Optimal Control of Multi-Agent Systems with Processing Delays

Mruganka Kashyap* Laurent Lessard†

Abstract

In this article, we consider a cooperative control problem involving a heterogeneous network of dynamically decoupled continuous-time linear plants. The (output-feedback) controllers for each plant may communicate with each other according to a fixed and known transitively closed directed graph. Each transmission incurs a fixed and known time delay. We provide an explicit closed-form expression for the optimal decentralized controller and its associated cost under these communication constraints and standard linear quadratic Gaussian (LQG) assumptions for the plants and cost function. We find the exact solution without discretizing or otherwise approximating the delays. We also present an implementation of each sub-controller that is efficiently computable, and is composed of standard finite-dimensional linear time-invariant (LTI) and finite impulse response (FIR) components, and has an intuitive observer-regulator architecture reminiscent of the classical separation principle.

1 Introduction

In multi-agent systems such as swarms of unmanned aerial vehicles, it may be desirable for agents to cooperate in a decentralized fashion without receiving instructions from a central coordinating entity. Each agent takes local measurements, performs computations, and may communicate its measurements with a given subset of the other agents, with a time delay. In this article, we investigate the problem of optimal control under the aforementioned communication constraints.

We model each agent as a continuous-time linear time-invariant (LTI) system. We make no assumption of homogeneity across agents; each agent may have different dynamics. We assume the aggregate dynamics of all agents are described by the state-space equations

$$\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix}, \tag{1}$$

where x is the global state, z is the regulated output, y is the measured output, w is the exogenous disturbance, and u is the controlled input. The decoupled nature of the agents imposes a sparsity structure on the plant. Namely, if we partition x, y, w, u each into N pieces corresponding to the

*M. Kashyap is with the Department of Electrical and Computer Engineering at Northeastern University, Boston, MA 02115, USA. (e-mail: kashyap.mru@northeastern.edu).

†L. Lessard is with the Department of Mechanical and Industrial Engineering at Northeastern University, Boston, MA 02115, USA. (e-mail: l.lessard@northeastern.edu).

N agents, the conformally partitioned state space matrices A, B_1, B_2, C_2, D_{21} are block-diagonal. The regulated output z, however, couples all agents' states and inputs, so in general C_1 and D_{12} will be dense. The matrix transfer function $(w, u) \to (z, y)$ is a standard four-block plant that takes the form[1]

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} \mathcal{P}_{11}(s) & \mathcal{P}_{12}(s) \\ \mathcal{P}_{21}(s) & \mathcal{P}_{22}(s) \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix}, \tag{2}$$

where $\mathcal{P}_{21}$ and $\mathcal{P}_{22}$ are block-diagonal.

We assume information sharing is mediated by a fixed and known directed graph. Specifically, if there is a (possibly multi-hop) directed path from Agent i to Agent j, then Agent j can observe the local measurements of Agent i with a delay τ. We further assume there are no self-delays, so agents can observe their local measurements instantaneously.

In practice, our setting corresponds to a network where the chief source of latency is due to processing and transmission delays [12, §1.4] (the encoding, decoding, and transmission of information). Therefore, we neglect propagation delays (proportional to distance traveled) and queuing delays (related to network traffic and hops required to reach the destination).

We assume τ is fixed and known and homogeneous across all communication paths, as it is determined by the physical capabilities (e.g., underlying hardware and software) of the individual agents rather than external factors. Thus, Agent i's feedback policy (in the Laplace domain) is of the form[2]

$$u_i = \mathcal{K}_{ii}(s) y_i + \sum_{j \to i} e^{-s\tau} \mathcal{K}_{ij}(s) y_j, \tag{3}$$

where the sum is over all agents j for which there is a directed path from j to i in the underlying communication graph.

Given the four-block plant (2), the directed communication graph, and the processing delay τ, we study the problem of finding a structured controller that is internally stabilizing and minimizes the $\mathcal{H}_2$ norm of the closed-loop map $w \to z$.

In spite of the non-classical information structure present in this problem, it is known that there is a convex Youla-like parameterization of the set of stabilizing structured controllers, and the associated $\mathcal{H}_2$ synthesis problem is a convex, albeit infinite-dimensional, optimization problem.

Main contribution. We provide a complete solution to this structured cooperative control problem that is computationally tractable and intuitively understandable. Specifically, the optimal controller can be implemented with a finite memory and transmission bandwidth that does not grow over time. Moreover, the controller implementations at the level of individual agents have separation structures between the observer and regulator reminiscent of classical $\mathcal{H}_2$ synthesis theory.

In the remainder of this section, we give context to this problem and relate it to works in optimal control, delayed control, and decentralized control. In Section 2, we cover some mathematical preliminaries and give a formal statement of the problem. In Section 3, we give a convex parameterization of all structured suboptimal controllers, and present the $\mathcal{H}_2$-optimal controller for the

[1] In a slight abuse of notation, the vectors z, y, w, and u now refer to the Laplace transforms of the corresponding time-domain signals in (1).

[2] There is no loss of generality in assuming a linear control policy; see Section 1.1 for details.

non-delayed ($\tau = 0$) and delayed ($\tau > 0$) cases. In Section 4, we describe the optimal controller architecture at the level of the individual agents, and give intuitive interpretations of the controller architecture. In Section 5, we present case studies that highlight the trade-offs between processing delay, connectivity of the agents, and optimal control cost. Finally, we conclude in Section 6 and discuss future directions.

1.1 Literature review

If we remove the structural constraint (3) and allow each u_i to have an arbitrary causal dependence on all y_j with no delays, the optimal controller is linear and admits an observer–regulator separation structure [34]. This is the classical $\mathcal{H}_2$ (LQG) synthesis problem, solved for example in [37].

The presence of structural constraints generally leads to an intractable problem [1]. For example, linear compensators can be strictly suboptimal, even under LQG assumptions [33]. Moreover, finding the best *linear* compensator also leads to a non-convex infinite-dimensional optimization problem.

However, not all structural constraints lead to intractable synthesis problems. For LQG problems with *partially nested* information, there is a linear optimal controller [4]. If the information constraint is *quadratically invariant* with respect to the plant, the problem of finding the optimal LTI controller can be convexified [26,27]. The problem considered in this article is both partially nested and quadratically invariant, so there is no loss in assuming a linear policy as we do in (3).

Once the problem is convexified, the optimal controller can be computed exactly using approaches like vectorization [28, 32], or approximated to arbitrary accuracy using Galerkin-style numerical approaches [25, 29]. However, these approaches lead to realizations of the solution that are neither minimal nor easily interpreted. For example, a numerical solution will not reveal a separation structure in the optimal controller, nor will it provide an interpretation of controller states or the signals communicated between agents' controllers. Indeed, the optimal controller may have a rich structure, reminiscent of the centralized separation principle. Such *explicit solutions* were found for broadcast [15], triangular [18, 31], and dynamically decoupled [6, 8, 9] cases.

The previously mentioned works do not consider time delays. In the presence of delays, we distinguish between discrete and continuous time. In discrete time, the delay transfer function z^{-1} is rational. Therefore, the problem may be reduced to the non-delayed case by absorbing each delay into the plant [14]. However, this reduction is not possible in continuous time because the continuous-time delay transfer function $e^{-s\tau}$ is irrational. A Padé approximation may be used for the delays [35], but this leads to approximation error and a larger state dimension.

Although the inclusion of continuous-time delays renders the state space representation infinite-dimensional, the optimal controller may still have a rich structure. For systems with a *dead-time delay* (the entire control loop is subject to the same delay), a loop-shifting approach using finite impulse response (FIR) blocks can transform the problem into an equivalent delay-free LQG problem with a finite-dimensional LTI plant [20, 24]. A similar idea was used in the discrete-time case to decompose the structure into dead-time and FIR components, which can be optimized separately [13].

The loop-shifting technique can be extended to the *adobe delay* case, where the feedback path contains both a delayed and a non-delayed path [21–23]. The loop-shifting technique was also

extended to specific cases like bilateral teleoperation problems that involve two stable plants whose controllers communicate across a delayed channel [2, 10], and haptic interfaces that have two-way communication with a shared virtual environment [11]. Another example is the case of homogeneous agents coupled via a diagonal-plus-low-rank cost [19]. All three of these examples are special cases of the information structure (3).

In the present work, we solve a general structured $\mathcal{H}_2$ synthesis problem with N agents that communicate using a structure of the form (3). We present explicit solutions that show an intuitive observer-regulator structure at the level of each individual sub-controller. Preliminary versions of these results that only considered stable or non-delayed plants were reported in [6,7]. In this article, we consider the general case of an unstable plant, we find an agent-level parameterization of all stabilizing controllers, and we obtain explicit closed-form expressions for the optimal cost.

2 Preliminaries

Transfer matrices. Let $\mathbb{C}_\alpha := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > \alpha\}$ and $\bar{\mathbb{C}}_\alpha := \{s \in \mathbb{C} \mid \mathrm{Re}(s) \geq \alpha\}$. A transfer matrix $\mathcal{G}(s)$ is said to be *proper* if there exists an $\alpha > 0$ such that $\sup_{s\in\mathbb{C}_\alpha} \|\mathcal{G}(s)\| < \infty$. We call this set $\mathcal{L}_{\mathrm{prop}}$. Similarly, a transfer matrix $\mathcal{G}(s)$ is said to be *strictly proper* if this supremum vanishes as $\alpha \to \infty$. The Hilbert space $\mathcal{L}_2$ consists of analytic functions $\mathcal{F} : i\mathbb{R} \to \mathbb{C}^{m\times n}$ equipped with the inner product $\langle \mathcal{F}, \mathcal{G} \rangle := \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{tr}\big(\mathcal{F}(i\omega)^* \mathcal{G}(i\omega)\big)\, \mathrm{d}\omega$, where the inner product induced norm $\|\mathcal{F}\|_2 := \langle \mathcal{F}, \mathcal{F} \rangle^{1/2}$ is bounded. A function $\mathcal{F} : \bar{\mathbb{C}}_0 \to \mathbb{C}^{m\times n}$ is in $\mathcal{H}_2$ if $\mathcal{F}(s)$ is analytic in $\mathbb{C}_0$, $\lim_{\sigma\to 0^+} \mathcal{F}(\sigma + i\omega) = \mathcal{F}(i\omega)$ for almost every $\omega \in \mathbb{R}$, and $\sup_{\sigma\geq 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{tr}\big(\mathcal{F}(\sigma + i\omega)^* \mathcal{F}(\sigma + i\omega)\big)\, \mathrm{d}\omega < \infty$. This supremum is always achieved at $\sigma = 0$ when $\mathcal{F} \in \mathcal{H}_2$. The set $\mathcal{H}_2^\perp$ is the orthogonal complement of $\mathcal{H}_2$ in $\mathcal{L}_2$. The set $\mathcal{RH}_2$ refers to the subspace of strictly proper rational transfer functions with no poles in $\bar{\mathbb{C}}_0$. Similarly, the set $\mathcal{RH}_2^\perp$ refers to the subspace of strictly proper rational transfer functions with all poles in $\mathbb{C}_0$. The set $\mathcal{L}_\infty$ consists of matrix-valued functions $\mathcal{F} : i\mathbb{R} \to \mathbb{C}^{m\times n}$ for which $\sup_{\omega\in\mathbb{R}} \|\mathcal{F}(i\omega)\| < \infty$. $\mathcal{H}_\infty$ and $\mathcal{RH}_\infty$ are defined analogously to $\mathcal{H}_2$ and $\mathcal{RH}_2$.

The state-space notation for transfer functions is

$$\mathcal{G}(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := D + C(sI - A)^{-1}B. \tag{4}$$

A square matrix A is *Hurwitz* if none of its eigenvalues belong to $\mathbb{C}_0$. If A is Hurwitz in (4), then $\mathcal{G} \in \mathcal{RH}_\infty$. If A is Hurwitz and $D = 0$, then $\mathcal{G} \in \mathcal{RH}_2$. The *conjugate* of $\mathcal{G}$ is

$$\mathcal{G}^\sim(s) = \mathcal{G}^\mathsf{T}(-s) = \left[\begin{array}{c|c} -A^\mathsf{T} & C^\mathsf{T} \\ \hline -B^\mathsf{T} & D^\mathsf{T} \end{array}\right].$$

The dynamics (1) and four-block plant $\mathcal{P}$ from (2) satisfy

$$\mathcal{P}(s) := \begin{bmatrix} \mathcal{P}_{11}(s) & \mathcal{P}_{12}(s) \\ \mathcal{P}_{21}(s) & \mathcal{P}_{22}(s) \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]. \tag{5}$$

If we use the feedback policy $u = \mathcal{K}y$, then we can eliminate u and y from (2) to obtain the closed-loop map $w \to z$, which is given by the lower linear fractional transformation (LFT) defined

as $\mathcal{F}_l(\mathcal{P},\mathcal{K}) := \mathcal{P}_{11} + \mathcal{P}_{12}\mathcal{K}(I - \mathcal{P}_{22}\mathcal{K})^{-1}\mathcal{P}_{21}$. LFTs can be inverted: if $\mathcal{K} = \mathcal{F}_l(\mathcal{J},\mathcal{Q})$ and $\mathcal{J}$ has a proper inverse, then $\mathcal{Q} = \mathcal{F}_u(\mathcal{J}^{-1},\mathcal{K})$, where $\mathcal{F}_u$ is the upper linear fractional transformation: $\mathcal{F}_u(\mathcal{P},\mathcal{K}) := \mathcal{P}_{22} + \mathcal{P}_{21}\mathcal{K}(I - \mathcal{P}_{11}\mathcal{K})^{-1}\mathcal{P}_{12}$.

Block indexing. Ordered lists of indices are denoted using $\{\ldots\}$. The total number of agents is N and $[N] := \{1,\ldots,N\}$. The i^{th} subsystem has state dimension n_i, input dimension m_i, and measurement dimension p_i. The global state dimension is $n := n_1 + \cdots + n_N$ and similarly for m and p. The matrix I_k is the identity of size k and $\mathrm{blkd}(\{X_i\})$ is the block-diagonal matrix formed by the blocks $\{X_1,\ldots,X_n\}$. The zeros used throughout are matrix or vector zeros and their sizes are dependent on the context.

