Q domain sub/super-optimization linear programming methods for MIMO l_1 control problems

Alessandro Casavola and Domenico Famularo
DEIS - Università degli studi della Calabria
Rende (CS), 87036, ITALY
{casavola,famularo}@deis.unical.it

Abstract

In this paper, the MIMO multi-block l_1-optimal control problem is considered. It is shown that it can be converted via polynomial equation techniques to an infinite dimensional linear programming (LP) problem. Finite dimensional sub/super approximations can be determined by considering two sequences of modified finite dimensional linear programming problems derived directly from the YJBK parameterization by exploiting the underlying algebraic structure. This approach induces the application of a consistent truncation strategy that leads to a redundancy-free constraint formulation and, as a consequence, to linear programming problems less affected by degeneracy. Further, more insight on the algebraic structure of the problem and on the achievement of exact rational solutions is provided, allowing the development of a simple and conceptually attractive theory.

Keywords: l_1 Optimal Control, Linear Programming, Matrix Polynomials, Discrete-time Systems.

Definitions and notations

$Z(\cdot)$**:** The Z-transform operator. Given a matrix $\hat{H} = \{\hat{H}(k)\}_{k=0}^{\infty}$ of causal real sequences: $Z(\hat{H}) = H(d) := \sum_{k=0}^{\infty} \mid \hat{H}(k) \mid d^k$, in the complex variable d.

$l_{1,p\times m}$**:** The real normed linear space of all $p \times m$ matrices $\hat{H}$ of absolutely summable real causal sequences $\hat{H}_{ij} = \{\hat{H}_{ij}(k)\}_{k=0}^{\infty}$ with norm $\|\hat{H}\|_1 := \max_{i\in \underline{p}} \sum_{j=1}^{m} \sum_{k=0}^{\infty} \mid \hat{H}_{ij}(k) \mid < \infty$, where $\underline{p} := \{1,2,...,p\}$.

$A_{p\times m}$**:** The real normed linear space of all $p \times m$ matrices $H(d)$ which are Z-transform of some matrix sequence $\hat{H} \in l_{1,p\times m}$. The A-norm is defined as $\|H\|_A := \|\hat{H}\|_1$.

$R\,l_{1,p\times m}$**:** The subspace of $l_{1,p\times m}$ of all real causal matrix sequences each of whose entries has a rational Z-transform. Any sequence with finite support belongs to this subspace because its Z-transform is a polynomial.

$R\,A_{p\times m}$**:** The subspace of $A_{p\times m}$ consisting of elements each of whose entries is a stable stable real rational functions (Elements of $R\,A_{p\times m}$ are Z-transform of some element of $R\,l_{1,p\times m}$).

$\partial(X)$**:** Denotes the degree of the matrix X.

In the sequel the dimensions of the spaces are not indicated either if they are not important or clear from the context. Further, with the *hat* we denote (matrix) sequences, whereas the same *unhatted* variable denotes its corresponding Z-transform; the space of $(p \times m)$ polynomial matrices will be denoted by $R_{p\times m}[d]$ whereas rational transfer matrices by $R_{p\times m}(d)$. With $R^t_{p\times m}[d]$ will be denoted all polynomial matrices with degree lower than or equal to t.

1 Introduction

The discrete-time model matching l_1-optimization problem [1],[2] amounts to the minimization of the A-norm of the closed-loop error transfer matrix

$$E(d) = H(d) - U(d)Q(d)V(d) \tag{1}$$

where H, U, and V are given stable rational matrices ($\in R\,A$) of appropriate dimensions and $Q \in A$ is the free parameter. If U has full row rank and V has full column rank, the problem is usually denoted as one-block or good-rank problem. If one or both of the above conditions are violated, the problem is denoted as multi-block or bad-rank problem. It is now well understood [3, 4] that one-block problems admit rational solutions, that is there exist minimizers $Q_o \in R\,A$ provided that U and V have not transmission zeros on the unit circle. Further, it has been shown [1] that the optimal closed-loop error transfer matrix is a polynomial matrix of some finite degree. As a consequence, a practicable strategy of computation is to restrict the set of feasible Q's to a subset of $R\,A$ such that the corresponding $E(d)$ is a polynomial matrix of finite degree δ, increasing δ until the solution is reached. In bad-rank cases such a strong result of existence is not so far available and minimizers were shown to exist only in A. In this case, finite dimensional approximations are of interest, along with an estimate of the gap existing between the corresponding sub-optima and the true optimum. Various strategies have been proposed to that purpose: Q-design [5], FMV-FME [6], Delay Augmentation Method [7], Semidefinite (Quadratic) programming methods [8]. All of them, except the last method, consist of adding additional constraints to the original optimization problem in order to achieve sub-optimal finite dimensional solutions. Further, in order to have an estimate of the quality of the approximated solution, in [6] was firstly introduced the idea of building a dual sequence of linear programming problems whose solutions form a not-decreasing sequence of super-optima converging from the below to the optimum. The latter was obtained by dropping some of the structural constraints existing between E and Q in (1). The combined use of such a sub/super-optimization scheme allows one to obtain sequences of lower and upper bounds both converging to the optimum. The Semidefinite programming approach [8] instead, embeds the original l_1 problem into two finite dimensional H_2 quadratic programming problems (sub- and super-optimal) and uses the LMI paradigm to numerically approach the problem. All of the above cases, exhausting thoroughly the matter, are based on the transformation of an infinite dimensional linear programming problem into a sequence of finite dimensional ones. In [6] it was pointed out that such an operation is highly not unique and, as consequence, a not secondary task is to investigate how to formulate the finite linear programming problems in order to avoid unnecessary conditions and redundancies in the constraints formulation that lead to linear programming problems affected by degeneracy. In fact, in [4] it was shown that linear programming schemes based on the classical Dahleh-Pearson primal/dual approach, there referred to as FMV-FME schemes, presents some drawbacks and numerical misbehavior just for scalar mixed-sensitivity examples. In particular, numerical misbehavior was due to the existence of linear dependence among the constraints. The proposed remedy was to drop those constraints that caused the loss of rank, a cure that seems unpractical for large dimension problems. Also the Delay Augmentation [7] method seems to be impractical for large problems because of the need to perform some heuristic rows/columns permutation in order the scheme to converge. Moreover, the derivation of the rank/zeros interpolation conditions is quite cumbersome just for simple zeros and potentially numerically instable. The more recent Semidefinite programming approach [8, 16] looks more appealing because it solves the drawbacks of the Delay Augmentation method, moreover the use of a quadratic programming scheme and LMI algorithms are computationally heavier with respect to linear programming methods. The main goal of this paper is to present a different characterization of the closed-loop error transfer function (1). This is done by resorting the polynomial equation approach of Kučera [9]. This allows a more direct achievement of unconstrained suboptimal and superoptimal linear programming problems that are free of most of the above defects. The key idea consists of parameterizing both the closed-loop error E and the free parameter Q in terms of a polynomial matrix, that really represents the available degrees of freedom existing once that the closed-loop stability and feedback structural constraints underlying (1) have been satisfied. As a consequence, the original optimization problem can be expressed in terms of this new polynomial matrix, resulting in an *unconstrained* linear programming problem. Such a scheme was considered in [11] for scalar mixed-sensitivity problems. Here the general MIMO theory is developed. The paper is organized as follows: in section 2 the problem is stated together with structural and stability conditions. In the sections 3,4 sub/super approximation algorithms together with convergence results are described and in the last section a Mixed Sensitivity l_1 optimization problem taken from literature is analyzed.

