Proceedings of the
American Control Conference
Baltimore, Maryland • June 1994

# Biaffine Matrix Inequality Properties and Computational Methods *†

K. C. Goh‡ L. Turan M. G. Safonov
G. P. Papavassilopoulos J. H. Ly
Dept. of Electrical Engineering - Systems, University of Southern California,
Los Angeles, CA 90089-2563.


## Abstract

Many robust control synthesis problems, including $\mu/k_m$– Synthesis, have been shown to be reducible to the problem of finding a feasible point under a Biaffine Matrix Inequality (BMI) constraint. The paper discusses the related problem of minimizing the maximum eigenvalue of a biaffine combination of symmetric matrices, a biconvex, nonsmooth optimization problem. Various properties of the problem are examined and several local optimization approaches are presented, although the problem requires a global optimization approach in general.


## 1 Introduction

In [35], it is shown that a wide variety of key robust control synthesis problems may be reduced to BMI feasibility problems.

**Definition 1.1 (The BMI Feasibility Problem)** *Given matrices* $F_{i,j} = F_{i,j}^T \in \mathbf{R}^{m\times m}$, *for* $i \in \{0,\dots n_x\}$, $j \in \{0,\dots n_y\}$. *Define the biaffine function* $F : \mathbf{R}^{n_x} \times \mathbf{R}^{n_y} \to \mathbf{R}^{m\times m}$:

$$F(x,y) := F_{0,0} + \sum_{i=1}^{n_x} x_i F_{i,0} + \sum_{j=1}^{n_y} y_j F_{0,j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{i,j} \quad (1)$$

*Find, if it exists,* $(x,y) \in \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$ *such that*

$$F(x,y) > 0 \quad (2)$$

Clearly, a Linear or Affine Matrix Inequality (LMI/AMI) constraint (see [6] and references therein) is a special case of (2).

Define the function

$$\Lambda(x,y) := -\underline{\lambda}\{F(x,y)\} \quad (3)$$

and consider the following biconvex nonsmooth optimization problem:

$$\min\ \Lambda(x,y),\ (x,y) \in \mathbf{R}^{n_x} \times \mathbf{R}^{n_y} \quad (4)$$

It is obvious that the solution set to the BMI feasibility problem (2) is nonempty if and only if

$$0 > \min_{(x,y)} \Lambda(x,y)$$

This paper discusses the properties of the function $\Lambda(x,y)$ and various preliminary approaches for minimizing it. The organization is as follows: basic properties and trivial cases for (4) are discussed, and upper and lower bounds are derived. A geometrical viewpoint is then presented. Various local optimization approaches are discussed. A brief overview of applicable global optimization approaches is also included.

*Research supported in part by AFOSR grants F49620-92-J-0014, F45620-93-1-0505.

†The authors gratefully acknowledge many valuable discussions with Dr. T.C. Tsao of U.S.C.

‡Email: goh@nyquist.usc.edu.

| | |
|---|---|
| $\mathbf{R},\mathbf{R}^p,\mathbf{R}^{p\times q}$ | The sets of real numbers, of $p$-dimensional real vectors, and of $p\times q$ real matrices, respectively. |
| $X(i,j)$ | $ij$th element or submatrix of matrix $X$. |
| $[x]_{ij}$ | A matrix with $ij$th element given by $x \in \mathbf{R}$. |
| $x^T$, $X^T$ | The vector and matrix transposes, respectively. |
| $X > 0$ | $X \in \mathbf{R}^{p\times p}$ is symmetric positive definite. |
| $\underline{\lambda}\{X\}$ | The least eigenvalue of square matrix $X$. |
| $\overline{\lambda}\{X\}$ | The greatest eigenvalue of square matrix $X$. |
| $\mathrm{tr}\{X\}$ | $\mathrm{tr}\{X\} := \sum_i X(i,i)$. |
| $\langle X,Y\rangle$ | The inner product of matrices $X$, $Y$ of same dimension, $\langle X,Y\rangle := \mathrm{tr}\{Y^TX\}$. |
| $\|x\|_2$ | The Euclidean norm on vector $x$, $:= \sqrt{x^Tx}$. |
| $\mathbf{SR}^p$ | The unit sphere in $\mathbf{R}^p$, $\{x \in \mathbf{R}^p\ :\ \|x\|_2 = 1\}$. |
| $\|X\|_F$ | The Frobenius norm on matrix $X$, $:= \sqrt{\langle X,X\rangle}$. |
| co $\mathcal{X}$ | The convex hull of a set $\mathcal{X}$. |

Table 1: Notation

### 1.1 Background: Eigenvalue Optimization Problems

A brief survey of previous work on eigenvalue optimization is presented. An extensive survey is found in [6].

Standard early references on perturbation theory for eigenvalues of matrices are [38, 25]. An early investigation into the problem of optimizing eigenvalues of matrices is given in [8], which established the nonsmoothness of the problem and proposed a generalized gradient descent type algorithm. It is of interest to note that considerable attention was devoted to singular value (and hence eigenvalue) optimization based on generalized gradients in the early robust control literature [10, 33].