We write $\underline{i}$ to denote the *descendants* of node i, i.e., the set of nodes j such that there is a directed path from i to j for all $i \in [N]$. By convention, we list i first, and then the remaining indices in increasing order. The directed path represents the direction of information transfer between the agents. Similarly, $\bar{i}$ denotes the *ancestors* of node i (again listing i first). We also use $\bar{\underline{i}}$ and $\underline{\underline{i}}$ to denote the *strict* ancestors and descendants, respectively, which excludes i. For example, in Fig. 1, we have $\underline{2} = \{2,5\}$ and $\bar{\underline{3}} = \{1,4\}$.

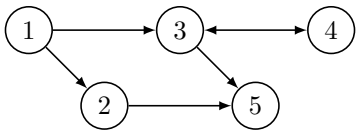

Fig. 1: Directed graph representing five interconnected systems.

We also use this notation to index matrices. For example, if X is a 5×5 block matrix, then $X_{1\underline{2}} = \begin{bmatrix} X_{12} & X_{15} \end{bmatrix}$. We will use specific partitions of the identity matrix throughout: $I_n := \mathrm{blkd}(\{I_{n_i}\})$, and for each agent $i \in [N]$, we define $E_{n_i} := (I_n)_{:i}$ (the i^{th} block column of I_n). We have $n_{\underline{i}} = \sum_{k\in\underline{i}} n_k$ and $n_{\bar{i}} = \sum_{k\in\bar{i}} n_k$, akin to the descendant and ancestor definitions above. The dimensions of $E_{n_{\bar{i}}}$ and $E_{n_{\underline{i}}}$ are determined by the context of use. We also use the notations $X_{:i}$ and $X_{\bar{i}:}$ to indicate the i^{th} block column and $\bar{i}^{\text{th}}$ block rows respectively for a matrix X. Similar notations 1_n is the $n \times 1$ matrix of 1's. Further notations are defined at their points of first use.

2.1 Delay

We follow the notation conventions set in [23]. The *adobe delay* matrix $\Lambda^i_m := \mathrm{blkd}(I_{m_i}, e^{-s\tau} I_{m_{\underline{\underline{i}}}})$ leaves block i unchanged and imposes a delay of τ on all strict descendants of i. We define $\Gamma : (\mathcal{P}, \Lambda^i_m) \mapsto (\tilde{\mathcal{P}}, \Pi_u, \Pi_b)$ that maps the plant $\mathcal{P}$ in (5) and adobe delay matrix Λ^i_m to a modified plant $\tilde{\mathcal{P}}$ and FIR systems Π_u and Π_b. This loop-shifting transformation reported in [21–23] shown in Fig. 2 transforms a loop with adobe input delay into a modified system involving a rational plant $\tilde{\mathcal{P}}$. See Appendix A for details on the definition of Γ.

In this decomposition, $\langle \Delta, \Psi \rangle = 0$ and Ψ is inner (if $\Psi \in \mathcal{RH}_\infty$ and $\Psi^\sim \Psi = I$), so the closed-loop map satisfies $\|\mathcal{F}_l(\mathcal{P}, \Lambda^i_m \mathcal{K})\|^2 = \|\Delta\|^2 + \|\mathcal{F}_l(\tilde{\mathcal{P}}, \tilde{\mathcal{K}})\|^2$. Thus, we can find the $\mathcal{H}_2$-optimal $\mathcal{K}$ by first solving a standard $\mathcal{H}_2$ problem with $\tilde{\mathcal{P}}$ to obtain $\tilde{\mathcal{K}}$, and then transforming back using $\mathcal{K} = \Pi_u \tilde{\mathcal{K}}(I - \Pi_b \tilde{\mathcal{K}})^{-1}$. This transformation, illustrated in the bottom left panel of Fig. 2, has the form of a *modified Smith predictor*, where the FIR blocks Π_u and Π_b compensate for the effect of the adobe delay in the original loop. See [22, §III.C] for further detail.

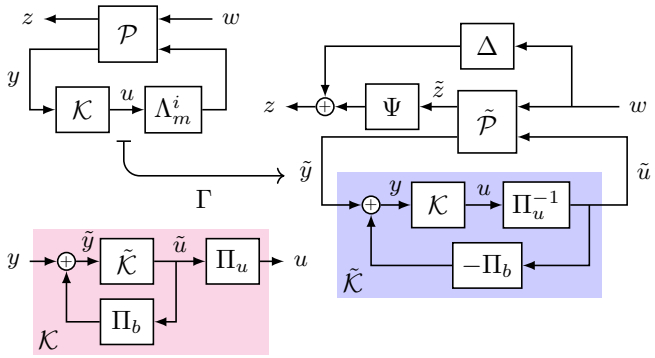

Fig. 2: The loop-shifting approach [21–23] transforms a loop with adobe input delay (top left) into a modified system involving a rational plant $\tilde{\mathcal{P}}$ and FIR blocks Π_u and Π_b (right). This transformation Γ is defined in Appendix A. We can recover $\mathcal{K}$ from $\tilde{\mathcal{K}}$ via the inverse transformation (bottom left).

2.2 Problem statement

Consider a four-block plant (5) representing the aggregated dynamics of N agents as described in Section 1, which we label using indices $i \in [N]$. Suppose $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $y \in \mathbb{R}^p$, partitioned conformally with the N subsystems as $n = n_1 + \cdots + n_N$ and similarly for m and p.

Consider a directed graph on the nodes $[N]$, and let $\mathcal{S}_\tau$ be the set of compensators of the form (3). For example, for the directed graph of Fig. 1, every controller takes the form

$$\begin{bmatrix} \mathcal{K}_{11} & 0 & 0 & 0 & 0 \\ e^{-s\tau}\mathcal{K}_{21} & \mathcal{K}_{22} & 0 & 0 & 0 \\ e^{-s\tau}\mathcal{K}_{31} & 0 & \mathcal{K}_{33} & e^{-s\tau}\mathcal{K}_{34} & 0 \\ e^{-s\tau}\mathcal{K}_{41} & 0 & e^{-s\tau}\mathcal{K}_{43} & \mathcal{K}_{44} & 0 \\ e^{-s\tau}\mathcal{K}_{51} & e^{-s\tau}\mathcal{K}_{52} & e^{-s\tau}\mathcal{K}_{53} & e^{-s\tau}\mathcal{K}_{54} & \mathcal{K}_{55} \end{bmatrix}$$

where $\mathcal{K}_{ij} \in \mathcal{L}_{\text{prop}}$. So each agent may use its local measurements with no delay, and measurements from its ancestors with a delay of τ. An output-feedback policy $u = \mathcal{K}y$ (internally) *stabilizes* $\mathcal{P}$ if

$$\begin{bmatrix} I & -\mathcal{P}_{22} \\ -\mathcal{K} & I \end{bmatrix}^{-1} \in \mathcal{H}_\infty.$$

For further background on stabilization, we refer the reader to [3, 37]. We consider the problem of finding a structured controller that is stabilizing and minimizes the $\mathcal{H}_2$ norm of the closed-loop map. Specifically, we seek to

$$\begin{array}{ll} \underset{\mathcal{K}}{\text{minimize}} & \left\|\mathcal{F}_l(\mathcal{P}, \mathcal{K})\right\|_2^2 \\ \text{subject to} & \mathcal{K} \in \mathcal{S}_\tau \text{ and } \mathcal{K} \text{ stabilizes } \mathcal{P}. \end{array} \tag{6}$$

In the remainder of this section, we list our technical assumptions and define control and estimation gains that will appear in our solution. The assumptions we make ensure that relevant estimation and control subproblems are non-degenerate. We make no assumptions regarding the open-loop stability of $\mathcal{P}$.

Assumption 1 (System assumptions) *For the N interacting agents, the Riccati assumptions defined in Definition 2 hold for (A, B_2, C_1, D_{12}) and for $(A_{ii}^\mathsf{T}, C_{2_{ii}}^\mathsf{T}, B_{1_{ii}}^\mathsf{T}, D_{21_{ii}}^\mathsf{T})$ for all $i \in [N]$.*

Definition 2 (Riccati assumptions) *Matrices (A, B, C, D) satisfy the* Riccati assumptions *[8, 23] if:*

R1. $D^\mathsf{T} D \succ 0$.

R2. (A, B) *is stabilizable.*

R3. $\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$.

If the Riccati assumptions hold, there is a unique stabilizing solution for the corresponding algebraic Riccati equation. We write this as $(X, F) = \mathrm{Ric}(A, B, C, D)$. Thus, $X \succ 0$ satisfies

$$A^\mathsf{T} X + XA + C^\mathsf{T} C - (XB + C^\mathsf{T} D)(D^\mathsf{T} D)^{-1}(B^\mathsf{T} X + D^\mathsf{T} C) = 0,$$

with $A + BF$ Hurwitz and $F := -(D^\mathsf{T} D)^{-1}(B^\mathsf{T} X + D^\mathsf{T} C)$.

2.2.1 Riccati equations

The algebraic Riccati equations (AREs) corresponding to the centralized linear quadratic regulator (LQR) and Kalman filtering are

$$(X_{\mathrm{cen}}, F_{\mathrm{cen}}) := \mathrm{Ric}(A, B_2, C_1, D_{12}), \tag{7a}$$

$$(Y_{\mathrm{cen}}, L_{\mathrm{cen}}^\mathsf{T}) := \mathrm{Ric}(A^\mathsf{T}, C_2^\mathsf{T}, B_1^\mathsf{T}, D_{21}^\mathsf{T}). \tag{7b}$$

Consider controlling the descendants of Agent i using only measurements y_i. The associated four-block plant is

$$\mathcal{P}_i := \begin{bmatrix} \mathcal{P}_{11_{:\underline{i}}} & \mathcal{P}_{12_{:\underline{i}}} \\ \mathcal{P}_{21_{i\underline{i}}} & \mathcal{P}_{22_{i\underline{i}}} \end{bmatrix} := \left[\begin{array}{c|cc} A_{\underline{i}\underline{i}} & B_{1_{\underline{i}i}} & B_{2_{\underline{i}\underline{i}}} \\ \hline C_{1_{:\underline{i}}} & 0 & D_{12_{:\underline{i}}} \\ C_{2_{i\underline{i}}} & D_{21_{ii}} & 0 \end{array}\right], \tag{8}$$

and we define the corresponding ARE solutions as

$$(X^i, F^i) := \mathrm{Ric}(A_{\underline{i}\underline{i}}, B_{2_{\underline{i}\underline{i}}}, C_{1_{:\underline{i}}}, D_{12_{:\underline{i}}}), \tag{9a}$$

$$(Y^i, {L^i}^\mathsf{T}) := \mathrm{Ric}(A_{ii}^\mathsf{T}, C_{2_{ii}}^\mathsf{T}, B_{1_{ii}}^\mathsf{T}, D_{21_{ii}}^\mathsf{T}). \tag{9b}$$

Note that the block-diagonal structure of the estimation subproblems implies $Y_{\mathrm{cen}} = \mathrm{blkd}(\{Y^i\})$ and $L_{\mathrm{cen}} = \mathrm{blkd}(\{L^i\})$. Existence of the matrices defined in (7) and (9) follows from Assumption 1 and the fact that A, B_1, B_2, C_2, and D_{21} are block-diagonal. If we apply the loop-shifting transformation Γ described in Section 2.1 and Fig. 2, we obtain the modified plant

$$\tilde{\mathcal{P}}_i := \begin{bmatrix} \tilde{\mathcal{P}}_{11_{:\underline{i}}} & \tilde{\mathcal{P}}_{12_{:\underline{i}}} \\ \mathcal{P}_{21_{i\underline{i}}} & \tilde{\mathcal{P}}_{22_{i\underline{i}}} \end{bmatrix} := \left[\begin{array}{c|cc} A_{\underline{i}\underline{i}} & B_{1_{\underline{i}i}} & \tilde{B}_{2_{\underline{i}\underline{i}}} \\ \hline \tilde{C}_{1_{:\underline{i}}} & 0 & D_{12_{:\underline{i}}} \\ C_{2_{i\underline{i}}} & D_{21_{ii}} & 0 \end{array}\right].$$

This modified plant has the same estimation ARE as in (9b), but a new control ARE, which we denote

$$(\tilde{X}^i, \tilde{F}^i) := \mathrm{Ric}(A_{\underline{i}\underline{i}}, \tilde{B}_{2_{\underline{i}\underline{i}}}, \tilde{C}_{1_{:\underline{i}}}, D_{12_{:\underline{i}}}), \tag{10}$$

Existence of the matrices defined in (10) also follows from Assumption 1 [23, Lem. 4 and Rem. 1].

3 Optimal Controller

We now present our solution to the structured optimal control problem described in Section 2.2. We begin with a convex parameterization of all structured stabilizing controllers.

3.1 Parameterization of stabilizing controllers

This parameterization is similar to the familiar state-space parameterization of all stabilizing controllers [3, 37], but with an additional constraint on the parameter $\mathcal{Q}$ to enforce the required controller structure.

Lemma 3 *Consider the structured optimal control problem described in Section 2.2with $\mathcal{P}$ given by (5) and suppose Assumption 1 holds. Pick F_d and L_d block-diagonal such that $A + B_2F_d$ and $A + L_dC_2$ are Hurwitz. The following are equivalent:*

(i) $\mathcal{K} \in \mathcal{S}_\tau$ and $\mathcal{K}$ stabilizes $\mathcal{P}$.

