1 \text{ or } \mathbf{M}_{2,1} = 1]$ and $\kappa_2 = \mathbb{1}[\mathbf{M}_{1,2} \notin \{(0,0,0), (0,1,0)\} \text{ or } \mathbf{M}_{2,2} \neq (0,0,0)]$. Thus $(0,1,0)$ shares only $\mathbf{U}_{1,1}$ at zero cost, while every other nonempty sharing incurs cost one. The information available to each agent at each timestep is defined as $\forall i \in [2], \mathbf{I}_{i,1^-} = \{\mathbf{Y}_{i,1}\}, \mathbf{I}_{i,1^+} = \mathbf{I}_{i,1^-} \cup \mathbf{Z}_1^a$, and $\mathbf{I}_{1,2^-} = \mathbf{I}_{1,1^+} \cup \{\mathbf{U}_{1,1}, \mathbf{Y}_{1,2}\}, \mathbf{I}_{2,2^-} = \mathbf{I}_{2,1^+} \cup \{\mathbf{U}_{2,1}, \mathbf{Y}_{2,2}\}, \mathbf{I}_{i,2^+} = \mathbf{I}_{i,2^-} \cup \mathbf{Z}_2^a$.

Then, we can verify that:

- $\mathcal{D}_1$ satisfies Assumption II.1.
- $\mathcal{D}_1$ has PN IS: If there is no additional sharing, then for any $i_1, i_2 \in [2], \mathbf{I}_{i_1,1^-} \subseteq \mathbf{I}_{i_2,2^-}$ if $i_1 = i_2$, and otherwise agent $(i_1, 1)$ does not influence agent $(i_2, 2)$.

- $\mathcal{D}_1$ does not satisfy Assumption III.2: $B_{1,1} = B_{2,1} = 0$, but $\mathbf{U}_{1,1} \in \mathbf{I}_{1,2^-}$ and $\mathbf{U}_{2,1} \in \mathbf{I}_{2,2^-}$.
- $\mathcal{D}_1$ satisfies Assumption III.4: For any $i \in [2]$, $E_{-i,2} B_{i,1} = B_{i,1} = 0$.

Now, we aim to show that the optimal control strategy is nonlinear. Firstly, we can construct a strategy $(g_{1:2}^{m,*}, g_{1:2}^{a,*})$ as

$$g_{1,1}^{m,*} = g_{2,1}^{m,*} = 0, g_{1,2}^{m,*} = (0, 1, 0), g_{2,2}^{m,*} = (0, 0, 0), g_{1,1}^{a,*}(\mathbf{I}_{1,1^+}) = \varphi_{bi}(\mathbf{Y}_{1,1}), g_{2,2}^{a,*}(\mathbf{I}_{2,2^+}) = \varphi_{bi}^{-1}(\mathbf{U}_{1,1}),$$

where φ_{bi} is a Borel isomorphism that pushes $\mathcal{N}(0, I_2)$ to $\mathcal{N}(0, 1)$ (such an isomorphism exists between atomless standard probability spaces), and the other two controls are zero. Hence the strategy is square-integrable and has total cost zero almost surely. Since every cost is nonnegative, it is team-optimal. Such a Borel bijection cannot be linear: every linear $\mathbb{T} : \mathbb{R}^2 \rightarrow \mathbb{R}$ has a nontrivial kernel by the rank-nullity theorem [25].

However, we will show that for any strategy $(g_{1:2}^m, g_{1:2}^a)$ such that $g_{1:2}^a$ is linear, $(g_{1:2}^m, g_{1:2}^a)$ cannot be team-optimal, i.e., $J_{\mathcal{D}_1}(g_{1:2}^m, g_{1:2}^a) > 0$. We prove it by contradiction, and suppose that $(g_{1:2}^{m'}, g_{1:2}^{a'})$ satisfies that $J_{\mathcal{D}_1}(g_{1:2}^{m'}, g_{1:2}^{a'}) = 0$ and $g_{1:2}^{a'}$ is linear. Then, it must hold that $g_{1,1}^{m'} = g_{2,1}^{m'} = 0, g_{2,2}^{m'} = (0, 0, 0), g_{1,2}^{m'} = (0, 0, 0)$ or $(0, 1, 0)$, since otherwise the communication cost is at least 1. If $g_{1,2}^{m'} = (0, 0, 0)$, changing it to $(0, 1, 0)$ preserves the original controls as feasible prescriptions, adds no communication cost, and therefore preserves the zero value. We may thus assume $g_{1,2}^{m'} = (0, 1, 0)$. Then $\mathbf{I}_{i,1^+} = \{\mathbf{Y}_{i,1}\}$ for $i \in [2]$ and $\mathbf{I}_{2,2^+} = \{\mathbf{Y}_{2,1}, \mathbf{U}_{2,1}, \mathbf{Y}_{2,2}, \mathbf{U}_{1,1}\}$. Since $g_{1:2}^{a'}$ is linear, we can write:

$$\mathbf{U}_{1,1} = G_{1,1}^{1,1,y} \mathbf{Y}_{1,1}, \quad \mathbf{U}_{2,1} = G_{2,1}^{2,1,y} \mathbf{Y}_{2,1}, \quad \mathbf{U}_{2,2} = G_{2,2}^{2,1,y} \mathbf{Y}_{2,1} + G_{2,2}^{2,2,y} \mathbf{Y}_{2,2} + G_{2,2}^{1,1,u} \mathbf{U}_{1,1} + G_{2,2}^{2,1,u} \mathbf{U}_{2,1},$$

for some matrices $G_{1,1}^{1,1,y}, G_{2,1}^{2,1,y}, G_{2,2}^{2,1,y}, G_{2,2}^{2,2,y}, G_{2,2}^{1,1,u}, G_{2,2}^{2,1,u}$ with proper dimensions. Then, we can write $\mathbf{U}_{2,2}$ as $\mathbf{U}_{2,2} = G_{2,2}^{1,1,u} G_{1,1}^{1,1,y} \mathbf{Y}_{1,1}$. Then, we can write $J_{\mathcal{D}_1}(g_{1:2}^{m'}, g_{1:2}^{a'}) = \mathbb{E}[(\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y}) \mathbf{X}_1]^\top (\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y}) \mathbf{X}_1] = \text{Tr}((\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y})^\top (\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y}))$. Note that $G_{2,2}^{1,1,u}$ is a 2×1 matrix and $G_{1,1}^{1,1,y}$ is a 1×2 matrix. Therefore, we have $\text{rank}(G_{2,2}^{1,1,u} G_{1,1}^{1,1,y}) \leq 1$, and then $\text{Tr}((\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y})^\top (\mathbb{I} - G_{2,2}^{1,1,u} G_{1,1}^{1,1,y})) \geq 1$ due to Eckart-Young Theorem [26]. This means that $J_{\mathcal{D}_1}(g_{1:2}^{m'}, g_{1:2}^{a'}) > 0$ and leads to the contradiction. Hence, we know that there exists a team-optimal strategy of $\mathcal{D}_1$ and the control strategy of any team-optimal strategy is nonlinear, which completes the proof of **Part 1**.

Part 2: Consider an $n = 2, H = 2$ JCCO problem $\mathcal{D}_2$ with the system dynamics being defined as

$$\begin{aligned} \forall h \in [2], i \in [2], \mathbf{X}_h, \mathbf{X}_3, \mathbf{Y}_{i,h}, \mathbf{U}_{i,h} \in \mathbb{R}, \mathbf{X}_2 = \mathbf{X}_1, \mathbf{X}_3 = \mathbf{X}_2 - \mathbf{U}_{2,2}, \\ \mathbf{Y}_{1,1} = \mathbf{X}_1, \mathbf{Y}_{1,2} = \mathbf{Y}_{2,1} = \mathbf{Y}_{2,2} = 0, \mathbf{X}_1 \sim \mathcal{N}(0, \frac{1}{4}), c_1 = \mathbf{U}_{1,1}^2, c_2 = 0, c_3 = \mathbf{X}_3^2. \end{aligned}$$

For the communication part, similarly to **Part 1**, suppose the baseline sharing is null and $\mathcal{M}_{i,1} = \{0, 1\}, \mathcal{M}_{i,2} = \{0, 1\}^3, \forall i \in [2]$, where $\mathbf{M}_{i,1}$ represents whether agent i shares $\mathbf{Y}_{i,1}$ and each digit of $\mathbf{M}_{i,2}$ represents whether agent i shares $\mathbf{Y}_{i,1}, \mathbf{U}_{i,1}$, and $\mathbf{Y}_{i,2}$, respectively. We set the communication costs as $\kappa_1 = \mathbb{1}[\mathbf{M}_{1,1} = 1 \text{ or } \mathbf{M}_{2,1} = 1], \kappa_2 = \mathbb{1}[\mathbf{M}_{1,2} \notin \{(0, 0, 0), (0, 1, 0)\} \text{ or } \mathbf{M}_{2,2} \neq (0, 0, 0)]$. This communication cost means that if any information except $\mathbf{U}_{1,1}$ is shared, then all agents will incur a communication cost of 1. The information available to each agent at each timestep is defined as $\forall i \in [2], \mathbf{I}_{i,1^-} = \{\mathbf{Y}_{i,1}\}, \mathbf{I}_{i,1^+} = \mathbf{I}_{i,1^-} \cup \mathbf{Z}_1^a, \mathbf{I}_{i,2^-} = \mathbf{I}_{i,1^+} \cup \{\mathbf{U}_{i,1}, \mathbf{Y}_{i,2}\}, \mathbf{I}_{i,2^+} = \mathbf{I}_{i,2^-} \cup \mathbf{Z}_2^a$.

Then, we can verify that:

- $\mathcal{D}_2$ satisfies Assumption II.1.
- $\mathcal{D}_2$ has PN IS: If there is no additional sharing, then for any $i_1, i_2 \in [2]$, $I_{i_1,1^-} \subseteq I_{i_2,2^-}$ if $i_1 = i_2$, and otherwise agent $(i_1, 1)$ does not influence agent $(i_2, 2)$.
- $\mathcal{D}_2$ does not satisfy Assumption III.2: $B_{1,1} = B_{2,1} = 0$, but $U_{1,1} \in I_{1,2^-}$ and $U_{2,1} \in I_{2,2^-}$.
- $\mathcal{D}_2$ satisfies Assumption III.4: For any $i \in [2]$, $E_{-i,2} B_{i,1} = B_{i,1} = 0$.

We prove it by contradiction. We assume $(g_{1:2}^{m,*}, g_{1:2}^{a,*})$ is a team-optimal strategy of $\mathcal{D}_2$. Firstly, it holds that $g_{1,1}^{m,*} = g_{2,1}^{m,*} = 0, g_{2,2}^{m,*} = (0, 0, 0), g_{1,2}^{m,*} = (0, 0, 0)$ or $(0, 1, 0)$, since otherwise $J_{\mathcal{D}_2}(g_{1:2}^{m,*}, g_{1:2}^{a,*}) \geq \mathbb{E}[\kappa_1 + \kappa_2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] \geq 1$; however, if we consider strategy $(g_{1:2}^m, g_{1:2}^a)$ that shares nothing and chooses $U_{i,h} = 0, \forall i \in [2], h \in [2]$, then $J_{\mathcal{D}_2}(g_{1:2}^m, g_{1:2}^a) = \mathbb{E}[X_3^2 | g_{1:2}^m, g_{1:2}^a] = \mathbb{E}[X_1^2] = \frac{1}{4}$. Secondly, we can assume $g_{1,2}^{m,*} = (0, 1, 0)$, otherwise we can change it to be $(0, 1, 0)$ and it is still a team-optimal strategy, since additionally sharing $U_{1,1}$ enlarges the $I_{2,2^+}$ but incurs no communication cost. Therefore, we know that under $g_{1:2}^{m,*}$, we have $I_{i,1^+} = \{Y_{i,1}\}, \forall i \in [2], I_{2,2^+} = \{Y_{2,1}, U_{2,1}, Y_{2,2}, U_{1,1}\}$. Thirdly, if $\mathbb{P}(U_{1,1} = 0 | g_{1:2}^{m,*}, g_{1:2}^{a,*}) = 1$, then we consider the realization $I_{2,2^+} = \{Y_{2,1} = 0, U_{2,1} = g_{2,1}^{a,*}(0), Y_{2,2} = 0, U_{1,1} = 0\}$, and let $U_{2,2} = g_{2,2}^{a,*}(I_{2,2^+})$ be the realization of $U_{2,2}$ when $g_{2,2}^{a,*}$ takes the realization $I_{2,2^+}$ as input. Then, it holds that $\mathbb{P}(U_{2,2} = U_{2,2} | g_{1:2}^{m,*}, g_{1:2}^{a,*}) = 1$, and we have

$$J_{\mathcal{D}_2}(g_{1:2}^{m,*}, g_{1:2}^{a,*}) \geq \mathbb{E}[X_3^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] = \mathbb{E}[(X_1 - U_{2,2})^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] = \mathbb{E}[(X_1 - U_{2,2})^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] \geq \mathbb{E}[X_1^2] \geq \frac{1}{4}.$$

However, we consider the strategy $(g_{1:2}^{m,*}, g_{1:2}^{a'})$ with $g_{1:2}^{a'}$ being defined as $g_{1,1}^{a'}(I_{1,1^+}) = \frac{Y_{1,1}}{2}, g_{2,2}^{a'}(I_{2,2^+}) = 2U_{1,1}$ if $U_{1,1} \in I_{2,2^+}$ and otherwise 0, and choose arbitrary $g_{2,1}^{a'}, g_{1,2}^{a'}$. Then, we can verify that

$$J_{\mathcal{D}_2}(g_{1:2}^{m,*}, g_{1:2}^{a'}) = \mathbb{E}[c_1 + c_3 | g_{1:2}^{m,*}, g_{1:2}^{a'}] = \mathbb{E}[U_{1,1}^2 + X_3^2 | g_{1:2}^{m,*}, g_{1:2}^{a'}] = \mathbb{E}[\frac{1}{4}Y_{1,1}^2 + (X_1 - 2 \cdot \frac{Y_{1,1}}{2})^2] = \mathbb{E}[\frac{1}{4}Y_{1,1}^2] = \frac{1}{16}.$$

Therefore, we know that it holds that $\mathbb{P}(U_{1,1} = 0 | g_{1:2}^{m,*}, g_{1:2}^{a,*}) < 1$, and then $\mathbb{E}[U_{1,1}^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] > 0$. Lastly, we can construct a new strategy $g_{1:2}^a$ based on $g_{1:2}^{a,*}$ as

$$\begin{aligned} g_{1,1}^a(I_{1,1^+}) &= \frac{1}{2}g_{1,1}^{a,*}(I_{1,1^+}), \\ g_{2,2}^a(I_{2,2^+}) &= g_{2,2}^a(Y_{2,1}, U_{2,1}, Y_{2,2}, U_{1,1}) = g_{2,2}^{a,*}(Y_{2,1}, U_{2,1}, Y_{2,2}, 2U_{1,1}) \\ &\quad \text{if } I_{2,2^+} = \{Y_{2,1}, U_{2,1}, Y_{2,2}, U_{1,1}\}, \text{ otherwise } 0. \end{aligned}$$

All other control rules remain unchanged. This construction means that $g_{1,1}^a$ will choose $U_{1,1}$ to be half of the $U_{1,1}$ chosen by $g_{1,1}^{a,*}$, but $g_{2,2}^{a,*}, g_{2,2}^a$ choose the same $U_{2,2}$. Therefore, we can verify that

$$\begin{aligned} J_{\mathcal{D}_2}(g_{1:2}^{m,*}, g_{1:2}^a) &= \mathbb{E}[U_{1,1}^2 | g_{1:2}^{m,*}, g_{1:2}^a] + \mathbb{E}[(X_1 - U_{2,2})^2 | g_{1:2}^{m,*}, g_{1:2}^a] \\ &= \mathbb{E}[(U_{1,1}/2)^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] + \mathbb{E}[(X_1 - U_{2,2})^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] \\ &< \mathbb{E}[U_{1,1}^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] + \mathbb{E}[(X_1 - U_{2,2})^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] = J_{\mathcal{D}_2}(g_{1:2}^{m,*}, g_{1:2}^{a,*}), \end{aligned}$$

where the inequality holds because $\mathbb{E}[U_{1,1}^2 | g_{1:2}^{m,*}, g_{1:2}^{a,*}] > 0$ as proved before. This means that for any team-optimal strategy $g_{1:2}^{m,*}, g_{1:2}^{a,*}$, we can construct a strategy $(g_{1:2}^{m,*}, g_{1:2}^a)$ that is strictly better than it. Therefore, the team-optimal strategy does not exist. $\square$

B-3 Proof of Lemma III.5

Proof. Consider a JCCO problem $\mathcal{D}$ with $n = 2, H = 2$, parameters k, σ_0 specified in Proposition B.2, and the system dynamics being defined as

$$\begin{aligned} \forall h \in [2], i \in [2], X_h, X_3, Y_{i,h}, U_{i,h} \in \mathbb{R}, X_2 = X_1 + U_{1,1}, X_3 = X_2 - U_{2,2}, \\ Y_{1,1} = X_1, Y_{1,2} = X_2 + W_{1,2}, Y_{2,1} = Y_{2,2} = 0, X_1 \sim \mathcal{N}(0, \sigma_0^2), W_{1,2} \sim \mathcal{N}(0, 1). \end{aligned}$$

For the communication part, baseline sharing is null and

$$\mathcal{M}_{i,1} = \{0, 1\}, \mathcal{M}_{1,2} = \{0, 1\}^3, \mathcal{M}_{2,2} = \{0, 1\}^2,$$

where $\mathbf{M}_{i,1}$ represents whether agent i shares $\mathbf{Y}_{i,1}$. Each digit of $\mathbf{M}_{1,2}$ represents whether agent 1 shares $\mathbf{Y}_{1,1}, \mathbf{U}_{1,1}$, and $\mathbf{Y}_{1,2}$. Each digit of $\mathbf{M}_{2,2}$ represents whether agent 2 shares $\mathbf{Y}_{2,1}$ and $\mathbf{Y}_{2,2}$.

For the control costs, we set $c_1 = k^2 \mathbf{U}_{1,1}^2, c_2 = 0, c_3 = \mathbf{X}_3^2$; for the communication costs, we set $\kappa_1 = (\sigma_0^2 + 1) \mathbb{1}[\mathbf{M}_{1,1} = 1 \text{ or } \mathbf{M}_{2,1} = 1], \kappa_2 = (\sigma_0^2 + 1) \mathbb{1}[\mathbf{M}_{2,2} \neq (0, 0) \text{ or } \mathbf{M}_{1,2} \notin \{(0, 0, 0), (0, 0, 1)\}]$. This communication cost means that if any information except $\mathbf{Y}_{1,2}$ is shared through additional sharing, it will incur a communication cost of $\sigma_0^2 + 1$. The information available to each agent at each timestep is defined as: $\forall i \in [2], \mathbf{I}_{i,1-} = \{\mathbf{Y}_{i,1}\}, \mathbf{I}_{i,1+} = \mathbf{I}_{i,1-} \cup \mathbf{Z}_1^a$, and $\mathbf{I}_{1,2-} = \mathbf{I}_{1,1+} \cup \{\mathbf{U}_{1,1}, \mathbf{Y}_{1,2}\}, \mathbf{I}_{2,2-} = \mathbf{I}_{2,1+} \cup \{\mathbf{Y}_{2,2}\}, \mathbf{I}_{i,2+} = \mathbf{I}_{i,2-} \cup \mathbf{Z}_2^a$. Then, we can verify that:

- $\mathcal{D}$ satisfies Assumption II.1.
- $\mathcal{D}$ has PN IS: If there is no additional sharing, then for any $i_1, i_2 \in [2], \mathbf{I}_{i_1,1-} \subseteq \mathbf{I}_{i_2,2-}$ if $i_1 = i_2$; otherwise, agent $(i_1, 1)$ does not influence agent $(i_2, 2)$.
- $\mathcal{D}$ satisfies Assumption III.2: The only zero input coefficient relevant to a later decision is $B_{2,1} = 0$, and $\mathbf{U}_{2,1}$ does not appear in any later information set.
- $\mathcal{D}$ does not satisfy Assumption III.4: It holds that $E_{2,2}B_{1,1} = 0$ but $B_{1,1} \neq 0$.

First, no team-optimal communication strategy shares any information other than $\mathbf{Y}_{1,2}$ through additional sharing, since otherwise it will incur a communication cost $\kappa_h \geq \sigma_0^2 + 1$, which is larger than the total cost when each agent i chooses $\mathbf{U}_{i,h} = 0, \forall h \in [2]$ and shares nothing via additional sharing.

Second, we analyze the optimal control strategy under two types of communication strategies.

