Sparsity-Promoting Optimal Control for a Class of Distributed Systems

Makan Fardad, Fu Lin, and Mihailo R. Jovanović

Abstract— We consider a linear quadratic optimal control problem with an additional penalty on the number of communication links in the distributed controller. We reformulate this combinatorial optimization problem as a sequence of weighted ℓ_1 problems, where the weighted ℓ_1 norm approximates the counting of the communication links. We identify a class of systems for which the weighted ℓ_1 problem can be formulated as a semidefinite program and therefore its solution can be computed efficiently. Application of the developed algorithm to the optimal control of vehicular formations reveals communication topologies that become sparser as the price of inter-vehicular communications is increased.

Index Terms— Communication architecture, convex optimization, ℓ_1 minimization, reweighting, semidefinite program, sparsity-promoting optimal control, vehicular formations.

I. INTRODUCTION

Traditional optimal control design does not take into account the price and feasibility of communication among system components. Consequently, the resulting optimal controllers are 'centralized', implying that the actuation signals are formed using information from all subsystems. As the size of the system becomes larger, such expectations become increasingly unrealistic.

The synthesis of distributed controllers for interconnected systems has received considerable attention in recent years [1]–[12]. Particular motivation to search for 'localized' controllers comes from [1], where it was shown that optimal controllers for spatially-invariant systems, although centralized, have an inherent spatial-decay property. For these systems, the dependence of the distributed controller on information coming from other parts of the system decays exponentially as one moves away from that controller. Similar results were generalized to systems over graphs in [10].

In this paper we consider an optimal control problem which, in addition to the standard performance requirements, penalizes the number of communication links. The controllers that result from this formulation are 'sparse' feedback matrices, where the level of sparsity is determined by the relative importance of performance versus communication cost in the objective function. We note the contrast between this framework and approaches in which the communication architecture, and thus the sparsity

M. Fardad is with the Department of Electrical Engineering and Computer Science, Syracuse University, NY 13244. F. Lin and M. R. Jovanović are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455. E-mails: makan@syr.edu, fu@umn.edu, mihailo@umn.edu.

structure, are specified *a priori*.

Counting the number of communication links leads to a combinatorial optimization problem, which is intractable in general. We therefore propose a meaningful relaxation of the problem using the ℓ_1 norm. For a special class of distributed systems, encountered in the control of vehicular formations, we demonstrate that the ℓ_1-relaxed problem can be formulated as a semidefinite program (SDP) and thus efficiently solved.

Our presentation is organized as follows. We formulate the sparsity-promoting optimal control problem in Section II. Motivated by the reweighted ℓ_1 algorithm [13], that is reviewed in Section III, we propose an algorithm for our optimal control problem in Section IV. For a class of distributed systems, we demonstrate in Section V that the optimal control problem can be formulated as a semidefinite program. An illustrative example for the control of vehicular formations is provided in Section VI. The paper is concluded with a summary of our contributions in Section VII.

II. MOTIVATION AND PROBLEM FORMULATION

Consider the following control problem

$$
\begin{aligned}
\dot{\psi} &= A\psi + B_1 d + B_2 u, \\
z &= C\psi + D u, \\
u &= -K\psi,
\end{aligned}
\tag{1}
$$

where $C = \begin{bmatrix} Q^{1/2} & 0 \end{bmatrix}^T$ and $D = \begin{bmatrix} 0 & R^{1/2} \end{bmatrix}^T$. The matrix K denotes a state feedback gain, and the closed-loop system is given by

$$
\begin{aligned}
\dot{\psi} &= (A - B_2 K)\psi + B_1 d, \\
z &= \begin{bmatrix} Q^{1/2} \\ -R^{1/2} K \end{bmatrix} \psi.
\end{aligned}
\tag{CL}
$$

Here, the input d denotes exogenous signals and the performance output z encapsulates both state and control penalties.

The design of the optimal state feedback gain K, subject to structural constraints that dictate its zero entries, was recently considered in [14], [15]. Let the subspace $\mathcal{S}$ embody these constraints and let us assume that there exists a stabilizing $K \in \mathcal{S}$. References [14], [15] then search for $K \in \mathcal{S}$ that minimizes the H_2 norm from d to z in (CL). Mathematically, the problem of H_2 norm minimization for the system in (CL) subject to $K \in \mathcal{S}$ can be reformulated as

$$
\begin{aligned}
\text{minimize} \quad & \text{trace}\left(P B_1 B_1^T\right) \\
\text{subject to} \quad & (A - B_2 K)^T P + P(A - B_2 K) \\
& = -(Q + K^T R K), \ K \in \mathcal{S}.
\end{aligned}
\tag{SH2}
$$

In the absence of structural constraints on K, problem (SH2) simplifies to the standard LQR problem.

