# A Quadratic Programming Approach for Solving the $\ell_1$ Multiblock Problem

Nicola Elia and Munther A. Dahleh, *Senior Member, IEEE*

***Abstract*— The authors present a new method to compute solutions to the general multiblock $\ell_1$ control problem. The method is based on solving a standard $\mathcal{H}_2$ problem and a finite-dimensional semidefinite quadratic programming problem of appropriate dimension. The new method has most of the properties that separately characterize many existing approaches. In particular, as the dimension of the quadratic programming problem increases, this method provides converging upper and lower bounds on the optimal $\ell_1$ norm and, for well posed multiblock problems, ensures the convergence in norm of the suboptimal solutions to an optimal $\ell_1$ solution. The new method does not require the computation of the interpolation conditions, and it allows the direct computation of the suboptimal controller.**

***Index Terms*—Computational methods, $l_1$ control, optimal control, quadratic programming.**

## I. Introduction

The general multiblock $\ell_1$ problem has been shown to be equivalent to an infinite-dimensional linear program. Several methods have been proposed in the literature to provide approximate solutions to this problem. The main ones are: $Q$-design, finitely many variables–finitely many equations approximation (FMV–FME), delay augmentation (DA), the convex programming approach using mixed $\ell_1/\mathcal{H}_\infty$, a geometric dynamic programming approach, and finally, a state-space approach. Bellow is a brief account of the advantages and disadvantages of these methods.

The first approach, the $Q$-design method [3], is based on approximating the optimal controller by approximating the stable $Q$ parameter (in the standard Youla parameterization) by a finite impulse response (FIR) system. The $\ell_1$ problem then becomes a finite-dimensional linear program in the parameters of $Q$. This provides an upper bound on the optimal solution which is guaranteed to converge to the actual optimal as the length of the FIR increases to infinity. The controller is automatically derived from the $Q$ parameter. In this paper, we will show that, for a large class of multiblock problems, this method guarantees that a subsequence of suboptimal solutions converges in norm to an optimal $\ell_1$ solution. The disadvantage of this procedure is that it has no stopping criterion and may cause order inflation in the controller due to the FIR approximation.

The FMV method is based on approximating the closed-loop map by an FIR and thus has similar properties to the above method. The FME method is based on approximating the dual of the original $\ell_1$ problem and that provides converging lower bounds of the optimal value. This resolves the issue of finding a stopping criterion; however, it requires computing interpolation conditions that characterize the closed-loop map.

In the DA method, the multiblock problem is transformed into a one-block problem by introducing fictitious delayed inputs and outputs to the controller. This method has three important properties: 1) the convergence of the lower bound is generally fast; 2) there exists a particular ordering of the input and output channels that ensures the convergence of the upper bound together with the convergence in norm of the computed suboptimal solutions to an optimal $\ell_1$ solution for all well-posed multiblock problems; and 3) for many problems, there exists an ordering of the input and output channels that will generate a sequence of suboptimal controllers without order inflation. Although properties 2) and 3) are very interesting from a theoretical point of view, in practice, finding the right ordering can be difficult even for problems of moderate size in terms of number of input and output channels. The computation of a suboptimal controller as well as the construction of the interpolation conditions are the main practical limitations to the application of DA and, in general, of all the methods based on direct optimization on the space of closed-loop maps; for details on the above methods, see [6] and [7]. A method that avoids the computation of the interpolation conditions tightly related to the $Q$-design approach has been recently proposed in [13].

The geometric methods based on dynamic programming arguments [1] provide a recursive algorithm for computing FMV. It has a direct relation to the state-space methods [16], [17], [2] as shown in [9]. Nevertheless, their computational properties are still under investigation.

Finally, a mixed objective approach based on solving an $\ell_1/\mathcal{H}_\infty$ was suggested in [12]. The method provides converging suboptimal solutions but does not provide a stopping criterion. The solutions are based on solving convex optimization problems.

The proposed method in this paper can be seen as a $Q$-design method based on a mixed objective optimization. It therefore provides directly computable suboptimal controllers, converging upper bounds to the optimal $\ell_1$ cost, and norm convergence of the suboptimal solutions to the optimal $\ell_1$

Manuscript received March 20, 1996; revised November 11, 1997. Recommended by Associate Editor, J. Shamma. This work was supported by the NSF under Grant 9157306-ECS, Draper Laboratory under Grant DL-H-441636, and AFOSR under Grant F49620-95-0219.

The authors are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Publisher Item Identifier S 0018-9286(98)06592-1.

solution. Moreover, since the method is based on closed-loop map approximation, it provides a converging lower bound to the optimal $\ell_1$ cost.

The main idea in this paper can be better understood in the simple single-input/single-output (SISO) case. Instead of minimizing the $\ell_1$ norm of the closed-loop impulse response, we minimize the square of the $\ell_1$ norm of the first $N$ samples plus the square of the $\mathcal{H}_2$ norm of the tail (after $N$) of the sequence. For $N = 0$ we have a standard $\mathcal{H}_2$ problem, and as $N \to \infty$ the problem becomes a standard $\ell_1$ problem. This problem has a special property in that the solutions are achieved by FIR $Q$'s of order $N-1$, when the Youla parameterization is derived from the optimal $\mathcal{H}_2$ controller. In order to solve the mixed objective optimization, we first solve the standard $\mathcal{H}_2$ problem. We then use the optimal $\mathcal{H}_2$ controller to compute $H$, $U$, and $V$ in the Youla parameterization ($\Phi = H - UQV$). Finally, we solve the convex optimization problem, which can be shown to be equivalent to a finite-dimensional problem. The paper also discusses in detail all the convergence issues.

Given the computational nature of the $\ell_1$ solutions, several researchers have extended these results into constrained $\ell_1$ problems [8], [18], [11], [12], [15]. The results in this paper utilize solutions for mixed objectives to solve the $\ell_1$ problem. In addition, these results can be used in mixed objectives; however, we will not discuss this here.

The paper is organized as follows. Section II contains preliminaries and the problem formulation; Section III presents the mixed problem that approximates the $\ell_1$ problem; Section IV discusses the various convergence issues; and Section V discusses computational issues and provides an example. Finally, Section VI contains the conclusions.

## II. Notation and Problem Setup

In this section we establish the notation that will be used throughout the paper. Apart from some minor differences, we follow quite closely the notation in [6].

Given a complex-valued $m \times n$ matrix $A$, $\bar{A}$ denotes the conjugate of $A$.

$\ell_2^{m\times n}(\mathbf{Z})$ denotes the Hilbert space of sequences of complex-valued $m \times n$ matrices, with inner product defined as

$$\langle H, G\rangle = \sum_{k=-\infty}^{\infty} \mathrm{trace}(\bar{G}(k)^T H(k)).$$

$\ell_2^{m\times n}(\mathbf{Z})$ can be written as the direct sum of two spaces of one-sided sequences

$$\ell_2^{m\times n}(\mathbf{Z}_+) \oplus \ell_2^{m\times n}(\mathbf{Z}_-)$$

with $0 \in \mathbf{Z}_+$. The Fourier transform of $G$ in $\ell_2^{m\times n}(\mathbf{Z})$ is defined as

$$\hat{G}(e^{-i\theta}) = \sum_{k=-\infty}^{\infty} G(k) e^{-ik\theta}.$$

$\mathcal{L}_2^{m\times n}[0, 2\pi)$ denotes the space whose elements are the Fourier Transform of elements in $\ell_2^{m\times n}(\mathbf{Z})$. The decomposition of $\ell_2^{m\times n}(\mathbf{Z})$ into $\ell_2^{m\times n}(\mathbf{Z}_+)$ and $\ell_2^{m\times n}(\mathbf{Z}_-)$ induces, through the Fourier Transform, an analogous separation in $\mathcal{L}_2^{m\times n}[0, 2\pi)$.

$$\mathcal{L}_2^{m\times n}[0, 2\pi) = \mathcal{H}_2^{m\times n} \oplus \mathcal{H}_2^{m\times n^\perp}$$

where $\mathcal{H}_2^{m\times n}$ contains all the matrix-valued functions in $\mathcal{L}_2^{m\times n}[0, 2\pi)$ that are analytic inside the open unit disc, and $\mathcal{H}_2^{m\times n^\perp}$ contains all the matrix-valued functions in $\mathcal{L}_2^{m\times n}[0, 2\pi)$ analytic in the complement of the unit disc. $\mathcal{RH}_2^{m\times n}$ is the space of rational transfer function matrices in $\mathcal{H}_2^{m\times n}$. $\mathcal{RH}_2^{m\times n^\perp}$ is defined analogously.

Since we will mostly work with real-valued unilateral matrix sequences supported on the positive integers, for notational convenience we will denote $\ell_2^{m\times n}(\mathbf{Z}_+)$ as $\ell_2^{m\times n}$.

