Unified Approach to Convex Robust Distributed Control given Arbitrary Information Structures

Author(s):
Furieri, Luca (iD); Kamgarpour, Maryam (iD)

Unified Approach to Convex Robust Distributed Control given Arbitrary Information Structures

Luca Furieri, Maryam Kamgarpour

Abstract—We consider the problem of computing optimal linear control policies for linear systems in finite-horizon. The states and the inputs are required to remain inside pre-specified safety sets at all times despite unknown disturbances. In this technical note, we focus on the requirement that the control policy is distributed, in the sense that it can only be based on partial information about the history of the outputs. It is well-known that when a condition denoted as Quadratic Invariance (QI) holds, the optimal distributed control policy can be computed in a tractable way. Our goal is to unify and generalize the class of information structures over which quadratic invariance is equivalent to a test over finitely many binary matrices. The test we propose certifies convexity of the output-feedback distributed control problem in finite-horizon given any arbitrarily defined information structure, including the cases of time-varying communication networks and forgetting mechanisms. Furthermore, the framework we consider allows for including polytopic constraints on the states and the inputs in a natural way, without affecting convexity.

Index Terms—constrained & distributed optimal control, quadratic invariance, arbitrary information structures, finite-horizon robust control.

I. INTRODUCTION

CRITICAL emerging large-scale systems, such as the electric power grid, autonomous vehicles, the internet and financial systems feature autonomous and interacting decision making agents. Due to geographic distance or privacy concerns, the agents can only base their decisions on local partial information. This lack of full information has a significant impact on the design of optimal decisions.

It is well-known that the design of optimal control policies given an information structure can be extremely challenging even in simple cases [1]. Several instances of this problem were shown to be intractable [2], [3]. Significant work has thus been directed towards identifying special cases of this problem for which efficient algorithms can be derived. The work in [4] showed that when the information structure is Partially Nested (PN) the linear-quadratic Gaussian stochastic control problem can be reduced to a static problem, whose optimal solution is linear. More recently, the work [5] established convexity of the problem of optimal disturbance rejection in broader classes of controlled systems, such as those being "nested, chained, hierarchical, delayed interaction and communication, or symmetric" [5]. In [6], optimal distributed controller design for spatially invariant systems was shown to be convex given "funnel causality" of the information structure.

The cases [4]–[6] were unified in the celebrated works [7]–[10], where the authors established necessary and sufficient conditions on the plant and the information structure for convexity of optimal control referred to as *quadratic invariance* (QI). Quadratic invariance revealed that full sensor information is needed when all the states dynamically influence each other. Such limitation encouraged including a communication network in the distributed control system design to restore convexity. The works [11], [12] proposed a system level approach, which is based on allowing controllers to share past control inputs, thus improving overall convexity and performance. Input sharing might raise privacy concerns in competitive scenarios where controllers need to keep their control strategies unknown to others. Instead, [13] considered a time-invariant communication network to share local output measurements with fixed delays. The authors of [13] used the concepts of propagation and transmission delays to derive algebraic conditions that are equivalent to QI for these information structures.

Once a convex problem is obtained under the QI condition, one possible approach is to compute optimal controllers in the form of transfer functions by approximating the corresponding infinite-dimensional convex optimization problem [14]. In addition to relying on approximations, these techniques might lead to controllers whose order is too high for practical implementation. Moreover, operations at the transfer function level can be numerically unstable [15] and may not offer sufficient insight into the dynamics of the internal states of controllers. To deal with these issues, another approach is to explicitly compute distributed optimal controllers within state space representations by using dynamic programming. This approach led to several works, each addressing a specific class of dynamical systems and information structures as follows. The authors in [15] considered the case of poset-causal systems given that controllers can only access partial sensor measurements at each time and showed that the problem reduces to a set of decoupled centralized problems. The case of output sharing between controllers with one-timestep delays was solved in [16]–[18]. Combining sparse sensor measurements and fixed delays in the communication between controllers was considered in [19]. In [20], the authors solved the more general case of time-varying delays given a special class of systems consisting of two interconnected plants.

In several control scenarios of practical interest, it is important to ensure that states and control inputs remain inside specified safety sets at all times. However, the past works on optimal distributed control based on convex programming [5]–[9] and on explicit dynamic programming [15]–[20] did

not address the case where constraints on the state and input trajectories must be enforced. A technique to include state and input constraints is distributed model predictive control [21], [22]. Typically, this approach involves dividing a large-scale system into subsystems that communicate in order to minimize a given cost function, while being robust to the dynamical couplings. Convexity of the resulting optimization problems is obtained at the cost of potential conservativeness by optimizing over open-loop control policies. However, repeating the optimization at every time step mitigates this suboptimality. A different approach is to consider closed-loop predictions that allow for longer prediction horizons without running into infeasibility of the optimization problem [23]. Following this line of work, [24] has recently considered the problem of convex robust distributed control with closed-loop predictions given a fixed sensing topology. In this technical note, we adopt the closed-loop prediction approach.

The literature review above reveals that, when distributed control is considered, different instances of dynamical systems [5], [15], time-invariant and time-varying information structures [15]–[20], and the inclusion of state and input constraints [21]–[24] are treated as separate challenges, to be addressed with a variety of different approaches and mathematical tools. The contributions of this paper are as follows. First, we provide a unified and generalized approach to convex design of optimal distributed controllers. We show that sparsity constraints, delay constraints, arbitrary time-varying sensing and communication networks and forgetting mechanisms can all be combined within the framework of robust distributed control in finite-horizon. Second, we characterize convexity given any arbitrary information structure as per above, in terms of a test on finitely many binary matrices. The test we propose generalizes those proposed in [8], [13], [24], [25] for particular information structures, in the sense that they can all be alternatively proved as direct corollaries of our main result. To show generality of our approach, we provide an example of a quadratically invariant information structure which was not captured in the previous works. Third, our framework allows for including polytopic constraints on the state and input trajectories in a natural and forthright way.

The paper is structured as follows. In Section II, after introducing the problem under study we explain how to encode arbitrary information structures using binary matrices whose dimensions grow with the prediction horizon. In Section III we present our main result on a generalized binary test for quadratic invariance, and show an example of a QI information structure which was not captured in previous work. In Section IV, we specialize our result to the case of combining a fixed sensing topology and a fixed communication network. In doing so, we present alternative proofs to the results of [13], [24] as direct corollaries of our main theorem. These new proofs offer additional insight on the results of [13], [24] in terms of the minimal set of inequalities that must be verified for convexity.

II. Preliminaries

In this section, we introduce some notation on sparsity structures, and present the problem of distributed optimal control in finite-horizon with hard constraints on the states and the inputs.

