Optimal Sparse Output Feedback Controller Design: 
A Rank Constrained Optimization Approach

Reza Arastoo† Nader Motee‡ Mayuresh V. Kothare‡

Abstract

We consider the problem of optimal sparse output feedback controller synthesis for continuous linear time invariant systems when the feedback gain is static and subject to specified structural constraints. Introducing an additional term penalizing the number of non-zero entries of the feedback gain into the optimization cost function, we show that this inherently non-convex problem can be equivalently cast as a rank constrained optimization, hence, it is an NP-hard problem. We show that our problem reformulation allows us to incorporate additional implementation constraints, such as norm bounds on the control inputs or system output, by assimilating them into the rank constraint. We use a version of the Alternating Direction Method of Multipliers (ADMM) as an efficient method to sub-optimally solve the equivalent rank constrained problem. As a special case, we study the problem of designing the sparsest stabilizing output feedback controller, and show that it is, in fact, a structured matrix recovery problem where the matrix of interest is simultaneously sparse and low rank. Furthermore, we show that this matrix recovery problem can be equivalently cast in the form of a canonical and well-studied rank minimization problem. We finally illustrate performance of our proposed methodology using numerical examples.

I. INTRODUCTION

The problem of optimal linear quadratic controller design has been extensively studied for several decades. In conventional control, it is usually assumed that all measurements are accessible to a centralized controller, while in large scale interconnected systems this assumption is not practical, since it is often desirable that subsystems only communicate with a few neighboring components due to the high cost and, sometimes, infeasibility of communication links. Therefore, the need to exploit a particular controller structure, obtained based on the layout of the system network, seems undeniable. Furthermore, the traditional controller synthesis methods, which are closely related to solving the Algebraic Riccati Equation no longer work when additional constraints are imposed on the structure of the controller.

In general, the problem of designing constant gain feedback controllers subject to additional constraints is NP-hard [1]. In recent years, numerous attempts have been made to provide distributed controller synthesis approaches for different classes of systems [2], [3], [4], [5], [6]. Bamieh et al, in [7], [8], investigated the distributed control of

† R. Arastoo and N. Motee are with Department of Mechanical Engineering and Mechanics, Packard Lab., Lehigh University, Bethlehem, PA. Email Addresses: {reza.arastoo,nader.motee}@lehigh.edu
‡ M. V. Kothare is with the Department of Chemical and Biomolecular Engineering, Iaccoca Hall, Lehigh University, Bethlehem, PA. Email Address: mayuresh.kothare@lehigh.edu
spatially invariant systems, then the work in [9] has proved that the solution of Riccati and Lyapunov equations for systems consisting of Spatially Decaying (SD) operators has SD property, which lends credibility to the search for controllers that have access only to local measurements. The design of optimal state feedback gain in the presence of an *a priori* specified structure, usually in the form of sparsity patterns, is considered in [5]. In their recent papers, Lavaei *et al.* [10], [11], [12] cast the problem of optimal decentralized control for discrete time systems as a rank constrained optimization problem, developed results on the possible rank of the resulting feasible set, and introduced several rank-reducing heuristics as well. Frequency domain approaches to design optimal decentralized controllers are also presented in [13], [14].

In the design of linear feedback controllers for interconnected systems, a common desired structure is that the controller matrices are sparse, which could correspond to a simpler controller topology and fewer sensors/actuators. However, fewer measurement/communication links leads to performance deterioration and sometimes even instability of the overall system. Therefore, there exists a trade off between the stability and performance of the system and minimizing the number of non-zero entries of the feedback gain matrices. The problem of minimizing the number of nonzero elements of a vector/matrix subject to a set of constraints in inherently NP-hard and arises in many fields, such as Compressive Sensing (CS) where the inherent sparseness of signals is exploited in determining them from relatively few measurements [15]. Since the advent of Compressive Sensing, considerable work has been done on the design of compressive measurement matrices based on different criteria such as sparse signal support detection and estimation [16], [17], sparse signal detection and classification [18], [19], etc.

To alleviate the issues caused by the combinatorial nature of cardinality functions, several convex/non-convex functions have been proposed as surrogates for the cardinality functions in optimization problems. For example, in cases where the optimization constraint is affine, $\ell_1$-norm, as a convex relaxation of $\ell_0$-norm, has proved to work reliably under certain conditions, namely Restricted Isometry Property (RIP) [20], [21], [22]. Thus, $\ell_1$-norm and its weighted versions have been extensively used in signal processing and control applications [23], [24], [25]. Non-convex relaxations of the cardinality function, such as $\ell_q$-quasi-norm ($0 < q < 1$), have also received considerable attention recently [26], [27]. In [28], [29], [30], it is shown that, for a large class of SD systems, the quadratically-optimal feedback controllers inherit spatial decay property from the dynamics of the underlying system. Moreover, the authors have proposed a method, based on new notions of $q$-Banach algebras, by which sparsity and spatial localization features of the same class can be studied when $q$ is chosen sufficiently small.

In the present paper, we consider the problem of optimal sparse feedback controller synthesis for linear time invariant system, in which convex constraints are imposed on the structure of the controller feedback gain. The main contribution of our paper is to propose a novel approach which allows us to equivalently represent the intrinsically nonlinear constraints, such as closed loop stability condition and enforcement of controller structure, with a *single* rank constraint in an otherwise convex optimization program. Having all non-linearities encapsulated in only one rank constraint allows us to employ one of several existing algorithms to efficiently solve the resulting problem.

Our results are different from those reported in [23] in that we present an alternative formulation which not only solves the regular sparse controller design problem, but also enables us to solve the output feedback control problem,
while integrating many different types of nonlinear system constraints, such as constraints on the controller matrix and its norms, into the existing rank constraint. Furthermore, the rank constraint emerging in our approach originates from the positive definiteness of the Lyapunov matrix and the properties of fixed rank matrices. In contrast, the rank one constraint in [12] results from utilizing the auxiliary variable introduced by self multiplying the vector formed by augmenting the states, inputs, and outputs. Thus, compared to [12], the gap between the matrix dimension and its rank is not very large in our approach, and hence the convergence is faster.

We start by augmenting the $\ell_0$-norm of the feedback gain matrix to the quadratic cost function of our optimization problem. This additional term penalizes the extra communication links in the feedback pathway. We then reformulate it into an equivalent optimization problem where the non-convex constraints are lumped into a rank constraint. Based on the notions of holdable ellipsoid, we propose a reformulation of the problem to incorporate norm bounds on the control inputs and outputs of the system, which usually appear in controller implementations. Employing a convex relaxation of the added cardinality term, based on the $\ell_1$-norm, we argue that Alternating Direction Method of Multipliers (ADMM) is well-suited to solve our problem, since our search is to obtain a solution with an a priori known rank. ADMM iteratively solves the rank-unconstrained problem and projects the solution into the space of the matrices with the desired rank until the convergence criteria are met. We further investigate the special case of designing the sparsest stabilizing controller, and show that this problem can be rewritten as a rank minimization problem. Rank minimization problems have received considerable attention in recent years [31], [32], [33]. In [34], it is shown that if a certain Restricted Isometry Property holds for the linear transformation defining the constraints, the minimum rank solution can be recovered by solving the minimization of the nuclear norm over the feasible space. Therefore, the nuclear norm may be used as a proxy for the rank minimization in our problem.

The remainder of this paper is organized as follows. In Section II, the general optimal sparse output feedback control problem setup is defined. Section III we reformulate the optimal sparse output feedback control problem as a rank constrained problem, and develop several results based on the proposed reformulation. In Section IV we study the convex relaxation of this problem, and discuss the application of ADMM in solving the problem. The special case where the sparsity penalizing factor dominates the quadratic terms in the cost function is described in Section V. Numerical examples illustrating the proposed methods are provided in VI. Finally, Section VII concludes the paper.