2 Problem Formulation

Consider a n_w+n_u-inputs, n_z+n_y-outputs discrete LTI dynamical system described by a real-rational transfer matrix $G(d)$. The system has two (vector) inputs and two (vector) outputs: w is a vector of n_w exogenous disturbance inputs, z is a vector of n_z outputs which are to be regulated, y is a vector of n_y outputs which are measurable, and u is a vector of n_u control inputs. The control input u is assumed to be the output of a causal linear finite-dimensional discrete-time controller C whose input is the measured output y. If the matrix G is partitioned accordingly with these inputs and outputs, we obtain the following closed-loop system

$$\begin{aligned} \begin{pmatrix} z \\ y \end{pmatrix} &= \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{pmatrix} w \\ u \end{pmatrix}, \\ u &= C y, \end{aligned} \tag{2}$$

It is well known that if the system is well-posed, the set of all asymptotically stable closed loop error maps E from w to z, which may be achieved by any stabilizing C, are described by the following YJBK parametrization

$$E = H - UQV, \tag{3}$$

where:

$$\begin{array}{lllllll} H & := & G_{11}+G_{12}M\tilde{Y}_G G_{21}, & U & := & G_{12}M, \\ V & := & \tilde{M}G_{21}, & Q & \in & A_{n_u\times n_y}, \end{array} \tag{4}$$

and $G_{22} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ expressed via right and left coprime factorizations with N, M, $\tilde{N}$ and $\tilde{M}$ polynomial matrices. Moreover, there exist polynomial matrices $\tilde{X}_G$, $\tilde{Y}_G$, X_G and Y_G of proper dimension such that the following Bezout identity is satisfied

$$\begin{bmatrix} \tilde{X}_G & -\tilde{Y}_G \\ \tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & Y_G \\ -N & X_G \end{bmatrix} = I_{n_u+n_y}.$$

The transfer matrices $U \in RA_{n_z\times n_u}$, $V \in RA_{n_y\times n_w}$, and $H \in RA_{n_z\times n_w}$ in (4) are A-stable and the transfer matrix Q is the free design parameter. Then, $Q \in RA_{n_u\times n_y}$ implies both $E \in RA_{n_z\times n_w}$ and C rational given by

$$C = (Y_G - MQ)(X_G+NQ)^{-1} = (\tilde{X}_G + Q\tilde{N})^{-1}(\tilde{Y}_G - Q\tilde{M}).$$

In this setting the problem of minimizing $\|\hat{z}\|_\infty$ with respect to all stabilizing C when $\hat{w}$ belongs to the unitary l_∞-norm ball is equivalent to the following minimum distance problem in A:

$$\text{(OPT):} \quad \inf_{K\in S_A} \|H-K\|_{A_{n_z\times n_w}} =: \mu_{\text{opt}} \tag{5}$$

where $S_A := \left\{K \in A_{n_z\times n_w} : \exists Q \in A_{n_u\times n_y} \text{ s.t. } K = UQV\right\}$ is the feasible set. In this paper we consider the situations $n_u < n_z$ (more objective outputs than control inputs) and $n_y < n_w$ (more exogenous disturbances than regulated outputs) which lead to so-called *bad rank* or *four-block* problems [3]. It follows that U is a transfer matrix with more rows than columns ($n_u < n_z$) and V has more columns than rows ($n_y < n_w$). Before continuing we will state two preliminary assumptions which we assume to hold true through the paper.

Assumption 1 *U has full column rank and V have full row rank.*

This assumption can be stated w.l.o.g and ensures that, in the bad rank case, all the control inputs and all the regulated outputs are linearly independent each other. Notice in fact that, if this assumption would not be true, it would be possible to remove one or more of the control inputs (columns of U) and/or one or more regulated outputs (rows of V) so that the remaining control inputs and/or regulated outputs satisfy the assumption. The following assumption on the zero structure of U and V on the unit circle is also required [6]:

Assumption 2 *The rank of the matrices U and V is constant on the unit circle $|d| = 1$.*