A. Notation and sparsity structures

We use $\mathbb{R}$ and $\mathbb{N}$ to denote the sets of real numbers and positive integers respectively. For any $n \in \mathbb{N}$, we use $\mathbb{N}_{[a,b]} \subseteq \mathbb{N}$ to denote the set of integers from a to b. The (i, j)-th element in a matrix $Y \in \mathbb{R}^{m \times n}$ is referred to as $Y(i, j)$. Throughout the paper we indicate the i-th component of a vector $v \in \mathbb{R}^n$ as $v^i \in \mathbb{R}$. We use I_n to denote the identity matrix of size $n \times n$ and $0_{m \times n}$ to denote the zero matrix of size $m \times n$. The sparsity structure of a matrix can be conveniently represented by a binary matrix. A binary matrix is a matrix with entries from the set $\{0, 1\}$, and we use $\{0, 1\}^{m \times n}$ to denote the set of $m \times n$ binary matrices. Given a binary matrix $X \in \{0, 1\}^{m \times n}$, we define the sparsity subspace $\mathrm{Sparse}(X) \subseteq \mathbb{R}^{m \times n}$ as

$$\mathrm{Sparse}(X) := \{Y \in \mathbb{R}^{m \times n} \mid Y(i, j) = 0, \text{ if } X(i, j) = 0, \forall i, j\}.$$

Similarly, given $Y \in \mathbb{R}^{m \times n}$, we define a binary matrix $X := \mathrm{Struct}(Y)$ as

$$X(i, j) = \begin{cases} 1 & \text{if } Y(i, j) \neq 0, \\ 0 & \text{otherwise}. \end{cases}$$

Let $X, \hat{X} \in \{0, 1\}^{m \times n}$ and $Z \in \{0, 1\}^{n \times p}$ be binary matrices. Throughout the paper, we adopt the following conventions:

$$X + \hat{X} := \mathrm{Struct}(X + \hat{X}),$$
$$XZ := \mathrm{Struct}(XZ).$$

We state that $X \leq \hat{X}$ if and only if $X(i, j) \leq \hat{X}(i, j) \, \forall i, j$, and $X < \hat{X}$ if and only if $X \leq \hat{X}$ and there exist indices i, j such that $X(i, j) < \hat{X}(i, j)$. Also, we denote $X \nleq \hat{X}$ if and only if there exist indices i, j such that $X(i, j) > \hat{X}(i, j)$.

B. Problem formulation

We consider a discrete time system

$$x_{k+1} = Ax_k + Bu_k + Dw_k, \quad y_k = Cx_k + Hw_k, \quad (1)$$

where $k \geq 0$ is an integer number, $y_k \in \mathbb{R}^p$, $u_k \in \mathbb{R}^m$, $w_k \in \mathcal{W}$ and $\mathcal{W} \subseteq \mathbb{R}^n$ is a closed and bounded set of possible disturbances. The system starts from a known initial condition $x_0 \in \mathbb{R}^n$. Let us define a prediction horizon of length $N \in \mathbb{N}$. Our goal is to minimize a cost function of the history of states and inputs $J(x_0, \cdots, x_N, u_0, \cdots, u_{N-1})$. Furthermore, the states and inputs need to satisfy

$$\begin{bmatrix} x_k \\ u_k \end{bmatrix} \in \Gamma \subseteq \mathbb{R}^{n+m}, \quad x_N \in \mathcal{X}_f \subseteq \mathbb{R}^n, \quad (2)$$

for all $k \in \mathbb{N}_{[0, N-1]}$ and for all realizations of disturbances taken from set $\mathcal{W}$. Each control input can base its decisions on a given subset of the present and past outputs. This subset is denoted as the *information structure*. An information structure can be arbitrary, in the sense that controllers may measure, receive, memorize or forget any output measurements at any time. The search in the class of all output-feedback policies

is intractable. Hence, a possible approach is to restrict the search to the class of controllers that are affine in the history of the outputs. The time-varying output-feedback affine control policy is expressed as

$$u_k = \sum_{j=0}^{k} L_{k,j} y_j + g_k \,, \qquad (3)$$

for all time instants $k \in \mathbb{N}_{[0,N-1]}$. In (3), the matrices $L_{k,j} \in \mathbb{R}^{m \times p}$ represent the output-feedback part of the control policy, while the vectors $g_k \in \mathbb{R}^m$ represent a feed-forward term that can be synthesized to improve performance. For every $j \in \mathbb{N}_{[0,k]}$, we consider binary matrices $S_{k,j} \in \{0,1\}^{m \times p}$ describing the information known at time k to controllers about the outputs at time $j \leq k$. Specifically, $S_{k,j}(a,b) = 1$ if and only if the a-th scalar control input at time k knows the b-th scalar output at time j. A given information structure on the input can thus be enforced as

$$L_{k,j} \in \mathrm{Sparse}(S_{k,j}) \,, \quad \forall k \in \mathbb{N}_{[0,N-1]}, \; \forall j \in \mathbb{N}_{[0,k]} \,. \qquad (4)$$

Arbitrary information structures: The general description we suggest in (4) allows for unifying the treatment of several classes of information structures considered in the literature. We provide a list below.
1) Time-invariant sensing topologies described by $S \in \{0,1\}^{m \times p}$ as per [24] can be encoded by selecting $S_{k,j} = S$ for all $k \in \mathbb{N}_{[0,N-1]}$ and $j \in \mathbb{N}_{[0,k]}$.
2) Networked systems with fixed communication delays between subsystems as considered in [19] can be encoded as follows. Let $d_{a,b} \in \mathbb{N}$ be the time steps it takes for the information about the b-th output to reach the a-th input, where the case $d_{a,b} = \infty$ is denoted in [19] as a sparsity constraint. Then, for any given time k, we set $S_{k,l}(a,b) = 1$ for all $l \in \mathbb{N}_{[0,k-d_{a,b}]}$ and $S_{k,l}(a,b) = 0$ for all $l \in \mathbb{N}_{[k-d_{a,b}+1,k]}$.
3) Time-varying delays in a networked control system [20] can be encoded as follows. Let $e_{a,b}(k)$ be the number of time steps it takes at time k for the information about the b-th output to reach the a-th input. Then, for any given time k, we set $S_{k,l}(a,b) = 1$ for each $l \in \mathbb{N}_{[0,k-e_{a,b}(k)]}$ and $S_{k,l}(a,b) = 0$ for all $l \in \mathbb{N}_{[k-e_{a,b}(k)+1,k]}$.
4) Scheduled intermittent observations as per [25] can be encoded in terms of time-varying sensing and communication topologies by choosing the matrices $S_{k,j}$ as per [25, Lemma 1].
5) Random information structures as considered in [26] can be encoded by letting $S_{k,j}$ be taken at random from a pre-specified set of possible sparsities.

We are now ready to state the optimization problem under study given an arbitrary information structure.