In contrast to the optimal control problem (SH2), where the communication architecture of the controller is determined *a priori*, in the present work our emphasis is on identifying favorable communication structures without any prior assumptions on the sparsity patterns in the matrix K. We propose an optimization framework, using the ℓ_0 norm, in which the sparsity of the feedback gain is directly incorporated into the objective function. The ℓ_0 norm of a vector/matrix is equal to the number of nonzero elements of that vector/matrix.

We consider the following optimization problem

$$
\begin{aligned}
\text{minimize} \quad & \text{trace}\left(PB_1B_1^T\right) + \gamma \|K\|_{\ell_0} \\
\text{subject to} \quad & (A - B_2K)^T P + P(A - B_2K) \\
& = -(Q + K^T RK),
\end{aligned} \quad \text{(L0)}
$$

where $\|K\|_{\ell_0}$ denotes the *number of nonzero entries* of the matrix K. Thus, the incorporation of $\|K\|_{\ell_0}$ into the objective function promotes sparsity of the communication architecture. The positive scalar γ characterizes our emphasis on the sparsity of K; a larger γ encourages a sparser K, while $\gamma = 0$ renders a centralized gain that is the solution of the standard LQR problem. Via examples we demonstrate that as γ increases, the elements of K that correspond to long-range and less-critical links become identically zero.

III. ENHANCING SPARSITY BY REWEIGHTED ℓ_1 MINIMIZATION

Problem (L0), as formulated, is not convex both because of its nonconvex constraint and because the ℓ_0 norm is a nonconvex function of its argument. In fact, problem (L0) is a combinatorial optimization problem and thus intractable in general. In this section, we give a brief overview of the use of ℓ_1 minimization as a proxy for ℓ_0 minimization. The interested reader is referred to [13] for additional details.

The ℓ_1 norm, although nondifferentiable, is a convex function of its argument. In optimization problems where sparsity is desired, a common approach is to relax the ℓ_0 norm of the optimization variable to its ℓ_1 norm [13]. Recently, it was proved in [16] that for the specific problem of reconstructing a sparse signal from a small number of measurements, under certain conditions an optimization problem formulated using the ℓ_1 norm is capable of recovering the sparse signal exactly.

To better approximate the ℓ_0 norm, [13] uses a *weighted ℓ_1 norm*,

$$
\sum_i w_i \, |v_i|,
$$

where w_i are positive weights. The weighted ℓ_1 norm reduces to the regular ℓ_1 norm when $w_i = 1$ for all i. On the other hand, if the weights w_i are chosen to be inversely proportional to the magnitude of v_i,

$$
\begin{cases}
w_i = 1/|v_i|, & v_i \neq 0, \\
w_i = \infty, & v_i = 0,
\end{cases}
$$

then the weighted ℓ_1 norm of v and the ℓ_0 norm of v coincide,

$$
\sum_i w_i \, |v_i| \; = \; \|v\|_{\ell_0}.
$$

Now consider an optimization problem in which the objective function includes the ℓ_0 norm of the variable v,

$$
\text{minimize} \quad f(v) + \|v\|_{\ell_0}.
$$

Let v^* be a solution of this optimization problem. Then the solution of the problem

$$
\text{minimize} \quad f(v) + \sum_i w_i \, |v_i|, \quad \text{(RW)}
$$

with

$$
\begin{cases}
w_i = 1/|v_i^*|, & v_i^* \neq 0, \\
w_i = \infty, & v_i^* = 0,
\end{cases}
$$

has the following properties:
  (a) it has the same sparsity structure as v^*, and
  (b) for $v = v^*$ the value of the objective function in (RW) coincides with that of the original optimization problem.

To see why property (a) holds, suppose $v_j^* = 0$ for some j. Then $w_j = \infty$ and thus the solution of (RW) is forced to satisfy $v_j = 0$, otherwise the objective function incurs a value of infinity.

The above weighting scheme, however, cannot be implemented, since the weights themselves depend on the solution we seek. Reference [13] proposes an algorithm that attempts to produce these weights.