(ii) $\mathcal{K} = \mathcal{F}_l(\mathcal{J}, \mathcal{Q})$ for some $\mathcal{Q} \in \mathcal{H}_\infty \cap \mathcal{S}_\tau$, where

$$\mathcal{J} := \left[\begin{array}{c|cc} A + B_2F_d + L_dC_2 & -L_d & B_2 \\ \hline F_d & 0 & I \\ -C_2 & I & 0 \end{array}\right]. \tag{11}$$

Proof. A similar approach was used in [16, Thm. 11] to parameterize the set of stabilizing controllers when $\mathcal{K} \in \mathcal{S}_0$ (no delays). In the absence of the constraint $\mathcal{K} \in \mathcal{S}_\tau$, the set of stabilizing controllers is given by $\{\mathcal{F}_l(\mathcal{J}, \mathcal{Q}) \mid \mathcal{Q} \in \mathcal{H}_\infty\}$ [37, Thm. 12.8]. It remains to show that $\mathcal{K} \in \mathcal{S}_\tau$ if and only if $\mathcal{Q} \in \mathcal{S}_\tau$. Expanding the definition of the lower LFT, we have

$$\mathcal{K} = \mathcal{J}_{11} + \mathcal{J}_{12}\mathcal{Q}\left(I - \mathcal{J}_{22}\mathcal{Q}\right)^{-1}\mathcal{J}_{21}. \tag{12}$$

The matrices A, B_2, C_2, F_d, L_d are block-diagonal, so $\mathcal{J}_{ij}$ is block-diagonal and therefore $\mathcal{J}_{ij} \in \mathcal{S}_\tau$. The delays in our graph satisfy the triangle inequality, so $\mathcal{S}_\tau$ is closed under multiplication (whenever the matrix partitions are compatible). Moreover, $\mathcal{S}_\tau$ is quadratically invariant with respect to $\mathcal{J}_{22}$ [26]. Therefore, if $\mathcal{Q} \in \mathcal{S}_\tau$, then $\mathcal{Q}\left(I - \mathcal{J}_{22}\mathcal{Q}\right)^{-1} \in \mathcal{S}_\tau$ [17, 26], and we conclude from (12) that $\mathcal{K} \in \mathcal{S}_\tau$. Applying the inversion property of LFTs, we have $\mathcal{Q} = \mathcal{F}_u(\mathcal{J}^{-1}, \mathcal{K})$. Now

$$\mathcal{J}^{-1} = \left[\begin{array}{c|cc} A & B_2 & -L_d \\ \hline C_2 & 0 & I \\ -F_d & I & 0 \end{array}\right],$$

so we can apply a similar argument to the above to conclude that $(\mathcal{J}^{-1})_{ij} \in \mathcal{S}_\tau$ and $\mathcal{K} \in \mathcal{S}_\tau \implies \mathcal{Q} \in \mathcal{S}_\tau$. ∎

We refer to $\mathcal{Q}$ in Lemma 3 as the *Youla parameter*, due to its similar role as in the classical Youla parameterization [36].

Remark 4 *Although the problem we consider is quadratically invariant (QI), the existing approaches for convexifying a general QI problem [27] or even a QI problem involving sparsity and*

delays [26] require strong assumptions, such as $\mathcal{P}_{22}$ being stable or strongly stabilizable. Due to the particular delay structure of our problem, the parameterization presented in Lemma 3 does not require any special assumptions and holds for arbitrary (possibly unstable) $\mathcal{P}$.

Remark 5 *In the special case where A is Hurwitz (so $\mathcal{P}$ is stable), we can substitute $F_d = 0$ and $L_d = 0$ in (11) to obtain a simpler parameterization of stabilizing controllers.*

Using the parameterization of Lemma 3, we can rewrite the synthesis problem (6) in terms of the Youla parameter $\mathcal{Q}$. After simplification, we obtain the convex optimization problem

$$\begin{aligned} \underset{\mathcal{Q}}{\text{minimize}} \quad & \left\|\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}\mathcal{T}_{21}\right\|_2^2 \\ \text{subject to} \quad & \mathcal{Q} \in \mathcal{H}_\infty \cap \mathcal{S}_\tau. \end{aligned} \tag{13}$$

where $\mathcal{T} = \begin{bmatrix} \mathcal{T}_{11} & \mathcal{T}_{12} \\ \mathcal{T}_{21} & 0 \end{bmatrix}$

$$= \left[\begin{array}{cc|cc} A+B_2F_d & -B_2F_d & B_1 & B_2 \\ 0 & A+L_dC_2 & B_1+L_dD_{21} & 0 \\ \hline C_1+D_{12}F_d & -D_{12}F_d & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right]. \tag{14}$$

Remark 6 *The convex problem (13)–(14) is similar to its unstructured counterpart [37, Thm. 12.16], except we have the additional constraint $\mathcal{Q} \in \mathcal{S}_\tau$ on the Youla parameter.*

Remark 7 *We use $L := L_{\text{cen}} = L_d = \text{blkd}(\{L^i\})$ throughout the rest of the article. This choice of L yields a $\mathcal{Q}_{\text{opt}}$ with reduced state dimension and simplifies our exposition.*

3.2 Optimal controller without delays

When there are no processing delays ($\tau = 0$), the optimal structured controller is rational. We now provide an explicit state-space formula for this optimal $\mathcal{K}$.

Theorem 8 *Consider the structured optimal control problem described in Section 2.2 and suppose Assumption 1 holds. Choose a block-diagonal F_d such that $A + B_2F_d$ is Hurwitz. A realization of the $\mathcal{Q}_{\text{opt}}$ that solves (13) in the case $\tau = 0$ is*

$$\mathcal{Q}_{\text{opt}} = \left[\begin{array}{c|c} \bar{A}+\bar{B}\bar{F} & -\bar{L}\bar{\mathbf{1}}_p \\ \hline \bar{\mathbf{1}}_m^{\mathsf{T}}(\bar{F}-\bar{F}_d) & 0 \end{array}\right] \tag{15}$$

and a corresponding $\mathcal{K}_{\text{opt}}$ that solves (6) is

$$\mathcal{K}_{\text{opt}} = \left[\begin{array}{c|c} \bar{A}+\bar{B}\bar{F}+\bar{L}\bar{C}\bar{\mathbf{1}}_n\bar{\mathbf{1}}_n^{\mathsf{T}} & -\bar{L}\bar{\mathbf{1}}_p \\ \hline \bar{\mathbf{1}}_m^{\mathsf{T}}\bar{F} & 0 \end{array}\right]. \tag{16}$$

In (15)–(16), we defined the new symbols

$$\begin{aligned} &\bar{A} := I_N \otimes A, \quad \bar{B} := I_N \otimes B_2, \quad \bar{C} := I_N \otimes C_2, \quad \bar{F}_d := I_N \otimes F_d, \\ &\bar{\mathbf{1}}_n := 1_N \otimes I_n, \qquad \bar{\mathbf{1}}_m := 1_N \otimes I_m, \qquad \bar{\mathbf{1}}_p := 1_N \otimes I_p. \end{aligned}$$

Matrices $\bar{L}$ and $\bar{F}$ are block-diagonal concatenations of zero-padded LQR and Kalman gains for each agent. Specifically, $\bar{F} := \text{blkd}(\{E_{m_{\underline{i}}}F^iE_{n_{\underline{i}}}^{\mathsf{T}}\})$ and $\bar{L} := \text{blkd}(\{E_{n_i}L^iE_{p_i}^{\mathsf{T}}\})$ for all $i \in [N]$, where F^i and L^i are defined in (9).

Proof. See Appendix C. ∎

Remark 9 *The optimal controller* (16) *can also be expressed explicitly in terms of the adjacency matrix; see for example [6, 30]. We opt for the realization* (16) *as this expression generalizes more readily to the case with delays.*

Remark 10 *Since agents can act as relays, any cycles in the communication graph can be collapsed and the associated nodes can be aggregated when there are no delays. For example, the graph of Fig. 1 would become the four-node diamond graph* $\{1\} \to \{3,4\} \to \{5\}$*, and* $\{1\} \to \{2\} \to \{5\}$*. So in the delay-free setting, there is no loss of generality in assuming the communication graph is acyclic.*

Remark 11 *Although the optimal* $\mathcal{Q}_{\text{opt}}$ (15) *and associated* $\mathcal{J}$ (11) *depend explicitly on* F_d*, the optimal* $\mathcal{K}_{\text{opt}}$ (16) *does not.*

3.3 Optimal controller with delays

In this section, we generalize Theorem 8 to include an arbitrary but fixed processing delay $\tau > 0$. To this end, we introduce a slight abuse of notation to aid in representing non-rational transfer functions. We generalize the notation of (4) to allow for A, B, C, D that depend on s. So we write:

$$\left[\begin{array}{c|c} A(s) & B(s) \\ \hline C(s) & D(s) \end{array}\right] := D(s) + C(s)\left(sI - A(s)\right)^{-1} B(s).$$

Theorem 12 *Consider the setting of Theorem 8. The transfer function of* $\mathcal{Q}_{\text{opt}} \in \mathcal{H}_\infty \cap \mathcal{S}_\tau$ *that solves* (13) *for any* $\tau \geq 0$ *is*

$$\mathcal{Q}_{\text{opt}} = \left[\begin{array}{cc|c} \bar{A}+\bar{L}\bar{C} & \tilde{B}\tilde{F}-\bar{L}\bar{\Pi}_b\tilde{F}-\bar{B}\bar{\Pi}_u\tilde{F} & 0 \\ \bar{L}\bar{C} & \bar{A}+\tilde{B}\tilde{F}-\bar{L}\bar{\Pi}_b\tilde{F} & -\bar{L}\bar{\mathbf{1}}_p \\ \hline \bar{\mathbf{1}}_m^{\mathsf{T}}\bar{\Lambda}_m\bar{F}_d & \bar{\mathbf{1}}_m^{\mathsf{T}}\bar{\Lambda}_m(\bar{\Pi}_u\tilde{F}-\bar{F}_d) & 0 \end{array}\right] \tag{17}$$

and a corresponding $\mathcal{K}_{\text{opt}}$ *that solves* (6) *is*

$$\mathcal{K}_{\text{opt}} = \left[\begin{array}{cc|c} \bar{A}+\tilde{B}\tilde{F}+\bar{L}\bar{C}\bar{\mathbf{1}}_n\bar{\mathbf{1}}_n^{\mathsf{T}}\bar{\Lambda}_n-\bar{L}\bar{\Pi}_b\tilde{F} & & -\bar{L}\bar{\mathbf{1}}_p \\ \hline \bar{\mathbf{1}}_m^{\mathsf{T}}\bar{\Lambda}_m\bar{\Pi}_u\tilde{F} & & 0 \end{array}\right], \tag{18}$$

where $\bar{A}$*,* $\bar{L}$*,* $\bar{F}_d$*,* $\bar{\mathbf{1}}_n$*,* $\bar{\mathbf{1}}_m$*,* $\bar{\mathbf{1}}_p$*, are defined in Theorem 8. The remainder of the symbols are defined as follows. We apply the loop-shifting transformation* $(\tilde{\mathcal{P}}_i, \Pi_{u_i}, \Pi_{b_i}) = \Gamma(\mathcal{P}_i, \Lambda_m^i)$*, where* $\mathcal{P}_i$*,* $\tilde{\mathcal{P}}_i$*,* $\tilde{F}^i$ *are defined in Section 2.2.1, and*

$$\begin{aligned} \tilde{F} &:= \text{blkd}(\{E_{m_{\underline{i}}}\tilde{F}^i E_{n_{\underline{i}}}^{\mathsf{T}}\}), \quad \bar{\Pi}_b := \text{blkd}(\{E_{p_{\underline{i}}}\Pi_{b_i}E_{m_{\underline{i}}}^{\mathsf{T}}\}), \\ \tilde{B} &:= \text{blkd}(\{E_{n_{\underline{i}}}\tilde{B}_{2_{\underline{i}\underline{i}}}E_{m_{\underline{i}}}^{\mathsf{T}}\}), \quad \bar{\Pi}_u := \text{blkd}(\{E_{m_{\underline{i}}}\Pi_{u_i}E_{m_{\underline{i}}}^{\mathsf{T}}\}), \\ \bar{\Lambda}_k &:= \text{blkd}(\{E_{k_{\underline{i}}}\Lambda_k^i E_{k_{\underline{i}}}^{\mathsf{T}}\}), \quad \textit{for } k \in \{m, n\}. \end{aligned}$$

Proof. See Appendix D. ∎

The transfer matrices $\mathcal{Q}_{\text{opt}}$ in (17) and $\mathcal{K}_{\text{opt}}$ in (18) are not rational, due to the presence of the FIR blocks $\bar{\Pi}_u$, $\bar{\Pi}_b$, and delay blocks $\bar{\Lambda}_m$ and $\bar{\Lambda}_n$. Consequently, we cannot write standard state-space realizations as in Theorem 8. When $\tau = 0$, we have $\bar{\Pi}_u = I$, $\bar{\Pi}_b = 0$, $\bar{\Lambda}_m = I$, $\tilde{F} = \bar{F}$, and $\tilde{B} = \bar{B}$, and we recover the results of Theorem 8.

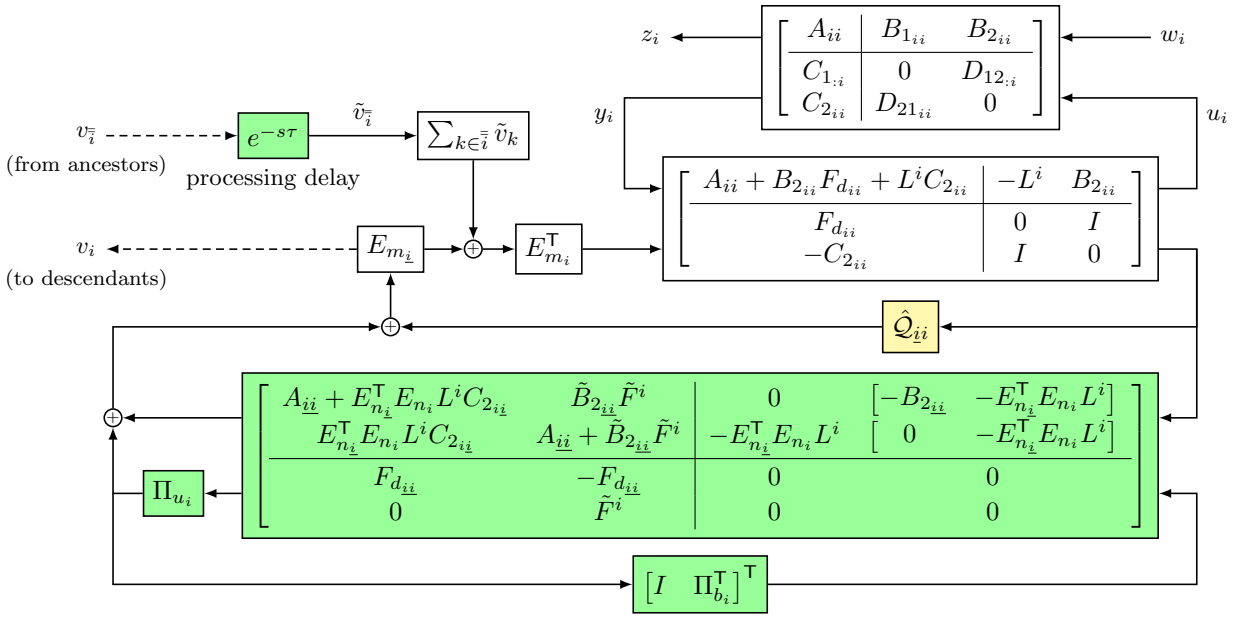

Fig. 3: Agent-level implementation of all structured stabilizing controllers, parameterized by $\hat{\mathcal{Q}} \in \mathcal{H}_\infty \cap \mathcal{S}_0$. Here, F_d is any block-diagonal matrix such that $A_{ii} + B_{2_{ii}} F_{d_{ii}}$ is Hurwitz. The $\mathcal{H}_2$-optimal controller is achieved when $\hat{\mathcal{Q}} = 0$, and results in the simplified diagram of Fig. 4. The blocks that depend on the processing delay τ are colored in green. All symbols are defined in Theorem 13.