Problem 1

minimize $\quad J(x_0, \cdots, x_N, u_0, \cdots, u_{N-1})$

subject to $(1),(2) \quad \forall w_k \in \mathcal{W} \,,$

$\qquad\qquad (3),(4) \quad \forall k \in \mathbb{N}_{[0,N-1]} \,, \forall j \in \mathbb{N}_{[0,k]} \,.$

In the problem above, the decision variables are the output-feedback matrices $L_{k,j}$ and the feed-forward vectors g_k as in

(3) for all $k \in \mathbb{N}_{[0,N-1]}$ and $j \in \mathbb{N}_{[0,k]}$. As we optimize over the matrices defining the closed-loop policy (3), we say that we perform closed-loop predictions. Without any knowledge about the statistics of the realization of w_k at each time in the set $\mathcal{W}$, we consider $J(\cdot)$ to be a function of the disturbance-free state and input trajectories. For computational tractability, we assume that $J(\cdot)$ is a convex function and for simplicity we assume that the sets Γ, $\mathcal{X}_f$ are polytopes $\Gamma = \{(x,u) \in (\mathbb{R}^n, \mathbb{R}^m)$ s.t. $Ux + Vu \leq b\}$, where $U \in \mathbb{R}^{s \times n}$, $V \in \mathbb{R}^{s \times m}$ and $b \in \mathbb{R}^s$, and $\mathcal{X}_f = \{x \in \mathbb{R}^n$ s.t. $Rx \leq z\}$, where $R \in \mathbb{R}^{r \times n}$ and $z \in \mathbb{R}^r$. Despite these assumptions, it was shown in [27] that Problem 1 is non-convex in general. This is due to the nonlinear propagation of the output-feedback matrices through the linear model (1).

III. CONVEXITY FOR ARBITRARY INFORMATION STRUCTURES

Our goal is then to find a convex program which is equivalent to the intractable Problem 1. It is convenient to define the vectors of stacked variables as

$$\mathbf{x} = \begin{bmatrix} x_0^\mathsf{T} & \cdots & x_N^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n(N+1)} \,,$$
$$\mathbf{y} = \begin{bmatrix} y_0^\mathsf{T} & \cdots & y_N^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{p(N+1)} \,,$$
$$\mathbf{u} = \begin{bmatrix} u_0^\mathsf{T} & \cdots & u_{N-1}^\mathsf{T} & 0_{m \times 1}^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{m(N+1)} \,,$$
$$\mathbf{w} = \begin{bmatrix} w_0^\mathsf{T} & \cdots & w_{N-1}^\mathsf{T} & 0_{n \times 1}^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n(N+1)} \,.$$

Equation (1) can be succinctly expressed as

$$\mathbf{x} = \mathbf{A}x_0 + \mathbf{B}\mathbf{u} + \mathbf{E}_D \mathbf{w} \,, \quad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{H}\mathbf{w} \,, \qquad (5)$$

where matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{E}_D$, $\mathbf{C}$ and $\mathbf{H}$ are defined in the Appendix. Their derivation is straightforward from the recursive application of (1). Similarly, considering (3), the control input can be expressed as $\mathbf{u} = \mathbf{L}\mathbf{y} + \mathbf{g}$, where $\mathbf{L} \in \mathbb{R}^{m(N+1) \times p(N+1)}$ and $\mathbf{g} \in \mathbb{R}^{m(N+1)}$ are defined in the Appendix. In order to satisfy (4) matrix $\mathbf{L}$ must lie in a subspace $\mathcal{S} \subseteq \mathbb{R}^{m(N+1) \times p(N+1)}$, where $\mathcal{S} = \mathrm{Sparse}(\mathbf{S})$ and $\mathbf{S}$ stacks the matrices $S_{k,j}$'s as per (18) in the Appendix.

A. Disturbance-feedback parametrization

It is known that when an information structure is not enforced, parametrizing the controller as a disturbance-feedback affine policy restores tractability of Problem 1 [27]. Letting $\mathbf{P} = \mathbf{C}\mathbf{E}_D + \mathbf{H}$ we define the disturbance-feedback controller

$$\mathbf{u} = \mathbf{Q}\mathbf{P}\mathbf{w} + \mathbf{v} \,. \qquad (6)$$

The decision variable $\mathbf{Q} \in \mathbb{R}^{m(N+1) \times p(N+1)}$ is causal as per (18). It is easy to verify that we can map a disturbance-feedback controller $(\mathbf{Q}, \mathbf{v})$ to the unique corresponding output-feedback controller $(\mathbf{L}, \mathbf{g})$ and vice versa as follows.

$$\mathbf{L} = \mathbf{Q}(\mathbf{C}\mathbf{B}\mathbf{Q} + I_{p(N+1)})^{-1} \,, \qquad (7)$$
$$\mathbf{g} = \mathbf{v} - \mathbf{Q}(\mathbf{C}\mathbf{B}\mathbf{Q} + I_{p(N+1)})^{-1}(\mathbf{C}\mathbf{B}\mathbf{v} + \mathbf{C}\mathbf{A}x_0) \,,$$
$$\mathbf{Q} = \mathbf{L}(I_{p(N+1)} - \mathbf{C}\mathbf{B}\mathbf{L})^{-1} \,, \qquad (8)$$
$$\mathbf{v} = \mathbf{L}(I_{p(N+1)} - \mathbf{C}\mathbf{B}\mathbf{L})^{-1}(\mathbf{C}\mathbf{B}\mathbf{g} + \mathbf{C}\mathbf{A}x_0) + \mathbf{g} \,.$$

Motivated by [7]–[10] we introduce the closed-loop map operator for systems in finite-horizon.

Definition 1 (Closed-loop map): Let $\mathbf{X} \in \mathbb{R}^{m(N+1) \times p(N+1)}$ and $\mathbf{Y} \in \mathbb{R}^{p(N+1) \times m(N+1)}$. The closed-loop map function $h : \mathbb{R}^{m(N+1) \times p(N+1)} \times \mathbb{R}^{p(N+1) \times m(N+1)} \longrightarrow \mathbb{R}^{m(N+1) \times p(N+1)}$ is defined such that $h(\mathbf{X}, \mathbf{Y}) = -\mathbf{X}(I_{p(N+1)} - \mathbf{Y}\mathbf{X})^{-1}$.