A quadratically convergent algorithm for minimizing the largest (in absolute value) eigenvalue of an affine combination of symmetric matrices is given in [31]. The numerical range approach is used in [11, 2], the key concept being the reduction of the LMI problem to the problem of finding the minimum norm point in the convex hull of a numerical range [16].

Approaches based on interior point methods [22, 12, 24] have recently gained prominence, e.g., [28, 23, 5, 37], particularly with regard to generating polynomial time algorithms. In fact, a commercially available software package, LMI–LAB, for solving LMIs based on interior point methods has been introduced [14].

## 2 Preliminaries

We will now discuss the properties of a BMI. Table 1 provides a summary of our notation, which is standard.

There are certain cases where the solution to the BMI (2) is trivial, or where the BMI reduces to an AMI. For example, if

$F_{i,j} = 0$ for all $i = 1, \ldots, n_x, j = 1, \ldots, n_y$ then $F(x,y) > 0$ is in fact an AMI in $(x,y)$. Also, obtaining a feasible point for the BMI (2) is trivial whenever any one of the $F_{i,j}$ is sign definite. Further, note that if any of the AMIs in $x$: $F_{0,0}+\sum_{i=1}^{n_x} x_i F_{i,0} > 0$, $F_{0,0}+F_{0,j}+\sum_{i=1}^{n_x} x_i(F_{i,0}+F_{i,j}) > 0$, $j = 1, \ldots, n_y$, has a solution, then the solution to the BMI problem (2) trivially follows. The same holds for the corresponding AMIs in $y$.

If the BMI constraint is linear in at least one of the variables, say $x$, then $F(x,y) > 0 \Leftrightarrow F(\kappa x, y) > 0$ for all $\kappa > 0$, i.e. the feasible set of (2) will be unbounded if it is non-empty and $\max \underline{\lambda}\{F(x,y)\}$ is either unbounded or negative.

Further, if the BMI is actually linear in both $x$ and $y$, then $F(\bar{x},\bar{y}) > 0 \Leftrightarrow F(\kappa_x \bar{x}, \kappa_y \bar{y}) > 0 \ \forall \ \kappa_x \kappa_y > 0$. Hence if the feasible set is non-empty, it is unbounded.

## 2.1 Upper and Lower Bounds

The next few facts arise from the inequalities

$$\underline{\lambda}\{A\} + \underline{\lambda}\{B\} \leq \underline{\lambda}\{A+B\} \leq \overline{\lambda}\{A\} + \underline{\lambda}\{B\} \tag{5}$$

which hold for all symmetric $A$, $B$, [38]. We first have the following facts about the affine matrix function $H(x) := H_o + \sum_{i=1}^{n} x_i H_i$ from [23] (all maximizations being over $x \in \mathbf{R}^n$):

$$\underline{\lambda}\{H_o\} \leq \max \underline{\lambda}\{H(x)\} \tag{6}$$

$$\max \underline{\lambda}\{H(x)\} < \infty \Leftrightarrow \max \underline{\lambda}\left\{\sum_{i=1}^{n} x_i H_i\right\} = 0 \tag{7}$$

$$\max \underline{\lambda}\{H(x)\} < \infty \Leftrightarrow \max \underline{\lambda}\{H(x)\} \leq \overline{\lambda}\{H_o\} \tag{8}$$

Similarly, for the function $F(x,y)$ of (1), we have:

**Proposition 2.1** *It holds that* $\max_{x,y} \underline{\lambda}\{F(x,y)\}$ *is unbounded if and only if*

$$\max_{x,y} \underline{\lambda}\left\{\sum_{i=1}^{n_x} x_i F_{i,0} + \sum_{j=1}^{n_y} y_j F_{0,j} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{i,j}\right\} > 0 \tag{9}$$

We now present some useful lower and upper bounds for $\max_{x,y} \underline{\lambda}\{F(x,y)\}$. Define the affine matrix function

$$F_L(x,y,w) := F_{0,0} + \sum_{j=1}^{n_y} y_j F_{0,j} + \sum_{i=1}^{n_x} x_i F_{i,0} + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_k F_{i,j} \tag{10}$$

where $w \in \mathbf{R}^{n_x n_y}$, and the index $k := (i-1)n_y + j$.

