Feedback Controller Sparsification for a Class of Linear Systems with Parametric Uncertainties

Reza Arastoo† and MirSaleh Bahavarnia‡

Abstract

We consider the problem of output feedback controller sparsification for systems with parametric uncertainties. We develop an optimization scheme that minimizes the performance deterioration caused by the sparsification process, while enhancing sparsity pattern of the feedback gain. In order to improve temporal proximity of an existing closed-loop system and its sparsified counterpart, we also incorporate an additional constraint into the problem formulation so as to bound the variation in the system output pre and post sparsification. We also show that the resulting non-convex optimization problem can be equivalently reformulated into a rank-constrained optimization problem. We then formulate a bi-linear minimization program along with an iterative algorithm to obtain a sub-optimal solution which satisfies the rank constraint with arbitrary tolerance. Lastly, a sub-optimal sparse controller design for IEEE 39-bus New England power network is utilized to showcase the effectiveness of our proposed method.

I. INTRODUCTION

In control theory, there has always been a desire to achieve the best possible performance, while taking into account the feasibility and cost of the communication between the subsystems. The reason behind such yearning is mostly rooted in the fact that unlike small-scale dynamical systems, where centralized control methodologies can efficiently be applied owing to the availability of information from all subsystems, the subsystem level information is not globally accessible throughout the network in medium to large scale systems. With the emergence and growth of ultra large-scale interconnected systems, such as power grids, transportation systems, and wireless data networks, exploiting the characteristics of the underlying structure of the system in the control design has become inescapable. Hence, the concepts of distributed and decentralized controllers have received increasing attention in recent years [1], [2], [3], [4].

It has been known that optimal controller design under controller structural constraints is a challenging problem. Nonetheless, numerous studies have been carried out to either propose controller design frameworks or reveal inherent structural properties of controllers for special classes of systems [5], [6], [7], [8]. Another concern in the design of large-scale control systems is the number of communication links between the subsystems which poses major issues especially when establishing links between nodes is very costly. Sparsifying the controller gain leads to fewer information pathways as well as fewer controller sensors and actuators. As a result, the design of controller gains with minimum number of non-zero entries can mitigate the communication overflow issues emergent in large

† R. Arastoo and M. Bahavarnia are with the Department of Mechanical Engineering and Mechanics, Packard Lab., Lehigh University, Bethlehem, PA. Email Addresses: {reza.arastoo,msbmo2008}@gmail.com
interconnected systems. In the sparsity-promoting control problem, the ultimate objective is to minimize the number of feedback links without losing much performance. This is achieved by incorporating additional functions into the optimization cost function to penalize the number of communication links. The problem has been addressed by a number of researchers, who opted for various techniques to tackle the inherently non-convex problem [1], [9], [10], [11], [12], [13].

For example, in [9], [12] the Alternating Direction Method of Multipliers is utilized to handle the non-convex terms in the problem formulation. In [10], the authors proposed a novel framework in which all non-convexities are lumped into a rank constraint further enabling it to address output feedback problems with norm constraints on the input/output signals. In some recent papers an unconventional approach to synthesize near optimal sparse controllers has been adopted. This proposed sparse controller design framework is founded based on the assumption that a pre-designed well-performing controller is available and the ultimate goal is to obtain a sparse feedback controller approximating the attributes and qualities of the original well-preforming controller [14], [15], [16], [17].

In this paper, we extend the work published in [14], [15]. In [15], deterministic linear systems are considered and it is shown that the optimal sparse control problem for this class of systems can be solved by applying an ADMM algorithm to its equivalent rank-constrained optimization problem. The work published in [14] expanded the applicability of the method introduced in [15] to linear systems with parametric time-varying uncertainties. We showed that by utilizing $H_2$ and $H_\infty$ control theorems [18], [19], this novel approach can be easily applied to design sparse robust controllers by equivalently reformulating the original problem into a rank-constrained optimization where all non-convexities are collected into a rank constraint. The next notable improvement in [14] is that the ADMM algorithm, used in [15] with no proof of convergence, is replaced by our novel convergent algorithm, which employs a bi-linear optimization to reach the sub-optimal solution of the rank-constrained optimization problem. In this paper, we provide detailed proof of all lemmas and theorems presented in [14] without proofs. We also show that the optimization parameters can be tuned such that the optimal solution of our proposed minimization problem satisfies the rank constraint with arbitrary tolerance. Furthermore, we present discussions on the convergence of the proposed solving algorithm and provide remarks on the choice of the weight matrix to improve the sparsification performance without adversely affecting the convergence of the algorithm. The simulation section is also substantially extended to provide more vivid insights into advantages of applying the proposed procedure to synthesize sparse controllers for power networks. It further examines in more detail the impact of the structure and magnitude of the power network uncertainties on the robustness of the closed-loop systems as well as the sparsity level of the controllers.

This paper is structured as follows. Section II provides key definitions and notations used throughout the paper. In Section III, we formally state the problem we aim to solve. In Sections IV and V, we elaborate how our problem can equivalently be reformulated into an optimization problem constrained to several linear matrix inequalities and a fixed rank constraint. Section VI provides insight into our proposed algorithm and states several related results. The results of our numerical simulations are presented in Section VII. Finally, we end with concluding remarks in Section VIII.
II. Notations

Throughout this paper, matrices are customarily referred to with capital letters, and the entries are referred to with the corresponding lower-case letters, using subscripts. The vectors, on the other hand, are symbolized by lower-case letters with components denoted by the same letter using subscripts. A unit vector with its $i^{th}$ entry equal to one is denoted by $e_i$. The set of real numbers is denoted by $\mathbb{R}$. The space of $n$ by $m$ matrices with real entries is indicated by $\mathbb{R}^{n \times m}$. The set of real matrices with non-negative (positive) entries is represented by $\mathbb{R}^{n \times m}_+(\mathbb{R}^{n \times m}_{++})$. The $n$ by $n$ identity matrix is denoted by $I_n$. The vector of singular values of matrix $X$ is denoted by $\sigma(X)$. The entry-wise product of two matrices, i.e., Hadamard product, is represented by $\circ$. If $X = [x_{ij}]$, then the matrix $|X|$ is the entry-wise absolute value of $X$, i.e., $|X| = [|x_{ij}|]$. The number of non-zero entries of a matrix is denoted by $\| \cdot \|_0$ while $\| \cdot \|_1$ denotes $\ell_1$ norm, and the $L_2$-norm is defined by

$$
\|x\|_{L_2(\mathbb{R}^n)} := \left( \int_0^\infty \|x(t)\|^2_2 \, dt \right)^{1/2}.
$$

Whenever it is not confusing, we use $L_2$ instead of $L_2(\mathbb{R}^n)$. $\operatorname{Tr}(\cdot)$ and $\operatorname{rank}(\cdot)$ denote the trace and rank of the matrix operands, respectively. The operator $\operatorname{diag}(\cdot)$ constructs block diagonal matrix from input arguments.

**Definition 1**: For a given $\epsilon > 0$ and matrix $X$, we say that rank of $X$ is $k$ with tolerance $\epsilon$, and it is denoted by $\operatorname{rank}(X; \epsilon)$, if exactly $k$ singular values of $X$ are larger than or equal to $\epsilon$.