4 Agent-level controllers

The optimal controller presented in Theorem 8 is generally not minimal. For example, $\mathcal{K}_{\text{opt}}$ in (16) has a state dimension of Nn, which means a copy of the global plant state for each agent. However, if we extract the part of $\mathcal{K}_{\text{opt}}$ associated with a particular agent, there is a dramatic reduction in state dimension. So in a distributed implementation of this controller, each agent would only need to store a small subset of the controller's state. A similar reduction exists for the optimal controller for the delayed problem presented in Theorem 12.

Our next result presents reduced implementations for these *agent-level* controllers and characterizes the information each agent should store and communicate with their neighbors. We find that Agent i simulates its descendants' dynamics, and so has dimension $n_{\underline{i}}$, which is at least N times smaller than the dimension Nn of the aggregate optimal controller from Theorem 8.

Theorem 13 *Consider the setting of Theorem 8 with $\tau \geq 0$. The agent-level implementation of all structured stabilizing controllers, parameterized by $\hat{\mathcal{Q}} \in \mathcal{H}_\infty \cap \mathcal{S}_0$, is shown in Fig. 3. Here, the optimal controller is achieved when $\hat{\mathcal{Q}} = 0$. In this case, we obtain the simpler structure of Fig. 4. All symbols used are defined in Theorems 8 and 12.*

Proof. See Appendix E. ∎

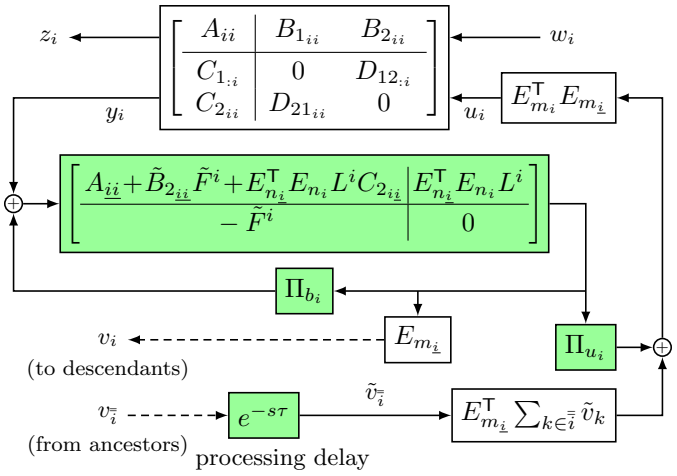

Fig. 4: Agent-level implementation of the $\mathcal{H}_2$-optimal controller with processing delays. This is the result of setting $\hat{\mathcal{Q}} = 0$ in Fig. 3. The blocks that depend on the processing delay τ are colored in green. All symbols are defined in Theorem 13.

4.1 Interpretation of optimal controller

Fig. 3 shows that Agent i transmits the same signal v_i to each of its strict descendants. When an agent receives the signals $v_{\bar{\bar{i}}}$ from its strict ancestors $\bar{\bar{i}}$, it selectively extracts and sums together certain components of the signals. To implement the optimal controller, each agent only needs to know the dynamics and topology of its descendants.

If the network has the additional property that there is at most one directed path connecting any two nodes[3], then the communication scheme can be further simplified. Since Agent i's decision u_i is a sum of terms from all ancestors, but each ancestor has exactly one path that leads to i, the optimal controller can be implemented by transmitting all information to *immediate descendants only* and performing recursive summations. This scheme is illustrated for a four-node chain graph in Fig. 5.

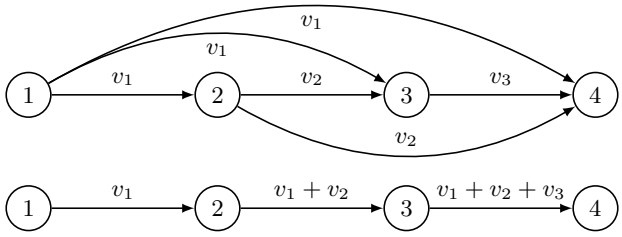

Fig. 5: Four-agent chain graph with standard broadcast (top) and efficient immediate-neighbor implementation (bottom), which is possible because this graph is a multitree.

Remark 14 *The agent-level controller from Fig. 4 can be represented as the combination of an observer with transfer matrix* $\mathcal{T}_{\underline{ii}} := (sI - A_{\underline{ii}} - E_{n_{\underline{i}}}^\mathsf{T} E_{n_i} L^i C_{2_{i\underline{i}}})^{-1}$*, and a regulator with an LQR gain* $\tilde{F}^i$ *in Fig. 6. This yields a separation structure reminiscent of standard LQG theory [37].*

[3]Also known as a *multitree* or a *diamond-free poset*.

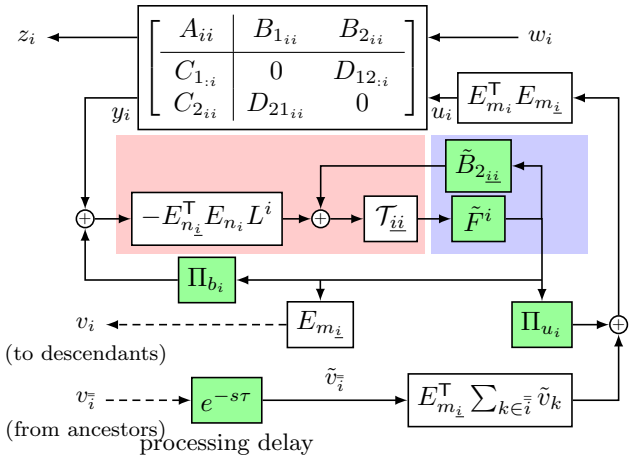

Fig. 6: Agent-level implementation of the $\mathcal{H}_2$-optimal controller with processing delays, featuring an observer (red) and regulator (blue) separation structure. Here $\mathcal{T}_{\underline{ii}} := (sI - A_{\underline{ii}} - E_{n_{\underline{i}}}^\mathsf{T} E_{n_i} L^i C_{2_{\underline{i}}})^{-1}$ is the transfer matrix of the observer dynamics.

Remark 15 *Compared to the architecture proposed in [7, Fig. 4], the agent-level optimal controller in Fig. 4 is more efficient because each agent transmits a single vector v_i to its descendants, instead of two.*

Remark 16 *The controller in Fig. 4 has the form of a feed-forward Smith predictor, similar to Fig. 2 (bottom left). The FIR block Π_{u_i} compensates for the effect of adobe delay. Similarly, the FIR block Π_{b_i} resembles the internal feedback in traditional dead-time controllers.*

5 Characterizing the cost

In this section, we characterize the cost of any structured stabilizing controller. The cost is defined as $J := \left\|\mathcal{F}_l(\mathcal{P}, \mathcal{K})\right\|_2^2 = \left\|\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}\mathcal{T}_{21}\right\|_2^2$, where $\mathcal{K}$ is feasible for (6) or equivalently, $\mathcal{Q} = \mathcal{F}_u(\mathcal{J}^{-1}, \mathcal{K})$ is feasible for (13) (see Lemma 3). We show how to interpret the cost in different ways, and how to compute it efficiently. We illustrate our result using an example with $N = 4$ agents.

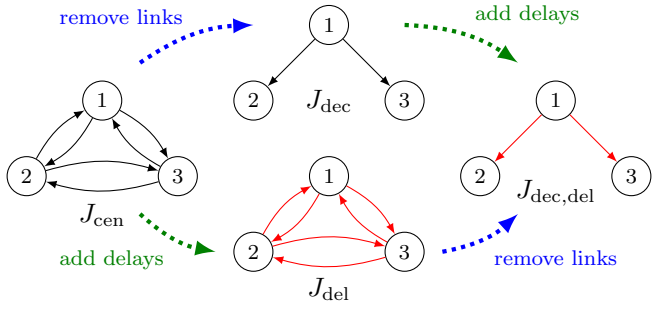

Fig. 7: Hierarchy of optimal costs for different communication patterns in a three-agent example. Additional cost is incurred if links are removed (blue dotted arrows), or if processing delay is added (green dottted arrows). Delayed edges are red. In this example, $J_{\text{cen}} \leq J_{\text{dec}} \leq J_{\text{dec,del}}$ and $J_{\text{cen}} \leq J_{\text{del}} \leq J_{\text{dec,del}}$ but J_{dec} and J_{del} are not comparable.

Theorem 17 *Consider the setting of Theorem 8. The optimal (minimal) costs for the cases: a fully connected graph with no delays, a decentralized graph with no delays, a fully connected graph*

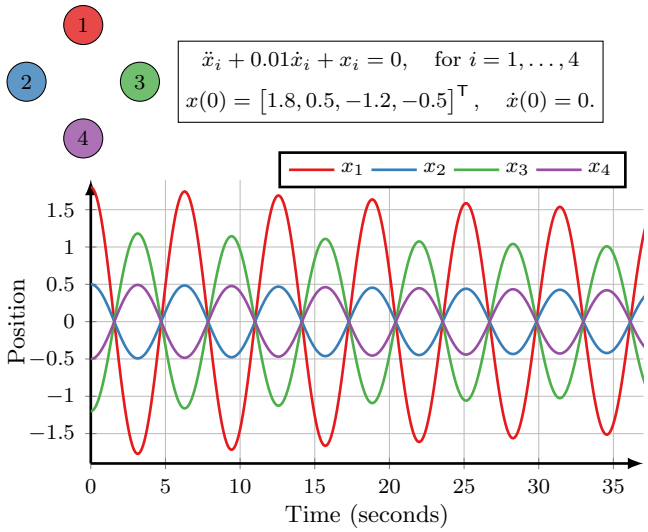

Fig. 8: Open-loop zero-input response of a network of four lightly damped oscillators.

with delays, and a decentralized graph with delays are:

$$J_{\text{cen}} = \text{tr}(Y_{\text{cen}} C_1^\mathsf{T} C_1) + \text{tr}(X_{\text{cen}} L D_{21} D_{21}^\mathsf{T} L^\mathsf{T}), \tag{19a}$$
$$J_{\text{dec}} = \text{tr}(Y_{\text{cen}} C_1^\mathsf{T} C_1) + \text{tr}(X_{\text{dec}} L D_{21} D_{21}^\mathsf{T} L^\mathsf{T}), \tag{19b}$$
$$J_{\text{del}} = \text{tr}(Y_{\text{cen}} C_1^\mathsf{T} C_1) + \text{tr}(X_{\text{del}} L D_{21} D_{21}^\mathsf{T} L^\mathsf{T}), \tag{19c}$$
$$J_{\text{dec,del}} = \text{tr}(Y_{\text{cen}} C_1^\mathsf{T} C_1) + \text{tr}(X_{\text{dec,del}} L D_{21} D_{21}^\mathsf{T} L^\mathsf{T}), \tag{19d}$$

respectively. If a feasible but sub-optimal $\mathcal{Q}$ is used in any of the above cases, write $\mathcal{Q}_\Delta := \mathcal{Q} - \mathcal{Q}_{\text{opt}}$. The cost of this sub-optimal $\mathcal{Q}$ is found by adding $J_Q := \left\| \mathcal{T}_{12} \mathcal{Q}_\Delta D_{21} \right\|_2^2$ to (19a)–(19d)*. The various symbols are defined as*

$$X_{\text{dec}} := \text{blkd}(\{X^i(1,1)\}), \quad X_{\text{del}} := \text{blkd}(\{\Xi^i_{c_\tau}(1,1)\}),$$
$$X_{\text{dec,del}} := \text{blkd}(\{\Xi^i_\tau(1,1)\}), \qquad \textit{and satisfy}$$

$$\text{blkd}(\{X_{\text{cen}}(i,i)\}) \preceq X_{\text{dec}} \preceq X_{\text{dec,del}}, \tag{20a}$$
$$\text{blkd}(\{X_{\text{cen}}(i,i)\}) \preceq X_{\text{del}} \preceq X_{\text{dec,del}}. \tag{20b}$$

$X_{\text{cen}}, Y_{\text{cen}}, F_{\text{cen}}$, and L are defined in Section 2.2.1. Ξ^i_τ and $\Xi^i_{c_\tau}$ are defined in Appendices F.6 and F.7, respectively.

