Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99) Haifa, Israel - June 28-30, 1999

Computation of ℓ_1 Optimal Controllers using $\mathcal{H}_2$ Projections

Héctor Rotstein*
Rafael
and
Dept. of Electrical Engineering†
Technion — Israel Institute of Technology
Haifa 32000, Israel

Alfredo Desages
Dept. of Electrical Engineering
Universidad Nacional del Sur
8000 Bahia Blanca, Argentina

Abstract

Although the ℓ_1 or peak-to-peak norm can be used to capture a number of desirable closed-loop specifications, few practical applications of the criterion have been reported to date. Arguably, the single main reason for this is that, unlike the situation with the close $\mathcal{H}_2$ and $\mathcal{H}_\infty$ relatives, efficient numerical algorithms for ℓ_1 are still not available. The purpose of the present paper is to present an algorithm for computing sub-optimal ℓ_1 controllers using sequential $\mathcal{H}_2$ projections. As opposed to previous approaches, the algorithm does not use interpolations constraints nor attempts to solve an infinite optimization problem via finite approximation. Instead, sequential projections onto convex sets are performed to decide whether a given sub-optimal ℓ_1-norm level can be achieved or not. The present algorithm has several key advantages over previous methods:

1. At each stage, a finite optimization problem must be solved. This finite dimensionality is not due to truncation but results from the *exact* application of the algorithm.
2. The finite optimization problems are $\mathcal{H}_2$ projections and can be solved efficiently.
3. The approach does not rely on interpolation constraints. The same algorithm that works for the simplest version of the ℓ_1 problem (e.g., 1-block, SISO), can be modified in a straightforward manner to yield a solution to the general linear time-invariant case (e.g., 4-block, MIMO).

1 Introduction

The nominal ℓ_1 performance problem consists of minimizing the maximum amplitude of the regulated output of a closed-loop system, under the assumption that the exogenous input is bounded in amplitude by 1 but otherwise unknown. More specifically, Fig. 1 shows the interconnection of a plant P, assumed to be linear, time-invariant and discrete time, with a controller K. The n_i dimensional vector signal w entering into the system verifies the condition $|w_j(t)| \leq 1$ for $j = 1, \cdots, n_i$ and every positive t, but is otherwise unknown. A controller K is said to have an ℓ_1 performance level of γ if the n_o dimensional controlled signal z verifies $|z_j(t)| \leq \gamma$ for $j = 1, \cdots, n_o$ and every positive t. Optimality is obtained when γ is the smallest possible number achievable by a controller.

The ℓ_1 criterion was first proposed by Vidyasagar (1986) and solved for the SISO and MIMO one-block case in (Dahleh and Pearson, 1986, 1987) respectively. The criterion can be used to

*E-mail: *hector@ee.technion.ac.il*
†All correspondence should be sent to this address

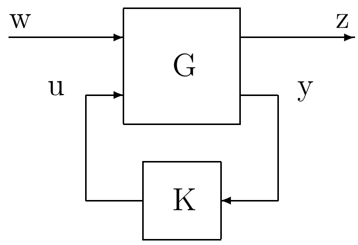

Figure 1: Plant-controller interconnection

impose a natural performance specification on numerous practical systems (Elia and Dahleh, 1994). For instance, ℓ_1 can be applied in a natural manner for designing active vision control systems (Rivlin *et al.*, 1997; Rivlin and Rotstein, 1995). Motivated by its practical and theoretical interest, several researchers have investigated ℓ_1 optimization and produced a considerable body of relevant results (Dahleh and Pearson, 1987; Mendlovitz, 1989; Diaz-Bobillo and Dahleh, 1993; Staffans, 1993)). The book (Dahleh and Diaz-Bobillo, 1994) contains a good introduction to the subject, and a description of the main results until 1995. In spite of this activity, to the best of the authors' knowledge, few applications of ℓ_1 control have been reported to date. Arguably, the single main reason for this lack is that the algorithms proposed for solving ℓ_1 are not as attractive as the ones developed, e.g., for the close $\mathcal{H}_2$ and $\mathcal{H}_\infty$ relatives. Original methods for solving ℓ_1 optimization (Dahleh and Pearson, 1987; Diaz-Bobillo and Dahleh, 1993; Staffans, 1993) were based on using the interpolation constraints for characterizing the set of all closed-loop transfer matrices resulting from stable controllers. Since this characterization results on a system of linear equality constraints, and the ℓ_1 norm can be formulated as a set of linear inequalities, the problem can be transformed into an infinite dimensional linear programming optimization. In a remarkable work, Dahleh and Pearson (1986, 1987) showed that for the simplest so-called "one-block" instance, the problem can be *exactly* reduced to finite dimensional. Unfortunately this property does not hold true for the general "four-block" case; consequently, the problems needs to be truncated for computing a numerical solution. The resulting algorithms use truncation for computing upper bounds on the optimal performance, and invoque duality arguments to generate lower bounds. The difference between the two bounds can be used as a stopping criterion. Alternatively, delay-augmentation can be employed (Dahleh and Diaz-Bobillo, 1994).

The combination of interpolation constraints and converging optimization problems proves to be rather cumbersome to implement and prone to numerical difficulties. This has motivated researches to look for alternative solutions. Two new approaches (Khammash, 1996; Elia and Dahleh, 1994), have been recently introduced which eliminate the interpolation constraints so as to partially remedy the drawbacks mentioned above. In (Khammash, 1996), a solution is computed considering the linear programming problem resulting from the direct truncation of the "Q" parameter of a Youla parameterization. In (Elia and Dahleh, 1996) this parameter is also truncated, but a quadratic programming problem is solved.

The purpose of the present paper is to present an algorithm for computing sub-optimal ℓ_1 controllers based on a significantly different idea. As opposed to the methods mentioned above, the algorithm is not based on interpolations constraints or increasing truncation for convergence. Rather, the solution is found by sequentially computing the ℓ_2 projections onto convex sets, as described in the sequel. The main features of this approach are:

1. At each stage, a finite optimization problem must be solved. This finite dimensionality is

not due to truncation but results from the *exact* application of the algorithm.

2. The finite optimization problems are $\mathcal{H}_2$ projections and can be solved efficiently.

3. The approach does not rely on interpolation constraints. The same algorithm that works for the simplest version of the ℓ_1 problem (e.g., 1-block, SISO), can be modified in a straightforward manner to yield a solution to the general linear time-invariant case (e.g., 4-block, MIMO).

2 Iterative Solutions for Induced-norm Problems

The starting point of the present work is the interesting work by Sideris (1990) and Kavranoglu and Sideris (1989), concerning the computation of a solution to $\mathcal{H}_\infty$ optimal control problem by considering a sequence of $\mathcal{H}_2$ optimizations. The idea presented in these papers is to solve a sequence of *weighted* $\mathcal{H}_2$ problems, where the weights are also sequentially adjusted, until convergence to a solution of the original $\mathcal{H}_\infty$ problem is achieved. The success of the approach hinges upon the following two facts: (a) the $\mathcal{H}_2$ problem can be efficiently solved; and (b) calculations are performed by manipulating finite dimensional systems.

The following example shows that a similar approach can, in principle, be used for solving ℓ_1 optimization. Suppose that the interconnection transfer matrix $G(z)$ in Fig. 1 has two inputs and two outputs so that $K(z)$ is single input-single output. As discussed in (Francis, 1987), introducing the parameterization of all stabilizing controllers, the resulting ℓ_1 problem can be formulated as

$$\min_{\mathbf{q}\ \text{stable}} \|\mathbf{t}_1 - \mathbf{t}_2\mathbf{q}\|_1$$

where $\mathbf{t}_1,\ \mathbf{t}_2$ are stable transfer functions constructed from the problem data. As mentioned above, (Dahleh and Pearson, 1987) proved that this instance of the problem can be solved elegantly by using linear programming, but unfortunately the extension to the general four-block case has proved elusive. Instead, an algorithm inspired in (Kavranoglu and Sideris, 1989) would proceed as follows:

1. Set k=0, and pick an initial (stable, with stable inverse) weight $\mathbf{w}^k(z)$.

2. Solve the optimization
$$\min_{\mathbf{q}\ \text{stable}} \|\mathbf{w}^k \star (\mathbf{t}_1 - \mathbf{t}_2\mathbf{q})\|_2.$$
Let $\mathbf{q}^k(z)$ denote the solution to this problem.

3. Select $\mathbf{w}^{k+1}$ based on the transfer function $\mathbf{w}^k \star (\mathbf{t}_1 - \mathbf{t}_2\mathbf{q}^k)$.