1) Suppose the optimal communication strategy $g_{1:2}^{m,*}$ yields that agents share $\mathbf{Y}_{1,2}$ through additional sharing, i.e. $\mathbf{Z}_1^a = \emptyset, \mathbf{Z}_2^a = \{\mathbf{Y}_{1,2}\}$. Then, finding a team-optimal strategy of $\mathcal{D}$ can be reduced to finding an optimal strategy of the Decentralized LQG problem in Proposition B.2, where the optimal control strategy is nonlinear.

2) Suppose the optimal communication strategy $g_{1:2}^{m,*}$ yields that agents share nothing through additional sharing, i.e., $\mathbf{Z}_1^a = \mathbf{Z}_2^a = \emptyset$. Then, let $g_{1:2}^{a,*}$ be the optimal control strategy, and let $g_{1:2}^{m,'}$ be the communication strategy that additionally shares $\mathbf{Y}_{1,2}$. Since both $g_{1:2}^{m,'}$ and $g_{1:2}^{m,*}$ lead to no communication cost and $g_{1:2}^{m,'}$ enlarges $\mathbf{I}_{2,2+}$, it follows that $J_{\mathcal{D}}(g_{1:2}^{m,*}, g_{1:2}^{a,*}) \geq J_{\mathcal{D}}(g_{1:2}^{m,'}, g_{1:2}^{a,*})$, which means $(g_{1:2}^{m,'}, g_{1:2}^{a,*})$ is an optimal strategy. Therefore, from case 1) we know that $g_{1:2}^{a,*}$ is nonlinear, which completes the proof. $\square$

B-4 Proof of Theorem III.6

Proof. Fix any communication strategy $g_{1:H}^m \in \mathcal{G}_{1:H}^m$ and let $M_{1:H}$ be the communication actions chosen by $g_{1:H}^m \in \mathcal{G}_{1:H}^m$. Then, $\mathcal{D}$ becomes a decentralized LQG problem denoted by $\mathcal{D}(g_{1:H}^m)$, with the same system dynamics as $\mathcal{D}$ and the following information evolution: for any $h \in [H]$

- (a) The common information in $\mathcal{D}(g_{1:H}^m)$ evolves as $\bar{\mathcal{C}}_h = \cup_{t=1}^h \bar{\mathcal{Z}}_t$, and $\bar{\mathcal{Z}}_h = \bar{\chi}_h(\bar{\mathcal{P}}_{h-1}, \bar{\mathcal{U}}_{h-1}, \bar{\mathcal{Y}}_h)$ for some fixed projection function $\bar{\chi}_h$. Here, $\bar{\chi}_h$ is defined as

$$\bar{\chi}_h(\bar{\mathcal{P}}_{h-1}, \bar{\mathcal{U}}_{h-1}, \bar{\mathcal{Y}}_h) = \chi_h(\bar{\mathcal{P}}_{h-1}, \bar{\mathcal{U}}_{h-1}, \bar{\mathcal{Y}}_h) \cup \phi_h(M_h, \zeta_h(\bar{\mathcal{P}}_{h-1}, \bar{\mathcal{U}}_{h-1}, \bar{\mathcal{Y}}_h)).$$

Recall that χ_h, ζ_h, ϕ_h are defined in Assumption II.1. Such a $\bar{\chi}_h$ is a fixed projection function since $\chi_h, \{\phi_{i,h}(M_{i,h}, \cdot)\}_{i \in [n]}$ are fixed projection functions, and for any $P_{h-} \in \mathcal{P}_{h-}$, $\phi_h(M_h, P_{h-}) = \cup_{i=1}^n \phi_{i,h}(M_{i,h}, P_{i,h-})$.

- (b) The private information in $\mathcal{D}(g_{1:H}^m)$ evolves as follows. For any $i \in [n]$, $\bar{\mathcal{P}}_{i,h} = \bar{\zeta}_{i,h}(\bar{\mathcal{P}}_{i,h-1}, \bar{\mathcal{U}}_{i,h-1}, \bar{\mathcal{Y}}_{i,h})$ for some fixed projection function $\bar{\zeta}_{i,h}$ and thus $\bar{\mathcal{P}}_h = \bar{\zeta}_h(\bar{\mathcal{P}}_{h-1}, \bar{\mathcal{U}}_{h-1}, \bar{\mathcal{Y}}_h)$ for some fixed projection function $\bar{\zeta}_h$. Here, $\bar{\zeta}_{i,h}$ is defined as

$$\bar{\zeta}_{i,h}(\bar{\mathcal{P}}_{i,h-1}, \bar{\mathcal{U}}_{i,h-1}, \bar{\mathcal{Y}}_{i,h}) = \zeta_{i,h}(\bar{\mathcal{P}}_{i,h-1}, \bar{\mathcal{U}}_{i,h-1}, \bar{\mathcal{Y}}_{i,h}) \setminus \phi_{i,h}(M_{i,h}, \zeta_{i,h}(\bar{\mathcal{P}}_{i,h-1}, \bar{\mathcal{U}}_{i,h-1}, \bar{\mathcal{Y}}_{i,h})).$$

Recall that $\zeta_{i,h}, \phi_{i,h}$ are defined in Assumption II.1. Such a $\bar{\zeta}_{i,h}$ is a fixed projection function since $\zeta_{i,h}, \phi_{i,h}(M_{i,h}, \cdot)$ are fixed projection functions.

- (c) For each agent $i \in [n]$, $\bar{\mathcal{Y}}_{i,h} \in \bar{\mathcal{I}}_{i,h}$ for $h \in [H]$, and $\bar{\mathcal{I}}_{i,h} \subseteq \bar{\mathcal{I}}_{i,h+1}$ for $h \in [H-1]$. These properties still hold since additional sharing only moves some private information into common information but does not remove any information.

Then, we know that $\mathcal{D}(g_{1:H}^m)$ is a decentralized LQG problem with some information sharing due to both the baseline sharing and the communication actions $M_{1:H}$. It remains to prove partial nestedness. Use $\check{\cdot}$ for the baseline problem $\check{\mathcal{D}}$. Suppose that $\bar{\mathcal{U}}_{i_1, h_1}$ influences $\bar{\mathcal{I}}_{i_2, h_2}$, where $h_1 < h_2$. If $B_{i_1, h_1} = 0$, the action has no state effect, and Assumption III.2 excludes its later appearance as a shared label, a contradiction. Hence $B_{i_1, h_1} \neq 0$. Assumption III.4 gives some $j \neq i_1$ with $E_{j, h_1+1} B_{i_1, h_1} \neq 0$, so $\check{\mathcal{U}}_{i_1, h_1}$ influences $\check{\mathcal{Y}}_{j, h_1+1} \in \check{\mathcal{I}}_{j, h_1+1}$. Baseline partial nestedness yields $\check{\mathcal{I}}_{i_1, h_1} \subseteq \check{\mathcal{I}}_{j, h_1+1}$. Since $j \neq i_1$, every label private to agent i_1 can enter agent j 's information only through common information due to Assumption II.1, then it holds $\check{\mathcal{I}}_{i_1, h_1} \subseteq \check{\mathcal{C}}_{h_1+1}$. Moreover, every label in $\bar{\mathcal{I}}_{i_1, h_1} \setminus \check{\mathcal{I}}_{i_1, h_1}$ arrived through additional sharing and lies in $\bar{\mathcal{C}}_{h_1}$, and then it holds

$$\bar{\mathcal{I}}_{i_1, h_1} \subseteq \bar{\mathcal{C}}_{h_1+1} \subseteq \bar{\mathcal{C}}_{h_2} \subseteq \bar{\mathcal{I}}_{i_2, h_2}.$$

Thus $\mathcal{D}(g_{1:H}^m)$ is partially nested. Corollary A.2 supplies a linear team optimum for every fixed schedule. Since the open-loop schedule set is finite, a minimizing schedule exists, and its associated linear control strategy proves the final assertion. $\square$

C. Deferred Details of §IV

For the appendix derivations only, we use the following cost-free terminal convention after the last decision:

$$\widetilde{\mathcal{C}}_{H+1} := \widetilde{\mathcal{C}}_H \cup \{\widetilde{\mathcal{U}}_H, \widetilde{\mathcal{X}}_{H+1}\}, \quad \widetilde{\mathcal{P}}_{i, H+1} := \emptyset.$$

Equivalently, set $\tilde{\mathbf{Y}}_{H+1} := \tilde{\mathbf{X}}_{H+1}$, $\tilde{\mathbf{E}}_{H+1} := \mathbf{I}$, $\tilde{\mathbf{W}}_{1:n,H+1} := \mathbf{0}$, and $\tilde{\mathbf{Z}}_{H+1} := (\tilde{\mathbf{U}}_H, \tilde{\mathbf{Y}}_{H+1})$. Then $\tilde{\mathbf{S}}_{H+1} = \tilde{\mathbf{\Theta}}_{H+1} = \tilde{\mathbf{X}}_{H+1}$, $\tilde{\Sigma}_{H+1} = \mathbf{0}$, and $\mathbb{I}_{x,H+1} = \mathbf{I}$. This convention is only a device for writing the terminal-cost calculation in the same form as the preceding backward steps; it introduces no additional decision and changes neither the strategy space nor the cost.

C-1 Proof of Lemma IV.1

Proof. From the construction, the system dynamics and control cost are the same for both decentralized LQG problems, and it holds that $\tilde{\mathbf{I}}_{i,h} \subseteq \tilde{\mathbf{I}}_{i,h}$ for any $i \in [n], h \in [H]$. Therefore, the agents in $\tilde{\mathcal{D}}(g_{1:H}^m)$ have larger strategy spaces than those in $\mathcal{D}(g_{1:H}^m)$, which implies $\min_{\tilde{g}_{1:H} \in \tilde{\mathcal{G}}_{1:H}} J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}) \leq \min_{\tilde{g}_{1:H} \in \tilde{\mathcal{G}}_{1:H}} J_{\mathcal{D}(g_{1:H}^m)}(\tilde{g}_{1:H})$.

Now, we want to show that for any optimal $\tilde{g}_{1:H}^*$ of problem $\tilde{\mathcal{D}}(g_{1:H}^m)$, we can construct $\tilde{g}_{1:H}^* = \varphi(\tilde{g}_{1:H}^*, \mathcal{D}(g_{1:H}^m))$ to be a linear optimal strategy of $\mathcal{D}(g_{1:H}^m)$, and $J_{\mathcal{D}(g_{1:H}^m)}(\tilde{g}_{1:H}^*) = J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}^*)$. For any optimal linear strategy $\tilde{g}_{1:H}^* \in \arg\min_{\tilde{g}_{1:H} \in \tilde{\mathcal{G}}_{1:H}} J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H})$ of $\tilde{\mathcal{D}}(g_{1:H}^m)$, we recursively construct a strategy $\tilde{g}_{1:H}^* = \varphi(\tilde{g}_{1:H}^*, \mathcal{D}(g_{1:H}^m))$ of $\mathcal{D}(g_{1:H}^m)$. For $h = 1$ and any $i \in [n]$, from the construction of $\tilde{\mathcal{D}}(g_{1:H}^m)$, we know that $\tilde{\mathbf{I}}_{i,1} = \tilde{\mathbf{I}}_{i,1}$ always holds. Then we can define $\tilde{g}_{i,1}^*(\tilde{\mathbf{I}}_{i,1}) = \tilde{g}_{i,1}^*(\tilde{\mathbf{I}}_{i,1})$ for any $i \in [n]$, $\tilde{\mathbf{I}}_{i,1} \in \tilde{\mathbf{I}}_{i,1} = \tilde{\mathbf{I}}_{i,1}$. Then, for any $i \in [n]$, $\tilde{g}_{i,1}^*$ and $\tilde{g}_{i,1}^*$ output the same control action, namely, $\tilde{\mathbf{U}}_{i,1} = \tilde{\mathbf{U}}_{i,1}$, and $\tilde{g}_{i,1}^*$ is linear.

Now, for any $h \in [2 : H]$, assume that we have already constructed the linear strategy $\tilde{g}_{1:h-1}^*$ from $\tilde{g}_{1:h-1}^*$ such that for any $i \in [n]$ and $t < h$, $\tilde{g}_{i,t}^*$ and $\tilde{g}_{i,t}^*$ output the same actions. Now, we aim to construct $\tilde{g}_h^*$ from $\tilde{g}_h^*$. From the construction of problem $\tilde{\mathcal{D}}(g_{1:H}^m)$ from $\mathcal{D}(g_{1:H}^m)$, for any $i \in [n]$, $\tilde{\mathbf{I}}_{i,h} \subseteq \tilde{\mathbf{I}}_{i,h}$ and the only additional variables are the actions $\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h} \setminus \tilde{\mathbf{I}}_{i,h}$, where $j \in [n]$ and $t < h$. Furthermore, if $\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h} \setminus \tilde{\mathbf{I}}_{i,h}$, then $\tilde{\mathbf{B}}_{j,t} \neq \mathbf{0}$ and $\tilde{\mathbf{I}}_{j,t} \subseteq \tilde{\mathbf{C}}_h$. Since $\tilde{g}_{1:h-1}^*$ outputs the same actions as $\tilde{g}_{1:h-1}^*$, we can write $\tilde{\mathbf{U}}_{j,t} = \tilde{\mathbf{U}}_{j,t} = \tilde{g}_{j,t}^*(\tilde{\mathbf{I}}_{j,t})$, which means that we can obtain $\tilde{\mathbf{U}}_{j,t}$ from $\tilde{g}_{j,t}^*$ and $\tilde{\mathbf{I}}_{j,t}$. Since $\tilde{g}_{j,t}^*$ is linear, we write $\tilde{g}_{j,t}^*(\tilde{\mathbf{I}}_{j,t}) = \sum_{\tilde{\mathbf{Y}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}} \tilde{\mathbf{G}}_{j,t}^{k,s,y} \tilde{\mathbf{Y}}_{k,s} + \sum_{\tilde{\mathbf{U}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}} \tilde{\mathbf{G}}_{j,t}^{k,s,u} \tilde{\mathbf{U}}_{k,s}$ for any $j \in [n]$ and $t < h$. Here, the sum $\sum_{\tilde{\mathbf{Y}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}}$ ranges over all $k \in [n]$ and $s \in [H]$ such that $\tilde{\mathbf{Y}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}$, and the other sums are interpreted similarly. Meanwhile, from the linearity of $\tilde{g}_{i,h}^*$ for any $i \in [n]$, we can write $\tilde{g}_{i,h}^*(\tilde{\mathbf{I}}_{i,h}) = \sum_{\tilde{\mathbf{Y}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,y} \tilde{\mathbf{Y}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \tilde{\mathbf{U}}_{j,t}$ for some matrices $\tilde{\mathbf{G}}_{i,h}^{j,t,y}, \tilde{\mathbf{G}}_{i,h}^{j,t,u}$. Then, we have

$$\begin{aligned} \tilde{g}_{i,h}^*(\tilde{\mathbf{I}}_{i,h}) &= \sum_{\tilde{\mathbf{Y}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,y} \tilde{\mathbf{Y}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \tilde{\mathbf{U}}_{j,t} = \sum_{\tilde{\mathbf{Y}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,y} \tilde{\mathbf{Y}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \tilde{\mathbf{U}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h} \setminus \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \tilde{\mathbf{U}}_{j,t} \\ &= \sum_{\tilde{\mathbf{Y}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,y} \tilde{\mathbf{Y}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \tilde{\mathbf{U}}_{j,t} + \sum_{\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h} \setminus \tilde{\mathbf{I}}_{i,h}} \tilde{\mathbf{G}}_{i,h}^{j,t,u} \left(\sum_{\tilde{\mathbf{Y}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}} \tilde{\mathbf{G}}_{j,t}^{k,s,y} \tilde{\mathbf{Y}}_{k,s} + \sum_{\tilde{\mathbf{U}}_{k,s} \in \tilde{\mathbf{I}}_{j,t}} \tilde{\mathbf{G}}_{j,t}^{k,s,u} \tilde{\mathbf{U}}_{k,s} \right). \end{aligned}$$

The second equality in the first line holds because $\tilde{g}_{1:h-1}^*$ and $\tilde{g}_{1:h-1}^*$ output the same control actions, and the system dynamics of both problems are the same, so the observations are the same before agents take actions at timestep h . The equality between the first and the second lines follows from $\tilde{\mathbf{U}}_{j,t} = \tilde{g}_{j,t}^*(\tilde{\mathbf{I}}_{j,t})$ and substitution of the explicit form of $\tilde{g}_{j,t}^*$. Also, if $\tilde{\mathbf{U}}_{j,t} \in \tilde{\mathbf{I}}_{i,h} \setminus \tilde{\mathbf{I}}_{i,h}$, then $\tilde{\mathbf{I}}_{j,t} \subseteq \tilde{\mathbf{I}}_{i,h}$,

and any $\bar{Y}_{k,s} \in \bar{I}_{j,t}, \bar{U}_{k,s} \in \bar{I}_{j,t}$ will also be included in $\bar{I}_{i,h}$. Therefore, we have

$$\begin{aligned}\bar{g}_{i,h}^*(\bar{I}_{i,h}) &= \sum_{\bar{Y}_{j,t} \in \bar{I}_{i,h}} \left(\bar{G}_{i,h}^{j,t,y} + \sum_{\bar{U}_{k,s} \in \bar{I}_{i,h} \setminus \bar{I}_{j,t}} \bar{G}_{i,h}^{k,s,u} \mathbb{1}[\bar{Y}_{j,t} \in \bar{I}_{k,s}] \bar{G}_{k,s}^{j,t,y} \right) \bar{Y}_{j,t} \\ &\quad + \sum_{\bar{U}_{j,t} \in \bar{I}_{i,h}} \left(\bar{G}_{i,h}^{j,t,u} + \sum_{\bar{U}_{k,s} \in \bar{I}_{i,h} \setminus \bar{I}_{j,t}} \bar{G}_{i,h}^{k,s,u} \mathbb{1}[\bar{U}_{j,t} \in \bar{I}_{k,s}] \bar{G}_{k,s}^{j,t,u} \right) \bar{U}_{j,t}.\end{aligned}$$

Note that in the equation above, if $\bar{Y}_{j,t} \notin \bar{I}_{k,s}$, then it means that $\bar{U}_{k,s}$ is not based on $\bar{Y}_{j,t}$, and we can just assign $\bar{G}_{k,s}^{j,t,y} = 0$. Also, if $\bar{U}_{j,t} \notin \bar{I}_{k,s}$, then we assign $\bar{G}_{k,s}^{j,t,u} = 0$. Therefore, we can construct $\bar{g}_{i,h}^*$ for any $i \in [n]$ as

$$\begin{aligned}\bar{g}_{i,h}^*(\bar{I}_{i,h}) &= \sum_{\bar{Y}_{j,t} \in \bar{I}_{i,h}} \bar{G}_{i,h}^{j,t,y} \bar{Y}_{j,t} + \sum_{\bar{U}_{j,t} \in \bar{I}_{i,h}} \bar{G}_{i,h}^{j,t,u} \bar{U}_{j,t}, \\ \forall j \in [n], t \leq h, \bar{G}_{i,h}^{j,t,y} &:= \bar{G}_{i,h}^{j,t,y} + \sum_{\bar{U}_{k,s} \in \bar{I}_{i,h} \setminus \bar{I}_{j,t}} \bar{G}_{i,h}^{k,s,u} \mathbb{1}[\bar{Y}_{j,t} \in \bar{I}_{k,s}] \bar{G}_{k,s}^{j,t,y}, \\ \forall j \in [n], t \leq h, \bar{G}_{i,h}^{j,t,u} &:= \bar{G}_{i,h}^{j,t,u} + \sum_{\bar{U}_{k,s} \in \bar{I}_{i,h} \setminus \bar{I}_{j,t}} \bar{G}_{i,h}^{k,s,u} \mathbb{1}[\bar{U}_{j,t} \in \bar{I}_{k,s}] \bar{G}_{k,s}^{j,t,u}.\end{aligned}$$

Then, $\bar{g}_h^*$ is linear and outputs the same actions as $\bar{g}_h^*$. Recursively, we can construct $\bar{g}_h^*$ from $h = 1$ to H based on $\bar{g}_{1:H}^*$ and the problem $\mathcal{D}(g_{1:H}^m)$, and we denote this construction by a function φ , so that $\bar{g}_{1:H}^* = \varphi(\bar{g}_{1:H}^*, \mathcal{D}(g_{1:H}^m))$. Since the dynamics and control costs are the same in the two problems, and $\bar{g}_{1:H}^*$ and $\bar{g}_{1:H}^*$ output the same actions, $J_{\bar{\mathcal{D}}(g_{1:H}^m)}(\bar{g}_{1:H}^*) = J_{\mathcal{D}(g_{1:H}^m)}(\bar{g}_{1:H}^*)$. Also, we know that

$$J_{\bar{\mathcal{D}}(g_{1:H}^m)}(\bar{g}_{1:H}^*) = \min_{\bar{g}_{1:H} \in \bar{\mathcal{G}}_{1:H}} J_{\bar{\mathcal{D}}(g_{1:H}^m)}(\bar{g}_{1:H}) \leq \min_{\bar{g}_{1:H} \in \bar{\mathcal{G}}_{1:H}} J_{\mathcal{D}(g_{1:H}^m)}(\bar{g}_{1:H}) \leq J_{\mathcal{D}(g_{1:H}^m)}(\bar{g}_{1:H}^*).$$

Therefore, $\bar{g}_{1:H}^*$ is a linear optimal strategy of $\mathcal{D}(g_{1:H}^m)$. $\square$

C-2 Proof of Theorem IV.2

Proof. The proof consists of three **Parts**: **Part 1**: $\bar{\mathcal{D}}(g_{1:H}^m)$ is (strictly) PN; **Part 2**: $\bar{\mathcal{D}}(g_{1:H}^m)$ has the information evolution rules; **Part 3**: $\bar{\mathcal{D}}(g_{1:H}^m)$ satisfies the SI-CIB condition.

