2012 American Control Conference
Fairmont Queen Elizabeth, Montréal, Canada
June 27-June 29, 2012

Sparse feedback synthesis via the alternating direction method of multipliers

Fu Lin, Makan Fardad, and Mihailo R. Jovanović

***Abstract*— We study the design of feedback gains that strike a balance between the $\mathcal{H}_2$ performance of distributed systems and the sparsity of controller. Our approach consists of two steps. First, we identify sparsity patterns of feedback gains by incorporating sparsity-promoting penalty functions into the $\mathcal{H}_2$ problem, where the added terms penalize the number of communication links in the distributed controller. Second, we optimize feedback gains subject to structural constraints determined by the identified sparsity patterns. In the first step, we identify sparsity structure of feedback gains using the alternating direction method of multipliers, which is a powerful algorithm well-suited to large optimization problems. This method alternates between optimizing the sparsity and optimizing the closed-loop $\mathcal{H}_2$ norm, which allows us to exploit the structure of the corresponding objective functions. In particular, we take advantage of the separability of sparsity-promoting penalty functions to decompose the minimization problem into sub-problems that can be solved analytically. An example is provided to illustrate the effectiveness of the developed approach.**

***Index Terms*— Alternating direction method of multipliers, cardinality minimization, communication architectures, distributed systems, homotopy, ℓ_1 minimization, sum-of-logs penalty, sparsity-promoting optimal control.**

I. Introduction

The design of distributed controllers for interconnected systems has received considerable attention in recent years [1]–[12]. Research efforts have focused on two major issues, namely, the design of communication architectures of distributed controllers and the design of optimal controllers under *a priori* specified structural constraints.

In this paper, we develop methods for the design of *sparse* optimal $\mathcal{H}_2$ feedback gains. Our approach consists of two steps. The first step, which can be viewed as a *structure identification step*, is aimed at finding sparsity patterns $\mathcal{S}$ that strike a balance between the $\mathcal{H}_2$ performance and the sparsity of controller. This is achieved by incorporating *sparsity-promoting* penalty functions into the optimal control problem, where the added sparsity-promoting terms penalize the number of communication links. Second, we solve an optimal control problem subject to the feedback gain belonging to the identified structure $\mathcal{S}$. This *polishing step* improves the $\mathcal{H}_2$ performance of structured controllers.

The main contributions of our paper can be summarized as follows. First, we solve the sparsity-promoting optimal control problem for general linear time-invariant systems. As a consequence, our approach accommodates homogeneous or heterogenous subsystems with coupled or decoupled dynamics in networks that have directed or undirected communication links.

Second, we demonstrate that the *alternating direction method of multipliers* (ADMM) [13] provides an effective tool for the design of optimal distributed controllers. This method alternates between optimizing the sparsity of the feedback matrix and optimizing the closed-loop $\mathcal{H}_2$ norm. The advantage of this alternating mechanism is threefold.

- It provides a flexible framework for incorporation of different penalty functions to promote sparsity.
- It allows us to exploit the *separability* of the sparsity-promoting penalty functions and to *decompose* the corresponding optimization problems into sub-problems that can be solved *analytically*.
- It facilitates the use of descent algorithms for the $\mathcal{H}_2$ optimization, in which a descent direction can be formed by solving two Lyapunov equations and one Sylvester equation.

Finally, we use several sparsity-promoting penalty functions including weighted ℓ_1 norm, nonconvex sum-of-logs, and cardinality functions; these have advantageous sparsity-promoting properties compared to the widely used ℓ_1 norm [14]. For all these penalty functions, analytical expressions for solution of the corresponding optimization problem in ADMM can be obtained. Furthermore, these analytical results are independent of the $\mathcal{H}_2$ norm objective. Therefore, they can be utilized in ADMM to design sparse feedback gains that are optimal with respect to other control objectives.

Our approach is motivated in part by the emerging field of compressive sensing. In controls community, recent work inspired by similar ideas includes [15], [16]. In [16], a cardinality induced gain was introduced to quantify the sparsity of the impulse response of a discrete-time system. In [15], the weighted ℓ_1 framework was used to design structured dynamic output feedback controllers subject to a given $\mathcal{H}_\infty$ performance.

Our presentation is organized as follows. We formulate the sparsity-promoting optimal control problem and compare several sparsity-promoting penalty functions in Section II. We present the ADMM algorithm, emphasize the separability of penalty functions, and provide analytical solutions to the sub-problems of ADMM in Section III. We demonstrate the effectiveness of the developed approach using a mass-spring system example in Section IV. We conclude the paper with a summary of our contributions in Section V.