With reference to mappings (7) and (8), notice that the operator $h(\cdot)$ maps an output-feedback controller $\mathbf{L}$ to the corresponding disturbance-feedback controller $-\mathbf{Q}$. In particular, the mappings (7) and (8) can be expressed as $\mathbf{L} = h(-\mathbf{Q}, \mathbf{CB})$ and $\mathbf{Q} = -h(\mathbf{L}, \mathbf{CB})$ respectively. It is easy to show that when $J(\cdot)$ is convex in the disturbance-free states and inputs, then it does not depend on $\mathbf{Q}$ and it is convex in $\mathbf{v}$. Similar to [27], the state and input constraints (2) are affine in $(\mathbf{Q}, \mathbf{v})$ and can be expressed as

$$F\mathbf{v} + \max_{\mathbf{w} \in \mathcal{W}^{N+1}} (F\mathbf{Q}P + G)\mathbf{w} \le c, \qquad (9)$$

where $F \in \mathbb{R}^{(Ns+r) \times m(N+1)}$, $G \in \mathbb{R}^{(Ns+r) \times n(N+1)}$, $c \in \mathbb{R}^{Ns+r}$ are reported in the Appendix for completeness, and $\max_{\mathbf{w} \in \mathcal{W}^{N+1}}(\cdot)$ denotes row-wise maximization. We remark that no assumption on the convexity of the set of disturbances $\mathcal{W}$ is needed for the set of disturbance-feedback controllers $(\mathbf{Q}, \mathbf{v})$ satisfying (9) to be convex [27]. Nonetheless, a non-convex set of disturbances $\mathcal{W}$ can be substituted by its convex hull without loss of generality or performance [28, Example 7.1.2.]. In this case, computing a policy that satisfies (9) is easily done by formulating the dual of the maximization problems corresponding to each row of $\max_{\mathbf{w} \in \mathcal{W}^{N+1}}(F\mathbf{Q}P + G)\mathbf{w}$. We refer the reader to [27, Section 4.2] for the specific cases of polytopic and norm-bounded disturbances.

By the discussion above, we conclude that Problem 1 can be equivalently formulated as the following optimization problem:

Problem 2

$$\underset{\mathbf{Q}, \mathbf{v}}{\text{minimize}} \quad J(x_0, \mathbf{v})$$

$$\text{subject to} \quad F\mathbf{v} + \max_{\mathbf{w} \in \mathcal{W}^{N+1}} (F\mathbf{Q}P + G)\mathbf{w} \le c,$$

$$\mathbf{Q} \in -h(\mathcal{S}, \mathbf{CB}). \qquad (10)$$

The optimal solution of Problem 2 can be translated into the optimal solution of Problem 1, and vice-versa, through the mappings (7), (8). However, Problem 2 is still an intractable problem due to $h(\mathcal{S}, \mathbf{CB})$ being a non-convex set in general.

B. Quadratic invariance for convexity given an arbitrary information structure

If the set $h(\mathcal{S}, \mathbf{CB})$ is convex, then the resulting constraints on $\mathbf{Q}$ are convex and Problem 2 is a convex program equivalent to Problem 1. Hence, the question arises as to when $h(\mathcal{S}, \mathbf{CB})$ is a convex set.

Lemma 1: The set $h(\mathcal{S}, \mathbf{CB})$ is convex if and only if it is equal to $\mathcal{S}$.

Lemma 1 can be proved similar to [10, Theorem 5] by using finite-dimensional operators, subspaces and convex sets whereas [10] uses Banach spaces over $\mathbb{R}$, double-cones and star sets. Lemma 1 motivates looking for conditions so that $h(\mathcal{S}, \mathbf{CB}) = \mathcal{S}$. Such conditions are based on adapting the notion of quadratic invariance [8] to the considered setting.

Definition 2: The set $\mathcal{S} = \text{Sparse}(\mathbf{S})$ is quadratically invariant (QI) with respect to $\mathbf{CB}$ if and only if

$$\mathbf{L}\mathbf{CB}\mathbf{L}' \in \mathcal{S}, \quad \forall \mathbf{L}, \mathbf{L}' \in \mathcal{S}. \qquad (11)$$

Based on the same proof technique as per [8, Theorem 14] the following result holds.

Lemma 2: The sets $h(\mathcal{S}, \mathbf{CB})$ and $\mathcal{S}$ are equivalent if and only if $\mathcal{S}$ is QI with respect to $\mathbf{CB}$.

One of the key insights into proving Lemma 2 is the power series expansion of the closed-loop map $h(\mathbf{X}, \mathbf{Y})$. Notice that in the infinite-dimensional setting of [8] additional assumptions for plant and controller are needed to ensure $\|\mathbf{X}\mathbf{Y}\| < 1$ and convergence of the power series [8, Section III-A], whereas in the finite-dimensional setting considered here this condition is satisfied by construction. Indeed, the diagonal of matrix $\mathbf{LCB}$ is null for any dynamical system and affine controller. The above discussion leads to the following result.

Proposition 1: Problem 2 is a convex program equivalent to Problem 1 if and only if $\mathcal{S}$ is QI with respect to $\mathbf{CB}$.

Proof: Problem 1 and Problem 2 are equivalent upon parametrizing the control input as an output-feedback policy (3) or a disturbance-feedback policy (6) respectively. By combining Lemma 2 and Lemma 1 above we have that $h(\mathcal{S}, \mathbf{CB})$ is convex if and only if $\mathcal{S}$ is QI with respect to $\mathbf{CB}$. ∎

When $h(\mathcal{S}, \mathbf{CB})$ is convex, it is equal to $\mathcal{S}$ by Lemma 2. Hence, given QI one can simply substitute (10) with $\mathbf{Q} \in \mathcal{S}$ and obtain a finite-dimensional convex program, efficiently solvable with standard convex optimization techniques. The optimal solution $(\mathbf{Q}^\star, \mathbf{v}^\star)$ is translated to the globally optimal affine output-feedback solution of Problem 1 $(\mathbf{L}^\star, \mathbf{g}^\star)$ through the mapping (7).

Notice that QI as per (11) cannot be tested in practice, as the subspace $\mathcal{S}$ contains infinite elements. Our main result is to translate convexity conditions into finitely many inequalities over binary matrices for any arbitrary information structure in finite-horizon.

Theorem 1: Let $\Delta_g = \text{Struct}(CA^gB)$. Then, $\mathcal{S}$ is QI with respect to $\mathbf{CB}$ if and only if

$$S_{k,h}\Delta_g S_{h-g-1,j} \le S_{k,j}, \qquad (12)$$

for all $k \in \mathbb{N}_{[1,N-1]}$, $j \in \mathbb{N}_{[0,k-1]}$, $h \in \mathbb{N}_{[j+1,k]}$ and $g \in \mathbb{N}_{[0,h-j-1]}$.