4. Increase k and repeat until convergence.

Notice that the precise meaning of the multiplication "$\star$" has not been given yet. In the $\mathcal{H}_\infty$ problem considered in (Kavranoglu and Sideris, 1989), this product is the standard transfer function multiplication, and the weighted function $\mathbf{w}^{k+1}$ is computed by solving a spectral factorization problem. A proper choice of initial solution coupled with an analytical model simplification step, leads to a state-space representation of the relevant transfer functions with a fixed number of states. The iterations can then be performed efficiently without state inflation. Unfortunately, no such nice structure has been found for the problem at hand. Indeed, if

$$\mathbf{s}^k = \mathbf{t}_1 - \mathbf{t}_2\mathbf{q}^k$$

and

$$\mathbf{s}^k(z) = \sum_{i\geq 0} s^k(i)z^{-i},$$

then the weight $\mathbf{w}^{k+1}(z) = \sum_{i\geq 0} w^{k+1}(i)z^{-i}$, can be computed using (Sideris and Rotstein, 1996):

$$w^{k+1}(i) = \begin{cases} w^k(i)/|s^k(i)| & s^k(i) \neq 0 \\ 0 & s^k(i) = 0 \end{cases}$$

In this case, the weight should be applied in the *time domain*, i.e., "$\star$" takes the form of point-wise multiplication in the time domain. By working out a simple example, it is easy to see that after the first iteration, the weight may easily become infinite dimensional! A different approach may be followed by formulating the problem as a linear programming optimization (Dahleh and Pearson, 1987), and then solving this problem iteratively by using an interior-point method. For instance, on each iteration one could solve a quadratic programming problem connected with $\mathcal{H}_2$ optimization. Unfortunately, this also gives rise to infinite dimensional systems after the first iteration. Moreover, the two methods outlined do not have a straightforward generalization to the four-block case.

2.1 Solution via Projections onto Convex Sets

Alternatively, the ℓ_1 suboptimal control problem can be formulated as one of finding an intersection point of two convex sets or showing that the intersection is empty. Indeed, consider the following two convex sets: S_2, consisting of all closed-loop transfer matrices resulting from stabilizing controllers, and $S_1(\gamma)$ consisting of all stable transfer matrices with ℓ_1 norm less than or equal to the performance level γ. Finding whether there exists a stabilizing controller achieving the performance level γ is equivalent to verifying whether $S_1 \cap S_2 \neq \emptyset$. By including an Euclidean structure into the space considered, this problem can in turn be solved by using a procedure called *Alternative Projections* (Bauschke and Borwein, 1993) or *Projection Onto Convex Sets* (POCS) (Bauschke and Borwein, 1996) which basically consists on projecting an original point recursively onto S_1 and S_2. Under suitable assumptions, the sequence of points generated by the algorithm will either converge to points in S_1 and S_2 lying at minimum distance or will have no accumulation point. Whenever a suboptimal controller achieving the desired performance exists, convergence is to a point in the intersection of S_1 and S_2, i.e., to a solution of the problem. The tutorial (Bauschke and Borwein, 1996) contains a detailed exposition of known and new results, including algorithms and convergence proofs for a much general instance of the problem. AP algorithms have been applied successfully to solve problems in numerous fields, including statistics, image reconstruction and processing, electron microscopy and image processing.

AP methods have also been exploited for solving the $\mathcal{H}_2/\mathcal{H}_\infty$ control problem in (Frazho *et al.*, 1995; Halikias *et al.*, 1997), in a formulation closely connected with the approach in the present work (compare their convex sets). However, there is an essential difference between their approach to $\mathcal{H}_2/\mathcal{H}_\infty$ and the current to ℓ_1. Namely, Frazho *et al.* (1995); Halikias *et al.* (1997) seek an approximate solution by considering a finite dimensional approximation to $\mathcal{H}_2/\mathcal{H}_\infty$ which allows for a numerical implementation. In contrast, the present work concentrates in solving the original infinite-dimensional ℓ_1 problem.

The AP or POCS algorithm can be used in principle to solve a large class of optimization problems. However, in order for the algorithm to be useful in practice, the projections onto the convex sets must be easy and inexpensive to compute. Also, finite-dimensionality should be maintained through the iterations, to guarantee that computations can be performed without having to resort to approximations. The main contribution of this work is to show that, in the case under consideration, this two properties which are crucial for the success of an iterative method, actually hold.

The algorithm in this paper has several advantages over previous methods for computing ℓ_1 controllers:

1. Truncation of the "Q" parameter is not required; rather, finite dimensional problems appear as a natural consequence of the approach.

2. The optimization problems are $\mathcal{H}_2$ projections, which are easy to implement and can be solved efficiently.

3. The approach does not rely on interpolation constraints. The same algorithm that works for the simplest version of the ℓ_1 problem (e.g., 1-block, SISO), can be modified in a straightforward manner to yield a solution to the general linear time-invariant case (e.g., 4-block, MIMO).

3 Preliminaries and Problem Setup

In this section the notation used in the paper is established and the problem setup is presented. The space $\ell_1^{m\times n}$ is the set of all sequences of $m \times n$ real matrices such that $\|\mathbf{H}\|_1 < \infty$, where

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

Similarly the $\ell_2^{m\times n}$ space is characterized by $\|\mathbf{H}\|_2 < \infty$ where

$$\|\mathbf{H}\|_2^2 \doteq \sum_{t=0}^{\infty} \mathrm{trace}(H(t)H(t)^T) = \sum_{t=0}^{\infty} \sum_{i,j} \|h_{ij}(t)\|^2 .$$

In the paper, systems are represented by their transfer matrices, and denoted using bold letter (e.g., $\mathbf{x}$, $\mathbf{H}$). Whenever dimensions are clear from the context, the supra-indices $m \times n$ will not be indicated explicitly.

Systems considered in this paper are MIMO discrete-time, described by their transfer matrices $\mathbf{G}(z)$. Stable systems have a single sided expansion of the form

$$\mathbf{G}(z) = \sum_{t\geq 0} G(t) z^{-t}. \tag{3-1}$$

The notation in (3-1) is used throughout the paper: a system $\mathbf{G}$ is expanded using the Markov coefficients $G(t)$. The truncation operator $\mathcal{P}_k : \ell_2 \to \ell_2$ is defined by

$$\mathcal{P}_k \left(\sum_{t=0}^{\infty} G(t) z^{-t} \right) = \sum_{t=0}^{k-1} G(t) z^{-t}. \tag{3-2}$$

With a slight abuse of notation, norms on a stable system $\mathbf{G}(z)$ are defined via the corresponding one of the sequence in the expansion (3-1). Also, given $\mathbf{G}(z)$, the adjoint system is defined as $\mathbf{G}(z)^{\sim} \doteq \mathbf{G}(1/z)^T$.

The ℓ_1 problem considered in this paper is defined in terms of the generalized interconnection G illustrated in Fig. 1. Partitioning G compatible with the inputs $w,\ u$ and the outputs $z,\ y$,

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}.$$

Systems of interest are real-rational, with a state space representation (notice the notation)

$$G = \left(\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right). \tag{3-3}$$

The dimension of these matrices are $A \in R^{n\times n},\ B_1 \in \mathbb{R}^{n\times n_w},\ B_2 \in \mathbb{R}^{n\times n_u},\ C_1 \in \mathbb{R}^{n_z\times n},\ C_2 \in \mathbb{R}^{n_y\times n}$. The γ-suboptimal ℓ_1 problem consists on finding, whenever possible, an internally stabilizing controller K such that the closed loop map $\mathbf{T} = \mathbf{G}_{11} + \mathbf{G}_{12}K(I - \mathbf{G}_{22}K)^{-1}\mathbf{G}_{21}$ verifies $\|\Phi\|_1 \le \gamma$.