Part 1: We aim to prove that $\bar{\mathcal{D}}(g_{1:H}^m)$ is (strictly) partially-nested.

Fix $i_1, i_2 \in [n]$ and $h_1 < h_2$ in $[H]$, and suppose that $\bar{U}_{i_1, h_1}$ influences $\bar{I}_{i_2, h_2}$. If $\bar{B}_{i_1, h_1} = 0$, the action has no state effect, while Assumption III.2 prevents it from appearing in any later information set, contradicting the assumed influence. Hence $\bar{B}_{i_1, h_1} = \bar{B}_{i_1, h_1} \neq 0$. Assumption III.4 gives some $i \neq i_1$ such that $\bar{U}_{i, h_1}$ influences $\bar{Y}_{i, h_1+1}$. Since $h_1 + 1 \leq h_2$ and information is retained, $\bar{U}_{i, h_1}$ influences $\bar{I}_{i, h_2}$. Theorem III.6 therefore gives $\bar{I}_{i, h_1} \subseteq \bar{I}_{i, h_2}$. Since $i \neq i_1$, every information of agent i_1 can be known by agent i 's information only if it lies in common information from Assumption II.1, so it holds $\bar{I}_{i_1, h_1} \subseteq \bar{C}_{h_2}$. Moreover, from Equation (IV.1), every information in $\bar{I}_{i_1, h_1} \setminus \bar{I}_{i_1, h_1}$ is common by construction. Thus, we have $\bar{I}_{i_1, h_1} \subseteq \bar{C}_{h_2} \subseteq \bar{I}_{i_2, h_2}$, and further have $\bar{U}_{i_1, h_1} \in \bar{C}_{h_2} \subseteq \bar{I}_{i_2, h_2}$ and then prove that $\bar{\mathcal{D}}(g_{1:H}^m)$ is strictly partially-nested.

Part 2: We want to prove $\widetilde{\mathcal{D}}(g_{1:H}^m)$ has the information evolution rules. From Theorem III.6, we know that the problem $\mathcal{D}(g_{1:H}^m)$ has the information evolution rules with some projection functions $\{\bar{X}_h\}_{h \in [H]}, \{\bar{C}_{i,h}\}_{i \in [n], h \in [H]}$.

For any $i \in [n], h \in [H]$, because the corresponding label sets in $\widetilde{\mathcal{Z}}_h$ and $\widetilde{\mathcal{P}}_{i,h}$ are fixed, it suffices to prove that $\widetilde{\mathcal{Z}}_h \subseteq \widetilde{\mathcal{P}}_{h-1} \cup \{\widetilde{U}_{h-1}, \widetilde{Y}_h\}$, $\widetilde{\mathcal{P}}_{i,h} \subseteq \widetilde{\mathcal{P}}_{i,h-1} \cup \{\widetilde{U}_{i,h-1}, \widetilde{Y}_{i,h}\}$.

Firstly, we show that the common information of $\widetilde{\mathcal{D}}(g_{1:H}^m)$ has the evolution rule. For any $h \in [H]$, consider the $\widetilde{\mathcal{Z}}_h$. If any $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{Z}}_h, i \in [n], t \leq h$, then $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{C}}_h$ and $\widetilde{Y}_{i,t} \notin \widetilde{\mathcal{C}}_{h-1}$. Then $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{C}}_h$ and $\widetilde{Y}_{i,t} \notin \widetilde{\mathcal{C}}_{h-1}$ since the construction does not change any observation in the information. Therefore, $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{Z}}_h \subseteq \widetilde{\mathcal{P}}_{h-1} \cup \{\widetilde{U}_{h-1}, \widetilde{Y}_h\}$ due to the evolution rule of problem $\mathcal{D}(g_{1:H}^m)$; since $\widetilde{Y}_{i,t}$ is an observation label, it belongs to $\widetilde{\mathcal{P}}_{h-1} \cup \{\widetilde{Y}_h\}$, and then $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{P}}_{h-1} \cup \{\widetilde{Y}_h\}$. If any $\widetilde{U}_{i,t} \in \widetilde{\mathcal{Z}}_h$, then $\bar{B}_{i,t} \neq \mathbf{0}$ from Assumption III.2 and construction of $\widetilde{\mathcal{D}}(g_{1:H}^m)$. If $t < h-1$, we know $\widetilde{U}_{i,t}$ influences $\bar{X}_{t+1}$ and thus influences $\bar{Y}_{j,t+1}$ for some $j \neq i$ due to Assumption III.4. Partial nestedness gives $\bar{I}_{i,t} \subseteq \bar{I}_{j,t+1}$. Due to $j \neq i$ and Assumption II.1, it holds that $\bar{I}_{i,t} \subseteq \bar{\mathcal{C}}_{t+1}$, and the strict expansion therefore places $\widetilde{U}_{i,t}$ in $\widetilde{\mathcal{C}}_{t+1}$. From $t < h-1$ we know $\widetilde{U}_{i,t} \notin \widetilde{\mathcal{Z}}_h$, which leads to a contradiction. Hence, if $\widetilde{U}_{i,t} \in \widetilde{\mathcal{Z}}_h$, it is only possible that $t = h-1$. In conclusion, $\widetilde{\mathcal{Z}}_h \subseteq \widetilde{\mathcal{P}}_{h-1} \cup \{\widetilde{U}_{h-1}, \widetilde{Y}_h\}$.

Secondly, we show that the private information of $\widetilde{\mathcal{D}}(g_{1:H}^m)$ satisfies the evolution rule. For any $i \in [n], h \in [H]$, consider the $\widetilde{\mathcal{P}}_{i,h}$. If any $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{P}}_{i,h}, t \leq h$, then $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{P}}_{i,h}$ since the construction does not change any observation in the information. Therefore, $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{P}}_{i,h-1} \cup \{\widetilde{Y}_{i,h}\}$ due to the evolution rule of problem $\mathcal{D}(g_{1:H}^m)$, and then $\widetilde{Y}_{i,t} \in \widetilde{\mathcal{P}}_{i,h-1} \cup \{\widetilde{Y}_{i,h}\}$. If any $\widetilde{U}_{i,t} \in \widetilde{\mathcal{P}}_{i,h}$, then $\bar{B}_{i,t} \neq \mathbf{0}$ by Assumption III.2. Then, we know $\widetilde{U}_{i,t}$ influences $\bar{X}_{t+1}$ and thus influences $\bar{Y}_{j,t+1}$ for some $j \neq i$ due to Assumption III.4. Partial nestedness gives $\bar{I}_{i,t} \subseteq \bar{I}_{j,t+1}$. Due to $j \neq i$ and Assumption II.1, it holds that $\bar{I}_{i,t} \subseteq \bar{\mathcal{C}}_{t+1}$, and the strict expansion therefore places $\widetilde{U}_{i,t}$ in $\widetilde{\mathcal{C}}_{t+1}$. Thus, we have $\widetilde{U}_{i,t} \notin \widetilde{\mathcal{P}}_{i,h}$, which leads to a contradiction. In conclusion, $\widetilde{\mathcal{P}}_{i,h} \subseteq \widetilde{\mathcal{P}}_{i,h-1} \cup \{\widetilde{Y}_{i,h}\}$.

Part 3: $\widetilde{\mathcal{D}}(g_{1:H}^m)$ satisfies the SI-CIB condition. We show this by induction. If $h = 1$, then the belief $\bar{B}_1$ does not depend on any strategy, and the SI-CIB condition holds automatically.

For each $h \geq 2$, the space $\widetilde{\mathcal{S}}_h = \widetilde{\mathcal{X}} \times \widetilde{\mathcal{P}}_{1,h} \times \cdots \times \widetilde{\mathcal{P}}_{n,h}$ is finite-dimensional because the collection of random variables comprising each $\widetilde{\mathcal{P}}_{i,h}, i \in [n]$, is fixed. For any fixed Borel set $\mathcal{Q}_h \subseteq \widetilde{\mathcal{S}}_h$ and any fixed common-information realization $\bar{C}_h \in \bar{\mathcal{C}}_h$ that can be reached by two control strategies $\widetilde{g}_{1:h-1}$ and $\widetilde{g}'_{1:h-1}$, define

$$\mathbb{P}_1 = \mathbb{P}(\bar{S}_h \in \mathcal{Q}_h | \bar{C}_h = \bar{C}_h, \widetilde{g}_{1:h-1}), \quad \mathbb{P}_2 = \mathbb{P}(\bar{S}_h \in \mathcal{Q}_h | \bar{C}_h = \bar{C}_h, \widetilde{g}'_{1:h-1}).$$

It suffices to prove that $\mathbb{P}_1 = \mathbb{P}_2$.

We first suppose that $\widetilde{g}_{1:h-1}$ and $\widetilde{g}'_{1:h-1}$ differ only at one pair (i_1, h_1) , where $i_1 \in [n]$ and $h_1 < h$.

If $\bar{B}_{i_1, h_1} = \bar{B}_{i_1, h_1} \neq \mathbf{0}$, then $\widetilde{U}_{i_1, h_1}$ influences $\bar{X}_{h_1+1}$ in $\mathcal{D}(g_{1:H}^m)$. By Assumption III.4, there exists $i_2 \neq i_1$ such that $\widetilde{U}_{i_1, h_1}$ influences $\bar{Y}_{i_2, h_1+1}$. Since $\mathcal{D}(g_{1:H}^m)$ has a PN IS, we have $\bar{I}_{i_1, h_1} \subseteq \bar{I}_{i_2, h_1+1}$. Due to $i_2 \neq i_1$ and Assumption II.1, we have $\bar{I}_{i_1, h_1} \subseteq \bar{\mathcal{C}}_{h_1+1}$. By the construction of $\widetilde{\mathcal{D}}(g_{1:H}^m)$,

$$\widetilde{U}_{i_1, h_1} \in \widetilde{\mathcal{C}}_{h_1+1} \subseteq \widetilde{\mathcal{C}}_h.$$

Moreover,

$$\bar{I}_{i_1, h_1} \setminus \bar{I}_{i_1, h_1} \subseteq \bar{\mathcal{C}}_{h_1}, \quad \bar{I}_{i_1, h_1} \subseteq \bar{\mathcal{C}}_{h_1+1} \subseteq \bar{\mathcal{C}}_{h_1+1} \subseteq \bar{\mathcal{C}}_h,$$

and hence $\bar{I}_{i_1, h_1} \subseteq \bar{\mathcal{C}}_h$. Conditioning on $\bar{C}_h = \bar{C}_h$ therefore fixes both the information at which $\widetilde{g}_{i_1, h_1}$ and $\widetilde{g}'_{i_1, h_1}$ are evaluated and the resulting action. Since both strategies reach the same common-

information realization, they produce the same recorded action from the same recorded information. With these quantities fixed, the system dynamics and information evolution are identical under the two strategies. Therefore, $\mathbb{P}_1 = \mathbb{P}_2$.

If $\tilde{B}_{i_1, h_1} = \bar{B}_{i_1, h_1} = \mathbf{0}$, then $\tilde{U}_{i_1, h_1}$ does not affect the system dynamics. Moreover, Assumption III.2 prevents $\tilde{U}_{i_1, h_1}$ from appearing in any later information set. Since the strict expansion only adds actions with nonzero input coefficients, $\tilde{U}_{i_1, h_1}$ does not appear in $\tilde{C}_{h'}$ or $\tilde{P}_{i, h'}$ for any $h' \in [H]$ with $h' > h_1$ and any $i \in [n]$. Thus, changing $\tilde{g}_{i_1, h_1}$ affects neither the state nor any later common or private information. It follows again that $\mathbb{P}_1 = \mathbb{P}_2$.

Finally, consider any two arbitrary $\tilde{g}_{1:h-1}$ and $\tilde{g}'_{1:h-1}$ that can both reach a realization $\tilde{C}_h$, we connect $\tilde{g}_{1:h-1}$ and $\tilde{g}'_{1:h-1}$ by finitely many intermediate strategies, i.e., $\tilde{g}_{1:h-1}^0 := \tilde{g}_{1:h-1}, \tilde{g}_{1:h-1}^1, \dots, \tilde{g}_{1:h-1}^K := \tilde{g}'_{1:h-1}$, replacing the control strategy of one agent at a time. We will show that all of the intermediate strategies can reach $\tilde{C}_h$ by induction. The claim holds for $k = 0$ trivially. Now suppose $\tilde{g}_{1:h-1}^{k-1}$ can reach $\tilde{C}_h$ and $\tilde{g}_{1:h-1}^k$ is constructed by replacing $\tilde{g}_{i_1, h_1}$ by $\tilde{g}'_{i_1, h_1}$ for some $i_1 \in [n], h_1 < h$. If $\tilde{B}_{i_1, h_1} \neq \mathbf{0}$, then from the results above, we know that $\tilde{I}_{i_1, h_1} \cup \{\tilde{U}_{i_1, h_1}\} \subseteq \tilde{C}_h$. Let $\tilde{I}_{i_1, h_1}$ and $\tilde{U}_{i_1, h_1}$ be the realization in $\tilde{C}_h$, then it holds that $\tilde{g}_{i_1, h_1}(\tilde{I}_{i_1, h_1}) = \tilde{g}'_{i_1, h_1}(\tilde{I}_{i_1, h_1}) = \tilde{U}_{i_1, h_1}$ since both $\tilde{g}_{1:h-1}$ and $\tilde{g}'_{1:h-1}$ can reach $\tilde{C}_h$, and thus $\tilde{g}_{1:h-1}^k$ can also reach $\tilde{C}_h$; Otherwise, $\tilde{U}_{i_1, h_1}$ can neither influence the state nor the information of other agents from the results above, so $\tilde{g}_{1:h-1}^k$ can also reach $\tilde{C}_h$. Therefore, it suffices to only consider the case that the two strategies only differ in the strategy of one agent, which completes the proof. $\square$

C-3 Proof of Lemma IV.3

The proof follows the arguments in Appendices B and C of [11], with the minor modification needed for possibly singular Gaussian distributions.

Proof. We first prove Gaussianity:

At timestep $h = 1$, the state, private information, and common information are linear functions of the primitive Gaussian variables. Hence, $(\tilde{S}_1, \tilde{C}_1)$ is jointly Gaussian, possibly singular, and the conditional belief $\tilde{B}_1$ is Gaussian.

Fix $h \geq 2$ and suppose that $\tilde{g}_{1:h-1}$ is affine. Due to the linearity of the system dynamics, the state $\tilde{X}_h$, private information $\tilde{P}_h$, and common information $\tilde{C}_h$ are jointly Gaussian. Consequently, for every common information realization $\tilde{C}_h$ that can be reached under $\tilde{g}_{1:h-1}$, the conditional distribution of $\tilde{S}_h$ given $\tilde{C}_h = \tilde{C}_h$ is Gaussian. This conclusion remains valid when the joint distribution is singular: the conditional-Gaussian formula, with the Moore-Penrose pseudo-inverse in place of the inverse, gives a Gaussian conditional distribution whose covariance does not depend on $\tilde{C}_h$.

If $\tilde{g}_{1:h-1}$ is not affine, then for every common information realization $\tilde{C}_h$ that can be reached under $\tilde{g}_{1:h-1}$, there exist realizations $\tilde{Y}_{1:h}, \tilde{U}_{1:h-1}$ such that $(\tilde{Y}_{1:h}, \tilde{U}_{1:h-1}, \tilde{C}_h)$ is reachable under $\tilde{g}_{1:h-1}$. Then, we construct affine strategies $\tilde{g}'_{1:h-1} \in \tilde{\mathcal{G}}_{1:h-1}$ as: $\forall i \in [n], t < h, \forall \tilde{I}_{i,t} \in \tilde{\mathcal{I}}_{i,t}, \tilde{g}'_{i,t}(\tilde{I}_{i,t}) \equiv \tilde{U}_{i,t}$. Then, $(\tilde{Y}_{1:h}, \tilde{U}_{1:h-1}, \tilde{C}_h)$ is also reachable under $\tilde{g}'_{1:h-1}$. Since the CIB belief $\tilde{B}_h$ is strategy independent, its value at $\tilde{C}_h$ is the same under the original and comparison strategies. Thus, $\tilde{B}_h$ is Gaussian under any admissible strategy $\tilde{g}_{1:h-1}$.

We next prove that $\tilde{\Theta}_h$ and $\tilde{\Sigma}_h$ satisfy the evolution rules. The system dynamics and Equation (IV.2) give the usual fixed CIB-belief update from the preceding belief, the past control laws, and

$\tilde{\mathbf{Z}}_h$. The SI-CIB condition removes the dependence on the control laws, so

$$(\tilde{\boldsymbol{\Theta}}_h, \tilde{\Sigma}_h) = \tilde{F}_{h-1}((\tilde{\boldsymbol{\Theta}}_{h-1}, \tilde{\Sigma}_{h-1}), \tilde{\mathbf{Z}}_h),$$

where $\tilde{F}_{h-1}$ is a fixed transformation that does not depend on the control strategies.

To determine the covariance component, again take affine control laws. The variables $\tilde{\mathbf{S}}_h$ and $\tilde{\mathbf{C}}_h$ are jointly Gaussian, and the conditional covariance of jointly Gaussian variables does not depend on the realized conditioning value, including in the singular case. Similarly to the argument above, for any strategy, we can replace it by a constant affine strategy. Constant control values change only the conditional mean, not the conditional covariance. Hence,

$$\tilde{\Sigma}_h = \tilde{\Psi}_h^2(\tilde{\Sigma}_{h-1})$$

for a fixed deterministic function $\tilde{\Psi}_h^2$.

Under the same affine control laws, $\mathbb{E}[\tilde{\mathbf{S}}_h | \tilde{\mathbf{C}}_h]$ is an affine function of $\tilde{\mathbf{C}}_h$, and $\mathbb{E}[\tilde{\mathbf{S}}_{h-1} | \tilde{\mathbf{C}}_{h-1}]$ is an affine function of $\tilde{\mathbf{C}}_{h-1}$. Combining these facts with the preceding fixed belief update gives

$$\tilde{\boldsymbol{\Theta}}_h = \tilde{\Psi}_h^1(\tilde{\boldsymbol{\Theta}}_{h-1}, \tilde{\mathbf{Z}}_h)$$

for a fixed affine function $\tilde{\Psi}_h^1$. Since the dynamics and observation equations are homogeneous and the relevant known controls are included in $\tilde{\mathbf{Z}}_h$, the constant term is zero, so $\tilde{\Psi}_h^1$ is linear. Neither update depends on the control strategies, which completes the proof. $\square$

C-4 Proof of Theorem IV.4

Proof. Part 1: For $h \in [H-1]$, given $(\tilde{\boldsymbol{\Theta}}_h, \tilde{\gamma}_h)$, draw $\tilde{\mathbf{S}}_h \sim \mathcal{N}(\tilde{\boldsymbol{\Theta}}_h, \tilde{\Sigma}_h)$ and the independent Gaussian noises and then apply the dynamics, $\tilde{\chi}_{h+1}$, and Lemma IV.3. The law of $\tilde{\boldsymbol{\Theta}}_{h+1}$ therefore depends only on $(\tilde{\boldsymbol{\Theta}}_h, \tilde{\gamma}_h)$. At $h = H$, the final-state cost has already been included in $c_H^\dagger$. Hence, $\tilde{\mathcal{D}}^\dagger(g_{1:H}^m)$ is Markov.