This assumption is required because the presence of zeros on the unit circle for U and V prevents existence of minimizers of OPT; some examples with boundary zeros where no minimizers exist are given in [13]. Without loss of generality, U and V can be rewritten, using right and left coprime factorizations,

$$U = UD_U^{-1}N_U, \tag{6}$$

$$V = N_V D_V^{-1} V, \tag{7}$$

where $U \in R_{n_z\times n_u}[d]$ and $V \in R_{n_y\times n_w}[d]$. Notice that the zeros of $D_U \in R_{n_u\times n_u}[d]$, $D_V \in R_{n_y\times n_y}[d]$ are respectively the poles of U and V. Moreover, the matrices $N_U \in R_{n_u\times n_u}[d]$ and $N_V \in R_{n_y\times n_y}[d]$ contain zeros that are common factors in the columns of U and, respectively, in the rows of V. Next, by considering (6) and (7), (3) can be rewritten as

$$E = H - UD_U^{-1}N_U Q N_V D_V^{-1} V. \tag{8}$$

Observe further that, by removing the terms $D_U^{-1}N_U$ from U and $N_V D_V^{-1}$ from V, these matrices are respectively right and left unimodular [15]. Hence, there exist two matrices $U_0 \in R_{n_z\times(n_z-n_u)}[d]$ and $V_0 \in R_{n_w\times(n_w-n_y)}[d]$ such that

$$\begin{bmatrix} U & U_0 \end{bmatrix}, \quad \begin{bmatrix} V \\ V_0 \end{bmatrix} \tag{9}$$

are unimodular and, consequently, their inverses exist too over $R_{n_z\times n_z}[d]$ and $R_{n_w\times n_w}[d]$ such that

$$\begin{bmatrix} \tilde{U} \\ \tilde{U}_0 \end{bmatrix} \begin{bmatrix} U & U_0 \end{bmatrix} = I_{n_z}, \qquad \begin{bmatrix} V \\ V_0 \end{bmatrix} \begin{bmatrix} \tilde{V} & \tilde{V}_0 \end{bmatrix} = I_{n_w} \tag{10}$$

with $\tilde{U} \in R_{n_u\times n_z}[d]$, $\tilde{U}_0 \in R_{(n_z-n_u)\times n_z}[d]$, $\tilde{V} \in R_{n_w\times n_y}[d]$ and $\tilde{V}_0 \in R_{n_w\times(n_w-n_y)}[d]$. As a conclusion, the problem we want to solve can equivalently be stated as

$$\text{(OPT):} \quad \inf_{Q\in A_{n_u\times n_y}} \|E\|_{A_{n_z\times n_w}} := \mu_{opt}, \tag{11}$$

2.1 Structural Conditions

In order to solve (11), we first needs to characterize the class of all admissible closed-loop maps E, viz. the ones compatible with the feedback structure (2). To this end, for the sake of clarity, we consider first the case in which the following assumption holds true

Assumption 3 *Let the quantities $H\tilde{V}_0$ and $\tilde{U}_0 H$ be polynomial matrices, viz. there exist $T_1 \in R_{n_z\times(n_w-n_y)}[d]$, $T_2 \in R_{(n_z-n_u)\times n_w}[d]$, such that*

$$T_1 := H\tilde{V}_0, \tag{12}$$
$$T_2 := \tilde{U}_0 H. \tag{13}$$

Then, the following result characterizes the structure of E.

Lemma 1 *Let Assumption 3 be fulfilled. Then, the one and the only closed-loop maps E that solve (8) can be parameterized in term of a possibly infinite degree polynomial $X \in A_{n_u\times n_y}$ as*

$$E = E^0 + UXV, \tag{14}$$

and a polynomial matrix E^0, jointly solution of the following pair of matrix polynomial equations

$$\begin{cases} E\tilde{V}_0 &= T_1, \\ \tilde{U}_0 E &= T_2. \end{cases} \tag{15}$$

Remark 1 - In order to avoid degree inflation in the closed-loop maps (14), it is desirable that E^o be the unique minimal degree solution of the pair (15), viz. $\partial(E^0) < \partial(U) + \partial(V)$. If this is not the case, left/right polynomial divisions lead to determine a matrix polynomial pair (F,R) such that $E^0 = UFV + R$, with $\partial(R) < \partial(U) + \partial(V)$. □

From now on, we suppose that the polynomial solution E^0 described in Lemma 1 is the minimum degree solution of (15).

2.2 Stability Constraints

From (15) and Assumption 3 we have

$$\left(H - E^0\right)\tilde{V}_0 = 0, \tag{16}$$
$$\tilde{U}_0\left(H - E^0\right) = 0, \tag{17}$$

and, in turn,

$$H - E^0 = UWV, \tag{18}$$

for some $W \in RA_{n_u\times n_y}$ computable by pre-/post- multiplying $H - E^0$ by $\tilde{U}$ and, respectively, $\tilde{V}$. Note also that, from (14) and (8), we have

$$H - E^0 = UXV + UD_U^{-1}N_U Q N_V D_V^{-1} V, \tag{19}$$

and, in order to derive a polynomial equation that relates Q to the free parameter X, the following result related to the structure of the poles of W can be proved

Lemma 2 *Let Assumption 3 be fulfilled. Then, the poles of W are the union of the poles of U and V (see eqs. (6), (7)).*

From (19) and (18), it follows that $W = X + D_U^{-1}N_U Q N_V D_V^{-1}$. Without loss of generality, we can assume that N_U and N_V collect only unstable factors. The stable factors of these matrices, by taking the Smith factorizations of N_U and N_V, can be embedded inside Q. Then, because of Lemma 2, one has that $D_U W D_V = W$ is a polynomial matrix and the stability can be ensured if X and Q are matrix polynomial solutions of

$$D_U X D_V + N_U Q N_V = W. \tag{20}$$

We have the following lemma

Lemma 3 *If W is from (16) then the diophantine equation (20) admits always a polynomial solution.*