**Notations:** Throughout the paper, the following notation is adopted. 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 $n$ by $n$ identity matrix is denoted $I_n$. Operators $\text{Tr}(\cdot)$ and $\text{rank}(\cdot)$ denote the trace and rank of the operands, which are matrices. A matrix is said to be Hurwitz if all its eigenvalues lie in the open left half of the complex plane. $\|\cdot\|_0$ represents the cardinality of a vector/matrix, while $\|\cdot\|_1$ and $\|\cdot\|_F$ denote $\ell_1$ and Frobenius norm operators. Also, the norm $\|\cdot\|_{L_\infty}^{\text{q}}(\mathbb{R}^n)$ is defined by

$$\|x\|_{L_\infty}^{\text{q}}(\mathbb{R}^n) \triangleq \sup_{t \geq 0} \|x(t)\|_q$$

The transpose and vectorization operators are denoted by $(\cdot)^T$ and $\text{vec}(\cdot)$, respectively.
A real symmetric matrix is said to be positive definite (semi-definite) if all its eigenvalues are positive (non-negative). $S_+^n$ ($S_+^n$) denotes the space of positive definite (positive semi-definite) real symmetric matrices, and the notation $X \succeq Y$ ($X \succ Y$) means $X - Y \in S_+^n$ ($X - Y \in S_+^n$).

II. Problem Formulation

Let a linear time invariant system be given by its state space realization

\[ \dot{x}(t) = Ax(t) + Bu(t) \]
\[ y(t) = Cx(t) \]

where $x(t) \in \mathbb{R}^n$ is the state vector, $y(t) \in \mathbb{R}^p$ is the output of the system, $u(t) \in \mathbb{R}^m$ is the control input, and matrices $A$, $B$ and $C$ have appropriate dimensions. We consider designing a constant gain output feedback stabilizing controller

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

with the minimum number of non-zero entries that minimizes a quadratic objective function. We assume the set $\mathcal{K}$, denoting the set of all acceptable feedback gains with a priori specified structure, is a convex set. Although it is a restricting assumption, convex controller constraints have broad real-world applications. One example of such convex structural constraints with practical applications is the set of admissible feedback matrices represented by a directed graph, i.e. the pair $(\mathcal{V}, \mathcal{E})$ of vertices and edges respectively, as shown in equation (3).

\[ \mathcal{K} = \{ K | K_{ij} = 0 \text{ if } (v_i, v_j) \notin \mathcal{E} \} \]

In addition, we consider an upper bound on the norm of the control input $u(t)$ and the closed loop system output $y(t)$. The search for such a controller can be formulated as an optimization problem, in which the sparsity of the feedback gain is incorporated by adding the $\ell_0$-norm of the gain matrix to the objective function. The $\ell_0$-norm denotes the cardinality of the feedback gain, hence, it penalizes the number of non-zero entries of the matrix. Therefore, we have the following optimization problem

\[
\min_{x,K} J = \int_0^\infty [x(t)^T Q x(t) + u(t)^T R u(t)] dt + \lambda \|K\|_0 \tag{P1}
\]

s.t. \[ \dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0 \]
\[ u(t) = KCx(t), \quad K \in \mathcal{K} \]
\[ \|u\|_{L^q_\infty(\mathbb{R}^m)} \leq u_{\text{max}}, \quad \|y\|_{L^q_\infty(\mathbb{R}^p)} \leq y_{\text{max}}, \]

where $Q \in \mathbb{R}^{n \times n}_+$ and $R \in \mathbb{R}^{m \times m}_+$ are performance weight matrices, $x_0$ is the initial state, and $\lambda > 0$ is the regularization parameter. Also, the value of $q$ in the norm $\|\cdot\|_{L^q_\infty(\mathbb{R}^n)}$ can be either infinity or two. It is possible to rewrite our
The feedback gain matrix $K$ derived from solving the above optimization problem depends on the value of the initial state $x_0$. To avoid resolving the minimization problem for every value of $x_0$, we design a state feedback controller which minimizes the expected value of the cost function assuming that the entries of $x_0$ are independent Gaussian random variables with zero mean and unit variance, i.e. $x_0 \in \mathcal{N}(0, I_n)$. Using Lyapunov stability theorem, it can be easily checked that the global asymptotic stability of the closed loop system is guaranteed if and only if the matrix $X_{11}$ is positive definite, thus we can rewrite the optimization problem as follows

$$
\begin{align*}
\min_{X_{11}, X_{12}, X_{22}, K} & \quad \text{Tr}[QX_{11}] + \text{Tr}[RKCX_{11} CT K^T] + \lambda\|K\|_0 \\
\text{s.t.} & \quad (A + BKC)X_{11} + X_{11}^T (A + BKC)^T + x_0x_0^T = 0, \\
& \quad (A + BKC) \text{ Hurwitz,} \\
& \quad K \in \mathcal{K}, \\
& \quad \|KCx\|_{L_\infty(\mathbb{R}^m)} \leq u_{\text{max}}, \quad \|Cx\|_{L_\infty(\mathbb{R}^p)} \leq y_{\text{max}}.
\end{align*}
$$

where $X_{11} \in \mathbb{R}^{n \times n}$, $X_{12} \in \mathbb{R}^{n \times m}$, and $X_{22} \in \mathbb{R}^{m \times m}$. In optimization problem (5), the constraints (5b-5d) are convex, nevertheless, the constraints (5e) are nonlinear and the control input/output constraints (5f) are in time domain, hence, the problem is non-convex.

### III. Rank Constrained Formulation

In traditional LQR problems, with no input/output constraints, the nonlinear constraints can be replaced by a linear matrix inequality to form an equivalent convex problem. However, the addition of the sparsity penalizing term to the cost function, the existence of structural constraints on the feedback gain matrix, and incorporation of input/output bounds differentiate our problem from the conventional LQR problem, making the conventional approach inapplicable. Here, we propose a controller synthesis approach based on the idea that the non-convex constraints can be replaced by a rank constraint. For legibility purposes, we first develop the equivalent formulation for the case with no constraint imposed on the control inputs/outputs, then, we incorporate the bounds on the input/output of the closed loop system. Before proceeding, lets state the following lemma.
Lemma III.1. 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 $W = Y U Y^T$ and $V^T = Y U$, where

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

Proof: If $\text{rank}(M) = \text{rank}(U) = n$, we have:

$$\text{rank}(M) = \text{rank} \begin{bmatrix}
I_n & U^{-1} V \\
V^T & W \\
I_n & Y^T
\end{bmatrix}$$

subtracting first block-row from the last one, we obtain

$$\text{rank}(M) = \text{rank} \begin{bmatrix}
I_n & U^{-1} V \\
V^T & W \\
0_n & Y^T - U^{-1} V
\end{bmatrix}$$

Clearly, the last block row of the matrix must be equal to zero to avoid adding to the row rank of matrix $M$, thus, $Y^T = U^{-1} V$ and as a result

$$\text{rank}(M) = \text{rank} \begin{bmatrix}
I_n & Y^T \\
Y U & W
\end{bmatrix}$$

Post-multiplying the first block-column by $Y^T$, then, subtracting the result from the second block-column gives us

$$\text{rank}(M) = \text{rank} \begin{bmatrix}
I_n & 0_{n \times m} \\
Y U & W - Y U Y^T
\end{bmatrix} = n$$

The rank equality is true if and only if $W = Y U Y^T$. This completes the proof of the lemma.

A. Rank Constraint Formulation with no Input/Output Constraint

Assuming that no upper bound is defined for the input/output of the controlled system, the next proposition states that the nonlinear Semidefinite Program can be cast as an optimization problem, where all constraints are convex except one, which is a rank constraint.
Proposition III.2. The optimization program (5a-5e) is equivalent to the following rank constrained problem

\[ \min_{X_{11}, X_{12}, X_{22}, K} \text{Tr}[QX_{11}] + \text{Tr}[RX_{22}] + \lambda \|K\|_0 \] (P2)

subject to

- \( AX_{11} + X_{11}^T A^T + BX_{12}^T + X_{12} B^T + I_n = 0 \)
- \( X_{11} > 0 \)
- \( K \in \mathcal{K} \)
- \( \text{rank} \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \\ I_n & (KC)^T \end{bmatrix} = n \)

**Proof:** Using lemma III.1, it can be seen that the constraints \( X_{22} = (KC)X_{11}(KC)^T \) and \( X_{12}^T = (KC)X_{11} \) are equivalent to the rank of matrix \( \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \\ I_n & (KC)^T \end{bmatrix} \) being equal to \( n \), since \( X_{11} \) is constrained to be positive definite.