Part 2: We aim to prove that for any strategy $\tilde{g}_{1:H} \in \tilde{\mathcal{G}}_{1:H}$, $J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}) = J_{\tilde{\mathcal{D}}^\dagger(g_{1:H}^m)}(g_{1:H}^\dagger)$, where $g_h^\dagger =$

$\varsigma_h(\tilde{g}_h), \forall h \in [H]$. For every $h \in [H-1]$,

$$\mathbb{E}[c_h^\dagger | g_{1:H}^\dagger] = \mathbb{E}\left[\mathbb{E}[c_h^\dagger(\tilde{\Theta}_h, g_h^\dagger(\tilde{C}_h)) | \tilde{C}_h, \tilde{\Theta}_h = \tilde{\Psi}_h^3(\tilde{C}_h)] | g_{1:h}^\dagger\right] \quad (C.1)$$

$$= \mathbb{E}\left[\mathbb{E}\left[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + [\tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h})^\top \quad \cdots \quad \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h})^\top] \tilde{Q}_h^2 \begin{bmatrix} \tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h}) \\ \cdots \\ \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h}) \end{bmatrix} \middle| \tilde{C}_h, \tilde{\gamma}_h = g_h^\dagger(\tilde{C}_h)\right] | g_{1:h}^\dagger\right] \quad (C.2)$$

$$= \mathbb{E}\left[\mathbb{E}\left[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + [\tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h})^\top \quad \cdots \quad \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h})^\top] \tilde{Q}_h^2 \begin{bmatrix} \tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h}) \\ \cdots \\ \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h}) \end{bmatrix} \middle| \tilde{C}_h, \tilde{\gamma}_h(\cdot) = \tilde{g}_h(\tilde{C}_h, \cdot)\right] | g_{1:h}^\dagger\right] \quad (C.3)$$

$$= \mathbb{E}\left[\mathbb{E}\left[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + [\tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h})^\top \quad \cdots \quad \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h})^\top] \tilde{Q}_h^2 \begin{bmatrix} \tilde{\gamma}_{1,h}(\tilde{\mathbf{P}}_{1,h}) \\ \cdots \\ \tilde{\gamma}_{n,h}(\tilde{\mathbf{P}}_{n,h}) \end{bmatrix} \middle| \tilde{C}_h, \tilde{\gamma}_h(\cdot) = \tilde{g}_h(\tilde{C}_h, \cdot)\right] | \tilde{g}_{1:h}\right] \quad (C.4)$$

$$= \mathbb{E}[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + \tilde{\mathbf{U}}_h^\top \tilde{Q}_h^2 \tilde{\mathbf{U}}_h | \tilde{g}_{1:H}]. \quad (C.5)$$

Equation (C.1) and Equation (C.5) are due to the tower property; Equation (C.2) is from the definition of $c_h^\dagger$; Equation (C.3) is from the definition of ς_h ; Equation (C.4) is because $\tilde{C}_h$ has the same distribution under both strategies $\tilde{g}_{1:h}$ and $g_{1:h}^\dagger$, i.e., for any Borel set $\tilde{\mathcal{C}}_h \subseteq \tilde{\mathcal{C}}_h$, it holds that

$$\begin{aligned} \mathbb{P}(\tilde{C}_h \in \tilde{\mathcal{C}}_h | g_{1:h}^\dagger) &= \mathbb{E}[\mathbb{1}[\tilde{C}_h \in \tilde{\mathcal{C}}_h] | g_{1:h}^\dagger] = \mathbb{E}\left[\mathbb{E}[\mathbb{1}[\tilde{C}_h \in \tilde{\mathcal{C}}_h] | \tilde{C}_{h-1}, \tilde{\gamma}_{h-1} = g_{h-1}^\dagger(\tilde{C}_{h-1})] | g_{1:h-1}^\dagger\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\cdots \mathbb{E}\left[\mathbb{E}[\mathbb{1}[\tilde{C}_h \in \tilde{\mathcal{C}}_h] | \tilde{C}_{h-1}, \tilde{\gamma}_{h-1} = g_{h-1}^\dagger(\tilde{C}_{h-1})] | \tilde{C}_{h-2}, \tilde{\gamma}_{h-2} = g_{h-2}^\dagger(\tilde{C}_{h-2})\right] \cdots | \tilde{C}_1, \tilde{\gamma}_1 = g_1^\dagger(\tilde{C}_1)\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\cdots \mathbb{E}\left[\mathbb{E}[\mathbb{1}[\tilde{C}_h \in \tilde{\mathcal{C}}_h] | \tilde{C}_{h-1}, \tilde{\gamma}_{h-1}(\cdot) = \tilde{g}_{h-1}(\tilde{C}_{h-1}, \cdot)] | \tilde{C}_{h-2}, \tilde{\gamma}_{h-2}(\cdot) = \tilde{g}_{h-2}(\tilde{C}_{h-2}, \cdot)\right] \cdots | \tilde{C}_1, \tilde{\gamma}_1(\cdot) = \tilde{g}_1(\tilde{C}_1, \cdot)\right]\right] \\ &= \mathbb{E}[\mathbb{1}[\tilde{C}_h \in \tilde{\mathcal{C}}_h] | \tilde{g}_{1:h}] = \mathbb{P}(\tilde{C}_h \in \tilde{\mathcal{C}}_h | \tilde{g}_{1:h}). \end{aligned}$$

The same calculation at $h = H$, now using the definition of $c_H^\dagger$, gives

$$\mathbb{E}[c_H^\dagger | g_{1:H}^\dagger] = \mathbb{E}[\tilde{\mathbf{X}}_H^\top \tilde{Q}_H^1 \tilde{\mathbf{X}}_H + \tilde{\mathbf{U}}_H^\top \tilde{Q}_H^2 \tilde{\mathbf{U}}_H + \tilde{\mathbf{X}}_{H+1}^\top \tilde{Q}_{H+1}^1 \tilde{\mathbf{X}}_{H+1} | \tilde{g}_{1:H}].$$

Therefore,

$$J_{\tilde{\mathcal{D}}^\dagger(g_{1:H}^m)}(g_{1:H}^\dagger) = \mathbb{E}\left[\sum_{h=1}^H c_h^\dagger | g_{1:H}^\dagger\right] = J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\tilde{g}_{1:H}),$$

which completes the proof. $\square$

C-5 Proof of Lemma IV.5

Proof. For convenience, given any matrices $\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}$ and $\{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n], h \in [H]}$ with appropriate dimensions, we denote by $\tilde{g}_{1:H} = \Lambda(\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n], h \in [H]})$ the linear strategy generated by $\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n], h \in [H]}$, namely for any $i \in [n], h \in [H]$, $\tilde{g}_{i,h}(\tilde{C}_h, \tilde{\mathbf{P}}_{i,h}) = \tilde{\mathcal{E}}_{i,h} \tilde{C}_h + \tilde{\mathcal{F}}_{i,h} \tilde{\mathbf{P}}_{i,h}$. Theorem IV.2 and Corollary A.2 give an optimal linear strategy $\tilde{g}_{1:H}^* = \Lambda(\{\tilde{\mathcal{E}}_{i,h}^*\}, \{\tilde{\mathcal{F}}_{i,h}^*\})$. If we

fix $\{\tilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]}$, then $\{\tilde{\mathcal{E}}_{i,h}^*\}_{i \in [n], h \in [H]}$ minimizes $J_{\tilde{\mathcal{D}}(g_{1:H}^m)}$ over $\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}$ with this private-information component fixed, i.e.,

$$\{\tilde{\mathcal{E}}_{i,h}^*\}_{i \in [n], h \in [H]} \in \underset{\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}}{\operatorname{argmin}} J_{\tilde{\mathcal{D}}(g_{1:H}^m)} \left(\Lambda(\{\tilde{\mathcal{E}}_{i,h}\}_{i \in [n], h \in [H]}, \{\tilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]}) \right).$$

From Theorem IV.2, we know that for any $h \in [H]$

$$\tilde{\mathbf{Z}}_h = \tilde{\chi}_h(\tilde{\mathbf{P}}_{h-1}, \tilde{\mathbf{U}}_{h-1}, \tilde{\mathbf{Y}}_h), \quad \tilde{\mathbf{P}}_h = \tilde{\zeta}_h(\tilde{\mathbf{P}}_{h-1}, \tilde{\mathbf{U}}_{h-1}, \tilde{\mathbf{Y}}_h),$$

where $\tilde{\chi}_h, \tilde{\zeta}_h$ are fixed projection functions. Therefore, there exist matrices such that

$$\tilde{\mathbf{Z}}_h = \tilde{\mathcal{T}}_h^1 \tilde{\mathbf{P}}_{h-1} + \tilde{\mathcal{T}}_h^2 \tilde{\mathbf{U}}_{h-1} + \tilde{\mathcal{T}}_h^3 \tilde{\mathbf{Y}}_h, \quad \tilde{\mathbf{P}}_h = \tilde{\mathcal{L}}_h^1 \tilde{\mathbf{P}}_{h-1} + \tilde{\mathcal{L}}_h^2 \tilde{\mathbf{U}}_{h-1} + \tilde{\mathcal{L}}_h^3 \tilde{\mathbf{Y}}_h,$$

for some matrices $\tilde{\mathcal{T}}_h^1, \tilde{\mathcal{T}}_h^2, \tilde{\mathcal{T}}_h^3, \tilde{\mathcal{L}}_h^1, \tilde{\mathcal{L}}_h^2, \tilde{\mathcal{L}}_h^3$.

As shown in [17], if we fix $\{\tilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]}$, we can construct an H -step centralized linear control problem $\check{\mathcal{D}}$ based on $\tilde{\mathcal{D}}(g_{1:H}^m)$, where $\check{\mathcal{D}}$ is an LQG system with linear dynamics, linear observations, quadratic costs, and Gaussian noises. For completeness, we include the construction below.

Specifically, we can construct the centralized LQG problem $\check{\mathcal{D}}$ based on $\tilde{\mathcal{D}}(g_{1:H}^m)$ and fixed $\{\tilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]}$, where the state $\check{\mathbf{X}}_h \in \check{\mathcal{X}}_h$, the action $\check{\mathbf{U}}_h \in \check{\mathcal{U}}_h$, and the observation $\check{\mathbf{Y}}_h \in \check{\mathcal{Y}}_h$ evolve as follows:

$$\begin{aligned} \forall h \in [H], \check{\mathcal{X}}_h &= \tilde{\mathcal{X}} \times \tilde{\mathcal{P}}_h, \quad \check{\mathcal{U}}_h = \tilde{\mathcal{U}}_h, \quad \check{\mathcal{Y}}_h = \tilde{\mathcal{Z}}_h, \quad \check{\mathbf{X}}_{h+1} = \check{\mathbf{A}}_h \check{\mathbf{X}}_h + \check{\mathbf{B}}_h \check{\mathbf{U}}_h + \check{\mathbf{W}}_h, \\ \check{\mathbf{Y}}_h &= \check{\mathbf{E}}_h \check{\mathbf{X}}_{h-1} + \check{\mathbf{F}}_h \check{\mathbf{U}}_{h-1} + \check{\mathbf{V}}_h \quad (\text{for } h > 1), \quad \check{\mathbf{Y}}_1 = \check{\mathbf{V}}_1, \\ \check{\mathbf{X}}_{H+1} &= \tilde{\mathbf{X}}_{H+1}, \quad \check{\mathbf{X}}_h = \begin{bmatrix} \tilde{\mathbf{X}}_h \\ \tilde{\mathbf{P}}_h \end{bmatrix}, \quad \check{\mathbf{U}}_h = \tilde{\mathbf{U}}_h - \operatorname{diag}(\tilde{\mathcal{F}}_{1,h}^*, \dots, \tilde{\mathcal{F}}_{n,h}^*) \tilde{\mathbf{P}}_h, \quad \check{\mathbf{Y}}_h = \tilde{\mathbf{Z}}_h. \end{aligned}$$

Defining $\tilde{\mathcal{F}}_h^* = \operatorname{diag}(\tilde{\mathcal{F}}_{1,h}^*, \dots, \tilde{\mathcal{F}}_{n,h}^*)$, the matrices $\check{\mathbf{A}}_h, \check{\mathbf{B}}_h, \check{\mathbf{E}}_h, \check{\mathbf{F}}_h$ and noises $\check{\mathbf{W}}_h, \check{\mathbf{V}}_h$ are given by

$$\begin{aligned} \forall h \in [H-1], \check{\mathbf{A}}_h &= \begin{bmatrix} \tilde{\mathbf{A}}_h & \tilde{\mathbf{B}}_h \tilde{\mathcal{F}}_h^* \\ \tilde{\mathcal{L}}_{h+1}^3 \tilde{\mathbf{E}}_{h+1} \tilde{\mathbf{A}}_h & \tilde{\mathcal{L}}_{h+1}^1 + \tilde{\mathcal{L}}_{h+1}^2 \tilde{\mathcal{F}}_h^* + \tilde{\mathcal{L}}_{h+1}^3 \tilde{\mathbf{E}}_{h+1} \tilde{\mathbf{B}}_h \tilde{\mathcal{F}}_h^* \end{bmatrix}, \quad \check{\mathbf{B}}_h = \begin{bmatrix} \tilde{\mathbf{B}}_h \\ \tilde{\mathcal{L}}_{h+1}^2 + \tilde{\mathcal{L}}_{h+1}^3 \tilde{\mathbf{E}}_{h+1} \tilde{\mathbf{B}}_h \end{bmatrix}; \\ \forall h \in [2:H], \check{\mathbf{E}}_h &= \begin{bmatrix} \tilde{\mathcal{T}}_h^3 \tilde{\mathbf{E}}_h \tilde{\mathbf{A}}_{h-1} & \tilde{\mathcal{T}}_h^1 + \tilde{\mathcal{T}}_h^2 \tilde{\mathcal{F}}_{h-1}^* + \tilde{\mathcal{T}}_h^3 \tilde{\mathbf{E}}_h \tilde{\mathbf{B}}_{h-1} \tilde{\mathcal{F}}_{h-1}^* \end{bmatrix}, \quad \check{\mathbf{F}}_h = \begin{bmatrix} \tilde{\mathcal{T}}_h^2 + \tilde{\mathcal{T}}_h^3 \tilde{\mathbf{E}}_h \tilde{\mathbf{B}}_{h-1} \end{bmatrix}; \\ \forall h \in [H-1], \check{\mathbf{W}}_h &= \begin{bmatrix} \tilde{\mathbf{W}}_{0,h} \\ \tilde{\mathcal{L}}_{h+1}^3 (\tilde{\mathbf{E}}_{h+1} \tilde{\mathbf{W}}_{0,h} + \tilde{\mathbf{W}}_{1:n,h+1}) \end{bmatrix}, \quad \check{\mathbf{V}}_{h+1} = \tilde{\mathcal{T}}_{h+1}^3 (\tilde{\mathbf{E}}_{h+1} \tilde{\mathbf{W}}_{0,h} + \tilde{\mathbf{W}}_{1:n,h+1}); \\ \check{\mathbf{A}}_H &= \begin{bmatrix} \tilde{\mathbf{A}}_H & \tilde{\mathbf{B}}_H \tilde{\mathcal{F}}_H^* \end{bmatrix}, \quad \check{\mathbf{B}}_H = \tilde{\mathbf{B}}_H, \quad \check{\mathbf{W}}_H = \tilde{\mathbf{W}}_{0,H}, \quad \check{\mathbf{V}}_1 = \tilde{\mathcal{T}}_1^3 (\tilde{\mathbf{E}}_1 \tilde{\mathbf{X}}_1 + \tilde{\mathbf{W}}_{1:n,1}), \end{aligned}$$

where we recall that $\forall h \in [H-1], \tilde{\mathbf{W}}_{1:n,h+1} = \begin{bmatrix} \tilde{\mathbf{W}}_{1,h+1} \\ \dots \\ \tilde{\mathbf{W}}_{n,h+1} \end{bmatrix}$. The cost of each stage $h \in [H]$ is $\check{c}_h =$

$\begin{bmatrix} \check{\mathbf{X}}_h^\top & \check{\mathbf{U}}_h^\top \end{bmatrix}^\top \begin{bmatrix} \check{\mathbf{Q}}_h^1 & \check{\mathbf{N}}_h \\ \check{\mathbf{N}}_h^\top & \check{\mathbf{Q}}_h^2 \end{bmatrix} \begin{bmatrix} \check{\mathbf{X}}_h \\ \check{\mathbf{U}}_h \end{bmatrix}$ and $\check{c}_{H+1} = \check{\mathbf{X}}_{H+1}^\top \check{\mathbf{Q}}_{H+1}^1 \check{\mathbf{X}}_{H+1}$, where the matrices $\check{\mathbf{Q}}_h^1, \check{\mathbf{Q}}_h^2, \check{\mathbf{N}}_h, \check{\mathbf{Q}}_{H+1}^1$ are defined as

$$\forall h \in [H], \check{\mathbf{Q}}_h^1 = \begin{bmatrix} \tilde{\mathbf{Q}}_h^1 & 0 \\ 0 & (\tilde{\mathcal{F}}_h^*)^\top \tilde{\mathbf{Q}}_h^2 \tilde{\mathcal{F}}_h^* \end{bmatrix}, \quad \check{\mathbf{Q}}_h^2 = \tilde{\mathbf{Q}}_h^2, \quad \check{\mathbf{N}}_h = \begin{bmatrix} 0 \\ (\tilde{\mathcal{F}}_h^*)^\top \tilde{\mathbf{Q}}_h^2 \end{bmatrix}, \quad \check{\mathbf{Q}}_{H+1}^1 = \tilde{\mathbf{Q}}_{H+1}^1.$$

The strategy of $\tilde{\mathcal{D}}$ at each stage $h \in [H]$ is defined by $\check{g}_h : \prod_{t=1}^h \check{\mathcal{Y}}_t \rightarrow \check{\mathcal{U}}_h$ and the objective is defined as $J_{\tilde{\mathcal{D}}}(\check{g}_{1:H}) = \mathbb{E}[\sum_{h=1}^{H+1} \check{c}_h | \check{g}_{1:H}]$. Thus, an optimal strategy satisfies $\check{g}_{1:H}^* \in \operatorname{argmin}_{\check{g}_{1:H}} J_{\tilde{\mathcal{D}}}(\check{g}_{1:H})$.

Applying the Bellman projection argument below recursively to this centralized construction shows that an optimal strategy can be chosen linear even when the induced noises are correlated and singular and the quadratic weights are semidefinite. Under this strategy, the aggregate variables of $\tilde{\mathcal{D}}$ are jointly Gaussian, possibly singular, so the Moore-Penrose conditional-Gaussian formula gives matrices $\check{K}_h$ such that $\widehat{\check{X}}_h = \mathbb{E}[\check{X}_h | \check{Y}_{1:h}] = \check{K}_h \check{Y}_{1:h}$. Hence there exist matrices $\check{\mathcal{E}}_h^*$ such that $\check{g}_h^*(\check{Y}_{1:h}) = \check{\mathcal{E}}_h^* \widehat{\check{X}}_h = \check{\mathcal{E}}_h^* \check{K}_h \check{Y}_{1:h}$ for every $h \in [H]$. Also, one can verify that for any linear

strategy $\check{g}_{1:H}$ with the form $\check{g}_h(\check{Y}_{1:h}) = \check{\mathcal{E}}_h \check{Y}_{1:h}, \forall h \in [H]$ and $\check{\mathcal{E}}_h = \begin{bmatrix} \check{\mathcal{E}}_{1,h} \\ \cdots \\ \check{\mathcal{E}}_{n,h} \end{bmatrix}$, we can construct a linear

strategy of $\tilde{\mathcal{D}}(g_{1:H}^m)$ as $\widetilde{g}_{1:H} := \Lambda(\{\check{\mathcal{E}}_{i,h}^*\}_{i \in [n], h \in [H]}, \{\widetilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]})$ such that $J_{\tilde{\mathcal{D}}}(\check{g}_{1:H}) = J_{\tilde{\mathcal{D}}(g_{1:H}^m)}(\widetilde{g}_{1:H})$, and vice versa. Therefore, $\widetilde{g}_{1:H}^* := \Lambda(\{\check{\mathcal{E}}_{i,h}^* \check{K}_h\}_{i \in [n], h \in [H]}, \{\widetilde{\mathcal{F}}_{i,h}^*\}_{i \in [n], h \in [H]})$ is an optimal strategy of $\tilde{\mathcal{D}}(g_{1:H}^m)$,

where $\check{\mathcal{E}}_h^* = \begin{bmatrix} \check{\mathcal{E}}_{1,h}^* \\ \cdots \\ \check{\mathcal{E}}_{n,h}^* \end{bmatrix}$ for any $h \in [H]$.