Proof. See Appendix F. ∎

In (19a) we recognize J_{cen} as the standard LQG cost (fully connected graph with no delays). Further, there are two intuitive interpretations for Theorem 17 that are represented in Fig. 7 for a 3-agents system. The intermediate graph topologies are different, but the starting and ending topologies are equal for both. Along the upper path, $J_{\text{dec}} - J_{\text{cen}}$ is the additional cost incurred due

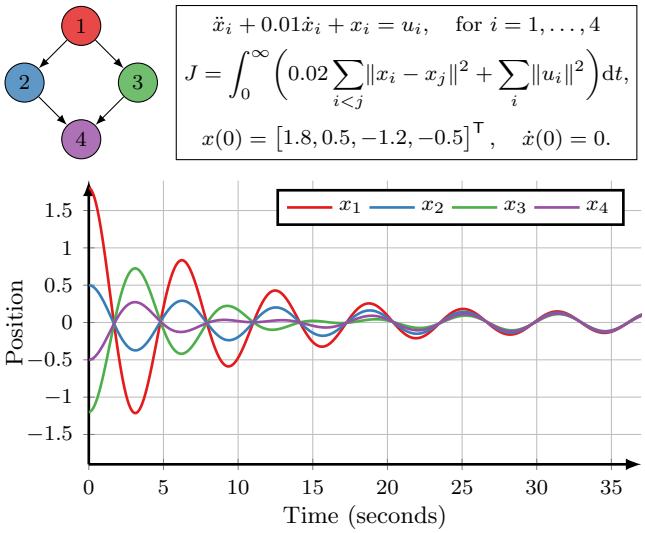

Fig. 9: Closed-loop response of the four-oscillator system from Fig. 8 using the optimal controller from Theorem 8 for a diamond-shaped communication graph with no processing delay. The oscillators leverage the communication network to achieve synchronization.

to decentralization alone, and $J_{\text{dec,del}} - J_{\text{dec}}$ is the further additional cost due to delays. Likewise, along the lower path, $J_{\text{del}} - J_{\text{cen}}$ is the additional cost due to delays alone and $J_{\text{dec,del}} - J_{\text{del}}$ is the further additional cost due to decentralization. Finally, J_Q is the additional cost incurred due to suboptimality. Theorem 17 unifies existing cost decomposition results for the centralized [37, §14.6], decentralized [18, Thm. 16], and delayed [23, Prop. 6] cases.

Remark 18 *Delay and decentralization do not contribute independently to the cost. Specifically, the marginal increase in cost due to adding processing delays depends on the graph topology. Likewise, the marginal increase in cost due to removing communication links depends on the processing delay. In other words,* $J_{\text{cen}} + J_{\text{dec,del}} \neq J_{\text{dec}} + J_{\text{del}}$.

Remark 19 *There is a dual expression for the cost* J_{cen} *in* (19a)*:*

$$J_{\text{cen}} = \operatorname{tr}(X_{\text{cen}} B_1 B_1^\mathsf{T}) + \operatorname{tr}(Y_{\text{cen}} F_{\text{cen}}^\mathsf{T} D_{12}^\mathsf{T} D_{12} F_{\text{cen}}).$$

The corresponding dual expressions for (19b)–(19d) *are unfortunately more complicated. See Appendix F.3 for details.*

5.1 Synchronization example

We demonstrate Theorem 8 via a simple structured LQG example. We consider $N = 4$ identical lightly damped oscillators. The oscillators begin with different initial conditions and the goal is to achieve synchronization. The oscillators have identical dynamics defined by the differential equations in Figs. 8 and 9. Fig. 8 shows the open-loop zero-input response for the four oscillators with given initial conditions. Due to the light damping, the states slowly converge to zero as $t \to \infty$.

Fig. 9 shows the closed-loop response using the optimal controller from Theorem 8 for a diamond-shaped communication network with no processing delay. The controller states are initialized to

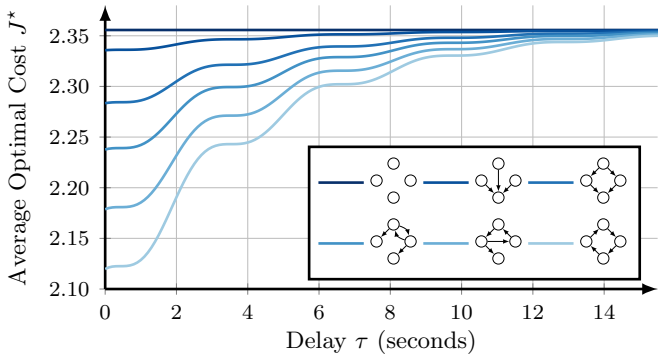

Fig. 10: The average optimal cost, as a function of processing delay, for the 4-agent system of Fig. 9 with different network topologies. For each topology, the cost is an increasing function of the processing delay.

match the initial state of the plant. Since the observer is an unbiased estimator, the LQG controller replicates the behavior of full-state feedback LQR. Fig. 9 shows the four oscillators leveraging their shared information to achieve synchronization to a common oscillation pattern.

In Fig. 10, we use the same system as in Fig. 9, but we plot the total average cost as a function of time delay for various network topologies. The highest cost corresponds to a fully disconnected network, while the lowest cost corresponds to a fully connected network. In the limit as $\tau \to \infty$ (infinite processing delay), the cost tends to that of the fully disconnected case.

6 Conclusion

We studied a structured optimal control problem where multiple dynamically decoupled agents communicate over a delay network. Specifically, we characterized the structure and efficient implementation of optimal controllers at the individual agent level. We now propose some possible future applications for our work.

First, our approach can be readily generalized to treat cases with a combination of processing delays and network latency, where the various delays are heterogeneous but known [5]. Next, the observer-regulator architecture elucidated in Fig. 6 could also be used to develop heuristics for solving cooperative control problems where the agents' dynamics are nonlinear or the noise distributions are non-Gaussian. Examples could include decentralized versions of the Extended Kalman Filter or Unscented Kalman Filter. Finally, our closed-form expressions for the optimal cost can serve as lower bounds to the cost of practical implementation that have additional memory, power, or bandwidth limitations.

Appendices

A Definition of the Γ function

The Γ function takes in a four-block plant $\mathcal{P}$ and adobe delay matrix Λ_m^i and returns a transformed plant $\tilde{\mathcal{P}}$ and FIR systems Π_u, Π_b. As in [23], we first consider the special case where $D_{12}^\mathsf{T} D_{12} = I$. The completion operator $\pi_\tau\{\cdot\}$ acts on a rational LTI system delayed by τ and returns the unique FIR system supported on $[0, \tau]$ that provides a rational completion:

$$\pi_\tau\left\{\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] e^{-s\tau}\right\} := \left[\begin{array}{c|c} A & e^{-A\tau}B \\ \hline C & 0 \end{array}\right] - \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] e^{-s\tau}.$$

The input matrices B_2 and D_{12} of $\mathcal{P}$ are partitioned according to the blocks of adobe delay matrix Λ_m^i. So, $B_2 = \begin{bmatrix} B_{2_0} & B_{2_\tau} \end{bmatrix}$, where the two blocks correspond to inputs with delay 0 and τ, respectively. D_{12} is partitioned in a similar manner. Define the Hamiltonian matrix

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} := \begin{bmatrix} A - B_{2_0} D_{12_0}^\mathsf{T} C_1 & -B_{2_0} B_{2_0}^\mathsf{T} \\ -C_1^\mathsf{T} P_\tau C_1 & -A^\mathsf{T} + C_1^\mathsf{T} D_{12_0} B_{2_0}^\mathsf{T} \end{bmatrix},$$

where $P_0 := D_{12_0} D_{12_0}^\mathsf{T}$ and $P_\tau := I - P_0$, and define its matrix exponential as $\Sigma := e^{H\tau}$. Define the modified matrices

$$\begin{aligned} \tilde{B}_2 &:= \begin{bmatrix} B_{2_0} & \Sigma_{12}^\mathsf{T} C_1^\mathsf{T} D_{12_\tau} + \Sigma_{22}^\mathsf{T} B_{2_\tau} \end{bmatrix} \\ \tilde{C}_1 &:= \left(P_\tau C_1 + P_0 C_1 \Sigma_{22}^\mathsf{T} - D_{12_0} B_{2_0}^\mathsf{T} \Sigma_{21}^\mathsf{T} \right) \Sigma_{22}^{-\mathsf{T}}, \end{aligned}$$

where the Σ_{ij} are partitioned the same way as the H_{ij}. The modified four-block plant output by Γ is then

$$\tilde{\mathcal{P}} := \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \tilde{\mathcal{P}}_{12} \\ \mathcal{P}_{21} & \tilde{\mathcal{P}}_{22} \end{bmatrix} := \left[\begin{array}{c|cc} A & B_1 & \tilde{B}_2 \\ \hline \tilde{C}_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right],$$

Finally, define the FIR systems

$$\begin{bmatrix} \tilde{\Pi}_u \\ \tilde{\Pi}_b \end{bmatrix} := \pi_\tau \left\{ \left[\begin{array}{cc|c} H_{11} & H_{12} & B_{2_\tau} \\ H_{21} & H_{22} & -C_1^\mathsf{T} D_{12_\tau} \\ \hline D_{12_0}^\mathsf{T} C_1 & B_{2_0} & 0 \\ C_2 & 0 & 0 \end{array}\right] e^{-s\tau} \right\}.$$

FIR outputs of Γ are $\Pi_u := \begin{bmatrix} I & \tilde{\Pi}_u \\ 0 & I \end{bmatrix}$ and $\Pi_b := \begin{bmatrix} 0 & \tilde{\Pi}_b \end{bmatrix}$.

In the general case $D_{12}^\mathsf{T} D_{12} \neq I$, we can use a standard change of variables to transform back to the case $D_{12}^\mathsf{T} D_{12} = I$. See [21, Rem. 2] for details.

B Gramian equations

Here we provide the set of Lyapunov equations that are uniquely associated with the multi-agent problem.

Lemma 20 *Suppose* $(X_{\text{cen}}, F_{\text{cen}})$ *and* (X^i, F^i) *are defined in* (7a) *and* (9a) *respectively. Then* $W_X^i := X^i - X_{\text{cen}_{\underline{ii}}} \succeq 0$ *is the unique solution to the Lyapunov equation*

$$(A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i)^{\mathsf{T}} W_X^i + W_X^i (A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i) + (E_{m_{\underline{i}}} F^i - F_{\text{cen}} E_{n_{\underline{i}}})^{\mathsf{T}} D_{12}^{\mathsf{T}} D_{12} (E_{m_{\underline{i}}} F^i - F_{\text{cen}} E_{n_{\underline{i}}}) = 0. \quad (21)$$

Proof. Left and right multiply the ARE for (7a) by $E_{n_{\underline{i}}}^{\mathsf{T}}$ and $E_{n_{\underline{i}}}$ respectively, and subtract it from (9a). The result follows from algebraic manipulation and applying the definitions of F^i and F_{cen}. Since the final term in (21) is positive semidefinite and $A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i$ is Hurwitz, it follows that $W_X^i := X^i - X_{\text{cen}_{\underline{ii}}} \succeq 0$ and is unique. ∎

We also have a dual analog to Lemma 20, provided below.

Lemma 21 *Consider the setting of Lemma 20. There exists a unique* $W_Y^i \succeq 0$ *that satisfies the Lyapunov equation*

$$(A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i) W_Y^i + W_Y^i (A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i)^{\mathsf{T}} + E_{n_{\underline{i}}}^{\mathsf{T}} \bar{L} \bar{\mathbf{1}}_p D_{21} D_{21}^{\mathsf{T}} \bar{\mathbf{1}}_p^{\mathsf{T}} \bar{L}^{\mathsf{T}} E_{n_{\underline{i}}} = 0. \quad (22)$$

Proof. Since $E_{n_{\underline{i}}}^{\mathsf{T}} \bar{L} \bar{\mathbf{1}}_p D_{21} D_{21}^{\mathsf{T}} \bar{\mathbf{1}}_p^{\mathsf{T}} \bar{L}^{\mathsf{T}} E_{n_{\underline{i}}} \succeq 0$ and the matrix $A_{\underline{ii}} + B_{2_{\underline{ii}}} F^i$ is Hurwitz, $W_Y^i \succeq 0$ and is unique. ∎

C Proof of Theorem 8

For the case $\tau = 0$, we can replace $\mathcal{Q} \in \mathcal{H}_\infty \cap \mathcal{S}_\tau$ by $\mathcal{Q} \in \mathcal{H}_2 \cap \mathcal{H}_\infty \cap \mathcal{S}_0$ in (13) because the closed-loop map must be strictly proper in order to have a finite $\mathcal{H}_2$ norm. Since $\mathcal{T}_{11}$ is strictly proper, this forces $\mathcal{Q}$ to be strictly proper as well, and hence $\mathcal{Q} \in \mathcal{H}_2 \cap \mathcal{H}_\infty$. Further, if $\mathcal{Q}$ is rational, we have $\mathcal{Q} \in \mathcal{RH}_2$. The optimization problem (13) is a least squares problem with a subspace constraint, so the necessary and sufficient conditions for optimality are given by the normal equations $\mathcal{T}_{12}^{\sim}(\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}\mathcal{T}_{21})\mathcal{T}_{21}^{\sim} \in (\mathcal{RH}_2 \cap \mathcal{S}_0)^{\perp}$ with the constraint that $\mathcal{Q} \in \mathcal{RH}_2 \cap \mathcal{S}_0$.

We can check membership $\mathcal{F} \in (\mathcal{RH}_2 \cap \mathcal{S}_0)^{\perp}$ by checking if $\mathcal{F}_{ij} \in \mathcal{RH}_2^{\perp}$ whenever there is a path $j \to i$. For example, consider the two-node graph $1 \to 2$. Then we have

$$\mathcal{RH}_2 \cap \mathcal{S}_0 = \begin{bmatrix} \mathcal{RH}_2 & 0 \\ \mathcal{RH}_2 & \mathcal{RH}_2 \end{bmatrix} \quad \text{and} \quad (\mathcal{RH}_2 \cap \mathcal{S}_0)^{\perp} = \begin{bmatrix} \mathcal{RH}_2^{\perp} & \mathcal{L}_2 \\ \mathcal{RH}_2^{\perp} & \mathcal{RH}_2^{\perp} \end{bmatrix}.$$

So here, $\mathcal{F} \in (\mathcal{RH}_2 \cap \mathcal{S}_0)^{\perp}$ if and only if $\mathcal{F}_{11}, \mathcal{F}_{21}, \mathcal{F}_{22} \in \mathcal{RH}_2^{\perp}$. We will show that the proposed $\mathcal{Q}_{\text{opt}}$ in (15) is optimal by directly verifying the normal equations.

Substituting $\mathcal{Q}_{\text{opt}}$ from (15) into $\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}_{\text{opt}}\mathcal{T}_{21}$ with $\mathcal{T}_{ij}$ defined in (14), we obtain the closed-loop map

$$\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}_{\text{opt}}\mathcal{T}_{21} = \left[\begin{array}{c|c} A_{\text{cl}} & B_{\text{cl}} \\ \hline C_{\text{cl}} & 0 \end{array}\right] := \left[\begin{array}{cc|c} \bar{A} + \bar{B}\bar{F} & -\bar{L}\bar{C}\bar{\mathbf{1}}_n & -\bar{L}\bar{\mathbf{1}}_p D_{21} \\ 0 & A_L & B_L \\ \hline C_1\bar{\mathbf{1}}_n^{\mathsf{T}} + D_{12}\bar{\mathbf{1}}_m^{\mathsf{T}}\bar{F} & C_1 & 0 \end{array}\right], \quad (23)$$

where $A_L := A + LC_2$ and $B_L := B_1 + LD_{21}$. Next, we show that the controllability Gramian for the closed loop map is block-diagonal.