Proof: Let $\Delta = \text{Struct}(\mathbf{CB})$. By [8, Theorem 26], QI is equivalent to

$$\mathbf{S}(k,i)\Delta(i,j)\mathbf{S}(j,l)(1 - \mathbf{S}(k,l)) = 0,$$
$$\forall k, j \in \mathbb{N}_{[0,mN]}, \ \forall i, l \in \mathbb{N}_{[0,pN]}.$$

By using the definition of the matrix product, the above is equivalent to

$$\mathbf{S}\Delta\mathbf{S} \le \mathbf{S}. \qquad (13)$$

It remains to decompose (13) into (12). Let us define $\varphi = \Delta\mathbf{S}$ and $\Phi = \mathbf{S}\varphi = \mathbf{S}\Delta\mathbf{S}$. Let $\varphi_{k,j}$ and $\Phi_{k,j}$ denote their $p \times p$ and $m \times p$ block sub-matrices respectively, located at block-row k and block-column j. Observe that

$$\varphi_{h,j} = \sum_{g=0}^{h-j-1} \Delta_g S_{h-g-1,j},$$

for $h \in \mathbb{N}_{[1,N]}$, $j \in \mathbb{N}_{[0,k-1]}$ and $\varphi_{h,j} = 0_{p \times p}$ otherwise. Pre-multiplying by $\mathbf{S}$, we obtain

$$\Phi_{k,j} = \sum_{h=j+1}^{k} S_{k,h} \varphi_{h,j}$$
$$= \sum_{h=j+1}^{k} S_{k,h} \sum_{g=0}^{h-j-1} \Delta_g S_{h-g-1,j}, \qquad (14)$$

for $k \in \mathbb{N}_{[1,N-1]}$, $j \in \mathbb{N}_{[0,k-1]}$ and $\Phi_{k,j} = 0_{m \times p}$ otherwise. Then, (13) is equivalent to $\Phi_{k,j} \leq S_{k,j}$ for every $k \in \mathbb{N}_{[1,N-1]}$ and $j \in \mathbb{N}_{[0,k-1]}$. Given (14), $\Phi_{k,j} \leq S_{k,j}$ if and only if each addend in its expression is less or equal than $S_{k,j}$. This concludes the proof. ∎

Our test (12) shows that convexity can be certified given any arbitrary information structure in finite-horizon, by verifying a finite number of inequalities over binary matrices. Instead, the work [8] introduced a binary test which is applicable to the special case of time-invariant sparsity constraints. Algebraic conditions on the propagation and transmission delays for the case of combining time-invariant sparsity constraints and delayed communication were proposed in [13]. However, finitely many inequalities equivalent to QI were not characterized for more general information structures, such as the intricate time-varying cases studied in [20] and [26].

Example 1: Here, we use our proposed binary test (12) to identify a complex QI information structure, that involves a combination of time-varying sensing, communication and the possibility for controllers to forget the outputs they have either measured or received. Consider the following three-dimensional system with two scalar controllers, whose dynamics (1) are defined by the matrices

$$A = \begin{bmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = I_3,$$

and matrices D and H are chosen arbitrarily. We consider a prediction horizon of $N = 3$. The information structure is defined as follows. At time $k = 0$, u^2 measures y_0^1. At time $k = 1$, both u^1 and u^2 measure y_1^1. Additionally, u^1 receives information about y_0^3 and u^2 receives information about y_0^2 and y_0^3. However, u^2 forgets y_0^1. At time $k = 2$, u^1 measures y_2^2 and receives information about y_1^3, y_0^1 and y_0^2, while u^2 does not perform any measurement and receives information about y_1^2. However, u^2 forgets y_0^2. This rather complicated information structure is conveniently encoded by the matrices

$$S_{0,0} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad S_{1,0} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

$$S_{1,1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad S_{2,0} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$S_{2,1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad S_{2,2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Our goal is to certify whether convex design of optimal robust controllers complying with such an information structure is possible. By letting $\Delta_0 = \text{Struct}(CB) = B$ and $\Delta_1 = \text{Struct}(CAB)$, we apply the generalized test for convexity developed in Theorem 1, which amounts to verifying the following set of inequalities:

$$S_{1,1} \Delta_0 S_{0,0} \leq S_{1,0}, \quad S_{2,1} \Delta_0 S_{0,0} \leq S_{2,0},$$
$$S_{2,2} \Delta_0 S_{1,0} \leq S_{2,0}, \quad S_{2,2} \Delta_1 S_{0,0} \leq S_{2,0},$$
$$S_{2,2} \Delta_0 S_{1,1} \leq S_{2,1}.$$

Since each of the inequalities above holds, we conclude that the considered information structure allows for convex design of optimal robust distributed controllers.

We remark that the convexity tests proposed in [8], [13] cannot be applied in this case, as the information structure cannot be described by a time-invariant sparsity pattern with delays. The cases of time-varying delays considered in [20] and that of intermittent observations recently considered in [25] assume that controllers never forget the information they have either observed or received over time. Hence, to the best of the authors' knowledge, the QI information structure considered here is not captured by other work. This example we consider reveals that convexity can be preserved even if controllers forget information in a finite-horizon.

Interpretation of Theorem 1. The binary test (12) naturally allows for a generalized interpretation of convexity in terms of the information which must be available to controllers as follows.

"Whenever controllers at time k know some information about the outputs y_h at time $h \leq k$, they must also know the output information which was available to those controllers whose decisions influenced these values of y_h through the evolution of the dynamics."

The above interpretation of (12) is a direct consequence of the fact that $S_{k,j}$ encodes what the controllers at time k know about the outputs at a past time j and that Δ_g encodes which outputs are affected after g time steps by the decisions of controllers. We remark that the above is consistent with the idea of *signaling* [29], that is the fact that controllers might need to infer information they don't have by inverting the system dynamics, thus compromising convexity of the corresponding optimization problem. Each binary inequality in (12) can thus be thought of as eliminating signaling between controllers at two specific different times.

IV. THE CASE OF FIXED SENSING AND COMMUNICATION TOPOLOGIES

In this section, we specialize the argument of Theorem 1 to specific information structures that are commonly found in practice. Our goal is to show that the convexity results of [13] can be alternatively proved as a direct corollary of Theorem 1, thus allowing for additional insight on the minimal number of inequalities to be tested for convexity. For simplicity, we start with the case where a communication network is not available.

A. Sensor information structures

We define a *sensing topology* that indicates which outputs are directly measured by which controllers. The sensing topology can be conveniently described by a binary matrix $S \in \{0,1\}^{m \times p}$ such that $S(i,j) = 1$ if and only if controller

i can measure y_k^j at every time instant k. As outlined in Section II, we can then set $S_{k,j} = S$ for all k, j to encode the corresponding information structure. Here, we show that [24, Theorem 3] about conditions on the system reachability matrix and the sensing topology for convexity can be alternatively proved as an immediate corollary of Theorem 1. First, we state the following lemma.

Lemma 3: Let X and Y be binary matrices of dimensions $m \times n$. Let Z, Z' be binary matrices of dimensions $p \times m$ and $n \times p$ respectively. Then, $X \leq Y$ implies that $ZX \leq ZY$ and $XZ' \leq YZ'$.