II. Sparsity-promoting Optimal Control Problem

Consider the following control problem

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

Financial support from the National Science Foundation under awards CMMI-0927509 and CMMI-0927720 and under CAREER Award CMMI-0644793 is gratefully acknowledged.

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.

where d is the exogenous input signal, z is the performance output, $C = \left[\, Q^{1/2} \;\; 0 \,\right]^T$, and $D = \left[\, 0 \;\; R^{1/2} \,\right]^T$. The matrix F is a state feedback gain, $Q = Q^T \geq 0$ and $R = R^T > 0$ are the state and control performance weights, and the closed-loop system is given by

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

The design of the optimal state feedback gain F, subject to structural constraints that dictate its zero entries, was recently considered by the authors in [9], [12]. Let the subspace $\mathcal{S}$ embody these constraints and let us assume that there exists a stabilizing $F \in \mathcal{S}$. References [9], [12] search for $F \in \mathcal{S}$ that minimizes the $\mathcal{H}_2$ norm of the transfer function from d to z. Since for stabilizing F, the closed-loop $\mathcal{H}_2$ norm

$$J(F) \,=\, \mathrm{trace}\left(B_1^T \int_0^\infty \mathrm{e}^{(A - B_2 F)^T t} (Q + F^T R F) \,\times\, \mathrm{e}^{(A - B_2 F) t} \,\mathrm{d}t\, B_1\right) \tag{2}$$

can be obtained from the solution of the Lyapunov equation

$$(A - B_2 F)^T P \,+\, P (A - B_2 F) \,=\, -\left(Q + F^T R F\right),$$

the $\mathcal{H}_2$ problem subject to structural constraint on the feedback matrix F can be formulated as

$$\begin{array}{ll} \text{minimize} & J(F) \,=\, \mathrm{trace}\left(B_1^T P(F) B_1\right) \\ \text{subject to} & F \in \mathcal{S}. \end{array} \tag{SH2}$$

In the absence of the constraint $F \in \mathcal{S}$, problem (SH2) simplifies to the standard LQR problem.

Note that the communication architecture of the controller is *a priori* specified in (SH2). In contrast, in this paper our emphasis shifts to identifying favorable communication structures without any prior assumptions on the sparsity patterns of matrix F. We propose an optimization framework in which the sparsity of feedback gain is directly incorporated into the objective function.

Consider the following optimization problem

$$\text{minimize} \;\; J(F) \,+\, \gamma\, g_0(F), \tag{3}$$

where

$$g_0(F) \,=\, \mathbf{card}\,(F) \tag{4}$$

denotes the cardinality function, i.e., the *number of nonzero elements* of a matrix. Note that, in contrast to problem (SH2), no structural constraint is imposed on F; instead, our goal is to promote sparsity of the feedback gain by incorporating cardinality function into the optimization problem. The positive scalar γ characterizes our emphasis on the sparsity of F; a larger γ encourages a sparser F, while $\gamma = 0$ renders a centralized gain that is the solution of the standard LQR problem. For $\gamma = 0$, the solution to (3) is given by $F_c = R^{-1} B_2^T P$, where P is the unique positive definite solution of the algebraic Riccati equation

$$A^T P \,+\, P A \,+\, Q \,-\, P B_2 R^{-1} B_2^T P \,=\, 0. \tag{5}$$

A. Sparsity-promoting penalty functions

Problem (3) is a combinatorial optimization problem whose solution usually requires an intractable combinatorial search. In optimization problems where sparsity is desired, cardinality function is typically replaced by the ℓ_1 norm of the optimization variable [17, Chapter 6],

$$g_1(F) \,=\, \|F\|_{\ell_1} \,=\, \sum_{i,j} |F_{ij}|. \tag{6}$$

Recently, a *weighted* ℓ_1 norm was used to enhance sparsity in signal recovery [14],

$$g_2(F) \,=\, \sum_{i,j} W_{ij} |F_{ij}|, \tag{7}$$

where $W_{ij} \in \mathbb{R}$ are positive weights. Weighted ℓ_1 norm tries to bridge the difference between ℓ_1 norm and cardinality function. In contrast to cardinality function that assigns the *same* cost to any nonzero element, ℓ_1 norm penalizes more heavily elements of larger magnitudes. The positive weights can be chosen to counteract this magnitude dependence of ℓ_1 norm. For example, if W_{ij} is chosen to be *inversely proportional* to the magnitude of F_{ij},

$$\left\{ \begin{array}{ll} W_{ij} = 1/|F_{ij}|, & F_{ij} \neq 0, \\ W_{ij} = \infty, & F_{ij} = 0, \end{array} \right.$$

then weighted ℓ_1 norm and cardinality function of F coincide,

$$\sum_{i,j} W_{ij}\, |F_{ij}| \,=\, \mathbf{card}\,(F)\,.$$

The above scheme for weights, however, cannot be implemented, since weights depend on the unknown feedback gain. A reweighted algorithm that solves a sequence of weighted ℓ_1 optimization problems in which the weights are determined by the solution of weighted ℓ_1 problem in the previous iteration was proposed in [14]. This reweighted scheme was recently employed by the authors to design sparse feedback gains for a class of distributed systems [18].