$\ell_1^{m\times n}$ is the space of all sequences $H$ of $m\times n$ real matrices such that

$$\|H\|_1 \triangleq \max_{1\le i\le m} \sum_{j=1}^{n} \sum_{t=0}^{\infty} |h_{ij}(t)| < \infty.$$

The dual of a Banach space $X$ is denoted by $X^*$. The dual space of $\ell_1^{n\times m}$ is $\ell_\infty^{m\times n}$ which is the space of all sequences, $G$, of $m \times n$ real matrices such that

$$\|G\|_\infty \triangleq \sum_{i=1}^{m} \max_{1\le j\le n} \max_{0\le t<\infty} |g_{ij}(t)| < \infty.$$

$c_0^{m\times n}$ denotes the subspace of $\ell_\infty^{m\times n}$ consisting of all the sequences of $m \times n$ real matrices for which

$$\lim_{t\to\infty} g_{ij}(t) = 0 \quad \forall i = 1, \cdots, m, \qquad \forall j = 1, \cdots, n.$$

It is well known that $(c_0^{m\times n})^* = \ell_1^{m\times n}$.

Given $\mathcal{A}$, a bounded linear operator from $X$ to $Z$, $\mathcal{A}^* : Z^* \to X^*$ denotes its adjoint. Given $\mathcal{T}$, a bounded linear operator from $Z^*$ to $X^*$, ${}^*\mathcal{T} : X \to Z$ denotes the pre-adjoint, when it exists, of $\mathcal{T}$. $({}^*\mathcal{T})^* = \mathcal{T}$.

For notational simplicity, we will often drop the superscripts $m \times n$ when no confusion arises.

We consider discrete time multi-input/multi-output (MIMO) systems. Given a system $G$, its state-space representation is denoted by

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

$\hat{G}(\lambda)$ denotes the $\lambda$-transform of $G$ which is given by

$$\hat{G}(\lambda) = D + C\lambda(I - \lambda A)^{-1}B.$$

$\hat{G}^\sim(\lambda)$ denotes $\hat{G}^T(\frac{1}{\lambda})$. It is often called the Hilbert space adjoint of $\hat{G}$.

We consider the standard generalized system $M$ shown in Fig. 1

$$M = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right].$$

We make the following assumptions.

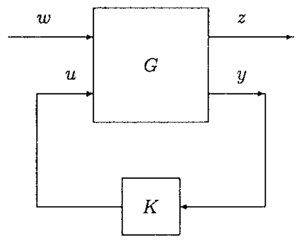


Fig. 1. General setup.

*Assumption 2.1:*

A1) $(A, B_2)$ is stabilizable and $(C_2, A)$ is detectable.

A2) $D_{12}$ has full column rank and $D_{21}$ has full row rank.

A3) $\begin{bmatrix} A-Ie^{-i\theta} & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all $\theta \in [0, 2\pi)$.

A4) $\begin{bmatrix} A-Ie^{-i\theta} & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all $\theta \in [0, 2\pi)$.

A5) $D_{11} = 0$ and $D_{22} = 0$.

Assumptions A1) to A4) are standard assumptions; see [6] and [19]. Assumption A5) is made only to simplify the derivations and can be removed.

All stable closed-loop maps from $w$ to $z$ can be parameterized as $\Phi = H - UQV$ where $Q \in \ell_1^{n_u \times n_y}$ is arbitrary, and $H$, $U$, and $V$ are stable systems that can be computed from the problem data [6]. The $\ell_1$ problem is given by

$$\mu^o = \inf_{\substack{\text{subject to:} \\ \Phi = H - UQV \\ Q \in \ell_1^{n_u \times n_y}}} \|\Phi\|_1. \tag{1}$$

Next, we are going to decompose the above expression in a special way. Let $\ell^{n \times m}$ be the space of all sequences of $n \times m$ matrices. For a fixed positive integer $N$, let $P_N : \ell^{n \times m} \to \mathbb{R}^{n \times m} \times \mathbb{R}^N$ denote the *truncation* operator of order $N$

$$P_N x = [x(0) \quad x(2) \quad \cdots \quad x(N-1)]$$

and $T_N : \ell^{n \times m} \to \ell^{n \times m}$ denote the *tail* operator of order $N$ $T_N x = [x(N) \ x(N+1) \ \cdots]$. Also let $\bar{P}_N : \mathbb{R}^{n \times m} \times \mathbb{R}^N \to \ell^{n \times m}$ denote the operator that makes an infinite sequence from a finite sequence of length $N$ by adding zeros.

$$\bar{P}_N y = [y(0) \quad y(1) \quad \cdots \quad y(N-1) \quad 0 \quad 0 \quad \cdots]$$

and $\bar{T}_N : \ell^{n \times m} \to \ell^{n \times m}$ denote the operator that puts $N$ zeros at the beginning of a sequence

$$\bar{T}_N y = [\underbrace{0 \quad 0 \quad \cdots \quad 0}_{N} \quad y(0) \quad y(1) \quad \cdots].$$

Using these operators, $\Phi$ can be expressed as $\Phi = \bar{P}_N \Phi_1 + \bar{T}_N \Phi_2$ where $\Phi_1 = P_N \Phi$, $\Phi_2 = T_N \Phi$. Similarly, $Q = \bar{P}_N P_N Q + \bar{T}_N T_N Q = \bar{P}_N Q_1 + \bar{T}_N Q_2$ and $H = \bar{P}_N P_N H + \bar{T}_N T_N H = \bar{P}_N H_1 + \bar{T}_N H_2$. Also define

$$U_1 \triangleq P_N U \bar{P}_N$$
$$U_{12} \triangleq T_N U \bar{P}_N$$
$$U_2 \triangleq T_N U \bar{T}_N.$$

Note that $U_1 : \mathbb{R}^{n_u} \times \mathbb{R}^N \to \mathbb{R}^{n_z} \times \mathbb{R}^N$, $U_{12} : \mathbb{R}^{n_u} \times \mathbb{R}^N \to \ell_1^{n_z}$, and $U_2 : \ell_1^{n_u} \to \ell_1^{n_z}$. It follows from the Toeplitz structure of $U$ that $U_2 = U$. Since $U$ is causal, it follows that $U_{21} \triangleq P_N U \bar{T}_N = 0$.

Using these expressions, the decomposition of $\Phi$ can be written as

$$\Phi_1 = H_1 - U_1 Q_1 V$$
$$\Phi_2 = H_2 - U_{12} Q_1 V - U Q_2 V.$$

## III. Approximation Method

Consider the following optimization problem:

$$\mu_N = \inf_{\substack{\text{subject to:} \\ \Phi_1 = H_1 - U_1 Q_1 V \\ \Phi_2 = H_2 - U_{12} Q_1 V - U Q_2 V \\ Q_1 \in \mathbb{R}^{n_u \times n_y \times N}, Q_2 \in \ell_1^{n_u \times n_y}}} \sqrt{\|\Phi_1\|_1^2 + \|\Phi_2\|_2^2}. \tag{2}$$

*Properties of Problem (2):*

P1) For any $N$, denote

$$\|\Phi\|_N := \sqrt{\|\Phi_1\|_1^2 + \|\Phi_2\|_2^2}. \tag{3}$$

$\|\cdot\|_N$ is a norm on $X = \mathbb{R}^{n_z \times n_w \times N} \times \ell_2^{n_z \times n_w}$. Problem (2) is a constrained norm-minimization problem.

P2) For each $N$ the solution to Problem (2) exists.

P3) There exists a parameterization $\Phi = H - UQV$, such that, for any finite $N$, Problem (2) is equivalent to the finite-dimensional convex optimization

$$\mu_N = \inf_{\substack{\text{subject to:} \\ \Phi_1 = H_1 - U_1 Q_1 V \\ \Phi_2 = H_2 - U_{12} Q_1 V \\ Q_1 \in \mathbb{R}^{n_u \times n_y \times N}}} \sqrt{\|\Phi_1\|_1^2 + \|\Phi_2\|_2^2}. \tag{4}$$

Equivalently, the optimal solution with this parameterization satisfies $Q_2 = 0$.

*Proof (Sketch):* The first property is immediate. The second property follows from standard duality theory results, given the fact that Problem (2) is itself the dual of another convex optimization problem. For the third property, the special parameterization is the one obtained from the model-based optimal $\mathcal{H}_2$ controller. For that case, $U$ and $V$ are both inner. The details of the proof are in the Appendix. This property of the optimal $\mathcal{H}_2$ controller can also be derived from the principle of optimality in dynamic programming. □

The next theorem gives the dual formulations for (2) and (4). These formulations will be used in future sections.

