Convex Relaxation for Optimal Distributed Control Problem

Ghazal Fazelnia Ramtin Madani Javad Lavaei
Department of Electrical Engineering, Columbia University

***Abstract*— This paper is concerned with the optimal distributed control (ODC) problem. The objective is to design a fixed-order distributed controller with a pre-specified structure for a discrete-time system. It is shown that this NP-hard problem has a quadratic formulation, which can be relaxed to a semidefinite program (SDP). If the SDP relaxation has a rank-1 solution, a globally optimal distributed controller can be recovered from this solution. By utilizing the notion of treewidth, it is proved that the nonlinearity of the ODC problem appears in such a sparse way that its SDP relaxation has a matrix solution with rank at most 3. A near-optimal controller together with a bound on its optimality degree may be obtained by approximating the low-rank SDP solution with a rank-1 matrix. This convexification technique can be applied to both time-domain and Lyapunov-domain formulations of the ODC problem. The efficacy of this method is demonstrated in numerical examples.**

I. Introduction

The area of decentralized control is created to address the challenges arising in the control of real-world systems with many interconnected subsystems. The objective is to design a structurally constrained controller—a set of partially interacting local controllers—with the aim of reducing the computation or communication complexity of the overall controller. The local controllers of a decentralized controller may not be allowed to exchange information. The term *distributed control* is sometimes used in lieu of decentralized control in the case where there is some information exchange between the local controllers. It has been long known that the design of an optimal decentralized (distributed) controller is a daunting task because it amounts to an NP-hard optimization problem [1]. Great effort has been devoted to investigating this highly complex problem for special types of systems, including spatially distributed systems and dynamically decoupled systems [2]–[5]. Another special case that has received considerable attention is the design of an optimal static distributed controller [6]. Due to the recent advances in the area of convex optimization, the focus of the existing research efforts has shifted from deriving a closed-form solution for the above control synthesis problem to finding a convex formulation of the problem that can be efficiently solved numerically [7]–[14].

There is no surprise that the decentralized control problem is computationally hard to solve. This is a consequence of the fact that several classes of optimization problems, including polynomial optimization and quadratically-constrained quadratic program as a special case, are NP-hard in the worst case. Due to the complexity of such problems, various convex relaxation methods based on linear matrix inequality (LMI), semidefinite programming (SDP), and second-order cone programming (SOCP) have gained popularity [15], [16]. These techniques enlarge the possibly non-convex feasible set into a convex set characterizable via convex functions, and then provide the exact or a lower bound on the optimal objective value. The SDP relaxation usually converts an optimization with a vector variable to a convex optimization with a matrix variable, via a lifting technique. The exactness of the relaxation can then be interpreted as the existence of a low-rank (e.g., rank-1) solution for the SDP relaxation. We developed the notion of "nonlinear optimization over graph" in [17] and [18] to study the exactness of an SDP relaxation. By adopting this notion, the objective of the present work is to study the potential of the SDP relaxation for the optimal distributed control problem.

In this work, we cast the optimal distributed control (ODC) problem as a quadratically-constrained quadratic program (QCQP) in a non-unique way, from which an SDP relaxation can be obtained. The primary goal is to show that this relaxation has a low-rank solution whose rank depends on the topology of the controller to be designed. In particular, we prove that the design of a static distributed controller with a pre-specified structure amounts to a sparse SDP relaxation with a solution of rank at most 3. This result also holds true for dynamic controllers and stochastic systems. In other words, although the rank of the SDP matrix can be arbitrarily large in theory, it never becomes greater than 3. This positive result may be used to understand how much approximation is needed to make the ODC problem tractable. It is also discussed how to round the rank-3 SDP matrix to a rank-1 matrix in order to design a near-optimal controller. Note that this paper significantly improves our recent result in [19], stating that the ODC problem with diagonal Q, R and K has an SDP solution with rank at most 4.

Notations: $\mathbb{R}$, $\mathbb{S}^n$, and $\mathbb{S}^n_+$ denote the sets of real numbers, $n \times n$ symmetric matrices, and $n \times n$ symmetric positive semidefinite matrices,, respectively. $\text{rank}\{W\}$ and $\text{trace}\{W\}$ denote the rank and trace of a matrix W. The notation $W \succeq 0$ means that W is symmetric and positive semidefinite. Given a matrix W, its (l, m) entry is denoted as W_{lm}. The superscript $(\cdot)^{\text{opt}}$ is used to show the globally optimal value of an optimization parameter. The symbols $(\cdot)^T$ and $\|\cdot\|$ denote the transpose and 2-norm operators, respectively. $|x|$ shows the size of a vector x. The notation

Emails: {ghazal, madani, lavaei}@ee.columbia.edu
This work was supported by a Google Research Award, NSF CAREER Award and ONR YIP Award.

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denotes as a graph $\mathcal{G}$ with the vertex set $\mathcal{V}$ and the edge set $\mathcal{E}$.

II. Problem formulation

Consider the discrete-time system

$$\begin{cases} x[\tau+1] = Ax[\tau] + Bu[\tau] \\ y[\tau] = Cx[\tau] \end{cases} \qquad \tau = 0, 1, ..., p \tag{1}$$

with the known parameters $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{r\times n}$, and $x[0] \in \mathbb{R}^n$. The goal is to design a decentralized (distributed) controller minimizing a quadratic cost function. With no loss of generality, we focus on the static case where the objective is to design a static controller of the form $u[\tau] = Ky[\tau]$ under the constraint that the controller gain K must belong to a given linear subspace $\mathcal{K} \subseteq \mathbb{R}^{m\times r}$ (see Section VI for various extensions). The set $\mathcal{K}$ captures the sparsity structure of the unknown constrained controller $u[\tau] = Ky[\tau]$ and, more specifically, it contains all $m \times r$ real-valued matrices with forced zeros in certain entries. This paper is mainly concerned with the following problem.

Optimal Distributed Control (ODC) problem: Design a static controller $u[\tau] = Ky[\tau]$ to minimize the cost function

$$\sum_{\tau=0}^{p} \left(x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau] \right) + \alpha \operatorname{trace}\{KK^T\} \tag{2}$$

subject to the system dynamics (1) and the controller requirement $K \in \mathcal{K}$, given the terminal time p, positive-definite matrices Q and R, and a constant α.

