Sparse Linear-Quadratic Feedback Design Using Affine Approximation

MirSaleh Bahavarnia

Abstract—We consider a class of $\ell_0$-regularized linear-quadratic (LQ) optimal control problems. This class of problems is obtained by augmenting a penalizing sparsity measure to the cost objective of the standard linear-quadratic regulator (LQR) problem in order to promote sparsity pattern of the state feedback controller. This class of problems is generally NP hard and computationally intractable. First, we apply a $\ell_1$-relaxation and consider the $\ell_1$-regularized LQ version of this class of problems, which is still nonconvex. Then, we convexify the resulting $\ell_1$-regularized LQ problem by applying affine approximation techniques. An iterative algorithm is proposed to solve the $\ell_1$-regularized LQ problem using a series of convexified $\ell_1$-regularized LQ problems. By means of several numerical experiments, we show that our proposed algorithm is comparable to the existing algorithms in the literature, and in some cases it even returns solutions with superior performance and sparsity pattern.

I. INTRODUCTION

The area of distributed control systems has been growing rapidly in the past decade and it has been applied to various real-world problems such as formation control of autonomous vehicles, power networks, transportation networks, mobile wireless networks, only to name a few. In several important applications, the centralized control methodologies cannot be applied due to the lack of access to global information in subsystem level throughout the network. This design constraint has been motivated researchers to consider the possibility of designing near-optimal sparse feedback controllers for large-scale dynamical networks [1], [2].

In this paper, we consider a class of $\ell_0$-regularized linear-quadratic (LQ) optimal control problems. This class of problems can be formulated by considering the standard linear-quadratic regulator (LQR) problem and augmenting it with an additional term in the cost objective in order to promote sparsity pattern of the state feedback controller. This class of $\ell_0$-regularized LQ problems is generally nonconvex and computationally intractable. This is basically due to the presence of $\ell_0$-measure in the cost objective and nonlinear terms in the Lyapunov equation that corresponds to the closed-loop stability of the system. The nonconvex sparsity-promoting term in the cost objective can be convexified by replacing it with $\ell_1$-norm [3] in order to relax the problem as a $\ell_1$-regularized LQ problem. This is still a nonconvex problem. Our main contributions in this paper are twofold. First, we convexify the $\ell_1$-regularized LQ problem by applying affine approximation techniques to convexify the nonlinear terms in the corresponding Lyapunov equation. Second, we propose an iterative algorithm to solve the $\ell_1$-regularized LQ problem using a series of convexified $\ell_1$-regularized LQ problems.

In the literature review, we discuss that our results are close in spirit to [4]. In the numerical experiments section, we analyze the performance of our proposed algorithm by means of several simulations and compare our results to the algorithm proposed in [2]. Our simulation results reveal that our proposed algorithm is comparable to the existing algorithms in the literature, and in some cases it even returns solutions with superior performance and sparsity pattern.

Some works have been done in order to tract problems similar to sparsity-promoting LQR problem. One of the main ideas has been the change of $\ell_0$-norm with convex $\ell_1$-norm and then solving the penalized problem with ADMM method [1], [2]. Another method which has been used to get some bounds on the optimal value of such sparsity-constrained problems works based on projection and gives some useful intuition about optimal value of the corresponding sparsity-constrained problem [5]. All the methods mentioned so far (except the [4]), have been proposed for the continuous-time sparsity-promoting state feedback gain controller. In addition to such continuous-time state feedback gain controllers, some methods have been presented by considering the discrete-time standard LQR problem which some of them propose sparsity-promoting state feedback gain controllers [6], [7], [4], [8],[9] and [10]. In one of these recent methods, decentralized state feedback gain controller is presented by using some convex relaxations where some graph theoretic proofs are provided to determine the upper bound for rank of such a relaxed SDP solution and if such a rank is equal to 1, then the globally optimal solution can be reconstructed from the relaxed SDP solution [7] and [8]. However, the solution obtained by such a relaxation-based method is decentralized and is not presented for general sparse controllers. But, it is spanning both finite and infinite horizon discrete-time sparsity-promoting LQR problems [7] and [8].

Recently, a method has been revealed which yields suboptimal structured and sparse feedback gains on the basis of iterative convex programming [4]. In such a method, the objective function of optimization problem has an $\ell_2$-norm form. Then, the non-convexity of problem is extracted as an inversion of an optimization variable. After that, the equality condition having such a non-convexity is replaced by two complementary inequalities. One of them is characterized by Schur complement and the other one is rewritten as an equality having some penalized term [11]. It should be noted that the non-convex part appeared on such an equality is linearized around an optimal estimate. The achieved results are really notable in comparison with the simplicity of the
used iterative convex programming. However, no convergence proof is presented [4]. In [10] as well as [4], convex optimization has been used as an effective tool to design a sub-optimal sparse static output feedback controller.

One of the well-known systems studied in Control Theory is spatially decaying systems which has been discussed in recent years [12], [13], [14], [15]. In these recent works, a large class of spatially decaying systems is classified where their quadratically-optimal feedback controllers inherit spatial decay property from the dynamics of the underlying system. Moreover, they propose a method based on new notions of q-Banach algebras where sparsity and spatial localization features of spatially decaying systems can be studied when q is chosen sufficiently small. In this paper, in first subsection of numerical examples, we employ such a class of systems for verifying our proposed design method. The results are considerably consistent with our expectations. The state feedback gain controllers have patterns very similar to the structure of the underlying spatially decaying systems with reasonable sparsification.