Both ℓ_1 norm and its weighted version are *convex* relaxations of cardinality function. We also consider *nonconvex* alternatives that could be more aggressive in promoting sparsity. Suppose that we wish to find the sparsest feedback gain that provides a given level of $\mathcal{H}_2$ performance $\sigma > 0$,

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\,(F) \\ \text{subject to} & J(F) \,\leq\, \sigma. \end{array}$$

Approximating $\mathbf{card}\,(F)$ with a penalty function yields

$$\begin{array}{ll} \text{minimize} & g(F) \\ \text{subject to} & J(F) \,\leq\, \sigma. \end{array} \tag{8}$$

Solution to (8) is the intersection of the constraint set $\mathcal{C} = \{F \,|\, J(F) \,\leq\, \sigma\}$ and the smallest sub-level set of g that touches $\mathcal{C}$; see Fig. 1. In contrast to ℓ_1 norm whose sub-level sets are determined by the *convex* ℓ_1 ball, the sub-level sets of the *nonconvex* function (e.g., ℓ_p norm with $0 < p < 1$) have a star-like shape; see Fig. 1d. The sum-of-logs function is another example of nonconvex functions with similar geometry of sub-level sets,

$$g_3(F) \,=\, \sum_{i,j} \log\left(1 \,+\, \frac{|F_{ij}|}{\varepsilon}\right), \quad 0 \,<\, \varepsilon \,\ll\, 1. \tag{9}$$

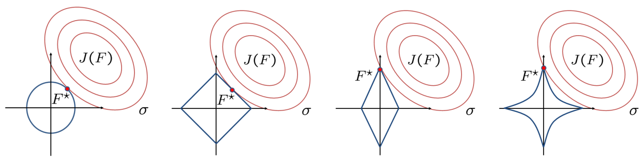


Fig. 1: Solution $F^\star$ of the constrained problem (8) is the intersection of constraint set $\mathcal{C} = \{F \,|\, J(F) \leq \sigma\}$ and the smallest sub-level set of g that touches $\mathcal{C}$. The penalty function g is (a) ℓ_2 norm (i.e., Frobenius norm), (b) ℓ_1 norm, (c) weighted ℓ_1 norm with appropriate weights, and (d) nonconvex function such as ℓ_p norm with $0 < p < 1$ or the sum-of-logs function (9).

B. *Sparsity-promoting optimal control problem*

Our approach to sparse feedback synthesis makes use of the above discussed penalty functions. In order to obtain sparse state feedback gains that yield satisfactory $\mathcal{H}_2$ performance, we consider the following optimal control problem

$$\text{minimize} \;\; J(F) \,+\, \gamma\, g(F) \tag{SP}$$

where J is the closed-loop $\mathcal{H}_2$ norm (2) and g is a sparsity-promoting penalty function, e.g., given by (4), (6), (7), or (9). When cardinality function (4) is replaced by (6), (7), or (9), problem (SP) can be viewed as a relaxation of the combinatorial problem (3)-(4), obtained by approximating cardinality function with the corresponding penalty function.

As parameter γ varies over $[0, +\infty)$, the solution of (SP) traces the optimal trade-off path between $\mathcal{H}_2$ performance J and feedback gain sparsity g. When $\gamma = 0$, the solution is the centralized feedback gain, which can be computed from the solution of the algebraic Riccati equation (5). We then slightly increase γ and employ an iterative algorithm – the alternating direction method of multipliers – initialized by the optimal feedback matrix at the previous γ. The solution of (SP) becomes sparser as γ increases. After a desired level of sparsity is achieved, we fix the sparsity structure and find the optimal structured feedback gain by solving the structured $\mathcal{H}_2$ problem (SH2).

Remark 1: We employ Newton's method in conjunction with conjugate gradient scheme to solve the structured $\mathcal{H}_2$ problem (SH2). Due to space limitation, however, we refer the reader to [19] for a detailed discussion of this approach.