Note that the third term in the objective function of the ODC problem is a soft penalty term aimed at avoiding a high-gain controller.

III. SDP Relaxation for Quadratic Optimization

The objective of this section is to study the SDP relaxation of a QCQP problem using a graph-theoretic approach. Before proceeding with this part, some notions in graph theory will be reviewed.

A. Graph Theory Preliminaries

Definition 1: For two simple graphs $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, the notation $\mathcal{G}_1 \subseteq \mathcal{G}_2$ means that $\mathcal{V}_1 \subseteq \mathcal{V}_2$ and $\mathcal{E}_1 \subseteq \mathcal{E}_2$. $\mathcal{G}_1$ is called a subgraph of $\mathcal{G}_2$ and $\mathcal{G}_2$ is called a supergraph of $\mathcal{G}_1$.

Definition 2 (Treewidth): Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a tree $\mathcal{T}$ is called a tree decomposition of $\mathcal{G}$ if it satisfies the following properties:

1) Every node of $\mathcal{T}$ corresponds to and is identified by a subset of $\mathcal{V}$. Alternatively, each node of $\mathcal{T}$ is regarded as a group of vertices of $\mathcal{G}$.
2) Every vertex of $\mathcal{G}$ is a member of at least one node of $\mathcal{T}$.
3) $\mathcal{T}_k$ is a connected graph for every $k \in \mathcal{V}$, where $\mathcal{T}_k$ denotes the subgraph of $\mathcal{T}$ induced by all nodes of $\mathcal{T}$ containing the vertex k of $\mathcal{G}$.
4) The subgraphs $\mathcal{T}_i$ and $\mathcal{T}_j$ have a node in common for every $(i, j) \in \mathcal{E}$.

The width of a tree decomposition is the cardinality of its biggest node minus one (recall that each node of $\mathcal{T}$ is indeed a set containing a number of vertices of $\mathcal{G}$). The treewidth of $\mathcal{G}$ is the minimum width over all possible tree decompositions of $\mathcal{G}$ and is denoted by $\operatorname{tw}(\mathcal{G})$.

Definition 3: The representative graph of an $n \times n$ symmetric matrix W, denoted by $\mathcal{G}(W)$, is a simple graph with n vertices whose edges are specified by the locations of the nonzero off-diagonal entries of W. In other words, two arbitrary vertices i and j are connected if W_{ij} is nonzero.

Given a graph $\mathcal{G}$ accompanied by a tree decomposition $\mathcal{T}$ of width t, there exists a supergraph of $\mathcal{G}$ with certain properties that is called **enriched supergraph** of $\mathcal{G}$ derived by $\mathcal{T}$. The reader may refer to [20], [21] for a precise definition.

B. SDP Relaxation

Consider the standard nonconvex QCQP problem

$$\min_{x\in\mathbb{R}^n} \quad f_0(x) \tag{3a}$$

$$\text{s.t.} \quad f_k(x) \leq 0 \quad \text{for} \quad k = 1, \ldots, p \tag{3b}$$

where $f_k(x) = x^T A_k x + 2b_k^T x + c_k$ for $k = 0, \ldots, p$. Define

$$F_k \triangleq \begin{bmatrix} c_k & b_k^T \\ b_k & A_k \end{bmatrix} \quad \text{and} \quad w \triangleq [x_0 \quad x^T]^T, \tag{4}$$

where x_0=1. Given $k \in \{0, 1, ..., p\}$, the function $f_k(x)$ is a homogeneous polynomial of degree 2 with respect to w. Hence, $f_k(x)$ has a linear representation as $f_k(x) = \operatorname{trace}\{F_k W\}$, where

$$W \triangleq ww^T \tag{5}$$

Conversely, an arbitrary matrix $W \in \mathbb{S}^{n+1}$ can be factorized as (5) with $w_1 = 1$ if and only if it satisfies the three properties: $W_{11} = 1$, $W \succeq 0$, and $\operatorname{rank}\{W\} = 1$. Therefore, the general QCQP (3) can be reformulated as below:

$$\min_{W\in\mathbb{S}^{n+1}} \quad \operatorname{trace}\{F_0 W\} \tag{6a}$$

$$\text{s.t.} \quad \operatorname{trace}\{F_k W\} \leq 0 \quad \text{for} \quad k = 1, \ldots, p \tag{6b}$$

$$W_{11} = 1 \tag{6c}$$

$$W \succeq 0 \tag{6d}$$

$$\operatorname{rank}\{W\} = 1. \tag{6e}$$

This optimization is called a **rank-constrained formulation** of the QCQP (3). In the above representation of QCQP, the constraint (6e) carries all the nonconvexity. Neglecting this constraint yields a convex problem, which is called an **SDP relaxation** of the QCQP (3). The existence of a rank-1 solution for the SDP relaxation guarantees the equivalence between the original QCQP and its relaxed problem.

C. Connection Between Rank and Sparsity

To explore the rank of the minimum-rank solution of the SDP relaxation, define $\mathcal{G} = \mathcal{G}(F_0) \cup \cdots \cup \mathcal{G}(F_p)$ as the **sparsity graph** associated with the rank-constrained problem (6). The graph $\mathcal{G}$ describes the zero-nonzero pattern of the matrices $F_0, \ldots, F_p$, or alternatively captures the sparsity level of the QCQP problem (3). The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ has the following properties:

1) Each vertex of $\mathcal{V}$ corresponds to one of the entries of w or equivalently one of the elements of the set $\{x_0, x_1, ..., x_n\}$ (note that $x_0 = 1$). Let the vertex associated with the variable x_i be denoted as v_{x_i} for $i = 0, 1, ..., n$.
2) Given two distinct indices $i, j \in \{0, 1, \ldots, n\}$, the pair (v_{x_i}, v_{x_j}) is an edge of $\mathcal{G}$ if and only if the monomial $x_i x_j$ has a nonzero coefficient in at least one of the polynomials $f_0(x), f_1(x), \ldots, f_p(x)$.

Let $\bar{\mathcal{G}}$ be an enriched supergraph of $\mathcal{G}$ with $\bar{n}$ vertices that is obtained from a tree decomposition of width t (note that $\bar{n} \geq n$).