All closed loop maps resulting from stabilizing controllers can be written as (Elia and Dahleh, 1994)

$$\mathbf{T} = \mathbf{H} - \mathbf{UQV}$$

where $\mathbf{H},\ \mathbf{U}$ and $\mathbf{V}$ are stable systems computed from the problem data, and $\mathbf{Q} \in \ell_1^{n_u\times n_y}$ is the free parameter of the parameterization. Under mild assumption over the realization (3-3), the transfer matrices $\mathbf{U}$ and $\mathbf{V}$ can be chosen to be inner and co-inner respectively. Moreover, they can be completed to unitary transfer matrices, e.g., there exist $\mathbf{U}_\perp,\ \mathbf{V}_\perp$ such that:

$$\mathbf{T} = \mathbf{H} - [\mathbf{U}\ \mathbf{U}_\perp] \begin{bmatrix} \mathbf{Q} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{V}_\perp \end{bmatrix},$$

where $\mathbf{U}_a = [\mathbf{U}\ \mathbf{U}_\perp],\ \mathbf{V}_a = \begin{bmatrix} \mathbf{V} \\ \mathbf{V}_\perp \end{bmatrix}$ verify $\mathbf{U}_a^\sim \mathbf{U}_a = \mathbf{U}_a \mathbf{U}_a^\sim = I$ and $\mathbf{V}_a^\sim \mathbf{V}_a = \mathbf{V}_a \mathbf{V}_a^\sim = I$. With this selection,

$$\|\mathbf{T}\|_2 = \|\mathbf{H} - \mathbf{UQV}\|_2 = \|\mathbf{U}_a^\sim \mathbf{H} \mathbf{V}_a^\sim - \mathbf{Q}_a\|_2, \quad \mathbf{Q}_a = \begin{bmatrix} \mathbf{Q} & 0 \\ 0 & 0 \end{bmatrix}.$$

By a standard projection argument, minimization of $\|\Phi\|_2$ is achieved by selecting the $1-1$ block of the stable part of $\mathbf{U}^\sim\mathbf{H}\mathbf{V}^\sim$. The special selection of the parameterization results on the optimal $\mathbf{Q}^*$ being a constant matrix, namely the feed-through term in a state-space realization of $\mathbf{U}^\sim\mathbf{H}\mathbf{V}^\sim$. The reader is referred to (Rotstein, 1992) for a general way of computing these transfer matrices.

4 Projections Algorithms for Convex Problems

This section contains a brief review on projections algorithms for solving convex feasibility problems; the reader is referred to (Bauschke and Borwein, 1993, 1996) for a complete discussion and a list of references. Let $\mathcal{X}$ be a possibly infinite dimensional Hilbert space and S_i, $i = 1,\ 2$, two closed convex subsets with possibly empty intersection S:

$$S = S_1 \cap S_2.$$

Given a convex set S_i and a point $x \notin S_i$, there exist a unique point $P_i(x) \in S_i$ such that $d(x, P_i(x)) = \inf_{s \in S_i} d(x, s)$, where $d(\cdot, \cdot)$ denotes the distance defined by the Hilbert space norm. The point $P_i(x)$ is called the projection of x onto S_i.

The convex feasibility problem is to find a point x in the intersection set S assuming this set is non-empty. Bauschke and Borwein (1993) suggested a generalization of the problem for the case of S possibly empty, which is followed in this paper. Let $d(S_1, S_2)$ denote the distance between S_1 and S_2:

$$d(S_1, S_2) \doteq \inf \{\|s_1, s_2\|_2 \text{ s.t. } s_1 \in S_1,\ s_2 \in S_2\}.$$

Then, the intersection $S = S_1 \cap S_2$ is generalized as

$$\begin{aligned} E &\doteq \{s_1 \in S_1 \text{ s.t. } d(s_1, S_2) = d(S_1, S_2)\}, \\ F &\doteq \{s_2 \in S_2 \text{ s.t. } d(S_1, S_2) = d(S_1, S_2)\}. \end{aligned}$$

Notice that if $S_1 \cap S_2 \neq \emptyset$ then $S_1 \cap S_2 = E = F$. As discussed in (Bauschke and Borwein, 1993), there exists a unique vector called the *displacement vector* v such that $\|v\| = d(S_1, S_2)$ and $E + v = F$.

In general, the convex feasibility problem must be solved iteratively. An algorithm often used for solving this problem is based on successfully projecting a point onto the two convex sets, which can be traced back to the work of von Neumann (1950). The general projection algorithm works as follows. Given the current iterate $x^{(n)}$, the next iterate $x^{(n+1)}$ is constructed by setting

$$x^{(n+1)} = \left(\sum_{i=1}^{2} \lambda_i^{(n)} \left[(1 - \alpha_i^{(n)}) I + \alpha_i^{(n)} P_i \right] \right) x^{(n)}. \tag{4-1}$$

The coefficients $\alpha_i^{(n)} \in [0, 2]$ are relaxation parameters, and the $\lambda_i^{(n)}$ are positive weights verifying $\lambda_1^{(n)} + \lambda_2^{(n)} = 1$. According to the way these parameters are chosen, the algorithm takes several special forms. In particular, the selection of interest in the present paper is

$$\alpha_i^{(n)} = 1 \tag{4-2}$$

and

$$\begin{aligned} \lambda_1^{(n)} &= \begin{cases} 1. & n \text{ odd} \\ 0. & n \text{ even} \end{cases} \\ \lambda_2^{(n)} &= \begin{cases} 0. & n \text{ odd} \\ 1. & n \text{ even} \end{cases} \end{aligned} \tag{4-3}$$

For this case, the algorithm can be described as:

$$b_0 \doteq x, \qquad a_n \doteq P_1(b_{n-1}), \qquad b_n \doteq P_2(a_{n-1}),$$

where (a_n), (b_n) are called "von Neumann sequences." The following result holds for the very general version of the problem considered so far.

Lemma 1

1. *$b_n - a_n,\ b_n - a_{n+1} \to v$.*
2. *If $E,\ F$ are nonempty, then $a_n \to e^* \in E$ and $b_n \to f^* = e^* + v \in F$, where the convergence is weak.*

Recall that $v = 0$ if S_1 and S_2 have a nonempty intersection.

Proof. See Fact 1.2 in (Bauschke and Borwein, 1993) □

To establish convergence in norm, additional properties on the convex sets are required. For instance, consider the following definition.

Definition 1
The N-tuple of closed convex sets $(S_1,\ S_2, \cdots,\ S_N)$ is regular *if (Bauschke and Borwein, 1996)*

$$\underset{\epsilon > 0}{\forall} \quad \underset{\delta > 0}{\exists} \quad \underset{\substack{x \in X \\ \max\{d(x, S_j) : j = 1, \cdots, N\} \leq \delta}}{\forall} \quad d(x, S) \leq \epsilon.$$

If this property only holds on bounded sets, then the N-tuple is called boundedly regular*.*

The intuition behind this definition is clear: if a point lies close to each set of a (boundedly) regular N-tuple, then it cannot be far away from its intersection. Under additional mild assumption, it is possible to prove (see Theorem 5.2 in (Bauschke and Borwein, 1996)), that boundedly regularness is a sufficient condition for the convergence of POCS in norm. A stronger property, namely linear regularity, will give rise to stronger convergence results.

Definition 2
The N-tuple of closed convex sets $(S_1,\ S_2, \cdots,\ S_N)$ is linearly regular *if (Bauschke and Borwein, 1996)*

$$\underset{\kappa > 0}{\exists} \quad \underset{x \in X}{\forall} \quad d(x, S) \leq \kappa \max\{d(x, S_j) : j = 1, \cdots, N\}.$$

If this property only holds on bounded sets, then the N-tuple is called boundedly linearly regular*.*

As reviewed bellow, linear regularity provides convergence in the following sense.

Definition 3
A sequence $\{x^{(n)}\}$ converges **linearly** *to x if for some $\rho \in [0, 1],\ M > 0$, then $\|x^{(n)} - x\|_2 \leq M\rho^n$.*

Theorem 1

1. *If S_1 or S_2 is boundedly regular, then the von Neumann sequences converge in norm.*
2. *If S_1 or S_2 is linearly regular, then the von Neumann sequences converge linearly.*

Proof. See (Bauschke and Borwein, 1993). □

As shown in Section 6, the iteration proposed for computing a solution to the suboptimal ℓ_1 problem is based on the projection onto two linear regular sets; consequently Theorem 1 guarantees linear convergence.

5 Projections for ℓ_1 Optimization

The approach used in this paper is to use projections onto convex sets to solve the ℓ_1 suboptimal control problem. Projections are performed onto the two sets S_1 and S_2:

$$\begin{aligned} S_1(\gamma) &= \left\{\mathbf{S} \in \ell_1^{n_z \times n_w} \text{ s.t. } \|\mathbf{S}\|_1 \leq \gamma\right\} \\ S_2 &= \left\{\mathbf{S} = \mathbf{H} - \mathbf{UQV},\ \mathbf{Q} \in \ell_1^{n_u \times n_y}\right\}. \end{aligned}$$

Without loss of generality take $\gamma = 1$, which amounts to scaling the ℓ_1 performance objective, and call $S_1 \doteq S_1(1.)$. These sets are subset of the Hilbert space $\mathcal{X} = \ell_2^{n_z \times n_w}$, so that projections are to be computed using the distance induced by the 2-norm.

Lemma 2
The sets S_1 and S_2 are closed as subsets of ℓ_2.