The rank constraint in (P3) is imposed on a tall matrix. In general, dealing with symmetric square matrices is computationally more convenient. Also, we plan to use a positive semidefinite relaxation of the rank constraint later in the paper, thus, it is important to find a way to associate the rank of the tall matrix to that of a symmetric matrix. The next lemma helps resolving this issue.

**Lemma III.3.** Assuming \( X_{11} > 0 \), the constraint

\[ \text{rank} \begin{bmatrix} X_{11} & X_{12} & I_n \\ X_{12}^T & X_{22} & (KC) \\ I_n & (KC)^T & Z \end{bmatrix} = n \]

is equivalent to

\[ \begin{cases} \text{rank} \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \\ I_n & (KC)^T \end{bmatrix} = n \\ Z = X_{11}^{-1} \end{cases} \]

**Proof:** Applying Lemma III.1 to \( \text{rank}(X) = n \) immediately yields the desired result.

Therefore, augmenting the matrix \( \begin{bmatrix} I_n & KC & Z \end{bmatrix}^T \) to the original rank constrained matrix only adds some redundant constraints along with a free matrix variable \( Z \). It should be noted that although we are increasing the number of variables by introducing the new \( n \)-by-\( n \) variable \( Z \), having a symmetric rank constrained matrix has proved to be helpful when relaxing the non-convex problem into a convex problem. We can rewrite the problem in the following
equivalent form, where the rank constraint is imposed on a symmetric square matrix.

\[
\begin{align*}
\min_X & \quad \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|K\|_0 \\
\text{s.t.} & \quad AX_{11} + X_{11}^TA^T + BX_{12}^T + X_{12}B^T + I_n = 0, \\
& \quad X_{11} \succ 0, \\
& \quad K \in \mathcal{K}, \\
& \quad \text{rank}(X) = n,
\end{align*}
\]

where

\[
X = \begin{bmatrix}
X_{11} & X_{12} & I_n \\
X_{12}^T & X_{22} & (KC) \\
I_n & (KC)^T & Z
\end{bmatrix}.
\]

**Remark III.4.** The optimal value of \(Z\) in problem (P3) is the inverse of the optimal \(X_{11}\), i.e. \(Z^* = X_{11}^{-1}\).

The above remark implies that the matrix \(Z\) can be replaced by \(X_{11}^{-1}\) in the optimization problem (P3). Since multiplying the rows and columns of a matrix by a non-zero scalar does not affect its rank, we can state the following proposition.

**Proposition III.5.** Assuming the set of all admissible controller gains \(\mathcal{K}\) is invariant under positive scaling. The following optimization solves the optimal sparse state feedback control problem, i.e. (P3) with \(C = I_n\).

\[
\begin{align*}
\min_X & \quad \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|\tilde{K}\|_0 \\
\text{s.t.} & \quad AX_{11} + X_{11}^TA^T + BX_{12}^T + X_{12}B^T + I_n = 0, \\
& \quad X_{11} \succ 0, \\
& \quad \tilde{K} \in \mathcal{K}, \\
& \quad \Gamma = \text{diag}(\alpha_1, \cdots, \alpha_n) \succ 0, \\
& \quad X \succeq 0,
\end{align*}
\]

where

\[
X = \begin{bmatrix}
X_{11} & X_{12} & \Gamma \\
X_{12}^T & X_{22} & \tilde{K} \\
\Gamma & \tilde{K}^T & \Gamma X_{11}^{-1} \Gamma
\end{bmatrix}.
\]

Furthermore, the optimal \(\tilde{K}^*\) satisfies \(\tilde{K}^* = K^* \Gamma^*\), where \(K^*\) represents the optimal controller obtained from (P3).

**Proof:** Scaling the last block-row and column of the matrix \(X\) in (P3), assuming the scaler is not zero, does
not affect the rank constraint, thus we can rewrite it as

\[
\begin{bmatrix}
X_{11} & X_{12} & \Gamma \\
X_{12}^T & X_{22} & K \Gamma \\
\Gamma & \Gamma K^T & \Gamma X_{11}^{-1} \Gamma
\end{bmatrix}
= n
\]

The \(\ell_0\)-norm is invariant under positive scaling, hence \(\|K\Gamma\|_0 = \|K\|_0\). Also, Since the set \(K\) is assumed to be invariant under positive scaling, the constraint \(K \in K\) is the same as the matrix \(K \Gamma\) belonging to the set of admissible controller structures. Therefore, it is possible to rewrite the optimization problem in terms of the new variable, defined as \(\tilde{K} = K \Gamma\). In optimization problem (6), due to the positive definiteness of \(X_{11}\), the constraint \(X \succeq 0\) is equivalent to positive definiteness of its Schur complement, that is

\[
\begin{pmatrix}
X_{22} & \tilde{K} \\
\tilde{K}^T & \Gamma X_{11}^{-1} \Gamma
\end{pmatrix} - \begin{pmatrix}
X_{12}^T \\
\Gamma
\end{pmatrix} X_{11}^{-1} \begin{pmatrix}
X_{12} & \Gamma
\end{pmatrix} \succeq 0
\]

which holds if and only if \(\tilde{K} = X_{12}^T X_{11}^{-1} \Gamma\), i.e. \(K = X_{12}^T X_{11}^{-1}\), and \(X_{22} = X_{12}^T X_{11}^{-1} X_{12} + M\), where \(M \succeq 0\). Therefore, the feasible set of (P3) is a subset of the feasible set of (6). To conclude our proof, it suffices to show that the optimal value of \(M\) in the optimization problem (6) is zero.

Let’s assume \(X^*\) is the optimal solution to (6), where \(M^*\) is not zero. The optimal cost corresponding to this optimum, namely \(J^*\), becomes

\[
J^* = \text{Tr}[R(X_{12}^T X_{11}^{-1} X_{12} + M^*)] + \text{Tr}[Q X_{11}^*] + \lambda \|K^*\|_0
\]

\[
= \text{Tr}[R(X_{12}^T X_{11}^{-1} X_{12}^*)] + \text{Tr}[Q X_{11}^*] + \lambda \|K^*\|_0 + \text{Tr}[QM^*]
\]

Since \(\text{Tr}[QM^*] \geq 0\), setting \(M^* = 0\), along with the same values of \(X_{11}\), \(X_{12}\), and \(K^*\), also belonging to the feasible set, generates a lower cost. This contradicts the optimality of \(X^*\), hence the optimal value of \(M\) must be zero.

The stated proposition may not provide us with a formulation that can be solved in a timely manner, however, it helps us find an upper bound for the optimal cost of the optimization problem (P3), for \(C = I_n\), through the next theorem.

**Theorem III.6.** Assuming the set \(K\) is invariant under positive scaling, the optimal cost of the following optimization
problem provides an upper bound for the solution of the rank constrained problem \( P3 \) when \( C = I_n \).

\[
\begin{align*}
\min_X & \quad \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|\tilde{K}\|_0 \\
\text{s.t.} & \quad AX_{11} + X_{11}^T A + BX_{12} + X_{12}B^T + I_n = 0, \\
& \quad X_{11} \succ 0, \quad \tilde{K} \in \mathcal{K}, \\
& \quad \Gamma = \text{diag}(\alpha_1, \cdots, \alpha_n) \succ 0, \\
& \quad X \succeq 0,
\end{align*}
\]

where

\[
X = \begin{bmatrix} X_{11} & X_{12} & \Gamma \\ X_{12}^T & X_{22} & \tilde{K} \\ \Gamma & \tilde{K}^T & 2\Gamma - X_{11} \end{bmatrix}.
\]

**Proof:** For the positive definite matrix \( X_{11} \) and the positive scaler \( \Gamma \), we have \( \Gamma^{-\frac{1}{2}}X_{11}^{-\frac{1}{2}}\Gamma^{-\frac{1}{2}}X_{11}^{-\frac{1}{2}} \preceq 2I \).