Lemma 22 *The controllability Gramian for the closed-loop map* (23) *is given by*

$$\Theta := \mathrm{blkd}(\{E_{n_{\underline{i}}} W_Y^i E_{n_{\underline{i}}}^\mathsf{T}\}_{i\in[N]}, Y_{\mathrm{cen}}),$$

where Y_{cen} *and* W_Y^i *are defined in Eq.* (7b) *and Lemma 21, respectively. In other words,* $\Theta \succeq 0$ *is the unique solution to* $A_{\mathrm{cl}}\Theta + \Theta A_{\mathrm{cl}}^\mathsf{T} + B_{\mathrm{cl}} B_{\mathrm{cl}}^\mathsf{T} = 0$.

Proof. A_{cl} is Hurwitz and $B_{\mathrm{cl}} B_{\mathrm{cl}}^\mathsf{T} \succeq 0$ so the Lyapunov equation has a unique solution and $\Theta \succeq 0$. We can verify the solution by direct substitution using Lemma 21 and the ARE associated with (7b). ∎

Lemma 22 has the following statistical interpretation. If the controlled system (23) is driven by standard Gaussian noise, its state in these coordinates will have a steady-state covariance Θ, so each block component will be mutually independent.

C.1 Proof of optimality

Let $\Omega := \mathcal{T}_{12}^{\sim}(\mathcal{T}_{11} + \mathcal{T}_{12}\mathcal{Q}_{\mathrm{opt}}\mathcal{T}_{21})\mathcal{T}_{21}^{\sim}$. Substituting $\mathcal{Q}_{\mathrm{opt}}$ from (15) and using (23), we obtain

$$\Omega = \left[\begin{array}{ccc|c} -A_K^\mathsf{T} & -C_K^\mathsf{T} C_{\mathrm{cl}} & 0 & 0 \\ 0 & A_{\mathrm{cl}} & B_{\mathrm{cl}} B_L^\mathsf{T} & B_{\mathrm{cl}} D_{21}^\mathsf{T} \\ 0 & 0 & -A_L^\mathsf{T} & -C_2^\mathsf{T} \\ \hline B_2^\mathsf{T} & D_{12}^\mathsf{T} C_{\mathrm{cl}} & 0 & 0 \end{array}\right], \tag{24}$$

where $A_K := A + B_2 F_d$, $C_K := C_1 + D_{12} F_d$, and A_{cl}, B_{cl}, C_{cl}, are defined in (23). Apply the state transformation

$$T = \begin{bmatrix} I & \begin{bmatrix}\bar{\mathbf{1}}_n^\mathsf{T}\bar{X} & 0\end{bmatrix} & \bar{\mathbf{1}}_n^\mathsf{T}\bar{X}\bar{W}\bar{\mathbf{1}}_p \\ 0 & I & \Theta\bar{\mathbf{1}}_p \\ 0 & 0 & I \end{bmatrix}$$

to (24), where we defined $\bar{W} := \mathrm{blkd}(\{E_{n_{\underline{i}}} W_Y^i E_{n_{\underline{i}}}^\mathsf{T}\}_{i\in[N]})$ and $\bar{X} := \mathrm{blkd}(\{E_{n_{\underline{i}}} W_X^i E_{n_{\underline{i}}}^\mathsf{T} + X_{\mathrm{cen}}\}_{i\in[N]})$, and Θ is defined in Lemma 22. The transformed Ω is

$$\Omega = \left[\begin{array}{ccccc} -A_K^\mathsf{T} & \star_1 & \star & \star & \star \\ 0 & \bar{A} + \bar{B}\bar{F} & -\bar{L}\bar{C}\bar{\mathbf{1}}_n & \star_2 & \star_3 \\ 0 & 0 & A_L & \star_5 & \star_6 \\ 0 & 0 & 0 & -A_L^\mathsf{T} & -C_2^\mathsf{T} \\ \hline B_2^\mathsf{T} & \star_4 & D_{12}^\mathsf{T} C_1 & \star & 0 \end{array}\right],$$

where we have defined the symbols

$$\begin{aligned}
\star_1 &:= -A_K^\mathsf{T}\bar{\mathbf{1}}_n^\mathsf{T}\bar{X} - C_K^\mathsf{T}(C_1\bar{\mathbf{1}}_n^\mathsf{T} + D_{12}\bar{\mathbf{1}}_m^\mathsf{T}\bar{F}) - \bar{\mathbf{1}}_n^\mathsf{T}\bar{X}(\bar{A} + \bar{B}\bar{F}) \\
\star_2 &:= -\bar{L}\bar{\mathbf{1}}_p D_{21} B_L^\mathsf{T} - \bar{L}\bar{C}\bar{\mathbf{1}}_n Y_{\mathrm{cen}} + (\bar{A} + \bar{B}\bar{F})\bar{W}\bar{\mathbf{1}}_p + \bar{W}\bar{\mathbf{1}}_p A_L^\mathsf{T} \\
\star_3 &:= -\bar{L}\bar{\mathbf{1}}_p D_{21}^\mathsf{T} + \bar{W}\bar{\mathbf{1}}_p C_2^\mathsf{T} \\
\star_4 &:= D_{12}^\mathsf{T}(C_1\bar{\mathbf{1}}_n^\mathsf{T} + D_{12}\bar{\mathbf{1}}_m^\mathsf{T}\bar{F}) + B_2^\mathsf{T}\bar{\mathbf{1}}_n^\mathsf{T}\bar{X} \\
\star_5 &:= A_L Y_{\mathrm{cen}} + B_L B_L^\mathsf{T} + Y_{\mathrm{cen}} A_L^\mathsf{T} \\
\star_6 &:= B_L D_{21}^\mathsf{T} + Y_{\mathrm{cen}} C_2^\mathsf{T}.
\end{aligned}$$

A $\star$ without subscript denotes an unimportant block. Simplifying using Riccati and Lyapunov equations from Section 2.2.1 and Appendix B respectively, we get $\star_5 = \star_6 = 0$; the mode A_L is uncontrollable. Removing it, we obtain

$$\Omega = \left[\begin{array}{ccc|c} -A_K^\mathsf{T} & \star_1 & \star & \star \\ 0 & \bar{A}+\bar{B}\bar{F} & \star_2 & \star_3 \\ 0 & 0 & -A_L^\mathsf{T} & -C_2^\mathsf{T} \\ \hline B_2^\mathsf{T} & \star_4 & \star & 0 \end{array}\right]. \tag{25}$$

Now consider a block Ω_{ij} for which there is a path $j \to i$.

$$\Omega_{ij} = \left[\begin{array}{ccc|c} -A_{K_{ii}}^\mathsf{T} & \star_{1_{i:}} & \star & \star \\ 0 & \bar{A}+\bar{B}\bar{F} & \star_{2_{:j}} & \star_{3_{:j}} \\ 0 & 0 & -A_{L_{jj}}^\mathsf{T} & -C_{2_{jj}}^\mathsf{T} \\ \hline B_{2_{ii}}^\mathsf{T} & \star_{4_{i:}} & \star & 0 \end{array}\right]. \tag{26}$$

Let $\star_1^k$ and $\star_4^k$ denote the k^{th} block column and let $\star_2^k$ and $\star_3^k$ denote the k^{th} block row. Algebraic manipulation reveals that

(i) If $i \in \underline{k}$ and $\ell \in \underline{k}$, then $\star_{1_{i\ell}}^k = \star_{4_{i\ell}}^k = 0$.

(ii) If $\ell \notin \underline{k}$ or $j \notin \underline{k}$, then $\star_{2_{\ell j}}^k = \star_{3_{\ell j}}^k = 0$.

Consider the k^{th} diagonal block of $\bar{A}+\bar{B}\bar{F}$ in (26), which is $A+E_{n_{\underline{k}}}B_{2_{\underline{k}\underline{k}}}F^kE_{n_{\underline{k}}}^\mathsf{T}$. This block is itself block-diagonal; it contains the block $A_{\underline{k}\underline{k}}+B_{2_{\underline{k}\underline{k}}}F^k$ and smaller blocks $A_{\ell\ell}$ for all $\ell \notin \underline{k}$. We have three cases.

1. If $k \in \bar{i}$, then for all $\ell \in \underline{k}$, we have $\star_{1_{i\ell}}^k = \star_{4_{i\ell}}^k = 0$ from Item (i) above, so the mode $A_{\underline{k}\underline{k}}+B_{2_{\underline{k}\underline{k}}}F^k$ is unobservable.
2. If $k \in \bar{i}$, but instead $\ell \notin \underline{k}$, we have $\star_{2_{\ell j}}^k = \star_{3_{\ell j}}^k = 0$ from Item (ii) above, so the modes $A_{\ell\ell}$ are uncontrollable.
3. If $k \notin \bar{i}$, then $k \notin \bar{j}$ because $j \to i$ by assumption. Then from Item (ii) above, all such modes are uncontrollable.

Consequently every block of $\bar{A}+\bar{B}\bar{F}$ is either uncontrollable or unobservable, leading us to the reduced realization

$$\Omega_{ij} = \left[\begin{array}{cc|c} -A_{K_{ii}}^\mathsf{T} & \star & \star \\ 0 & -A_{L_{jj}}^\mathsf{T} & -C_{2_{jj}}^\mathsf{T} \\ \hline B_{2_{ii}}^\mathsf{T} & \star & 0 \end{array}\right]. \tag{27}$$

Therefore, $\Omega_{ij} \in \mathcal{RH}_2^\perp$ whenever $j \to i$, as required.

D Proof of Theorem 12

Start with the convexified optimization problem (13). Based on the structured realization (14), we see that $\mathcal{T}_{21}$ is block-diagonal. Therefore, the optimal cost can be split by columns:

$$\left\|\mathcal{T}_{11}+\mathcal{T}_{12}\mathcal{Q}\mathcal{T}_{21}\right\|_2^2 = \sum_{i=1}^{N}\left\|\mathcal{T}_{11_{:i}}+\mathcal{T}_{12_{:\bar{i}}}\mathcal{Q}_{\bar{i}\bar{i}}\mathcal{T}_{21_{ii}}\right\|_2^2.$$

Since $\mathcal{Q} \in \mathcal{H}_\infty \cap \mathcal{S}_\tau$, we can factor each block column of $\mathcal{Q}$ as $\mathcal{Q}_{\underline{i}i} = \Lambda_m^i \tilde{\mathcal{Q}}_{\underline{i}i}$, where $\tilde{\mathcal{Q}}_{\underline{i}i} \in \mathcal{H}_\infty$ has no structure or delay, and Λ_m^i is the adobe delay matrix (defined in Section 2.1). We can therefore optimize for each block column $\tilde{\mathcal{Q}}_{\underline{i}i}$ separately. Thus, each subproblem is to

$$\underset{\tilde{\mathcal{Q}}_{\underline{i}i} \in \mathcal{H}_\infty}{\text{minimize}} \qquad \left\| \mathcal{T}_{11_{:i}} + \mathcal{T}_{12_{:\underline{i}}} \Lambda_m^i \tilde{\mathcal{Q}}_{\underline{i}i} \mathcal{T}_{21_{ii}} \right\|_2^2, \tag{28}$$

Define $\mathcal{T}_i := \begin{bmatrix} \mathcal{T}_{11_{:i}} & \mathcal{T}_{12_{:\underline{i}}} \\ \mathcal{T}_{21_{ii}} & 0 \end{bmatrix}$. Comparing to (13)–(14), we observe that (28) is a special case of the problem (13), subject to the transformations $\mathcal{P} \mapsto \mathcal{P}_i$ (defined in (8)) and $F_d \mapsto F_{d_{\underline{i}i}}$, $L_d \mapsto E_{n_{\underline{i}}}^\mathsf{T} E_{n_i} L^i$, and $\mathcal{Q} \mapsto \Lambda_m^i \tilde{\mathcal{Q}}_{\underline{i}i}$. If we define the associated $\mathcal{J}_i$ for this subproblem (according to (11)), we view the subproblem as that of finding the $\mathcal{H}_2$-optimal controller for the plant $\mathcal{P}_i$ subject to an adobe input delay, as illustrated in the left panel of Fig. 11. The key difference between this problem and (6) is that we no longer have a sparsity constraint.

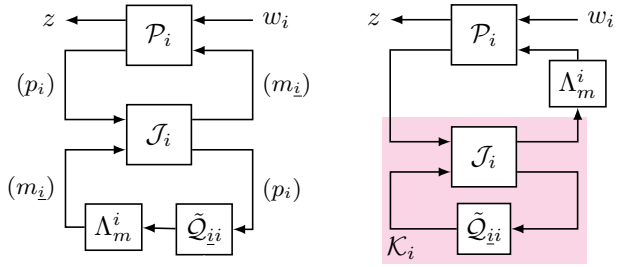

Fig. 11: Equivalent subproblems via commuting Λ_m^i and $\mathcal{J}_i$. Dimensions of signals are indicated along the arrows.

The adobe delay Λ_m^i can be shifted to the input channel, shown in the right panel of Fig. 11. This follows from leveraging state-space properties and the block structure of certain blocks of $\mathcal{J}_i$. Examples include $B_{2_{\underline{i}i}} \Lambda_m^i = \Lambda_n^i B_{2_{\underline{i}i}}$ and $\Lambda_n^i E_{n_{\underline{i}}}^\mathsf{T} E_{n_i} L^i = E_{n_{\underline{i}}}^\mathsf{T} E_{n_i} L^i$.

The remainder of the proof proceeds as follows: we define $\mathcal{K}_i$ to be the shaded system in Fig. 11 (right panel). This is a standard adobe delayed problem, so we can apply the Γ transformation illustrated in Fig. 2. Specifically, we define $(\tilde{\mathcal{P}}_i, \Pi_{u_i}, \Pi_{b_i}) = \Gamma(\mathcal{P}_i, \Lambda_m^i)$, and obtain Fig. 12.