Another class of spatially distributed systems is cyclic systems which arises in biochemical reactions. Such a class of systems have been investigated mathematically in [16], [17], [18], [19], [20], [21]. The second subsection of numerical examples is devoted to considering such well-known systems and their sub-optimal sparse controllers design. Also, we state how the sub-optimal sparsest solution for such a special case can be derived analytically. The interesting fact about such a sub-optimal sparsest solution is that it just has 1 nonzero term.

This paper is organized as follows: Section II is an Introduction to the Distributed Optimal Sparse Controllers, our main contributions and diverse methods proposed to design such controllers. Section III is devoted to formulate the ℓ₀-regularized LQR design problem. In Section III, we see the main part of the paper which shows how ℓ₁-norm, affine approximation and Schur complement help us to cast the ℓ₁-regularized LQR problem as a simple convex SDP form which is solved with CVX toolbox [22]. Section IV contains the proposed algorithm which finds the sub-optimal sparse state feedback gain controller. Section V gives some numerical experiments of spatially decaying systems and cyclic systems which show how useful the new Affine Approximation-based formulation presented in Section III is. Finally, Section VI concludes the paper together with mentioning some future work at the end of it.

II. PROBLEM FORMULATION

We consider the class of linear time-invariant systems

\[ \dot{x} = Ax + Bu \quad \text{with} \quad x(0) = x_0, \]

where \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{n \times m} \). It is assumed that the pair \((A, B)\) is controllable and the initial condition \(x_0\) is drawn from a standard normal distribution (i.e., with zero mean and unit standard deviation). Let us define the following ℓ₀-regularized Linear-Quadratic (LQ) optimal control problem:

### ℓ₀-Regularized LQ (ℓ₀-RLQ) Problem:

Minimize \( \mathbb{E} \left\{ \int_0^\infty (x^T Q x + u^T R u) dt \right\} + \alpha_1 \|K\|_{\ell_0} \)

subject to:

\[ \dot{x} = Ax + Bu, \]

\[ u = K x, \]

\[ x(0) = x_0, \]

\[ K : \text{stabilizing}. \]

In this formulation, \( K \in \mathbb{R}^{m \times n} \) is the state feedback gain matrix, \( R > 0 \) is the control weight matrix, \( Q \succeq 0 \) is the state weight matrix, and \( \|K\|_{\ell_0} \) is the number of nonzero elements of matrix \( K \) [23]. The ℓ₀-RLQ problem can be viewed as a penalized standard LQR problem, where the penalty term \( \|K\|_{\ell_0} \) is augmented to improve sparsity pattern of the state feedback gain \( K \).

By incorporating constraints of ℓ₀-RLQ problem, we can cast the objective functional in this problem in a simpler form using the following standard procedure

\[ \int_0^\infty (x^T Q x + u^T R u) dt = \int_0^\infty \text{Tr}(x^T (Q + K^T R K)x) dt. \]

\[ = \int_0^\infty \text{Tr}(x_0^T e^{(A+BK)t} (Q + K^T R K)e^{(A+BK)^T t} x_0) dt = \text{Tr}(X_0 X), \]

where \( X_0 = x_0 x_0^T \) and

\[ X = \int_0^\infty e^{(A+BK)t} (Q + K^T R K)e^{(A+BK)^T t} dt. \]

In order to satisfy the last constraint in ℓ₀-RLQ problem, we apply the standard Lyapunov theorem and conclude that \( K \) is stabilizing if and only if \( X \) is the unique positive definite solution of the following Lyapunov equation [24]

\[ (A + B K)^T X + X (A + B K) + Q + K^T R K = 0. \]

In the final step, by taking expectation from the equivalent cost (2) it follows that \( \mathbb{E}(\text{Tr}(X_0 X)) = \text{Tr}(X) \). This can be done as trace is a linear operator and the expectation operator distributes over it. As a result of our simplifications, ℓ₀-RLQ problem can be rewritten in the following compact form:

### Equivalent Form of the ℓ₀-RLQ Problem:

Minimize \( \text{Tr}(X) + \alpha_1 \|K\|_{\ell_0} \)

subject to:

\[ (A + B K)^T X + X (A + B K) + Q + K^T R K = 0, \]

\[ X > 0. \]
measure with its best convex approximation, i.e., ℓ₁-norm [3] in order to get the following relaxed optimal control problem:

\[ \ell_1\text{-Regularized LQ (ℓ₁-RLQ) Problem:} \]

Minimize \( \text{Tr}(X) + \alpha_1 \| K \|_{\ell_1} \) \hspace{1cm} (6)
subject to:
\[ (A + BK)^T X + X(A + BK) + Q + K^T R K = 0, \hspace{1cm} (7) \]
\[ X \succ 0. \hspace{1cm} (8) \]

The ℓ₁-RLQ problem is still nonconvex due to nonlinear constraint (7). This constraint has differentiable algebraic forms. In the following section, we employ an affine approximation method to convexify the ℓ₁-RLQ problem. Then, we use our results to propose an iterative algorithm to compute sub-optimal solutions for the original ℓ₀-RLQ problem.

**Remark 1:** Alternative forms of ℓ₁-RLQ problem have been considered before in [2], where the authors employ an ADMM-based approach to handle nonlinear inequality (7). In our numerical experiments in Section VI, we will repeatedly compare the performance of our proposed algorithm to that of [2].

### III. SDP-Restriction of the REℓ₁-RLQ Problem

In this section, we derive a SDP-restriction of a relaxed form of ℓ₁-RLQ problem. In the first step using the following lemma, we rewrite ℓ₁-RLQ problem in a new equivalent form where the source of nonlinearity in ℓ₁-RLQ problem is extracted and isolated as a new constraint, i.e., inequality (7). Then, we will show how to utilize affine approximations to estimate that nonlinear constraint and come up with a SDP-restriction of a relaxed form of ℓ₁-RLQ problem. Let us denote \( S^n_{++} \) and \( S^n_+ \) to be the positive definite cone and positive semi-definite cone of \( n \times n \) matrices, respectively.