Thus

\[
\Gamma X_{11}^{-\frac{1}{2}} \gamma \preceq 2\Gamma - X_{11}
\]

Therefore, the feasible set of (7) is a subset of the feasible set of the optimization problem (6). The rest of the proof is straightforward.

**B. Rank Constraint Formulation in Presence of Input/Output Constraints**

Next, we present how an upper bound on the norm of the control input/output can be incorporated into our rank constrained formulation. It known that for the positive scaler \( \gamma \) satisfying \( x_0^T X_{11}^{-1} x_0 \leq \gamma^{-1} \), where \( x_0 \) is the initial state of the system and \( X_{11} \) is the solution to the Lyapunov stability condition, the set

\[
\mathcal{M} = \{ x \in \mathbb{R}^n \mid x^T X_{11}^{-1} x \leq \gamma^{-1} \}
\]

is an invariant set for the closed loop system. Employing the concept of invariant sets for linear systems, we can develop the rank constraint formulation of control system with bounded input norms. The details for two choices of norms utilize to bound the control input in given in the sequel.

- **System Norm:** Based on the lines in [36, p. 103], we have

\[
\|u\|_{L^\infty(\mathbb{R}^m)} = \sup_{t \geq 0} \|u(t)\|_2 = \sup_{t \geq 0} \|KCx(t)\|_2 \\
\leq \sup_{x \in \mathcal{M}} \|KCx\|_2 \\
= \sup_{x \in \mathcal{M}} \|KCX_{11}^{1/2}X_{11}^{-1/2}x\|_2 \\
= \sqrt{\lambda_{\text{max}}(X_{11}^{1/2}(KC)^T(KC)X_{11}^{1/2})} \gamma^{-1}
\]


Thus, the input constraint $\|u\|_{L_\infty^2(\mathbb{R}^m)} \leq u_{\text{max}}$ holds for all $t \geq 0$ if

$$
\begin{bmatrix}
\gamma X_{11}^{-1} & (KC) \\
(KC) & u_{\text{max}}^2 I_m
\end{bmatrix} \succeq 0,
$$

$$
x_0^T \gamma X_{11}^{-1} x_0 \leq 1.
$$

The existence of the term $\gamma X_{11}^{-1}$ in the above matrix inequality makes it nonlinear, however, we show that, with the help of rank constrained formulation, it is possible to convert it into a linear matrix inequality.

**Theorem III.7.** The optimization problem (P3) can be modified to conservatively incorporate an upper bound on the system norm of the control input, i.e. $\|u\|_{L_\infty^2(\mathbb{R}^m)} \leq u_{\text{max}}$, as follows

$$
\min_X \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|K\|_0
$$

**s.t.**

$$
AX_{11} + X_{12}^T A^T + BX_{12} + X_{13} B^T + I_n = 0,
$$

$$
X_{11} \succ 0,
$$

$$
K \in K,
$$

$$
\begin{bmatrix}
W & (KC)^T \\
(KC) & u_{\text{max}}^2 I_m
\end{bmatrix} \succeq 0,
$$

$$
x_0^T W x_0 \leq 1,
$$

$$
\text{rank}(X) = n,
$$

where

$$
X = \begin{bmatrix}
X_{11} & X_{12} & I_n \\
X_{12}^T & X_{22} & (KC) \\
I_n & (KC)^T & Z \\
\gamma I_n & Y & W
\end{bmatrix}.
$$

**Proof:** Using Lemma III.1, it can be verified that the rank constraint $\text{rank}(X) = n$, applied on the modified matrix $X$, is equivalent to introducing the variables $W = \gamma X_{11}^{-1}$. The rest of the proof is straightforward. ■

- **Infinity Norm:** If the constraint on the control input is in the form of $\|u(t)\|_{L_\infty^2(\mathbb{R}^m)} \leq u_{\text{max}}$, it can be represented using the following matrix inequalities [36] p. 104].

$$
\begin{bmatrix}
V & KC \\
(KC)^T & \gamma X_{11}^{-1}
\end{bmatrix} \succeq 0,
$$

$$
V_{ii} \leq u_{\text{max}}^2
$$

$$
x_0^T \gamma X_{11}^{-1} x_0 \leq 1.
$$

Therefore, this problem can also be posed as a rank constrained problem through the next theorem.
Theorem III.8. The optimization problem (P3) can be modified to conservatively incorporate an upper bound on the infinity norm of the control input, i.e. \(|u|_{\infty}(\mathbb{R}^m) \leq u_{\text{max}}\), as follows

\[
\begin{align*}
\min_X & \quad \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|K\|_0 \\
\text{s.t.} & \quad AX_{11} + X_{11}^T A^T + BX_{12} + X_{12} B^T + I_n = 0, \\
& \quad X_{11} \succ 0, \\
& \quad K \in \mathcal{K}, \\
& \quad \begin{bmatrix} V & KC \\ (KC)^T & W \end{bmatrix} \succeq 0, \\
& \quad V_{ii} \leq u_{\text{max}}^2, \\
& \quad x_0^T W x_0 \leq 1, \\
& \quad \text{rank}(X) = n,
\end{align*}
\]

where

\[
X = \begin{bmatrix}
X_{11} & X_{12} & I_n \\
X_{12}^T & X_{22} & (KC) \\
I_n & (KC)^T & Z \\
\gamma I_n & Y & W
\end{bmatrix},
\]

Remark III.9. Other norms such as element-wise bound on the control input or the norm bounds on the system outputs can also be assimilated into the rank constraint using similar techniques. The details are omitted with the purpose of improving the readability of the manuscript.

All of the optimization problems posed so far are NP-hard due to the existence of the \(\ell_0\)-norm in the cost function and the rank constraint. Therefore, no polynomial time algorithm capable of solving it in its general form, exists. In the next two sections, we propose a method to sub-optimally solve the problem, then, discuss an special case of the problem where only the sparsity of the controller is of importance.

IV. CONVEX RELAXTIONS OF THE OPTIMAL CONTROL PROBLEM

In this section, we study the general problem of designing a sparse optimal feedback controller. Although the results we present in the sequel are applicable to the optimization problem (P3) in its general form, to enhance the legibility of the paper, we choose to state them in the absence of the constraints on the control inputs and system outputs. Hence, we consider the problem (P3), which is a combinatorial problem, due to the existence of the \(\ell_0\)-norm, in fact a quasi-norm, in the cost and the rank constraint. The \(\ell_1\)-norm minimization problem is a well-known heuristic for cardinality minimization [15], [37], [22]. Although \(\ell_1\)-norm relaxation does not guarantee the exact optimal controller recovery, it reduces the complexity of the problem substantially. Substituting the cardinality...
penalizing term with the $\ell_1$-norm of the controller gain matrix, we obtain the following relaxed optimization problem

$$\min_X \text{Tr}[QX_{11}] + \text{Tr}[RX_{22}] + \lambda \|K\|_1$$  \hfill (C1)

s.t.  
\[
AX_{11} + X_{11}^T A^T + BX_{12}^T + X_{12} B^T + I_n = 0, \\
X_{11} \succ 0, \\
K \in \mathcal{K}, \\
\text{rank}(X) = n.
\]

The combinatorial nature of the our rank constrained problem, makes the search for the optimal point computationally intractable. Therefore, a systematic solution to general rank constrained problem has remained open \[38], \[39]. Nonetheless, attempts have been made to solve specific rank constrained problems, and algorithms proposed to locally solve such problems. Here, we propose to use a particular form of Alternating Direction Method of Multipliers (ADMM) to solve our rank constrained problem.