Proof: Let us suppose $X \leq Y$. We want to show that $ZX(i,j) = 1$ implies that $ZY(i,j) = 1$ for all indices i, j. If $ZX(i,j) = 1$, then there exists a k such that $Z(i,k) = X(k,j) = 1$. By hypothesis, also $Y(k,j) = 1$. Then $ZY(i,j) = 1$. $XZ' \leq YZ'$ is proved analogously. ∎

Corollary 1 (of Theorem 1): Let $S \in \{0,1\}^{m \times p}$ denote the sensing topology matrix and let $S_{k,j} = S$ for every $k \in \mathbb{N}_{[0,N-1]}$ and $j \in \mathbb{N}_{[0,k]}$. Let $\Delta_g = \text{Struct}(CA^g B)$. Then, Problem 2 is a convex program equivalent to Problem 1 if and only if

$$S\Delta_g S \leq S, \quad \forall g \in \mathbb{N}_{[0,n-1]}. \tag{15}$$

Proof: By Proposition 1 and Theorem 1, Problem 2 is a convex program equivalent to Problem 1 if and only if (12) holds. Clearly, (12) reduces to (15). It remains to be proven that it is sufficient to consider $g \in \mathbb{N}_{[0,n-1]}$. Suppose that $v \geq n$. Then, by Cayley-Hamilton we have that $CA^v B = C \sum_{i=0}^{n-1} \lambda_i A^i B$ for some coefficients $\lambda_i \in \mathbb{R}$. It follows that $\Delta_v \leq \sum_{i=0}^{n-1} \Delta_i$, because the coefficients λ_i might lead to some entries of Δ_v being null despite the same entry being 1 in Δ_i for some $i \in \mathbb{N}_{[0,n-1]}$. But then, if (15) holds, we have by Lemma 3 that

$$S\Delta_v S \leq S \sum_{i=0}^{n-1} \Delta_i S \leq S.$$

In other words, $S\Delta_g S \leq S$ for all $g \in \mathbb{N}_{[0,n-1]}$ implies $S\Delta_v S \leq S$ for all $v \in \mathbb{N}$. ∎

By [8, Theorem 26], the conditions (15) imply that $\text{Sparse}(S)$ is QI with respect to $CA^g B$ for every $g \in \mathbb{N}_{[0,n-1]}$, thus extending [24, Theorem 3] to the output-feedback case. We remark that while our binary test for convexity (12) involves verifying a number of inequalities that is polynomial in N in general, the corollary above shows that only n inequalities are necessary for the class of fixed sensor information structures.

An inherent limitation of time-invariant sensor information structures appears evident from (15). Indeed, when the system matrix A describes a strongly connected topology, we have that $CA^{n-1}B$ is a dense matrix in general. It follows that the product $S\Delta_{n-1}S$ is dense whenever S contains at least a 1 in each row and column, and hence that S must also be dense in general to satisfy (15). This makes convex design of distributed controllers impossible given a fixed sensor information structure. We conclude that a communication network that propagates the output measurements across controllers must be available in order to restore a convex optimization problem for

strongly connected systems. In the next section, we exploit Theorem 1 to investigate the value of communication in allowing for convex design of distributed controllers.

B. Sensor and communication information structures

Suppose that a time-invariant sensing topology S is defined and that a communication network is also available. Controllers that can exchange information are encoded by a corresponding *communication topology*, where information sharing between controller i and controller j is unidirectional in general. It is convenient to encode the communication topology in a binary matrix $Z \in \{0,1\}^{m \times m}$ such that $Z(i,j) = 1$ if and only if controller i receives information from controller j at each time k. In many applications, controllers are endowed with memory and thus "receive" information from themselves. For this reason we pose $Z(i,i) = 1$ for all i.

Communication could be exploited by simply letting controllers exchange their direct measurements, obtained according to the sensing topology. In this case, we would have that $S_{k,j} = ZS$ for every k, j. However, similar to the case of sensor information structures, Z would need to be dense in general for convexity of Problem 2 given a strongly connected system. To go beyond this limitation, we let controllers memorize both their direct sensor measurements and the information they receive through communication, and then propagate everything they have stored in their memory to their neighbours in the communication topology, who receive this information one time step later.

In other words, at every time instant $k \in \mathbb{N}_{[0,N-1]}$, controller u^a knows y^l at time $k - r$ if and only if there is a controller u^b which can directly measure y^l according to S, and there is a path of length r in the communication topology Z from u^b to u^a. In the next lemma, we encode this requirement as a particular choice for the sparsity subspace $\mathcal{S}$. First, we recall the definition of diameter of graph.

Definition 3 (Diameter of a directed graph): Consider a directed graph $\mathcal{G}$. We define the diameter of $\mathcal{G}$ to be the largest number of nodes contained in any of its paths when backtracking, detouring and looping paths are excluded from consideration and we denote it as $\mathcal{D}(X)$, where X is the binary matrix describing which nodes of $\mathcal{G}$ are connected by an edge.

Lemma 4: Let $S \in \{0,1\}^{m \times p}$ be the sensing topology matrix and $Z \in \{0,1\}^{m \times m}$ be the communication topology matrix. The information structure defined above can be encoded by selecting the matrices $S_{k,j}$ in (4) as

$$S_{k,j} = Z^{\min(\mathcal{D}(Z), k-j)} S . \tag{16}$$

Proof: It is well-known that Z^r encodes the paths of length r in the communication topology [30]. Then, by the definition of the information structure as above, we have that u^a knows y_{k-r}^l if and only if there exists a controller u^b such that $S(b,l) = 1$ and $Z^r(a,b) = 1$, or equivalently $(Z^r S)(a,l) = 1$. Now, for any $k \in \mathbb{N}_{[0,N-1]}$ and $r \in \mathbb{N}_{[0,k]}$, by definition the matrix $S_{k,k-r}$ must encode the output measurements known to controllers at time k about the outputs at time $k - r$. Hence, $S_{k,k-r} = Z^r S$ or equivalently $S_{k,j} = Z^{k-j} S$ for every $j \in \mathbb{N}_{[0,k]}$. Now, if $\mathcal{D}(Z) \leq k - j$, then $Z^{k-j} S = Z^{\mathcal{D}(Z)} S$

because $Z \geq I_n$ and by the definition of diameter of a graph. Equation (16) follows. ∎

Conditions for QI given a similar information structure were derived in [13], based on the notions of propagation and transmission delays between subsystems. Here, we show that a test of convexity in finite-horizon can be alternatively derived as a direct corollary of our Theorem 1. We report the proof in the Appendix.

Corollary 2 (of Theorem 1): Let $S \in \{0,1\}^{m \times p}$ denote the sensing topology matrix and $Z \in \{0,1\}^{m \times m}$ denote the communication topology matrix. Let $S_{k,j}$ be chosen according to (16). Let $\Delta_g = \mathrm{Struct}(CA^gB)$. Then, Problem 2 is a convex program equivalent to Problem 1 if and only if

$$S\Delta_g Z^r S \leq Z^{g+r+1}S,$$
$$\forall g \in \mathbb{N}_{[0,n-1]}, \ \forall r \in \mathbb{N}_{[0,\mathcal{D}(Z)]} \text{ s.t. } g + r \leq N - 2. \quad (17)$$

Equation (17) shows in a purely algebraic way the reason why communication enables convexity for strongly connected systems. Indeed, notice that the growth of the left-hand-side of (17) as g grows can be accommodated by the growth of the right-hand-side thanks to Z^g. Instead, in the case of sensor information structures where $Z = I_m$, the right-hand side cannot grow, thus compromising convexity for strongly connected systems.