**Lemma 1:** The ℓ₁-RLQ problem is equivalent to the following optimization problem:

**Equivalent Form of the ℓ₁-RLQ Problem (Eℓ₁-RLQ):**

Minimize \( \text{Tr}(X) + \alpha_1 \| K \|_{\ell_1} \) \hspace{1cm} (9)
subject to:
\[ \begin{bmatrix} -Q - F + Y & 2X - P^T & P^T \\ 2X - P & I & 0 \\ P & 0 & \frac{1}{\delta} I \end{bmatrix} \preceq 0, \]
\[ P = X - \frac{A + BK}{2}, \]
\[ X \succeq \epsilon_1 I, \]
\[ \begin{bmatrix} Y & P^T \\ P & \frac{1}{1+\delta} I \end{bmatrix} \succeq 0, \]
\[ \begin{bmatrix} F & K^T \\ K & R^{-1} \end{bmatrix} \succeq 0, \]
\[ Y \preceq (1 + \delta) P^T P, \]

where \( X \in S^n_{++}, Y \in S^n_+, F \in S^n_+, K \in \mathbb{R}^{m \times n} \) and \( P \in \mathbb{R}^{n \times n} \).

**Proof:** See Appendix.

In the following proposition, we take an important step toward dealing with the nonconvex term (9) in REℓ₁-RLQ problem.

**Proposition 1:** The REℓ₁-RLQ problem can be relaxed as:

**Relaxed Eℓ₁-RLQ Problem (REℓ₁-RLQ):**

Minimize \( \text{Tr}(X) + \alpha_1 \| K \|_{\ell_1} \)
subject to:
\[ \begin{bmatrix} -Q - F + Y & 2X - P^T & P^T \\ 2X - P & I & 0 \\ P & 0 & \frac{1}{\delta} I \end{bmatrix} \succeq 0, \]
\[ P = X - \frac{A + BK}{2}, \]
\[ X \succeq \epsilon_1 I, \]
\[ \begin{bmatrix} Y & P^T \\ P & \frac{1}{1+\delta} I \end{bmatrix} \succeq 0, \]
\[ \begin{bmatrix} F & K^T \\ K & R^{-1} \end{bmatrix} \succeq 0, \]
\[ Y \preceq (1 + \delta) P^T P, \]

where \( \| . \|_* \) denotes the nuclear norm of \( (.) \) which is defined as sum of singular values of \( (.) \). And also for a given matrix \( U \), we have \( \| U \|_* = \text{Tr}(\sqrt{U^T U}) \) where \( \sqrt{U^T U} \) is a matrix for which we have \( U^T U = (\sqrt{U^T U})^2 \).

**Proof:** Our goal is to prove that REℓ₁-RLQ problem is a relaxation of REℓ₁-RLQ problem. It suffices to show that the feasible set of REℓ₁-RLQ problem, which is represented by \( F_1 \), contains the feasible set of Eℓ₁-RLQ problem, which is represented by \( F_2 \). Moreover, suppose that \( F_3 \) denotes the set of all feasible points specified by all constraints of REℓ₁-RLQ problem except constraint (10) and \( F_4 \) denotes the set of all feasible points determined only by constraint (10). It is straightforward to verify that \( F_2 \subseteq F_3 \). This is because the set of constraints specifying \( F_3 \) is a subset of set of constraints specifying \( F_2 \). This implies that for a feasible point \( (X_1, K_1, F_1, P_1, Y_1) \in F_2 \), it follows that \( (X_1, K_1, F_1, P_1, Y_1) \in F_3 \). Since \( (X_1, K_1, F_1, P_1, Y_1) \in F_2 \), \( Y_1 = (1 + \delta) P_1^T P_1 \) is satisfied. Therefore, \( \| Y_1 - (1 + \delta) P_1^T P_1 \|_* = 0 \) is also satisfied and consequently \( \| Y_1 - (1 + \delta) P_1^T P_1 \|_* \leq \epsilon \) is also true. Thus, \( (X_1, K_1, F_1, P_1, Y_1) \in F_4 \). Considering the fact that \( F_3 \cap F_4 = F_1 \) is satisfied, it implies that \( (X_1, K_1, F_1, P_1, Y_1) \in F_1 \). Thus, it proves our claim that \( F_1 \subseteq F_2 \).

**Remark 2:** It should be noted that there is no much difference between choice of \( \| . \|_* \) and \( \| . \| \) in (10). It just affects the choice of parameter \( \epsilon \).

In the following step, we employ a real analysis idea [25] to convexify the right hand side of inequality (10) in REℓ₁-RLQ problem, which is the only nonconvex term appearing in REℓ₁-RLQ problem. By performing some matrix calculus,
we calculate the convex affine approximation of \( M = (1 + \delta)P^TP \) as follows:
\[
N = (1 + \delta)(P^T \tilde{P} + \tilde{P}^T(P - \tilde{P}))
\]
where \( \tilde{P} \) is an estimate of \( P \). We refer to [22] for more details on affine approximations of functions of matrices.