A matrix is said to be Hurwitz if all of its eigenvalues lie within the open left half of the complex plane. A real symmetric matrix is said to be positive definite (semi-definite) if all of its eigenvalues are positive (non-negative). $\mathbb{S}^n_+(\mathbb{S}^n_{++})$ denotes the space of positive definite (positive semi-definite) real symmetric matrices, and the notation $X \succeq Y$ ($X \succ Y$) means $X - Y \in \mathbb{S}^n_+$ ($X - Y \in \mathbb{S}^n_{++}$).

**Remark 1**: For simplicity of our notations, we will use a new notation in statements of theorems, where we use an asterisk ‘*’ to represent the upper triangular sub-blocks of symmetric matrices. Moreover, in the occasions when the optimal solutions of the optimization problems in these theorems do not depend on some of the sub-blocks of matrices, we use a dash ‘-’ to represent such sub-blocks with no apparent utilization in the problem.

III. Problem Formulation

A. LTI Systems with Parametric Uncertainties

The focus of this paper is on the following class of uncertain linear time-invariant (LTI) systems that are defined by the state space realization

$$
\begin{align*}
\dot{x}(t) &= [A + \Delta_A]x(t) + [B_1 + \Delta_{B_1}]u(t) + B_2d(t) \\
y(t) &= Cx(t),
\end{align*}
$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input and $d(t) \in \mathbb{R}^p$ represents the exogenous disturbance input. We assume that the matrices $A \in \mathbb{R}^{n \times n}$, $B_1 \in \mathbb{R}^{n \times m}$, $B_2 \in \mathbb{R}^{n \times p}$, and $C \in \mathbb{R}^{q \times n}$ are constant real matrices describing the dynamics of the nominal system, whereas $\Delta_A$ and $\Delta_{B_1}$ represent the parameter

\footnote{It is assumed that the pair $(A, B_1)$ is controllable and $(A, C)$ is detectable.}
uncertainties of the matrices $A$ and $B_1$, respectively. In this paper, we consider a special uncertainty structure expressed by

$$
\begin{bmatrix}
\Delta_A & \Delta_{B_1}
\end{bmatrix} = D \Delta
\begin{bmatrix}
E_A & E_{B_1}
\end{bmatrix},
$$

where $D$, $E_A$ and $E_{B_1}$ are known constant real matrices with appropriate dimensions, which characterize the structure of the uncertainties, while $\Delta$ is an unknown $i$ by $j$ real matrix which is constrained by

$$
\Delta^T \Delta \preceq \rho^2 I_j.
$$

This class of uncertain linear systems was initially reported by Petersen in papers [20], [21] and later thoroughly addressed by Khargonekar et al. [22].

B. Controller Sparsification via $\mathcal{H}_p$ Approximations

Suppose that a pre-designed well-performing controller, namely $\hat{K}$, is readily available and the nominal system controlled by such a controller, represented by $\hat{S}$, has all the desired characteristics. The objective is to synthesize a constant gain output feedback controller of the form

$$
u(t) = Ky(t), \quad K \in \mathcal{K}
$$

with minimum number of non-zero entries, while minimizing the performance deterioration from that of the closed-loop system $\hat{S}$ under parametric uncertainties. In (4), $\mathcal{K}$ denotes a set of admissible feedback gains which holds desirable properties such as pre-defined communication layout.

**Assumption 1:** It is assumed that the set $\mathcal{K}$ is convex.

It should be emphasized that this assumption does not offer any premise on characterization of the set of all stabilizable output feedback controllers. In our follow-up discussions, we will show that if our proposed optimal control design is feasible, then the resulting output feedback controller will be stabilizing and satisfy the structural constraint $K \in \mathcal{K}$. There are numerous applications associated with such convexly constrained controller design, such as power grids or multi-UAV systems. It is sometimes practically infeasible to establish some specific communication links between particular nodes due to the nodes distant locations or security issues in networks. There are also cases where the attenuation/amplification in certain feedback paths is upper bounded, due to technological shortcomings. Such restrictions are addressed by forcing the corresponding controller entries to be contained in a convex set.

Our goal is to solve the following $\ell_0$-regularized optimal control problem to compute a sparse output feedback controller under parametric uncertainties

$$
\begin{align*}
\text{minimize} & \quad \varepsilon S + \lambda_1 \varepsilon y + \lambda_2 \|K\|_0 \\
\text{subject to} & \quad K \in \mathcal{K} \\
S & : \text{Stable} \\
\|y_S - y_{\hat{S}}\|_{\ell_2} < \varepsilon_y \|d\|_{\ell_2} \\
\|S - \hat{S}\|_{\mathcal{H}_2} < \varepsilon_S
\end{align*}
$$

(5a) (5b) (5c) (5d) (5e)
in which $\| \cdot \|_{\mathcal{H}_2}$ is the well-known $\mathcal{H}_2$ norm. Nominal closed-loop system $\hat{S}$ is a previously designed desired optimal closed-loop system with output signal $y_{\hat{S}}$ and $S$ is the resulting system by closing the loop using sparse feedback controller $K$. The output signal of $S$ is denoted by $y_{S}$. In order to promote sparsity of feedback gain matrix $K$, the $\ell_0$ measure of $K$, which is denoted by $\| K \|_0$, has been added to the cost function. Two design parameters $\lambda_1$ and $\lambda_2$ are introduced to achieve desired trade-off between performance loss and sparsity.

In the optimal control problem (5), constraint (5e) is included to ensure that the nominal closed-loop system $\hat{S}$ is well-approximated by a closed-loop system controlled by a sparse controller $K$. To enhance temporal features of our approximation, we also incorporate another requirement into our design scheme, characterized by constraint (5d). This constraint guarantees that the energy level of the difference between the output signals of the two closed-loop systems remains under a pre-specified level $\epsilon_y$ when both closed-loop systems are excited by a disturbance input $d$ with unit norm.

The goal of this paper is to study the effect of parametric uncertainties on the best achievable levels of sparsity. However, finding the optimal solution of the problem (5) is inherently NP-hard; see our discussion in Section V. In Section VI, we will propose a tractable approximation algorithm to solve this problem. The following sections discuss the equivalent problem reformulation exploited in numerically solving our optimization problem.