*Theorem 3.1:* The dual of Problem (2) with no duality gap is

$$\mu_N = \max_{\substack{\text{subject to:}\\ U_1^* x_1^* V^* + U_{12}^* x_2^* V^* = 0 \\ U_2^* x_2^* V^* = 0 \\ \|x_1^*\|_\infty^2 + \|x_2^*\|_2^2 \le 1 \\ x_1^* \in \mathbb{R}^N, x_2^* \in \ell_2^{n_z \times n_w}}} \langle H_1, x_1^* \rangle + \langle H_2, x_2^* \rangle \tag{5}$$

and the dual of Problem (4) with no duality gap is:

$$\mu_N = \max_{\substack{\text{subject to:}\\ U_1^* x_1^* V^* + U_{12}^* x_2^* V^* = 0 \\ \|x_1^*\|_\infty^2 + \|x_2^*\|_2^2 \le 1 \\ x_1^* \in \mathbb{R}^N, x_2^* \in \ell_2^{n_z \times n_w}}} \langle H_1, x_1^* \rangle + \langle H_2, x_2^* \rangle. \tag{6}$$

*Proof:* This is [10, Th. 1, p. 119]. The details are left to the reader. □

## IV. Convergence Properties

In this section, we discuss the convergence of several quantities. First $\mu_N$ is shown to converge to $\mu^o$. The solution to Problem (4) is a feasible solution to Problem (1) which has a cost denoted by $\bar{\mu}_N$. We will show that $\bar{\mu}_N$ converges to $\mu^o$. In addition, we will derive a simple sequence of lower bounds on $\mu^o$ and show its convergence. Finally, we discuss the convergence of the actual solutions.

### A. Convergence of the Cost $\mu_N$

In the next theorem, we will show that the sequence $\mu_N$ converges to $\mu^o$. The constraints in this problem can be rewritten as a function of $\Phi$ only by introducing a linear operator $\mathcal{A}_{\text{feas}}$, which is standard in the $\ell_1$ literature. The problem is, then, equivalent to the following one:

$$\mu^o = \inf_{\substack{\text{subject to:}\\ \mathcal{A}_{\text{feas}}\Phi = b_{\text{feas}} \\ \Phi \in \ell_1^{n_z \times n_w}}} \|\Phi\|_1 \tag{7}$$

where $\mathcal{A}_{\text{feas}} : \ell_1^{n_z \times n_w} \to \ell_1$ is a linear bounded operator with closed range.

We can write two dual problems for the above minimization: one in $\ell_\infty^{n_z \times n_w}$, which is the norm-dual of $\ell_1^{n_z \times n_w}$, and the other in $\ell_\infty$, the dual of the constraint space. They are respectively given by

$$\max_{\substack{\text{subject to:}\\ U^* x^* V^* = 0 \\ \|x^*\|_\infty \le 1 \\ x^* \in \ell_\infty^{n_z \times n_w}}} \langle H, x^* \rangle \tag{8}$$

and

$$\max_{\substack{\text{subject to:}\\ x^* = \mathcal{A}_{\text{feas}}^* z^* \\ \|x^*\|_\infty \le 1 \quad z^* \in \ell_\infty}} \langle b_{\text{feas}}, z^* \rangle. \tag{9}$$

It is worth mentioning that all $x^*$ satisfying $U^* x^* V^* = 0$ are in the range of $\mathcal{A}_{\text{feas}}^*$.

We can now prove the following result.

*Theorem 4.1:* Let $\mu_N$ be defined as in Problem (4). Then

$$\lim_{N \to \infty} \mu_N = \mu^o.$$

*Proof (Sketch):* From the optimal solution $x^o$ of the $\ell_1$ problem, which exists under standard assumptions, we compute the sequence $\{\|x^o\|_N\}$, where $\|\cdot\|_N$ is defined as in (3). This provides a sequence convergent to $\mu^o$ with $\|x^o\|_N \ge \mu_N$ for any $N \ge 0$. Thus $\limsup_{N\to\infty} \mu_N \le \mu^o$.

To show that $\liminf_{N\to\infty} \mu_N \ge \mu^o$, we consider Problem (9), dual of Problem (7). It is well known that there are finite support feasible dual sequences $z^*$ whose cost, $\mu$, is arbitrarily close to $\mu^o$ from below. The main step is to show that from each $z^*$ we can find feasible dual solutions to Problem (6) with cost approaching $\mu$ from below as $N$ goes to infinity. The details are in the Appendix. □

### B. Convergence of the Upper Bounds

For each $N$, the optimal solution of Problem (4) has $Q = [Q_1, 0] \in \ell_1^{n_u \times n_y}$ and thus the resulting optimal $\Phi^N = H - UQV = [\Phi_1^N \ \ \Phi_2^N]$ is a bounded-input/bounded-output (BIBO) stable closed-loop map. Define

$$\bar{\mu}_N = \|\Phi^N\|_1.$$

It follows that $\bar{\mu}_N \ge \mu^o$.

Next, we will show that $\bar{\mu}_N$ converges to $\mu^o$ as $N$ goes to infinity. To do this, we first show that if $\|\Phi_2^N\|_2$ is bounded, then $\|\Phi_2^N\|_1$ is also bounded. We then prove that the sequence $\|\Phi_2^N\|_2$ goes zero.

To prove the first statement, we need to recall the following result.

*Fact 4.1:* For the generalized system $M$ there exists a parameterization of all closed-loop stable maps $\Phi = H - UQV$ with polynomial $\hat{U}(\lambda)$ and $\hat{V}(\lambda)$.

*Lemma 4.1:* Consider the sequence of optimal solutions, $[\Phi_1^N \ \ \Phi_2^N]$, of Problem (4), or equivalently of Problem (2), as a function of $N$. Then there exists a fixed $M$ and two positive constants $c_1$ and $c_2$ such that

$$\left\|\Phi_2^N\right\|_1 \le c_1 \left\|\Phi_2^N\right\|_2 + c_2, \qquad \text{for all } N \ge M.$$

Moreover, $c_2$ goes to zero as $M$ approaches $\infty$.

*Proof:* For simplicity, we consider the case where $n_z = n$ and $n_u = n_w = 1$. Assume we have found a parameterization as in Fact 4.1. Consider Problem (2) with the resulting $\hat{U}(\lambda)$ polynomial. Note that $H$ may not be polynomial in general; however, it is an element of $\ell_1^{n \times 1}$ and hence of $\ell_2^{n \times 1}$.

Note that Problem (2) is jointly convex in $Q_1$ and $Q_2$. This implies that, to find the optimal solution, we can first find the optimal $Q_2$ for each fixed $Q_1$ and then we can minimize with respect to $Q_1$. In our case, for a fixed $Q_1$, the optimal $Q_2$ is the one that minimizes the norm of $\Phi_2$, i.e., the one that minimizes $\|H_2 - U_{12}Q_1 - UQ_2\|_2$. This is a standard $\mathcal{H}_2$ problem. However, instead of approaching it in the space $\mathcal{L}_2$ as done in the proof of P3 (see Appendix), it is convenient to look at it as a minimization on the space of one-sided sequences

$\ell_2^{n_z \times n_w}(\mathbf{Z}_+)$. Thus, $H_2$ and $Q_2$ can be seen as sequences in $\ell_2^{n\times 1}(\mathbf{Z}_+)$, $Q$ as an infinite block Toeplitz matrix, and $Q_{12}$ as a matrix with infinite rows and $N$ columns.

Since this is a standard least squares minimization in an Hilbert space, the optimal $Q_2$ is given by

$$Q_2 = U^+(H_2 - U_{12}Q_1)$$

and the optimal $\Phi_2^N$ is

$$\Phi_2^N = (I - UU^+)(H_2 - U_{12}Q_1)$$

where $U^+$ denotes the pseudo-inverse of $U$. It is well known from the standard $\mathcal{H}_2$ optimization problem that the optimal $Q_2$ is in $\ell_1$ and $\Phi_2^N$ is in $\ell_1^{n\times 1}$. If we denote by $A = I - UU^+$, then

$$\Phi_2^N = AH_2 - AU_{12}Q_1.$$

If $\hat{U}(\lambda)$ is a polynomial matrix, say of order $M-1$, then for all the $N \geq M$, $U_{12}$ has all zeros in the first $N-M$ columns. Moreover, the last $M$ columns of $U_{12}$ stay the same for all $N \geq M$ and their $k$ elements are zeros for $k \geq M$.

Thus, for $N \geq M$, the optimal $\Phi_2^N$ has a representation of the form

$$\Phi_2^N = AH_2 + [0, B]Q_1$$

where $[0, B]Q_1 = AU_{12}Q_1$ and $B$ maps $\mathbb{R}^M$ into $\ell_1^{n\times 1}$. In other words, $\Phi_2^N$ lies in a fixed translated subspace of $\ell_1^{n\times 1}$ and depends only on the last $M-1$ elements of $Q_1$.