The general solution of equation (20) has the following expression

$$\begin{cases} X &= X^0 - \tilde{D}_U T \tilde{D}_V, \\ Q &= Q^0 + \tilde{N}_U T \tilde{N}_V, \end{cases} \tag{21}$$

where $T \in R_{n_u\times n_y}[d]$ is a free polynomial matrix and $\tilde{D}_U$, $\tilde{D}_V$, $\tilde{N}_U$, $\tilde{N}_V$ are polynomial matrices such that

$$\begin{bmatrix} D_U & N_U \\ \tilde{A}_U & -\tilde{B}_U \end{bmatrix}\begin{bmatrix} B_U & \tilde{N}_U \\ A_U & -\tilde{D}_U \end{bmatrix} = I,$$
$$\begin{bmatrix} \tilde{N}_V & \tilde{D}_V \\ B_V & -A_V \end{bmatrix}\begin{bmatrix} \tilde{A}_V & D_V \\ -\tilde{B}_V & N_V \end{bmatrix} = I.$$

As a consequence, all admissible and A-stable closed-loop error maps can be parameterized in terms of a possibly infinite degree free term $T \in A_{n_u\times n_y}$ as

$$E = E^0 + UX^0V - U\tilde{D}_U T \tilde{D}_V V. \tag{22}$$

If E^0 and X^0 are the minimal degree solutions, we know that $\partial(E^0) < \partial(U) + \partial(V)$ and $\partial(X^0) < \partial(\tilde{D}_U) + \partial(\tilde{D}_V)$. Then, $\partial(E^0 + UX^0V) < \partial(U) + \partial(V) + \partial(\tilde{D}_U) + \partial(\tilde{D}_V)$. If $\partial(T) = x$ we have that

$$\partial(E) < \partial(U) + \partial(V) + \partial(\tilde{D}_U) + \partial(\tilde{D}_V) + \partial(T) < m_0 + x. \tag{23}$$

where $m_0 := \partial(U) + \partial(V) + \partial(\tilde{D}_U) + \partial(\tilde{D}_V)$. The parameterization (22) hinges upon the limitative Assumption 3 which is equivalent to assume G_{11} polynomial. When this assumption doesn't hold true, one can adopt a truncation strategy. Several equivalent alternatives are possible. A simple idea is to approximate $G_{11} = N_{G_{11}} D_{G_{11}}^{-1}$, with $N_{G_{11}}$ and $D_{G_{11}}$ left coprime polynomial matrices and $D_{G_{11}}$ strictly Schur, by a N-th order polynomial approximation

$$G_{11} := N_{G_{11}} D_{G_{11}}^{-1} = E^{(N)} + \tilde{E}^{(N)} d^{N+1} D_{G_{11}}^{-1}, \tag{24}$$

with $E^{(N)}$, $\tilde{E}^{(N)}$ polynomial matrices which solve the following Diophantine equation $E^{(N)} D_{G_{11}} + \tilde{E}^{(N)} d^{N+1} = N_{G_{11}}$. The above truncation enables one to write the error sequence matrix E as

$$E = E^{(N)} + UXV - d^{N+1}\tilde{E}^{(N)} D_{G_{11}}^{-1}. \tag{25}$$

Correspondingly, (18) modifies in

$$H - E^{(N)} - d^{N+1}\tilde{E}^{(N)} D_{G_{11}}^{-1} = U\bar{W}V, \tag{26}$$

Again, it can be shown that the poles of $\bar{W}$ are the same of those of U and V in (4) and the stability issue is resolved by requiring that X and Q be solutions of the following Diophantine equation

$$D_U X D_V + N_U Q N_V = \bar{W}, \quad \text{with } \bar{W} = D_U \bar{W} D_V \in R_{n_y\times n_u}[d], \tag{27}$$

note that, due to the structure of $\bar{W}$, Lemma 3 holds true. Finally, for any integer $N > 0$, the parameterization of all admissible and A-stable closed-loop error maps (25) modifies in

$$E = E^{(N)} + UX^0V - U\tilde{D}_U T \tilde{D}_V V + d^{N+1}\tilde{E}^{(N)} D_{G_{11}}^{-1}. \tag{28}$$

3 Suboptimization

The parameterizations (22) and (28) allows one to directly construct suboptimization schemes by imposing that the closed-loop error maps are polynomials. In fact, the above conditions impose additional constraints to (11) and the corresponding solutions are of course sub-optimal. In order to simplify the description in the following, we will define the matrices $\hat{U} = U\tilde{D}_U \in R_{n_z\times n_u}[d]$, $\hat{V} = \tilde{D}_V V \in R_{n_y\times n_w}[d]$.

Then, any arbitrarily tight approximating solution to (11) can be obtained by solving the following finite-dimensional linear programming problem for a sufficient large value for $x := \partial(T)$:

SUB-OPT$_x$: $\quad \bar{\mu}_x := \min_{T\in R^x[d]} \left\| E^{(N)} + UX^0V - \hat{U}\,T\,\hat{V} \right\|_A .$

A convenient choice for N, in order to have all coefficients of the matrix $E^{(N)} + UX^0V$ influenced by T, is

$$N = N(x) = \left\{\partial(\hat{U}) + \partial(\hat{V}) + \partial(T)\right\}, \tag{29}$$

Then, by denoting with

$$\tilde{\mu}_x := \left\| d^{N+1}\,\tilde{E}^{(N)} D^{-1}_{G_{11}} \right\|_A , \tag{30}$$

the A-norm of the truncation tail, a link with the OPT problem is established by the following Lemma.

Lemma 4 *Let Assumptions 1 and 2 be fulfilled and $T^{(x)}$ denote a solution of* **SUB-OPT**$_x$. *Then, the sequence $\bar{\mu}_x$ is non-increasing and*

$$\bar{\mu}_x + \tilde{\mu}_x \geq \bar{\mu}_{x+1} + \tilde{\mu}_{x+1} \geq \mu_{opt},\ \forall\, x \geq 0 \quad \textit{and} \lim_{x\to\infty} \bar{\mu}_x = \mu_{opt}, \quad \lim_{x\to\infty} \tilde{\mu}_x = 0$$

Further, the sequence of solutions $T^{(x)}$ admits a subsequence $T^{(x_s)}$ that converges in the A-norm (component-wise) to an optimal solution of the OPT *problem as $x \to \infty$. If such a solution is unique, the whole sequence converges to it.*