From the construction of $\tilde{\mathcal{D}}$, we know that $\check{X}_h = \begin{bmatrix} \widetilde{X}_h \\ \widetilde{P}_h \end{bmatrix}, \check{Y}_h = \widetilde{Z}_h$ for any $h \in [H]$, so $\widehat{\check{X}}_h = \mathbb{E}[\check{X}_h | \check{Y}_{1:h}] = \mathbb{E}[\begin{bmatrix} \widetilde{X}_h \\ \widetilde{P}_h \end{bmatrix} | \widetilde{Z}_{1:h}] = \begin{bmatrix} \mathbb{E}[\widetilde{X}_h | \widetilde{C}_h] \\ \mathbb{E}[\widetilde{P}_h | \widetilde{C}_h] \end{bmatrix} = \widetilde{\Theta}_h$, and $\widetilde{\Psi}_h^3(\widetilde{C}_h) = \check{K}_h \widetilde{C}_h$. The centralized construction proves global optimality. The stronger pointwise Bellman assertion follows by backward induction. Under the appendix terminal convention, $\widetilde{V}_{H+1}(\widetilde{\Theta}) = \widetilde{\Theta}^\top \widetilde{Q}_{H+1}^1 \widetilde{\Theta}$. Suppose $\widetilde{V}_{h+1}$ is quadratic with positive-semidefinite quadratic part, fix a reachable mean $\Theta_h = \theta$, and write $\widetilde{S}_h = \theta + \widetilde{\omega}_h$, where $\widetilde{\omega}_h \sim \mathcal{N}(0, \widetilde{\Sigma}_h)$. For an arbitrary admissible prescription, let $\delta \widetilde{P}_{i,h} := \widetilde{P}_{i,h} - \mathbb{E}[\widetilde{P}_{i,h} | \widetilde{\Theta}_h = \theta]$ and take the L^2 projection of $\widetilde{U}_{i,h}$ onto the affine functions of $\delta \widetilde{P}_{i,h}$

$$\widetilde{U}_{i,h} = a_{i,h} + \widetilde{\mathcal{F}}_{i,h} \delta \widetilde{P}_{i,h} + r_{i,h}, \quad \mathbb{E}[r_{i,h}] = 0, \quad \mathbb{E}[r_{i,h} \delta \widetilde{P}_{i,h}^\top] = 0.$$

Since $(\widetilde{\omega}_h, \delta \widetilde{P}_{i,h})$ is jointly Gaussian, possibly singular, its conditional mean is linear on its support. Hence $\mathbb{E}[r_{i,h} \widetilde{\omega}_h^\top] = 0$, and therefore $r_{i,h}$ is also orthogonal to every $\delta \widetilde{P}_{j,h}$; it is orthogonal to the fresh independent noises as well. The Bellman objective is a positive-semidefinite quadratic in the state, controls, and fresh noises. Expanding it makes all terms linear in the stacked projection residual r_h vanish, while the remaining residual term is nonnegative. Thus an affine prescription dominates every admissible prescription at this fixed mean. Stacking the affine projections gives $\widetilde{U}_h = a_h + \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} \widetilde{\omega}_h$ with block-diagonal $\widetilde{\mathcal{F}}_h$. Its Bellman objective separates into a mean quadratic in a_h and a centered quadratic in $\widetilde{\mathcal{F}}_h$. Both are bounded-below finite-dimensional convex quadratics, so their generalized normal equations are consistent. The centered problem is independent of θ . Writing the mean-dependent part as $a_h^\top G_h a_h + 2a_h^\top C_h \theta + \theta^\top D_h \theta$, the mean quadratic is positive semidefinite in (a_h, θ) , so its block matrix is positive semidefinite. Hence $\ker(G_h) \subseteq \ker(C_h^\top)$, equivalently $\operatorname{range}(C_h) \subseteq \operatorname{range}(G_h)$; therefore $a_h^* = -G_h^\dagger C_h \theta =: \widehat{\mathcal{E}}_h^* \theta$. Thus $\widetilde{\mathcal{F}}_h^*$ and $\widehat{\mathcal{E}}_h^*$ minimize simultaneously for every reachable mean. Since $\widetilde{P}_h = \mathbb{I}_{p,h}(\theta + \widetilde{\omega}_h)$, set $\widetilde{\mathcal{E}}_h^* := \widehat{\mathcal{E}}_h^* - \widetilde{\mathcal{F}}_h^* \mathbb{I}_{p,h}$. The same prescription minimizes at every CIB mean reachable under a preceding strategy, and its linear formula defines the strategy elsewhere. Its quadratic value has a positive-semidefinite quadratic part, so backward induction yields $\widetilde{g}_{i,h}^*(\widetilde{C}_h, \widetilde{P}_{i,h}) = \widetilde{\mathcal{E}}_{i,h}^* \widetilde{\Psi}_h^3(\widetilde{C}_h) + \widetilde{\mathcal{F}}_{i,h}^* \widetilde{P}_{i,h}$, simultaneously linear, CIB-

Markovian, globally optimal, and Bellman optimal over all prescriptions at every such reachable mean. $\square$

C-6 Proof of Theorem IV.6

Proof. The proof consists of two **Parts**.

Part 1: we aim to show the evolution of $\tilde{\Theta}_h$.

For every $h \in [H]$, we define $\bar{X}_h := \tilde{S}_h = \begin{bmatrix} \tilde{X}_h \\ \tilde{P}_h \end{bmatrix}$, $\bar{Y}_h := \tilde{Z}_h$, and

$$\bar{U}_h := \tilde{U}_h^e = \begin{bmatrix} \mathbb{1}[\tilde{B}_{1,h} \neq \mathbf{0}] \tilde{U}_{1,h} \\ \mathbb{1}[\tilde{B}_{2,h} \neq \mathbf{0}] \tilde{U}_{2,h} \\ \dots \\ \mathbb{1}[\tilde{B}_{n,h} \neq \mathbf{0}] \tilde{U}_{n,h} \end{bmatrix},$$

where $\tilde{U}_h^e$ is defined by replacing the $\tilde{U}_{i,h}$ part of $\tilde{U}_h$ by $\mathbf{0}$ for all $i \in [n]$ such that $\tilde{B}_{i,h} = \mathbf{0}$. We claim that such a system is a linear system with the evolution rules given by

$$\begin{aligned} \bar{X}_{h+1} &= \bar{A}_h \bar{X}_h + \bar{B}_h \bar{U}_h + \bar{W}_h, & h \in [H-1], \\ \bar{Y}_h &= \bar{E}_h \bar{X}_{h-1} + \bar{F}_h \bar{U}_{h-1} + \bar{V}_h, & t \in [2:H], \end{aligned}$$

for some matrices $\bar{A}_h, \bar{B}_h, \bar{E}_h, \bar{F}_h$ and Gaussian noises $\bar{W}_h, \bar{V}_h$ defined below.

Firstly, from the system evolution, we know that

$$\forall h \in [H], \quad \tilde{X}_{h+1} = \tilde{A}_h \tilde{X}_h + \tilde{B}_h \tilde{U}_h + \tilde{W}_{0,h}, \quad \tilde{Y}_h = \tilde{E}_h \tilde{X}_h + \tilde{W}_{1:n,h}, \quad \tilde{E}_h = \begin{bmatrix} \tilde{E}_{1,h} \\ \dots \\ \tilde{E}_{n,h} \end{bmatrix}, \quad \tilde{W}_{1:n,h} = \begin{bmatrix} \tilde{W}_{1,h} \\ \dots \\ \tilde{W}_{n,h} \end{bmatrix}.$$

Also, from Theorem IV.2, we know that $\tilde{Z}_h, \tilde{P}_h$ evolve as

$$\tilde{Z}_h = \tilde{T}_h^1 \tilde{P}_{h-1} + \tilde{T}_h^2 \tilde{U}_{h-1} + \tilde{T}_h^3 \tilde{Y}_h, \quad \tilde{P}_h = \tilde{\mathcal{L}}_h^1 \tilde{P}_{h-1} + \tilde{\mathcal{L}}_h^2 \tilde{U}_{h-1} + \tilde{\mathcal{L}}_h^3 \tilde{Y}_h,$$

for some matrices $\{\tilde{T}_h^k\}_{k \in [3]}$ and $\{\tilde{\mathcal{L}}_h^k\}_{k \in [3]}$. Then, we can show that the system dynamics are defined for $h \in [H-1]$ as

$$\bar{A}_h := \begin{bmatrix} \tilde{A}_h & 0 \\ \tilde{\mathcal{L}}_{h+1}^3 \tilde{E}_{h+1} \tilde{A}_h & \tilde{\mathcal{L}}_{h+1}^1 \end{bmatrix}, \quad \bar{B}_h := \begin{bmatrix} \tilde{B}_h \\ \tilde{\mathcal{L}}_{h+1}^2 + \tilde{\mathcal{L}}_{h+1}^3 \tilde{E}_{h+1} \tilde{B}_h \end{bmatrix}, \quad \bar{W}_h = \begin{bmatrix} \tilde{W}_{0,h} \\ \tilde{\mathcal{L}}_{h+1}^3 (\tilde{E}_{h+1} \tilde{W}_{0,h} + \tilde{W}_{1:n,h+1}) \end{bmatrix}.$$

The observation quantities are defined for $t \in [2:H]$ by

$$\bar{E}_t := [\tilde{T}_t^3 \tilde{E}_t \tilde{A}_{t-1} \quad \tilde{T}_t^1], \quad \bar{F}_t := \tilde{T}_t^2 + \tilde{T}_t^3 \tilde{E}_t \tilde{B}_{t-1}, \quad \bar{V}_t := \tilde{T}_t^3 (\tilde{E}_t \tilde{W}_{0,t-1} + \tilde{W}_{1:n,t}).$$

Note that based on Assumption III.2, one can verify that if $\tilde{B}_{i,h} = \mathbf{0}$, then $\tilde{T}_{i,h+1}^2 = \mathbf{0}, \tilde{\mathcal{L}}_{i,h+1}^2 = \mathbf{0}$, where $\tilde{T}_{i,h+1}^2$ and $\tilde{\mathcal{L}}_{i,h+1}^2$ are the corresponding agent i 's column blocks of $\tilde{T}_{h+1}^2, \tilde{\mathcal{L}}_{h+1}^2$. Therefore, it holds that for any $h \in [H-1]$, $\bar{B}_h \tilde{U}_h = \bar{B}_h \tilde{U}_h^e$, $\bar{F}_{h+1} \tilde{U}_h = \bar{F}_{h+1} \tilde{U}_h^e$.

Defining $\bar{\Theta}_h := \mathbb{E}[\bar{X}_h | \bar{Y}_{1:h}, \bar{U}_{1:h-1}] = \mathbb{E}[\tilde{S}_h | \tilde{Z}_{1:h}, \tilde{U}_{1:h-1}]$, the standard linear-system formulas

give for $h \in [H-1]$

$$\begin{aligned}
\bar{\Theta}_{h+1} &= \bar{A}_h \bar{\Theta}_h + \bar{B}_h \bar{U}_h + (\bar{A}_h \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \bar{S}_h) [\bar{E}_{h+1} \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \Sigma_{h+1}^v]^\dagger (\bar{Y}_{h+1} - \bar{E}_{h+1} \bar{\Theta}_h - \bar{F}_{h+1} \bar{U}_h),^4 \\
\text{with } \bar{\Sigma}_{h+1}^p &:= \bar{A}_h \bar{\Sigma}_h^p \bar{A}_h^\top + \Sigma_h^w - (\bar{A}_h \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \bar{S}_h) [\bar{E}_{h+1} \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \Sigma_{h+1}^v]^\dagger (\bar{E}_{h+1} \bar{\Sigma}_h^p \bar{A}_h^\top + \bar{S}_h^\top), \\
\Omega_1 &:= \text{diag}(\Sigma_1, \tilde{\Sigma}_{1,1}, \dots, \tilde{\Sigma}_{n,1}), \quad M_1 := \begin{bmatrix} I & 0 \\ \tilde{\mathcal{L}}_1^3 \tilde{E}_1 & \tilde{\mathcal{L}}_1^3 \end{bmatrix}, \quad N_1 := [\tilde{\mathcal{T}}_1^3 \tilde{E}_1 \quad \tilde{\mathcal{T}}_1^3], \\
\bar{\Theta}_1 &= M_1 \Omega_1 N_1^\top (N_1 \Omega_1 N_1^\top)^\dagger \tilde{Z}_1, \quad \bar{\Sigma}_1^p = M_1 \Omega_1 M_1^\top - M_1 \Omega_1 N_1^\top (N_1 \Omega_1 N_1^\top)^\dagger N_1 \Omega_1 M_1^\top, \\
\Sigma_t^y &:= \begin{cases} \text{diag}(\tilde{\Sigma}_{1,t}, \dots, \tilde{\Sigma}_{n,t}), & t \in [H], \\ 0_{d_x \times d_x}, & t = H+1, \end{cases} \\
\bar{S}_h &:= \text{cov}(\bar{W}_h, \bar{V}_{h+1}) = \begin{bmatrix} \tilde{\Sigma}_{0,h} \tilde{E}_{h+1}^\top \\ \tilde{\mathcal{L}}_{h+1}^3 (\tilde{E}_{h+1} \tilde{\Sigma}_{0,h} \tilde{E}_{h+1}^\top + \Sigma_{h+1}^y) \end{bmatrix} (\tilde{\mathcal{T}}_{h+1}^3)^\top, \quad h \in [H-1], \\
\Sigma_h^v &:= \text{cov}(\bar{V}_h) = \tilde{\mathcal{T}}_h^3 (\tilde{E}_h \tilde{\Sigma}_{0,h-1} \tilde{E}_h^\top + \text{diag}(\tilde{\Sigma}_{1,h}, \dots, \tilde{\Sigma}_{n,h})) (\tilde{\mathcal{T}}_h^3)^\top, \quad h \in [2:H], \\
\Sigma_h^w &:= \text{cov}(\bar{W}_h) = \begin{bmatrix} \tilde{\Sigma}_{0,h} & \tilde{\Sigma}_{0,h} \tilde{E}_{h+1}^\top (\tilde{\mathcal{L}}_{h+1}^3)^\top \\ \tilde{\mathcal{L}}_{h+1}^3 \tilde{E}_{h+1} \tilde{\Sigma}_{0,h} & \tilde{\mathcal{L}}_{h+1}^3 (\tilde{E}_{h+1} \tilde{\Sigma}_{0,h} \tilde{E}_{h+1}^\top + \Sigma_{h+1}^y) (\tilde{\mathcal{L}}_{h+1}^3)^\top \end{bmatrix}, \quad h \in [H-1].
\end{aligned}$$

For $h \in [H-1]$, strict partial nestedness implies that if $\tilde{B}_{i,h} \neq 0$, then both $\tilde{I}_{i,h}$ and $\tilde{U}_{i,h}$ are contained in $\tilde{C}_{h+1}$. Hence a projection $\text{Pr}_h^{z,u}$ satisfies $\tilde{U}_h^e = \text{Pr}_h^{z,u} \tilde{Z}_{h+1}$ for $h < H$. For every such effective action, its information is also revealed at $h+1$; after deleting redundant action coordinates and subtracting their known contribution, the remaining signal is a linear Gaussian observation with noise independent of the past. Thus the known-input Kalman formula below does not treat a private, state-correlated action as exogenous. At $h = H$, the synthetic terminal observation is $\tilde{Y}_{H+1} = \tilde{X}_{H+1}$, so directly set $K_{H+1}^1 = K_{H+1}^2 = K_{H+1}^3 = 0$ and $K_{H+1}^4 = I$; no known-input Kalman argument is used.

Finally, for $h \in [H-1]$, we can write

$$\bar{\Theta}_{h+1} = \bar{A}_h \bar{\Theta}_h + \bar{B}_h \text{Pr}_h^{z,u} \bar{Y}_{h+1} + (\bar{A}_h \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \bar{S}_h) [\bar{E}_{h+1} \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \Sigma_{h+1}^v]^\dagger (\bar{Y}_{h+1} - \bar{E}_{h+1} \bar{\Theta}_h - \bar{F}_{h+1} \text{Pr}_h^{z,u} \bar{Y}_{h+1}). \quad (\text{C.6})$$

Based on Equation (C.6), we know that $\bar{\Theta}_h$ is a linear combination of $\bar{Y}_{1:h} = \tilde{Z}_{1:h} = \tilde{C}_h$. Then, we have $\bar{\Theta}_h = \mathbb{E}[\bar{X}_h | \bar{Y}_{1:h}, \bar{U}_{1:h-1}] = \mathbb{E}[\bar{X}_h | \bar{Y}_{1:h}] = \mathbb{E}[\tilde{S}_h | \tilde{C}_h] = \tilde{\Theta}_h$. Furthermore, we obtain $\tilde{\Sigma}_h = \bar{\Sigma}_h^p$. Therefore, we can conclude

$$\begin{aligned}
\tilde{\Theta}_{h+1} &= \bar{A}_h \tilde{\Theta}_h + \bar{B}_h \text{Pr}_h^{z,u} \tilde{Z}_{h+1} + (\bar{A}_h \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \bar{S}_h) [\bar{E}_{h+1} \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \Sigma_{h+1}^v]^\dagger (\tilde{Z}_{h+1} - \bar{E}_{h+1} \tilde{\Theta}_h - \bar{F}_{h+1} \text{Pr}_h^{z,u} \tilde{Z}_{h+1}), \\
\tilde{\Sigma}_{h+1} &= \bar{A}_h \tilde{\Sigma}_h \bar{A}_h^\top + \Sigma_h^w - (\bar{A}_h \tilde{\Sigma}_h \bar{E}_{h+1}^\top + \bar{S}_h) [\bar{E}_{h+1} \tilde{\Sigma}_h \bar{E}_{h+1}^\top + \Sigma_{h+1}^v]^\dagger (\bar{E}_{h+1} \tilde{\Sigma}_h \bar{A}_h^\top + \bar{S}_h^\top). \quad (\text{C.7})
\end{aligned}$$

Note that for $h < H$, Equation (C.7) gives precisely the functions $\tilde{\Psi}_{h+1}^1, \tilde{\Psi}_{h+1}^2$ in Lemma IV.3.

For $h < H$, from Theorem IV.2, we know that $\tilde{Z}_{h+1} = \tilde{\chi}_{h+1}(\tilde{P}_h, \tilde{U}_h, \tilde{Y}_{h+1})$ for some projection function $\tilde{\chi}_{h+1}$. Then, we can replace $\tilde{Z}_{h+1}$ by a linear combination of $\tilde{P}_h, \tilde{U}_h, \tilde{Y}_{h+1}$ and then obtain $\{K_{h+1}^k\}_{k \in [4]}$ directly. The terminal matrices are the direct values stated above. Substituting those terminal values gives $\tilde{L}_H^1 = \tilde{L}_H^3 = \tilde{B}_H$, $\tilde{L}_H^2 = \tilde{L}_H^4 = \tilde{A}_H \mathbb{I}_{x,H}$, and $\tilde{Y}_H = \Sigma_{0,H}$, as stated in Theorem IV.6. For $h < H$, the two fresh-noise trace terms below combine as $\text{Tr}(\tilde{Y}_h \tilde{R}_{h+1})$.

⁴Let $\Omega_{h+1} := \bar{E}_{h+1} \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \Sigma_{h+1}^v$. Positivity of the joint innovation/state-error covariance implies $\ker(\Omega_{h+1}) \subseteq \ker(\bar{A}_h \bar{\Sigma}_h^p \bar{E}_{h+1}^\top + \bar{S}_h)$. Hence, these Moore-Penrose gains give the canonical singular Kalman update [27].

Part 2: we aim to show that the value function $\tilde{V}_h(\tilde{\Theta}_h)$ has a quadratic form, and the value function and the matrices $\tilde{\mathcal{E}}_h^*, \tilde{\mathcal{F}}_h^*$ of the optimal control strategy can be computed via Riccati Equations.