Consider now the sequence of optimal $Q_1$ for each $N$, denoted by $Q_1^N$. From the triangle inequality it follows that

$$\left\|\Phi_2^N\right\|_2 \geq -\|AH_2\|_2 + \left\|[0, B]Q_1^N\right\|_2$$

or equivalently

$$\left\|[0, B]Q_1^N\right\|_2 \leq \left\|\Phi_2^N\right\|_2 + \|AH_2\|_2.$$

From the fact that the range of $B$ is fixed and finite dimensional in $\ell_1^{n\times 1}$ for each $N \geq M$, it follows that there exists a positive constant $c_1$ such that

$$c_1\left\|[0, B]Q_1^N\right\|_2 \geq \left\|[0, B]Q_1^N\right\|_1.$$

Therefore, we have that

$$\left\|[0, B]Q_1^N\right\|_1 \leq c_1\left\|\Phi_2^N\right\|_2 + c_1\|AH_2\|_2$$

or

$$\left\|\Phi_2^N\right\|_1 \leq c_1\left\|\Phi_2^N\right\|_2 + c_1\|AH_2\|_2 + \|AH_2\|_1$$

where we added and subtracted $AH_2$ in the norm on the left-hand side, and we used the triangle inequality. Notice that $\epsilon_N = c_1\|AH_2\|_2 + \|AH_2\|_1$ is a monotonically nonincreasing sequence; therefore, $\epsilon_N \leq \epsilon_M$ for all $N \geq M$. Besides, $\epsilon_M \to 0$ as $M \to \infty$. The result then follows if we let $c_2 = \epsilon_M$. □

The previous result together with next theorem determine the convergence of $\bar{\mu}_N$ to the optimal $\ell_1$ cost.

*Theorem 4.2:* Consider the sequence of optimal solutions to Problem (4) as a function of $N$ $[\Phi_1^N \;\; \Phi_2^N]$. Then

$$\lim_{N\to\infty} \left\|\Phi_2^N\right\|_2 = 0.$$

*Proof:* See the Appendix. □

As an immediate result from the above theorem and Lemma 4.1 we have the following.

*Corollary 4.1:* The sequence of upper bounds $\bar{\mu}_N$ converges to $\mu^o$.

*Proof:* For each $N$, let $\Phi^N = [\Phi_1^N, \Phi_2^N]$ be the optimal solution and denote $\mu_{1N}$ and $\mu_{2N}$ as follows:

$$\mu_{1N} = \left\|\Phi_1^N\right\|_1$$
$$\mu_{2N} = \left\|\Phi_2^N\right\|_2.$$

Clearly, $\mu_N = \sqrt{\mu_{1N}^2 + \mu_{2N}^2}$.

From Theorem 4.2 we have that $\mu_{2N} \to 0$ as $N \to \infty$. From Theorem 4.1 we have that $\mu_N \to \mu^o$ as $N \to \infty$. As consequence of these results, it follows that

$$\lim_{N\to\infty} \mu_{1N} = \mu^o.$$

From Lemma 4.1 it follows that $\|\Phi_2^N\|_1 \to 0$ as $N \to \infty$. Since it is always true that

$$\mu_{1N} \leq \bar{\mu}_N \leq \mu_{1N} + \left\|\Phi_2^N\right\|_1$$

we have that $\bar{\mu}_N \to \mu^o$ as $N \to \infty$. □

### C. A Sequence of Lower Bounds

Next we derive an easily computable sequence of lower bounds for $\mu^o$ which will be denoted by $\underline{\mu}_N$. Assume, without loss of generality, that $\mu^o$ is strictly greater than zero; since $\mu^o = 0$ if and only if $\mu_0 = \mu_{1N} = \mu_{2N} = 0$.

Consider the dual problem in (5). For a given $N$, $x^* = [x_1^*, x_2^*]$ are the dual variables. Let $\gamma_1^N = \|x_1^*\|_\infty$ and $\gamma_2^N = \|x_2^*\|_2$. From the alignment conditions it follows that

$$\gamma_1^N = \frac{\mu_{1N}}{\mu_N}$$
$$\gamma_2^N = \frac{\mu_{2N}}{\mu_N}. \tag{10}$$

Given that $\|x_2^*\|_\infty \leq \sqrt{n_z}\|x_2^*\|_2$, it follows that the element $y^* = \frac{x^*}{\gamma_1^N + \sqrt{n_z}\gamma_2^N}$ has $\|y^*\|_\infty \leq 1$ and therefore, it is feasible for Problem (8), the dual of the $\ell_1$ problem in (1). Hence

$$\mu^o \geq \frac{\mu_N}{\gamma_1^N + \sqrt{n_z}\gamma_2^N}.$$

Substituting the expressions in (10) we obtain

$$\mu^o \geq \frac{\mu_N^2}{\mu_{1N} + \sqrt{n_z}\mu_{2N}} := \underline{\mu}_N.$$

Clearly $\underline{\mu}_N$ converges to $\mu^o$ as $N \to \infty$, since the numerator and the denominator converge to $(\mu^o)^2$ and $\mu^o$, respectively.

*Remark 4.1:* Notice that the results on the convergence of the upper and the lower bounds do not indicate that the sequences $\{\bar{\mu}_N\}$ and $\{\underline{\mu}_N\}$ are, respectively, monotonically nonincreasing and monotonically nondecreasing. In fact, they may not be as such, as shown in the example of Section V.

### D. Strong Convergence of the Suboptimal Solutions

We now address the convergence of the solutions of Problem (2), $\Phi^N$. The results in this section are similar to the ones presented in [7].

*Theorem 4.3:* For each $N$, let $\Phi^N$ be an optimal solution to Problem (2). Then, the sequence $\{\Phi^N\}$ contains a weak* convergent subsequence whose weak* limit is an optimal solution, $\Phi^o$, for Problem (1). Moreover, if the optimal solution is unique, then the whole sequence converges to it.

*Proof:* Since $\|\Phi^N\| = \bar{\mu}_N$ is a convergent sequence, $\{\Phi^N\}$ contains a weak* converging subsequence. Denote the weak* limit by $\Phi^{w^*}$. For each $N$, $\Phi^N$ is feasible to Problem (1). This implies that $\Phi^N$ belongs to the set $\mathcal{S} = \{\Phi \mid \mathcal{A}_{\text{feas}}\Phi = b_{\text{feas}}\}$. However, $\mathcal{S}$ is weak* closed, and thus it contains all its weak* limit points. Therefore, $\Phi^{w^*}$ is a feasible solution to Problem (1).

$\Phi^{w^*}$ is also an optimal solution to Problem (1). This follows from [7, Lemma 2.5.5]:

$$\|\Phi^{w^*}\|_1 \leq \liminf_{N_s \to \infty} \|\Phi^{N_s}\|_1 \leq \mu^o.$$

If $\Phi^o$ is unique, then all subsequences must converge weak* to it. Thus the whole sequence converges weak* to $\Phi^o$. □

We recall a standard result [7] useful in the development of the main convergence theorem.

*Lemma 4.2:* If $x_n \in \ell_1$ converges weak* to $x$, and $\|x_n\|_1$ converges to $\|x\|_1$, then $x_n$ converges to $x$ in the $\ell_1$ norm.

From the previous theorem and the above result it follows that, if any row of $\Phi^o$, $(\Phi^o)_i$ achieves the optimal norm, $\mu^o$, then the respective row of $\Phi^{N_s}$, $\Phi^{N_s}$, will converge in norm to $(\Phi^o)_i$.

Let $I \subset \{1, \cdots, n_z\}$ be the set of row indexes for which $\|(\Phi^o)_i\|_1 = \mu^o$. $I = \{i \mid \|(\Phi^o)_i\|_1 = \mu^o\}$. Let $\text{card(I)}$ denote the cardinality of $I$, i.e., the number of elements in $I$. Given any $\Phi \in \ell_1^{n_z \times n_w}$ and any index set $I \subset \{1, \cdots, n_z\}$ with cardinality $\text{card(I)}$, we can construct $\Phi_I \in \ell_1^{\text{card(I)} \times \text{n}_\text{w}}$ by collecting only the rows on $\Phi$ whose index is in $I$. Define $H_I$ and $U_I$ analogously so that $\Phi_I = H_I - U_I QV$.

*Definition 4.1:* A multiblock problem with an optimal solution such that $\text{card(I)} = \text{n}_\text{u}$ is referred to as a well-posed problem.

As noted in [7], for most multiblock problems the optimal solution achieves the optimal norm on at least $n_u$ rows. However, it is possible to construct multiblock problems for which $n_I < n_u$. Such problems are not well-posed in the sense of Definition 4.1.

We then have the following result.