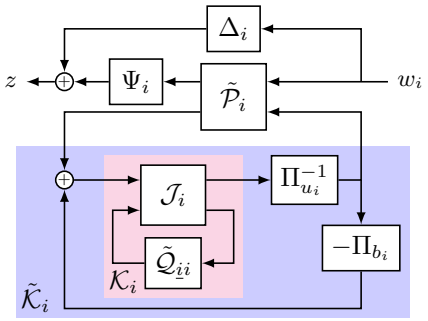

Fig. 12: Transformation of the right panel of Fig. 11 using the loop-shifting transformation illustrated in Fig. 2.

By the properties of the loop-shifting transformation discussed in Section 2.1, the optimal $\tilde{K}_i$ is found by solving a standard non-delayed LQG problem in the (rational) plant $\tilde{\mathcal{P}}_i$, whose solution

is

$$\tilde{\mathcal{K}}_i = \left[\begin{array}{c|c} A_{\underline{i}\underline{i}} + \tilde{B}_{2_{\underline{i}\underline{i}}}\tilde{F}^i + E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i C_{2_{i\underline{i}}} & -E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i \\ \hline \tilde{F}^i & 0 \end{array}\right].$$

Inverting each transformation, $\mathcal{K}_i = \Pi_{u_i}\tilde{\mathcal{K}}_i(I - \Pi_{b_i}\tilde{\mathcal{K}}_i)^{-1}$, and we can recover the Youla parameter via $\tilde{\mathcal{Q}}_{\underline{i}i} = \mathcal{F}_u(\mathcal{J}_i^{-1}, \mathcal{K}_i)$, which leads to (29). Now zero-pad, reintroduce delays, and concatenate,

$$\tilde{\mathcal{Q}}_{\underline{i}i} = \left[\begin{array}{cc|c} A_{\underline{i}\underline{i}} & B_{2_{\underline{i}\underline{i}}}\Pi_{u_i}\tilde{F}^i & -E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i \\ -E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i C_{2_{i\underline{i}}} & A_{\underline{i}\underline{i}} + \tilde{B}_{2_{\underline{i}\underline{i}}}\tilde{F}^i - E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i \Pi_{b_i}\tilde{F}^i + E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i C_{2_{i\underline{i}}} & -E_{n_{\underline{i}}}^{\mathsf{T}} E_{n_i} L^i \\ \hline -F_{d_{\underline{i}\underline{i}}} & \Pi_{u_i}\tilde{F}^i & 0 \end{array}\right]. \quad (29)$$

to obtain the global Youla parameter (17) via $\mathcal{Q}_{\text{opt}} = \sum_{i=1}^{N} E_{m_{\underline{i}}} \Lambda_m^i \tilde{\mathcal{Q}}_{\underline{i}i} E_{p_i}^{\mathsf{T}}$ and recover the optimal controller (18) via $\mathcal{K}_{\text{opt}} = \mathcal{F}_l(\mathcal{J}, \mathcal{Q}_{\text{opt}})$.

E Proof of Theorem 13

From Lemma 3, the set of sub-optimal controllers is parameterized as $\mathcal{K} = \mathcal{F}_l(\mathcal{J}, \mathcal{Q})$, where $\mathcal{Q} \in \mathcal{S}_\tau$. Equivalently, write $\mathcal{K} = \mathcal{F}_l(\mathcal{J}, \mathcal{Q}_{\text{opt}} + \mathcal{Q}_\Delta)$, where $\mathcal{Q}_\Delta \in \mathcal{S}_\tau$ and $\mathcal{Q}_{\text{opt}}$ is given in Theorem 12. The controller equation $u = \mathcal{K}y$ can be expanded using the LFT as $\left(\begin{smallmatrix} u \\ \eta \end{smallmatrix}\right) = \mathcal{J}\left(\begin{smallmatrix} y \\ v \end{smallmatrix}\right)$ with $v = \mathcal{Q}\eta$. If $\mathcal{J}$ has state ξ, the state-space equation for $\mathcal{J}$ decouples as

$$\begin{aligned} \dot{\xi}_i &= (A_{ii} + B_{2_{ii}}F_{d_{ii}} + L^i C_{2_{ii}})\xi_i - L^i y_i + B_{2_{ii}} v_i, \\ u_i &= F_{d_{ii}}\xi_i + v_i, \\ \eta_i &= -C_{2_{ii}}\xi_i + y_i, \quad \text{for } i = 1, \ldots, N. \end{aligned}$$

Note that we replaced $L_{d_{ii}}$ by L^i from (9b). This leads to simpler algebra, but is in principle not required. Meanwhile, the $\mathcal{Q}$ equation is coupled: $v = (\mathcal{Q}_{\text{opt}} + \mathcal{Q}_\Delta)\eta$. Now consider Agent i. Since we are interested in the agent-level implementation, we begin by extracting u_i, which requires finding v_i. Separate $\mathcal{Q}$ by columns as in Appendix D to obtain

$$\begin{aligned} v_i &= E_{m_i}^{\mathsf{T}} \left(\mathcal{Q}_{\text{opt}} + \mathcal{Q}_\Delta\right) \eta \\ &= \sum_{k \in [N]} E_{m_i}^{\mathsf{T}} E_{m_{\underline{k}}} \Lambda_m^k \left(\tilde{\mathcal{Q}}_{\underline{k}k} + \hat{\mathcal{Q}}_{\underline{k}k}\right) \eta_k \\ &= \left(\tilde{\mathcal{Q}}_{ii} + \hat{\mathcal{Q}}_{ii}\right) \eta_i + e^{-s\tau} \sum_{k \in \overline{\overline{i}}} \left(\tilde{\mathcal{Q}}_{ik} + \hat{\mathcal{Q}}_{ik}\right) \eta_k, \end{aligned} \quad (30)$$

where $\tilde{\mathcal{Q}}_{\underline{i}i}$ is given in (29), and $\hat{\mathcal{Q}} \in \mathcal{S}_0$ is the delay-free component of $\mathcal{Q}_\Delta$. A possible distributed implementation is to have Agent i simulate ξ_i locally. Since y_i is available locally, then so is η_i. We further suppose Agent i computes $v_{i,\underline{i}} := (\tilde{\mathcal{Q}}_{\underline{i}i} + \hat{\mathcal{Q}}_{\underline{i}i})\eta_i$ locally. Component $v_{i,i}$ is used locally, while component $v_{i,j}$ for $j \in \underline{\underline{i}}$ is transmitted to descendant j. Each agent then computes v_i by summing its local $v_{i,i}$ with the delayed $e^{-s\tau}v_{i,k}$ received from strict ancestors $k \in \overline{\overline{i}}$. The complete agent-level implementation is shown in Fig. 3.

When $\hat{\mathcal{Q}} = 0$, we recover the optimal controller. In this case, the equations simplify considerably; standard state-space manipulations reduce Fig. 3 to the simpler Fig. 4. It is worth noting that the optimal controller does not depend on the choice of nominal gain F_d.

F Proof of Theorem 17

All the estimation, control gains and Riccati solutions used here are defined in Section 2.2.1. The additional cost incurred due to suboptimality is $J_Q := \left\|\mathcal{T}_{12}\mathcal{Q}_\Delta\mathcal{T}_{21}\right\|_2^2$ [37, §14.6]. Using [37, Lem. 14.3], we have $J_Q := \left\|\mathcal{T}_{12}\mathcal{Q}_\Delta D_{21}\right\|_2^2$.

F.1 $J_{\mathbf{cen}}$ (19a)

The optimal cost for a fully connected graph [37, Thm. 14.7] is

$$\begin{aligned} J_{\text{cen}} &:= \left\| \left[\begin{array}{c|c} A+B_2F_{\text{cen}} & B_1 \\ \hline C_1+D_{12}F_{\text{cen}} & 0 \end{array}\right] \right\|_2^2 + \left\| \left[\begin{array}{c|c} A_L & B_L \\ \hline D_{12}F_{\text{cen}} & 0 \end{array}\right] \right\|_2^2, \\ &= \text{tr}(Y_{\text{cen}}C_1^\mathsf{T}C_1) + \text{tr}(X_{\text{cen}}LD_{21}D_{21}^\mathsf{T}L^\mathsf{T}), \\ &= \text{tr}(X_{\text{cen}}B_1B_1^\mathsf{T}) + \text{tr}(Y_{\text{cen}}F_{\text{cen}}^\mathsf{T}D_{12}^\mathsf{T}D_{12}F_{\text{cen}}), \end{aligned}$$

where A_L, B_L are defined in Appendix C for (23).

F.2 $J_{\mathbf{dec}}$ (19b)

Consider that $\mathcal{K}_{\text{opt}}$ in (16) is a sub-optimal centralized controller for $\left\|\mathcal{T}_{11}+\mathcal{T}_{12}\mathcal{Q}\mathcal{T}_{21}\right\|_2^2$, subject to $\mathcal{Q} \in \mathcal{RH}_2$. Centralized $\mathcal{H}_2$ theory [37] implies that $J_{\text{dec}} = J_{\text{cen}} + \Delta$, where $\Delta := \left\|D_{12}\mathcal{Q}_{\text{you}}D_{21}\right\|_2^2$ and $\mathcal{Q}_{\text{you}}$ is the centralized Youla parameter. Here, $\mathcal{Q}_{\text{you}} = \mathcal{F}_u(\mathcal{J}^{-1}, \mathcal{K}_{\text{opt}})$, where

$$\mathcal{J}^{-1} = \left[\begin{array}{c|cc} A & B_2 & -L \\ \hline C_2 & 0 & I \\ -F_{\text{cen}} & I & 0 \end{array}\right].$$

After simplifications, we obtain

$$\mathcal{Q}_{\text{you}} = \left[\begin{array}{c|c} \bar{A}+\bar{B}\bar{F} & -\bar{L}\bar{\mathbf{1}}_p \\ \hline \bar{\mathbf{1}}_m^\mathsf{T}(\bar{F}-\bar{F}_{\text{cen}}) & 0 \end{array}\right].$$

We substitute $\mathcal{Q}_{\text{you}}$ into the expression for Δ, using $\left\|D_s + C_s(sI-A_s)^{-1}B_s\right\|_2^2 = \text{tr}(C_sW_cC_s^\mathsf{T})$, where W_c is the controllability Gramian given by Lyapunov equation $A_sW_c + W_cA_s^\mathsf{T} + B_sB_s^\mathsf{T} = 0$. Based on the Lemma 20 and using the identity $L_i = E_{n_i}L^iE_{p_i}^\mathsf{T}$, we evaluate

$$\begin{aligned} \Delta &= \sum_{i=1}^{N} \text{tr}(D_{21}^\mathsf{T}L_i^\mathsf{T}E_{n_{\underline{i}}}\{X^i - X_{\text{cen}_{\underline{ii}}}\}E_{n_{\underline{i}}}^\mathsf{T}L_iD_{21}) \\ &= \text{tr}((\text{blkd}(\{X^i(1,1)\}) - X_{\text{cen}})LD_{21}D_{21}^\mathsf{T}L^\mathsf{T}). \end{aligned}$$

We obtain (19b) by substituting Δ into $J_{\text{dec}} = J_{\text{cen}} + \Delta$.

F.3 Alternative formulas for the cost

We obtained an alternative formula for J_{cen} in Appendix F.1. Similarly, in Appendix F.2 for J_{dec}, $\left\|D_s + C_s(sI - A_s)^{-1}B_s\right\|_2^2$ is also equal to $\text{tr}(B_s B_s^\mathsf{T} W_o)$, where W_o is the observability Gramian given by the dual Lyapunov equation $A_s^\mathsf{T} W_o + W_o A_s + C_s^\mathsf{T} C_s = 0$. Based on Lemma 21, we can evaluate $\Delta = \sum_{i=1}^N \text{tr}(D_{12}(E_{m_{\underline{i}}} F_i - F_{\text{cen}} E_{n_{\underline{i}}}) W_Y^i (E_{m_{\underline{i}}} F_i - F_{\text{cen}} E_{n_{\underline{i}}})^\mathsf{T} D_{12}^\mathsf{T})$. Similar alternative formulas exist for (19c), and (19d) as well.

F.4 $J_{\text{dec,del}}$ (19c)

We can split the cost in (13) into a sum of N separate terms because $\mathcal{T}_{21}$ is block-diagonal. Using [23, Prop. 6] on each of these N problems, we write $J_{\text{dec,del}}$ as a combination of a non-delayed cost J_{dec} and a Δ incurred by adding delays to that system: $J_{\text{dec,del}} = J_{\text{dec}} + \Delta$, where $\Delta := \sum_{i=1}^N \text{tr}(D_{21_{\underline{i}i}}^\mathsf{T} L_i^\mathsf{T} E_{n_{\underline{i}}}^\mathsf{T} (\Xi_\tau^i - X^i) E_{n_{\underline{i}}} L_i D_{21_{\underline{i}i}})$. Also, $\Delta = \text{tr}(\text{blkd}(\{\Xi_\tau^i(1,1) - X^i(1,1)\}) L D_{21} D_{21}^\mathsf{T} L^\mathsf{T})$ since $L_i = E_{n_i} L^i E_{p_i}^\mathsf{T}$. We obtain (19c) by substituting Δ into $J_{\text{dec,del}} = J_{\text{dec}} + \Delta$. See Appendix F.6 below for explanation on Ξ_τ^i.

F.5 J_{del} (19d)

Derivation is analogous to that of $J_{\text{dec,del}}$. See Appendix F.7 below for explanation on $\Xi_{c_\tau}^i$.

F.6 Proofs for (20a)

We have $X^i - X_{\text{cen}_{\underline{i}\underline{i}}} \succeq 0$ in Lemma 20 for all $i \in [N]$. The properties of a positive semi-definite matrix give us $X^i(1,1) - X_{\text{cen}_{\underline{i}\underline{i}}}(1,1) \succeq 0$, and hence $\text{blkd}(\{X_{\text{cen}}(i,i)\}) \preceq X_{\text{dec}}$.