III. Identification of Sparsity-patterns via the Alternating Direction Method of Multipliers

The alternating direction method of multipliers has been studied extensively since the 1970s. This simple but powerful algorithm blends the *separability* of the dual decomposition with the superior convergence of the method of multipliers. Reference [13] provides an excellent survey of ADMM with emphasis on its application to large-scale distributed optimization problems. ADMM has been used in a wide range of applications including sparse signal recovery, image restoration and denoising, and sparse inverse covariance selection; see [13] and the references therein.

Consider the following constrained optimization problem

$$\begin{array}{ll} \text{minimize} & J(F) \,+\, \gamma\, g(G) \\ \text{subject to} & F \,-\, G \,=\, 0, \end{array} \tag{10}$$

which is clearly equivalent to (SP). The augmented Lagrangian [20] associated with the constrained problem (10) is given by

$$\begin{aligned} \mathcal{L}_\rho(F, G, \Lambda) \;&=\; J(F) \,+\, \gamma\, g(G) \,+\, \text{trace}\left(\Lambda^T (F \,-\, G)\right) \\ &+\; \frac{\rho}{2} \| F \,-\, G \|_F^2, \end{aligned}$$

where Λ is the dual variable (i.e., the *Lagrange multiplier*), ρ is a positive scalar, and $\|\cdot\|_F$ is the Frobenius norm. It might appear that we have complicated the problem by introducing an additional variable G and an additional constraint $F - G = 0$. By doing this, however, we have in effect simplified (SP) by decoupling the objective function into two parts that depend on two different variables. As discussed below, this allows us to exploit structures of J and g.

In order to find a minimizer of the constrained problem (10), the ADMM algorithm uses a sequence of iterations

$$F^{k+1} \;:=\; \arg\min_{F} \, \mathcal{L}_\rho(F, G^k, \Lambda^k) \tag{11a}$$

$$G^{k+1} \;:=\; \arg\min_{G} \, \mathcal{L}_\rho(F^{k+1}, G, \Lambda^k) \tag{11b}$$

$$\Lambda^{k+1} \;:=\; \Lambda^k \,+\, \rho(F^{k+1} \,-\, G^{k+1}), \tag{11c}$$

until

$$\|F^{k+1} \,-\, G^{k+1}\|_F \,\leq\, \epsilon \;\; \text{and} \;\; \|G^{k+1} \,-\, G^k\|_F \,\leq\, \epsilon.$$

In contrast to the *method of multipliers* [20], in which F and G are *minimized jointly*

$$(F^{k+1}, G^{k+1}) \;:=\; \arg\min_{F,\, G} \, \mathcal{L}_\rho(F, G, \Lambda^k),$$

ADMM consists of an F-minimization step (11a), a G-minimization step (11b), and a dual variable update step (11c). Thus, the optimal F and G are determined in an alternating fashion, which motivates the name *alternating direction*. Note that the dual variable update (11c) uses a step-size equal to ρ, which guarantees the dual feasibility of (G^{k+1}, Λ^{k+1}) in each ADMM iteration [13].

ADMM brings two major benefits to the sparsity-promoting optimal control problem (SP):

- *Separability of g.* The penalty function g is *separable* with respect to *individual* elements of a matrix. In contrast, the closed-loop $\mathcal{H}_2$ norm cannot be decomposed into componentwise functions of the feedback gain. By separating g and J in the minimization of the augmented Lagrangian $\mathcal{L}_\rho$, we can decompose G-minimization problem (11b) into sub-problems that only involve *scalar* variables. This allows us to determine *analytically* the solution of (11b) for different penalty functions including weighted ℓ_1 norm, nonconvex sum-of-logs function, and even cardinality function.
- *Differentiability of J.* The closed-loop $\mathcal{H}_2$ norm J is a *differentiable* function of the feedback gain matrix [12]; this is in contrast to g which is *non-differentiable*. By separating g and J in the minimization of the augmented Lagrangian $\mathcal{L}_\rho$, we can utilize descent al-

gorithms that rely on the differentiability of J to solve F-minimization problem (11a). To find the minimizer of (11a), we employ one such algorithm that alternates between solving two Lyapunov equations and a Sylvester equation.

In Section III-A, we derive analytical expressions for solutions of G-minimization problem (11b). In Section III-B, we describe the Anderson-Moore method to solve F-minimization problem (11a).