Theorem 1: Consider an arbitrary solution $\widehat{W} \in \mathbb{S}_+^n$ to the SDP relaxation of (6) and let $Z \in \mathbb{S}^{\bar{n}}$ be a matrix with the property that $\mathcal{G}(Z) = \bar{\mathcal{G}}$. Let $\overline{W}^{\text{opt}}$ denote an arbitrary solution of the optimization

$$\min_{\overline{W} \in \mathbb{S}^{\bar{n}}} \quad \text{trace}\{Z\overline{W}\} \tag{7a}$$

$$\text{s.t.} \quad \overline{W}_{kk} = \widehat{W}_{kk} \quad \text{for} \quad k = 0, 1, ..., n, \tag{7b}$$

$$\overline{W}_{kk} = 1 \quad \text{for} \quad k = n+1, ..., \bar{n}, \tag{7c}$$

$$\overline{W}_{ij} = \widehat{W}_{ij} \quad \text{for} \quad (i, j) \in \mathcal{E}_{\mathcal{G}}, \tag{7d}$$

$$\overline{W} \succeq 0. \tag{7e}$$

Define $W^{\text{opt}} \in \mathbb{S}^{n+1}$ as a matrix obtained from $\overline{W}^{\text{opt}}$ by deleting the last $\bar{n} - n - 1$ rows and $\bar{n} - n - 1$ columns. Let t denote the treewidth of the graph. If there exists a positive definite matrix $\overline{W}$ satisfying the constraints of the above optimization, then W^{opt} has two properties:

a) W^{opt} is an optimal solution to the SDP relaxation of (6).
b) $\text{rank}\{W^{\text{opt}}\} \leq t + 1$.

Proof: See [20] for the proof. ■

Assume that a tree decomposition of $\mathcal{G}$ with a small width is known. Theorem 1 states that an arbitrary (high-rank) solution to the SDP relaxation problem can be transformed into a low-rank solution by solving the convex program (7). Note that the above theorem requires the existence of a positive-definite feasible point, but this assumption can be removed after a small modification to (7) (see [20]).

IV. Quadratic Formulations of ODC Problem

The primary objective of the ODC problem is to design a structurally constrained gain K. Assume that the matrix K has l free entries to be designed. Denote these parameters as $h_1, h_2, ..., h_l$. To formulate the ODC problem, the space of permissible controllers can be characterized as

$$\mathcal{K} \triangleq \left\{ \sum_{i=1}^{l} h_i M_i \;\middle|\; h \in \mathbb{R}^l \right\}, \tag{8}$$

for some (fixed) 0-1 matrices $M_1, ..., M_l \in \mathbb{R}^{m \times r}$. Now, the ODC problem can be stated as follows.

ODC Problem: Minimize

$$\sum_{\tau=0}^{p} \left(x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau] \right) + \alpha \,\text{trace}\{KK^T\} \tag{9a}$$

subject to

$$x[\tau+1] = Ax[\tau] + Bu[\tau] \quad \text{for} \quad \tau = 0, 1, \ldots, p \tag{9b}$$

$$y[\tau] = Cx[\tau] \quad \text{for} \quad \tau = 0, 1, \ldots, p \tag{9c}$$

$$u[\tau] = Ky[\tau] \quad \text{for} \quad \tau = 0, 1, \ldots, p \tag{9d}$$

$$K = h_1 M_1 + \ldots + h_l M_l \tag{9e}$$

$$x[0] = \text{given} \tag{9f}$$

over the variables

$$x[0], x[1], \ldots, x[p] \in \mathbb{R}^n \tag{9g}$$

$$y[0], y[1], \ldots, y[p] \in \mathbb{R}^r \tag{9h}$$

$$u[0], u[1], \ldots, u[p] \in \mathbb{R}^m \tag{9i}$$

$$h \in \mathbb{R}^l. \tag{9j}$$

To cast the ODC problem as a quadratic optimization, define

$$w_d \triangleq \begin{bmatrix} z & h^T & x^T & u^T \end{bmatrix}^T, \quad w_s \triangleq \begin{bmatrix} z & h^T & x^T & u^T & y^T \end{bmatrix}^T \tag{10}$$

where

$$x \triangleq \begin{bmatrix} x[0]^T & \cdots & x[p]^T \end{bmatrix}^T \tag{11a}$$

$$u \triangleq \begin{bmatrix} u[0]^T & \cdots & u[p]^T \end{bmatrix}^T \tag{11b}$$

$$y \triangleq \begin{bmatrix} y[0]^T & \cdots & y[p]^T \end{bmatrix}^T, \tag{11c}$$

and z is a scalar auxiliary variable playing the role of number 1 (equivalently, the quadratic constraint $z^2 = 1$ can be posed instead of $z = 1$ without affecting the solution). The objective and constraints of the ODC problems are all homogeneous functions of w_s with degree 2. For example, (9c) is equivalent to $z \times y[\tau] - C \times z \times x[\tau] = 0$, which is a second-order equality constraint. Hence, the ODC problem can be cast as a QCQP with a sparsity graph $\mathcal{G}$, where the vertices of $\mathcal{G}$ correspond to the entries of w_s. In particular, the vertex set $\mathcal{V}_{\mathcal{G}}$ can be partitioned into five vertex subsets, where subset 1 consists of a single vertex associated with the variable z and subsets 2-5 correspond to the vectors x, u, y and h, respectively.

Sparsity graph for diagonal Q, R, and K: Consider the case where the matrices Q and R are diagonal and the controller K to be designed needs to be diagonal as well. The underlying sparsity graph $\mathcal{G}$ is drawn in Figure 1, where each vertex of the graph is labeled by its corresponding variable. To increase the readability of the graph, some edges of vertex z are not shown in the picture. Indeed, z is connected to all vertices corresponding to the elements of x, u and y. This is due to the linear terms $x[\tau]$, $u[\tau]$ and $y[\tau]$ in (9) that are equivalent to $z \times x[\tau]$, $z \times u[\tau]$ and $z \times y[\tau]$.

Theorem 2: The sparsity graph of the ODC problem (9) has treewidth 2 in the case of diagonal Q, R, and K.