So, we substitute the nonconvex condition \( \| Y - M \| \leq \epsilon \) with convex condition \( \| Y - N \| \leq \frac{\epsilon}{n} \) where \( \| \cdot \| \) denotes operator norm which is defined as largest singular value of \( (\cdot) \). Finally, using the proof of proposition 2.1. in [26], we get the following proposition:

**Proposition 2:** The RE\( \ell_1 \)-RLQ problem can be restricted as the following SDP form:

**SDP Restriction of the RE\( \ell_1 \)-RLQ Problem (SDP-R-RE\( \ell_1 \)-RLQ):**

\[
\begin{align*}
\text{Minimize} & \quad \text{Tr}(X) + \alpha_1 \| K \|_{\ell_1} \\
\text{subject to:} & \quad \begin{bmatrix}
-Q - F + Y & 2X - P^T \\
2X - P & I \\
P & 0 \\
0 & \frac{1}{\delta} I
\end{bmatrix} \succeq 0, \\
P = X - \frac{A + BK}{2}, \\
X \succeq \epsilon I, \\
Y \begin{bmatrix} P^T \\ P \end{bmatrix} \succeq 0, \\
\begin{bmatrix} F & K^T \\ K & R^{-1} \end{bmatrix} \succeq 0, \\
\begin{bmatrix} \frac{\epsilon}{n} I & Y - N \\ Y - N & \frac{\epsilon}{n} I \end{bmatrix} \succeq 0.
\end{align*}
\]

**Proof:** Let us suppose that \( F_0 \) denotes the set specified by constraint \( \| Y - N \| \leq \frac{\epsilon}{n} \) and \( F_0 = F_3 \cap F_5 \).

So, we should prove that \( F_0 \subseteq F_1 \). Suppose that \( (X_2, K_2, F_2, P_2, Y_2) \) be a feasible point in \( F_0 \). Hence, we have \( (X_2, K_2, F_2, P_2, Y_2) \in F_3 \) and \( (X_2, K_2, F_2, P_2, Y_2) \in F_5 \). Using the fact indicated in inequality (2.1) in [26], we have
\[
\| Y - N \| \leq \frac{\epsilon}{n}\| Y - N \| \leq \epsilon.
\]
Thus, with this approach, \( Y \succeq M \), it implies that \( Y \succeq M \succeq N \). Now, \( Y - M \preceq Y - N \) and \( (Y - N) - (Y - M) = M - N \succeq 0 \) implies that \( \text{Tr}(Y - N) - \text{Tr}(Y - M) \geq 0 \). Since \( Y - M \) and \( Y - N \) are both symmetric matrices, their traces are same as their nuclear norms. Hence, we will have
\[
\| Y - M \|_n \leq \| Y - N \|_n.
\]
Thus, \( \| Y - M \|_n \leq \epsilon \) and subsequently \( (X_2, K_2, F_2, P_2, Y_2) \in F_4 \) are resulted. Consequently, \( (X_2, K_2, F_2, P_2, Y_2) \in F_4 \cap F_3 = F_1 \). Subsequently, \( F_0 \subseteq F_1 \) is satisfied and proof is done.

**Remark 3:** In the following, we try to explain the existing technical issue with the method proposed by [4]. In problem (RLX) in [4], very briefly, the authors have considered the \( Y - N \) as difference of two positive semi-definite matrices \( Z_+ \) and \( Z_- \) and forced trace of \( Z_+ \) to zero by minimizing some multiplier of it in objective function, then they have made a conclusion that \( Y - N \) becomes negative semi-definite and this together with \( Y \succeq X^{-1} \) forces \( Y = X^{-1} \). We know that \( Y \succeq N \), thus \( Y - N \) cannot be negative semi-definite. We are just able to minimize \( \| Y - N \| \). The point is that when we force trace of \( Z_+ \) to zero, it cannot be necessarily implied that \( Z_+ - Z_- \) tends to \(-Z_- \) and so it is negative semi-definite. Because, \( Z_- \) is an optimization variable and its trace can be either tend to zero and if the rate of tending be greater than of \( Z_+ \), then \( Z_+ - Z_- \) cannot be negative semi-definite. Also, if we run algorithm proposed by [4], we will see that not only \( Z_+ - Z_- \) is not negative semi-definite, but also it is positive semi-definite which is previously theoretically guaranteed. Thus, we minimize \( \| Y - N \| \) instead of using such a method presented in [4].

**Remark 4:** The other point is that we have characterized the approximated constraint via some parameter called \( \epsilon \) in feasible set. However, authors in [4], has done such a goal by considering some parameter called \( \alpha_2 \) in objective function. Although, these two approaches are equivalent by ignoring the issue arisen in \( Z_+ - Z_- \), our proposed method gives much better insight. Because, in proposition 2, we show how \( \| Y - N \| \) and \( \| Y - M \| \) are related to each other.

We consider the SDP-R-RE\( \ell_1 \)-RLQ problem and solve it iteratively by using CVX toolbox. The procedure is as follows:

1. Firstly, we start to solve the SDP-R-RE\( \ell_1 \)-RLQ problem for initial guess \( P^0 \). Then, at each step we update the previous value of \( \tilde{P} \) by its current computed value and continue this process until we reach to the desired level of accuracy.
2. It should be mentioned that the parameter \( \epsilon \) is firstly chosen as a sufficiently large number to find a primitive solution and then it decreases in a geometric progression form in order to achieve the sub-optimal solution. In section IV we propose and discuss such an algorithm in detail.