We prove this by induction. Under the appendix terminal convention, $\tilde{\Theta}_{H+1} = \tilde{\mathbf{X}}_{H+1}$, $\tilde{\Sigma}_{H+1} = 0$, and $\mathbb{I}_{x,H+1} = I$. Thus $\tilde{V}_{H+1}(\tilde{\Theta}_{H+1}) = \tilde{\Theta}_{H+1}^\top \tilde{Q}_{H+1}^1 \tilde{\Theta}_{H+1}$, so $\tilde{R}_{H+1} = \tilde{Q}_{H+1}^1$ and $\tilde{c}_{H+1} = 0$. At $h = H$, this is exactly the final-state term already included in $c_H^\dagger$ in the main text. For the calculation below, we equivalently decompose $c_H^\dagger$ into the ordinary stage- H cost and the auxiliary terminal value $\tilde{V}_{H+1}(\tilde{\mathbf{X}}_{H+1})$. Now, for any given $h \in [H]$, we assume that $\tilde{V}_{h+1}(\tilde{\Theta}_{h+1}) = \tilde{\Theta}_{h+1}^\top \tilde{R}_{h+1} \tilde{\Theta}_{h+1} + \tilde{c}_{h+1}$ for some $\tilde{R}_{h+1} \geq 0$ and $\tilde{c}_{h+1}$. From Lemma IV.3, we know that the CIB belief $\tilde{\mathbf{B}}_h$ is a Gaussian distribution $\mathcal{N}(\tilde{\Theta}_h, \tilde{\Sigma}_h)$, so we can write $\tilde{\mathbf{S}}_h = \tilde{\Theta}_h + \tilde{\omega}_h$, where $\tilde{\omega}_h \sim \mathcal{N}(\mathbf{0}, \tilde{\Sigma}_h)$. From the definition of $\tilde{\mathbf{S}}_h$, we can write $\tilde{\mathbf{X}}_h = \mathbb{I}_{x,h} \tilde{\mathbf{S}}_h$, $\tilde{\mathbf{P}}_h = \mathbb{I}_{p,h} \tilde{\mathbf{S}}_h$ for the projection matrices $\mathbb{I}_{x,h}$ and $\mathbb{I}_{p,h}$. Also, for any $i \in [n]$, since the action of agent i has the form $\tilde{\mathbf{U}}_{i,h} = \tilde{\mathcal{E}}_{i,h} \tilde{\Theta}_h + \tilde{\mathcal{F}}_{i,h} \tilde{\mathbf{P}}_{i,h}$, we can write the joint control

action as $\tilde{\mathbf{U}}_h = \tilde{\mathcal{E}}_h \tilde{\Theta}_h + \tilde{\mathcal{F}}_h \tilde{\mathbf{P}}_h$ with $\tilde{\mathcal{E}}_h := \begin{bmatrix} \tilde{\mathcal{E}}_{1,h} \\ \vdots \\ \tilde{\mathcal{E}}_{n,h} \end{bmatrix}$ and $\tilde{\mathcal{F}}_h := \text{diag}(\tilde{\mathcal{F}}_{1,h}, \dots, \tilde{\mathcal{F}}_{n,h})$. Therefore, $\tilde{\mathbf{U}}_h = \tilde{\mathcal{E}}_h \tilde{\Theta}_h +$

$\tilde{\mathcal{F}}_h \mathbb{I}_{p,h}(\tilde{\Theta}_h + \tilde{\omega}_h) = (\tilde{\mathcal{E}}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h}) \tilde{\Theta}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\omega}_h$. Now, we define $\hat{\mathcal{E}}_h := \tilde{\mathcal{E}}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h}$, and then optimizing $(\tilde{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]})$ is equivalent to optimizing $(\hat{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]})$ with respect to the corresponding spaces. This is because for any $(\tilde{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]})$, we can find the corresponding $(\hat{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]})$, and vice versa. Then, we can write the stage-cost as

$$\begin{aligned} \min_{\tilde{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]}} \mathbb{E}[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + \tilde{\mathbf{U}}_h^\top \tilde{Q}_h^2 \tilde{\mathbf{U}}_h | \tilde{\Theta}_h = \tilde{\Theta}_h, \tilde{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]}] \\ = \min_{\hat{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]}} \mathbb{E}[\tilde{\mathbf{X}}_h^\top \tilde{Q}_h^1 \tilde{\mathbf{X}}_h + \tilde{\mathbf{U}}_h^\top \tilde{Q}_h^2 \tilde{\mathbf{U}}_h | \tilde{\Theta}_h = \tilde{\Theta}_h, \hat{\mathcal{E}}_h, \tilde{\mathcal{F}}_h] \\ = \min_{\hat{\mathcal{E}}_h, \{\tilde{\mathcal{F}}_{i,h}\}_{i \in [n]}} \tilde{\Theta}_h^\top \mathbb{I}_{x,h}^\top \tilde{Q}_h^1 \mathbb{I}_{x,h} \tilde{\Theta}_h + \tilde{\Theta}_h^\top \hat{\mathcal{E}}_h^\top \tilde{Q}_h^2 \hat{\mathcal{E}}_h \tilde{\Theta}_h + \text{Tr}(\mathbb{I}_{x,h} \tilde{\Sigma}_h \mathbb{I}_{x,h}^\top \tilde{Q}_h^1) + \text{Tr}(\tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\Sigma}_h \mathbb{I}_{p,h}^\top \tilde{\mathcal{F}}_h^\top \tilde{Q}_h^2). \end{aligned}$$

For $h < H$, Theorem IV.2 gives the information evolution $\tilde{\mathbf{Z}}_{h+1} = \tilde{\mathbf{X}}_{h+1}(\tilde{\mathbf{P}}_h, \tilde{\mathbf{U}}_h, \tilde{\mathbf{Y}}_{h+1})$, and Lemma IV.3 gives $\tilde{\Theta}_{h+1} = \tilde{\Psi}_{h+1}^1(\tilde{\Theta}_h, \tilde{\mathbf{Z}}_{h+1})$. Then, we can write $\tilde{\Theta}_{h+1} = \tilde{\Psi}_{h+1}^4(\tilde{\Theta}_h, \tilde{\mathbf{P}}_h, \tilde{\mathbf{U}}_h, \tilde{\mathbf{Y}}_{h+1})$ for some linear function $\tilde{\Psi}_{h+1}^4$. Therefore, we can write $\tilde{\Theta}_{h+1} = K_{h+1}^1 \tilde{\Theta}_h + K_{h+1}^2 \tilde{\mathbf{P}}_h + K_{h+1}^3 \tilde{\mathbf{U}}_h + K_{h+1}^4 \tilde{\mathbf{Y}}_{h+1}$ for some matrices $K_{h+1}^1, K_{h+1}^2, K_{h+1}^3, K_{h+1}^4$. At $h = H$, the same expression follows from the appendix terminal convention and the terminal matrices stated above. Also, we know that $\tilde{\mathbf{P}}_h = \mathbb{I}_{p,h}(\tilde{\Theta}_h + \tilde{\omega}_h)$, $\tilde{\mathbf{U}}_h = (\tilde{\mathcal{E}}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h}) \tilde{\Theta}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\omega}_h$, $\tilde{\mathbf{Y}}_{h+1} = \tilde{E}_{h+1} \tilde{\mathbf{X}}_{h+1} + \tilde{\mathbf{W}}_{h+1} = \tilde{E}_{h+1}(\tilde{A}_h \tilde{\mathbf{X}}_h + \tilde{B}_h \tilde{\mathbf{U}}_h + \tilde{\mathbf{W}}_{0,h}) + \tilde{\mathbf{W}}_{1:n,h+1}$, where

$$\tilde{E}_{h+1} = \begin{bmatrix} \tilde{E}_{1,h+1} \\ \vdots \\ \tilde{E}_{n,h+1} \end{bmatrix}, \tilde{\mathbf{W}}_{1:n,h+1} = \begin{bmatrix} \tilde{\mathbf{W}}_{1,h+1} \\ \vdots \\ \tilde{\mathbf{W}}_{n,h+1} \end{bmatrix}. \text{ Therefore, we can write } \tilde{\Theta}_{h+1} \text{ as}$$

$$\begin{aligned} \tilde{\Theta}_{h+1} &= K_{h+1}^1 \tilde{\Theta}_h + K_{h+1}^2 \mathbb{I}_{p,h}(\tilde{\Theta}_h + \tilde{\omega}_h) + K_{h+1}^3 (\hat{\mathcal{E}}_h \tilde{\Theta}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\omega}_h) \\ &\quad + K_{h+1}^4 (\tilde{E}_{h+1}(\tilde{A}_h \mathbb{I}_{x,h}(\tilde{\Theta}_h + \tilde{\omega}_h) + \tilde{B}_h(\hat{\mathcal{E}}_h \tilde{\Theta}_h + \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} \tilde{\omega}_h) + \tilde{\mathbf{W}}_{0,h}) + \tilde{\mathbf{W}}_{1:n,h+1}), \\ &= (K_{h+1}^1 + K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^3 \hat{\mathcal{E}}_h + K_{h+1}^4 \tilde{E}_{h+1} \tilde{A}_h \mathbb{I}_{x,h} + K_{h+1}^4 \tilde{E}_{h+1} \tilde{B}_h \hat{\mathcal{E}}_h) \tilde{\Theta}_h + K_{h+1}^4 \tilde{\mathbf{W}}_{1:n,h+1} \\ &\quad + (K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^3 \tilde{\mathcal{F}}_h \mathbb{I}_{p,h} + K_{h+1}^4 \tilde{E}_{h+1}(\tilde{A}_h \mathbb{I}_{x,h} + \tilde{B}_h \tilde{\mathcal{F}}_h \mathbb{I}_{p,h})) \tilde{\omega}_h + K_{h+1}^4 \tilde{E}_{h+1} \tilde{\mathbf{W}}_{0,h}. \end{aligned}$$

Since $\widetilde{\omega}_h$ depends only on primitive randomness through time h , it is independent of $\widetilde{\mathbf{W}}_{0,h}$ and $\widetilde{\mathbf{W}}_{1:n,h+1}$. Then, we can define the parameters as

$$\begin{aligned}\widetilde{L}_h^1 &:= K_{h+1}^3 + K_{h+1}^4 \widetilde{E}_{h+1} \widetilde{B}_h, & \widetilde{L}_h^2 &:= K_{h+1}^1 + K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^4 \widetilde{E}_{h+1} \widetilde{A}_h \mathbb{I}_{x,h}, \\ \widetilde{L}_h^3 &:= K_{h+1}^3 + K_{h+1}^4 \widetilde{E}_{h+1} \widetilde{B}_h, & \widetilde{L}_h^4 &:= K_{h+1}^2 \mathbb{I}_{p,h} + K_{h+1}^4 \widetilde{E}_{h+1} \widetilde{A}_h \mathbb{I}_{x,h},\end{aligned}$$

and have

$$\widetilde{\Theta}_{h+1} = (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2) \widetilde{\Theta}_h + (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4) \widetilde{\omega}_h + K_{h+1}^4 \widetilde{E}_{h+1} \widetilde{\mathbf{W}}_{0,h} + K_{h+1}^4 \widetilde{\mathbf{W}}_{1:n,h+1}.$$

Thus, from the induction hypothesis, we have

$$\begin{aligned}\mathbb{E}[\widetilde{V}_{h+1}(\widetilde{\Theta}_{h+1}) | \widetilde{\Theta}_h, \widetilde{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h] &= \mathbb{E}[\widetilde{V}_{h+1}(\widetilde{\Theta}_{h+1}) | \widetilde{\Theta}_h, \widehat{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h] = \mathbb{E}[\widetilde{\Theta}_{h+1}^\top \widetilde{R}_{h+1} \widetilde{\Theta}_{h+1} + \widetilde{c}_{h+1} | \widetilde{\Theta}_h, \widetilde{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h] \\ &= \widetilde{\Theta}_h^\top (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2)^\top \widetilde{R}_{h+1} (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2) \widetilde{\Theta}_h + \text{Tr}((\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4) \widetilde{\Sigma}_h (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4)^\top \widetilde{R}_{h+1}) \\ &\quad + \text{Tr}(K_{h+1}^4 \widetilde{E}_{h+1} \Sigma_{0,h} \widetilde{E}_{h+1}^\top (K_{h+1}^4)^\top \widetilde{R}_{h+1}) + \text{Tr}(K_{h+1}^4 \Sigma_{h+1}^y (K_{h+1}^4)^\top \widetilde{R}_{h+1}) + \widetilde{c}_{h+1}.\end{aligned}$$

If we define the following quantities:

$$\begin{aligned}J_h^1(\widetilde{\Theta}_h) &:= \widetilde{\Theta}_h^\top (\mathbb{I}_{x,h}^\top \widetilde{Q}_h^1 \mathbb{I}_{x,h} + \widehat{\mathcal{E}}_h^\top \widetilde{Q}_h^2 \widehat{\mathcal{E}}_h + (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2)^\top \widetilde{R}_{h+1} (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2)) \widetilde{\Theta}_h, \\ J_h^2 &:= \text{Tr}(\widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top \widetilde{\mathcal{F}}_h^\top \widetilde{Q}_h^2) + \text{Tr}((\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4) \widetilde{\Sigma}_h (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4)^\top \widetilde{R}_{h+1}), \\ J_h^3 &:= \text{Tr}(\mathbb{I}_{x,h} \widetilde{\Sigma}_h \mathbb{I}_{x,h}^\top \widetilde{Q}_h^1) + \text{Tr}(K_{h+1}^4 \widetilde{E}_{h+1} \Sigma_{0,h} \widetilde{E}_{h+1}^\top (K_{h+1}^4)^\top \widetilde{R}_{h+1}) + \text{Tr}(K_{h+1}^4 \Sigma_{h+1}^y (K_{h+1}^4)^\top \widetilde{R}_{h+1}) + \widetilde{c}_{h+1},\end{aligned}$$

then we can write $\mathbb{E}[\widetilde{\mathbf{X}}_h^\top \widetilde{Q}_h^1 \widetilde{\mathbf{X}}_h + \widetilde{\mathbf{U}}_h^\top \widetilde{Q}_h^2 \widetilde{\mathbf{U}}_h + \widetilde{V}_{h+1}(\widetilde{\Theta}_{h+1}) | \widetilde{\Theta}_h, \widetilde{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h] = J_h^1(\widetilde{\Theta}_h) + J_h^2 + J_h^3$. Note that $J_h^1(\widetilde{\Theta}_h)$ only depends on $\widetilde{\Theta}_h$ and $\widehat{\mathcal{E}}_h$, J_h^2 only depends on $\widetilde{\mathcal{F}}_h$, and J_h^3 is just a constant. Both minimizations below admit solutions. Indeed, the coefficient matrix $G_h := \widetilde{Q}_h^2 + (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1$ is positive semidefinite and Lemma B.1 gives $\text{range}((\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^2) \subseteq \text{range}(G_h)$. Hence the canonical least-squares gain $-G_h^\dagger (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^2$ minimizes the first quadratic for every mean vector simultaneously. The private-gain objective is a bounded-below finite-dimensional convex quadratic, so its normal equations are consistent and it also has a minimizer. Therefore, if $\widehat{\mathcal{E}}_h^*, \widetilde{\mathcal{F}}_h^*$ satisfy

$$\widehat{\mathcal{E}}_h^* \in \underset{\widehat{\mathcal{E}}_h}{\text{argmin}} J_h^1(\widetilde{\Theta}_h), \quad \forall \widetilde{\Theta}_h \in \widetilde{\mathcal{S}}_h, \quad \widetilde{\mathcal{F}}_h^* \in \underset{\widetilde{\mathcal{F}}_h}{\text{argmin}} J_h^2,$$

then $\widetilde{\mathcal{E}}_h^* := \widehat{\mathcal{E}}_h^* - \widetilde{\mathcal{F}}_h^* \mathbb{I}_{p,h}$ and $\widetilde{\mathcal{F}}_h^*$ will satisfy

$$(\widetilde{\mathcal{E}}_h^*, \widetilde{\mathcal{F}}_h^*) \in \underset{\widetilde{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h}{\text{argmin}} \mathbb{E}[\widetilde{\mathbf{X}}_h^\top \widetilde{Q}_h^1 \widetilde{\mathbf{X}}_h + \widetilde{\mathbf{U}}_h^\top \widetilde{Q}_h^2 \widetilde{\mathbf{U}}_h + \widetilde{V}_{h+1}(\widetilde{\Theta}_{h+1}) | \widetilde{\Theta}_h, \widetilde{\mathcal{E}}_h, \widetilde{\mathcal{F}}_h].$$

Firstly, we intend to minimize $J_h^1(\widetilde{\Theta}_h)$ and find an optimal $\widehat{\mathcal{E}}_h^*$. Since $\widetilde{Q}_h^1, \widetilde{Q}_h^2, \widetilde{R}_{h+1} \geq 0$, we can write

$$J_h^1(\widetilde{\Theta}_h) = \|(\widetilde{Q}_h^1)^{\frac{1}{2}} \mathbb{I}_{x,h} \widetilde{\Theta}_h\|_2^2 + \|(\widetilde{Q}_h^2)^{\frac{1}{2}} \widehat{\mathcal{E}}_h \widetilde{\Theta}_h\|_2^2 + \|\widetilde{R}_{h+1}^{\frac{1}{2}} (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h + \widetilde{L}_h^2) \widetilde{\Theta}_h\|_2^2,$$

which is a convex function of $\widehat{\mathcal{E}}_h$ given $\widehat{\Theta}_h$. Therefore, we can compute the optimal $\widehat{\mathcal{E}}_h$ via the first-order condition. Taking the gradient of $J_h^1(\widehat{\Theta}_h)$ with respect to $\widehat{\mathcal{E}}_h$, we obtain

$$\nabla_{\widehat{\mathcal{E}}_h} J_h^1(\widehat{\Theta}_h) = (2(\widetilde{Q}_h^2 + (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1) \widehat{\mathcal{E}}_h + 2(\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1) \widehat{\Theta}_h \widehat{\Theta}_h^\top.$$

Since it holds that $\ker(\widetilde{Q}_h^2 + (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1) \subseteq \ker((\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1)$ due to Lemma B.1, and $\widetilde{Q}_h^2 + (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1 \geq 0$, we know that this gradient is zero if we choose

$$\widehat{\mathcal{E}}_h^* = -(\widetilde{Q}_h^2 + (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^1)^\dagger (\widetilde{L}_h^1)^\top \widetilde{R}_{h+1} \widetilde{L}_h^2.$$

Then, we know that such a $\widehat{\mathcal{E}}_h^*$ minimizes $J_h^1(\widehat{\Theta}_h)$, and we can compute

$$\widetilde{R}_h = \mathbb{I}_{x,h}^\top \widetilde{Q}_h^1 \mathbb{I}_{x,h} + (\widehat{\mathcal{E}}_h^*)^\top \widetilde{Q}_h^2 \widehat{\mathcal{E}}_h^* + (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h^* + \widetilde{L}_h^2)^\top \widetilde{R}_{h+1} (\widetilde{L}_h^1 \widehat{\mathcal{E}}_h^* + \widetilde{L}_h^2).$$

This is a sum of positive-semidefinite Gram terms, so $\widetilde{R}_h \geq 0$, closing the induction. Note that the use of pseudo-inverse here precisely corresponds to the proof of *generalized* Riccati Equations in [28, 29].

Secondly, we intend to minimize J_h^2 and find an optimal $\widetilde{\mathcal{F}}_h^*$. Hence, the optimal $\widetilde{\mathcal{F}}_h^*$ can be found via the following optimization problem

$$\widetilde{\mathcal{F}}_h^* \in \underset{\widetilde{\mathcal{F}}_h = \text{diag}(\widetilde{\mathcal{F}}_{1,h}, \dots, \widetilde{\mathcal{F}}_{n,h})}{\text{argmin}} \left\{ \text{Tr}(\widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top \widetilde{\mathcal{F}}_h^\top \widetilde{Q}_h^2) + \text{Tr}((\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4) \widetilde{\Sigma}_h (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4)^\top \widetilde{R}_{h+1}) \right\}. \quad (\text{C.8})$$

Then, plugging in the optimal $\widetilde{\mathcal{F}}_h^*$, we can get

$$\begin{aligned} \widetilde{c}_h &= \text{Tr}(\widetilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top (\widetilde{\mathcal{F}}_h^*)^\top \widetilde{Q}_h^2) + \text{Tr}((\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} + \widetilde{L}_h^4) \widetilde{\Sigma}_h (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h^* \mathbb{I}_{p,h} + \widetilde{L}_h^4)^\top \widetilde{R}_{h+1}) + \text{Tr}(K_{h+1}^4 \widetilde{E}_{h+1} \Sigma_{0,h} \widetilde{E}_{h+1}^\top (K_{h+1}^4)^\top \widetilde{R}_{h+1}) \\ &\quad + \text{Tr}(K_{h+1}^4 \Sigma_{h+1}^y (K_{h+1}^4)^\top \widetilde{R}_{h+1}) + \text{Tr}(\mathbb{I}_{x,h} \widetilde{\Sigma}_h \mathbb{I}_{x,h}^\top \widetilde{Q}_h^1) + \widetilde{c}_{h+1}, \end{aligned}$$

which completes **Part 2**.