We remark that Corollary 2, while being in accordance with [13, Theorem 2], offers new insight. First, it allows treating sensing and communication as two completely independent time-invariant topologies. Second, it identifies the minimal number of inequalities to be verified for convexity in terms of the system dimension n and the communication network diameter $\mathcal{D}(Z)$, whereas [13] requires verifying one inequality for each pair of subsystems in terms of the propagation and transmission delays.

V. CONCLUSIONS

We have unified and generalized the class of systems and information structures for which convexity of the robust distributed control problem in finite-horizon can be certified by finitely many algebraic conditions. In particular, we have shown that the information structures separately treated in the past works can all be described within the framework we suggest, and we derived a test for QI consisting of finitely many inequalities over certain binary matrices. Given the generality of the considered framework, we recovered previous results about specific classes of information structures as direct corollaries of our generalized test for convexity. In doing so, we provided new insight on the conditions for convexity given a fixed sensing and communication topology.

This work can be extended in several directions. First, it would be interesting to use the tools developed in this paper to address the case of time-varying system dynamics and study their interaction with event-dependent information structures such as those considered in [25]. Second, Example 1 suggests that forgetting mechanisms can also be included in the information structure without necessarily compromising convexity. Hence, it would be relevant to address the topic of designing information structures which are robustly convex

despite random packet dropouts and/or disorderings [31] and despite faults in the internal memory of controllers.

APPENDIX A

A. Additional Definitions

We define the following matrices and vectors.

$$\mathbf{L} = \begin{bmatrix} L_{0,0} & 0_{m\times p} & \cdots & 0_{m\times p} \\ \vdots & \vdots & \ddots & \vdots \\ L_{N-1,0} & \cdots & L_{N-1,N-1} & 0_{m\times p} \\ 0_{m\times p} & \cdots & 0_{m\times p} & 0_{m\times p} \end{bmatrix}, \ \mathbf{g} = \begin{bmatrix} g_0 \\ \vdots \\ g_{N-1} \\ 0_{m\times 1} \end{bmatrix}. \quad (18)$$

The matrix blocks above are $L_{k,j} \in \mathbb{R}^{m\times p}$, $g_k \in \mathbb{R}^m$ as in (4), and the $0_{m\times p}$ blocks enforce causality of the controller. We also define the following matrices, where $\otimes$ denotes the Kronecker product.

$$\mathbf{A} = \begin{bmatrix} I_n & A^\mathsf{T} & \cdots & A^{N\mathsf{T}} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n(N+1)\times n},$$

$$\mathbf{E} = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} & 0_{n\times n} \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} & 0_{n\times n} \\ A & I_n & \cdots & 0_{n\times n} & 0_{n\times n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A^{N-1} & A^{N-2} & \cdots & I_n & 0_{n\times n} \end{bmatrix} \in \mathbb{R}^{n(N+1)\times n(N+1)},$$

$$\mathbf{B} = \mathbf{E}(I_{N+1} \otimes B), \quad \mathbf{E}_D = \mathbf{E}(I_{N+1} \otimes D),$$

$$\mathbf{C} = I_{N+1} \otimes C, \quad \mathbf{H} = I_{N+1} \otimes H,$$

$$\mathbf{U} = \begin{bmatrix} I_N \otimes U & 0_{Ns\times n} \\ 0_{r\times nN} & R \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} I_N \otimes V & 0_{Ns\times m} \\ 0_{r\times mN} & 0_{r\times m} \end{bmatrix},$$

$$F = \mathbf{UB} + \mathbf{V}, \quad G = \mathbf{UE}_D, \quad c = \begin{bmatrix} \mathbf{1}_N \otimes b \\ z \end{bmatrix} - \mathbf{UA}x_0.$$

B. Proof of Corollary 2

By Proposition 1, Problem 2 is a convex program equivalent to Problem 1 if and only if $\mathcal{S}$ is QI with respect to $\mathbf{CB}$. We thus prove necessity and sufficiency of (17) for QI.

1) Proof of sufficiency: First, observe that if (17) holds for all $g \in \mathbb{N}_{[0,n-1]}$, then (17) holds for any $v \geq n$. Indeed

$$S\Delta_v Z^r S \leq S \sum_{i=0}^{n-1} \Delta_i Z^r S \quad \text{(Cayley-Hamilton and Lemma 3)}$$
$$\leq Z^{n+r}S \quad \text{(by (17) and } Z \geq I_m)$$
$$\leq Z^{v+r+1}S \quad \text{(by } Z \geq I_m).$$

Also notice that it is redundant to consider $r > \mathcal{D}(Z)$ by the definition of diameter and $Z \geq I_m$, and that we can restrict to $g + r \leq N - 2$ due to the horizon length. Now, let us pose $r = h - g - j - 1$ and pre-multiply both sides of the inequality above by Z^{k-h}. We obtain that (17) implies the following set of inequalities by Lemma 3:

$$Z^{k-h}S\Delta_g Z^{h-g-j-1}S \leq Z^{k-j}S, \quad (19)$$

for all $k \in \mathbb{N}_{[1,N-1]}$, $j \in \mathbb{N}_{[0,k-1]}$, $h \in \mathbb{N}_{[j+1,k]}$ and $g \in \mathbb{N}_{[0,h-g-1]}$. By definition of the information structure under consideration, we have $S_{k,h} = Z^{k-h}S$, $S_{h-g-1,j} = Z^{h-g-j-1}S$ and $S_{k,j} = Z^{k-j}S$. Hence, (19) is equivalent to (12), which is equivalent to QI by Theorem 1. Sufficiency is thus proved.

2) Proof of necessity: Let $\mathbf{L}, \mathbf{L}' \in \mathcal{S}$ and define $\boldsymbol{\varphi} = \mathbf{CBL}'$ and $\boldsymbol{\Phi} = \mathbf{L}\boldsymbol{\varphi} = \mathbf{LCBL}'$. Let $\varphi_{i,i-j}$ and $\Phi_{i,i-j}$ denote their $p \times p$ and $m \times p$ block sub-matrices located at block-row i and block-column $i - j$. Similar to (14) we have

$$\Phi_{i,i-j} = \sum_{h=i-j+1}^{i} L_{i,h} \sum_{g=0}^{h-(i-j)-1} CA^g BL'_{h-g-1,i-j}, \quad (20)$$

for every $i \in \mathbb{N}_{[0,N-1]}$, $j \in \mathbb{N}_{[1,i]}$. The case $j \leq 0$ is not considered for causality. QI requires that $\boldsymbol{\Phi} \in \mathcal{S}$ for arbitrary $\mathbf{L}, \mathbf{L}' \in \mathcal{S}$. We proceed into proving necessity of (17) by contrapositive. Let us suppose that (17) is violated. Take $g^\star, r^\star$ inside the ranges defined in (17) such that