IV. AN ITERATIVE ALGORITHM TO SOLVE FOR SUB-OPTIMAL SPARSE SOLUTIONS

In order to compute a near-optimal sparse feedback controller, we propose an iterative algorithm which works based on solving a sequence of SDP forms [27]. The abstract notion of our proposed algorithm is as follows:

1. Given \( \alpha > 0 \), \( \beta < 1 \), \( \epsilon_2 > 0 \), \( \tilde{P} := P^0 \).
2. For \( i=1,2,... \) do
3. Set \( \epsilon := n\alpha \beta^{i-1} \).
4. Solve SDP-R-RE\( \ell_1 \)-RLQ problem to find optimal \( P^* \).
5. If \( \| P^* - P_\epsilon \|_F \geq \epsilon_2 \) or \( \| Y - (\tilde{Y} + (\delta + 1)P^* P^T) \|_F \geq \epsilon_2 \) then, set \( \tilde{P} \) equal to optimal \( P^* \), else, set \( P_{opt} = P^* \) break.
6. End for.

where \( \| \cdot \|_F \) denotes the Frobenius norm of \( (\cdot) \) which is defined as \( \| \cdot \|_F = \sqrt{\text{Tr}((\cdot)^T (\cdot))} \).

Our selection of Frobenius norm doesn’t provide any major advantage and it can be replaced by operator norm or nuclear norm [23]. It should be considered that \( P^0 \) is initiated.
as $X^0 = \frac{A + BK^0}{A + BK^0}$ where $(K^0)$ is the solution of standard LQR problem and $X^0$ is equal to $\int_0^\infty e^{(A + BK^0)t}(Q + K^{0T}RK^0)e^{(A + BK^0)t}dt$ which can be computed via solving a Lyapunov equation. The inequality appeared on item $5$ of the proposed algorithm are considered in a normalized form to reach higher accuracy.

As we will see in section $\text{V}$ our numerous experiments show that our proposed algorithm converges and suggests a sub-optimal sparse solution for $\ell_1$-LQR problem, however, we have no proof on its convergence.

V. NUMERICAL EXPERIMENTS

A. Spatially Decaying Systems

Here we give an example of spatially decaying systems which have been considered in [28], [12], [13], [14]. In such systems the absolute value of entries of $A$-matrix decreases spatially, i.e. when we start from an entry on main diagonal and get away from it, the absolute value of entries falls off.

As a sub-class of such spatially decaying systems, let us consider the sub-exponentially decaying systems. The $ij^{th}$ element of a random sub-exponentially decaying system $A$ is defined by $A_{ij} = c_i e^{-\alpha |i-j|^\beta A}$ where $c_i$ is the real random number corresponded to the $i^{th}$ row and generated by $\text{randn}$ command of MATLAB, $\alpha_A$ is a positive real number and $\beta_A$ is a real number between $0$ and $1$ [29]. Note that $\alpha_A$ determines the width of the band of the matrix $A$ and $\beta_A$ specifies the rate of decaying in sub-exponentially decaying system modeled by $A$. For large values of $\alpha_A$ we get narrow band width and for large values of $\beta$ we get high rate of spatially decaying. Since we have $A_{ii} = c_i$, $c_i$ appears on $i^{th}$ entry of main diagonal of matrix $A$. By taking $\alpha_A = 5$, $\beta_A = 0.5077$ and some randomly generated $c_i$'s we get the following sub-exponentially decaying $A$-matrix:

$$A = \begin{bmatrix}
0.5377 & 0.0036 & 0.0004 & 0.0001 & 0.0000 & 0.0000 \\
0.0124 & 1.8339 & 0.0124 & 0.0015 & 0.0003 & 0.0001 \\
-0.0018 & -0.0152 & -2.2588 & -0.0152 & -0.0018 & -0.0004 \\
0.0001 & 0.0007 & 0.0058 & 0.8622 & 0.0058 & 0.0007 \\
0.0000 & 0.0001 & 0.0003 & 0.0021 & 0.3188 & 0.0021 \\
-0.0000 & -0.0001 & -0.0002 & -0.0011 & -0.0088 & -1.3077
\end{bmatrix}$$

The set of eigenvalues of $A$ is $\{-2.2588, 1.8339, -1.3077, 0.5376, 0.8622, 0.3187\}$. Thus, $A$ is an unstable system. Assuming the $\alpha_1 = 0.005$, $\alpha = 0.00001$, $\beta = 0.99$, $\delta = 0.001$, $\epsilon_1 = 0.000001$, $\epsilon_2 = 0.00005$ and $B = Q = R = I$, Algorithm in section IV proposes the following sub-optimal sparse state feedback gain controller:

$$K^{opt} = \begin{bmatrix}
-1.6675 & -0.0666 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
-0.0113 & -3.9123 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & -0.1992 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & -2.1760 & -0.0014 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & -0.0001 & -1.3632 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.3304
\end{bmatrix}$$

The set of eigenvalues of $A + BK^{opt}$ consists of $-2.4575, -2.0789, -1.1298, -1.0444, -1.3139$ and $-1.6380$. Thus, the proposed sub-optimal sparse feedback gain controller is stabilizing. Also, the optimal value is $9.6968$. The optimal value of the standard LQR problem is $9.6965$ in this case.

In a similar way, we consider a $16 \times 16$ sub-exponentially decaying system and plot its corresponding matrices $A$ and $K$ in a logarithmic gray-scale spectrum. The corresponding parameters are as follows:

$$\alpha_1 = 0.005, \alpha = 0.00001, \beta = 0.99, \alpha_A = 5, \beta_A = 0.8903, \delta = 0.001, \epsilon_1 = 0.000001, \epsilon_2 = 0.00005$$

As we will see in section $\text{V}$, our numerous experiments show that our proposed algorithm converges and suggests a sub-optimal sparse solution for $\ell_1$-LQR problem, however, we have no proof on its convergence.

As it is depicted in Fig. 2, the sub-optimal sparse state feedback gain controller $K^{opt}$ has a sub-exponentially-like structure. In [30], it has been shown that if matrices $A$, $B$, $Q$ and $R$ belong to an operator algebra structure, the corresponding standard LQR problem solution inherits such a structure. So, here we see some similar intuitive observation which illustrates such a fact.