Combining **Part 1** and **Part 2**, we complete the proof. $\square$

Solution of Equation (C.8): From the fact that $\widetilde{\Sigma}_h, \widetilde{Q}_h^2, \widetilde{R}_{h+1} \geq 0$, we can write

$$J_h^2 = \|(\widetilde{Q}_h^2)^{\frac{1}{2}} \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} (\widetilde{\Sigma}_h)^{\frac{1}{2}}\|_F^2 + \|(\widetilde{R}_{h+1})^{\frac{1}{2}} (\widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} + \widetilde{L}_h^4) (\widetilde{\Sigma}_h)^{\frac{1}{2}}\|_F^2,$$

where $\|X\|_F$ denotes the Frobenius norm of X for any matrix X . Therefore, we know that J_h^2 is a convex function with respect to $\widetilde{\mathcal{F}}_h$, and the optimal $\widetilde{\mathcal{F}}_h^*$ can be computed via the first-order condition.

By taking the gradient of J_h^2 with respect to $\widetilde{\mathcal{F}}_h$, we obtain

$$\nabla_{\widetilde{\mathcal{F}}_h} J_h^2 = 2(\widetilde{Q}_h^2 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top + (\widetilde{L}_h^3)^\top \widetilde{R}_{h+1} \widetilde{L}_h^3 \widetilde{\mathcal{F}}_h \mathbb{I}_{p,h} \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top + (\widetilde{L}_h^3)^\top \widetilde{R}_{h+1} \widetilde{L}_h^4 \widetilde{\Sigma}_h \mathbb{I}_{p,h}^\top).$$

Then, we partition $\nabla_{\widetilde{\mathcal{F}}_h} J_h^2$ into $n \times n$ blocks and set its diagonal blocks to zero, namely, let $[\nabla_{\widetilde{\mathcal{F}}_h} J_h^2]_{i,i} = \mathbf{0}$ for all $i \in [n]$. Here the block partitions follow their natural row and column variables: $[\widetilde{Q}_h^2]_{i,k}$ and $[\widetilde{L}_h^1]_{i,k}$ have dimensions $\dim(\widetilde{\mathcal{U}}_i) \times \dim(\widetilde{\mathcal{U}}_k)$, $[\widetilde{\Sigma}_h]_{k,i}$ has dimension $\dim(\widetilde{\mathcal{P}}_k) \times \dim(\widetilde{\mathcal{P}}_i)$, and $[\widetilde{L}_h^2]_{i,i}$

and $[\nabla_{\tilde{\mathcal{F}}_h} J_h^2]_{i,i}$ have dimensions $\dim(\tilde{\mathcal{U}}_i) \times \dim(\tilde{\mathcal{P}}_i)$. We define $\widehat{\Sigma}_h := \mathbb{I}_{p,h} \tilde{\Sigma}_h \mathbb{I}_{p,h}^\top$, $T_h^1 := (\tilde{L}_h^3)^\top \tilde{R}_{h+1} \tilde{L}_h^3$, and $T_h^2 := (\tilde{L}_h^3)^\top \tilde{R}_{h+1} \tilde{L}_h^4 \tilde{\Sigma}_h \mathbb{I}_{p,h}^\top$, where $\tilde{\Sigma}_h$ is the covariance matrix of $\tilde{\mathcal{P}}_h$. Then, for any $i \in [n]$, it holds

$$[\nabla_{\tilde{\mathcal{F}}_h} J_h^2]_{i,i} = 2 \left(\sum_{k=1}^n ([\tilde{Q}_h^2]_{i,k} + [T_h^1]_{i,k}) \tilde{\mathcal{F}}_{k,h} [\widehat{\Sigma}_h]_{k,i} + [T_h^2]_{i,i} \right),$$

and we can obtain the optimal $\tilde{\mathcal{F}}_h$ by solving the equations

$$\sum_{k=1}^n ([\tilde{Q}_h^2]_{i,k} + [T_h^1]_{i,k}) \tilde{\mathcal{F}}_{k,h} [\widehat{\Sigma}_h]_{k,i} + [T_h^2]_{i,i} = \mathbf{0}.$$

D. Deferred Details of §V

For the backward recursion in this appendix only, we use a cost-free post-decision terminal convention:

$$\mathbf{C}_{(H+1)^-} := \mathbf{C}_{H^+} \cup \{\mathbf{U}_H, \mathbf{X}_{H+1}\}, \quad \mathbf{P}_{(H+1)^-} := \emptyset,$$

and the same convention with every quantity tilded for the strict expansion. Thus $\tilde{\mathbf{B}}_{(H+1)^-} = \delta_{\tilde{\mathbf{X}}_{H+1}}$, $\tilde{\boldsymbol{\Theta}}_{(H+1)^-} = \tilde{\mathbf{X}}_{H+1}$, and $\tilde{\Sigma}_{(H+1)^-} = \mathbf{0}$. This is only terminal bookkeeping after all decisions and changes neither the feasible strategies nor the objective.

D-A Deferred details of results

For $h \in [H]$ and $M_{1:h} \in \mathcal{M}_{1:h}$, let $\Gamma_{i,h}(M_{1:h})$ denote the admissible prescriptions $\gamma_{i,h}^a : \mathcal{P}_{i,h^+}(M_{1:h}) \rightarrow \mathcal{U}_i$, and write γ_h^a and $\Gamma_h^a(M_{1:h})$ for their joint versions.

Lemma D.1. At a realized common history, let $\mu_{i,h}(P_{i,h^-}) := g_{i,h}^m(C_{h^-} \cup P_{i,h^-}, M_{1:h-1})$ and let μ_h be the joint communication prescription. If $\mathcal{D}$ satisfies Assumptions II.1 and III.2, then

$$\mathbf{B}_{h^-} = \Phi_h^1(\mathbf{B}_{(h-1)^+}, \mathbf{Z}_h^b, \gamma_h^a), \quad \mathbf{B}_{h^+} = \Phi_h^2(\mathbf{B}_{h^-}, \mathbf{Z}_h^a, \mathbf{M}_h, \mu_h). \quad (\text{D.1})$$

Under Assumption V.1, μ_h is constant in the private information; conditional on $\mathbf{M}_h$, it drops out, giving $\mathbf{B}_{h^+} = \Phi_h^2(\mathbf{B}_{h^-}, \mathbf{Z}_h^a, \mathbf{M}_h)$.

Proof. Given $\mathbf{B}_{(h-1)^+}$, the fixed model maps, independent noises, and γ_{h-1}^a determine the joint law of $(\mathbf{X}_h, \mathbf{P}_{h^-}, \mathbf{Z}_h^b)$; conditioning on $\mathbf{Z}_h^b$ gives Φ_h^1 . For the post-sharing update, impose $\mathbf{M}_h = \mu_h(\mathbf{P}_{h^-})$ and $\mathbf{Z}_h^a = \phi_h(\mathbf{M}_h, \mathbf{P}_{h^-})$; conditioning on $(\mathbf{M}_h, \mathbf{Z}_h^a)$ gives $\Phi_h^2(\mathbf{B}_{h^-}, \mathbf{Z}_h^a, \mathbf{M}_h, \mu_h)$. Under Assumption V.1, μ_h is constant in $\mathbf{P}_{h^-}$, so conditioning on $\mathbf{M}_h$ adds no private-information likelihood and μ_h may be omitted. $\square$

Given any JCCO problem $\mathcal{D}$ with closed-loop communication strategies, we can expand it into another JCCO problem $\tilde{\mathcal{D}}$ (with notation system $\sim$). First, for any $h \in [H]$, we define the set $\Xi_{h^-} := \{(i, t) | i \in [n], t < h, \mathbf{I}_{i,t^+} \subseteq \mathbf{C}_{h^-} \text{ under any communication strategy, and } B_{i,t} \neq \mathbf{0}\}$ and $\Xi_{h^+} := \{(i, t) | i \in [n], t < h, \mathbf{I}_{i,t^+} \subseteq \mathbf{C}_{h^+} \text{ under any communication strategy, and } B_{i,t} \neq \mathbf{0}\}$. The system variables are identified across the two problems; in particular, $\tilde{\mathbf{X}}_{1:H+1} := \mathbf{X}_{1:H+1}$, $\tilde{\mathbf{Y}}_{j,h} := \mathbf{Y}_{j,h}$, and $\tilde{\mathbf{U}}_{j,h} := \mathbf{U}_{j,h}$.

for every $j \in [n]$ and $h \in [H]$. Then, we can expand $\mathcal{D}$ to $\widetilde{\mathcal{D}}$ as follows: for any $h \in [H]$

$$\begin{aligned}\widetilde{\mathcal{C}}_{h^-} &= \mathcal{C}_{h^-} \cup \{\widetilde{\mathcal{U}}_{i,t} \mid i \in [n], t < h, (i,t) \in \Xi_{h^-}\}, \widetilde{\mathcal{C}}_{h^+} = \mathcal{C}_{h^+} \cup \{\widetilde{\mathcal{U}}_{i,t} \mid i \in [n], t < h, (i,t) \in \Xi_{h^+}\}, \\ \forall i \in [n], \widetilde{\mathcal{P}}_{i,h^-} &= \mathcal{P}_{i,h^-} \setminus \{\widetilde{\mathcal{U}}_{j,t} \mid j \in [n], t < h, (j,t) \in \Xi_{h^-}\}, \widetilde{\mathcal{P}}_{i,h^+} = \mathcal{P}_{i,h^+} \setminus \{\widetilde{\mathcal{U}}_{j,t} \mid j \in [n], t < h, (j,t) \in \Xi_{h^+}\}.\end{aligned}\tag{D.2}$$

Then, we have the following lemma.

Lemma D.2. Let $\mathcal{D}$ be a JCCO problem with PN IS that satisfies Assumptions II.1, III.2, III.4, and V.1, and let $\widetilde{\mathcal{D}}$ be the JCCO expanded from $\mathcal{D}$ according to Equation (D.2). Then, $\widetilde{\mathcal{D}}$ satisfies Assumption II.1, and the two problems have the same set of achievable costs. Hence either problem admits a team-optimal strategy if and only if the other does. Moreover, there exists a function φ^c such that for any team-optimal strategy $\widetilde{g}_{1:H}^{m,*}, \widetilde{g}_{1:H}^{a,*}$ of $\widetilde{\mathcal{D}}$, $(g_{1:H}^{m,*}, g_{1:H}^{a,*}) = \varphi^c(\widetilde{g}_{1:H}^{m,*}, \widetilde{g}_{1:H}^{a,*}, \mathcal{D})$ is a team-optimal strategy of $\mathcal{D}$, and $J_{\mathcal{D}}(g_{1:H}^{m,*}, g_{1:H}^{a,*}) = J_{\widetilde{\mathcal{D}}}(\widetilde{g}_{1:H}^{m,*}, \widetilde{g}_{1:H}^{a,*})$.

Proof. This proof consists of two **Parts**: **Part 1**: $\widetilde{\mathcal{D}}$ satisfies Assumption II.1; **Part 2**: we can construct optimal strategies of $\mathcal{D}$ from the optimal strategies of $\widetilde{\mathcal{D}}$, and the optimal values of the two problems are the same.

Part 1: To begin with, we discuss a property of the problem $\widetilde{\mathcal{D}}$. For any $i \in [n], h \in [H]$, if $B_{i,h} = \mathbf{0}$, then from Assumption III.2, $\mathbf{U}_{i,h} \notin \mathbf{I}_{j,(h')^-}, \mathbf{U}_{i,h} \notin \mathbf{I}_{j,(h')^+}, \forall h' \in [H]$ with $h' > h, j \in [n]$. Also from expansion, we know that $\widetilde{\mathbf{U}}_{i,h} = \mathbf{U}_{i,h}$ will never be added into common information. Therefore, $\widetilde{\mathbf{U}}_{i,h} \notin \widetilde{\mathbf{I}}_{j,(h')^-}, \widetilde{\mathbf{U}}_{i,h} \notin \widetilde{\mathbf{I}}_{j,(h')^+}, \forall h' \in [H]$ with $h' > h, j \in [n]$.

For $h \in [H-1]$, if $B_{i,h} \neq \mathbf{0}$, Assumption III.4 gives $j \neq i$ such that $\mathbf{U}_{i,h}$ influences $\mathbf{Y}_{j,h+1}$ and hence $\mathbf{I}_{j,(h+1)^-}$. Partial nestedness gives $\mathbf{I}_{i,h^-} \subseteq \mathbf{I}_{j,(h+1)^-}$. Since $j \neq i$, every label private to agent i can enter agent j 's information only through common information due to Assumption II.1, and then it holds $\mathbf{I}_{i,h^-} \subseteq \mathcal{C}_{(h+1)^-}$. Moreover, $\mathbf{I}_{i,h^+} \setminus \mathbf{I}_{i,h^-} \subseteq \mathcal{Z}_h^a \subseteq \mathcal{C}_{(h+1)^-}$, so $\mathbf{I}_{i,h^+} \subseteq \mathcal{C}_{(h+1)^-}$. Consequently $(i,h) \in \Xi_{(h')^-} \cap \Xi_{(h')^+}$ for every $h' > h$, and the expansion adds $\widetilde{\mathbf{U}}_{i,h}$ to $\widetilde{\mathcal{C}}_{(h+1)^-}$ immediately. No such argument is needed for $h = H$, because Equation (D.2) only adds actions with $t < h \leq H$.

Then, based on the property discussed above, we show that $\widetilde{\mathcal{D}}$ satisfies Assumption II.1.

- (a) Since $\mathcal{D}$ satisfies Assumption II.1, $\mathbf{Z}_h^b = \chi_h(\mathbf{P}_{(h-1)^+}, \mathbf{U}_{h-1}, \mathbf{Y}_h)$ for a fixed projection function χ_h . We claim that $\widetilde{\mathbf{Z}}_h^b \subseteq \widetilde{\mathcal{P}}_{(h-1)^+} \cup \{\widetilde{\mathbf{U}}_{h-1}, \widetilde{\mathbf{Y}}_h\}$: This is because it holds that $\mathbf{Z}_h^b \subseteq \mathbf{P}_{(h-1)^+} \cup \{\mathbf{U}_{h-1}, \mathbf{Y}_h\}$; from the property discussed above, $\widetilde{\mathbf{Z}}_h^b \setminus \mathbf{Z}_h^b$ only consists of some control actions at timestep $h-1$; from expansion, $(\mathbf{P}_{(h-1)^+} \setminus \widetilde{\mathcal{P}}_{(h-1)^+}) \subseteq \widetilde{\mathcal{C}}_{(h-1)^+}$, so $(\mathbf{P}_{(h-1)^+} \setminus \widetilde{\mathcal{P}}_{(h-1)^+}) \cap \widetilde{\mathbf{Z}}_h^b = \emptyset$. Also, since χ_h is a projection function, and the sets Ξ_{h^-}, Ξ_{h^+} are predefined and are not affected by the realization of each random variable, there exists a fixed projection function $\widetilde{\chi}_h$ such that $\widetilde{\mathbf{Z}}_h^b = \widetilde{\chi}_h(\widetilde{\mathcal{P}}_{(h-1)^+}, \widetilde{\mathbf{U}}_{h-1}, \widetilde{\mathbf{Y}}_h)$.
- (b) For each realization $M_{i,h}$, obtain $\widetilde{\phi}_{i,h}(M_{i,h}, \cdot)$ from the projection $\phi_{i,h}(M_{i,h}, \cdot)$ by deleting exactly the labels already moved to $\widetilde{\mathcal{C}}_{h^-}$ by the expansion. The retained labels lie in $\widetilde{\mathcal{P}}_{i,h^-}$, while the deleted set is fixed by Ξ_{h^-}, Ξ_{h^+} and is independent of all realizations. Hence $\widetilde{\phi}_{i,h}(M_{i,h}, \cdot)$ is a fixed projection and $\widetilde{\mathbf{Z}}_{i,h}^a = \widetilde{\phi}_{i,h}(\widetilde{\mathbf{M}}_{i,h}, \widetilde{\mathcal{P}}_{i,h^-})$. Stacking over agents gives $\widetilde{\mathbf{Z}}_h^a = \widetilde{\phi}_h(\widetilde{\mathbf{M}}_h, \widetilde{\mathcal{P}}_{h^-})$.
- (c) Since $\mathcal{D}$ satisfies Assumption II.1, for any $i \in [n], \mathbf{P}_{i,h^-} = \zeta_{i,h}(\mathbf{P}_{i,(h-1)^+}, \mathbf{U}_{i,h-1}, \mathbf{Y}_{i,h})$ for a fixed projection function $\zeta_{i,h}$. We claim that $\widetilde{\mathcal{P}}_{i,h^-} \subseteq \widetilde{\mathcal{P}}_{i,(h-1)^+} \cup \{\widetilde{\mathbf{U}}_{i,h-1}, \widetilde{\mathbf{Y}}_{i,h}\}$: This is because it holds that $\mathbf{P}_{i,h^-} \subseteq \mathbf{P}_{i,(h-1)^+} \cup \{\mathbf{U}_{i,h-1}, \mathbf{Y}_{i,h}\}$; from expansion, it holds that $\widetilde{\mathcal{P}}_{i,h^-} \subseteq \mathbf{P}_{i,h^-}$, and for any $t < h$ such that $(i,t) \in \Xi_{(h-1)^+}$, it must hold that $(i,t) \in \Xi_{h^-}$, which means if $\widetilde{\mathbf{U}}_{i,t} \in \mathbf{P}_{i,(h-1)^+} \setminus \widetilde{\mathcal{P}}_{i,(h-1)^+}$,

then $\tilde{U}_{i,t} \in \mathbf{P}_{i,h^-} \setminus \tilde{\mathbf{P}}_{i,h^-}$, so $\tilde{U}_{i,t} \notin \tilde{\mathbf{P}}_{i,h^-}$. Also, since $\zeta_{i,h}$ is a projection function, and the sets Ξ_{h^-}, Ξ_{h^+} are predefined and are not affected by the realization of each random variable, there exists a fixed projection function $\tilde{\zeta}_{i,h}$ such that $\tilde{\mathbf{P}}_{i,h^-} = \tilde{\zeta}_{i,h}(\tilde{\mathbf{P}}_{i,(h-1)^+}, \tilde{U}_{i,h-1}, \tilde{Y}_{i,h})$. Therefore, there exists a function $\tilde{\zeta}_h$ such that $\tilde{\mathbf{P}}_{h^-} = \tilde{\zeta}_h(\tilde{\mathbf{P}}_{(h-1)^+}, \tilde{U}_{h-1}, \tilde{Y}_h)$.

(d) For each $i \in [n]$, $\tilde{\mathbf{P}}_{i,h^+} = \tilde{\mathbf{P}}_{i,h^-} \setminus \tilde{\mathbf{Z}}_{i,h}^a$.

(e) For each $i \in [n]$, the construction of $\tilde{\mathcal{D}}$ only adds some actions into $\tilde{\mathbf{C}}_{h^-}, \tilde{\mathbf{C}}_{h^+}$ and does not change the observations. Thus $\tilde{\mathbf{I}}_{i,h^-} \subseteq \tilde{\mathbf{I}}_{i,h^+}$ and $\tilde{\mathbf{Y}}_{i,h} \in \tilde{\mathbf{I}}_{i,h^-}$ for $h \in [H]$, while $\tilde{\mathbf{I}}_{i,h^+} \subseteq \tilde{\mathbf{I}}_{i,(h+1)^-}$ for $h \in [H-1]$.

Part 2: From the construction of $\tilde{\mathcal{D}}$, we know that the system dynamics and communication cost are the same for both $\mathcal{D}$ and $\tilde{\mathcal{D}}$, and it holds that $\mathbf{I}_{i,h^-} \subseteq \tilde{\mathbf{I}}_{i,h^-}$ and $\mathbf{I}_{i,h^+} \subseteq \tilde{\mathbf{I}}_{i,h^+}$ for any $i \in [n], h \in [H]$. So the agents in $\tilde{\mathcal{D}}$ have larger strategy spaces than those in $\mathcal{D}$. Thus every cost achievable in $\mathcal{D}$ is achievable in $\tilde{\mathcal{D}}$.

More strongly, for every feasible strategy $(\tilde{g}_{1:H}^m, \tilde{g}_{1:H}^a)$ of $\tilde{\mathcal{D}}$, we recursively construct a feasible strategy $(g_{1:H}^m, g_{1:H}^a)$ of $\mathcal{D}$ with the same samplewise communication and control actions. For $h = 1$ and any $i \in [n]$, note that $\mathbf{I}_{i,1^-} = \tilde{\mathbf{I}}_{i,1^-}, \mathbf{I}_{i,1^+} = \tilde{\mathbf{I}}_{i,1^+}$ always holds. Then, set $g_{i,1}^m := \tilde{g}_{i,1}^m$ and $g_{i,1}^a := \tilde{g}_{i,1}^a$. It is immediate that these two strategies output the same communication and control actions.