Reweighted ℓ_1 minimization algorithm
  1. Set the iteration count μ to zero and set $w_i^{(0)} = 1$, $i = 1, \ldots, N$. Choose the positive scalar ε sufficiently small.
  2. Solve the minimization problem (RW) to find the optimal solution $v^{(\mu)}$.
  3. Update the weights,
  $$
  w_i^{(\mu+1)} \; = \; \frac{1}{|v_i^{(\mu)}| + \varepsilon}, \quad i = 1, \ldots, N.
  $$
  4. Terminate on convergence. Otherwise, increment μ and go to Step 2.

The small positive number ε is introduced to ensure that the weights are well-defined when $v_i^{(\mu)} = 0$. It is also possible to decrease the value of ε at every iteration or select a different ε for different entries of the weighting matrix.

IV. SPARSITY-PROMOTING OPTIMAL CONTROL

Motivated by the reweighted minimization framework presented in the previous section, we propose an iterative algorithm for solving (L0). For any matrix $K = [k_{ij}]$, we define its ℓ_1 norm as $\|K\|_{\ell_1} = \sum_{i,j} |k_{ij}|$.

We now introduce the main algorithm of the paper.

Sparsity–promoting optimal control algorithm

1. Set the iteration count μ to zero and form the matrix $W^{(0)} = [w_{ij}^{(0)}]$ with its ijth entry given by

$$w_{ij}^{(0)} = 1.$$

   Choose the positive scalar $\varepsilon^{(0)}$ sufficiently small.

2. Solve the weighted ℓ_1 minimization problem

   minimize $\quad \text{trace}\left(PB_1 B_1^T\right) + \gamma \, \|W^{(\mu)} \circ K\|_{\ell_1}$

   subject to $\quad (A - B_2 K)^T P + P(A - B_2 K)$
   $$= -(Q + K^T R K),$$
   (WL1)

   to find the optimal solution $K^{(\mu)}$. The symbol $\circ$ denotes elementwise matrix multiplication.

3. Update ε,
   $$\varepsilon^{(\mu+1)} = \alpha \, \varepsilon^{(\mu)},$$

   with $0 < \alpha \le 1$.

4. Update the weights,
   $$w_{ij}^{(\mu+1)} = \frac{1}{|k_{ij}^{(\mu)}| + \varepsilon^{(\mu)}},$$

   where $k_{ij}^{(\mu)}$ is the ijth element of the matrix $K^{(\mu)}$. Form the matrix $W^{(\mu+1)} = [w_{ij}^{(\mu+1)}]$.

5. Terminate on convergence. Otherwise, increment μ and go to Step 2.

Even though the nonconvexity caused by the ℓ_0 norm in problem (L0) has been removed by using the reweighted ℓ_1 norm, the optimization problem (WL1) in Step 2 of the algorithm may still be nonconvex, due to its nonconvex constraint equation. Numerical optimization schemes can still be employed to obtain local optima of (WL1). In the next section we restrict our attention to a particular class of systems, encountered in the control of vehicular formations, for which (WL1) is indeed convex.

Finally, we point out that the reweighted minimization algorithm presented here can also be implemented using the ℓ_2 norm; see Appendix for details.

V. SDP Formulation for a Class of Distributed Systems

In this section, we identify a class of distributed systems for which the optimization problem (WL1) becomes an SDP and can therefore be solved efficiently. These systems are encountered in network consensus problems [17]–[19], control of vehicular formations [18], [20], and distributed averaging [21].

Consider a network of N single-integrator systems

$$\dot{x}_i = u_i + d_i, \qquad i = 1, \ldots, N, \tag{2}$$

where u_i and d_i are the control input and the disturbance acting on the ith system. Here,

$$u_i = -k_{ii}x_i - \sum_{j \ne i} k_{ij}(x_i - x_j),$$

implying that the control of the ith system is formed using its own state and a weighted average of the relative difference between its own state and the states of other systems. In matrix form, the dynamics of the network is described by

$$\dot{x} = -Kx + d,$$

where K is a state feedback gain. Furthermore, for bi-directional communication links we have $k_{ij} = k_{ji}$, and therefore the matrix K is symmetric.