IV. Equivalent Reformulation

The first two terms in the cost function of the optimization problem (5) can be simplified into the $\mathcal{H}_2/\mathcal{H}_\infty$ norms of an augmented system, namely $\bar{S}$, constructed by the following state space realization matrices

$$
\bar{A} = \text{diag}(\bar{A}_{11}, A + B_1 \hat{K} C),
$$

$$
\bar{B} = \begin{bmatrix} B_2^T & B_2^T \end{bmatrix}^T, \quad \bar{C} = \begin{bmatrix} C & -C \end{bmatrix},
$$

where $\bar{A}_{11} = [A + \Delta_A] + [B_1 + \Delta_{B_1}] K C$. As it can be seen, the system $\bar{S}$ represents the difference between the nominal system controlled by the pre-designed controller and the uncertain system, stabilized by closing its feedback loop using a sparse controller. Hence, we can re-formulate our problem into the $\mathcal{H}_2/\mathcal{H}_\infty$ norm minimization of the augmented system as follows:

$$
\min K, \epsilon_S, \epsilon_y \quad \max \epsilon_S + \lambda_1 \epsilon_y + \lambda_2 \| K \|_0
$$

subject to $K \in \mathcal{K}$, $\bar{A}_{11}$ Hurwitz,

$$
\| \bar{C}(sI - \bar{A})^{-1} \bar{B} \|_{\mathcal{H}_\infty} < \epsilon_y,
$$

$$
\| \bar{C}(sI - \bar{A})^{-1} \bar{B} \|_{\mathcal{H}_2} < \epsilon_S.
$$

In problem (7), the attempt is to minimize the worst case gap between the frequency response of the systems in the sense of a weighted sum of the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms. Therefore, unlike the design schemes introduced in [9], [10], the approach proposed in this paper allows us to exploit the advantages offered by other controller design schemes in the sparse controller design. In the next section, we show that the optimization problem (7) includes bi-linear
matrix inequality constraints mainly due to the existence of the Lyapunov stability conditions. Here, we intend to employ the idea of lumping all nonlinear constraints into a rank-constrained problem, proposed in [10], to rewrite problem as a rank-constrained optimization. Based on the obtained reformulation, it is possible to either develop heuristics to sub-optimally solve the problem or provide necessary and sufficient conditions for the feasibility of the points with particular desired costs.

V. FIXED RANK OPTIMIZATION REFORMULATION

The approach adopted in this paper is based on solving the problem of sparse controller approximation via rank-constrained optimization. Hence, we start by stating the main lemmas which helps us cast the constraints of the optimization problem as rank-constrained linear matrix inequalities.

Lemma 1 ([10]): Let $U \in \mathbb{R}^{n \times n}$, $V \in \mathbb{R}^{n \times m}$, $W \in \mathbb{R}^{m \times m}$, and $Y \in \mathbb{R}^{m \times n}$, with $U \succ 0$. Then, $\text{rank}(M) = n$ if and only if $W = YU^T$, $V^T = YU$, and $Z = U^{-1}$ where

$$M = \begin{bmatrix} U & V & I_n \\ V^T & W & Y \\ I_n & Y^T & Z \end{bmatrix}.$$  

The above lemma can be utilized to collect almost all non-convex terms of the optimization problems in one and only one constraint in the form of a rank constraint. There are a number of algorithms proposed to solve rank-constrained optimization problems [23], [24], [25], [26], [27]. In this manuscript, we aim to render such algorithms applicable in solving our inherently nonlinear controller sparsification problem by collecting various forms of non-convex/combinatorial constraints into a fixed rank constraint.

As a first step, we show how the $H_2$ norm of an uncertain system can be formulated by rank-constrained linear matrix inequalities.

Lemma 2: Given a strictly proper uncertain linear system $P$ with state space realization $(A + \Delta A, B, C)$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{q \times n}$, $\Delta A = \mathcal{D}\Delta \mathcal{E}$ and $\Delta^T\Delta \preceq \rho^2 I_j$, then $P$ is stable and $\|P\|_2^2 < \gamma$ if and only if there exists a positive definite matrix $X \succ 0$ and a positive scalar $\varepsilon$ such that

$$\text{Tr}(CXC^T) < \gamma,$$

$$\begin{bmatrix} \sqrt{\rho} Y_1^T + BB^T + \varepsilon D^T \rho \sqrt{\rho} Y_2 \\ \varepsilon I_j \end{bmatrix} < 0,$$

$$\begin{bmatrix} X & * & * \\ * & * & * \\ * & * & * \\ I_n & A^T & \varepsilon^T \end{bmatrix} = n.$$

Proof 1: The system $P$ is stable with $H_2$ norm less than $\gamma$ if and only if there exists a positive definite matrix $X$ such that [28, p. 210]

$$\text{Tr}(CXC^T) < \gamma,$$

$$(A + \Delta A)X + X(A + \Delta A) + BB^T < 0.$$
Substituting $\Delta_A = D\Delta E$ into the second equation, we have
\[ AX + XA^T + BB^T + D\Delta EX + X(D\Delta E)^T \prec 0. \]
Since the term $AX + XA^T + BB^T$ is symmetric and $\Delta^T \Delta \preceq \rho^2 I_j$, the above linear matrix inequality holds if and only if there exists a positive scalar $\varepsilon > 0$ such that [18]
\[ AX + XA^T + BB^T + \varepsilon \rho DD^T + \varepsilon^{-1} \rho X^T EX \prec 0. \]
This is equivalent to having
\[
\begin{bmatrix}
AX + XA^T + BB^T + \varepsilon \rho DD^T & \sqrt{\rho}(EX)^T \\
\sqrt{\rho}(EX) & \varepsilon I_j
\end{bmatrix} \prec 0.
\]
Applying Lemma 1, the last LMI can equivalently be rewritten as shown below.
\[
\begin{bmatrix}
y_1 + y_1^T + BB^T + \varepsilon \rho DD^T & \sqrt{\rho}y_2 \\
\sqrt{\rho}y_2^T & -\varepsilon I_j
\end{bmatrix} \prec 0,
\]
\[
\begin{bmatrix}
X & y_1 & y_2 \\
I_n & A^T & E^T
\end{bmatrix} = n.
\]
Augmenting proper rows and columns to the rank-constrained matrix to make it symmetric completes our proof.

Similar to Lemma 2, which paves the way in casting the $H_2$ norm term in our optimal controller sparsification problem, as a rank-constrained optimization problem, the $H_\infty$ norm term of problem (7) can also be equivalently represented with a set of rank-constrained linear matrix inequalities. In the next lemma, we prove such equivalence, which later helps in accommodating the whole problem of controller sparsification under parametric uncertainties into the framework of rank-constrained optimization.

Lemma 3: Suppose a strictly proper uncertain LTI plant $P$, represented in the state space triplet $(A + \Delta_A, B, C)$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{q \times n}$, $\Delta_A = D\Delta E$ and $\Delta^T \Delta \preceq \rho^2 I_j$, then the system is stable with $H_\infty$ norm less than $\gamma$ if and only if there exists a positive definite matrix $X > 0$ and a positive scalar $\varepsilon > 0$ satisfying
\[
\begin{bmatrix}
y_1 + y_1^T + \varepsilon \rho DD^T & * & * \\
B^T & -\gamma I_m & * \\
(CX) & 0 & -\gamma I_q & * \\
\sqrt{\rho}y_2^T & 0 & 0 & -\varepsilon I_j
\end{bmatrix} \prec 0,
\]
\[
\begin{bmatrix}
X & * & * \\
y_1^T & \gamma_1^T & * \\
y_2^T & \gamma_2^T & * \\
I_n & A^T & E^T
\end{bmatrix} = n.
\]

Proof 2: Employing Lemma 7.4 in [28, p. 221], the plant $P$ is stable with $\|P\|_\infty < \gamma$ if and only if there exists a positive definite matrix $Z$ such that
\[ Z(A + \Delta_A) + (A + \Delta_A)^T Z + C^T C + \gamma^{-2} ZBB^T Z \prec 0. \]
Pre and post multiplying the above LMI by the inverse of \( Z \), namely \( Y \), we have

\[
(A + \Delta A)Y + Y(A + \Delta A)^T + YC^TCY + \gamma^{-2}BB^T) < 0,
\]

\[
\Delta AY + YA^T + \gamma^{-2}BB^T + YC^TCY + \Delta AY + AY^T < 0.
\]