A. Separable solution to the G-minimization problem (11b)

Completion of squares with respect to G in the augmented Lagrangian $\mathcal{L}_\rho$ can be used to show that G-minimization problem (11b) is equivalent to

$$\text{minimize} \quad \phi(G) \,=\, \gamma\, g(G) \,+\, (\rho/2)\|G \,-\, V^k\|_F^2, \qquad (12)$$

where

$$V^k \,=\, (1/\rho)\Lambda^k \,+\, F^{k+1}.$$

To simplify notation, we drop the superscript in V^k throughout this section. Since both g and the square of Frobenius norm can be written as a summation of componentwise functions of a matrix, we can decompose (12) into sub-problems expressed in terms of *individual* elements of G. For example, if g is ℓ_1 norm, then

$$\phi(G) \,=\, \sum_{i,j} \Big(\gamma |G_{ij}| + (\rho/2)(G_{ij} - V_{ij})^2 \Big).$$

This facilitates the conversion of (12) to minimization problems that only involve *scalar* variables. By doing so, we can determine *analytically* the solution of (12) for different penalty functions including weighted ℓ_1 norm, sum-of-logs function, and cardinality function.

1) Weighted ℓ_1 norm: In this case, the objective function in (12) is a summation of strictly convex functions

$$\phi(G) \,=\, \sum_{i,j} \Big(\gamma W_{ij} |G_{ij}| \,+\, (\rho/2)(G_{ij} \,-\, V_{ij})^2 \Big).$$

Therefore, problem (12) is decomposed into sub-problems,

$$\text{minimize} \quad \phi_{ij}(G_{ij}) \,=\, \gamma W_{ij} |G_{ij}| \,+\, (\rho/2)(G_{ij} \,-\, V_{ij})^2$$

whose unique solution is given by the *shrinkage* operator (e.g., see [13, Section 4.4.3])

$$G_{ij}^\star \,=\, \begin{cases} V_{ij} \,-\, a, & V_{ij} \,\in\, (a,\, +\infty) \\ 0, & V_{ij} \,\in\, [-a,\, a] \\ V_{ij} \,+\, a, & V_{ij} \,\in\, (-\infty,\, -a), \end{cases} \qquad (13)$$

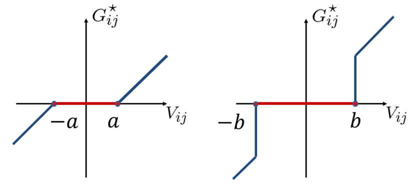
where $a \,=\, (\gamma/\rho)W_{ij}$; see Fig. 2a. For given V_{ij}, $G_{ij}^\star$ is obtained by moving V_{ij} towards zero with the amount $(\gamma/\rho)W_{ij}$; in particular, $G_{ij}^\star$ is set to zero if $|V_{ij}| \,\leq\, (\gamma/\rho)W_{ij}$. Therefore, a more aggressive scheme for driving $G_{ij}^\star$ to zero can be obtained by increasing γ and W_{ij} and by decreasing ρ.

2) Cardinality function: In this case, problem (12) is decomposed into sub-problems,

$$\text{minimize} \quad \phi_{ij}(G_{ij}) \,=\, \gamma\, \mathbf{card}\,(G_{ij}) \,+\, (\rho/2)(G_{ij} \,-\, V_{ij})^2$$

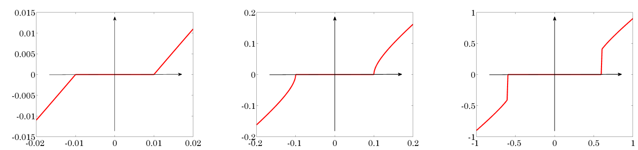
whose unique solution is given by the *truncation* operator [19]

$$G_{ij}^\star \,=\, \begin{cases} 0, & |V_{ij}| \,\leq\, b \\ V_{ij}, & |V_{ij}| \,>\, b, \end{cases} \qquad (14)$$



(a) (b)

Fig. 2: (a) Shrinkage operator (13) with $a = (\gamma/\rho)W_{ij}$; (b) truncation operator (14) with $b = \sqrt{2\gamma/\rho}$. The slope of lines in both (a) and (b) is equal to one.



(a) $\gamma = 0.1$ (b) $\gamma = 1$ (c) $\gamma = 10$

Fig. 3: Characteristics of operator (15) with $\{\rho = 100,\, \varepsilon = 0.1\}$ for different γ values. For $\gamma \,=\, 0.1$, (15) resembles the shrinkage operator (13) in Fig. 2a; for $\gamma \,=\, 10$, (15) resembles the truncation operator (14) in Fig. 2b; for $\gamma = 1$, (15) bridges the difference between (13) and (14).

where $b = \sqrt{2\gamma/\rho}$; see Fig. 2b. For given V_{ij}, $G_{ij}^\star$ is set to V_{ij} if $|V_{ij}| > \sqrt{2\gamma/\rho}$ and to zero if $|V_{ij}| \leq \sqrt{2\gamma/\rho}$.