4 Superoptimization

In order to derive linear programming problems whose solutions provide a sequence of lower-bounds to μ_{opt}, it is necessary to rule out some constraints from (11). Because one cannot eliminate constraints related to stability, the only possibility is to relax some structural constraints. This can be done easily by adding for each element E_{ij} $i = 1,\ldots,n_w$, $j = 1,\ldots,n_z$ of E in (22) and (28), the following extra degrees of freedom $T_{ij} \in R_{n_y\times n_u}[d]$, viz.

$$\dot{E}_{ij} = E^{(N)}_{ij} + \underline{u}^T_i\, X^0\, \underline{v}_j - \underline{\hat{u}}^T_i\, (T + d^{x+1}T_{ij})\,\underline{\hat{v}}_j \tag{31}$$

where $\underline{u}^T_i$, $\underline{v}_j$ ($\underline{\hat{u}}^T_i$, $\underline{\hat{v}}_j$) denote the i-th row and the j-th column of $\hat{U}$ and, respectively, $\hat{V}$. It is evident that this choice removes some structural conditions, the extent of which depending on the degrees of T and T_{ij}.Obviously, the corresponding closed-loop maps

$$\dot{E} = E^{(N)} + U\,X^0\,V - \hat{U}\,T\,\hat{V} - d^{x+1}\hat{\mathbf{U}}\,\tilde{T}\,\hat{\mathbf{V}}, \tag{32}$$

are any longer admissible for OPT. Notice that in (32), $\hat{\mathbf{U}} = \operatorname{diag}\left(\underline{\hat{u}}^T_i\right) \in R_{n_z\times(n_u n_z)}[d]$, $i = 1,\ldots,n_z$ and $\hat{\mathbf{V}} = \operatorname{diag}(\underline{\hat{v}}_i) \in R_{(n_w n_y)\times n_w}[d]$, $i = 1,\ldots,n_w$ and let $\tilde{T} \in A_{(n_z n_u)\times(n_w n_y)}[d]$ have polynomial entries denoted by $T_{ij} \in R\,[d]$. For the purpose of obtaining super-optimization schemes, (32) can be simplified furtherly. To this end, let be given the following subspace of RA

$$(U,V)_{RA} = \left\{ L(d) \in RA \;\middle|\; \begin{array}{c} L(d) = U(d)\,T(d)\,V(d) \in (U,V)_{RA}\,, \\ \forall\, T(d) \in (U,V)_{RA} \end{array} \right\}$$

where $U(d) \in R_{n\times m}[d]$, $(n \leq m)$, and $V(d) \in R_{p\times q}[d]$, $(p \geq q)$, are matrices with full normal rank. Using a well known result in literature due to Vidyasagar [13] it can be shown that the closure of $(U,V)_{RA}$ is given by $(U^*,V^*)_{RA}$ where U^* and V^* are respectively left and right divisors of U and V such that: (1) every zero of U and V in the open unit disk is a zero of U^* and, respectively, V^* of the same multiplicity; (2) the zeros of U and V on the unit circle are simple zeros in U^* and, respectively, V^*; (3) if z is a zero on the unit circle both for U and V, such a zero must be a simple zero either for U^* or V^*. So we can conclude that $U(d)^*T(d)V(d)^* \in RA$ if and only if $T(d) \in RA$. In particular, $U(d)^*T(d)V(d)^*$ is a polynomial matrix if and only if $T(d)$ is a polynomial matrix as well. In our context, the above result permits the use the following expression

$$\dot{E} = E^{(N)} + UX^0V - \hat{U}T\hat{V} - d^{x+1}\,\hat{\mathbf{U}}^*\,\tilde{T}\,\hat{\mathbf{V}}^*, \tag{33}$$

in place of (32), with $\hat{\mathbf{U}}^* = \operatorname{diag}\left(\underline{\hat{u}}^{T,*}_i\right) \in R_{n_z\times(n_u n_z)}[d]$, $i = 1,\ldots,n_z$ and $\hat{\mathbf{V}}^*\operatorname{diag}(\underline{\hat{v}}^*_i) \in R_{n_z\times(n_u n_z)}[d]$, $i = 1,\ldots,n_z$, each entries of which determined as described in Lemma 5. The above discussion leads to the following finite dimensional, *unconstrained*, superoptimization sequence of linear programming problems in the unknown matrices (T,T_{ij}), $i = 1,\ldots,n_w$, $j = 1,\ldots,n_z$, indexed by $x = \partial(T) \geq 0$, for given fixed $t_{ij} = \partial(T_{ij}) \geq 0$ and N selected as in (29).

SUP-OPT$_x$:

$$\underline{\nu}_x := \min_{T\in R^x[d];\,\tilde{T}\in R^{t_{ij}}[d]} \left\| E^{(N)} + UX^0V - \hat{U}T\hat{V} - d^{x+1}\,\hat{\mathbf{U}}^*\,\tilde{T}\,\hat{\mathbf{V}}^* \right\|_A ,$$

Remark 2 - The key feature of this approach is that the number of equations retained with respect to the number of variables considered can be modified by varying the degrees of T and T_{ij}. In particular, if t_{ij} are too small, the required superoptimal behavior for the sequence of $\underline{\nu}_x$ cannot be ensured because the quantity of constraints relaxed is not still significant if compared with the cost improvement achieved by increasing x. On the other hand, t_{ij} should be chosen as small as possible because the more the constraints are relaxed, the worse the lower bounds are achieved. Values for t_{ij} which suffice for priming the superoptimal behavior in **SUP-OPT**$_x$ can be obtained by solving the latter with the extra condition $T \equiv 0$. In fact, $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ are matrices which have more columns than rows and, respectively, more rows than columns. So, in such a case ($T \equiv 0$) **SUP-OPT**$_x$ reduces to a good-rank problem in the unknowns $\tilde{T}$

$$\min_{T_{ij}\in R^{t_{ij}}[d]} \left\| E^{(N)}_{ij} + \underline{u}^T_i\, X^0\, \underline{v}_j - \underline{\hat{u}}^T_i\, T_{ij}\,\underline{\hat{v}}_j \right\|_A ,$$

for which a finite degree polynomial solutions T^0_{ij} always exist [1]. Therefore, when $\bar{t}_{ij} = \partial(T^0_{ij})$ is used in **SUP-OPT**$_x$, the increment of x doesn't improve any longer the cost because the finite dimension of the admissible set is just spanned by the unknowns T_{ij} of degree $\bar{t}_{ij}$ and the only effect is that of adding more and more constraints. □