Plugging \( \Delta A = D\Delta E \) into the previous inequality, we get

\[
AY + YA^T + \gamma^{-2}BB^T + YC^TCY + (D\Delta E)Y + Y(D\Delta E)^T < 0.
\]

Having \( \Delta^T\Delta \preceq \rho^2 I_j \), the above inequality is valid for all acceptable values of \( \Delta \) if and only if there exists a positive \( \bar{\varepsilon} > 0 \) for which the following LMI holds.

\[
AY + YA^T + \gamma^{-1}BB^T + \gamma^{-1}XC^TCX + \bar{\varepsilon}\rho DD^T + \bar{\varepsilon}^{-1}\rho\varepsilon^T\varepsilon Y < 0.
\]

Multiplying both sides of the above inequality by \( \gamma \), we will have

\[
AX + XA^T + \gamma^{-1}BB^T + \gamma^{-1}XC^TCX + \bar{\varepsilon}\rho DD^T + \bar{\varepsilon}^{-1}\rho\varepsilon^T\varepsilon X < 0,
\]

where \( X = \gamma Y \) and \( \varepsilon = \bar{\varepsilon}\gamma \). It can effortlessly be verified that the last LMI is the Schur complement of the negative definite constraint

\[
\begin{bmatrix}
\gamma_1 + \gamma_1^T + \varepsilon\rho DD^T & * & * & * \\
\ast & -\gamma I_m & * & * \\
\ast & 0 & -\gamma I_q & * \\
\ast & \sqrt{\rho}\varepsilon X & 0 & -\varepsilon I_j \\
\end{bmatrix} < 0,
\]

where \( \gamma_1 = XA^T \). The rest of the proof is a mere application of Lemma 1; hence, omitted.

Consequently, we can reformulate the problem (7) into a rank-constrained problem, as described in the sequel.

**Theorem 4:** The optimization problem (7) is equivalent to the following rank-constrained optimization problem

\[
\begin{align*}
\text{minimize} & \quad \varepsilon_S + \lambda_1\varepsilon_y + \lambda_2\|K\|_0 \\
\text{subject to} & \quad K \in \mathcal{K}, \\
& \quad X_r > 0, \quad r = 1, 2, \\
& \quad \varepsilon_r > 0, \quad r = 1, 2, \\
& \quad \text{Tr}(CX_1C^T) < \varepsilon_S, \\
& \quad \begin{bmatrix}
P_1 + B\bar{B}^T & * \\
\sqrt{\rho}Y_2^T & -\varepsilon_1 I_j \\
\end{bmatrix} < 0, \\
& \quad \begin{bmatrix}
P_2 & * & * & * \\
B^T & -\varepsilon_y I_p & * & * \\
(CX_2) & 0 & -\varepsilon_y I_q & * \\
\sqrt{\rho}Y_4^T & 0 & 0 & -\varepsilon_2 I_j \\
\end{bmatrix} < 0, \\
& \quad \text{rank}(M_1) = 2n,
\end{align*}
\]
where

\[ P_r = Y_{2r-1} + Y_{2r-1}^T + \varepsilon_r \rho \bar{D} \bar{D}^T, \quad r = 1, 2, \]

\[
M_1 = \begin{bmatrix}
X_1 & * & * & * \\
Y_1^T & * & * & * \\
Y_2^T & * & * & * \\
X_2 & Y_3 & Y_4 & * \\
I_{2n} & A_{cl}^T & E_{cl}^T & -
\end{bmatrix},
\]

\[
A_{cl} = \text{diag}(A + B_1 KC, A + B_1 \tilde{K} C) \in \mathbb{R}^{2n \times 2n},
\]

\[
\bar{D} = \begin{bmatrix} D^T & 0 \end{bmatrix}^T \in \mathbb{R}^{2n \times 1},
\]

\[
E_{cl} = \begin{bmatrix} E_A + E_B KC & 0 \end{bmatrix} \in \mathbb{R}^{j \times 2n}.
\]

**Proof 3:** It can be observed that the closed-loop system can be represented using the state representation \((A_{cl} + \Delta A, \bar{B}, \bar{C}, 0)\), where \(\bar{B}\) and \(\bar{C}\) are defined in (6) and \(\bar{A} = \bar{D} \Delta E_{cl}\). Therefore, applying the results from lemmas 2 and 3 yields the desired result.

The next corollary is now immediate.

**Corollary 5:** The optimization problem (7) can equivalently be cast as the following rank-constrained optimization problem

\[
\begin{align*}
\text{minimize} & \quad \varepsilon_S + \lambda_1 \varepsilon_y + \lambda_2 \|K\|_0 \\
\text{subject to} & \quad K \in \mathcal{K} \\
& \quad X_r > 0, \quad r = 1, 2, \\
& \quad \varepsilon_r > 0, \quad r = 1, 2, \\
& \quad \text{Tr}(\bar{C} X_1 \bar{C}^T) < \varepsilon_S, \\
& \quad \begin{bmatrix} Q_1 + \bar{B} \bar{B}^T + \varepsilon_1 \rho \bar{D} \bar{D}^T & * \\
\sqrt{\rho} R_1 & -\varepsilon_1 I_j \end{bmatrix} \prec 0, \\
& \quad \begin{bmatrix} Q_2 + \varepsilon_2 \rho \bar{D} \bar{D}^T & * & * \\
\bar{B}^T & -\varepsilon_y I_p & * \\
(\bar{C} X_2) & 0 & -\varepsilon_y I_q & * \\
\sqrt{\rho} R_2 & 0 & 0 & -\varepsilon_2 I_j \end{bmatrix} \prec 0, \\
& \quad \text{rank}(M_2) = 2n.
\end{align*}
\]
where
\[ Q_r = X_rA_o^T + A_oX_r + Y_rB_K^T + B_K^TY_r^T, \quad r = 1, 2 \]  \hspace{1cm} (10a)
\[ R_r = E_oX_r + E_{B_1}Y_r^T, \quad r = 1, 2 \]  \hspace{1cm} (10b)

\[ M_2 = \begin{bmatrix}
  X_1 & * & * & *
  \\
  Y_1^T & & & \\
  X_2 & Y_2 & & *
  \\
  I_{2n} & (KC_K)^T & & \\
\end{bmatrix} \]  \hspace{1cm} (10c)

\[ A_o = \text{diag}(A, A + B_1\hat{K}C) \in \mathbb{R}^{2n \times 2n}, \]  \hspace{1cm} (10d)
\[ D = \begin{bmatrix}
  D^T & 0
\end{bmatrix}^T \in \mathbb{R}^{2n \times i}, \]  \hspace{1cm} (10e)
\[ E_o = \begin{bmatrix}
  E_A & 0
\end{bmatrix} \in \mathbb{R}^{j \times 2n}, \]  \hspace{1cm} (10f)
\[ C_K = \begin{bmatrix}
  C & 0
\end{bmatrix} \in \mathbb{R}^{q \times 2n}, \]  \hspace{1cm} (10g)
\[ B_K = \begin{bmatrix}
  B_{1}^T & 0
\end{bmatrix}^T \in \mathbb{R}^{2n \times m}. \]  \hspace{1cm} (10h)

VI. A Tractable Approximation Algorithm for Computing Sparse Feedback Controllers

Although both optimizations (8) and (9) can be utilized to solve our controller sparsification problem, we choose to only implement the one formulated in (9). The terms in our optimization problem are all convex except the sparsity-promoting term in the cost function and the rank constraint. This section intends to shed light on our approach in dealing with these two non-convex and combinatorial terms.