Proof. That S_2 is closed is obvious. See the Appendix for a proof that also S_1 is closed. □
As could be expected, computing the projection onto S_2 is relatively straightforward, since it is equivalent to solving an optimal approximation problem on the Hilbert norm. The remarkable fact is that the projection onto S_1 can be characterized in a computational attractive manner.

5.1 Projection onto S_1

The SISO case $m = n = 1$ is discussed first. Let $\mathbf{x}$ be a real-rational stable transfer function, $\|\mathbf{x}\|_1 > 1$. The purpose of this section is to show how the projection $P_1(\mathbf{x})$ onto S_1 can be computed. Consider first the case where $\mathbf{x}$ has a finite expansion, i.e., $\mathbf{x}(z)$ is a FIR, of length k:

$$\mathbf{x}(z) = \sum_{i=0}^{k-1} x(i) z^{-i}.$$

The projection onto S_1 can be characterized as follows.

Theorem 2
Let $\mathbf{x}(z)$ be an FIR of length k, $\|\mathbf{x}\|_1 > 1$. If $\mathbf{s} = P_1(\mathbf{x})$, $\mathbf{s}(z) = \sum_{i=0}^{k-1} s(i) z^{-i}$, then there exists $\lambda > 0$ such that

$$s(i) = \begin{cases} 0 & \text{if } |x(i)| \leq \lambda/2 \\ x(i) - \lambda/2 & \text{if } |x(i)| > \lambda/2 \end{cases} \tag{5-1}$$

Proof. Without loss of generality, assume that $x(i) \geq 0$. Note that otherwise one can consider $\hat{\mathbf{x}}(z) = \sum_{i=0}^{k-1} |x(i)| z^{-i}$, and if $\hat{\mathbf{s}} = P_1(\hat{\mathbf{x}})$, then $\sum_{i=0}^{k-1} \text{sign}(x(i)) \hat{s}(i) z^{-i}$ is an optimal approximation to $\mathbf{x}$. Under this assumption, the Markov parameters of the projection $s(i)$ are also non-negative, and the computation of the projection can be formulated as:

$$\min \left\{ \sum_{i=0}^{k-1} (x(i) - s(i))^2 \text{ s.t. } \sum_{i=0}^{k-1} s(i) = 1,\ s(i) \geq 0 \right\}.$$

Introducing Lagrange multipliers:

$$L(s, \lambda, \mu) = \sum_{i=1}^{k-1} (x(i) - s(i))^2 + \lambda \left[\left(\sum_{i=1}^{k-1} s(i) \right) - 1 \right] + \sum_{i=0}^{k-1} \mu(i) s(i), \qquad \mu(i) \geq 0$$

The Kuhn-Tucker necessary conditions (Franklin, 1980) for optimality give

$$-2(x(i) - s(i)) + \lambda + \mu(i) = 0$$

together with

$$\mu(i) s(i) = 0,$$

for $i = 0, \cdots, k-1$. If $s(i) > 0$, then $\mu(i) = 0$, so that $x(i) - s(i) = \lambda/2$, as required by the Theorem. □

Corollary 1

Assume $\|\mathbf{x}\|_1 > 1$. The transfer function $\mathbf{s}$ is such that $\mathbf{s} = P_1(\mathbf{x})$ if and only if it verifies the condition (5-1) and also $\|\mathbf{s}\|_1 = 1$.

Proof. Notice that for every $\mathbf{s}$ verifying (5-1), then the norm $\|\mathbf{s}\|_1$ strictly decreases for increasing λ. Optimality is then achieved when $\|\mathbf{s}\|_1 = 1$. □

Theorem 2 and its Corollary can be used to formulate a numerical algorithm for computing the optimal projection in the FIR case. Indeed, one can select a value for λ, and compute an FIR verifying the conditions (5-1). The value of λ can then be adjusted iteratively, until the ℓ_1 constraint be verified.

Consider now the case with $\mathbf{x}$ real rational with an infinite expansion:

$$\mathbf{x}(z) = \sum_{i \geq 0} x(i) z^{-i}$$

The projection onto S_1 can be characterized as follows.

Theorem 3

Let $\mathbf{x}(z)$ be a stable transfer function, with $\mathbf{x}(z) = \sum_{i\geq 0} x(i)z^{-i}$ and $\|\mathbf{x}\|_1 > 1$. Let $\mathbf{s} = P_1(\mathbf{x})$. Then:

1. *The transfer function $\mathbf{s}$ is an FIR.*
2. *There exists a constant $\lambda > 0$ such that either $s(i) = 0$ or $|x(i) - s(i)| = \lambda/2$ (i.e., condition (5-1) holds).*

Proof. As in the proof of Theorem 2, assume without loss of generality that $x(i) \geq 0$. Let $\mathbf{x}^k = \mathcal{P}_k(\mathbf{x})$ be the truncation of $\mathbf{x}$ of length k, where k is selected so that $\|\mathbf{x}^k\|_1 > 1$. Since $\mathbf{x}^k$ is a FIR, the projection $\mathbf{s}^k = P_1(\mathbf{x}^k)$ can be characterized using the previous theorem:

$$s^k(i) = \begin{cases} 0 & \text{if } x(i) \leq \lambda^k/2 \\ x^k(i) - s^k(i) = \lambda^k/2 & \text{otherwise} \end{cases}$$

Since $\mathbf{s} \in S_1$ but $\|\mathbf{x}^k\|_1 > 1$, the constant λ^k verifies

$$\lambda^k/2 \geq \frac{\|\mathbf{x}^k\|_1 - 1}{k}.$$

It is claimed that the sequence $\{\lambda^k\}$ is non-decreasing. To see this take $k_1 < k_2$. Since the FIR $\mathcal{P}_{k_1}\left(P_1(\mathbf{x}^{k_2})\right)$ verifies the conditions (5-1) for x^{k_1}, the inequality $\lambda^{k_1} > \lambda^{k_2}$ would contradict Corollary 1. Since $\mathbf{x} \in \ell_2$, the coefficients $x(i) \to 0$ and consequently there exists K_1, K_2 such that $\mathbf{x}(k) < \lambda^{K_2}/2$ for each $k \geq K_1$. This implies that $s^l(k) = 0$ for each $k \geq K_1$ and each $l \geq K_2$, so that $\mathbf{s}^k = \mathbf{s}^{K_2}$, for each $k \geq K_2$. It is claimed next that $d\left(\mathbf{s}^{K_2}, \mathbf{x}\right) = d\left(S_2), \mathbf{x}\right)$, so that $\mathbf{s}^{K_2} = P_1(\mathbf{x})$. This claim can easily be established by using a contradiction argument. □

The construction in the proof can be implemented for computing the projection onto S_1. Indeed, if $\mathbf{x}$ is a stable, real-rational transfer function, the coefficients $x(i)$ in the series expansion converge to zero exponentially fast, i.e., there exist M and $\rho < 1$ such that $|x(i)| < M\rho^i$. Selecting a k such that $\|\mathcal{P}_k(\mathbf{x})\|_1 > 1$, compute a lower bound for the optimal λ and use the exponential upper bound for $x(i)$ to truncate $\mathbf{x}$. Now use the algorithm for projecting an FIR to conclude the computation.

The MIMO case

The extension from SISO to MIMO is relatively straightforward. By definition of the ℓ_1-norm, $\mathbf{s} \in \ell^{m\times n}$ if and only if

$$\sum_{j=1}^{n}\sum_{t=0}^{\infty}|s_{ij}(t)| \le 1, \quad \forall\, i = 1, \cdots, m.$$

Given $\mathbf{x} \in \mathcal{X}$, the square of the norm of $\mathbf{x}$ can also be computed "row-wise", namely

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n}\sum_{t=0}^{\infty}|x_{ij}(t)|^2\right),$$

so that minimizing the norm is equivalently to minimizing each row and then summing up. Consequently, the projection $P_1(\mathbf{x})$ can be computed by the m projections $P_1(\mathbf{x}_{i\cdot})$, where $\mathbf{x}_{i\cdot}$ represents the i-th row of $\mathbf{x}$. The projection of each row can be performed using the same algorithm described for the SISO case, with only minor changes to accommodate notation.