$$S\Delta_{g^\star} Z^{r^\star} S \not\leq Z^{g^\star+r^\star+1} S .$$

We find matrices $\mathbf{L}, \mathbf{L}' \in \mathcal{S}$ and indices such that there is a block of $\boldsymbol{\Phi}$ located at block-row i and block-column $i - j$ for some i, j for which $\text{Struct}(\Phi)_{i,i-j} \not\leq Z^j S$. First, let us pose $L_{a,b} = 0_{m \times p}$ for all integers a, b such that $a \neq g^\star + r^\star + 1$ or $b \neq g^\star + r^\star + 1$ and $L'_{a,b} = 0_{m \times p}$ for all integers a, b such that $a \neq r^\star$ or $b \neq 0$. Then we have

$$\Phi_{g^\star+r^\star+1,0} = \sum_{h=1}^{g^\star+r^\star+1} L_{g^\star+r^\star+1,h} \sum_{g=0}^{h-1} CA^g BL'_{h-g-1,0}$$
$$= L_{g^\star+r^\star+1,g^\star+r^\star+1} CA^{g^\star} BL'_{r^\star,0} .$$

Finally, we choose matrices $L_{g^\star+r^\star+1,g^\star+r^\star+1}$ and $L'_{r^\star,0}$ such that $\text{Struct}(L_{g^\star+r^\star+1,g^\star+r^\star+1}) \leq S$, $\text{Struct}(L'_{r^\star,0}) \leq Z^{r^\star} S$ and

$$\text{Struct}(L_{g^\star+r^\star+1,g^\star+r^\star+1} CA^{g^\star} BL'_{r^\star,0}) = S\Delta_{g^\star} Z^{r^\star} S .$$

This can be done by making sure that whenever an entry of $S\Delta_{g^\star} Z^{r^\star} S$ is one, the corresponding entry of $L_{g^\star+r^\star+1,g^\star+r^\star+1} CA^{g^\star} BL'_{r^\star,0}$ is different from zero by avoiding finitely many specific values in the entries of $L_{g^\star+r^\star+1,g^\star+r^\star+1}$ and $L'_{r^\star,0}$. Thus, there are infinitely many suitable choices for $L_{g^\star+r^\star+1,g^\star+r^\star+1}$ and $L'_{r^\star,0}$. As a result, $\text{Struct}(\Phi_{g^\star+r^\star+1,0}) = S\Delta_{g^\star} Z^{r^\star} S \not\leq Z^{g^\star+r^\star+1}$, and $\boldsymbol{\Phi} \notin \mathcal{S}$, thus concluding the proof. ∎

REFERENCES

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

[2] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, vol. 36, no. 9, pp. 1249–1274, 2000.

[3] C. H. Papadimitriou and J. Tsitsiklis, "Intractable problems in control theory," *SIAM journal on control and optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[4] Y.-C. Ho and C. K'Ai-Ching, "Team decision theory and information structures in optimal control problems–Part I," *IEEE Transactions on Automatic control*, vol. 17, no. 1, pp. 15–22, 1972.

[5] P. G. Voulgaris, "A convex characterization of classes of problems in control with specific interaction and communication structures," in *IEEE American Control Conference (ACC)*, 2001, pp. 3128–3133.

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

[7] M. C. Rotkowitz, *Tractable problems in optimal decentralized control*. Stanford University, 2005.

[8] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 274–286, 2006.

[9] M. Rotkowitz, "On information structures, convexity, and linear optimality," in *IEEE Conference on Decision and Control (CDC)*, 2008, pp. 1642–1647.

[10] L. Lessard and S. Lall, "Quadratic invariance is necessary and sufficient for convexity," in *IEEE American Control Conference (ACC)*, 2011, pp. 5360–5362.

[11] Y.-S. Wang, N. Matni, S. You, and J. C. Doyle, "Localized distributed state feedback control with communication delays," in *IEEE American Control Conference (ACC)*, 2014, pp. 5748–5755.

[12] Y.-S. Wang, N. Matni, and J. C. Doyle, "A system level approach to controller synthesis," *IEEE Transactions on Automatic Control*, 2019.

[13] 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*, vol. 2, no. 1-3, pp. 30–54, 2010.

[14] A. Alavian and M. C. Rotkowitz, "Q-parametrization and an SDP for $\mathcal{H}_\infty$-optimal decentralized control," *IFAC Proceedings Volumes*, vol. 46, no. 27, pp. 301–308, 2013.

[15] 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*, vol. 58, no. 12, pp. 3084–3096, 2013.

[16] B.-Z. Kurtaran and R. Sivan, "Linear-quadratic-gaussian control with one-step-delay sharing pattern," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 571–574, 1974.

[17] N. Sandell and M. Athans, "Solution of some nonclassical LQG stochastic decision problems," *IEEE Transactions on Automatic Control*, vol. 19, no. 2, pp. 108–116, 1974.

[18] T. Yoshikawa, "Dynamic programming approach to decentralized stochastic control problems," *IEEE Transactions on Automatic Control*, vol. 20, no. 6, pp. 796–797, 1975.

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

[20] N. Matni, A. Lamperski, and J. C. Doyle, "Optimal Two Player LQR State Feedback With Varying Delay," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 2854–2859, 2014.

[21] C. Conte, C. N. Jones, M. Morari, and M. N. Zeilinger, "Distributed synthesis and stability of cooperative distributed model predictive control for linear systems," *Automatica*, vol. 69, pp. 117–125, 2016.

[22] P. Giselsson and A. Rantzer, "On feasibility, stability and performance in distributed model predictive control," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 1031–1036, 2014.

[23] A. Bemporad, "Reducing conservativeness in predictive control of constrained systems with disturbances," in *IEEE Conference on Decision and Control (CDC)*, 1998, pp. 1384–1389.

[24] L. Furieri and M. Kamgarpour, "Robust control of constrained systems given an information structure," in *IEEE Conference on Decision and Control (CDC)*, 2017, pp. 3481–3486.

[25] P. U. Abara, V. Causevic, and S. Hirche, "Quadratic invariance for distributed control system with intermittent observations," in *IEEE Conference on Decision and Control (CDC)*, 2018, pp. 2854–2860.

[26] Y. Ouyang, S. M. Asghari, and A. Nayyar, "Stochastic teams with randomized information structures," in *IEEE Conference on Decision and Control (CDC)*, 2017, pp. 4733–4738.

[27] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

[28] D. P. Bertsekas, A. Nedi, A. E. Ozdaglar *et al.*, "Convex analysis and optimization," 2003.

[29] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *IEEE Conference on Decision and Control (CDC)*, 2012, pp. 1291–1306.

[30] C. Godsil and G. F. Royle, *Algebraic graph theory*. Springer Science & Business Media, 2013, vol. 207.

[31] S. Wang, X. Meng, and T. Chen, "Wide-area control of power systems through delayed network communication," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 2, pp. 495–503, 2012.