As for the sparsity-promoting term of the objective function, since the \( \ell_0 \) norm is an integer-valued function, utilizing it in our formulation introduces the complications of combinatorial optimization. In order to reduce the complexity of sparse vector/matrix recovery problems, we employ the \( \ell_1 \) norm and its weighted versions. This is because convex surrogates of the \( \ell_0 \) norm are among the most common functions used to measure the sparsity and have been utilized in diverse applications [9], [29]. Therefore, we have

\[
\begin{align*}
\text{minimize} & \quad \varepsilon_S + \lambda_1\varepsilon_y + \lambda_2\|W \circ K\|_1 \\
\text{subject to} & \quad (9b) - (9h), \\
& \quad (10a) - (10b),
\end{align*}
\]  \hspace{1cm} (11)

where the weight matrix \( W = [w_{ij}] \in \mathbb{R}^{m \times q} \) is entry-wise positive and chosen according to the objectives of the problem.

The convex relaxation of the sparsity-promoting term in the cost function of (11) leaves us with an optimization problem in which non-convexity only arises in the form of a rank constraint, i.e., \( \text{rank}(M_2) = 2n \). It is known that presence of the rank constraint still causes our optimization problem to become NP-hard. Therefore, we propose a technique, which is built upon the method studied in [30], to solve the rank constraint optimization problem. In
a nutshell, this method is based on substituting the rank constraint on the symmetric matrix $M_2$ with a positive semidefinite constraint while introducing extra convex constraints along with a bi-linear term to the cost function.

**Theorem 6:** Let us consider the rank-constrained optimization problem (11) and define the following auxiliary optimization problem

$$
\begin{align*}
\text{minimize} & \quad \varepsilon S + \lambda_1 \varepsilon_y + \lambda_2 \| W \circ K \|_1 + \nu \text{Tr}(YM_2) \\
\text{subject to} & \quad (9b) - (9g), \\
& \quad (10a) - (10h), \\
& \quad 0 \preceq Y \preceq I_{6n+m}, \\
& \quad \text{Tr}(Y) = 4n + m, \\
& \quad M_2 \succeq 0,
\end{align*}
$$

in which $\lambda_1, \lambda_2, \nu > 0$ and the entry-wise positive matrix $W$ are some given design parameters. If problem (11) is feasible, then there exists a constant $\eta > 0$ for which the optimal solution $M_2$ from solving (12) satisfies

$$\text{rank}(M_2; \eta \nu^{-1}) \leq 2n,$$

i.e., rank of $M_2$ is less than or equal to $2n$ with tolerance threshold $\eta \nu^{-1}$ according to Definition 1.

**Proof 4:** Examining the optimization problem

$$
\begin{align*}
\text{minimize} & \quad \varepsilon S + \lambda_1 \varepsilon_y + \lambda_2 \| W \circ K \|_1 + \nu \text{Tr}(YM_2) \\
\text{subject to} & \quad (9b) - (9g), \\
& \quad (10a) - (10h), \\
& \quad 0 \preceq Y \preceq I_{6n+m}, \\
& \quad \text{Tr}(Y) = 4n + m,
\end{align*}
$$

and considering the assumption of feasibility of (11), it is straightforward to show that the optimal cost of this problem is equal to the optimal cost of (11) for all values of $\nu$. This is due to the fact that the optimal value of $\text{Tr}(YM_2)$, which can be shown to be the sum of the $4n + m$ smaller singular values of matrix $M_2$ [31, p.266], is always equal to zero, since $\text{rank}(M_2) = 2n$.

Also, it is easy to establish that the feasible set of the optimization problem (12) is a super-set of that of the (13). Hence, the optimal cost of (12) is bounded above by the optimal cost of (13). Therefore, denoting the optimal cost of (11) by $J^*$, we can write

$$
J^* \geq \varepsilon_S^*(\nu) + \lambda_1 \varepsilon_y^*(\nu) + \lambda_2 \| W \circ K^*(\nu) \|_1 \\
+ \nu \text{Tr}(Y^*(\nu)M_2^*(\nu)),
$$
Algorithm 1: Solution to problem (12)

Inputs: \(A, B_1, B_2, C, Q, R, \lambda_1, \lambda_2, \nu, K, W, \rho, \) and \(\varepsilon^*.\)

1: **Initialization:**
   Set \(Y^{(0)} = I_{6n+m}, \varepsilon^{(0)} > \varepsilon^*, K^{(0)} = 0_{m \times q}\) and \(k = 0\).

2: **While** \(\varepsilon^{(k)} > \varepsilon^*\) **do**
   3: Update \(Z^{(k+1)}\) by solving (15),
   4: Update \(Y^{(k+1)}\) using the equation (16),
   5: Update \(\varepsilon^{(k+1)}\) using the equation (17),
   6: \(k \leftarrow k + 1\),
   7: **end while**

8: Truncate \(K\).

**Output:** \(K\)

where \(\varepsilon^*_x(\nu), \varepsilon^*_y(\nu), \nu^*(\nu), Y^*(\nu), \) and \(M^*_2(\nu)\) are the optimal solutions to (12) for that particular value of \(\nu\).

Since the first three terms on the right hand side of the above inequality are non-negative, we can deduce

\[
\text{Tr}(Y^*(\nu)M^*_2(\nu)) \leq J^*\nu^{-1}.
\]

Therefore, the sum of \(4n + m\) smaller singular values of matrix \(M_2\) is less or equal to \(J^*\nu^{-1}\). Taking the constant \(\eta\) greater than or equal to \(J^*\), we have

\[
\text{rank}(M_2; \eta\nu^{-1}) \leq 2n.
\]

Hence, the proof is complete.

We should remind that according to (10c) and the specific structure of matrix \(M_2\) it is always true that \(\text{rank}(M_2) \geq 2n\). As a result of the previous theorem, we can now solve the optimization problem (12) for an appropriately-chosen parameter \(\nu\) to obtain a sub-optimal solution to the problem (11). For the simplicity of our notations, the letter \(Z\) is used to denote the stack of all optimization variables excluding variable \(Y\). The optimization problem (12) can be rewritten as follows.

\[
\begin{align*}
\text{minimize} & \quad F(Z, Y) \\
\text{subject to} & \quad Z \in C_z, \quad Y \in C_y,
\end{align*}
\]

where \(C_z\) is the convex set defined by the constraints (9b)-(9g), (10a)-(10h), along with \(M_2 \succeq 0\), and the convex set \(C_y\) is generated by \(\text{Tr}(Y) = 4n + m\) and \(0 \preceq Y \preceq I_{6n+m}\). Needless to say that \(F(Z, Y)\) represents the bi-linear objective function in the minimization problem (12). The above reformulation allows us to carry out this problem by iteratively optimizing the objective function for \(Z\) and \(Y\). As a result, the main steps of this iterative method can be divided into two sub-problems, \(Z\)-minimization and \(Y\)-minimization problems. As both \(Z\)-minimization and \(Y\)-minimization steps are convex optimizations, they can be performed in a computationally efficient manner. However, for the \(Y\)-minimization, there also exists an analytic solution, stated in the next theorem.
Theorem 7: The optimal solution to the $Y$-minimization step is given by

$$Y^* = I_{6n+m} - \sum_{i=1}^{2n} u_i u_i^T,$$

where vectors $u_i$ for $i = 1, \ldots, 2n$ are the singular vectors corresponding to the $2n$ larger singular values of $M_2$.