Thus, we consider a dynamic network with its state-space representation given by (1) and

$$\begin{aligned} A = 0, \quad B_1 = I, \quad B_2 = I, \\ R = rI, \quad Q > 0, \\ K = K^T, \end{aligned} \tag{3}$$

where r is a positive scalar. Problem (WL1) then becomes

minimize $\quad \text{trace}(P) + \gamma \, \|W \circ K\|_{\ell_1}$

subject to $\quad KP + PK = Q + rKK.$

Since $A = 0$ and $K = K^T$, the stability of the closed-loop system is equivalent to K being a positive definite matrix. Multiplying both sides of the Lyapunov equation from the right by K^{-1} and using

$$\text{trace}\left(KPK^{-1}\right) = \text{trace}\left(PK^{-1}K\right) = \text{trace}(P),$$

we obtain

$$\text{trace}(P) = (1/2)\,\text{trace}\left(QK^{-1} + rK\right).$$

The optimization problem thus simplifies to

minimize $\quad J = (1/2)\,\text{trace}\left(QK^{-1} + rK\right) + \gamma \, \|W \circ K\|_{\ell_1}$

subject to $\quad K > 0.$
(J1)

The objective function in (J1) can be rewritten as

$$J = (1/2)\,\text{trace}\left(QK^{-1} + rK\right) \\ + \gamma \, \text{trace}\left(M(W \circ K \circ \text{sign}(K))\right),$$

where M is a matrix whose entries are all equal to one,

$$[M]_{ij} = 1,$$

and the ijth entry of the matrix $\text{sign}(K)$ is determined by

$$[\text{sign}(K)]_{ij} = \begin{cases} 1, & k_{ij} > 0, \\ 0, & k_{ij} = 0, \\ -1, & k_{ij} < 0. \end{cases}$$

Proposition 1: The optimization problem (J1) is equivalent to the semidefinite program

minimize $\quad \bar{J} = (1/2)\,\text{trace}\left(X + rK\right) + \gamma \, \text{trace}(MY)$

subject to $\quad K > 0$
$$\begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & K \end{bmatrix} \ge 0$$
$$-Y \preceq W \circ K \preceq Y,$$
(P1)

where $\preceq$ denotes *elementwise* matrix inequality, and the

optimization variables are the symmetric matrices K and X, and the matrix Y whose entries are nonnegative.

Proof: Using $K > 0$ and the Schur complement [22, Appendix 5.5], we have

$$\begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & K \end{bmatrix} \geq 0 \iff X \geq Q^{1/2} K^{-1} Q^{1/2},$$

and thus

$$\text{trace}\,(X) \geq \text{trace}\,\left(QK^{-1}\right).$$

The elementwise inequality

$$-Y_{ij} \leq W_{ij} K_{ij} \leq Y_{ij},$$

yields

$$W_{ij}|K_{ij}| \leq Y_{ij},$$

and thus

$$\text{trace}\,(MY) \geq \text{trace}\,(M(W \circ K \circ \text{sign}(K))).$$

Therefore, for any feasible point (K, X, Y) of (P1), we have

$$\bar{J}(K, X, Y) \geq J(K), \tag{4}$$

and thus minimum of $\bar{J}$ is an *upper* bound of the minimum of $J(K)$.

Let K^* be the global minimizer of (J1). Then, we can choose

$$X^* = Q^{1/2}\left(K^*\right)^{-1} Q^{1/2}$$

and

$$Y^* = W \circ K^* \circ \text{sign}(K^*)$$

which ensure that X^* and Y^* satisfy the constraints in (P1) and $\bar{J}(K^*, X^*, Y^*) = J(K^*)$. Let $(\bar{K}, \bar{X}, \bar{Y})$ be the global minimizer of $\bar{J}$. It follows that

$$\bar{J}(\bar{K}, \bar{X}, \bar{Y}) \leq \bar{J}(K^*, X^*, Y^*) = J(K^*)$$
$$\leq J(\bar{K}) \leq \bar{J}(\bar{K}, \bar{X}, \bar{Y}),$$

where the last inequality is from (4). Thus, the above inequalities are indeed all equalities. Since (P1) aims to minimize a linear function over a convex set [22], the global minimizer of (P1) is unique. Therefore, $\bar{K} = K^*$ and thus (P1) is equivalent to (J1). ∎

The SDP reformulation presented here corresponds only to Step 2 of the iterative algorithm introduced in Section IV. For the system (1)-(3), this amounts to replacing the optimization problem (WL1) with (P1), with the rest of the algorithm remaining the same. The SDP reformulation can be solved efficiently using available SDP solvers.