### A. ADMM for Solving the Relaxed Problem

ADMM was originally developed in 1970s \[40], \[41], and has been used for optimization purposes since. Boyd et al., in \[42], argued that this method can be efficiently applied to large-scale optimization problems. For non-convex problems, the convergence of ADMM is not guaranteed, also, it may not reach the global optimum when it converges, thus, the convergence point should be considered as a local optimum.

For the optimization problem (C1), one way to perform convex relaxation is replacing the rank constraint on matrix $X$ with a positive semi-definite constraint, i.e. $X \succeq 0$. Since $X_{11}$ is positive definite, using lemma III.1 it can be seen that the rank constraint in (C1) is equivalent to

\[
\begin{bmatrix}
X_{22} & (KC) \\
(KC)^T & Z
\end{bmatrix}
- \begin{bmatrix}
X_{12}^T \\
I_n
\end{bmatrix}
X_{11}^{-1}
\begin{bmatrix}
X_{12} & I_n
\end{bmatrix} = 0,
\]

which implies that the Schur complement of the matrix $X$ should be equal to zero, while $X \succeq 0$ is the same as positive semi-definiteness of its Schur complement. Therefore, the set defined by the PSD constraint is a super-set for the one defined by the rank constraint. Now, if we define the convex set

$$\mathcal{C} = \{X | AX_{11} + X_{11}^T A^T + BX_{12}^T + X_{12} B^T + I_n = 0, \quad X_{11} \succ 0, \quad K \in \mathcal{K}, \quad X \succeq 0\}$$

and $\mathcal{S}$ denotes the set of $(2n + m) \times (2n + m)$ symmetric matrices with rank equal to $n$, the minimization (C1) can be represented as

$$\min_x f(x)$$  \hfill (10)

s.t.  
\[ x \in \mathcal{C} \cap \mathcal{S} \]
where

\[ f(x) = \text{Tr}[RX_{22}] + \text{Tr}[QX_{11}] + \lambda \|K\|_1 \]

Considering the above formulation, the ADMM algorithm can be carried out by repeatedly performing the steps stated in the sequel till certain convergence criteria is satisfied [42, p. 74].

\[
x^{k+1} = \arg \min_{x \in \mathcal{C}} f(x) + \left(\frac{\rho}{2}\right)\|x - w^k + u^k\|_F^2
\]

\[
w^{k+1} = \Pi_S(x^{k+1} + u^k)
\]

\[
u^{k+1} = u^k + x^{k+1} - w^{k+1}
\]

The convexity of the cost function and the constraints makes (11) a convex problem, hence, it can be solved by various computationally efficient methods. The operator \( \Pi_S(.) \), in (12), denotes projection onto the set \( \mathcal{S} \). Although the projection on a non-convex set is generally not an easy task, it can be carried out exactly in the case of projecting on the set of matrices with pre-defined rank. In our case, the set \( \mathcal{S} \) is the set of matrices with rank \( n \), thus, \( \Pi_S(.) \) can be determined by carrying out Singular Value Decomposition (SVD) and keeping the top dyads, i.e.

\[
\Pi_S(X) \triangleq \sum_{i=1}^{n} \sigma_i u_i v_i^T
\]

where \( \sigma_i, i = 1, \cdots, n \) are the \( n \) largest singular values of matrix \( x \), and the vectors \( u_i \in \mathbb{R}^{(2n+m)} \) and \( v_i \in \mathbb{R}^{(2n+m)} \) are their corresponding left and right singular vectors. However, executing (12) using equation (14) may cause some undesirable effects on the convergence behaviour of our ADMM algorithm. Since eq. (14) modifies all sub-blocks of matrix \( X \), it may interfere with the sparsity promoting objective of our problem, furthermore, the output of (14) is likely to violate the pre-defined structure of the controller matrix \( K \).

The last step in the algorithm is a simple matrix manipulation to update the auxiliary variable \( u \), which is exploited in the next iteration. Initializing with the stabilizing LQR controller along with its corresponding Lyapunov matrix, a sub-optimal minimizer to the problem (C1) can be obtained by iterating the steps (11) until the convergence is achieved. The algorithm’s stopping criteria is \( \varepsilon < \varepsilon^* \), where \( \varepsilon \) update is performed using the following equation.

\[
\varepsilon^{k+1} \triangleq \max\{\|x^{k+1} - w^{k+1}\|_F, \|w^{k+1} - w^k\|_F\}
\]

The small enough entries of the generated controller gain can then be truncated to yield a sparse controller matrix, while considering the extent of its adverse effect on the stability and performance of the closed loop system. The step-by-step procedure is described in Algorithm 1.

Remark IV.1. For the problem of optimal sparse state feedback control design, i.e \( C = I_n \), if there exists no a priori defined controller structure or the constraint on the controller matrix is in the form of sparsity pattern, one way to perform the truncation is to solve the the minimization problem, assuming that all of the variables have
already converged to their optimal values except the controller matrix. Thus, we will have

$$\min_K \lambda \|K\| + (\rho/2)\|K - (K_{w*} - K_{u*})\|_F^2,$$

s.t. $K \in \mathcal{K}$.\hfill (16)

where $K_{w*}$ and $K_{u*}$ are the sub-blocks of $w^*$ and $u^*$, respectively, which correspond to the controller gain matrix, and $\|\cdot\|$ can be chosen as either $\ell_1$ or $\ell_0$-norm. Moreover, in such problems, the problem (16) has a unique solution that can be obtained analytically as follows [23], [42]. For example, if the norm used in (16) is $\ell_0$-norm, the optimal values of the elements, not constrained to zero, can be obtained through the following element-wise truncation operator

$$K^*(i, j) = \begin{cases} K_{w*}(i, j) - K_{u*}(i, j), & |K_{w*}(i, j) - K_{u*}(i, j)| > \sqrt{2\lambda/\rho} \\ 0, & \text{otherwise.} \end{cases}$$

Remark IV.2. A weighted $\ell_1$-norm can also be used as a proxy for the $\ell_0$-norm. The re-weighted $\ell_1$ minimization method, introduced in [22], can be used to enhance the sparse controller recovery, where the optimization is run iteratively with the weight, assigned to each controller entry, inversely proportional to the value of the corresponding matrix entry recovered from the previous iteration.

V. SPARSEST STABILIZING OUTPUT FEEDBACK CONTROLLER DESIGN

Next, we study the special case in which obtaining a stabilizing constant gain feedback controller with the sparsest feasible structure, i.e. considering the constraints, is desirable. To this end, we eliminate the terms which penalize the system performance from the cost, i.e. both $R$ and $Q$ are zero. One of the applications that can be addressed using this problem setup is the problem of stabilizing controller synthesis for networks/systems where establishing
communication links between nodes are so costly that the control effort and error cost are almost negligible. Having \( R = 0 \), it can be seen the variable \( X_{22} \) is irrelevant in this case, so its corresponding constraints can be removed from the optimization program. Therefore, we will have

\[
\min_{X_{11}, X_{12}, K, N} \| K \|_0 \\
\text{s.t. } AX_{11} + X_{11}^T A^T + BX_{12}^T + X_{12} B^T + N = 0, \\
X_{11} \succ 0, \; N \succ 0, \\
K \in \mathcal{K}, \\
\text{rank} \left[ \begin{array}{cc} X_{11} & X_{12} \\ I_n & (KC)^T \end{array} \right] = n,
\]

The following lemma helps us convert rank constrained cardinality minimization problem (P5) into an affine rank minimization problem.