Proof 5: The optimal value of $\text{Tr}(Y M_2)$, subject to $0 \preceq Y \preceq I_{6n+m}$ and $Y \succeq 0$, is the sum of the $4n + m$ smaller singular values of matrix $M_2$ [31, p.266]. The rest of the proof is straightforward.

A. Summary of the Algorithm

We utilize the following sequence of iterations to obtain the minimizer of the constrained problem (12). First, we solve the $Z$-minimization and $Y$-minimization subproblems

$$Z^{(k+1)} = \arg \min_{Z \in C} F(Z, Y^{(k)}),$$

$$Y^{(k+1)} = I_{6n+m} - \sum_{i=1}^{2n} u_i^{(k+1)} u_i^{(k+1)T},$$

where $M_2^{(k+1)} = \sum_{i=1}^{6n+m} \sigma_i^{(k+1)} u_i^{(k+1)} u_i^{(k+1)T}$ is the singular value decomposition of $M_2^{(k+1)}$. The stopping criterion is established by $\varepsilon^{(k+1)} \leq \varepsilon^*$, where $\varepsilon^*$ is the given desired precision, with the following update law

$$\varepsilon^{(k+1)} = \frac{\|K^{(k+1)} - K^{(k)}\|}{\|K^{(k+1)}\|}.$$

In the last step of the algorithm, we truncate negligible entries of the resulting feedback gain $K$, e.g., those smaller than $5 \times 10^{-5}$. These small entries show very weak couplings between the nodes in the information structure of the controller. A summary of our proposed algorithm is described in Algorithm 1.

Remark 2: Let $\{(Z^{(k)}, Y^{(k)})\}$ be the sequence generated by Algorithm 1. Then, $\{F(Z^{(k)}, Y^{(k)})\}$ is a monotonically decreasing sequence and, hence, convergent.

Remark 3: The choice of the weight matrix $W$ plays an important role in the sparsity-promoting properties of our method. When a proper weight matrix is not accessible, the weighted $\ell_1$ norm technique can also be employed to enhance the sparse controller recovery. In this method, the weight assigned to each controller entry is updated inversely proportional to the value of the corresponding matrix entry recovered from the previous iteration, i.e.

$$w_{ij}^{(k+1)} = 1/(|k_{ij}^{(k)}| + \xi), \quad \forall i, j,$$

where the constant $\xi > 0$ which is chosen as a relatively small constant, is augmented to the denominator of the update law (18) to guarantee the stability of the algorithm, especially, when $k_{ij}^{(k)}$ turns out to be zero in the previous iteration [32]. It should be noted that, our simulation results are obtained by performing this update law only in the first few iterations to avoid adversely affecting the convergence of the algorithm.
VII. Simulation Results

In this section, we examine our proposed method by utilizing the IEEE 39-Bus New England power system which consists of $N_G = 10$ synchronous generators. Specifically, we take advantage of the state-space model provided by [33], which is a linearized state-space model of swing equations is characterized by (1), where

$$x = \begin{bmatrix} \theta^T & \omega^T \end{bmatrix}^T,$$

$$A = \begin{bmatrix} 0 & I \\ -\tilde{M}^{-1}L & -\tilde{M}^{-1}\tilde{D} \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0 \\ \tilde{M}^{-1} \end{bmatrix},$$

$$B_2 = B_1,$$

$$\tilde{M} = \text{diag}(\tilde{M}_1, \cdots, \tilde{M}_{N_G}),$$

$$\tilde{D} = \text{diag}(\tilde{D}_1, \cdots, \tilde{D}_{N_G}),$$

$$u = Kx,$$

$$K = \begin{bmatrix} K_\theta & K_\omega \end{bmatrix},$$

$$\omega = \dot{\theta}.$$  

The Laplacian or admittance matrix $L$ satisfies the following equations.

$$l_{ij} = -b_{ij}^{\text{Kron}},$$

$$l_{ii} = \sum_{k=1,k\neq i}^{N_G} b_{ik}^{\text{Kron}},$$

where $B^{\text{Kron}}$ is the susceptance matrix of the corresponding Kron reduced admittance matrix.
The power network utilized in our simulation is depicted in Figure 1, and its parameters, in per unit system, are presented in Table I.

We define the following performance metrics which quantify the deviation in $H_2$ and $H_\infty$ norms caused by the sparsification process. They also allow for comparison of the sparsification performance in the absence and presence of uncertainty on the system matrices.

$$R_2 = \frac{\|S - \hat{S}\|_{H_2}}{\|\hat{S}\|_{H_2}},$$
$$R_\infty = \frac{\|S - \hat{S}\|_{H_\infty}}{\|\hat{S}\|_{H_\infty}}.$$  

(19)

(20)

Now, we assume that the susceptance corresponding to the link between two randomly-chosen nodes $i_1$ and $i_2$ is affected by an uncertainty of the form

$$\rho = \rho_{rel}(b_{i_1i_2}^{\text{Kron}}),$$

where $\rho_{rel}$ is called the relative uncertainty and $b_{ij}^{\text{Kron}}$ is assumed to take non-zero values. Should we relax this assumption, the uncertainty and relative uncertainty will have to be defined in a different way, e.g.

$$\rho = \rho_{rel} \min \left\{ \sum_{k=1,k\neq i_1}^{N_G} b_{i_1k}^{\text{Kron}}, \sum_{k=1,k\neq i_2}^{N_G} b_{i_2k}^{\text{Kron}} \right\}.$$

To study the effect of adding uncertainty to the link between generators $i_1$ and $i_2$, the matrices $D$, $E_A$, and $E_{B_1}$ are chosen as follows.

$$D = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{M}^{-1} \end{bmatrix} (\epsilon_{i_1+N_G} - \epsilon_{i_2+N_G}),$$

$$E_A = \epsilon_{i_1+N_G}^{T} - \epsilon_{i_2+N_G}^{T},$$

$$E_{B_1} = 0.$$

| Bus | Generator | $\tilde{M}$ | $\tilde{D}$ | $\theta_0$ | $\omega_0$ |
|-----|-----------|-------------|-------------|------------|------------|
| 30  | $G_{10}$  | 4           | 5           | −0.0839    | 1          |
| 31  | $G_2$     | 3           | 4           | 0.0000     | 1          |
| 32  | $G_3$     | 2.5         | 4           | 0.0325     | 1          |
| 33  | $G_4$     | 4           | 6           | 0.0451     | 1          |
| 34  | $G_5$     | 2           | 3.5         | 0.0194     | 1          |
| 35  | $G_6$     | 3.5         | 3           | −0.0073    | 1          |
| 36  | $G_7$     | 3           | 7.5         | 0.1304     | 1          |
| 37  | $G_8$     | 2.5         | 4           | 0.0211     | 1          |
| 38  | $G_9$     | 2           | 6.5         | 0.1270     | 1          |
| 39  | $G_1$     | 6           | 5           | −0.2074    | 1          |

TABLE I: Power parameters used in our simulations.
Fig. 2: (a) Sparsity Pattern of $K$ for $\rho_{rel} = 0\%$; Blue and red bullets are used to depict diagonal and off-diagonal entries of $K$, respectively (b) Sparsity Pattern of $K$ for $\rho_{rel} = 30\%$ ($i_1, i_2 = (2, 3)$) (c) Sparsity Graph of $|K_\theta| + |K_\omega|$ for $\rho_{rel} = 0\%$; Blue solid lines, red dashed lines, and black self-loops are used to depict doubly-connected, singly-connected, and self-connected edges of $|K_\theta| + |K_\omega|$, respectively (d) Sparsity Graph of $|K_\theta| + |K_\omega|$ for $\rho_{rel} = 30\%$ ($i_1, i_2 = (2, 3)$).