A further characteristic of **SUP-OPT**$_x$ is that a suboptimal admissible solution, in general worse than that achieved by **SUB-OPT**$_x$, is directly available. In fact, denoting as $T^{(x)} + d^{x+1}T^{(x)}_{ij}$ the optimal solution of **SUP-OPT**$_x$ for each x, one has that the sequence of

$$\bar{\mu}_x \leq \left\| E^{(N)} - UX^0V + \hat{\mathbf{U}}T^{(x)}\hat{\mathbf{V}} \right\|_A =: \bar{\nu}_x$$

provides upper-bounds to μ_{opt}. The sequence $\bar{\nu}_x$ denotes a new suboptimal problem, which will be defined as **SUB-OPT**$^*_x$.

Finally it can be shown that **SUP-OPT**$_x$ is well-posed for each $x \geq 0$ and a link with the OPT problem is established by the following result.

Lemma 5 *Let Assumptions 1 and 2 be fulfilled. Then, provided that t_{ij} are sufficiently large, the sequence $\underline{\nu}_x$ is non-decreasing and bounded from the above by μ_{opt}, that is*

$$\underline{\nu}_x \leq \underline{\nu}_{x+1} \leq \mu_{opt} \quad \forall\, x \geq 0$$

Further, regardless of values used for t_{ij}, the sequences $\underline{\nu}_x$ and $\bar{\nu}_x$ converge to μ_{opt}, that is

$$\lim_{x\to\infty} \underline{\nu}_x = \mu_{opt}, \quad \textit{and} \lim_{x\to\infty} \bar{\nu}_x = \mu_{opt}$$

and the sequence of matrix polynomial solutions $T^{(x)}$ admits subsequences $T^{(x_s)}$ that converges in the A-norm (component-wise) to an optimal solution of OPT *as $x \to \infty$, while $T_{ij}^{(x)}$ converge strongly to zero. If such a solution is unique, the whole sequence $T^{(x)}$ converge to it.*

Because of the result from Lemma 5, it is possible to use **SUP-OPT**$_x$ to achieve both lower and upper bounds on the optimum along with an admissible approximating minimizer. In special cases, however, the exact minimizer can be determined. In particular, this is the case when the problem has a $n_u \times n_y$ block in E that is dominant, in the sense that the optimal solution of the reduced one-block problem applied to the original multiblock problem it is also optimal. This also means that the structural constraints corresponding to the inactive blocks can be removed without changing the solution.

Lemma 6 *Assume that $(E)_{hk}$, denotes a dominant block in OPT. Further, let $\left\{T^{(x_s)}, \tilde{T}^{(x_s)}\right\}$ be a subsequence of solutions to* **SUP-OPT**$_x$. *Then, $\lim_{x\to\infty} T_{x_s}^{(x)} = T^{\text{opt}}$, where T^{opt} denotes the optimal solution of the problem made by the dominant rows/columns, while $\lim_{x\to\infty}\left(T^{(x_s)}\right)_{hk} = 0$.*

5 An Example: Mixed Sensitivity MIMO Scheme

Given the following 1-input, 2-outputs plant

$$P(d) = \begin{bmatrix} \dfrac{d-3}{(d-0.1)(d+4)} \\ \dfrac{d-2}{(d-0.2)(d-6)} \end{bmatrix} = N_P D_P^{-1} = \tilde{D}_P^{-1}\tilde{N}_P\,, \qquad (34)$$

The objective is to minimize the A-norm of

$$E = \begin{bmatrix} (I_2+PC)^{-1}W_d & PC(I_2+PC)^{-1}W_n \\ \alpha C(I_2+PC)^{-1}W_d & \alpha C(I_2+PC)^{-1}W_n \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha\left(\tilde{Z}_2 - N_P\tilde{Q}\right)\tilde{D}_P W_d & \alpha\left(-I_{n_y} + \left(\tilde{Z}_2 - N_P\tilde{Q}\right)\tilde{D}_P W_n\right) \\ -\left(\tilde{Z}_1 + D_P\tilde{Q}\right)\tilde{D}_P W_d & -\left(\tilde{Z}_1 + D_P\tilde{Q}\right)\tilde{D}_P W_n \end{bmatrix},$$

where $N_P, D_P, \tilde{N}_P, \tilde{D}_P, \tilde{Z}_1, \tilde{Z}_2$ satisfy the bezoutian equation

$$\begin{bmatrix} Z_1 & Z_2 \\ -\tilde{D}_P & \tilde{N}_P \end{bmatrix}\begin{bmatrix} N_P & -\tilde{Z}_2 \\ D_P & \tilde{Z}_1 \end{bmatrix} = I_3\,,$$

and have the following expression

$$N_P(d) = \begin{bmatrix} (d-3)(d-0.2)(d-6) \\ (d-2)(d-0.1)(d+4) \end{bmatrix},$$

$$D_P(d) = (d-0.1)(d+4)(d-0.2)(d-6)\,,\; \tilde{N}_P(d) = \begin{bmatrix} d-3 \\ d-2 \end{bmatrix},$$

$$\tilde{D}_P(d) = \begin{bmatrix} (d-0.1)(d+4) & 0 \\ 0 & (d-0.2)(d-6) \end{bmatrix},$$

$$Z_1(d) = \begin{bmatrix} -0.14d - 0.57 & 0.22d - 1.36 \end{bmatrix},\; Z_2(d) = -0.08\,,$$

$$\tilde{Z}_1(d) = \begin{bmatrix} -0.14(d-0.2) & 2.28(d-0.1) \end{bmatrix}(d-6)(d+4)\,,$$