Proof: It follows from the graph drawn in Figure 1 that removing vertex z from the sparsity graph $\mathcal{G}$ makes the remaining subgraph acyclic. This implies that the treewidth of $\mathcal{G}$ is at most 2. On the other hand, the treewidth cannot be 1 in light of the cycles of the graph. ■

It follows from Theorems 1 and 2 that the SDP relaxation of the ODC problem has a matrix solution with rank 1, 2 or 3 in the diagonal case.

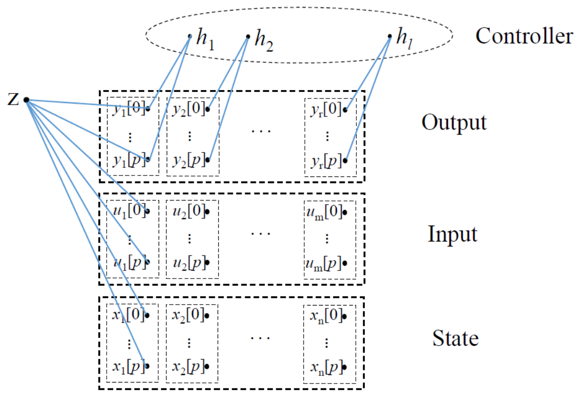

Fig. 1: Sparsity graph of the ODC problem for diagonal Q, R, and K (some edges of vertex z are not shown to increase the legibility of the graph).

Sparsity graph for non-diagonal Q, R, and K: As can be seen in Figure 1, there is no edge in the subgraph of $\mathcal{G}$ corresponding to the entries of x, as long as Q is diagonal. However, if Q has nonzero off-diagonal elements, certain edges (and probably cycles) may be created in the subgraph of $\mathcal{G}$ associated with the aggregate state x. Under this circumstance, the treewidth of $\mathcal{G}$ could be much higher than 2. The same argument holds for a non-diagonal R or K.

A. *Various Quadratic Formulations of ODC Problem*

The ODC problem has been cast as a QCQP in (9) although it has infinitely many quadratic formulations. Since every QCQP formulation of the ODC problem will be ultimately convexified in this work, the following question arises: what QCQP formulation of the ODC problem has a better SDP relaxation? Note that two important factors for an SDP relaxation are: (i) optimal objective value of the ODC problem serving as a lower bound on the globally optimal cost of the ODC problem, and (ii) the rank of the minimum-rank solution of the SDP relaxation. By taking these two factors into account, four different quadratic formulations of the ODC problem will be proposed below.

Consider an arbitrary QCQP formulation of the ODC problem. Assume that it is possible to design a set of redundant quadratic constraints whose addition to the QCQP problem would not affect its feasible set. These redundant constraints may lead to the shrinkage of the feasible set of the SDP relaxation of the original QCQP problem, leading to a tighter lower bound on the optimal cost of the ODC problem. More precisely, reducing the unnecessary (dependent) parameters of the QCQP problem and yet imposing additional redundant constraints help with Factor (i) mentioned above. Roughly speaking, this requires designing a quadratic formulation whose edge set is as large as possible. To this end, two modifications can be made on the QCQP problem (9):

1) The constraints (9c) and (9d) can be combined into a single equation $u[\tau] = KCx[\tau]$ and accordingly the variables $y[0], y[1], \ldots, y[p]$ can be removed from the optimization.
2) A multiplication of the constraint (9b) to itself for two different times τ_1 and τ_2 leads to a redundant quadratic equation, which can be imposed on the problem.

The above modifications yield two dense ODC problems.

First Dense Formulation of ODC: Minimize

$$\sum_{\tau=0}^{p}\left(x[\tau]^T Q x[\tau] + u[\tau]^T R u[\tau]\right) + \alpha\ \mathrm{trace}\{KK^T\} \quad (12a)$$

subject to

$$x[\tau+1] = Ax[\tau] + Bu[\tau] \quad (12b)$$
$$u[\tau] = KCx[\tau] \quad (12c)$$
$$K = h_1 M_1 + \ldots + h_l M_l \quad (12d)$$
$$x[0] = \text{given} \quad (12e)$$

for every $\tau \in \{0, 1, \ldots, p\}$, and subject to

$$x[\tau_1+1]x[\tau_2+1]^T = (Ax[\tau_1] + Bu[\tau_1]) \times (Ax[\tau_2] + Bu[\tau_2])^T, \quad (12f)$$

$$x[\tau_1+1](Ax[\tau_2] + Bu[\tau_2])^T = (Ax[\tau_1] + Bu[\tau_1]) \times x[\tau_2+1]^T \quad (12g)$$

for every $\tau_1, \tau_2 \in \{0, 1, \ldots, p\}$, over the optimization variables (9g), (9i) and (9j).

First Dense Formulation of ODC is a QCQP problem with a dense sparsity graph. Note that the redundant constraints (12f) and (12g) impose a high number of constraints on the entries of the associated SDP matrix W, leading to a possible shrinkage of the feasible set of the SDP relaxation. The SDP relaxation of First Dense Formulation of ODC aims to offer a good lower bound on the optimal cost of the ODC problem, but the rank of its SDP solution may be high. To reduce the rank, Theorem 1 suggests designing a quadratic formulation whose sparsity graph has a lower treewidth.

Second Dense Formulation of ODC: This optimization is obtained from First Dense Formulation of ODC (12) by dropping its redundant constraints (12f) and (12g).

As discussed before, non-diagonal Q and R result in a sparsity graph with a large treewidth. To remedy this issue, define a new set of variables as follows:

$$\bar{x}[\tau] \triangleq Q_d x[\tau], \qquad \bar{u}[\tau] \triangleq R_d u[\tau], \quad (14)$$

where $Q_d \in \mathbb{R}^{n\times n}$ and $R_d \in \mathbb{R}^{m\times m}$ are the respective eigenvector matrices of Q and R, i.e.,

$$Q = Q_d^T \Lambda_Q Q_d, \qquad R = R_d^T \Lambda_R R_d, \quad (15)$$

where $\Lambda_Q \in \mathbb{R}^{n\times n}$ and $\Lambda_R \in \mathbb{R}^{m\times m}$ are diagonal matrices. Define also

$$\bar{A} \triangleq Q_d A Q_d^T, \quad \bar{B} \triangleq Q_d B R_d^T, \quad \bar{C}_1 \triangleq C Q_d^T. \quad (16)$$

The ODC problem can be reformulated as below.