Assuming $C = I$, $Q = I$, $R = 10I$, $\lambda_1 = 0.5$, $\lambda_2 = 0.1$, $\nu = 100$, $\xi = 10^{-6}$, and $\varepsilon^* = 10^{-2}$, we randomly choose two generators, $i_1 = 2$ and $i_2 = 3$, and consider the uncertainty cases $\rho_{rel} \in \{0\%, 30\\%\}$. The results of the static state feedback controller design using our method are presented in Table II. According to this table, an increase in uncertainty increases $R_2$ and $R_\infty$, and worsens the sparsification of the controller.

Figures 2a and 2b visualize the corresponding sparsity patterns for both cases $\rho_{rel} = 0\%$ and $\rho_{rel} = 30\%$, respectively, and Figures 2c and 2d visualize the corresponding sparsity graphs for both cases $\rho_{rel} = 0\%$ and $\rho_{rel} = 30\%$, respectively. It should be noted that entries $k_{22}$, $k_{23}$, $k_{2(12)}$, $k_{2(13)}$, $k_{3(12)}$, and $k_{3(13)}$ take non-zero

| $\rho_{rel}$ | $R_2$ | $R_\infty$ | $\|K\|_0/\|\hat{K}\|_0$ |
|-------------|-------|----------|----------------------|
| 0%          | 21.35%| 49.42%   | 4.5%                 |
| 30%         | 36.31%| 88.71%   | 7.5%                 |

Table II: Performance and cardinality quantities for the case $\rho_{rel} \in \{0\%, 30\%\}$. 

16
values after applying the 30% relative uncertainty. The interpretation is that, since uncertainty causes interference to
the link between two randomly-chosen generators, the construction of communication links between such generators
is vital.

Furthermore, additional plots are presented in Figure 6 to show the similarity of the frequency behavior of the
sparsely-controlled system to that of the LQR-controlled system. The upper left sub-figure, i.e., Figure 6a, depicts
the largest and smallest singular values of $S$ and $\hat{S}$ for the case of $\rho_{rel} = 0\%$. It can be seen that the smallest
singular values of the systems match for almost the whole frequency range and largest singular values achieve the
same values for higher frequencies. Similar plots for the case of uncertain system with $\rho_{rel} = 30\%$ are depicted in
Figure 6b. The plots depict that the deviation of the maximum singular value, caused by increasing the magnitude
of the uncertainties, is much larger compared to the deviation of the minimum singular value. Also, the plots of
Schatten 2-norm of the systems $S$ and $\hat{S}$, are depicted in lower sub-figures of 6 for both cases, i.e., $\rho_{rel} = 0\%$
and $\rho_{rel} = 30\%$. It is noteworthy that in neither of the cases, does the sparsification process seem to affect the
higher frequency content of the closed-loop systems. This is desirable, since the controller sparsification will not
be amplifying the harmonics in power grids, which are the main cause of power quality degradation.

In order to verify the relationship between the magnitude of the susceptance of each link, visualized in Figure 3,
and the density level of the corresponding entries in the controller design, we consider all cases with the $\rho_{rel} = 30\%$
uncertainty on one link at a time, which results in 45 cases. We then, compute $f(K_\theta)$ and $f(K_\omega)$, the sub-blocks
of the controller matrix $K$, in which the matrix-valued function $f(X) = [f(X)_{ij}]$ is defined as

$$f(X)_{ij} = \begin{cases} 
\|x_{ii}\|_0 + \|x_{ij}\|_0 + \|x_{ji}\|_0 + \|x_{jj}\|_0 & \text{if } i \neq j, \\
0 & \text{otherwise.}
\end{cases}$$

$f(K_\theta)$ and $f(K_\omega)$ are used to visualize the number of controller links, necessary to be added to the generators
connected with the uncertain link. Figures 4a and 4b show this visualization.

As depicted in Figs. 4a and 4b, in the case of links with higher susceptance, more communication links in
controller design need to be established. This can be interpreted as the effective uncertainty of each link being
Fig. 4: (a) Gray scale Pattern of \( f(K_\theta) + f(K_\omega) \) for \( \rho_{rel} = 30\% \) (b) Gray scale Pattern of \( f(|K_\theta| + |K_\omega|) \) for \( \rho_{rel} = 30\% \)

Fig. 5: (a) \( f(K_\theta) + f(K_\omega) \) vs \( \rho_{rel}\% \) (b) \( f(|K_\theta| + |K_\omega|) \) vs \( \rho_{rel}\% \).

proportional to the susceptance of that link. Therefore, an increase in susceptance of a link magnifies the uncertainty of that link, which results in establishment of more links in the designed controller to compensate for the fragility of the network on that link. This leads to similar patterns in Figures 3, 4a, and 4b.

We furthermore showcase the effect of increasing the relative uncertainty of the network links on the cardinality of their corresponding controller entries for two randomly chosen links, connecting generator 4 to generator 5 and generators 2 to 3. As seen in figures 5a and 5b, the increase of relative uncertainty, leads to construction of more communication links between two corresponding generators in the designed controller gain.

VIII. Discussion

We have explored a new optimization framework for the design of optimal sparse controllers for LTI systems under parametric uncertainties. The idea entails pruning the links of a centralized controller towards a sparse controller,
while heeding the performance deterioration caused by this sparsification process. The design procedure is built upon constructing an optimization problem which seeks a sparse structured controller capable of exhibiting similar time and frequency characteristics of the previously designed controller, in the sense of $H_2$ and $H_\infty$ norms. We then, propose a computationally tractable algorithm, utilizing a bi-linear rank penalizing technique, to sub-optimally solve a fixed-rank reformulation of the aforementioned optimization problem. It should be noted one of the two steps in the bi-linear optimization algorithm, namely $Y$-minimization step, has an analytical solution, which immensely enhances the run-time of our algorithm. Also, the outstanding performance of our proposed algorithm has been demonstrated through our extensive numerical simulations, run against various types of networks and systems. Although proving the global convergence of this algorithm still remains an open problem, we have shown that the sequence generated by the optimal cost at each iteration is monotonically decreasing and as a result, convergent.