Finally, we note that a similar treatment can be developed for the case where the ℓ_1 norm is replaced by the ℓ_2 norm; see Appendix for details.

VI. VEHICULAR FORMATION EXAMPLE

In this section we apply the sparsity-promoting algorithm of Section IV to the optimal control of vehicular formations.

We consider the network of N single-integrator vehicles given by (2). We will henceforth refer to the 1st and the Nth vehicles as *leaders* and to the rest of the vehicles as *followers*. We assume that each of the followers uses the *relative* position errors between itself and all the other vehicles,

$$u_i = -\sum_{j=1, j \neq i}^{N} k_{ij}(x_i - x_j), \qquad i = 2, \ldots, N-1.$$

In addition, the 1st and the Nth vehicles have access to their own *absolute* position errors,

$$u_1 = -\sum_{j=2}^{N} k_{1j}(x_1 - x_j) - k_{11}x_1,$$

and

$$u_N = -\sum_{j=1}^{N-1} k_{Nj}(x_N - x_j) - k_{NN}x_N.$$

Let the relative feedback gains between two vehicles be equal,

$$k_{ij} = k_{ji}, \quad \text{for all} \ \ i \neq j.$$

Then, in matrix form we have

$$\dot{x} = u + d = -Kx + d,$$

where x, u, and d are the state, control, and disturbance vectors, respectively, and K is a symmetric matrix. For example, for $N = 3$,

$$K = \begin{bmatrix} k_{11} + k_{12} + k_{13} & -k_{12} & -k_{13} \\ -k_{12} & k_{12} + k_{23} & -k_{23} \\ -k_{13} & -k_{23} & k_{13} + k_{23} + k_{33} \end{bmatrix}.$$

Note that the state feedback gain K is a *structured* matrix that can be rewritten as

$$K = EK_r E^T + K_a$$

where $K_r \in \mathbb{R}^{m \times m}$ [with $m = N(N-1)/2$] and $K_a \in \mathbb{R}^{N \times N}$ are both diagonal matrices, and the incidence matrix E is defined as

$$E = \begin{bmatrix} e_{12} & e_{13} & \cdots & e_{(N-1)N} \end{bmatrix} \in \mathbb{R}^{N \times m},$$

where $e_{ij} \in \mathbb{R}^N$ takes 1 and -1 at the ith and jth entries, respectively, and 0 otherwise. The main diagonal of K_a is

$$\begin{bmatrix} k_{11} & 0 & \cdots & 0 & k_{NN} \end{bmatrix} \in \mathbb{R}^N,$$

and the main diagonal of K_r is

$$\{k_{ij}\} \ \ \text{for all} \ \ i \in \{1, \ldots, N\} \ \ \text{and} \ \ j \in \{i+1, \ldots, N\},$$

where i goes from 1 to N, and for a fixed i, j goes from $i+1$ to N. For example, for $N = 4$,

$$\begin{bmatrix} k_{12} & k_{13} & k_{14} & k_{23} & k_{24} & k_{34} \end{bmatrix}.$$

We consider the global performance measure that penal-

Fig. 1: Sparsity structures of K for $Q = Q_g$

(a) $\gamma = 0.1$ (b) $\gamma = 1$

(a) $\gamma = 0.001$ (b) $\gamma = 0.002$ (c) $\gamma = 0.004$

(d) $\gamma = 0.01$ (e) $\gamma = 0.02$ (f) $\gamma = 0.1$

Fig. 2: Sparsity structures of K for $Q = Q_l$

(a) $\gamma = 0.005$ (b) $\gamma = 0.01$ (c) $\gamma = 0.05$

(d) $\gamma = 0.1$ (e) $\gamma = 0.5$ (f) $\gamma = 1$
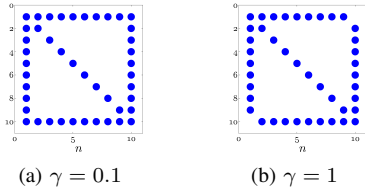
Fig. 3: Sparsity structures of K for $Q = Q_g + Q_l$

izes the *absolute* position errors of the vehicles,

$$x^T Q_g\, x = \sum_{i=1}^{N} x_i^2,$$

where the performance matrix is $Q_g = I$. We also consider the local performance measure that penalizes the *relative* position errors of the vehicles in addition to the absolute position errors of the 1st and the Nth vehicles,