B. Cyclic Systems

Here, we consider cyclic systems to investigate our proposed algorithm in section IV. Cyclic systems arise in various applications such as biochemical reactions which have been widely studied from mathematical point of view in different works such as [16], [17], [18], [19], [20], [21]. The $A$-matrix of a cyclic system is specified satisfying the following properties:
• $A_{ii}$ is negative for all $i \in \{1, \ldots, n\}$.
• $A_{i(i-1)}$ is positive for all $i \in \{2, \ldots, n\}$.
• $A_{1n}$ is negative.

Let us assume that $A$ is a $10 \times 10$ random cyclic matrix generated by MATLAB. Then, the sparsity pattern of sparse state feedback gain controller for such a system with following parameters: $\alpha_1 = 0.5$, $\alpha = 0.00001$, $\beta = 0.99$, $\delta = 0.001$, $\epsilon_1 = 0.000001$, $\epsilon_2 = 0.00005$, $B = Q = R = I$ has been depicted in Fig. 3. As it is pointed in Fig. 3 the number of nonzero terms of the proposed sub-optimal state feedback gain controller is equal to 18 i.e. 82% of entries of the proposed sub-optimal state feedback gain controller $K_{opt}$ is equal to 0. Thus, it is sparse. Also, the optimal value of the standard LQR and optimal value of our proposed sub-optimal state feedback gain controller are 16.3760 and 17.8602 respectively. So, for $\alpha_1 = 0.5$, if the quadratic cost gets increased by 9.06%, the sparsity is improved by 82%.

Table I demonstrates the effect of penalizing parameter $\alpha_1$ on optimal value $J_{opt}$ and cardinality of sub-optimal controller $\|K_{opt}\|_{l_0}$ in our proposed method (i.e. CVX) and the method proposed by [2] (i.e. ADMM). Based on data set shown in Table I, Fig. 4 depicts how penalizing parameter, sparsity level and performance degradation relate to each other. As it is seen intuitively, Fig. 4 confirms this fact that as penalizing parameter $\alpha_1$ increases, the sparsity level decreases, however, performance degradation increases. Thus, there is a trade-off between sparsity level and performance degradation. As an instance for $\alpha_1 = 0.1$, by increasing 1.17% and 1.12% in quadratic cost for CVX and ADMM, respectively, 73% and 72% of sub-optimal sparse state feedback gain controller entries get equal to 0, respectively.

According to the second condition stated in [17] for the stability of the linear cyclic systems, we have the following sufficient condition:

$$\frac{(\prod_{i=1}^{n}(A_{i(i-1)}^{c})(-A_{in})}{\prod_{i=1}^{n}(-A_{ii})} < \left(\sec\left(\frac{\pi}{n}\right)\right)^{n},$$

where by superscript $c$, we mean the closed loop.

By doing some basic math, it yields that we have:

| $\log\alpha$ | CVX $J_{opt}$ | CVX $\|K_{opt}\|_{l_0}$ | ADMM $J_{opt}$ | ADMM $\|K_{opt}\|_{l_0}$ |
|---|---|---|---|---|
| 0.001 | 16.3761 | 56 | 16.3760 | 94 |
| 0.002 | 16.3762 | 55 | 16.3760 | 64 |
| 0.003 | 16.3765 | 55 | 16.3761 | 58 |
| 0.004 | 16.3768 | 55 | 16.3763 | 57 |
| 0.005 | 16.3773 | 54 | 16.3765 | 55 |
| 0.006 | 16.3778 | 52 | 16.3768 | 54 |
| 0.007 | 16.3784 | 48 | 16.3772 | 51 |
| 0.008 | 16.3790 | 47 | 16.3777 | 48 |
| 0.009 | 16.3797 | 47 | 16.3782 | 48 |
| 0.01 | 16.3805 | 45 | 16.3789 | 46 |
| 0.02 | 16.3914 | 40 | 16.3875 | 40 |
| 0.03 | 16.4070 | 38 | 16.4020 | 39 |
| 0.04 | 16.4200 | 37 | 16.4198 | 38 |
| 0.05 | 16.4451 | 35 | 16.4389 | 35 |
| 0.06 | 16.4663 | 32 | 16.4590 | 32 |
| 0.07 | 16.4896 | 31 | 16.4812 | 31 |
| 0.08 | 16.5152 | 30 | 16.5057 | 30 |
| 0.09 | 16.5422 | 28 | 16.5318 | 28 |
| 0.1 | 16.5677 | 27 | 16.5597 | 28 |
| 0.2 | 16.8308 | 22 | 16.7944 | 22 |
| 0.3 | 17.1527 | 20 | 17.1126 | 21 |
| 0.4 | 17.4952 | 19 | 17.4584 | 19 |
| 0.5 | 17.8602 | 18 | 17.8118 | 18 |
| 0.6 | 18.2200 | 15 | 18.1611 | 15 |
| 0.7 | 18.5817 | 15 | 18.5123 | 15 |
| 0.8 | 18.9483 | 15 | 18.8688 | 15 |
| 0.9 | 19.3169 | 15 | 19.2274 | 15 |
| 1 | 19.6853 | 14 | 19.5860 | 14 |
| 2 | 23.1799 | 13 | 22.9991 | 13 |
| 3 | 25.5533 | 13 | 26.0798 | 13 |
| 4 | 26.9975 | 12 | 28.8795 | 12 |
| 5 | 28.6990 | 11 | 31.3772 | 10 |
| 6 | 30.1690 | 11 | 33.6764 | 10 |
| 7 | 31.5150 | 10 | 35.8389 | 10 |
| 8 | 32.9970 | 10 | 37.8935 | 10 |
| 9 | 34.4342 | 10 | 39.8488 | 10 |
| 10 | 35.2497 | 10 | 41.7020 | 10 |