3) Sum-of-logs function: In this case, problem (12) is decomposed into sub-problems,

$$\text{minimize} \quad \phi_{ij}(G_{ij}) = \gamma \log\left(1 + \frac{|G_{ij}|}{\varepsilon}\right) + \frac{\rho}{2}(G_{ij} - V_{ij})^2$$

whose solution is given by

$$G_{ij}^\star \,=\, \begin{cases} G_p^+, & V_{ij} \,\in\, (c, +\infty) \\ G_p^0, & V_{ij} \,\in\, [0, c] \\ G_m^0, & V_{ij} \,\in\, [-c, 0) \\ G_m^-, & V_{ij} \,\in\, (-\infty, -c), \end{cases} \qquad (15)$$

where $c = (\gamma/\rho)\varepsilon^{-1}$ and

$$G_p^0 \,:=\, \arg\min \{\phi_{ij}(G_p^+), \phi_{ij}(G_p^-), \phi_{ij}(0)\}$$
$$G_m^0 \,:=\, \arg\min \{\phi_{ij}(G_m^+), \phi_{ij}(G_m^-), \phi_{ij}(0)\}.$$

Since we need to compare the value of ϕ_{ij} at three points $\{G_p^+,\, G_p^-,\, 0\}$ for $V_{ij} \,\in\, [0, c]$ and at another three points $\{G_m^+,\, G_m^-,\, 0\}$ for $V_{ij} \,\in\, [-c, 0)$ to determine the solution $G_{ij}^\star$, operator (15) is more complex than shrinkage and truncation operators. Here,

$$G_p^\pm \,=\, \frac{1}{2}\left(V_{ij} \,-\, \varepsilon \,\pm\, \sqrt{(V_{ij} + \varepsilon)^2 - 4(\gamma/\rho)}\right)$$
$$G_m^\pm \,=\, \frac{1}{2}\left(V_{ij} \,+\, \varepsilon \,\pm\, \sqrt{(V_{ij} - \varepsilon)^2 - 4(\gamma/\rho)}\right)$$

are solutions of quadratic equations; see [19].

For fixed ρ and ε, operator (15) is determined by the value of γ; see Fig. 3. For small γ values, operator (15) resembles the shrinkage operator (cf. Fig. 3a and Fig. 2a) and for large

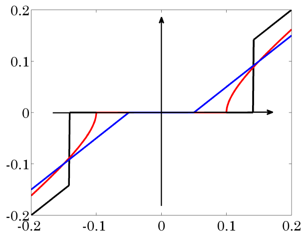


Fig. 4: A comparison between shrinkage operator (13) shown in blue, truncation operator (14) shown in black, and operator (15) shown in red, for $\{\gamma = 1, \rho = 100, W_{ij} = 5, \varepsilon = 0.1\}$. These operators (13), (14), and (15) correspond to weighted ℓ_1 norm, cardinality function, and sum-of-logs function, respectively.

γ values, it resembles the truncation operator (cf. Fig. 3c and Fig. 2b). In other words, operator (15) can be viewed as an intermediate step between shrinkage and truncation operators. For example, for $\{\gamma = 1, \rho = 100, W_{ij} = 5, \varepsilon = 0.1\}$, operator (15) has a bigger (resp. smaller) dead-zone interval compared to the shrinkage (resp. truncation) operator; see Fig. 4.

B. Anderson-Moore method for the F-minimization problem (11a)

We next employ the Anderson-Moore method to solve the F-minimization problem (11a). The advantage of this algorithm lies in its fast convergence (compared to the gradient method) and in its simple implementation (compared to Newton's method); e.g., see [21]. When applied to the F-minimization problem (11a), this method requires solutions of two Lyapunov equations and one Sylvester equation in each iteration.

By completing squares with respect to F in the augmented Lagrangian $\mathcal{L}_\rho$, we obtain the following equivalent problem

$$\text{minimize} \quad \varphi(F) \,=\, J(F) \,+\, (\rho/2)\|F \,-\, U^k\|_F^2,$$

where

$$U^k \,=\, G^k \,-\, (1/\rho)\Lambda^k.$$

Using standard techniques [12], [21], we obtain the necessary conditions for optimality

$$(A - B_2 F)L \,+\, L(A - B_2 F)^T \,=\, -\, B_1 B_1^T \qquad (16a)$$

$$(A - B_2 F)^T P \,+\, P(A - B_2 F) \,=\, -\,(Q + F^T R F) \qquad (16b)$$

$$\nabla\varphi(F) \,=\, 2RFL \,+\, \rho F \,-\, 2B_2^T P L \,-\, \rho U^k \,=\, 0. \qquad (16c)$$