$$x^T Q_l\, x = \sum_{i=1}^{N-1}(x_i - x_{i+1})^2 + x_1^2 + x_N^2.$$

Here, Q_l is a symmetric Toeplitz matrix whose first row is given by $[2\ -1\ 0 \dots 0] \in \mathbb{R}^N$. For example, for $N = 4$,

$$Q_l = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

For $Q = Q_g$ and small values of γ, the sparsity-promoting algorithm converges to a matrix K populated with non-zero elements. As γ increases, the followers *only* use information from the leaders; on the other hand, the leaders use relative information from all other vehicles (see Fig. 1a). For $\gamma \geq 1$, the sparsity structure of K is given in Fig. 1b, indicating that the leaders do not use relative position exchange between each other.

For $Q = Q_l$, as γ is increased the number of nonzero sub- and super-diagonals of K decreases and eventually K becomes a tridiagonal matrix (for $\gamma = 0.1$ and larger values); see Fig. 2. In other words, each vehicle interacts with a smaller number of neighbors as γ is increased, and the *nearest* neighbor interaction is selected by the algorithm for large values of γ.

Finally, the sparsity structures of K with $Q = Q_g + Q_l$ for different values of γ are illustrated in Fig. 3. For $\gamma = 1$ (and larger values), the nearest neighbor communication architecture is observed. We note *non-local* interaction patterns employed by the leaders for $\gamma = 0.5,\ 0.2,$ and 0.1. It turns out that the leaders communicate with the vehicles that are *uniformly* distributed along the formation. In particular, for $\gamma = 0.5$, the 1st vehicle exchanges information with the 7th vehicle; for $\gamma = 0.2$, it communicates with the 5th and the 7th vehicles; and for $\gamma = 0.1$, it uses information from all the *even* numbered vehicles.
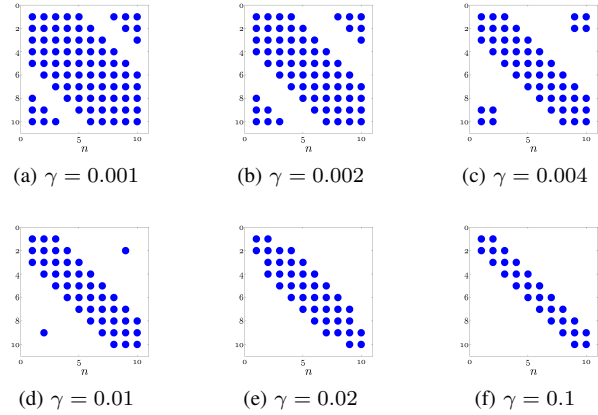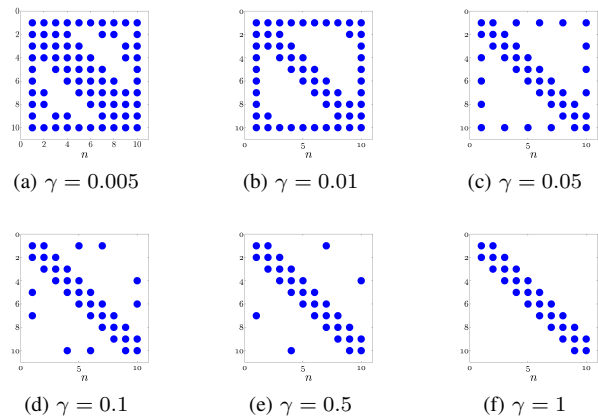
VII. Concluding Remarks

We consider the linear quadratic optimal control problem in which the objective function is modified to penalize the number of nonzero elements in the feedback gain. This promotes sparsity and translates to fewer communication links used by the distributed controller. We propose an iterative reweighted ℓ_1 algorithm, and for a class of distributed systems, we demonstrate that the algorithm amounts to solving a sequence of semidefinite programs. We apply this design method to the control of vehicular formations composed of single-integrator vehicles. A family of nontrivial communication patterns is revealed by our optimization algorithm as the price of communication is increased.

We are currently examining extension of the tools presented here to more general classes of distributed optimal control problems. One ongoing direction involves the *gradient projection* method proposed in [23], in conjunction with homotopy, to solve general nonconvex problems (WL1). Our preliminary results indicate the effectiveness of this approach in finding sparse feedback gains even for nonconvex problems.