First Sparse Formulation of ODC: Minimize

$$\sum_{\tau=0}^{p}\left(\bar{x}[\tau]^T \Lambda_Q \bar{x}[\tau] + \bar{u}[\tau]^T \Lambda_R \bar{u}[\tau]\right) + \alpha\ \mathrm{trace}\{KK^T\} \quad (17a)$$

subject to

$$\bar{x}[\tau+1] = \bar{A}\bar{x}[\tau] + \bar{B}\bar{u}[\tau] \tag{17b}$$
$$y[\tau] = \bar{C}_1\bar{x}[\tau] \tag{17c}$$
$$\bar{u}[\tau] = R_d K y[\tau] \tag{17d}$$
$$K = h_1 M_1 + \ldots + h_l M_l \tag{17e}$$
$$x[0] = \text{given} \tag{17f}$$

for $\tau = 0, 1, \ldots, p$, over the optimization variables

$$\bar{x}[0], \bar{x}[1], \ldots, \bar{x}[p] \in \mathbb{R}^n \tag{17g}$$
$$y[0], y[1], \ldots, y[p] \in \mathbb{R}^r \tag{17h}$$
$$\bar{u}[0], \bar{u}[1], \ldots, \bar{u}[p] \in \mathbb{R}^m \tag{17i}$$
$$h \in \mathbb{R}^l. \tag{17j}$$

By comparing the objective functions of First Sparse Formulation of ODC and First/Second Dense Formulation of ODC, it can be concluded that the arbitrary matrices Q and R have been substituted by two diagonal matrices Λ_Q and Λ_R. It is straightforward to verify that the treewidth of the sparsity graph of First Sparse Formulation of ODC is dependent only on the sparsity level of the to-be-designed controller K. To remedy this drawback, notice that there exist constant binary matrices $\Phi_1 \in \mathbb{R}^{m\times l}$ and $\Phi_2 \in \mathbb{R}^{l\times r}$ such that

$$\mathcal{K} = \left\{ \Phi_1 \text{diag}\{h\} \Phi_2 \;\middle|\; h \in \mathbb{R}^l \right\}, \tag{18}$$

where $\text{diag}\{h\}$ denotes a diagonal matrix whose diagonal entries are inherited from the vector h [22]. Now, define

$$\bar{y}[\tau] \triangleq \Phi_2 y[\tau] \tag{19a}$$
$$\bar{C}_2 \triangleq \Phi_2 C Q_d^T. \tag{19b}$$

The constraints (9c), (9d) and (9e) are equivalent to $\bar{y}[\tau] = \bar{C}_2\bar{x}[\tau]$ and $\bar{u}[\tau] = \Phi_1\text{diag}\{h\}\bar{y}[\tau]$. Hence, the matrix K can be diagonalized in the ODC problem as follows.

Second Sparse Formulation of ODC: Minimize

$$\sum_{\tau=0}^{p} \bar{x}[\tau]^T \Lambda_Q \bar{x}[\tau] + \bar{u}[\tau]^T \Lambda_R \bar{u}[\tau] + \alpha\ h^T h \tag{20a}$$

subject to

$$\bar{x}[\tau+1] = \bar{A}\bar{x}[\tau] + \bar{B}\bar{u}[\tau] \tag{20b}$$
$$\bar{y}[\tau] = \bar{C}_2\bar{x}[\tau] \tag{20c}$$
$$\bar{u}[\tau] = R_d \Phi_1 \text{diag}\{h\} \bar{y}[\tau] \tag{20d}$$
$$\bar{x}[0] = \text{given} \tag{20e}$$

for $\tau = 0, 1, \ldots, p$, over the variables (17g), (17i), (17j) and $\bar{y}[0], \bar{y}[1], \ldots, \bar{y}[p]$.

It should be mentioned that First Sparse Formulation of ODC (17), Second Sparse Formulation of ODC (20) and the ODC problem (9) are not only equivalent but also identical in the case of diagonal Q, R, and K.

Theorem 3: The sparsity graph of Second Sparse Formulation of ODC (20) has treewidth 2.

Proof: The proof is omitted due to its similarity to the proof of Theorem 2. ∎

So far, four equivalent formulations of the ODC problem have been introduced. In the next section, the SDP relaxations of these formulations will be contrasted with each other.

V. SDP Relaxations of ODC Problem

To streamline the presentation, the proposed formulations of the ODC problem will be renamed as:

- **Problem D-1:** First Dense Formulation of ODC
- **Problem D-2:** Second Dense Formulation of ODC
- **Problem S-1:** First Sparse Formulation of ODC
- **Problem S-2:** Second Sparse Formulation of ODC

As mentioned earlier, each of these problems is a QCQP formulation of the ODC problem. Hence, the technique delineated in Subsection III-B can be deployed to obtain an SDP relaxation for each of these problems. Let W_{D_1}, W_{D_2}, W_{S_1} and W_{S_2} denote the variables of the SDP relaxations of Problems D-1, D-2, S-1 and S-2, respectively. The exactness of the SDP relaxation for Problem D-1 is tantamount to the existence of an optimal rank-1 matrix $W_{D_1}^{\text{opt}}$. In this case, an optimal vector w_d^{opt} for the ODC problem can be recovered by decomposing $W_{D_1}^{\text{opt}}$ as $(w_d^{\text{opt}})(w_d^{\text{opt}})^T$ (note that w_d has been defined in (10)). A similar argument holds for Problems D-2, S-1 and S-2. The following observations can be made here:

- The computational complexity of a convex optimization problem, in the worst case, is related to the number of variables as well as the number of constraints of the problem. Under this measure of complexity, the SDP relaxation of Problem D-2 is simpler (easier to solve) than those of Problems D-1, S-1 and S-2. Similarly, the SDP relaxation of Problem S-1 is simpler than that of Problem S-2. It is expected that the SDP relaxation of Problem D-1 has the highest complexity among all four SDP relaxations.
- An SDP relaxation provides a lower bound on the optimal cost of the ODC problem. It is justifiable that the lower bounds obtained from Problems D-1, D-2, S-1 and D-2 form a descending sequence. This implies that Problem D-1 may offer the best lower bound on the globally optimal cost of the ODC problem. On the other hand, the treewidths of the sparsity graphs of Problems D-1, D-2, S-1 and D-2 would form a descending sequence as well. Hence, in light of Theorem 1, Problem S-2 may offer the best low-rank SDP solution.