*Theorem 4.4:* Assume that Problem (1) is a well-posed multiblock problem. Let $\Phi^{N_s}$ be a subsequence of solutions to Problem (2) that converges weak* to $\Phi^o$, an optimal solution to Problem (1) with $\text{card(I)} = \text{n}_\text{u}$. Assume further that $\hat{U}_I(\lambda)$ has full normal rank. Then $\|\Phi^{N_s} - \Phi^o\|_1 \to 0$ as $N \to \infty$.

*Proof:* We have that $\Phi_I^{N_s}$ converges strongly to $\Phi_I^o$. From the rank assumption on $\hat{U}_I$ we have that the map from $Q^{N_s}V$ to $\Phi_I^{N_s}$

$$\hat{Q}^{N_s}\hat{V} = \hat{U}_I^{-1}\left(\hat{H}_I - \Phi_I^{N_s}\right)$$

is continuous with continuous inverse. Therefore, $Q^{N_s}V$ converges strongly to $Q^oV$. The result follows from the continuity in $QV$ of the map $\Phi = H - UQV$. □

*Corollary 4.2:* The sequence $Q^{N_s}$ converges strongly to $Q^o$.

*Proof:* Since by assumption $\hat{V}$ has full normal rank, there is a set $J$ of $n_y$ columns of $\hat{V}$ such that $V_J \in \ell_1^{n_y \times n_y}$ and $\hat{V}_J$ is invertible with continuous inverse. Given that $Q^{N_s}V$ is converging strongly to $Q^oV$, then $Q^{N_s}V_J$ is also converging strongly to $Q^oV_J$. The result follows from the continuity of $V_J^{-1}$. □

## V. Example

In this section we briefly discuss some computational issues and present an example.

First, we describe how Problem (4) can be rewritten as a linear matrix inequality problem. To avoid notational complications, we only describe the case where $\Phi$ is SISO. The generalization to the case, where $\Phi$ is MIMO, although tedious, is straightforward and is left to the reader.

We start by removing the variable $\Phi_1$ and $\Phi_2$ from the problem by rewriting it as follows:

$$\mu_N^2 = \inf_{\substack{\text{subject to:} \\ \|H_1 - U_1Q_1V\|_1 \leq \gamma_1 \\ Q_1 \in \mathbb{R}^N}} \gamma_1^2 + \|H_2 - U_{12}Q_1V\|_2^2. \tag{11}$$

Using the linearity in $Q_1$, $U_1Q_1V$ can be rewritten in the following matrix form: $U_1Q_1V = A_f q_1$ where, $q_1 = [q_1(0), \cdots, q_1(N-1)]^T$ is the vector containing the first elements of the impulse response of $Q_1$, and $A_f \in \mathbb{R}^{N \times N}$.

The $\ell_1$ norm constraint can now be transformed into a set of linear constraints by a standard trick in linear programming. Namely, we have that

$$\left\{ \begin{array}{c} -\rho \leq H_1 - A_f q_1 \leq \rho \\ \rho(k) \geq 0 \\ A_{\ell_1} q_1 \triangleq \sum_{k=0}^{N-1} \rho(k) \leq \gamma_1 \end{array} \right\} \Leftarrow \|H_1 - A_f q_1\|_1 \leq \gamma_1.$$

On the other hand, if $\|H_1 - A_f q_1\|_1 = \gamma_1$ then there exists a vector $\rho$ such that

$$\begin{array}{c} -\rho \leq H_1 - A_f q_1 \leq \rho \\ \rho(k) \geq 0 \\ \sum_{k=0}^{N-1} \rho(k) = \gamma_1. \end{array}$$

Thus $\mu_n^2$ is given by

$$\mu_N^2 = \inf_{\substack{\text{subject to:} \\ A_{\ell_1}\rho \leq \gamma_1 \\ H_1 - A_f q_1 \leq \rho \\ -H_1 + A_f q_1 \leq \rho \\ \rho \geq 0,\ q_1 \in \mathbb{R}^N}} \gamma_1^2 + \|H_2 - U_{12}Q_1V\|_2^2. \tag{12}$$

Now, consider the term $\|H_2 - U_{12}Q_1V\|_2^2$. From simple state-space manipulations, $H_2 - U_{12}Q_1V$ can be represented as follows:

$$H_2 - U_{12}Q_1V = \left[\begin{array}{c|c} \tilde{A} & \tilde{B}\begin{bmatrix}1\\ q_1\end{bmatrix} \\ \hline \tilde{C} & 0 \end{array}\right].$$

Then

$$\|H_2 - U_{12}Q_1V\|_2^2 = \mathrm{Trace}\left(\begin{bmatrix}1 & q_1^T\end{bmatrix}\tilde{B}^T L_o \tilde{B}\begin{bmatrix}1\\ q_1\end{bmatrix}\right)$$

where $L_o$ is the observability gramian, which is given by unique positive definite solution of

$$\tilde{A}^T L_o \tilde{A} - L_o + \tilde{C}^T\tilde{C} = 0.$$

The dimension of $L_o$ is fixed for a given problem and does not depend on $N$, the order of the approximation. This is because $H_2 - U_{12}Q_1V$ can be realized with $\tilde{A}$ and $\tilde{C}$ fixed and independent of $N$. Moreover, the computation of $L_o$ is independent of $q_1$. Therefore, $L_o$ can be precomputed.

Finally, from [4], we have that the constraint

$$\mathrm{Trace}\left(\begin{bmatrix}1 & q_1^T\end{bmatrix}\tilde{B}^T L_o \tilde{B}\begin{bmatrix}1\\ q_1\end{bmatrix}\right) \le \gamma_2^2$$

is equivalent to the following linear matrix inequality (LMI) in $X(X = X^T)$, $q_1$, and $\gamma_2^2$:

$$\mathrm{Trace}(X) \le \gamma_2^2, \qquad \begin{bmatrix} X & \begin{bmatrix}1 & q_1^T\end{bmatrix}\tilde{B}^T \\ \tilde{B}\begin{bmatrix}1\\ q_1\end{bmatrix} & L_o^{-1}\end{bmatrix} \ge 0.$$

Thus, Problem 4 can be solved by solving the following LMI problem:

$$\mu_N^2 = \inf_{\substack{\text{subject to:} \\ \begin{bmatrix}\gamma^2-\gamma_2^2 & \gamma_1\\ \gamma_1 & 1\end{bmatrix}\ge 0 \\ A_{\ell_1}\rho \le \gamma_1 \\ \mathrm{Trace}(X)\le\gamma_2^2 \\ H_1 - A_f q_1 \le \rho \\ -H_1 + A_f q_1 \le \rho \\ \begin{bmatrix} X & \begin{bmatrix}1 & q_1^T\end{bmatrix}\tilde{B}^T \\ \tilde{B}\begin{bmatrix}1\\ q_1\end{bmatrix} & L_o^{-1}\end{bmatrix}\ge 0 \\ \rho\ge 0, \quad X = X^T, \quad q_1\in\mathbb{R}^N}} \gamma^2. \tag{13}$$

Note that, with a slight abuse of notation, we use $\ge$ and $\le$ to describe both matrix inequalities and component-wise inequalities. Note however, that since the linear inequalities are a special case of matrix inequalities, they can be interpreted as component-wise matrix inequalities.

We now apply the new method to solve the $\ell_1$ problem for the system described in Fig. 2, where the plant is given by

$$\hat{P}(\lambda) = \frac{\lambda(\lambda - 0.5)}{(\lambda - 0.1)(1 - 0.5\lambda)}$$

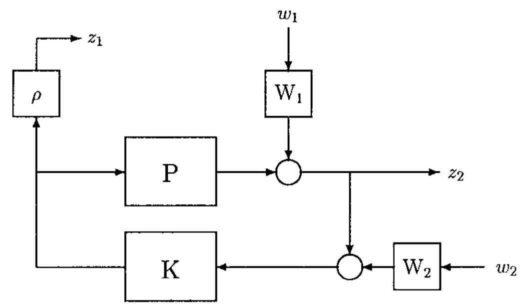


Fig. 2. System configuration.

and the weights are

$$\rho = 0.1, \qquad \hat{W}_1(\lambda) = \frac{0.4}{1 - 0.6\lambda}, \qquad \hat{W}_2(\lambda) = \frac{1 - 0.75\lambda}{1 - 0.25\lambda}.$$

The same problem has also been considered in [7].

The closed-loop system is the following:

$$\Phi = \begin{pmatrix} \rho K(1 - PK)^{-1}W_1 & \rho K(1 - PK)^{-1}W_2 \\ (1 - PK)^{-1}W_1 & PK(1 - PK)^{-1}W_2 \end{pmatrix}.$$

Fig. 3 shows the convergence of the upper and the lower bounds for increasing order of $Q_1(\lambda)$. Notice that the sequence of upper bounds is not monotonically nonincreasing.