$$\tilde{Z}_2(d) = \begin{bmatrix} 1.41d^2 - 1.28d + 2.63 & -0.22d^2 + 2.05d - 4.1 \\ 1.41d^2 + 0.28d - 1.14 & -0.22d^2 - 0.48d + 1.74 \end{bmatrix};$$

the weighting filtering functions have the expression

$$W_d(d) = \begin{bmatrix} \dfrac{1}{1-0.4d} & 0 \\ 0 & \dfrac{1}{1-0.5d} \end{bmatrix},$$

$$W_n(d) = \begin{bmatrix} \dfrac{0.1}{1-0.2d} & 0 \\ 0 & \dfrac{0.1}{1-0.1d} \end{bmatrix};$$

finally $\alpha = 0.1$ is a scalar that properly weights the first two rows of the error matrix in a manner that the block represented by the $(3,1)$ and $(3,2)$ elements will be dominant. From the YJBK parametrization we have that U and V from (6) and (6) are equal to

$$U = U = \begin{bmatrix} N_P \\ D_P \end{bmatrix},$$

$$V = \begin{bmatrix} W_d & W_n \end{bmatrix} = D_{dn}^{-1}\begin{bmatrix} N_d & N_n \end{bmatrix} = D_{dn}^{-1}V\,.,$$

$$D_{dn}(d) = \begin{bmatrix} (1-0.2d)(1-0.4d) & 0 \\ 0 & (1-0.1d)(1-0.5d) \end{bmatrix},$$

$$N_d(d) = \begin{bmatrix} (1-0.2d) & 0 \\ 0 & (1-0.1d) \end{bmatrix},$$

$$N_n(d) = \begin{bmatrix} 0.1(1-0.4d) & 0 \\ 0 & 0.1(1-0.5d) \end{bmatrix}.$$

Note that, as in the previous example, Assumption 3 is not satisfied and, from the structural constraints, the term to be truncated is

$$\begin{bmatrix} W_d & 0 \\ 0 & 0 \end{bmatrix}.$$

Due to the particular structure of U the diophantine equation (27) becomes

$$X D_{dn} + \hat{Q}\,\tilde{D}_P^{-} = -Z_1\,,$$

where $\tilde{D}_P^{-}$ collects the unstable factors of $\tilde{D}_P$ and is equal to

$$\begin{bmatrix} (d-0.1) & 0 \\ 0 & (d-0.2) \end{bmatrix}.$$

The solution of this equation is

$$X = X^0 + T K_X\,, \hat{Q} = Q^0 + T K_Q\,,$$

where

$$X^0(d) = \begin{bmatrix} 0.62 & 1.49 \end{bmatrix},$$

$$Q^0(d) = \begin{bmatrix} -0.04d + 0.5 & -0.07d + 0.65 \end{bmatrix},$$

and

$$K_X(d) = \begin{bmatrix} -12.5(d-0.1) & 0 \\ 0 & -20(d-0.2) \end{bmatrix},$$

$$K_Q(d) = \begin{bmatrix} d^2 - 7.5d + 12.5 & 0 \\ 0 & d^2 - 12d + 20 \end{bmatrix}.$$

Finally the error matrix becomes

$$\begin{bmatrix} E_{11}^N + \bar{E}_{11,l}\,T\,\bar{E}_{11,r} & E_{12}^N + \bar{E}_{12,l}\,T\,\bar{E}_{12,r} \\ E_{12}^N + \bar{E}_{21,l}\,T\,\bar{E}_{21,r} & E_{22}^N + \bar{E}_{22,l}\,T\,\bar{E}_{22,r} \end{bmatrix} =$$

$$\begin{bmatrix} W_d^N + N_P X^0 \tilde{N}_d + N_P T K_X \tilde{N}_d & N_P X^0 \tilde{N}_n + N_P T K_X \tilde{N}_n \\ D_P X^0 \tilde{N}_d + D_P T K_X \tilde{N}_d & D_P X^0 \tilde{N}_n + D_P T K_X \tilde{N}_n \end{bmatrix}$$

where W_d^N is the N-th order approximation polynomial matrix to W_d according to the scheme from eq. (26).

In table 1 the convergence of the norms is showed when the degree of T increases. Note that the optimum turns out to be 75.80755.

In table 2 the convergence of the error sequences for the components $(3,1)$ and $(3,2)$ in the dominant block is shown. It's evident that the super-optimal component remains identical when x varies. The sub-optimal component shows a tail of elements of increasing power whose coefficients become small as the approximation is refined (weak* convergent behaviour). The optimal Q can be retrieved from the expression of $E_{31,x}^0$, $E_{32,x}^0$ in the sup-optimal column in table 2 and is given by

$$Q(d) = \begin{bmatrix} \dfrac{0.05(d-8.58)(d+1.96)(d-2.78)}{(d+4)} \\ \dfrac{0.07(d-12.34)(d-2.23)(d+0.77)}{(d-6)} \end{bmatrix}^T \qquad (35)$$

	SUB-OPT$_x$	**SUP-OPT$_x$**	
x	$\tilde{\mu}_x$	$\underline{\nu}_x$	$\overline{\nu}_x$
0	79.81208	75.69411	80.20161
1	76.17749	75.78819	76.28033
2	75.99559	75.80458	76.06581
3	75.83362	75.80698	75.84200
4	75.81782	75.80738	75.82173
5	75.80915	75.80745	75.80962
6	75.80813	75.80746	75.80830
7	75.80763	75.80750	75.80778
8	75.80757	75.80752	75.80764
9	75.80756	75.80754	75.80762
10	75.80756	75.80755	75.80760
11	75.80755	75.80755	75.80758
12	75.80755	75.80755	75.80757
...	...	...	...