**Lemma V.1.** Consider the following rank constrained cardinality minimization problem

\[
\min_Y \| W_1 Y W_2 \|_0 \\
\text{s.t. } \mathcal{L}_1 (Y) = \mu, \\
\mathcal{L}_2 (Y) \succeq 0, \\
\text{rank}(Y) = \text{rank}(Y_{11}) = n,
\]

where \( Y \) is partitioned as \( Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \in \mathbb{R}^{p \times q} \), \( W_1 \in \mathbb{R}^{a \times p} \) and \( W_2 \in \mathbb{R}^{q \times b} \) are weight matrices, \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) are two arbitrary maps, and \( Y_{11} \in \mathbb{R}^{n \times n} \) is a full rank square matrix \((n < \min\{p, q\})\). If the optimization problem (18) is feasible, it can be equivalently formulated as

\[
\min_Y \| W_1 Y W_2 \|_0 + \nu \text{rank}(Y) \\
\text{s.t. } \mathcal{L}_1 (Y) = \mu, \\
\mathcal{L}_2 (Y) \succeq 0, \\
\text{rank}(Y_{11}) = n,
\]

for any \( \nu > ab \).

**Proof:** Let \( Y^* \) be the optimum of (18), then \( \text{rank}(Y_{11}^*) = n \) and it satisfies both equality and inequality constraints. Therefore, it belongs to the feasible set of (19). Furthermore, for every point \( Y \) in the feasible set of
with the rank greater than \(n\), we have
\[
J - J^* = \|W_1YW_2\|_0 + \nu \text{rank}(Y)
\]
\[
- (\|W_1Y^*W_2\|_0 + \nu \text{rank}(Y^*))
\]
\[
= (\|W_1YW_2\|_0 - \|W_1Y^*W_2\|_0)
\]
\[
+ \nu (\text{rank}(Y) - \text{rank}(Y^*))
\]
\[
\geq -\|W_1Y^*W_2\|_0 + \nu (\text{rank}(Y) - \text{rank}(Y^*))
\]
Since \(W_1Y^*W_2 \in \mathbb{R}^{a \times b}\), it is safe to bound the cardinality as \(\|W_1Y^*W_2\|_0 \leq ab\). Using \(\text{rank}(Y) - \text{rank}(Y^*) \geq 1\), we can write
\[
J - J^* > -ab + \nu
\]
Hence, the cost for all \(Y\), with rank greater than \(n\), is higher than the cost of \(Y^*\), if \(\nu > ab\). This means the optimum of (19) should be of rank \(n\). Knowing that \(Y^*\) has the minimum cardinality among the matrices with rank equal to \(n\), we conclude that \(Y^*\) is also the optimum for (19).

Conversely, let \(\bar{Y}\) be the optimal point for (19). As it is shown in the first part of the proof, the cost generated by matrices, with the rank higher than \(n\) is greater than that of rank \(n\) matrices, for \(\nu > ab\). Thus, the rank of \(\bar{Y}\) must be \(n\), unless no point with the rank equal to \(n\) exists in the feasible set of (19). However, this implies that (18) is infeasible, which contradicts the lemma’s assumption. Therefore, \(\bar{Y}\) is the minimizer of the cardinality term of the cost function among all rank \(n\) matrices in the feasible set of (19), i.e. \(\bar{Y}\) is the minimizer of (18).

**Remark V.2.** In the optimization problem (18), if the cost which is to be minimized is the rank of the matrix \(W_1XW_2\), instead of its cardinality, lemma V.1 can still be applied to the problem for any \(\nu > \min\{a, b\}\).

Applying lemma V.1 to (P5), we can equivalently write it as
\[
\min_{X_{11}, X_{12}, K, N} \|K\|_0 + \nu \text{rank} \begin{bmatrix} X_{11} & X_{12} \\ I_n & (KC)^T \end{bmatrix}
\]
\[
\text{s.t. } AX_{11} + X_{11}^T A^T + BX_{12}^T + X_{12} B^T + N = 0,
\]
\[
\text{diag}(X_{11}, N) \succ 0,
\]
\[
K \in \mathcal{K},
\]
with \(\nu > mn\). Note that the matrix \(X_{11}\) is full rank due to its positive definiteness, therefore, all of the requirements of lemma V.1 are satisfied.

**Remark V.3.** The solution to equation (20) falls into the category of the problem of recovery of simultaneously structured models where the matrix of interest is both sparse and low-rank \([43], [44]\). Oymak et al., in their recent paper, have shown that minimizing a combination of the known norm penalties corresponding to each structure (for example, \(\ell_1\)-norm for sparsity and nuclear norm for matrix rank) will not yield better results than an optimization.
exploiting only one of the structures. They have concluded that an entirely new convex relaxation is required in order to fully utilize both structures [43].

Without loss of generality, the following theorem is stated assuming \( m < n \).

**Theorem V.4.** The optimization problem \( (P_S) \), if feasible, is equivalent to

\[
\min_{X_{11}, X_{12}, C, K, N, \varepsilon} \text{rank} \left( \text{diag} \left( \text{vec}(K), \Psi_1, \ldots, \Psi_\nu, \Phi_1, \ldots, \Phi_\rho \right) \right) \\
s.t. \quad AX_{11} + X_{11}^T A^T + BX_{12} + X_{12} B^T + N = 0, \\
K \in \mathcal{K}, \\
\varepsilon > 0,
\]

where

\[
\Psi_i = \begin{bmatrix}
X_{11} & X_{12} \\
I_n & (KC)^T \\
\end{bmatrix} \begin{bmatrix}
0_{(2n \times (n-m))}
\end{bmatrix} \quad i = 1, \ldots, \nu
\]

\[
\Phi_i = \begin{bmatrix}
I_{2n} & D \\
D^T & \text{diag}(X_{11}, N) - \varepsilon I_{2n}
\end{bmatrix} \quad i = 1, \ldots, \rho
\]

and the parameters \( \nu \) and \( \rho \) are integers satisfying

\[
\rho > mn + \nu \max \{2n, (n + m)\}
\]

\[
\nu > mn
\]

**Proof:** For a function that maps matrices into \( q \times q \) symmetric matrices, positive semi-definiteness can be equivalently expressed as a rank constraint [34]

\[
f(X) \succeq 0 \iff \text{rank} \left[ \begin{bmatrix} I_q & U \\ U^T & f(X) \end{bmatrix} \right] \leq q
\]

for some \( U \in \mathbb{R}^q \). Since \( \text{diag}(X_{11}, N) \succeq 0 \) is equivalent to \( \text{diag}(X_{11}, N) \succeq \varepsilon I_{2n} \) for some \( \varepsilon > 0 \), it can be written as the following rank constraint

\[
\text{rank} \left[ \begin{bmatrix} I_{2n} & D \\ D^T & \text{diag}(X_{11}, N) - \varepsilon I_{2n} \end{bmatrix} \right] = 2n
\]

Noting that the cost function in (20) is bounded by \( mn + \nu \max \{2n, (n + m)\} \), we can use an argument similar to the one used in the proof of lemma [V.1] to to show that \( (P_S) \), if feasible, can be equivalently cast in the following
\[
\min_{X_{11}, X_{12}, K, N} \|K\|_0 + \nu \text{rank} \begin{bmatrix}
X_{11} & X_{12} \\ I_n & (KC)^T
\end{bmatrix} + \rho \text{rank} \begin{bmatrix}
I_{2n} & D \\ DT & M - \varepsilon I_{2n}
\end{bmatrix}
\]

subject to:

\[AX_{11} + X_{11}^TA^T + BX_{12}^T + X_{12}B^T + N = 0, \]

\[K \in \mathcal{K}, \]

\[\varepsilon > 0,\]

where

\[\nu > mn\]

\[\rho > mn + \nu \max\{2n, (n + m)\}.\]

Next, we are going to show that the cost function of (23) is equal to the cost function of (21) for \(\rho\) and \(\nu\) chosen to be integers satisfying the conditions. It can be easily verified that \(\|K\|_0 = \text{diag}(\text{vec}(K))\), also, the ranks of the square matrices \(\Psi_i\)'s are equal to the rank of \(\begin{bmatrix}
X_{11} & X_{12} \\ I_n & (KC)^T
\end{bmatrix}\).