For $N = 26$, which correspond to a polynomial $Q_1$ of order 25, we obtain that the suboptimal $\ell_1$ solution $\Phi^N$ with $\bar{\mu}^N = \|\Phi^N\|_1 = 71.1147$. The maximum value of the lower bound is 71.0884. Thus $\|\Phi^N\|_1$ is within 0.04% of the optimal $\ell_1$ norm, $\mu^o$. The $\ell_1$ norms of the single transfer functions are

$$\begin{bmatrix} \|\Phi_{11}\|_1 & \|\Phi_{12}\|_1 \\ \|\Phi_{21}\|_1 & \|\Phi_{22}\|_1 \end{bmatrix} = \begin{bmatrix} 1.8606 & 5.4428 \\ 26.0191 & 45.0956 \end{bmatrix}.$$

From this information, we see that the second row is dominant in the problem and, in particular, $\Phi_{22}$ is the element with higher norm. This immediately indicates the right reordering for DA, namely, we need to switch the two outputs and the two inputs in order for the DA upper bound to converge to $\mu^o$.

The coefficients of the suboptimal $Q_1$ for $N = 26$ are shown in Fig. 4.

Although $Q_1(k)$ is supported for all $N \le 26$, the coefficients $Q_1(k)$'s for $k > 13$ are all smaller than $3 \cdot 10^{-5}$. By neglecting these coefficients we obtain that the suboptimal controller $K$ has order 16. The resulting $\ell_1$ norm of the closed-loop system, using such a controller, is 71.1147. Thus, we do not lose much by considering $Q_1$ of order 12 instead of 25. If we only consider the first nine coefficients of $Q_1$, as the figure suggests, we obtain a suboptimal controller of order 12 with $\ell_1$ norm equal to 71.1296, still within the 0.06% of the optimal.

## VI. Conclusion

We have presented a new method to compute suboptimal solutions to the standard $\ell_1$ problem. The new approach has several advantages over existing methods. It provides converging upper and lower bounds to the optimal $\ell_1$ cost, and it guarantees the convergence in norm to the optimal $\ell_1$ solution of a subsequence of suboptimal solutions. In contrast with the delay augmentation method, the norm convergence

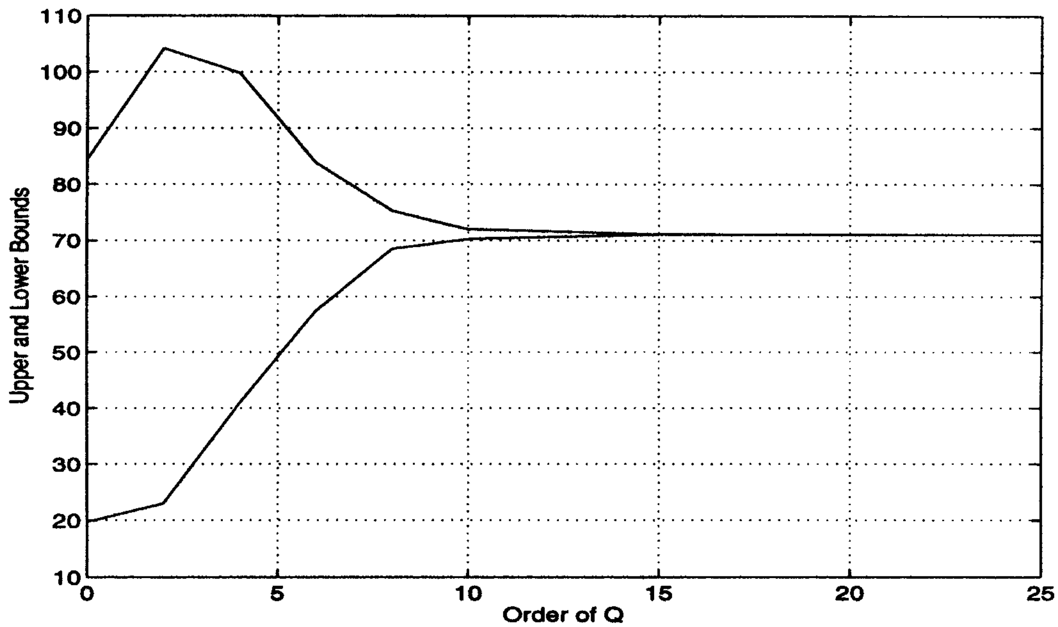


Fig. 3. Convergence of the upper and lower bounds to the optimal $\ell_1$ cost.

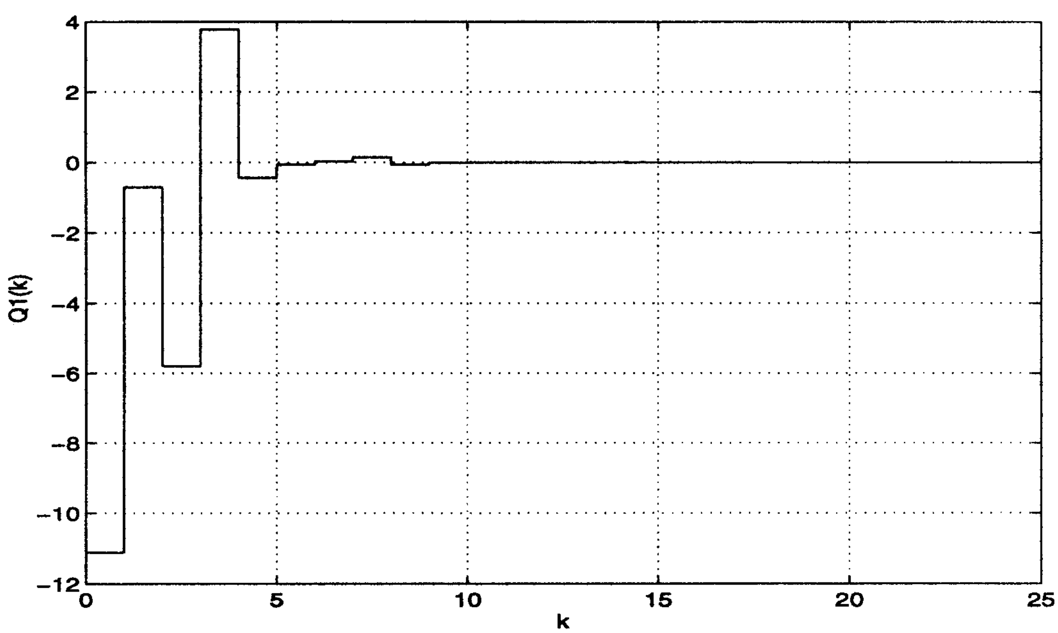


Fig. 4. Coefficients of $Q_1$ for $N = 26$.

is guaranteed independently of the inputs and outputs order. Moreover, the suboptimal $Q$ is directly computed, and therefore, the computation of the associated controller is much simplified in comparison with DA or any known closed-loop approximation method. Another computational advantage is that the computations of the interpolation conditions are completely avoided.

For each $N$, we have to solve a semidefinite quadratic programming problem instead of a linear programming problem. The complexity of these two problems is not fundamentally different. While the number of constraints is approximately the same for the two problems, the number of variables in the new method is greater than the number of variables used by DA by approximately $n_u \times n_y \times N$ (the number of elements in $Q$). However, for a fair and realistic comparison of the complexity of the two methods, we must include the complexity of the computation of the interpolation conditions required by DA method. Moreover, for the DA method, we must also include the overhead of computing the suboptimal $Q$. To perform such computations reliably, one must run a model reduction procedure, which is usually a time-consuming operation even for relatively small problems.

One property of the DA method is that for certain problems the method captures the structure of the optimal solution in a finite number of steps. In such cases, the order of the optimal control is derived from the sequence of suboptimal controllers. Unfortunately, there does not exist a characterization of these problems in terms of the problem data which limits the utility of this property. In addition, it may be necessary to reorder the inputs and outputs in the DA method to derive this information. The method presented in this paper does not possess this property. However, the norm convergence result may give an indication to the proper ordering for the DA problem. Deriving an algorithm that combines DA with the suggested algorithm is currently under investigation.

Finally, it is possible to extend the method in this paper to compute solutions to other mixed objective problems such as the $\ell_1$ problem with amplitude constraints on the response to fixed inputs.

## Appendix

### A. Proof of P3)

Here we show that for any $N$, Problem (2) is equivalent to a finite-dimensional convex optimization in the first $N$ elements of the impulse response of $Q$. This result follows from a property of the optimal $\mathcal{H}_2$ controller.