Table 1: MIMO Plant:Convergence of the Norms

x	**SUB-OPT$_x$**
$E^0_{31,0}$	$-0.295+4.436d-14.840d^2+1.038d^5-0.037d^6-0.011d^7$
$E^0_{32,0}$	$-0.719+1.080d-36.108d^2+1.844d^5-0.147d^6-0.000d^7$
$E^0_{31,1}$	$-0.294+4.418d-14.729d^2+0.083d^6+0.048d^7-0.008d^8$
$E^0_{32,1}$	$-0.715+1.073d-35.797d^2+0.115d^6-0.091d^7-0.007d^8$
$E^0_{31,2}$	$-0.294+4.418d-14.727d^2+0.058d^7-0.003d^8-7\,10^{-4}d^9$
$E^0_{32,2}$	$-0.716+1.074d-35.803d^2+0.089d^7-0.013d^8+4\,10^{-4}d^9$
$E^0_{31,3}$	$-14.727(0.02-0.3d+d^2)+0.003d^8+0.002d^9-5\,10^{-4}d^{10}$
$E^0_{32,3}$	$-35.802(0.02-0.3d+d^2)-0.011d^8+0.004d^9-4\,10^{-4}d^{10}$
...	...
$E^0_{31,6}$	$-14.727(0.02-0.3d+d^2)-2\,10^{-4}d^{11}+2\,10^{-4}d^{12}-8\,10^{-7}d^{13}$
$E^0_{32,6}$	$-35.802(0.02-0.3d+d^2)-2\,10^{-4}d^{11}-6\,10^{-6}d^{12}+3\,10^{-7}d^{13}$

x	**SUP-OPT$_x$**
$E^0_{31,0}$	$-14.72701(0.02-0.3d+d^2)$
$E^0_{32,0}$	$-35.80243(0.02-0.3d+d^2)$
$E^0_{31,1}$	$-14.72701(0.02-0.3d+d^2)$
$E^0_{32,1}$	$-35.80243(0.02-0.3d+d^2)$
$E^0_{31,2}$	$-14.72701(0.02-0.3d+d^2)$
$E^0_{32,2}$	$-35.80243(0.02-0.3d+d^2)$
$E^0_{31,3}$	$-14.72701(0.02-0.3d+d^2)$
$E^0_{32,3}$	$-35.80243(0.02-0.3d+d^2)$
...	...
$E^0_{31,6}$	$-14.72701(0.02-0.3d+d^2)$
$E^0_{32,6}$	$-35.80243(0.02-0.3d+d^2)$

Table 2: Convergence of the components of the error sequences in the dominant block

6 Conclusions

New sub/superoptimization schemes have been presented for the MIMO l_1 general four block control problem which result less affected by unnecessary redundant constraints and hence more efficiently solvable. This has been obtained by exploiting a matrix polynomial equations approach and the YJBK parameterization of the admissible maps. In this way the structural and stability constraints have been taken care by the parameterization and the resulting optimization problems are unconstrained. An example taken from literature has been fully analyzed and an approximation to the optimal solution, which is not rational, has been determined.

References

[1] M.A. Dahleh and J.B. Pearson. “l_1-Optimal Feedback Controllers for MIMO Discrete-time Systems”, *IEEE Trans. Automatic Contr.*, Vol. 32, pp. 314-322, (1987).

[2] M.A. Dahleh and J.B. Pearson. “Optimal Rejection of Persistent Disturbances, Robust Stability and Mixed Sensitivity Minimization”, *IEEE Trans. Automatic Contr.*, Vol. 33, pp. 924-930, (1988).

[3] J.S. MacDonald and J.B. Pearson. “l_1-Optimal Control of Multivariable Systems with Output Norm Constraints”, *Automatica*, Vol. 27, pp. 317-329, 1991.

[4] O.J. Staffans. “Mixed sensitivity minimization problems with rational l_1- optimal solutions”, *J. Optimiz. Theory Appl.*, Vol. 70, pp. 173-189, 1991.

[5] S.P. Boyd and C.H. Barratt. *Linear Controller Design: Limits of Performance*, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[6] O.J. Staffans. “The four-block model matching problem in l_1 and infinite-dimensional linear programming”, *SIAM J. Contr. Optimiz.*, Vol. 31, pp. 747-779, 1993.

[7] I.J. Diaz-Bobillo and M.A. Dahleh. “Minimization of the maximum peak-to-peak gain: The general multiblock problem”, *IEEE Trans. Automatic Control*, Vol. 38, pp. 1459-1482, 1993.

[8] N. Elia and M.A. Dahleh “A quadratic programming approach for solving the l_1 multiblock problem”, *IEEE Trans. Automatic Control*, Vol. 43, pp. 1242-1252, 1998.

[9] V. Kučera. *Discrete Linear Control*, Wiley, New York, 1979.

[10] I. Barrodale and F.D.K. Roberts. “An Improved Algorithm for Discrete l_1 Linear Approximation”, *SIAM J. Num. Anal.*, Vol. 10, pp. 839-848, 1973.

[11] A. Casavola. “A polynomial approach to the l_1-mixed sensitivity optimal control problem”, *IEEE Trans. Automatic Control*, Vol. 41, pp. 751-756, 1996.

[12] M. Vidyasagar. “Optimal rejection of persistent bounded disturbances”, *IEEE Trans. Autom. Contr.*, Vol. 31, pp. 527-534, 1986.

[13] M. Vidyasagar. “Further results on the optimal rejection of persistent bounded disturbances”, *IEEE Trans. Autom. Contr.*, Vol. 36, pp. 642-652, 1991.

[14] J. Dugundji. *Topology*, Allyn and Bacon, Boston, 1966.

[15] A. B. Özgüler. *Linear Multichannel Control. A System Matrix Approach*, Prentice Hall, New York, 1994.

[16] H. Yoon, B.H. Tongue and A.K. Packard. “Q domain optimization method for l_1-optimal controllers”, *Automatica*, Vol. 36, pp. 603-611, 2000.

[17] M. Khammash. “ A New Approach to the Solution of the ℓ_1 Control Problem: The Scaled-Q Method”, *IEEE Trans Autom. Contr.*, Vol. 45, pp. 180-187, 2000.