If the parameters \(\rho\) and \(\nu\) are integers, we can construct a block diagonal matrix in the following form

\[
\text{diag}\{\text{vec}(K), \Psi_1, \cdots, \Psi_\nu, \Phi_1, \cdots, \Phi_\rho\}
\]

Thus, the rank of such matrix is equal to the sum of the rank of its constructing block matrices. Therefore, it is equal to the cost function of the optimization problem (21), which completes our proof.

The above formulation is in the form of **Affine Rank Minimization Problem** (ARMP), which consists of minimizing the rank of a matrix subject to affine/convex constraints with the general form

\[
\min_X \text{rank}(X) \\
\text{s.t. } A(X) = b
\]

for a fixed infinitesimal \(\varepsilon > 0\). ARMP has been investigated thoroughly in the past decade and several heuristics have been proposed to solve it. For example, Recht et al. in [34] showed that nuclear norm relaxation of rank can recover the minimum rank solution if certain property, namely Restricted Isometry Property (RIP), holds for the linear mapping. A family of Iterative Re-weighted Least Squares algorithms which minimize Schatten-p norm, i.e. \(\|X\|_{S_p} = \text{Tr}(X^TX + \gamma I)^{p/2}\), of the matrix as a surrogate for its rank is also introduced in [45]. Singular Value Projection (SVP) algorithm is also guaranteed to recover the low rank solution for affine constraints which satisfy RIP [46].
Fig. 1: Sparsity pattern of (a) the network system (b) the optimal sparse feedback controller \( \{\lambda = 10, \rho = 100\} \). (c) represents the underlying graph of the sparse controller.

**Remark V.5.** The discrete-time counterpart of the optimization problem \([P5]\) can be formulated as

\[
\min_{X_{11}, X_{12}, K, N} \|K\|_0 \quad \text{subject to} \quad A^T X_{11} A + A^T X_{12} + X_{12}^T A + X_{22} - X_{11} + N = 0,
\]

\[
X_{11} \succ 0, \quad N \succ 0,
\]

\[
Y^T = B K C,
\]

\[
K \in \mathcal{K},
\]

\[
\text{rank} \begin{bmatrix}
X_{11} & X_{12} \\
X_{12}^T & X_{22} \\
I_n & Y^T
\end{bmatrix} = n.
\]

Hence, the results, developed in this section, are applicable to the problem of identifying the sparsest stabilizing controller for discrete-time linear time invariant systems.
VI. NUMERICAL EXAMPLES

In this section, we use several examples to demonstrate how our proposed rank constrained optimization approach can be exploited to solve the optimal sparse output feedback controller design problem considering the input/output constraints.

A. Unstable Lattice Network System

Here, we illustrate an example in which we design an optimal sparse state feedback controller for an unstable networked system defined on a $5 \times 5$ lattice. The entries of its corresponding system matrix are randomly generated scalars drawn from the standard uniform distribution on the open interval $(-1, 1)$, and it is assumed the state performance matrix $Q$ to be an identity matrix, while the control performance weight $R = 10I$. Here, we used the traditional LQR controller as the benchmark to measure the performance of our proposed algorithm. Performing standard LQR design method, our results show that the optimal cost, for the case of LQR control design, is $J^* = 211.173$.

Next, we applied Algorithm 1 to design an optimal sparse controller with the parameters values $\lambda = 10$ and $\rho = 100$, while keeping the performance weights unchanged. It can be observed that the optimal controller cost function increases to $J^* = 230.6989$, which is about 9.2% higher, comparing to that of the LQR design. On the other hand, the number of non-zero entries of the controller gain drops to 97, i.e. $\|K\|_0 = 97$. This means a major decrease in the number of non-zero entries of the controller gain. Figures 1a and 1b show the sparsity structure of the system network and the obtained sparse controller. In Figure 1c the graph representation of the generated sparse controller is depicted.

Additionally, we present a brief case study that compares our approach with the Sparsity Promoting Optimal...
Feedback Control (SPOFC) method, proposed in [47], [23]. The SPOFC method essentially solves a different control problem, since it solves the $\mathcal{H}_2$ problem, modified by adding a sparsity promoting penalty function to its cost function and obtain a sub-optimal sparse state feedback controller, while our proposed approach is built upon adjusting the $LQR$ problem to achieve a sparse output feedback controller. Moreover, the approach in SPOFC algorithm fails to directly incorporate the norms bounds on the inputs/outputs and the controller predefined structure. Nonetheless, for comparison purposes and demonstrating the comparable performance of our method, we have obtained the MATLAB source code for SPOFC from the website www.ece.umn.edu/mihailo/software/lqrsp, and applied both our method and SPOFC to design sparse state feedback controllers for the randomly generated system. Fig. 2 depicts the results of the simulations performed using both controller design methods. As predicted, the quadratic cost of the closed loop system increases, as the the parameter $\lambda$ becomes larger. Moreover, increasing the value of this parameters on the system promotes the sparsity level of feedback gain matrix. Figure 2 depicts the effect of the parameter $\lambda$ on the performance of the closed loop system and the number of non-zero entries of the controller gain. In Fig. 2a the $Y$-axis represents percentage of the performance loss, which is defined as $(J^* - J_{LQR}^*)/J_{LQR}^*$. The density level percentage of the controller gain is also shown in Fig. 2b when the parameter $\lambda$ varies from $10^{-3}$ to 10.

The simulation results, demonstrated in figure 2, show that the SPOFC approach compromises the performance for a sparser controller in comparison to the our method. Our proposed method assures less performance loss by obtaining denser feedback controller. The disagreement between the optimal solutions of the two algorithms is mainly due to convergence to different local optima. It should also be noted the optimization parameter $\rho$ plays an important, but different, role in adjusting the convergence properties in both of the methods. Hence, setting the parameter $\rho$ to the same value in both optimizations may not be the most accurate choice for the comparison purposes. Moreover, the choice of the system also affects the design performance of both methods. Overall, our extensive simulation results suggest the comparable performance of both approaches. Considering the fact that our problem formulation and solving procedure, which is completely different from the preceding method, generates roughly the same sparse controller, it can be concluded that the derived sparse controller is likely to be the best we can obtain.

B. Sub-exponentially Spatially Decaying System

To study the effects of parameters $\lambda$ and $\rho$ on the performance of our proposed method, we have run extensive simulations on a randomly generated sub-exponentially spatially decaying system [29]. In such systems, it is assumed the entries of the system matrices decay as they get further from the diagonal, thus we define the matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ as

$$
\begin{align*}
a_{ij} &= C_A a e^{-\alpha_A |i-j|^{\beta_A}} \\
b_{ij} &= C_B b e^{-\alpha_B |i-j|^{\beta_B}}
\end{align*}
$$

where $a$ and $b$ are uniformly distributed random variables on the open interval $(-1, 1)$. By employing Algorithm 1 till the rank constraint is satisfied, we have depicted the performance degradation and density level of the generated
controllers in figure 3 for different values of $\rho$ and $\lambda$. Although the proposed algorithm has converged for all choices of parameters in this simulation, It seems that the choice of the optimization parameter $\rho$ is needed to be at least one order of magnitude larger than the parameter $\lambda$ in order to guarantee the convergence to a proper sub-optimal minimum. In addition, since the main objective in designing a sparse controller is to obtain a controller with minimum number of nonzero entries and lowest performance decline, we have also presented the plot of the lowest performance loss obtained for particular values of density level in figure 3c. As expected, it can be observed the performance loss grows as the sparsity level of the controller increases.

C. Optimal Sparse Controller with Upper Bound Imposed on the Control Input Norm

In this example, we illustrate the effect of bounding the norm of the control input on the sparsity of the controller matrix. Considering a randomly generated sub-exponential spatially decaying systems, with the same parameter values used in section VI-B we first designed a sparse controller with no constraint on the control input. Our
results show that the controllers number of nonzero entries and its performance loss, with respect to the cost of the LQR controller which is 639.1912, are 55 and 9.3% respectively. It is also observed that for the generated controller, we have $\|u\|_{L^\infty(R^m)} = 228.66$.