1) For $K_{jj}$ where $j \in \{1, \ldots, n\}$, we get the following condition:

$$K_{jj} < \frac{\left(-\cos\left(\frac{\pi}{n}\right)\right)^{n} \prod_{i=2}^{n} (A_{ii} - A_{i1})}{\prod_{i=1}^{n}(-A_{ii})} - A_{jj},$$

2) For $K_{j(j-1)}$ where $j \in \{2, \ldots, n\}$, we get the following condition:

$$-A_{j(j-1)} < K_{j(j-1)}.$$
more sparse state feedback gain controllers and the num-
ber of nonzero entries of state feedback gain controller is
decreasing. And also, the optimal value is increasing. Thus,
there is a fundamental trade-off between sparsity level and
performance degradation. An advantage of our proposed
method is that just like algorithm 1 in [4], it can be solved
by using CVX toolbox because of its convex SDP form.
Although sparsifying rate of our proposed method for small
values of $\alpha_1$ is higher compared to the proposed method
in [2], the quadratic cost minimizing rate of the latter one
is higher than ours. For large values of $\alpha_1$, such a behavior
is getting reversed. Our proposed method is slower than the
method presented in [2], however, it has a simpler form.
Again, we insist on this point that the quadratic cost achieved
by our proposed method is an upper bound for the exact
optimal solution of the Sparsity-promoting LQR problem.

The improvement of the accuracy of our proposed method
can be considered as a future work. Also, using the norms
other than $\ell_1$-norm or utilizing the weighted version of the
$\ell_1$-norm could result in better achievements. We try to use
the Affine Approximation method as an appropriate tool in
sparsity problems for which the sparsification occurs in time
domain.

APPENDIX

The lemma is proved by using the following equivalent
problems:

Minimize $\text{Tr}(X) + \alpha_1\|K\|_{\ell_1}$
subject to:

\[ (A + BK)^TX + X(A + BK) + Q + K^TRK = 0, \]
\[ X > 0. \]

Minimize $\text{Tr}(X) + \alpha_1\|K\|_{\ell_1}$
subject to:

\[ (A + BK)^TX + X(A + BK) + Q + K^TRK \leq 0, \]
\[ X > 0, \]
\[ F \succeq K^TRK. \]

Minimize $\text{Tr}(X) + \alpha_1\|K\|_{\ell_1}$
subject to:

\[ -Q - F \succeq \begin{bmatrix} X + \frac{(A + BK)^T}{2} & X - \frac{(A + BK)^T}{2} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} X + \frac{A + BK}{2} \\ X - \frac{A + BK}{2} \end{bmatrix}, \]
\[ X > 0, \]
\[ \begin{bmatrix} F & K^T \\ K & R^{-1} \end{bmatrix} \succeq 0. \]  

VI. DISCUSSION AND CONCLUSION

As it is shown, we propose a sub-optimal sparsity-
promoting state feedback gain controller on the basis of
affine approximation, Schur complement and SDP form. The
numerical experiments show that our proposed method gives
an appropriate approximation of optimal sparse solution of
the $\ell_0$-regularized LQR problem. For $\alpha_1 = 0$ we get the
standard LQR solution and by increasing it, we achieve
more sparse state feedback gain controllers and the num-

Fig. 4. The Performance Degradation $\frac{\|K^{opt}\|_{\ell_0}}{\|K^{opt}\|_{\ell_0}} \times 100 \%$ vs Penalizing
Parameter $\alpha_1$, The Sparsity Level $\|K^{opt}\|_{\ell_0} \times 100 \%$ vs Penalizing
Parameter $\alpha_1$ and The Performance Degradation $\frac{\|K^{opt}\|_{\ell_0}}{\|K^{opt}\|_{\ell_0}} \times 100 \%$
vs Sparsity Level $\|K^{opt}\|_{\ell_0} \times 100 \%$ corresponded to data set shown in
Table I for both CVX and ADMM methods.

$K_{j(j-1)} < \left(\frac{\sec(\frac{\pi}{n})}{n}\right)^n \prod_{i=1}^n (-A_{ii})\prod_{i=2,i \neq j}^n (A_{i(i-1)})(-A_{1n}) - A_{j(j-1)}$.

3) For $K_{1n}$, we get the following condition:

$-A_{1n} > K_{1n}$,

$K_{1n} > -\left(\frac{\sec(\frac{\pi}{n})}{n}\right)^n \prod_{i=1}^n (-A_{ii})\prod_{i=2,i \neq j}^n (A_{i(i-1)}) - A_{1n}$.

In our special case, one of the above 2$n$ conditions
implies that the matrix with one nonzero term $K_{(10)(10)} =
-9577.1955$ is one of the sparsest solutions. However, its
corresponding quadratic cost is 10182.2224 which is an
extremely large number. In other words, such a solution
gives an upper bound for the sparse solution corresponded
to the extremely large value of $\alpha_1$ which is equivalent to the
minimizing of $\|K\|_{\ell_1}$. It is very interesting that the sparsity
level for both CVX and ADMM is equal to 10%, however, it
is equal to 1% for our previously mentioned cyclic system.
We do such a substitution, because it must have an SDP get problem (17) where $X$. In problem (16) we substitute complement to part (16) to get problem (17). During the problems (14) to (16) to apply the second Schur complement to problems (13) and (16) to achieve is same as one which has been stated in [4]. We apply Schur Complement to problems (13) and (16) to achieve an SDP form in problems (14) and (17) respectively which makes them applicable to CVX. Although the first Schur complement is applied easily, some operations must be done during the problems (13) to (16) to apply the second Schur complement to part (16) to get problem (17). In problem (16) we substitute $X \succ 0$ by $X \succeq \epsilon_1 I$ to get problem (17) where $\epsilon_1$ is sufficiently small number. We do such a substitution, because it must have an SDP form in order to be applied to CVX. Also, in problem (16), $Y = (\delta + 1)P^T P$ has been bifurcated to $Y \succeq (\delta + 1)P^T P$ and $Y \preceq (\delta + 1)P^T P$ where the Schur complement is applied to the first one to get the problem (17). We can do such a bifurcation. Because, they are equivalent conditions.