One future research direction would encompass modifying our method to study the effect of the structure and magnitude of the uncertainties on the robustness of the closed-loop systems as well as the sparsity level of the controllers. An important application of this study would be the analysis of the robustness of networks, such as power grids, against possible attacks on the critical nodes.

REFERENCES

[1] B. Polyak, M. Khlebnikov, and P. Shcherbakov, “An LMI approach to structured sparse feedback design in linear control systems,” ser. Proceeding of the 2013 European Control Conference, July 2013, pp. 833–838.
[2] B. Bamieh, F. Paganini, and M. A. Dahleh, “Distributed control of spatially-invariant systems,” IEEE Transactions on Automatic Control, vol. 47, pp. 1091–1107, 2002.
[3] J. Lavaei, “Optimal decentralized control problem as a rank-constrained optimization,” Oct 2013, pp. 39–45.
[4] F. Lin, M. Fardad, and M. Jovanović, “Augmented Lagrangian approach to design of structured optimal state feedback gains,” IEEE Transactions on Automatic Control, vol. 56, no. 12, pp. 2923–2929, Dec 2011.
[5] 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, Feb 2006.
[6] Y.-S. Wang and N. Matni, “Localized LQG optimal control for large-scale systems,” July 2016.
[7] Y. S. Wang, N. Matni, and J. C. Doyle, “Localized LQR control with actuator regularization,” July 2016.
[8] B. Bamieh and P. Voulgaris, “A convex characterization of distributed control problems in spatially invariant systems with communication constraints,” Systems and Control Letters, vol. 54, pp. 575–583, 2005.
[9] F. Lin, M. Fardad, and M. R. Jovanović, “Design of optimal sparse feedback gains via the Alternating Direction Method of Multipliers,” IEEE Transactions on Automatic Control, vol. 56, no. 9, pp. 2426–2431, 2013.
[10] R. Arastoo, N. Motee, and M. V. Kothare, “Optimal Sparse Output Feedback Control Design: a Rank Constrained Optimization Approach,” ArXiv e-prints, Dec. 2014.
[11] M. Bahavarnia, “Sparse Linear-Quadratic Feedback Design Using Affine Approximation,” ArXiv e-prints, Jul. 2015.
[12] F. Lin, M. Fardad, and M. R. Jovanović, “Sparse feedback synthesis via the Alternating Direction Method of Multipliers,” 2012, pp. 4765–4770.
[13] M. Wytock and J. Zico Kolter, “A Fast Algorithm for Sparse Controller Design,” ArXiv e-prints, Dec. 2013.
[14] R. Arastoo, M. Bahavarnia, M. V. Kothare, and N. Motee, “Closed-loop feedback sparsification under parametric uncertainties,” Dec 2016, pp. 123–128.
[15] R. Arastoo, M. Bahavarnia, M. Kothare, and N. Motee, “Output feedback controller sparsification via $H_2$ approximation,” IFAC-PapersOnLine, vol. 48, no. 22, pp. 112 – 117, 2015, 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems Philadelphia, 10-11 September 2015.
[16] S. Fattahi, G. Fazelnia, and J. Lavaei, “Transformation of optimal centralized controllers into near-global static distributed controllers,” Dec 2015, pp. 4915–4922.
Fig. 6: Frequency characteristics of the closed-loop systems controlled by the LQR (blue), the sparse controller (red) for the case $\rho_{\text{rel}} = 0\%$, and the sparse controller (green) for the case $\rho_{\text{rel}} = 30\%$. (a) and (b) depict maximum and minimum singular values for the cases of $\rho_{\text{rel}} = 0\%$ and $\rho_{\text{rel}} = 30\%$, respectively. (c) and (d) exhibit the Schatten 2-norm of the closed-loop system ((c) case $\rho_{\text{rel}} = 0\%$ (d) case $\rho_{\text{rel}} = 30\%$)

[17] S. Fattahi and J. Lavaei, “Theoretical guarantees for the design of near globally optimal static distributed controllers,” Sept 2016, pp. 582–589.

[18] L. Xie, M. Fu, and C. E. de Souza, “$H_\infty$ control and quadratic stabilization of systems with parameter uncertainty via output feedback,” *Automatic Control, IEEE Transactions on*, vol. 37, no. 8, pp. 1253–1256, 1992.

[19] L. Xie and C. E. de Souza, “Robust $H_\infty$ control for linear systems with norm-bounded time-varying uncertainty,” *Automatic Control, IEEE Transactions on*, vol. 37, no. 8, pp. 1188–1191, 1992.

[20] I. R. Petersen, “A stabilization algorithm for a class of uncertain linear systems,” *Systems & Control Letters*, vol. 8, no. 4, pp. 351–357, Mar. 1987.

[21] ——, “Stabilization of an uncertain linear system in which uncertain parameters enter into the input matrix,” *SIAM J. Control Optimization*, vol. 26, no. 6, pp. 1257–1264, Oct. 1988.

[22] P. Khargonekar, I. Petersen, and K. Zhou, “Robust stabilization of uncertain linear systems: quadratic stabilizability and $h$ infin; control theory,” *Automatic Control, IEEE Transactions on*, vol. 35, no. 3, pp. 356–361, Mar 1990.

[23] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre, “Fixed-rank matrix factorizations and reimannian low-rank optimization,” *CoRR*, vol. abs/1209.0430, 2012.

[24] H. Yu and V. K. N. Lau, “Rank-constrained Schur-convex optimization with multiple trace/log-det constraints,” *IEEE Transacation on Signal Processing*, vol. 59, pp. 304–314, 2011.
[25] S. Shalev-Shwartz, A. Gonen, and O. Shamir, “Large-scale convex minimization with a low-rank constraint,” CoRR, vol. abs/1106.1622, 2011.

[26] Y. Gao and D. Sun, “A majorized penalty approach for calibrating rank constrained correlation matrix problems,” Preprint available at http://www.math.nus.edu.sg/~matsundf/MajorPen.pdf, 2010.

[27] G. Zhou, W. Huang, K. A. Gallivan, P. Van Dooren, and P. A. Absil, “Rank-constrained optimization: A riemannian manifold approach,” U.C.Louvain, Tech. Rep. UCL-INMA-2015.02, January 2015.

[28] G. E. Dullerud and F. G. Paganini, A Course in Robust Control Theory: A Convex Approach. Springer Verlag, 2000, vol. 36.

[29] V. V. Kulkarni, R. Arastoo, A. Bhat, K. Subramanian, M. V. Kothare, and M. C. Riedel, “Gene regulatory network modeling using literature curated and high throughput data,” Systems and Synthetic Biology, vol. 6, no. 3-4, pp. 69–77, December 2012.

[30] R. Arastoo, Y. GhaedSharaf, M. V. Kothare, and N. Motee, “Optimal state feedback controllers with strict row sparsity constraints,” July 2016.

[31] J. Dattoro, Convex Optimization & Euclidean Distance Geometry. USA: Meboo Publishing, 2005.

[32] E. Candes, M. Wakin, and S. Boyd, “Enhancing sparsity by reweighted \(\ell_1\) minimization,” Journal of Fourier Analysis and Applications, vol. 14, pp. 877–905, July 2008.

[33] A. Kalbat, R. Madani, G. Fazelnia, and J. Lavaei, “Efficient convex relaxation for stochastic optimal distributed control problem.” IEEE, 2014, pp. 589–596.