5.2 Projection onto S_2

Let $\mathbf{x} \notin S_2$. The projection onto S_2 consists on finding the point $\mathbf{x}_2 \in S_2$ such that $\|\mathbf{x}-\mathbf{x}_2\|_2 = \min_{\mathbf{s}_2\in S_2}\|\mathbf{x}-\mathbf{s}_2\|_2 = \min_{\mathbf{Q}\in\ell_2}\|\mathbf{x}-\mathbf{H}+\mathbf{UQV}\|_2$. From the material reviewed in Section 3, the solution to this problem is:

$$\begin{aligned}\mathbf{Q}^* &= \text{stable part}\left(\mathbf{U}^{\sim}(\mathbf{x}-\mathbf{H})\,\mathbf{V}^{\sim}\right)\\ &= \text{stable part}\left(\mathbf{U}^{\sim}\mathbf{x}\mathbf{V}^{\sim}\right) - D_Q,\end{aligned}$$

where D_Q is the stable part of $\mathbf{U}^{\sim}\mathbf{H}\mathbf{V}^{\sim}$, which is known to be constant.

Of special interest is the case when $\mathbf{x}$ is an FIR, $\mathbf{x} = \sum_{k=0}^{N-1} x(k)z^{-k}$. Since $\mathbf{U}$ and $\mathbf{V}$ are stable (Rotstein, 1992), they can be expanded as

$$\begin{aligned}\mathbf{U}(z) &= \sum_{t\ge 0} U(t)z^{-t}\\ \mathbf{V}(z) &= \sum_{t\ge 0} V(t)z^{-t},\end{aligned}$$

so that

$$\begin{aligned}\mathbf{U}(z)^{\sim} &= \sum_{t\ge 0} U(t)^T z^{t}\\ \mathbf{V}(z)^{\sim} &= \sum_{t\ge 0} V(t)^T z^{t}.\end{aligned}$$

Since U and V are stable, the transfer matrices $\mathbf{U}^{\sim}$, $\mathbf{V}^{\sim}$ are anti-stable (all poles outside the unit disk). Consequently, the stable part of $\mathbf{U}^{\sim}\mathbf{x}\mathbf{V}^{\sim}$ is also an FIR. Some tedious but otherwise straightforward calculations show that

$$\mathbf{Q}^*(z) = \begin{bmatrix} U(n-1)^T & \cdots & U(1)^T & U(0)^T \end{bmatrix}\begin{bmatrix} x(n-1) & 0 & \cdots & 0\\ x(n-2) & x(n-1) & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ x(0) & x(1) & \cdots & x(n-1)\end{bmatrix}$$

$$\begin{bmatrix} V(0)^T & \cdots & \cdots & 0 \\ V(1)^T & V(0)^T & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 \\ V(n-1)^T & \cdots & \cdots & V(0)^T \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(n-1)} \end{bmatrix} - D_Q. \tag{5-2}$$

6 Algorithms and Convergence

From the material in the previous section, the following von Neumann algorithm can be used to solve the suboptimal ℓ_1 problem.

Algorithm 1

1. *Set $k = 0$. Compute an initial stabilizing solution, e.g., the solution to the optimal ℓ_2 problem. Let $\mathbf{T}^0$ denote the corresponding closed loop transfer matrix.*
2. *Compute $\mathbf{S}^k = P_1\left(\mathbf{T}^k\right)$, by using the algorithm outlined above. Note that $\mathbf{S}^k$ is a FIR.*
3. *Compute $\mathbf{T}^{k+1} = P_2\left(\mathbf{S}^k\right)$ by first computing the optimal $\mathbf{Q}$ using (5-2).*
4. *Set $k = k + 1$ and repeat until the convergence of $\mathbf{S}^k - \mathbf{T}^k$.*

Whenever $S_1 \cap S_2 \neq \emptyset$, convergence of the algorithm is based on the following fact.

Lemma 3
The pair $(S_1,\ S_2)$ is linearly regular

Proof. See Appendix B. □
The main convergence result can hence be formulated as follows.

Theorem 4 (Convergence of the AP algorithm for ℓ_1 Optimization)

Suppose that a stabilizing controller K exists for the configuration in Fig. 1, such that the closed loop transfer matrix $\mathbf{T}$ verifies $\|\mathbf{T}\|_1 \leq 1$. Then Algorithm 1 converges linearly to a solution of the suboptimal ℓ_1 problem.

Proof. From the previous Lemma, and the fact that the algorithm is of von Neumann type, the hypotheses in Theorem 1 hold. Convergence follows. □
Suppose now that no internally stabilizing controller achieves the performance level 1. Then, from Lemma 1 the difference between successive projections onto S_1 and S_2 will converge to the displacement vector v, such that $\|v\|_2 = d(S_1, S_2)$. This fact can be used to establish the following lemma.

Lemma 4 *Suppose that $S_1 \cap S_2 = \emptyset$ so that the optimal ℓ_1 norm is larger than 1, and let $0 < \delta < d(S_1, S_2)$. Then, the optimal ℓ_1 norm is also larger than $1 + \frac{\delta}{\sqrt{m}}$.*

Proof. It will be shown that if δ is as in the Lemma, then $S_1(1 + \delta/\sqrt{m}) \cap S_2 = \emptyset$. To see this, define

$$S_1^\delta \doteq \{\tilde{s} = s + r \text{ s.t. } s \in S_1,\ \|r\|_2 \leq \delta\}.$$

By the hypotheses of the Lemma, $S_1^\delta \cap S_2 = \emptyset$; consequently, it suffices to show that $S_1(1 + \delta/\sqrt{m}) \subset S_1^\delta$. Let $s \in S_1(1+\delta/\sqrt{m})$ with $\|s\|_1 = \eta > 1$ (the other case is trivial). Write:

$$s = \frac{1}{\eta}s + \frac{\eta - 1}{\eta}s.$$

Notice that $\frac{1}{\eta}s \in S_1$ and also $\left\|\frac{\eta-1}{\eta}s\right\|_1 = \eta - 1 \leq \delta/\sqrt{m}$. Now:

$$\left\|\frac{\eta-1}{\eta}s\right\|_2 \leq \sqrt{m}\left\|\frac{\eta-1}{\eta}s\right\|_1 = \sqrt{m}(\eta-1) \leq \delta$$

which establishes that $S_1(1+\delta/\sqrt{m}) \subset S_1^\delta$, thus concluding the proof. □

6.1 Computation of an optimal ℓ_1 solution

An optimal solution to the ℓ_1 problem can be computed using a bisection algorithm in the spirit of the γ-iterations for computing a solution to the optimal $\mathcal{H}_\infty$ control problem. Begin by computing the parameterization of all stabilizing controllers, and set:

$$\begin{aligned} \gamma_{min} &= \frac{1}{\sqrt{m}}\|\mathbf{H} - D_Q\|_2 \\ \gamma_{max} &= \|\mathbf{H} - D_Q\|_1. \end{aligned}$$

Recall that m is the number of controlled outputs of the system. Since $\sqrt{m}\gamma_{min}$ is the optimal $\mathcal{H}_2$ norm, γ_{min} provides a lower bound over the achievable ℓ_1 norm. The value γ_{max} is an upper bound since it is the ℓ_1 norm of a transfer matrix resulting from a stabilizing controller.

Algorithm 2

1. *Set* $\gamma = \frac{\gamma_{min}+\gamma_{max}}{2}$.
2. *Use Algorithm 1 to compute a solution in* $S_1 \cap S_2$.
3. *If Algorithm 1 terminates with a solution in* $S_1 \cap S_2$, *set* $\gamma_{max} = \gamma$ *and go to Step 5.*
4. *Otherwise, if Algorithm 1 terminates with* $0 < \delta = \left\|\mathbf{S}^K - \mathbf{T}^K\right\|_2$, *set:*
$$\begin{aligned} \gamma_{min} &= \gamma_{min} + \frac{1}{\sqrt{m}}\delta \\ \gamma_{max} &= \min\left\{\gamma_{max}, \left\|\mathbf{T}^K\right\|_1\right\}. \end{aligned}$$
5. *If* $\gamma_{max} - \gamma_{min} < \epsilon$, *stop. The transfer matrix* $\mathbf{T}^K$ *computed by Algorithm 1 is at most* ϵ *away from optimal. Otherwise, go to Step 1.*

6.2 Finite Convergence and Acceleration

Algorithm 1 may exhibit slow convergence, specially when the ℓ_1 performance level is close to optimal. This difficulty can be addressed in different ways. For instance, Algorithm 1 can be modified so that it exhibits guaranteed convergence in a finite number of steps. For instance, following (De Pierro and Iusem, 1988), one could compute $\mathbf{S}^k$ by projecting into the set $S_1(\gamma - \epsilon^k)$, $\epsilon^k > 0$, where the sequence $\left(\epsilon^k\right)$ verifies $\epsilon^k \to 0$ but $\sum_k \epsilon^k \to \infty$. See the Example section for further comments. Alternatively, instead of using the "cyclic" algorithm characterized by Eqns. (4-2) and (4), one could select different values for both the weights and the relaxation parameters, following, e.g., (Bauschke and Borwein, 1996).