Corollary 1: The SDP relaxation of Second Sparse Formulation of ODC has a matrix solution with rank at most 3.

Proof: This corollary is an immediate consequence of Theorems 2 and 3. ∎

Although Problem D-1 may offer the tightest lower bound in theory, Problem S-2 has a guaranteed low-rank SDP solution. A question arises as to which of the Problems D-1, D-2, S-1 and S-2 should be convexified? We have analyzed these problems for several thousand random systems with random control structures and observed that:

- The SDP relaxation of Problem D-1 often has high-rank solutions and is also computationally expensive to solve.
- The SDP relaxations of Problems D-1, D-2, S-1 and S-2 result in very similar lower bounds in almost all cases. To support this statement, the optimal values of the four SDP relaxations are plotted for 100 random systems in the technical report [21].

Based on the above observations, the transition from the highly-dense Problem D-1 to the highly-sparse Problem S-2 may change the optimal SDP cost insignificantly but improves the rank of the SDP solution dramatically (note that the sizes of the SDP matrices for these two problems are different).

A. Rounding of SDP Solution to Rank-1 Matrix

Let W^{opt} denote a low-rank SDP solution for one of the above mentioned SDP relaxations. If the rank of this matrix is 1, then W^{opt} can be mapped back into a globally optimal controller for the ODC problem through an eigenvalue decomposition $W^{\text{opt}} = w^{\text{opt}}(w^{\text{opt}})^T$. If W^{opt} has a rank greater than 1, there are multiple approaches to recover a controller:

- A near-optimal controller may be obtained from the first column of W^{opt} corresponding to the controller part h.
- First, W^{opt} is approximated by a rank-1 matrix by means of the eigenvector associated with its largest eigenvalue. Then, a near-optimal controller may be constructed from the first column of this approximate rank-1 matrix.
- Recall that the SDP relaxation was obtained by eliminating a rank constraint. In the case where this removal changes the solution, one strategy is to compensate for the rank constraint by incorporating an additive penalty function, denoted as $\mu(W)$, into the objective of the SDP relaxation. A common penalty function $\mu(\cdot)$ is $\varepsilon \times \text{trace}\{W\}$, where ε is a design parameter.

Note that by comparing the cost for the near-optimal controller with the lower bound obtained from the SDP relaxation, the optimality degree of the designed controller can be assessed.

VI. Extensions

In the case of designing an optimal fixed-order dynamic controller with a pre-specified structure, denote the unknown controller as:

$$\begin{cases} z_c[\tau+1] = A_c z_c[\tau] + B_c y[\tau] \\ u[\tau] = C_c z_c[\tau] + D_c y[\tau] \end{cases} \tag{21}$$

where $z_c[\tau] \in \mathbb{R}^{n_c}$ represents the state of the controller, n_c denotes its known degree, and the quadruple (A_c, B_c, C_c, D_c) needs to be designed. Since the controller is required to have a pre-determined distributed structure, the 4-tuple (A_c, B_c, C_c, D_c) must belong to a given polytope $\mathcal{K}$. This polytope enforces certain entries of A_c, B_c, C_c, and D_c to be zero. The above ODC problem is a nonlinear optimization because the dynamics of the controller has some unknown nonlinear terms such as $A_c z_c[\tau]$ and $B_c y[\tau]$. In order to convexify the above ODC problem, define a vector w as

$$w = \begin{bmatrix} 1 & h^T & x^T & u^T & y^T & z_c^T \end{bmatrix}^T \tag{22}$$

where z_c is a column vector consisting of the entries of $z_c[0], ..., z_c[p]$, and h denotes a vector including all free (nonzero) entries of the matrices A_c, B_c, C_c, and D_c. The ODC problem can be cast as a quadratic optimization with respect to the vector w, from which an SDP relaxation can be derived. Hence, the results derived earlier can all be naturally generalized to the above dynamic case.

Other extensions are to solve the ODC problem for $p = \infty$ and/or a stochastic system. The detail for these cases can be found in [23], [24].

A. Computationally-Cheap SDP Relaxation

Although the proposed SDP relaxations are convex, it may be difficult to solve them efficiently for a large-scale system. This is due to the fact that the size of the SDP matrix depends on the number of scalar variables at all times from 0 to p. It is possible to significantly simplify the SDP relaxations. For example, since the treewidth of the SDP relaxation of Problem S-2 is equal to 2, the complicating constraint $W_{S_2} \succeq 0$ can be replaced by positive semidefinite constraints on certain 3×3 submatrices of W_{S_2} (those induced by the nodes of the minimal tree decomposition of the sparsity graph of Problem S-2) [20]. After this simplification of the hard constraint $W_{S_2} \succeq 0$, a quadratic number of entries of W_{S_2} turn out to be redundant (not appearing in any constraint) and can be removed from the optimization.

There is a more efficient approach to derive a computationally-cheap SDP relaxation. This will be explained below for the case where Q and R are non-singular and $r, m \leq n$.

With no loss of generality, we assume that C has full row rank. There exists an invertible matrix Φ such that

$$C\Phi = \begin{bmatrix} I & 0 \end{bmatrix} \tag{23}$$

where I is the identity matrix and "0" is an $r \times (n-r)$ zero matrix. Define also $\mathcal{K}^2 = \{KK^T \mid K \in \mathcal{K}\}$. Indeed, $\mathcal{K}^2$ captures the sparsity pattern of the matrix KK^T. For example, if $\mathcal{K}$ consists of block-diagonal (rectangular) matrix, $\mathcal{K}^2$ will also include block-diagonal (square) matrices. Let $\mu \in \mathbb{R}$ be a positive number such that

$$Q \succ \mu \times \Phi^{-T}\Phi^{-1} \tag{24}$$

where Φ^{-T} denotes the transpose of the inverse of Φ. Define $\widehat{Q} := Q - \mu \times \Phi^{-T}\Phi^{-1}$.