Now we define Ξ_τ^i and establish that $\Xi_\tau^i - X^i \succeq 0$. The Hamiltonian for the control Riccati equation (10) is

$$H^i := \begin{bmatrix} A_{\underline{i}\underline{i}} - \tilde{B}_{2_{\underline{i}\underline{i}}} M^{-1} D_{12:i}^\mathsf{T} \tilde{C}_{1:\underline{i}} & -\tilde{B}_{2_{\underline{i}\underline{i}}} M^{-1} \tilde{B}_{2_{\underline{i}\underline{i}}}^\mathsf{T} \\ -\tilde{C}_{1:\underline{i}}^\mathsf{T} P_\tau \tilde{C}_{1:\underline{i}} & -A_{\underline{i}\underline{i}}^\mathsf{T} + \tilde{C}_{1:\underline{i}}^\mathsf{T} D_{12:i} M^{-1^\mathsf{T}} \tilde{B}_{2_{\underline{i}\underline{i}}}^\mathsf{T} \end{bmatrix},$$

where $M := D_{12:i}^\mathsf{T} D_{12:i}$, $P_0 := D_{12:i} M^{-1} D_{12:i}^\mathsf{T}$ and $P_\tau := I - P_0$, and define the corresponding symplectic matrix exponential as $\Sigma^i := e^{H^i \tau}$. The elements Σ_{22}^i, Σ_{21}^i of this modified Σ^i are used to define the Ξ_τ^i. For all $i \in [N]$, we define $\Xi_\tau^i := \tilde{X}^i - (\Sigma_{22}^{i^{-1}} \Sigma_{21}^i)^\mathsf{T}$. By solving the associated Differential Riccati Equation (DRE) [23, Eq. 16], we show $\Xi_\tau^i - X^i \succeq 0$ [23, §4.3]. This gives us $X_{\text{dec}} \preceq X_{\text{dec,del}}$.

F.7 Proofs for (20b)

Next we consider the case of a fully connected graph with delays. So Agent i's feedback policy looks like $u_i = \mathcal{K}_{ii}(s) y_i + \sum_{j \in [N] \setminus i} e^{-s\tau} \mathcal{K}_{ij}(s) y_j$. Since we solve for $\mathcal{Q}$ by solving for individual columns $\mathcal{Q}_{\underline{i}i}$, we define the associated state transition matrix for each column as $A_{\underline{i}\underline{i}}^c := \text{blkd}(\{A_{ii}, A_{\underline{i}\underline{i}}\})$, where $\underline{i} = [N] \setminus i$. We define the corresponding $B_{1_{\underline{i}i}}^c$, $B_{2_{\underline{i}\underline{i}}}^c$ $C_{1:\underline{i}}^c$, $D_{12:\underline{i}}^c$, $C_{2_{\underline{i}\underline{i}}}^c$, and $D_{21_{\underline{i}i}}^c$ in a similar

manner. We also define a centralized $\Xi^i_{c_\tau} := \tilde{X}^i_c - (\Sigma^{i^{-1}}_{22_c}\Sigma^i_{21_c})^\mathsf{T}$ for each Γ-modified plant

$$\tilde{\mathcal{P}}^c_i := \left[\begin{array}{c|cc} A^c_{\underline{ii}} & B^c_{1_{\underline{ii}}} & \tilde{B}^c_{2_{\underline{ii}}} \\ \hline \tilde{C}^c_{1_{:\underline{i}}} & 0 & D^c_{12_{:\underline{i}}} \\ C^c_{2_{\underline{ii}}} & D^c_{21_{ii}} & 0 \end{array}\right].$$

Each individual column $\mathcal{Q}_{\underline{ii}}$ has its own $\tilde{\mathcal{P}}^c_i$ as the associated adobe delay matrix is different. We have a corresponding control ARE $(\tilde{X}^i_c, \tilde{F}^i_c) := \mathrm{Ric}(A^c_{\underline{ii}}, \tilde{B}^c_{2_{\underline{ii}}}, \tilde{C}^c_{1_{:\underline{i}}}, D^c_{12_{:\underline{i}}})$. We solve DREs for each $\Xi^i_{c_\tau}$ as in [23, §V.C] to obtain $\Xi^i_{c_\tau} - X^c_{\mathrm{cen}_{\underline{ii}}} \succeq 0$ for all $i \in [N]$, where $X^c_{\mathrm{cen}_{\underline{ii}}}$ is a reshuffling of X_{cen} to mirror the ordering of $\underline{i} = \{i, [N] \setminus i\}$. This proves that $\mathrm{blkd}(\{X_{\mathrm{cen}}(i,i)\}) \preceq X_{\mathrm{cen,del}}$ for all $i \in [N]$.

Lemma 23 proves that $X_{\mathrm{cen,del}} \preceq X_{\mathrm{dec,del}}$ for all $i \in [N]$.

Lemma 23 *$\Xi^i_{c_\tau}$ and Ξ^i_τ are the solutions of the DREs for delayed fully connected and decentralized graphs respectively. Then, $W^i_\Xi := \Xi^i_\tau - \Xi^i_{c_{\underline{ii}}} \succeq 0$, where $\Xi^i_{c_{\underline{ii}}} := E^\mathsf{T}_{n_{\underline{i}}}\Xi^i_{c_\tau}E_{n_{\underline{i}}}$, and $\underline{i}$ corresponds to Ξ^i_τ.*

Proof. The DREs for Ξ^i_τ, and $\Xi^i_{c_\tau}$ are subtracted to obtain the differential Lyapunov equation

$$\begin{aligned}\dot{\Xi}^i_{c_{\underline{ii}}} - \dot{\Xi}^i_\tau = (A_{\underline{ii}} + B_{2_{\underline{ii}}}F^i_\Xi)^\mathsf{T}W^i_\Xi + W^i_\Xi(A_{\underline{ii}} + B_{2_{\underline{ii}}}F^i_\Xi) \\ + (E_{m_{\underline{i}}}F^i_\Xi - F^i_{\Xi_c}E_{n_{\underline{i}}})^\mathsf{T}D^\mathsf{T}_{12}D_{12}(E_{m_{\underline{i}}}F^i_\Xi - F^i_{\Xi_c}E_{n_{\underline{i}}}),\end{aligned}$$

where $F^i_\Xi := -(D^\mathsf{T}_{12_{:i}}D_{12_{:i}})^{-1}(\Xi^i_\tau B_{2_{\underline{ii}}} + C^\mathsf{T}_{1_{:\underline{i}}}D_{12_{:i}})^\mathsf{T}$, and $F^i_{\Xi_c} := -(D^{c^\mathsf{T}}_{12_{:i}}D^c_{12_{:i}})^{-1}(\Xi^i_{c_\tau}B^c_{2_{\underline{ii}}} + C^{c^\mathsf{T}}_{1_{:\underline{i}}}D^c_{12_{:i}})^\mathsf{T}$. The rest is analogous to the proof of Lemma 20. Finally, we obtain $\Xi^i_\tau - \Xi^i_{c_{\underline{ii}}} - X^i + X_{\mathrm{cen}_{\underline{ii}}} \succeq 0$. Using $X^i - X_{\mathrm{cen}_{\underline{ii}}} \succeq 0$ from Lemma 20, we obtain $\Xi^i_\tau - \Xi^i_{c_{\underline{ii}}} \succeq 0$. ∎

References

[1] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000. (Cited on p. 3)

[2] J. H. Cho and M. Kristalny. On the H^2 decentralized controller synthesis for delayed bilateral teleoperation systems. *IFAC Proceedings Volumes*, 45(22):393–398, 2012. (Cited on p. 4)

[3] G. E. Dullerud and F. Paganini. *A course in robust control theory: a convex approach*, volume 36. Springer Science & Business Media, 2013. (Cited on pp. 6 and 8)

[4] Y.-C. Ho and K.-C. Chu. Team decision theory and information structures in optimal control problems—Part I. *IEEE Transactions on Automatic Control*, 17(1):15–22, 1972. (Cited on p. 3)

[5] M. Kashyap. *Optimal Decentralized Control with Delays*. PhD thesis, Northeastern University, 2023. (Cited on p. 16)

[6] M. Kashyap and L. Lessard. Explicit agent-level optimal cooperative controllers for dynamically decoupled systems with output feedback. In *IEEE Conference on Decision and Control*, pages 8254–8259, 2019. (Cited on pp. 3, 4, and 10)

[7] M. Kashyap and L. Lessard. Agent-level optimal LQG control of dynamically decoupled systems with processing delays. In *IEEE Conference on Decision and Control*, pages 5980–5985, 2020. (Cited on pp. 4 and 13)

[8] J.-H. Kim and S. Lall. Explicit solutions to separable problems in optimal cooperative control. *IEEE Transactions on Automatic Control*, 60(5):1304–1319, 2015. (Cited on pp. 3 and 7)

[9] J.-H. Kim, S. Lall, and C.-K. Ryoo. Optimal cooperative control of dynamically decoupled systems. In *IEEE Conference on Decision and Control*, pages 4852–4857, 2012. (Cited on p. 3)

[10] M. Kristalny and J. H. Cho. On the decentralized H^2 optimal control of bilateral teleoperation systems with time delays. In *IEEE Conference on Decision and Control*, pages 6908–6914, 2012. (Cited on p. 4)

[11] M. Kristalny and J. H. Cho. Decentralized H^2 optimal control of haptic interfaces for a shared virtual environment. In *IEEE Conference on Decision and Control*, pages 5204–5209, 2013. (Cited on p. 4)

[12] J. F. Kurose and K. W. Ross. *Computer Networking: A top-down approach*. Pearson, 8 edition, 2021. (Cited on p. 2)

[13] A. Lamperski and J. C. Doyle. The $\mathcal{H}_2$ control problem for quadratically invariant systems with delays. *IEEE Transactions on Automatic Control*, 60(7):1945–1950, 2015. (Cited on p. 3)

[14] A. Lamperski and L. Lessard. Optimal decentralized state-feedback control with sparsity and delays. *Automatica*, 58:143–151, 2015. (Cited on p. 3)

[15] L. Lessard. Decentralized LQG control of systems with a broadcast architecture. In *IEEE Conference on Decision and Control*, pages 6241–6246, 2012. (Cited on p. 3)

[16] L. Lessard, M. Kristalny, and A. Rantzer. On structured realizability and stabilizability of linear systems. In *American Control Conference*, pages 5784–5790, 2013. (Cited on p. 8)

[17] L. Lessard and S. Lall. An algebraic approach to the control of decentralized systems. *IEEE Transactions on Control of Network Systems*, 1(4):308–317, 2014. (Cited on p. 8)

[18] L. Lessard and S. Lall. Optimal control of two-player systems with output feedback. *IEEE Transactions on Automatic Control*, 60(8):2129–2144, 2015. (Cited on pp. 3 and 15)

[19] D. Madjidian and L. Mirkin. H_2 optimal cooperation of homogeneous agents subject to delyed information exchange. *IFAC-PapersOnLine*, 49(10):147–152, 2016. (Cited on p. 4)

[20] L. Mirkin. On the extraction of dead-time controllers and estimators from delay-free parametrizations. *IEEE Transactions on Automatic Control*, 48(4):543–553, 2003. (Cited on p. 3)

[21] L. Mirkin, Z. J. Palmor, and D. Shneiderman. Loop shifting for systems with adobe input delay. *IFAC Proceedings Volumes*, 42(6):307–312, 2009. (Cited on pp. 3, 5, 6, and 17)

[22] L. Mirkin, Z. J. Palmor, and D. Shneiderman. Dead-time compensation for systems with multiple I/O delays: A loop-shifting approach. *IEEE Transactions on Automatic Control*, 56(11):2542–2554, 2011. (Cited on pp. 3, 5, and 6)

[23] L. Mirkin, Z. J. Palmor, and D. Shneiderman. H^2 optimization for systems with adobe input delays: A loop shifting approach. *Automatica*, 48(8):1722–1728, 2012. (Cited on pp. 3, 5, 6, 7, 15, 17, 24, and 25)

[24] L. Mirkin and N. Raskin. Every stabilizing dead-time controller has an observer–predictor-based structure. *Automatica*, 39(10):1747–1754, 2003. (Cited on p. 3)

[25] X. Qi, M. V. Salapaka, P. G. Voulgaris, and M. Khammash. Structured optimal and robust control with multiple criteria: A convex solution. *IEEE Transactions on Automatic Control*, 49(10):1623–1640, 2004. (Cited on p. 3)

[26] M. Rotkowitz, R. Cogill, and S. Lall. Convexity of optimal control over networks with delays and arbitrary topology. *International Journal of Systems, Control and Communications*, 2(1-3):30–54, 2010. (Cited on pp. 3, 8, and 9)

[27] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 50(12):1984–1996, 2005. (Cited on pp. 3 and 8)

[28] M. Rotkowitz and S. Lall. Convexification of optimal decentralized control without a stabilizing controller. In *International Symposium on Mathematical Theory of Networks and Systems*, pages 1496–1499, 2006. (Cited on p. 3)

[29] C. W. Scherer. Structured finite-dimensional controller design by convex optimization. *Linear Algebra and its Applications*, 351–352:639–669, 2002. (Cited on p. 3)

[30] P. Shah and P. A. Parrilo. $\mathcal{H}_2$-optimal decentralized control over posets: A state-space solution for state-feedback. *IEEE Transactions on Automatic Control*, 58(12):3084–3096, 2013. (Cited on p. 10)

[31] T. Tanaka and P. A. Parrilo. Optimal output feedback architecture for triangular LQG problems. In *American Control Conference*, pages 5730–5735, 2014. (Cited on p. 3)

[32] A. S. M. Vamsi and N. Elia. Optimal distributed controllers realizable over arbitrary networks. *IEEE Transactions on Automatic Control*, 61(1):129–144, 2016. (Cited on p. 3)

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

[34] W. Wonham. On the separation theorem of stochastic control. *SIAM Journal on Control*, 6(2):312–326, 1968. (Cited on p. 3)

[35] J. Yan and S. E. Salcudean. Teleoperation controller design using H_∞-optimization with application to motion-scaling. *IEEE Transactions on Control Systems Technology*, 4(3):244–258, 1996. (Cited on p. 3)

[36] D. Youla, H. Jabr, and J. Bongiorno. Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case. *IEEE Transactions on Automatic Control*, 21(3):319–338, 1976. (Cited on p. 8)

[37] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*, volume 40. Prentice Hall, New Jersey, 1996. (Cited on pp. 3, 6, 8, 9, 12, 15, and 23)