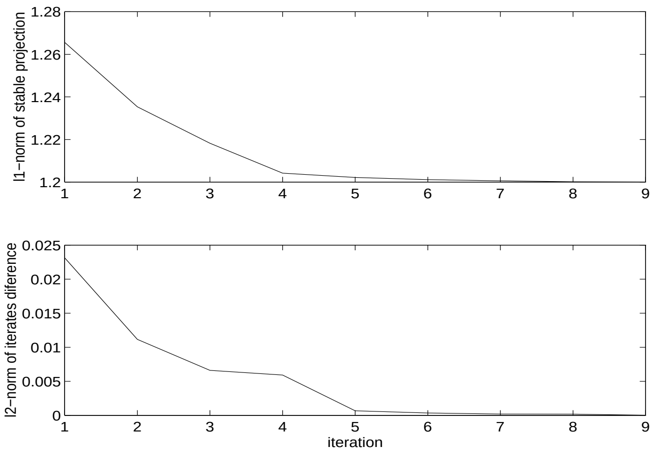

Figure 2: The case $\gamma = 1.2 > \gamma^*$

7 Example

This section considers the first example included in (Dahleh and Pearson, 1987), which is SISO and one-block. Given is the plant

$$\mathbf{p}(z) = \frac{.56z^2 - 1.5z + 1}{z^3 - 1.9z^2 + 1.18z - .24},$$

and the objective is to design a controller minimizing the weighted sensitivity function $\left\|\frac{\mathbf{w}}{1+\mathbf{pk}}\right\|_1$, where

$$\mathbf{w} = .5\frac{z - .99223}{z - .223}.$$

From (Dahleh and Pearson, 1987), the optimal ℓ_1 norm is .99286. The generalized interconnection of Fig. 1 then takes the form

$$G = \begin{bmatrix} \mathbf{w} & -\mathbf{wp} \\ 1 & -\mathbf{p} \end{bmatrix}.$$

Calculation of the parameterization of all stabilizing controllers and the optimal $\mathcal{H}_2$ solution was performed using the formulas in (Rotstein, 1992). The cyclic algorithm described in the previous section was used for searching for a suboptimal solution. Figures 2 and 3 show the performance of the algorithm for the case of performance specification higher and lower than the optimal. For each case, the ℓ_1 norm of the projections onto S_2, and the $\mathcal{H}_2$ norm of the difference between iterates is shown. Notice that the norm of the difference between the projections converge according to the theory presented in the previous section, and can be used as a stopping criteria. Next, an optimal ℓ_1 solution was computed using a bisection algorithm. Since the optimal $\mathcal{H}_2$ closed loop transfer function $\mathbf{T}_2$ verifies $\|\mathbf{T}_2\|_2 = .62$, $\|\mathbf{T}_2\|_1 = 1.34$, the optimal ℓ_1 norm lies within these two bounds. In order to highlight the computation of the lower and upper bounds, an initial value for the performance of $\gamma_1 = .85$ was selected. Figure 4 shows the upper and lower for the performance as a function of the iterations. In the course of the experiments, it was found that a faster convergence can be achieved if the projections are done onto the set $S_1(\gamma_1)$, where γ_1 is slightly smaller than the desired γ. For instance, Figure 5 shows the performance of the algorithm for the case $\gamma = 1$, both for the projections onto $S_1(1.)$ and $S_1(.98)$. The acceleration in convergence is readily noticeable.

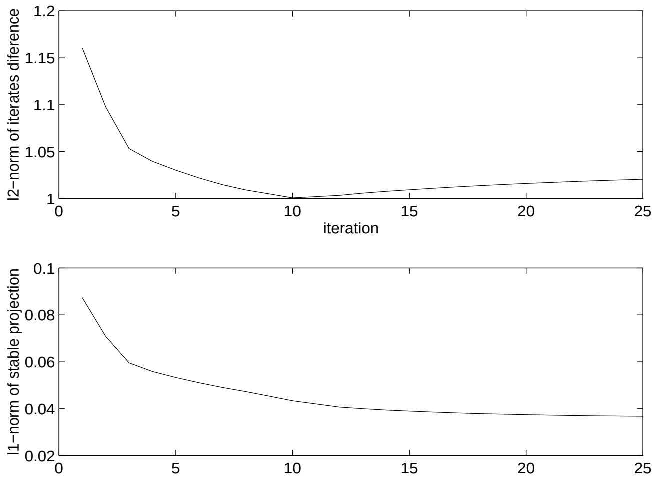

Figure 3: The case $\gamma = .9 < \gamma^*$

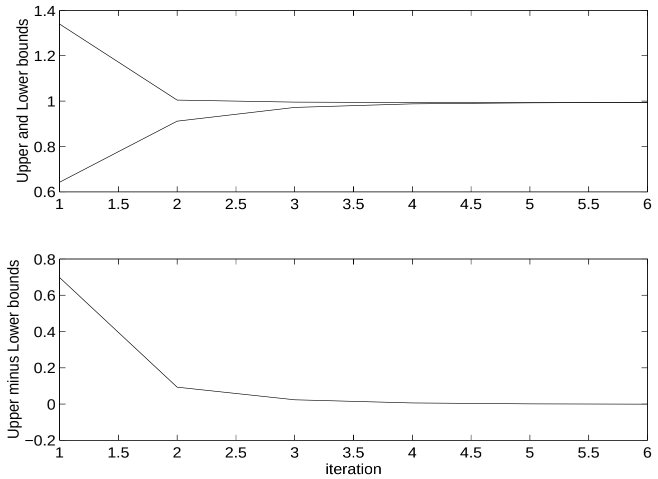

Figure 4: Upper and lower bounds over optimal performance

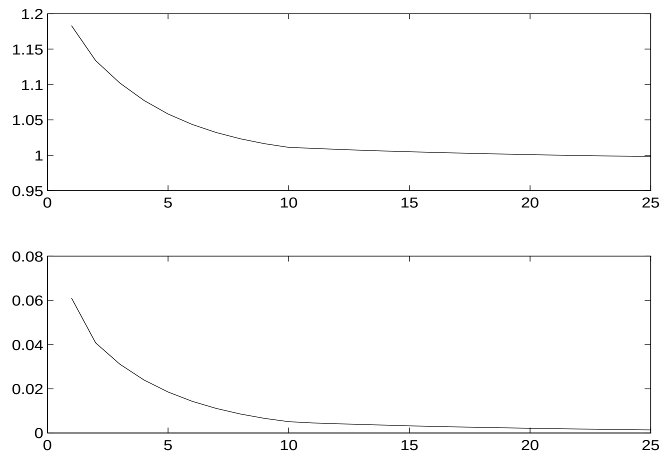

Figure 5: Convergence acceleration

8 Conclusions

In this paper, an algorithm for computing suboptimal ℓ_1 controllers has been presented. The algorithm is based on the alternative projection onto two convex sets: the set of stabilizing controllers and the unit ball in ℓ_1. The algorithm begins from an initial solution and constructs a sequence of iterates that converge to a stabilizing controller achieving the performance level, when such a controller exists. Otherwise the norm of the difference between the two projections on each iteration converge to the distance between the sets. An example has been given to illustrate the performance of the algorithm.

The computation of an optimal solution using iterations on the algorithm has also been discussed. Following this algorithm, one can think of the following hierarchy for the $\mathcal{H}_2$, $\mathcal{H}_\infty$ and ℓ_1 problems:

1. $\mathcal{H}_2$ optimal control problem. Can be solved in closed-form (e.g., using Riccati Equations).

2. $\mathcal{H}_\infty$ optimal control problem. Can be solved iterating over sub-optimal problems. Each sub-optimal problem can be solved in closed-form (e.g., by using Riccati Equations).

3. ℓ_1 optimal control problem. Can be solved iterating over sub-optimal problems. Each sub-optimal problem can be solved performing alternate projections onto two sets. Each projection can be solved in closed-form.

As described in the paper, the algorithm is straightforward and easy to implement, even for the more complicate four-block MIMO problem. The preliminary implementations suggest the following observations:

- Similarly to what happens with the γ iterations for optimal $\mathcal{H}_\infty$ control, verifying the existence of a controller achieving the performance becomes harder as γ converges to the optimal level. In the present approach, the number of iterations of the AP algorithm until convergence increases.