Now assume that $\tilde{g}_{1:h-1}^m, \tilde{g}_{1:h-1}^a$ and $g_{1:h-1}^m, g_{1:h-1}^a$ output the same actions. The only additional variables in $\tilde{\mathbf{C}}_{h^-} \setminus \mathbf{C}_{h^-}$ are actions $\tilde{U}_{j,t}$ with $j \in [n]$ and $t < h$. For each such action, the expansion gives $\mathbf{I}_{j,t^+} \subseteq \mathbf{C}_{h^-}$ under any additional sharing. Meanwhile, since $g_{j,t}^a$ outputs the same control action as $\tilde{g}_{j,t}^a$, we can recover $\tilde{U}_{j,t} = g_{j,t}^a(\mathbf{I}_{j,t^+}, \mathbf{M}_{1:t})$. Therefore, we can construct $\tilde{\mathbf{C}}_{h^-}$ from $\mathbf{C}_{h^-}, \mathbf{M}_{1:h-1}$ and $g_{1:h-1}^a$ by recovering all possible $\tilde{U}_{j,t} \in \tilde{\mathbf{C}}_{h^-} \setminus \mathbf{C}_{h^-}$. We then construct $g_{i,h}^m(\mathbf{C}_{h^-}, \mathbf{M}_{1:h-1}) = \tilde{g}_{i,h}^m(\tilde{\mathbf{C}}_{h^-}, \tilde{\mathbf{M}}_{1:h-1})$, and these two strategies will output the same communication action.

Similarly, assume that $\tilde{g}_{1:h}^m, \tilde{g}_{1:h-1}^a$ and $g_{1:h}^m, g_{1:h-1}^a$ output the same actions. For any i , the only additional variables in $\tilde{\mathbf{I}}_{i,h^+} \setminus \mathbf{I}_{i,h^+}$ are actions $\tilde{U}_{j,t}$ with $j \in [n]$ and $t < h$. For each such action, the expansion gives $\mathbf{B}_{j,t} \neq \mathbf{0}$ and $\mathbf{I}_{j,t^+} \subseteq \mathbf{C}_{h^+}$ under any additional sharing. Meanwhile, since $g_{j,t}^a$ outputs the same control action as $\tilde{g}_{j,t}^a$, we can recover $\tilde{U}_{j,t} = g_{j,t}^a(\mathbf{I}_{j,t^+}, \mathbf{M}_{1:t})$. Therefore, we can construct $\tilde{\mathbf{I}}_{i,h^+}$ from $\mathbf{I}_{i,h^+}, \mathbf{M}_{1:h}$ and $g_{1:h-1}^a$ by recovering all possible $\tilde{U}_{j,t} \in \tilde{\mathbf{I}}_{i,h^+} \setminus \mathbf{I}_{i,h^+}$. We then construct $g_{i,h}^a(\mathbf{I}_{i,h^+}, \mathbf{M}_{1:h}) = \tilde{g}_{i,h}^a(\tilde{\mathbf{I}}_{i,h^+}, \tilde{\mathbf{M}}_{1:h})$, and these two strategies will output the same control action.

Recursion defines $(g_{1:H}^m, g_{1:H}^a) = \varphi^c(\tilde{g}_{1:H}^m, \tilde{g}_{1:H}^a, \mathcal{D})$. The two problems then have identical samplewise actions, dynamics, and costs, so $J_{\mathcal{D}}(g_{1:H}^m, g_{1:H}^a) = J_{\tilde{\mathcal{D}}}(\tilde{g}_{1:H}^m, \tilde{g}_{1:H}^a)$ for every feasible expanded strategy. Thus every cost achievable in $\tilde{\mathcal{D}}$ is achievable in $\mathcal{D}$. Together with the preceding inclusion, the achievable cost sets coincide, and team optimality and its existence transfer in both directions. $\square$

Now, we introduce the following theorem as a full version of Theorem V.3.

Theorem D.3 (Full version of Theorem V.3). Let $\mathcal{D}$ be a JCCO problem with PN IS that satisfies Assumptions II.1, III.2, III.4, and V.1, and let $\tilde{\mathcal{D}}$ be the JCCO expanded from $\mathcal{D}$. Then $\tilde{\mathcal{D}}$ is a JCCO problem satisfying the SI-CIB condition, and Assumptions II.1, III.2, III.4, V.1. Moreover, for any $h \in [H]$, $\tilde{\mathbf{B}}_{h^-}, \tilde{\mathbf{B}}_{h^+}$ admit Gaussian distributions. Formally, $\tilde{\mathbf{B}}_{h^-} = \mathcal{N}(\tilde{\boldsymbol{\Theta}}_{h^-}, \tilde{\boldsymbol{\Sigma}}_{h^-}), \tilde{\mathbf{B}}_{h^+} = \mathcal{N}(\tilde{\boldsymbol{\Theta}}_{h^+}, \tilde{\boldsymbol{\Sigma}}_{h^+})$,

and

$$\begin{aligned}\tilde{\Theta}_{h^-} &= \tilde{\Psi}_{h^-}^1(\tilde{\Theta}_{(h-1)^+}, \tilde{\mathbf{M}}_{1:h-1}, \tilde{\mathbf{Z}}_h^b), & \tilde{\Theta}_{h^+} &= \tilde{\Psi}_{h^+}^1(\tilde{\Theta}_{h^-}, \tilde{\mathbf{M}}_{1:h}, \tilde{\mathbf{Z}}_h^a), \\ \tilde{\Sigma}_{h^-} &= \tilde{\Psi}_{h^-}^2(\tilde{\mathbf{M}}_{1:h-1}), & \tilde{\Sigma}_{h^+} &= \tilde{\Psi}_{h^+}^2(\tilde{\mathbf{M}}_{1:h}),\end{aligned}\quad (\text{D.3})$$

for some functions $\tilde{\Psi}_{h^-}^1, \tilde{\Psi}_{h^+}^1, \tilde{\Psi}_{h^-}^2, \tilde{\Psi}_{h^+}^2$ that do not depend on the strategies $\tilde{g}_{1:h}^m, \tilde{g}_{1:h}^a$. Therefore, it holds that

$$\tilde{\Theta}_{h^-} = \tilde{\Psi}_{h^-}^3(\tilde{\mathbf{C}}_{h^-}, \tilde{\mathbf{M}}_{1:h-1}), \tilde{\Theta}_{h^+} = \tilde{\Psi}_{h^+}^3(\tilde{\mathbf{C}}_{h^+}, \tilde{\mathbf{M}}_{1:h}), \quad (\text{D.4})$$

for some functions $\tilde{\Psi}_{h^-}^3, \tilde{\Psi}_{h^+}^3$ that do not depend on the strategies $\tilde{g}_{1:h}^m, \tilde{g}_{1:h}^a$.

Proof. This proof consists of three **Parts**: **Part 1**: $\tilde{\mathcal{D}}$ satisfies Assumptions II.1, III.2, III.4, and V.1; **Part 2**: $\tilde{\mathcal{D}}$ satisfies the SI-CIB condition; **Part 3**: Beliefs $\tilde{\mathbf{B}}_{h^-}, \tilde{\mathbf{B}}_{h^+}$ admit Gaussian distributions, and the mean and covariance satisfy Equation (D.3) and (D.4). Throughout this proof, untilded quantities belong to the original problem $\mathcal{D}$, whereas tilded quantities belong to its strict expansion $\tilde{\mathcal{D}}$.

Part 1: From Lemma D.2, we know that $\tilde{\mathcal{D}}$ satisfies Assumption II.1; Since for any $i \in [n], h \in [H], t < h$, we add control actions $\tilde{\mathbf{U}}_{j,t}$ into common information at timestep h in the expansion procedure only if $\tilde{B}_{j,t} = B_{j,t} \neq \mathbf{0}$, $\tilde{\mathcal{D}}$ satisfies Assumption III.2; Since expansion procedure does not change the system dynamics and $\tilde{E}_{i,h} = E_{i,h}, \forall i \in [n], h \in [H]$, $\tilde{\mathcal{D}}$ satisfies Assumption III.4; After expansion, $\tilde{g}_{i,h}^m, \forall i \in [n], h \in [H]$ can still only take $(\tilde{\mathbf{C}}_{h^-}, \tilde{\mathbf{M}}_{1:h-1})$ as input, and $\tilde{\mathcal{D}}$ satisfies Assumption V.1.

Part 2: Condition on a reachable common history and message sequence. Communication rules then add no private-information likelihood. If $\tilde{B}_{i,t} = 0$, Assumption III.2 makes the corresponding control irrelevant to later states and information. If $\tilde{B}_{i,t} \neq 0$, the argument in Part 1 gives $\tilde{\mathbf{I}}_{i,t^+} \subseteq \tilde{\mathbf{C}}_{(t+1)^-}$, and the strict expansion also places $\tilde{\mathbf{U}}_{i,t}$ there; the control rule's message-history input is among the separately conditioned messages. Thus past strategy rules impose only compatibility with conditioned variables and do not change either conditional belief, almost surely. Hence $\tilde{\mathcal{D}}$ satisfies SI-CIB. **Part 3**: For a fixed message sequence, conditioning on the common history fixes the strategy inputs and outputs identified in Part 2. After these compatibility relations are removed, the remaining model variables are affine functions of the primitive Gaussian variables. Hence $\tilde{\mathbf{B}}_{h^-}$ and $\tilde{\mathbf{B}}_{h^+}$ are Gaussian almost surely, possibly singular. Conditioned control values affect only affine offsets, so the covariances depend only on $\tilde{\mathbf{M}}_{1:h-1}$ and $\tilde{\mathbf{M}}_{1:h}$, respectively; thus $\tilde{\Sigma}_{h^-} = \tilde{\Psi}_{h^-}^2(\tilde{\mathbf{M}}_{1:h-1})$ and $\tilde{\Sigma}_{h^+} = \tilde{\Psi}_{h^+}^2(\tilde{\mathbf{M}}_{1:h})$. At $H+1$, the synthetic terminal disclosure gives $\tilde{\mathbf{B}}_{(H+1)^-} = \delta_{\tilde{\mathbf{X}}_{H+1}}, \tilde{\Theta}_{(H+1)^-} = \tilde{\mathbf{X}}_{H+1}$, and $\tilde{\Sigma}_{(H+1)^-} = \mathbf{0}$.

Because $\tilde{\mathcal{D}}$ satisfies Assumptions II.1, III.2, and V.1, Lemma D.1 implies that for each $h \in [H]$,

$$\tilde{\mathbf{B}}_{h^-} = \tilde{\Phi}_h^1(\tilde{\mathbf{B}}_{(h-1)^+}, \tilde{\mathbf{Z}}_h^b, \tilde{\gamma}_{h-1}^a), \quad \tilde{\mathbf{B}}_{h^+} = \tilde{\Phi}_h^2(\tilde{\mathbf{B}}_{h^-}, \tilde{\mathbf{Z}}_h^a, \tilde{\mathbf{M}}_h),$$

for some functions $\tilde{\Phi}_h^1, \tilde{\Phi}_h^2$. Meanwhile, we know that $\tilde{\mathbf{B}}_{h^-}, \tilde{\mathbf{B}}_{h^+}$ are Gaussian distributions that can be characterized by the means $\tilde{\Theta}_{h^-}, \tilde{\Theta}_{h^+}$ and the covariances $\tilde{\Sigma}_{h^-}, \tilde{\Sigma}_{h^+}$, respectively. Then, it holds that $(\tilde{\Theta}_{h^-}, \tilde{\Sigma}_{h^-}) = \tilde{\Phi}_{h^-}^3(\tilde{\Theta}_{(h-1)^+}, \tilde{\Sigma}_{(h-1)^+}, \tilde{\mathbf{Z}}_h^b, \tilde{\gamma}_{h-1}^a), (\tilde{\Theta}_{h^+}, \tilde{\Sigma}_{h^+}) = \tilde{\Phi}_{h^+}^3(\tilde{\Theta}_{h^-}, \tilde{\Sigma}_{h^-}, \tilde{\mathbf{Z}}_h^a, \tilde{\mathbf{M}}_h)$ for some functions $\tilde{\Phi}_{h^-}^3, \tilde{\Phi}_{h^+}^3$. From $\tilde{\Sigma}_{h^-} = \tilde{\Psi}_{h^-}^2(\tilde{\mathbf{M}}_{1:h-1}), \tilde{\Sigma}_{h^+} = \tilde{\Psi}_{h^+}^2(\tilde{\mathbf{M}}_{1:h})$, we have

$$\tilde{\Theta}_{h^-} = \tilde{\Psi}_{h^-}^1(\tilde{\Theta}_{(h-1)^+}, \tilde{\mathbf{M}}_{1:h-1}, \tilde{\mathbf{Z}}_h^b, \tilde{\gamma}_{h-1}^a), \tilde{\Theta}_{h^+} = \tilde{\Psi}_{h^+}^1(\tilde{\Theta}_{h^-}, \tilde{\mathbf{M}}_{1:h}, \tilde{\mathbf{Z}}_h^a),$$

for some functions $\tilde{\Psi}_{h^-}^1, \tilde{\Psi}_{h^+}^1$. Furthermore, since $\tilde{\mathcal{D}}$ satisfies the SI-CIB condition, $\tilde{\Theta}_{h^-}$ does not depend on $\tilde{\gamma}_{h-1}^a$, and $\tilde{\Psi}_{h^-}^1, \tilde{\Psi}_{h^+}^1$ do not depend on the strategies $(\tilde{g}_{1:h}^m, \tilde{g}_{1:h}^a)$. Finally, since $\tilde{\mathcal{C}}_{h^-} = (\cup_{t=1}^h \tilde{\mathcal{Z}}_t^b) \cup (\cup_{t=1}^{h-1} \tilde{\mathcal{Z}}_t^a)$, and $\tilde{\mathcal{C}}_{h^+} = (\cup_{t=1}^h \tilde{\mathcal{Z}}_t^b) \cup (\cup_{t=1}^h \tilde{\mathcal{Z}}_t^a)$, we know that $\tilde{\Theta}_{h^-} = \tilde{\Psi}_{h^-}^3(\tilde{\mathcal{C}}_{h^-}, \tilde{\mathcal{M}}_{1:h-1})$, $\tilde{\Theta}_{h^+} = \tilde{\Psi}_{h^+}^3(\tilde{\mathcal{C}}_{h^+}, \tilde{\mathcal{M}}_{1:h})$, for some functions $\tilde{\Psi}_{h^-}^3, \tilde{\Psi}_{h^+}^3$ that do not depend on the strategies $\tilde{g}_{1:h}^m, \tilde{g}_{1:h}^a$. This completes the proof. $\square$

For $\tilde{\mathcal{D}}$, the common-information coordinator states are $(\tilde{\Theta}_{h^-}, \tilde{\mathcal{M}}_{1:h-1})$ before sharing and $(\tilde{\Theta}_{h^+}, \tilde{\mathcal{M}}_{1:h})$ after sharing. Lemma D.1, SI-CIB, and the coordinator-policy correspondence give Algorithm 1. We assume that all displayed Bellman minima have choices defining admissible strategies.

D-B Algorithm to solve PN JCCO with closed-loop communication strategies

Expand $\mathcal{D}$ into $\tilde{\mathcal{D}}$ by Equation (D.2). Under the preceding assumption, Theorem D.3 and Algorithm 1 yield a team-optimal strategy of $\tilde{\mathcal{D}}$, which Lemma D.2 maps to one of $\mathcal{D}$.

Algorithm 1 Dynamic Programming for SI-CIB JCCO with Closed-loop Communication Strategies

Require: JCCO $\widetilde{\mathcal{D}}$ satisfying the SI-CIB condition and Assumptions II.1, III.2, III.4, and V.1.

```

1: for each  $\widetilde{M}_{1:H} \in \widetilde{\mathcal{M}}_{1:H}$ , and each realization  $\widetilde{\Theta}_{(H+1)^-} = \widetilde{X}_{H+1} \in \widetilde{\mathcal{X}}$  do
2:    $\widetilde{V}_{(H+1)^-}(\widetilde{\Theta}_{(H+1)^-}, \widetilde{M}_{1:H}) \leftarrow \widetilde{\Theta}_{(H+1)^-}^\top \widetilde{Q}_{H+1}^1 \widetilde{\Theta}_{(H+1)^-}$ 
3: end for
4: for  $h = H$  to  $1$  do
5:   for each  $\widetilde{M}_{1:h} \in \widetilde{\mathcal{M}}_{1:h}$ , and each realization  $\widetilde{\Theta}_{h^+} \in \widetilde{\mathcal{X}} \times \widetilde{\mathcal{P}}_{h^+}(\widetilde{M}_{1:h})$  do
6:     Choose  $\widetilde{\gamma}_h^{a,*} \in \operatorname{argmin}_{\widetilde{\gamma}_h^a \in \widetilde{\Gamma}_h^a(\widetilde{M}_{1:h})}$ 
7:        $\mathbb{E}[\widetilde{X}_h^\top \widetilde{Q}_h^1 \widetilde{X}_h + \widetilde{U}_h^\top \widetilde{Q}_h^2 \widetilde{U}_h + \widetilde{V}_{(h+1)^-}(\widetilde{\Theta}_{(h+1)^-}, \widetilde{M}_{1:h}) | \widetilde{\Theta}_{h^+} = \widetilde{\Theta}_{h^+}, \widetilde{M}_{1:h} = \widetilde{M}_{1:h}, \widetilde{\gamma}_h^a]$ 
8:      $\widetilde{V}_{h^+}(\widetilde{\Theta}_{h^+}, \widetilde{M}_{1:h}) \leftarrow$ 
9:        $\mathbb{E}[\widetilde{X}_h^\top \widetilde{Q}_h^1 \widetilde{X}_h + \widetilde{U}_h^\top \widetilde{Q}_h^2 \widetilde{U}_h + \widetilde{V}_{(h+1)^-}(\widetilde{\Theta}_{(h+1)^-}, \widetilde{M}_{1:h}) | \widetilde{\Theta}_{h^+} = \widetilde{\Theta}_{h^+}, \widetilde{M}_{1:h} = \widetilde{M}_{1:h}, \widetilde{\gamma}_h^{a,*}]$ 
10:    for each  $\widetilde{C}_{h^+}$  such that  $\widetilde{\Psi}_{h^+}^3(\widetilde{C}_{h^+}, \widetilde{M}_{1:h}) = \widetilde{\Theta}_{h^+}$  do
11:      for  $i \in [n]$  do
12:         $\widetilde{g}_{i,h}^{a,*}(\widetilde{C}_{h^+}, \cdot, \widetilde{M}_{1:h}) \leftarrow \widetilde{\gamma}_{i,h}^{a,*}(\cdot)$ 
13:      end for
14:    end for
15:  end for
16:  for each  $\widetilde{M}_{1:h-1} \in \widetilde{\mathcal{M}}_{1:h-1}$ , and each realization  $\widetilde{\Theta}_{h^-} \in \widetilde{\mathcal{X}} \times \widetilde{\mathcal{P}}_{h^-}(\widetilde{M}_{1:h-1})$  do
17:    Choose  $\widetilde{M}_h^* \in \operatorname{argmin}_{\widetilde{M}_h \in \widetilde{\mathcal{M}}_h} \mathbb{E}[\widetilde{\mathcal{K}}_h(\widetilde{M}_h) + \widetilde{V}_{h^+}(\widetilde{\Theta}_{h^+}, \widetilde{M}_{1:h}) | \widetilde{\Theta}_{h^-} = \widetilde{\Theta}_{h^-}, \widetilde{M}_{1:h-1} = \widetilde{M}_{1:h-1}, \widetilde{M}_h = \widetilde{M}_h^*]$ 
18:     $\widetilde{V}_{h^-}(\widetilde{\Theta}_{h^-}, \widetilde{M}_{1:h-1}) \leftarrow \mathbb{E}[\widetilde{\mathcal{K}}_h(\widetilde{M}_h^*) + \widetilde{V}_{h^+}(\widetilde{\Theta}_{h^+}, (\widetilde{M}_{1:h-1}, \widetilde{M}_h^*)) | \widetilde{\Theta}_{h^-} = \widetilde{\Theta}_{h^-}, \widetilde{M}_{1:h-1} = \widetilde{M}_{1:h-1}, \widetilde{M}_h = \widetilde{M}_h^*]$ 
19:    for each  $\widetilde{C}_{h^-}$  such that  $\widetilde{\Psi}_{h^-}^3(\widetilde{C}_{h^-}, \widetilde{M}_{1:h-1}) = \widetilde{\Theta}_{h^-}$  do
20:      for  $i \in [n]$  do
21:         $\widetilde{g}_{i,h}^{m,*}(\widetilde{C}_{h^-}, \widetilde{M}_{1:h-1}) \leftarrow \widetilde{M}_{i,h}^*$ 
22:      end for
23:    end for
24:  end for
25: return  $(\widetilde{g}_{1:H}^{m,*}, \widetilde{g}_{1:H}^{a,*})$ 

```