APPENDIX

The sparsity-promoting optimal control algorithm in Section IV, and its restriction to systems described by (1)-(3), can also be investigated using the ℓ_2 norm.

We consider the optimization problem in Step 2 of the sparsity-promoting optimal control algorithm, and replace (WL1) with

$$
\begin{aligned}
\text{minimize} \quad & \text{trace}\left(PB_1B_1^T\right) + \gamma \left\|W^{(\mu)} \circ K\right\|_{\ell_2}^2 \\
\text{subject to} \quad & (A - B_2K)^T P + P(A - B_2K) \qquad \text{(WL2)} \\
& \qquad = -(Q + K^T R K),
\end{aligned}
$$

where the ℓ_2 norm of the matrix K is defined as $\|K\|_{\ell_2}^2 = \sum_{i,j} |k_{ij}|^2$. All other steps of the algorithm remain the same. Once restricted to the class of systems described by (1)-(3), problem (WL2) can be rewritten as

$$
\begin{aligned}
\text{minimize} \quad & (1/2)\,\text{trace}\left(QK^{-1} + rK\right) + \gamma \left\|W \circ K\right\|_{\ell_2}^2 \\
\text{subject to} \quad & K > 0.
\end{aligned}
$$

$$\text{(J2)}$$

Proposition 2: The optimization problem (J2) is equivalent to the semidefinite program

$$
\begin{aligned}
\text{minimize} \quad & \bar{J} = (1/2)\,\text{trace}\left(X + rK\right) + \gamma\,\text{trace}\left(Z\right) \\
\text{subject to} \quad & K > 0 \\
& \begin{bmatrix} X & Q^{1/2} \\ Q^{1/2} & K \end{bmatrix} \geq 0 \\
& \begin{bmatrix} Z & W \circ K \\ W \circ K & I \end{bmatrix} \geq 0,
\end{aligned}
$$

$$\text{(P2)}$$

where the optimization variables are the symmetric matrices K, X, and Z.

Proof: The proof is similar to the proof of Proposition 1 and thus omitted for brevity. ■

REFERENCES

[1] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 7, pp. 1091–1107, 2002.

[2] G. A. de Castro and F. Paganini, "Convex synthesis of localized controllers for spatially invariant system," *Automatica*, vol. 38, pp. 445–456, 2002.

[3] P. G. Voulgaris, G. Bianchini, and B. Bamieh, "Optimal H^2 controllers for spatially invariant systems with delayed communication requirements," *Systems and Control Letters*, vol. 50, pp. 347–361, 2003.

[4] R. D'Andrea and G. E. Dullerud, "Distributed control design for spatially interconnected systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1478–1495, 2003.

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

[6] B. Bamieh and P. G. 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.

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

[8] A. Rantzer, "Linear quadratic team theory revisited," in *Proceedings of the 2006 American Control Conference*, 2006, pp. 1637–1641.

[9] ——, "A separation principle for distributed control," in *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006, pp. 3609–3613.

[10] 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.

[11] F. Borrelli and T. Keviczky, "Distributed LQR design for identical dynamically decoupled systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 8, pp. 1901–1912, Sept. 2008.

[12] M. Fardad and M. R. Jovanović, "On the state-space design of optimal controllers for distributed systems with finite communication speed," in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 5488–5493.

[13] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 877–905, 2008.

[14] M. Fardad, F. Lin, and M. R. Jovanović, "On the optimal design of structured feedback gains for interconnected systems," in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009, pp. 978–983.

[15] F. Lin, M. Fardad, and M. R. Jovanović, "Augmented lagrangian approach to design of structured optimal state feedback gains," submitted to *IEEE Transactions on Automatic Control*, 2011.

[16] E. J. Candès, J. 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.

[17] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[18] B. Bamieh, M. R. Jovanović, P. Mitra, and S. Patterson, "Coherence in large-scale networks: dimension dependent limitations of local feedback," *IEEE Transactions on Automatic Control*, 2010, provisionally accepted.

[19] D. Zelazo and M. Mesbahi, "Edge agreement: Graph-theoretic performance bounds and passivity analysis," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 544–555, 2011.

[20] F. Lin, M. Fardad, and M. R. Jovanović, "Optimal control of vehicular formations with nearest neighbor interactions," submitted to *IEEE Transactions on Automatic Control*, 2010.

[21] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[23] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.