Computationally-Cheap SDP Relaxation: This optimization problem is defined as the minimization of

$$\text{trace}\left\{X^T \widehat{Q} X + \mu\, \mathbf{W}_{22} + U^T R U + \alpha\, \mathbf{W}_{33}\right\} \tag{25}$$

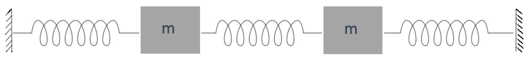

Fig. 2: Mass-spring system with two masses

subject to the constraints

$$x[\tau+1] = Ax[\tau] + Bu[\tau] \quad \text{for} \quad \tau = 0,1,\ldots,p, \tag{26a}$$

$$x[0] = \text{given}, \tag{26b}$$

$$\mathbf{W} := \left[\begin{array}{c|c|c} I_n & \Phi^{-1}X & \begin{bmatrix} K^T \\ 0 \end{bmatrix} \\ \hline X^T\Phi^{-T} & \mathbf{W}_{22} & U^T \\ \hline \begin{bmatrix} K & 0 \end{bmatrix} & U & \mathbf{W}_{33} \end{array}\right] \succeq 0, \tag{26c}$$

$$K \in \mathcal{K}, \tag{26d}$$

$$\mathbf{W}_{33} \in \mathcal{K}^2, \tag{26e}$$

with the optimization parameters

- $K \in \mathbb{R}^{m\times r}$
- $X = \big[x[0]\ x[1]\ \ldots\ x[p]\big] \in \mathbb{R}^{n\times(p+1)}$
- $U = \big[u[0]\ u[1]\ \ldots\ u[p]\big] \in \mathbb{R}^{m\times(p+1)}$
- $\mathbf{W} \in \mathbb{S}^{n+m+p+1}$.

Note that $\mathbf{W}_{22}$ and $\mathbf{W}_{33}$ are two blocks of $\mathbf{W}$ playing the role of auxiliary variables, and that I_n denotes the $n\times n$ identity matrix.

Theorem 4: The computationally-cheap SDP relaxation is a convex relaxation of the ODC problem. Furthermore, the relaxation is exact if and only if it possesses a solution $(K^{\text{opt}}, X^{\text{opt}}, U^{\text{opt}}, \mathbf{W}^{\text{opt}})$ such that $\text{rank}\{\mathbf{W}^{\text{opt}}\} = n$.

The reader may refer to [21] for the proof of the above theorem. The matrix $\mathbf{W}$ in the computationally-cheap SDP relaxation always has rank greater than or equal to n, due to its block submatrix I_n. Notice that the number of rows for the SDP matrix of Problem D-1, D-2, S-1 or S-2 is on the order of np, whereas the number of rows for the computationally-cheap SDP matrix is on the order of $n+p$. We have empirically observed that this cheap relaxation has the same solution as the relaxation of Problem D-2 for diagonal Q, R and C.

VII. Numerical Example

In this part, the aim is to evaluate the performance of the proposed controller design technique on the *Mass-Spring* system, as a classical physical system. Consider a mass-spring system consisting of N masses. This system is exemplified in Figure 2 for $N=2$. The system can be modeled in the continuous-time domain as

$$\dot{x}_c(t) = A_c x_c(t) + B_c u_c(t) \tag{27}$$

where the state vector $x_c(t)$ can be partitioned as $[o_1(t)^T\ \ o_2(t)^T]$ with $o_1(t) \in \mathbb{R}^n$ equal to the vector of positions and $o_2(t) \in \mathbb{R}^n$ equal to the vector of velocities of the N masses. In this example, we assume that $N=10$ and adopt the values of A_c and B_c from [25]. The goal is to design a static sampled-data controller with a pre-specified structure (i.e., the controller is composed of a sampler, a static discrete-time controller and a zero-order holder). To this end, we first discretize the system with the sampling time of 0.4 second and denote the obtained system as

$$x[\tau+1] = Ax[\tau] + Bu[\tau], \qquad \tau = 0,1,\ldots \tag{28}$$

It is aimed to design a constrained controller $u[\tau] = Kx[\tau]$ to minimize the cost function

$$\sum_{\tau=0}^{p} \left(x[\tau]^T x[\tau] + u[\tau]^T u[\tau]\right) \tag{29}$$

for $x[0]$ equal to the vector of 1's and $\alpha = 0$. Since it is assumed that all states of the system can be measured locally (i.e., $C = I$), Problem D-2 turns out to have a sparse graph. Hence, we convexify the ODC problem using the SDP relaxation of Problem D-2. We solve an SDP relaxation for the six different control structures shown in Figure 3. The free parameters of each controller are colored in red in this figure. For example, Structure (c) corresponds to a fully decentralized controller, where each local controller has access to the position and velocity of its associated mass. Structure (d) enables some communications between the local control of Mass 1 and the remaining local controllers. For each structure, the SDP relaxation of Problem D-2 is solved for four different terminal times $p = 5,\ 10,\ 15$ and 30. The results are tabulated in Table I. Four metrics are reported for each structure and terminal time:

- *Lower bound:* This number is equal to the optimal objective value of the SDP relaxation, which serves as a lower bound on the minimum value of the cost function (29).
- *Upper bound:* This number corresponds to the cost function (29) at a near-optimal controller K_{no} recovered from the first column of the SDP matrix. This number serves as an upper bound on the minimum value of the cost function (29).
- *Infinite-horizon performance:* This is equal to the infinite sum $\sum_{\tau=0}^{\infty}\left(x[\tau]^T x[\tau] + u[\tau]^T u[\tau]\right)$ associated with the system (28) under the designed near-optimal controller.
- *Stability:* This indicates the stability or instability of the closed-loop system.

It can be observed that the designed controllers are always stabilizing for $p = 30$. As demonstrated in Table I, the upper and lower bounds are very close to each other in many scenarios, in which cases the recovered controllers are almost globally optimal.