- As iteration progress, so can in general increase the McMillan degree of the iterates $\mathbf{T}^k$ and $\mathbf{S}^k$. Consequently, one could attempt to model reduce $\mathbf{T}^k$ after each iteration (notice that model reducing $\mathbf{S}^k$ destroys the nice structure of the projection onto S_2). However, this should be done with care since it may reduce the convergence rate.

- When looking for a solution with sub-optimal level γ significantly far from the optimal one, projecting over $S_1(\gamma - \epsilon)$ for some fix $\epsilon > 0$, results in a much faster convergence.

Current research focus in improving the convergence properties of the algorithm along the lines of the above comments, and of more general AP algorithms. It is worth stressing that even for the general algorithm described by (4-1), finite-dimensional of all transfer functions involved at each iteration is preserved. Application of the AP algorithm to other sub-optimal control problems is also being explored.

References

Bauschke, H. H. and J. M. Borwein (1993). "On the convergence of von neumann's alternating projection algorithm for two sets," *Set-Valued Analysis*, **1**, no. 2, pp. 185–212.

Bauschke, H. H. and J. M. Borwein (1996). "On projection algorithms for solving convex feasibility problems," *SIAM Reviews*, **38**, no. 3, pp. 367–426.

Dahleh, M. and I. J. Diaz-Bobillo (1994). *Control of uncertain systems: a linear programming approach*, Englewood Cliffs, N.J.: Prentice Hall.

Dahleh, M. A. and J. B. Pearson, Jr. (1986). "ℓ_1 optimal feedback controllers for discrete-time systems," in *Proceedings of the American Control Conference*, pp. 1964–1968.

Dahleh, M. and J. B. Pearson (1987). "l^1-optimal feedback controllers for MIMO discrete-time systems," *IEEE Transactions on Automatic Control*, **32**, no. 4, pp. 314–322.

De Pierro, A. R. and A. N. Iusem (1988). "A finitely convergent "row-action" method for the convex feasibility problem," *Applied Mathematics and Optimisation*, **17**, pp. 225–235.

Diaz-Bobillo, I. and M. Dahleh (1993). "Minimization of the maximum peak-to-peak gain: The general multiblock problem," *IEEE Transactions on Automatic Control*, **38**.

Elia, N. and M. A. Dahleh (1994). "Controller design with multiple objectives," in *Proceedings of the American Control Conference, Baltimore, Maryland*, pp. 1603–1607.

Elia, N. and M. A. Dahleh (1996). "A quadratic programming approach for solving the ℓ_1 multi-block problem," in *Proceedings of the Conference on Decision and Control, Kobe, Japan*, pp. 4028–4033.

Francis, B. (1987). *A Course in $\mathcal{H}_\infty$ Control Theory*, Lecture Notes in Control and Information Science, Springer Verlag, New York: Springer Verlag.

Franklin, J. N. (1980). *Methods of Mathematical Economics*, Springer-Verlag.

Frazho, A. E., K. M. Grigoriadis, and S. M. Kherat (1995). "Alternating projection methods for mixed $\mathcal{H}^2$ and $\mathcal{H}^\infty$ nehari problems," *IEEE Transactions on Automatic Control*, **40**, no. 12, pp. 2127–2131.

Halikias, G. D., I. M. Jaimoukha, and D. A. Wilson (1997). "A numerical solution to the matrix $\mathcal{H}_\infty/\mathcal{H}_2$ optimal control problem," *International Journal of Robust and Nonlinear Control*, **7**, pp. 711–726.

Kavranoglu, D. and A. Sideris (1989). "A simple solution to $\mathcal{H}_\infty$ optimization problems," in *Proceedings of American Control Conference, Pittsburgh, Pennsylvania.*

Khammash, M. (1996). "Solution of the ℓ_1 MIMO control problem without zero interpolation," in *Proceedings of the Conference on Decision and Control, Kobe, Japan*, pp. 4040–4045.

Mendlovitz, M. A. (1989). "A simple solution to the ℓ_1 optimization problem," *Systems & Control Letters*, **12**.

Rivlin, E. and H. Rotstein (1995). "Control of a camera for active vision: Foveal vision, smooth tracking and saccade," Submitted.

Rivlin, E., H. Rotstein, and Y. Zeevi (1997). "Two-mode control: An oculomotor-based approach to tracking systems," To appear in the *IEEE Transactions on Automatic Control.* Preliminary acceptance February 1996.

Rotstein, H. (1992). *Constrained $\mathcal{H}_\infty$-Optimization for Discrete-Time Control Systems*, Ph.D. thesis, California Institute of Technology.

Sideris, A. (1990). "$\mathcal{H}_\infty$ optimal control as a weighted Wiener-Hopf problem," *IEEE Transactions on Automatic Control*, **35**, no. 3, pp. 1272–1276.

Sideris, A. and H. Rotstein (1996). "Personal communication," .

Staffans, O. J. (1993). "The four-block model matching problem in ℓ^1 and infinite-dimensional linear programming," *SIAM Journal on Control and Optimization*, **31**, pp. 747–779.

Vidyasagar, M. (1986). "Optimal rejection of persistent bounded disturbances," *IEEE Transactions on Automatic Control*, **31**, no. 6, pp. 527–534.

von Neumann, J. (1950). *Functional Operators*, vol. II, Princeton University Press. Reprint of mimeographed lecture notes first distributed in 1933.

A Closedness of S_1

The fact that S_1 is defined via the ℓ_1 norm, but it is claimed to be closed in the Hilbert space ℓ_2, maybe somewhat hard to visualize. It is hence convenient to show that S_1 is indeed closed as a subset of ℓ_2. Consider a sequence $x_n \in S_1$ such that $\|x_n - x_m\|_2 \to 0$ as $m,\ n \to \infty$, and let $x_n \to x$ in the 2 norm. The set S_1 is closed if for any such sequence, $x \in S_1$. Assume by contradiction that this is not the case, then $\|x\|_1 > 1$, which implies that $\|\mathcal{P}_N(x)\|_1 > 1$ for some finite N. Take $\epsilon \doteq \|\mathcal{P}_N(x)\|_1 - 1 > 0$. The convergence in the 2-norm implies that $x_n(i) - x(i) \to 0$, although possibly not uniformly. Let M be such that $|x_n(i) - x(i)| < \epsilon/N \ \forall n \geq M,\ 0 \leq i < N$. Then

$$\|\mathcal{P}_N\left(x_m - x\right)\|_1 = \sum_{i=0}^{N-1} |x_M(i) - x(i)| < N\epsilon/N = \epsilon.$$

It follows that

$$\begin{aligned}
\sum_{i=0}^{N-1} |x(i)| &= \sum_{i=0}^{N-1} |x(i) - x_M(i) + x_M(i)| \\
&\leq \sum_{i=0}^{N-1} |x(i) - x_M(i)| + \sum_{i=0}^{N-1} |x_N(i)| \\
&< \epsilon + 1,
\end{aligned}$$

which is a contradiction. Hence $x \in S_1$ and consequently S_1 is closed.

B Proof of Linear Regularity

This section will establish the linear regularity of the pair $\{S_1, S_2\}$ in two different manners. First, a particular instance of the property is establish exploiting the specific properties of the pair. Then, the general case is established using a general result in linear regularity. The advantage of the former proof is that the more direct argument makes the result more plausible.

Lemma 5
Suppose the interconnection matrix G has two inputs and two outputs, and the performance level $\gamma = 1$ is larger than the optimal γ^*. Further assume that all the zeros $c_i,\ i = 1, \cdots, M$ of $\mathbf{U}$ are simple. Then the resulting pair $S_1,\ S_2$ is linearly regular.

Note that for the case described in the Lemma, the parameterization of all closed-loop transfer functions resulting from stable controllers reduces to $\mathbf{T} = \mathbf{H} - \mathbf{UQ}$ for some stable $\mathbf{Q}$.

Proof. Given $\mathbf{x} \in \ell_2$, assume that there exist $\mathbf{s}, \mathbf{T} \in \ell_2$ such that:

$$\begin{aligned}
\|\mathbf{x} - \mathbf{s}\|_2 \leq \delta_1 &\quad \text{and} \quad \|\mathbf{s}\|_1 \leq 1 \\
\|\mathbf{x} - \mathbf{T}\|_2 \leq \delta_2 &\quad \text{and} \quad \mathbf{T} = \mathbf{H} - \mathbf{UQ}
\end{aligned}$$

for some $\mathbf{Q} \in \ell_2$ (i.e., $\mathbf{s} \in S_1,\ \mathbf{T} \in S_2$).