Consider the following problem:

$$\mu_N^2 = \inf_{\substack{\text{subject to:}\\ \Phi_1 = H_1 - U_1 Q_1 V \\ \Phi_2 = H_2 - U_{12} Q_1 V - U_2 Q_2 V \\ Q_1 \in \mathbb{R}^{n_u \times n_y \times N}, Q_2 \in \ell_1^{n_u \times n_y}}} \|\Phi_1\|_1^2 + \|\Phi_2\|_2^2. \tag{14}$$

Notice that $Q_1$, $Q_2$ are minimizing solutions to Problem (14) if and only if they are minimizing solutions to Problem (2). Since Problem (14) is jointly convex in $Q_1$ and $Q_2$, it can be rewritten as follows:

$$\mu_N^2 = \inf_{\substack{\Phi_1 = H_1 - U_1 Q_1 V \\ Q_1 \in \mathbb{R}^{n_u \times n_y \times N}}} \|\Phi_1\|_1^2 + \inf_{\substack{\Phi_2 = H_2 - U_{12} Q_1 V - U_2 Q_2 V \\ Q_2 \in \ell_1^{n_u \times n_y}}} \|\Phi_2\|_2^2. \tag{15}$$

Given any $N \geq 0$, define

$$f^N(Q_1) = \inf_{\substack{\Phi_2 = H_2 - U_{12} Q_1 V - U_2 Q_2 V \\ Q_2 \in \ell_1^{n_u \times n_y}}} \|\Phi_2\|_2^2. \tag{16}$$

Then the following is true.

*Theorem 7.1:* Consider the parameterization for $\Phi$, $\hat{\Phi} = \hat{H} - \hat{U}\hat{Q}\hat{V}$ where $\hat{U}$ and $\hat{V}$ are inner (i.e., $\hat{U}^{\sim}\hat{U}$ and $\hat{V}\hat{V}^{\sim}$ are stable with stable inverse) and $\hat{H}$ is $\mathcal{H}_2$ optimal or equivalently, $\hat{U}^{\sim}\hat{H}\hat{V}^{\sim} \in \mathcal{RH}_2^{\perp}$. A parameterization with these properties can be obtained from the model-based optimal $\mathcal{H}_2$ controller.

Then, for any $N \geq 0$

$$f^N(Q_1) = \|H_2 - U_{12} Q_1 V\|_2^2$$

i.e., the minimum is achieved at $Q_2 = 0$.

The result of this theorem implies that Problem (14) is equivalent to Problem (4).

*Proof:* Consider $\hat{\Phi}_2 = \hat{H}_2 - (\widehat{U_{12}Q_1})\hat{V} - \hat{U}_2\hat{Q}_2\hat{V}$, where $(\widehat{U_{12}Q_1})$ is the $\lambda$-transform of the sequence $U_{12}Q_1$.

$\hat{\Phi}_2$ can be rewritten as follows:

$$\hat{\Phi}_2 = \lambda^{-N}[\hat{H} - \hat{U}\hat{Q}\hat{V} - \hat{\Phi}_1]$$

where $\Phi_1 = H_1 - U_1 Q_1 V$. Notice that $\hat{\Phi}_1(\lambda)$ is a polynomial in $\lambda$ of order $N-1$. Assume without loss of generality that $\hat{U}^{\sim}\hat{U} = I$ and $\hat{V}\hat{V}^{\sim} = I$. Then we have that

$$\begin{aligned} &\hat{U}^{\sim}\hat{\Phi}_2\hat{V}^{\sim} \\ &\quad = \lambda^{-N}\hat{U}^{\sim}\hat{H}\hat{V}^{\sim} - \lambda^{-N}\hat{Q}_1 - \hat{Q}_2 - \lambda^{-N}\hat{U}^{\sim}\hat{\Phi}_1 V^{\sim}. \end{aligned}$$

All the terms in the right-hand side of the above equation are in $\mathcal{RH}_2^{\perp}$ with the exception of $Q_2$. Thus we have that

$$\begin{aligned} &\|H_2 - U_{12}Q_1 - UQ_2V\|_2 \\ &= \|\hat{U}^{\sim}(\hat{H}_2 - (\widehat{U_{12}Q_1})\hat{V} - \hat{U}\hat{Q}_2\hat{V})\hat{V}^{\sim}\|_2 \\ &= \|\lambda^{-N}\hat{U}^{\sim}\hat{H}\hat{V}^{\sim} - \lambda^{-N}\hat{Q}_1 - \lambda^{-N}\hat{U}^{\sim}\hat{\Phi}_1 V^{\sim}\|_2 + \|\hat{Q}_2\|_2. \end{aligned}$$

Therefore, the minimum value of $f^N(Q_1)$ is achieved by $Q_2 = 0$. Thus

$$\min_{Q_2} \|H_2 - U_{12}Q_1 - UQ_2V\|_2 = \|H_2 - U_{12}Q_1V\|_2. \quad \square$$

### B. Proof of Theorem 4.1

Consider the optimal solution, $x^o$, of the $\ell_1$ problem, which exists under standard assumptions. The sequence $\{\|x^o\|_N\}$ converges to $\mu^o$, and for each $N$, $\|x^o\|_N$ is an upper bound on $\mu_N$. Thus, all the limit points of $\{\mu_N\}$ will have values greater than zero and less than $\mu^o$. Consider any subsequence $\{\mu_{N_s}\}$, converging to one of the limit points. To simplify the notation, we remove the subindex and denote the subsequence by $\{\mu_N\}$. Let $\mu^*$ be its limit. We are going to show that $\mu^* = \mu^o$. Fix any $\epsilon > 0$. Since $\{\|x^o\|_N\}$ converges to $\mu^o$, there exists an integer $M_1$ such that, for all $N \geq M_1$, $|\mu^o - \|x^o\|_N| < \epsilon$. Consider Problem (9). It is well known that there are finite support feasible dual sequences $z^*$ whose cost, $\mu$, is arbitrarily close to $\mu^o$, say $\mu^o - \mu < \epsilon_1$ for $\epsilon_1 < \epsilon/2$. Since $z^*$ is feasible, it is also true that $\|\mathcal{A}_{\text{feas}}^* z^*\|_\infty \leq 1$. Given $x^* = \mathcal{A}_{\text{feas}}^* z^*$, for any $N$, we have that

$$U_1^* P_N x^* V^* + U_{12}^* T_N x^* V^* = 0.$$

Now pick any $0 < \epsilon_2 < \min\{\epsilon/2, \mu\}$, and let $y^* = (1 - \epsilon_2/\mu)x^*$. Then $y^* \in \ell_2^{n_z \times n_w}$ and $\|y^*\|_\infty \leq 1 - \epsilon_2/\mu$. Moreover, there exists an $M_2$ large enough such that

$$\|P_N y^*\|_\infty^2 + \|T_N y^*\|_2^2 \leq 1$$

for all $N \geq M_2$. Hence, for all $N \geq M_2$, $y^*$ is feasible for Problem (6) with cost $\mu - \epsilon_2 < \mu_{M_2}$. Thus we have that, for all $N \geq \max\{M_1, M_2\}$

$$\mu^o - \mu_N \geq \mu^o - \|x^o\|_N > -\epsilon$$

and

$$\mu^o - \mu_N < \mu^o - \mu + \epsilon_2 < \epsilon_1 + \epsilon_2 < \epsilon.$$

Therefore

$$|\mu^o - \mu_N| < \epsilon \qquad \text{for all } N \geq \max\{M_1, M_2\}.$$

Thus, for any $\epsilon > 0$ there exists an $M$ such that $|\mu^o - \mu_N| < \epsilon$ for all $N \geq M$; hence, $\mu_N$ converges to $\mu^o$. Since any convergent subsequence converges to $\mu^o$, the whole sequence must converge to $\mu^o$, and the result is proved. □

### C. Proof of Theorem 4.2

For each $N$, let $\Phi^N = [\Phi_1^N, \Phi_2^N]$ be the optimal primal solution and $x^* = [x_1^*, x_2^*]$ the optimal dual solution. We know that they are aligned. Let $\mu_{1N} = \|\Phi_1^N\|_1$, $\mu_{2N} = \|\Phi_2^N\|_2$, $\gamma_1^N = \|x_1^*\|_\infty$, and $\gamma_2^N = \|x_2^*\|_2$. The alignment condition implies that

$$\begin{aligned} \gamma_1^N &= \frac{\mu_{1N}}{\mu_N} \\ \gamma_2^N &= \frac{\mu_{2N}}{\mu_N}. \end{aligned} \tag{17}$$

Moreover, $\Phi_2^N = \mu_N x_2^*$. Thus, $\|\Phi_1^N\|_2$ goes to zero if and only if $\gamma_2^N$ goes to zero.

To derive a contradiction, assume that $\gamma_2^N$ is not converging to zero. This implies that there exist a positive constant $1 \geq a > 0$ such that for any positive integer $M$ it is possible to find some $N_M \geq M$ for which

$$\gamma_2^{N_M} \geq a.$$

For simplicity, relabel the sequence $\gamma_2^{N_M}$ as $\gamma_2^N$. Then, from Lemma 4.1, also the sequence $\{\|\Phi_2^N\|_1\}$ is uniformly bounded by some constant $\alpha < \infty$.