We then redesigned the controller, using the re-weighted $\ell_1$ minimization method, by containing its control input norm in the interval $[0, 200]$, and obtained controller has the following characteristics: $\|K\|_0 = 105$ and $J = 737.16$. Although we bounded the control input norm to an approximately 10% lower value, the obtained controller demonstrates 50% less sparse pattern and 6% higher performance loss. The simulations results, depicted in figure 4, not only verifies the capability of our method to incorporate bounds on the control input, as well as the system output, but also reveals the adverse impact of sparsifying the controller matrix on the control input norm.

VII. CONCLUSIONS

In this paper, we have proposed a new framework for optimal sparse output feedback control design, which is capable of incorporating structural constraints on the feedback gain matrix as well as norm bounds on the inputs/outputs of the system. We have shown that problem can be converted to a rank constrained optimization problem with no other non-convex constraints. Using the proposed formulation, we have presented an optimization problem which yields an upper bound for the optimal value of the optimal sparse state feedback control problem. Exploiting the relaxation of the $\ell_0$-norm with the $\ell_1$-norm, we have also expressed that local optimum of the relaxed optimization problem, in its general form, can be obtained by performing ADMM algorithm, which is, in essence, iteratively solving the relaxed problem and projecting its solution to the space of matrices with rank $n$. For the special case, where the objective is merely sparsity pattern recognition of the controller gain, we have demonstrated that the problem can be reduced to an Affine Rank Minimization. The simulation results are also provided to
illustrate the utility and performance of our proposed approach. As compared to the results of [23], our results show that while our proposed method has the advantage of performing the output feedback control design restricted by various forms of nonlinear constraints, the performance of our approach is on a par with theirs when applied to the regular sparse state feedback controller design problem.

REFERENCES

[1] V. Blondel and J. N. Tsitsiklis, “NP-hardness of some linear control design problems,” *SIAM Journal on Control and Optimization*, vol. 35, pp. 2118 – 2127, 1997.

[2] S. Schuler, P. Li, J. Lam, and F. Allgöwer, “Design of structured dynamic output-feedback controllers for interconnected systems,” *International Journal of Control*, vol. 84, no. 12, pp. 2081–2091, 2011.

[3] B. Polyak, M. Khlebnikov, and P. Shcherbakov, “An LMI approach to structured sparse feedback design in linear control systems,” in *Proceeding of the 2013 European Control Conference*, July 2013, pp. 833–838.

[4] N. Motee and A. Jadbabaie, “Distributed multi-parametric quadratic programming,” *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2279–2289, Oct 2009.

[5] 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.

[6] K. Dvijotham, E. Todorov, and M. Fazel, “Convex structured controller design,” *CoRR*, vol. abs/1309.7731, 2013. [Online]. Available: http://arxiv.org/abs/1309.7731

[7] 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.

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

[10] J. Lavaei, “Optimal decentralized control problem as a rank-constrained optimization,” in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 39–45.

[11] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, “Convex relaxation for optimal distributed control problem – Part I: Time-domain formulation,” *optimization*, vol. 16, p. 17.

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

[13] 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.

[14] X. Qi, M. Salapaka, P. Voulgaris, and M. Khannouch, “Structured optimal and robust control with multiple criteria: A convex solution,” *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1623–1640, Oct 2004.

[15] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[16] R. A. DeVore, “Deterministic constructions of compressed sensing matrices,” *Journal of Complexity*, vol. 23, no. 46, pp. 918 – 925, 2007.

[17] R. Zahedi, L. Krakow, E. Chong, and A. Pezeshki, “Adaptive estimation of time-varying sparse signals,” *IEEE Access*, vol. 1, pp. 449–464, 2013.

[18] R. Zahedi, A. Pezeshki, and E. K. Chong, “Measurement design for detecting sparse signals,” *Physical Communication*, vol. 5, no. 2, pp. 64 – 75, 2012, compressive Sensing in Communications.

[19] M. Davenport, P. Boufounos, M. Wakin, and R. Baraniuk, “Signal processing with compressive measurements,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, April 2010.

[20] E. Candes and T. Tao, “Decoding by linear programming,” *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, Dec 2005.

[21] E. J. Candes, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
[22] 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.

[23] 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. 58, no. 9, pp. 2426–2431, 2013.

[24] S. Schuler, C. Ebenbauer, and F. Allgöwer, “$\ell_0$-system gain and $\ell_1$-optimal control,” in *Preprints of the 18th IFAC World Congress*, August 2011, pp. 9230–9235.

[25] 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.

[26] R. Chartrand, “Exact reconstruction of sparse signals via nonconvex minimization,” *Signal Processing Letters, IEEE*, vol. 14, no. 10, pp. 707–710, Oct 2007.

[27] M. Lai and J. Wang, “An unconstrained $\ell_q$ minimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems,” *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 82–101, 2011.

[28] N. Motee and Q. Sun, “Measuring sparsity in spatially interconnected systems,” in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, Dec 2013, pp. 1520–1525.

[29] ——, “Sparsity and spatial localization measures for spatially distributed systems;” *arXiv preprint arXiv:1402.4148*, 2014.

[30] ——, “Sparsity measures for spatially decaying systems,” in *Proceedings of the 2014 American Control Conference*, June 2014, pp. 5459–5464.

[31] B. Recht, W. Xu, and B. Hassibi, “Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization,” in *Proceedings of the 47th IEEE Conference on Decision and Control*. Cancun, Mexico, December 2008, pp. 180–185.

[32] ——, “Null space conditions and thresholds for rank minimization,” *Mathematical Programming*, vol. 127, pp. 175–211, July 2011.

[33] M. Mesbah and G. P. Papavassilopoulos, “On the rank minimization problem over a positive semidefinite Linear Matrix Inequality,” *IEEE Transactions on Automatic Control*, vol. 42, pp. 1239–243, February 1997].

[34] B. Recht, M. Fazel, and P. A. Parillo, “Guaranteed minimum rank solutions to Linear Matrix Equations via nuclear norm minimization,” *SIAM Review*, vol. 52, pp. 471–501, July 2007.

[35] E. Feron, V. Balakrishnan, S. Boyd, and L. El Ghaoui, “Numerical methods for $H_2$ related problems;” in *Proceedings of the 1992 American Control Conference*, June 1992, pp. 2921–2922.

[36] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[37] E. J. Candès, “Compressive sampling,” in *International Congress of Mathematicians*, vol. III. Zurich: Eur. Math. Soc., December 2006, pp. 1433–1452.

[38] H. Yu and V. K. N. Lau, “Rank-constrained Schur-convex optimization with multiple trace/log-det constraints;” *IEEE Transactions on Signal Processing*, vol. 59, pp. 304–314, 2011.

[39] M. X. Goemans and D. Williamson, “Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.” *J. ACM*, vol. 42, no. 6, pp. 1115–1145, Nov. 1995.

[40] M. A. Glowinski, R., “Sur l’approximation, par lments finis d’ordre un, et la rsolution, par pnalisation-dualit d’une classe de problmes de dirichlet non linaires,” *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, vol. 9, no. R2, pp. 41–76, 1975.

[41] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17 – 40, 1976.

[42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the Alternating Direction Method of Multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, pp. 1–124, 2011.

[43] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi, “Simultaneously structured models with application to sparse and low-rank matrices;” *CoRR*, vol. abs/1212.3753, 2012.

[44] Y. Chen, Y. Chi, and A. Goldsmith, “Estimation of simultaneously structured covariance matrices from quadratic measurements,” in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 7669–7673.

[45] K. Mohan and M. Fazel, “Iterative learning algorithms for matrix rank minimization,” *Journal of Machine Learning Research*, vol. 13.

[46] R. Meza, P. Jain, and I. S. Dhillon, “Guaranteed rank minimization via singular value projection,” *CoRR*, vol. abs/0909.5457, 2009.
[47] F. Lin, M. Fardad, and M. R. Jovanović, “Sparse feedback synthesis via the Alternating Direction Method of Multipliers,” in Proceedings of the 2012 American Control Conference, 2012, pp. 4765–4770.