Starting with a stabilizing feedback F, the Anderson-Moore method solves the Lyapunov equations (16a) and (16b), and then solves the Sylvester equation (16c) to obtain a new feedback gain $\bar{F}$. In other words, it alternates between solving (16a) and (16b) for L and P with F being fixed and solving (16c) for F with L and P being fixed. It can be shown that [19] the difference between two consecutive steps $\tilde{F} = \bar{F} - F$ forms a *descent direction* of $\varphi(F)$. As a consequence, standard step-size rules (e.g., Armijo rule [20, Section 1.2]) can be employed to determine s in $F + s\tilde{F}$ to guarantee the convergence to a stationary point of φ. Since φ is locally convex for sufficiently large ρ [19], the stationary point provides a local minimum of φ.

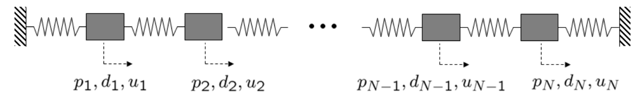


Fig. 5: Mass-spring system.

γ	0.04	0.27	1.00
card $(F^\star)/$**card** (F_c)	9.60%	3.92%	1.92%
$(J(F^\star) \,-\, J(F_c))/J(F_c)$	0.73%	4.14%	7.97%

TABLE I: Sparsity vs. performance for mass-spring system. Using only about 2% of nonzero elements, $\mathcal{H}_2$ performance of $F^\star$ is only about 8% worse than performance of the centralized gain F_c.

IV. Mass-spring System Example

We illustrate the utility of the developed approach using a mass-spring system example. Additional examples along with MATLAB source codes can be found at

www.ece.umn.edu/∼mihailo/software/lqrsp/

Consider a mass-spring system with N masses shown in Fig. 5. Let p_i be the displacement of the ith mass from its reference position and let the state variables be $\psi_1 := [\, p_1 \cdots p_N \,]$ and $\psi_2 := [\, \dot{p}_1 \cdots \dot{p}_N \,]$. For simplicity we consider unit masses and spring constant; note that our method can be used to design controllers for arbitrary values of these parameters. The state-space representation is then given by (1) with

$$A \,=\, \begin{bmatrix} O & I \\ T & O \end{bmatrix}, \quad B_1 \,=\, B_2 \,=\, \begin{bmatrix} O \\ I \end{bmatrix},$$

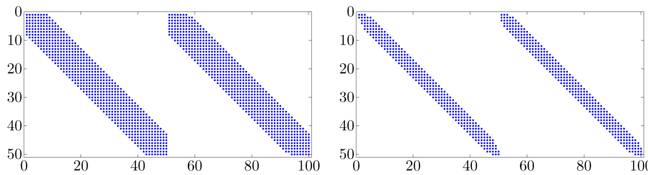
where T is an $N \times N$ symmetric tridiagonal matrix with -2 on its main diagonal and 1 on its first sub- and super-diagonal, and I and O are $N \times N$ identity and zero matrices. The state performance weight Q is the identity matrix and the control performance weight is $R = 10I$.

We use cardinality function (4) to promote sparsity. As γ increases, the number of nonzero sub- and super-diagonals of both position $F_p^\star$ and velocity $F_v^\star$ gains decreases; see Fig. 6. Eventually, both $F_p^\star$ and $F_v^\star$ become diagonal matrices. It is noteworthy that diagonals of both position and velocity feedback gains are nearly constant except for masses that are close to the boundary; see Fig. 7.

After sparsity structures of controllers are identified by solving (SP), we fix sparsity patterns and solve structured $\mathcal{H}_2$ problem (SH2) to obtain the optimal *structured* controllers. Comparing the sparsity level and the performance of these controllers to those of the centralized controller F_c, we see that using only a *fraction* of nonzero elements, the sparse feedback gain $F^\star$ achieves $\mathcal{H}_2$ performance comparable to the performance of F_c; see Fig. 8. In particular, using about 2% of nonzero elements, $\mathcal{H}_2$ performance of $F^\star$ is only about 8% worse than performance of F_c; see Table I.

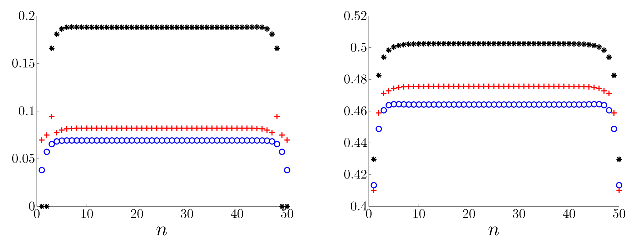
V. Concluding Remarks

We have designed sparse feedback gains that optimize the $\mathcal{H}_2$ performance of distributed systems. The ADMM



(a) $\gamma = 10^{-4}$ (b) $\gamma = 0.0233$

Fig. 6: Sparsity patterns of $F^\star = [F_p^\star \;\; F_v^\star] \in \mathbb{R}^{50\times 100}$ for the mass-spring system obtained using cardinality function to promote sparsity. As γ increases, the number of nonzero sub- and super-diagonals of $F_p^\star$ and $F_v^\star$ decreases.