Since $\gamma_2^N \geq a$ for all $N$, we have that $\gamma_1^N \leq \sqrt{1-a^2} < 1$ for all $N$.

Thus, from the alignment condition in (17), we have that $\|\Phi_1^N\|_1 \leq \mu_N\sqrt{1-a^2}$. The convergence of $\mu_N$ implies that the sequence $\{\|\Phi_1^N\|_1\}$ is uniformly bounded by some positive constant $b$. Thus, the sequence $\{\|\Phi^N\|_1\}$ is uniformly bounded because $\|\Phi^N\|_1 \leq |\Phi_1^N\|_1 + \|\Phi_2^N\|_1$.

It follows from the Banach Alouglu theorem that there is a subsequence, $\{\Phi^{N_s}\}$, converging weak* to some element $\Phi^{w^*}$.

We claim that $\|\Phi^{w^*}\|_1 \leq \mu^o\sqrt{1-a^2}$.

For each $N$, $\Phi_1^N \in \mathbb{R}^{n_z \times n_w \times N}$. $\Phi_1^N$ can be seen as an element of $\ell_1^{n_z \times n_w}$ by considering it as an FIR sequence in $\ell_1^{n_z \times n_w}$. We still denote this extension as $\Phi_1^N$. For any $\epsilon > 0$, there exists an $N_0$ such that $\|\Phi_1^{N_s}\|_1 \leq (\mu^o + \epsilon)\sqrt{1-a^2}$ for all $N_s \geq N_0$. Thus, for $N_s \geq N_0$ the sequence $\{\Phi_1^{N_s}\}$ contains a weak* convergent subsequence $\{\Phi_1^{N_{sr}}\}$. Let $\Phi_1^{w^*}$ denote its weak* limit point. Then $\Phi_1^{w^*} = \Phi^{w^*}$, since the sequence $\{\Phi_1^{N_{sr}} - \Phi^{N_{sr}}\}$ is weak* convergent to the zero element. Moreover, we have that

$$\|\Phi^{w^*}\|_1 \leq (\mu^o + \epsilon)\sqrt{1-a^2}.$$

Since $\epsilon$ can be arbitrarily small, we have that $\|\Phi^{w^*}\|_1 \leq \mu^o\sqrt{1-a^2}$.

We now show that $\Phi^{w^*}$ is a feasible solution to the $\ell_1$ problem. This immediately implies that $\mu^o$ is not the optimal cost, and this contradiction will prove the assertion of the theorem.

Let $R: \ell_1^{n_u \times n_y} \to \ell_1^{n_z \times n_w}$ be the linear operator mapping $Q$ to $RQ = UQV$. Under the current assumptions, $R$ is one-to-one with closed range in $\ell_1^{n_z \times n_w}$. The proof of the above statement is left to the reader.

Let $Y^N = RQ^N = H - \Phi^N$. Notice that the $\ell_1$ norm of $Y^N$ is uniformly bounded since $\|Y^N\|_1 \leq \|H\|_1 + \|\Phi^N\|_1$ and $\|\Phi^N\|_1$ is converging to some $\mu \leq \mu^o\sqrt{1-a^2}$.

From [10, Lemma 1, pp. 155], we have that if $R$ has closed range, then there is a positive constant $k$ such that for any $Y$ in the range of $R$, there is a $Q$ satisfying $Y = RQ$, with $\|Q\|_1 \leq k\|Y\|_1$. Thus, if we consider the sequence $\{Q^N\}$ of optimal $Q$'s, we have also that the sequence of norms, $\{\|Q^N\|_1\}$ is uniformly bounded. Therefore, it contains a subsequence which is weak* convergent to some $Q^{w^*} \in \ell_1^{n_u \times n_y}$, and moreover, $\Phi^{w^*} = H - UQ^{w^*}V$ since the sequence $\{\Phi^N - (H - UQ^N V)\}$ is weak* convergent to zero. Summarizing, we have a feasible solution $\Phi^{w^*}$ with $\|\Phi^{w^*}\|_1 \leq \mu^o\sqrt{1-a^2} < \mu^o$. But this is impossible; hence, $\|\Phi_2^N\|_2$ must go to zeros for $N \to \infty$. □

## ACKNOWLEDGMENT

The authors wish to thank Prof. A. E. Barabanov for suggesting a shorter proof to Theorem 7.1.

## REFERENCES

[1] A. E. Barabanov and A. A. Sokolov, "Geometrical approach to the $\ell_1$ optimization problem," in *Proc. 33rd Conf. on Decision and Control*, Florida, 1994, vol. 4, pp. 3143–3144.
[2] F. Blanchini and M. Sznaier, "Persistent disturbance rejection via static feedback," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1127–1131, 1995.
[3] S. P. Boyd and C. H. Barratt, *Linear Controller Design: Limits of Performance*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[4] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Studies in Applied Mathematics, vol. 15. Philadelphia, PA: SIAM, 1994.
[5] T. Chen and B. Francis, *Optimal Sampled-Data Control Systems*. London, U.K.: Springer-Verlag, 1995.
[6] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[7] I. J. Diaz-Bobillo and M. A. Dahleh, "Minimization of the maximum peak-to-peak gain: The general multiblock problem," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1459–1483, 1993.
[8] N. Elia and M. A. Dahleh, "Controller design with multiple objectives," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 596–613, May 1997.
[9] ——, "Minimization of the worst-case peak to peak gain via dynamic programming: State feedback case," in *Proc. Amer. Control Conf.*, Albuquerque, NM, 1997, vol. 4, pp. 3590–3595.
[10] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
[11] M. Sznaier, "Mixed $\ell_1/\mathcal{H}_\infty$ controllers for MIMO discrete-time systems," in *Proc. 33rd Conf. on Decision and Control*, Florida, 1994, pp. 3187–3191.
[12] M. Sznaier and J. Bu, "A solution to MIMO 4-block $\ell_1$ optimal control problems via convex optimization," in *Proc. Amer. Control Conf.*, 1995, pp. 951–955.
[13] M. Khammash, "Solution of the $\ell_1$ MIMO control problem without zero interpolation," in *Proc. 33rd Conf. on Decision and Control*, Kobe, Japan, 1996, vol. 4, pp. 4040–4045.
[14] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*. Cambridge, MA: MIT, 1985.
[15] P. M. Young and M. A. Dahleh, "Infinite dimensional convex optimization and robust control," in *IEEE Trans. Automat. Contr.*, to be published.
[16] J. S. Shamma, "Nonlinear state feedback for $\ell_1$ optimal control," *Syst. and Contr. Lett.*, pp. 265–270, Oct. 1993.
[17] ——, "Optimization of the $\ell_\infty$ induced norm under full state feedback," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 533–544, Apr. 1996.
[18] P. G. Voulgaris, "Optimal $\mathcal{H}_2/\ell_1$ control via duality theory," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1881–1888, Nov. 1995.
[19] M. V. Salapaka, M. Dahleh, and P. G. Voulgaris, "Mixed objective control synthesis: Optimal l1/H2 control,'' *SIAM J. Contr. and Optimiz.*, vol. 35, no. 5, p. 1672, 1997.

[20] K. Zhou, *Robust and Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

**Nicola Elia** received the Laurea degree in electrical engineering from Politecnico of Turin in 1987 and the Ph.D. degree in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge, in 1996.

He held a Control Engineer position at Fiat Research Center from 1987 to 1990. He is currently a Postdoctoral Associate at the Laboratory for Information and Decision Systems at MIT. His research interests include computational methods for controller design, control of hybrid systems, control oriented system identification, and control with communication constraints.

**Munther A. Dahleh** (S'84–M'87–SM'97) was born in 1962. He received the B.S. degree from Texas A&M University, College Station, in 1983 and the Ph.D. degree from Rice University, Houston, TX, in 1987, both in electrical engineering.

Since then, he has been with the Department of Electrical Engineering and Computer Science, MIT, where he is now an Associate Professor. He was a visiting Professor at the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, for the Spring of 1993. He has held consulting positions with several companies in the United States and abroad. His interests include robust control and identification, the development of computational methods for linear and nonlinear controller design, and applications for feedback control in several disiplines including modeling of biological systems.

Dr. Dahleh has been the recipient of the Ralph Budd Award in 1987 for the best thesis at Rice University, George Axelby Outstanding Paper Award (paper coauthored with J. B. Pearson in 1987), an NSF presidential young investigator award (1991), the Finmeccanica Career Development Chair (1992), and the Donald P. Eckman award from the American Control Council in 1993. He was a plenary speaker at the 1994 American Control Conference. He is currently serving as an Associate Editor for IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is the coauthor (with Ignacio Diaz-Bobillo) of the book *Control of Uncertain Systems: A Linear Programming Approach* (Englewood Cliffs, NJ: Prentice-Hall).