ACKNOWLEDGEMENT

The Author is thankful to Dr. Sadegh Bolouki for his fruitful comments.

REFERENCES

[1] F. Lin, M. Fardad, and M. R. Jovanovic, “Sparse feedback synthesis via the alternating direction method of multipliers,” in American Control Conference (ACC), 2012. IEEE, 2012, pp. 4765–4770.
[2] ——, “Design of optimal sparse feedback gains via the alternating direction method of multipliers,” Automatic Control, IEEE Transactions on, vol. 58, no. 9, pp. 2426–2431, 2013.
[3] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” Information Theory, IEEE Transactions on, vol. 52, no. 12, pp. 5406–5425, 2006.
[4] M. Fardad and M. R. Jovanovic, “On the design of optimal structured and sparse feedback gains via sequential convex programming,” in American Control Conference (ACC), 2014. IEEE, 2014, pp. 2426–2431.
[5] K. Martensson and A. Rantzer, “A scalable method for continuous-time distributed control synthesis,” in American Control Conference (ACC), 2012. IEEE, 2012, pp. 6308–6313.
[6] H. J. Chizeck, A. S. Willsky, and D. Castanon, “Discrete-time markovian-jump linear quadratic optimal control,” International Journal of Control, vol. 43, no. 1, pp. 213–231, 1986.
[7] J. Lavaei, “Optimal decentralized control problem as a rank-constrained optimization.” in Allerton, 2013, pp. 39–45.
[8] G. Fazelnia, R. Madani, and J. Lavaei, “Convex relaxation for optimal distributed control problem,” in Conference on Decision and Control, 2014.
[9] Y.-S. Wang, N. Matni, and J. C. Doyle, “Localized lqr optimal control,” arXiv preprint arXiv:1409.6404, 2014.
[10] Y. Wang, J. Lopez, and M. Szmider, “Sparse static output feedback controller design via convex optimization,” in Conference on Decision and Control, 2014.
[11] F. Zhang, The Schur complement and its applications. Springer, 2006, vol. 4.
[12] Y. Wang, J. Lopez, and M. Sznaier, “Sparse static output feedback control and information structures,” in American Control Conference (ACC), 2012. IEEE, 2012, pp. 4985–4990.
[13] E. J. Candes and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies,” in Conference on Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on. IEEE, 2013, pp. 1520–1525.
[14] ——, “Sparsity and spatial localization measures for spatially distributed systems,” arXiv preprint arXiv:1402.4148v3, 2014.
[15] ——, “Sparsity measures for spatially decaying systems,” in American Control Conference (ACC), 2014. June 2014, pp. 5459–5464.
[16] J. J. Tyson and H. G. Othmer, “The dynamics of feedback control circuits in biochemical pathways,” Prog. Theor. Biol, vol. 5, no. 1, p. 62, 1978.
[17] C. Thron, “The secant condition for instability in biochemical feedback control, the role of cooperativity and saturaability,” Bulletin of Mathematical Biology, vol. 53, no. 3, pp. 383–401, 1991.
[18] M. R. Jovanovic, M. Arcak, and E. D. Sontag, “Remarks on the stability of spatially distributed systems with a cyclic interconnection structure,” in American Control Conference, 2007. ACC’07. IEEE, 2007, pp. 2696–2701.
[19] ——, “A passivity-based approach to stability of spatially distributed systems with a cyclic interconnection structure,” Automatic Control, IEEE Transactions on, vol. 53, no. Special Issue, pp. 75–86, 2008.
[20] M. Siami and N. Motiee, “Scaling laws for disturbance propagation in cyclic dynamical networks,” arXiv preprint arXiv:1502.00926, 2015.
[21] ——, “Robustness and performance analysis of cyclic interconnected dynamical networks,”
[22] K. B. Petersen and M. S. Pedersen, “The matrix cookbook.” Technical University of Denmark, pp. 7–15, 2008.
[23] R. A. Horn and C. R. Johnson, Matrix analysis. Cambridge university press, 2012.
[24] G. Dullerud and F. Paganini, Course in Robust Control Theory. Springer-Verlag New York, 2000.
[25] H. L. Royden, P. Fitzpatrick, and P. Hall, Real analysis. Macmillan New York, 1988, vol. 32.
[26] B. Recht, M. Fazel, and A. P. Parrilo, “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,” SIAM review, vol. 52, no. 3, pp. 471–501, 2010.
[27] S. Boyd and L. Vandenberghe, Convex optimization. Cambridge university press, 2009.
[28] B. Bamieh, F. Fagnani, and M. A. Dahleh, “Distributed control of spatially invariant systems,” Automatic Control, IEEE Transactions on, vol. 47, no. 7, pp. 1091–1107, 2002.
[29] M. U. Guide, “The mathworks,” Inc., Natick, MA, vol. 5, 1998.
[30] N. Motiee, A. Jadbabaie, and B. Bamieh, “On decentralized optimal control and information structures;” in American Control Conference, 2008. IEEE, 2008, pp. 4985–4990.