(a) (b)

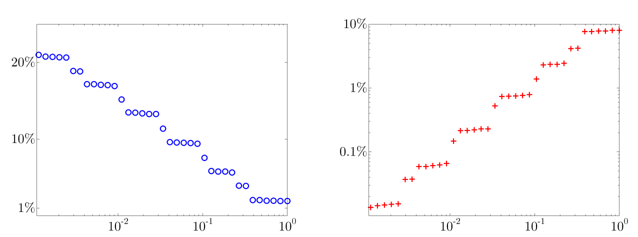
Fig. 7: (a) The diagonal of $F_p^\star$ and (b) the diagonal of $F_v^\star$ for different values of γ: 10^{-4} ($\circ$), 0.1526 ($+$), and 1 ($*$). Diagonals of centralized position and velocity gains are very similar to ($\circ$) for $\gamma = 10^{-4}$.

algorithm has been used to identify sparse communication architectures. We have then optimized the feedback gains subject to structural constraints imposed by the identified communication architectures. We have demonstrated the effectiveness of the developed approach via a simple mass-spring system example.

We have already extended our sparsity-promoting optimal control framework to the synthesis of *block* sparse feedback gains by incorporating penalty functions that promote block sparsity [19]. Furthermore, we have developed easy-to-use software

www.ece.umn.edu/∼mihailo/software/lqrsp/

and demonstrated its utility on several distributed control problems. Additionally, we have employed ADMM for selection of an *a priori* specified number of leaders in order to minimize the variance of stochastically forced dynamic networks [22]. We also aim to extend the presented framework to the observer-based sparse optimal feedback design.



(a) $\mathbf{card}\,(F^\star)/\mathbf{card}\,(F_c)$ (b) $(J(F^\star) - J(F_c))/J(F_c)$

Fig. 8: (a) The sparsity level and (b) the performance loss of $F^\star$ compared to the centralized gain F_c.

ACKNOWLEDGEMENTS

We would like to thank Stephen Boyd for stimulating discussions.

REFERENCES

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

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

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

[4] C. Langbort, R. S. Chandra, and R. D'Andrea, "Distributed control design for systems interconnected over an arbitrary graph," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1502–1519, 2004.

[5] B. Bamieh and P. G. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Syst. Control Lett.*, vol. 54, pp. 575–583, 2005.

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

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

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

[9] 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.

[10] M. R. Jovanović, "On the optimality of localized distributed controllers," *Int. J. Systems, Control and Communications*, vol. 2, no. 1/2/3, pp. 82–99, 2010.

[11] M. Fardad and M. R. Jovanović, "Design of optimal controllers for spatially invariant systems with finite communication speed," *Automatica*, vol. 47, no. 5, pp. 880–889, 2011.

[12] F. Lin, M. Fardad, and M. R. Jovanović, "Augmented Lagrangian approach to design of structured optimal state feedback gains," *IEEE Trans. Automat. Control*, vol. 56, no. 12, pp. 2923–2929, 2011.

[13] 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, no. 1, pp. 1–124, 2011.

[14] 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.

[15] S. Schuler, M. Gruhler, U. Münz, and F. Allgöwer, "Design of structured static output feedback controllers," in *Preprints of the 18th IFAC World Congress*, 2011, pp. 271–276.

[16] S. Schuler, C. Ebenbauer, and F. Allgöwer, "ℓ_0-system gain and ℓ_1-optimal control," in *Preprints of the 18th IFAC World Congress*, 2011, pp. 9230–9235.

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

[18] M. Fardad, F. Lin, and M. R. Jovanović, "Sparsity-promoting optimal control for a class of distributed systems," in *Proceedings of the 2011 American Control Conference*, 2011, pp. 2050–2055.

[19] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Automat. Control*, 2011, submitted; also arXiv:1111.6188v1.

[20] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[21] P. M. Mäkilä and H. T. Toivonen, "Computational methods for parametric LQ problems – a survey," *IEEE Trans. Automat. Control*, vol. 32, no. 8, pp. 658–671, 1987.

[22] F. Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in large dynamical networks: noise-corrupted leaders," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 2932–2937.