VIII. Conclusions

This paper studies the optimal distributed control (ODC) problem for discrete-time systems. The objective is to design a fixed-order distributed controller with a pre-determined structure to minimize a quadratic cost functional. This paper proposes a semidefinite program (SDP) as a convex relaxation for ODC. The notion of treewidth is exploited to study the rank of the minimum-rank solution of the SDP relaxation. This method is applied to the static distributed control case and it is shown that the SDP relaxation has

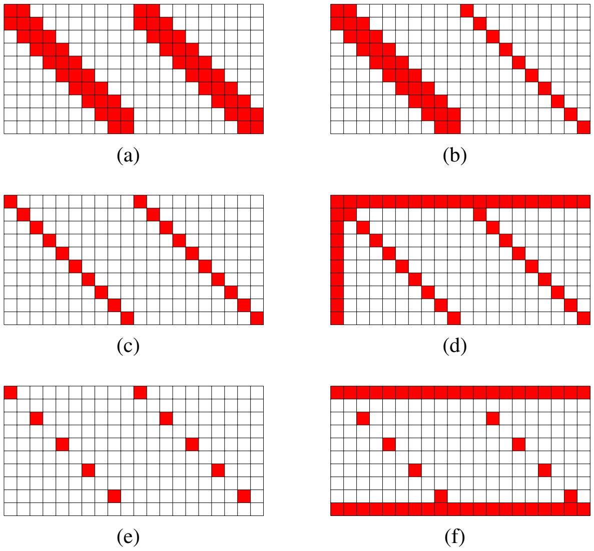

Fig. 3: Six different structures for the controller K: the free parameters are colored in red (uncolored entries are set to zero).

K	bounds	$p = 5$	$p = 10$	$p = 15$	$p = 30$
(a)	upper bound	126.752	140.105	140.681	140.691
	lower bound	126.713	140.080	140.660	140.690
	inf. horizon perf.	∞	∞	∞	140.691
	stability	unstable	unstable	unstable	stable
(b)	upper bound	126.809	140.183	140.685	140.702
	lower bound	126.713	140.080	140.661	140.690
	inf. horizon perf.	∞	∞	140.770	140.702
	stability	unstable	unstable	stable	stable
(c)	upper bound	127.916	140.762	140.792	140.795
	lower bound	126.713	140.080	140.660	140.690
	inf. horizon perf.	150.972	140.992	140.796	140.795
	stability	stable	stable	stable	stable
(d)	upper bound	127.430	140.761	140.762	140.761
	lower bound	126.713	140.080	140.661	140.690
	inf. horizon perf.	159.633	141.020	140.766	140.761
	stability	stable	stable	stable	stable
(e)	upper bound	175.560	235.240	240.189	242.973
	lower bound	167.220	215.202	222.793	226.797
	inf. horizon perf.	277.690	282.580	271.675	267.333
	stability	stable	stable	stable	stable
(f)	upper bound	175.401	230.210	231.022	230.382
	lower bound	164.114	208.484	214.723	216.431
	inf. horizon perf.	357.197	287.767	242.976	232.069
	stability	stable	unstable	stable	stable

TABLE I: The outcome of the SDP relaxation of Problem D-2 for the 6 different control structures given in Figure 3.

a matrix solution with rank at most 3. This result can be a basis for a better understanding of the complexity of the ODC problem because it states that almost all eigenvalues of the SDP solution are zero. It is also discussed that the same result holds true for the design of a dynamic controller for both deterministic and stochastic systems.

REFERENCES

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal of Control*, vol. 6, no. 1, 1968.

[2] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *Automatic Control, IEEE Transactions on*, vol. 53, no. 7, pp. 1616–1629, 2008.

[3] G. Dullerud and R. D'Andrea, "Distributed control of heterogeneous systems," *IEEE Transactions on Automatic Control*, vol. 49, no. 12, pp. 2113–2128, 2004.

[4] T. Keviczky, F. Borrelli, and G. J. Balas, "Decentralized receding horizon control for large scale dynamically decoupled systems," *Automatica*, vol. 42, no. 12, pp. 2105–2115, 2006.

[5] J. Lavaei, "Decentralized implementation of centralized controllers for interconnected systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1860–1865, 2012.

[6] M. Fardad, F. Lin, and M. R. Jovanovic, "On the optimal design of structured feedback gains for interconnected systems," *IEEE Conference on Decision and Control*, 2009.

[7] K. Dvijotham, E. Theodorou, E. Todorov, and M. Fazel, "Convexity of optimal linear controller design," *IEEE Conference on Decision and Control*, 2013.

[8] N. Matni and J. C. Doyle, "A dual problem in H_2 decentralized control subject to delays," *American Control Conference*, 2013.

[9] 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.

[10] P. Shah and P. A. Parrilo, "H_2-optimal decentralized control over posets: a state-space solution for state-feedback," *http://arxiv.org/abs/1111.1498*, 2011.

[11] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," *American Control Conference*, 2012.

[12] A. Lamperski and J. C. Doyle, "Output feedback H_2 model matching for decentralized systems with delays," *American Control Conference*, 2013.

[13] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2218–2223, 2011.

[14] A. Rantzer, "Distributed control of positive systems," *http://arxiv.org/abs/1203.0047*, 2012.

[15] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, 1996.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.

[17] S. Sojoudi and J. Lavaei, "On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part I," *IEEE Conference on Decision and Control*, 2013.

[18] S. Sojoudi and J. Lavaei, "On the exactness of semidefinite relaxation for nonlinear optimization over graphs: Part II," *IEEE Conference on Decision and Control*, 2013.

[19] J. Lavaei, "Optimal decentralized control problem as a rank-constrained optimization," *52th Annual Allerton Conference on Communication, Control and Computing*, 2013.

[20] R. Madani, G. Fazelnia, S. Sojoudi, and J. Lavaei, "Low-rank solutions of matrix inequalities with applications to polynomial optimization and matrix completion problems," *IEEE Conference on Decision and Control*, 2014.

[21] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, "Convex relaxation for optimal distributed control problem—part I: Time-domain formulation," *Technical Report*, 2014. [Online]. Available: http://www.ee.columbia.edu/~lavaei/Dec_Control_2014_PartI.pdf

[22] J. Lavaei and A. G. Aghdam, "Control of continuous-time LTI systems by means of structurally constrained controllers," *Automatica*, vol. 44, no. 1, 2008.

[23] A. Kalbat, R. Madani, G. Fazelnia, and J. Lavaei, "Efficient convex relaxation for stochastic optimal distributed control problem," *Allerton*, 2014.

[24] G. Fazelnia, R.Madani, A. Kalbat, and J. Lavaei, "Convex relaxation for optimal distributed control problem—part II: Lyapunov formulation and case studies," *Technical Report*, 2014. [Online]. Available: http://www.ee.columbia.edu/~lavaei/Dec_Control_2014_PartII.pdf

[25] F. Lin, M. Fardad, and M. R. Jovanovi, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, 2013.