Given $c \in \mathbb{C}$ such that $|c| > 1$, write the bound

$$\begin{aligned} |\mathbf{x}(c) - \mathbf{T}(c)| &= \left|\sum_i (x(i) - T(i))\, c^{-i}\right| \\ &\leq \|\mathbf{x} - \mathbf{T}\|_2 \left(\sum_{i \geq 0} c^{-2i}\right)^{1/2} \\ &\leq \delta_2 \frac{c}{\sqrt{c^2 - 1}}. \end{aligned}$$

Likewise,

$$\begin{aligned} |\mathbf{s}(c) - \mathbf{x}(c)| &\leq \|\mathbf{s} - \mathbf{x}\|_2 \frac{c}{\sqrt{c^2 - 1}} \\ &\leq \delta_1 \frac{c}{\sqrt{c^2 - 1}}. \end{aligned}$$

For each zero c_i of $\mathbf{U}$, $i = 1, \cdots, M$, define:

$$\Delta_i = \mathbf{s}(c_i) - \mathbf{T}(c_i) = \mathbf{s}(c_i) - \mathbf{H}(c_i).$$

From the bounds above

$$|\Delta_i| = c_i \frac{\delta_1 + \delta_2}{\sqrt{c_i^2 - 1}}.$$

Form the square matrix

$$C \doteq \begin{bmatrix} 1 & c_1^{-1} & c_1^{-2} & \cdots & c_1^{-(M-1)} \\ 1 & c_2^{-1} & c_2^{-2} & \cdots & c_2^{-(M-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & c_M^{-1} & c_M^{-2} & \cdots & c_M^{-(M-1)} \end{bmatrix}$$

which is invertible since $c_i \neq c_j$ for $i \neq j$. Consider now the vector

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(M-1) \end{bmatrix} = C^{-1} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_{M-1} \end{bmatrix}$$

and define the transfer function

$$\mathbf{y} = \sum_{i=0}^{M-1} y(i) z^{-i}$$

which verifies

$$\mathbf{y}(c_i) = \Delta_i.$$

Now, the ℓ_1 norm of $\mathbf{y}$ can be bounded by

$$\|\mathbf{y}\|_1 \leq \sigma\,(\delta_1 + \delta_2)\,.$$

Where σ depends on the c_i's (through the matrix C) and on the bound discussed above for $|\Delta_i|$. Take

$$\hat{\mathbf{s}} = \mathbf{s} - \mathbf{y}$$

so that

$$\hat{\mathbf{s}}(c_i) = \mathbf{s}(c_i) - \mathbf{y}(c_i) = \mathbf{H}(c_i).$$

Since the $\hat{\mathbf{s}}(c_i)$ satisfy the interpolation constraints, $\hat{\mathbf{s}} \in S_2$. The ℓ_1 norm of the transfer function $\hat{\mathbf{s}}$, can be larger than 1 but is bounded by

$$\|\hat{\mathbf{s}}\|_1 \leq 1 + \sigma(\delta_1 + \delta_2)$$

and also

$$\begin{aligned} \|\mathbf{x} - \hat{\mathbf{s}}\|_2 &\leq \|\mathbf{x} - \hat{\mathbf{s}}\|_2 + \|\mathbf{s} - \hat{\mathbf{s}}\|_2 \\ &\leq \delta_1 + \sigma(\delta_1 + \delta_2). \end{aligned}$$

The next step is to reduce the ℓ_1 norm of $\hat{\mathbf{s}}$ while remaining inside S_2 and close to $\hat{\mathbf{s}}$. To do this, take $\mathbf{x}^* \in S_2$ such that $\|\mathbf{x}^*\|_1 \leq 1 - \rho$ for some $0 < \rho < 1$. Existence of one such $\mathbf{x}^*$ is guaranteed by the assumption that $\gamma > \gamma^*$. Define

$$\mathbf{s}^* \doteq \lambda \mathbf{x}^* + (1 - \lambda)\hat{\mathbf{s}}$$

where λ is to be defined bellow. It is clear that $\mathbf{s}^*$ belongs to S_2. Also:

$$\begin{aligned} \|\mathbf{s}^*\|_1 &\leq \lambda(1 - \rho) + (1 - \lambda)\left(1 + \sigma(\delta_1 + \delta_2)\right) \\ &= 1 + \sigma(\delta_1 + \delta_2) - \lambda\left(\rho + \sigma(\delta_1 + \delta_2)\right). \end{aligned}$$

Consequently, taking

$$\lambda = \frac{\sigma(\delta_1 + \delta_2)}{\rho + \sigma(\delta_1 + \delta_2)}$$

guarantees $\|\mathbf{s}^*\|_1 \leq 1$, so that $\mathbf{s}^* \in S_1 \cap S_2$. The distance between $\mathbf{s}^*$ and $\mathbf{s}$ can be bounded as:

$$\begin{aligned} \|\mathbf{s}^* - \hat{\mathbf{s}}\|_1 &= \lambda\|\mathbf{x}^* - \hat{\mathbf{s}}\|_2 \leq \lambda\left(\|\mathbf{x}^*\|_1 + \|\hat{\mathbf{s}}\|_1\right) \\ &= \frac{\sigma(\delta_1 + \delta_2)}{\rho + \sigma(\delta_1 + \delta_2)}\left(1 - \rho + 1 + \sigma(\delta_1 + \delta_2)\right) \\ &< \frac{\sigma(\delta_1 + \delta_2)}{\rho + \sigma(\delta_1 + \delta_2)}(2 + \sigma(\delta_1 + \delta_2). \end{aligned}$$

Finally,

$$\begin{aligned} \|\mathbf{s}^* - \mathbf{x}\|_2 &\leq \|\mathbf{s}^* - \hat{\mathbf{s}}\|_2 + \|\hat{\mathbf{s}} - \mathbf{x}\|_2 \\ &\leq \frac{\sigma(\delta_1 + \delta_2)}{\rho + \sigma(\delta_1 + \delta_2)}(2 + \sigma(\delta_1 + \delta_2)) + \delta_1 + \sigma(\delta_1 + \delta_2). \end{aligned}$$

Taking $\delta = \max\{\delta_1, \delta_2\}$ and assuming without loss of generality that $\delta < \frac{1}{2}$, it follows that for some $\kappa > 0$,

$$\|\mathbf{s}^* - \mathbf{x}\|_2 \leq \kappa\delta.$$

Since κ is independent of $\mathbf{x}$, the proof follows. □

To give the general proof for Lemma 3, some additional notation is needed. Given the sets A, B in the Banach space X, $\mathrm{int}_A B$ denotes the interior of B with respect to A, $\overline{\mathrm{Aff}} A$ denotes the closed affine span of A, while $\mathrm{icr}(A)$ denotes the *intrinsic core* of A, i.e., $\mathrm{icr}(A) = \mathrm{int}_{\overline{\mathrm{Aff}}A} A$. The proof is then based on the following result.

Lemma 6
A pair (A, B) *is linearly regular if* $0 \in \ icr(A - B)$.

Proof. See Corollary 4.5 in (Bauschke and Borwein, 1993). □

Proof of Lemma 3
In order to use the previous lemma, define $B = \mathbf{H} + S_1$ and $A = \{\mathbf{UQV} \text{ s.t. } \mathbf{Q} \in \ell_2\}$. Being a geometrical property, it is clear that (S_1, S_2) is linearly regular if and only if the pair (A, B) is so.Then, except for the trivial case where $0 \in B$, $\|\mathbf{H}\|_1 > 1$ and $0 \in (A - B)$.

It is now claimed that $\overline{\text{Aff}}(A - B) = A$. Being A a close subspace, it is clear that $\overline{\text{Aff}}(A - B) \subset A$. Now let $a \in A$, $a \neq 0$, and assume that $a \notin A - B$. The $\|a - \mathbf{H}\|_1 \leq 1$. Now, for some $\lambda \in \mathbb{R}$ large enough, $\|\lambda a - \mathbf{H}\|_1 > 1$, so that $\lambda a \in A - B$, and consequently $a \in \overline{\text{Aff}}(A - B)$. This establishes the claim.

It follows from the claim that $0 \in \text{Icr}(A - B)$ if and only if $0 \in \text{Int}_A(A - B)$. To show this latter fact, take $\mathbf{x} \in A$. Let $\delta = d(0, B)$, which is larger than zero since $\|\mathbf{H}\|_1 > 1$, and consider the open ball $\mathcal{B}(0, \delta)$, centered at 0 and with radius δ. Take $\mathcal{B}_a \doteq \mathcal{B} \cap A$. It is clear that $0 \in \mathcal{B}_a$ and that $\mathcal{B}_a$ is open with respect to A. Moreover, by the definition of δ, $\mathcal{B}_a \cap B = \emptyset$. This concludes the proof.