**Theorem 2.1** *It holds that*

$$\begin{aligned} \underline{\lambda}\{F_{0,0}\} &\leq \left\{\max_y \underline{\lambda}\left\{F_{0,0} + \sum_{j=1}^{n_y} y_j F_{0,j}\right\}, \max_x \underline{\lambda}\left\{F_{0,0} + \sum_{i=1}^{n_x} x_i F_{i,0}\right\}\right\} \\ &\leq \max_{x,y} \underline{\lambda}\{F(x,y)\} \end{aligned} \tag{11}$$

*and*

$$\max_{x,y} \underline{\lambda}\{F(x,y)\} \leq \max_{x,y,w} \underline{\lambda}\{F_L(x,y,w)\} \tag{12}$$

*Further, if* $\max_{x,y,w} \underline{\lambda}\{F_L(x,y,w)\}$ *is bounded,*

$$\max_{x,y} \underline{\lambda}\{F(x,y)\} \leq \max_{x,y,w} \underline{\lambda}\{F_L(x,y,w)\} \leq \overline{\lambda}\{F_{0,0}\} \tag{13}$$

*Proof:* The proof for the lower bounds are of course trivial, and the proof for the upper bound (12) is obvious. The upper bound (13) comes from (8). *Q.E.D.*

Note that all the upper and lower bounds are at most AMI problems which are easily computable.

## 2.2 Biconvexity and the BMI

We now investigate the underlying geometry of the problem of finding a feasible point under a BMI constraint.

Consider some set $\mathcal{X} \times \mathcal{Y} \subset \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$, where $\mathcal{X}$ is convex in $\mathbf{R}^{n_x}$ and $\mathcal{Y}$ is convex in $\mathbf{R}^{n_y}$. Define the $x$ and $y$–sections of $\mathcal{B}$ as follows: $\mathcal{B}_x := \{y \in \mathcal{Y} : (x,y) \in \mathcal{B}\}$ and $\mathcal{B}^y := \{x \in \mathcal{X} : (x,y) \in \mathcal{B}\}$.

**Definition 2.1 (Biconvexity of a Set)** *The set $\mathcal{B} \subset \mathcal{X} \times \mathcal{Y}$ is biconvex if $\mathcal{B}_x$ is convex for every $x \in \mathcal{X}$ and $\mathcal{B}^y$ is convex for every $y \in \mathcal{Y}$.*

**Proposition 2.2** *The set $\mathcal{B} \subset \mathcal{X} \times \mathcal{Y}$ is biconvex if and only if for every quadruple of the form $(x_1,y_1),(x_1,y_2),(x_2,y_1),(x_2,y_2) \in \mathcal{B}$, it holds that*

$$(x_\beta, y_\gamma) := ((1-\beta)x_1 + \beta x_2, (1-\gamma)y_1 + \gamma y_2) \in \mathcal{B}$$

*for every $(\beta,\gamma) \in [0,1] \times [0,1]$.*

A biconvex set is not necessarily convex, e.g., consider the shape "L" on the $\mathbf{R} \times \mathbf{R}$ product space, which is biconvex but not convex and also the level sets for $f(x,y) = xy < 1$, $f(x,y) = xy < -1$, the latter of which is not even connected, but is still biconvex.

**Definition 2.2 (Biconvexity of a Function)** *Given a function $f : \mathcal{X} \times \mathcal{Y} \to \mathbf{R}$, $f(x,y)$ is biconvex in $(x,y)$ if it is convex in $x$ for every fixed $y \in \mathcal{Y}$ and convex in $y$ for every fixed $x \in \mathcal{X}$.*

**Proposition 2.3** *$f(x,y)$ is biconvex over $\mathcal{X} \times \mathcal{Y}$ if and only if for any $(x_1,y_1),(x_1,y_2),(x_2,y_1),(x_2,y_2) \in \mathcal{X} \times \mathcal{Y}$,*

$$\begin{aligned} f((x_\beta, y_\gamma)) \leq\ & (1-\beta)(1-\gamma)f(x_1,y_1) + (1-\beta)\gamma f(x_1,y_2) + \\ & \beta(1-\gamma)f(x_2,y_1) + \beta\gamma f(x_2,y_2) \end{aligned} \tag{14}$$

*for every $(\beta,\gamma) \in [0,1] \times [0,1]$, where $(x_\beta, y_\gamma) := ((1-\beta)x_1 + \beta x_2, (1-\gamma)y_1 + \gamma y_2)$.*

Clearly, just as one dimensional interpolation always overestimates a convex function, two dimensional interpolation always overestimates a biconvex function.

**Proposition 2.4** *If $f(x,y)$ is biconvex, then its level sets, $\mathcal{L}_c := \{(x,y) \in \mathcal{X} \times \mathcal{Y} : f(x,y) \leq c\}$, $c \in \mathbf{R}$, are biconvex for all $c$.*

One of the reasons for the study of convexity of functions is that if a function is convex, then every local minimum is a global minimum. However, convexity is not a necessary condition for the local-global property, and in fact *arcwise strict quasiconvexity* is a sufficient condition for every local minimum to be global minimum [20]. Unfortunately, biconvexity, as we have defined it, does not imply the local-global property.

Returning to the function $\Lambda(x,y)$ of (3), we have:

**Theorem 2.2** *The function $\Lambda(x,y)$ is biconvex over $\mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$.*

The proof follows trivially from the well established fact that AMI/LMIs are convex.

In general, BMI problems are multiextremal, and therefore global optimization techniques based on biconvexity will need to be considered in the minimization of $\Lambda(x,y)$. However, it is also quite easy to generate non-trivial BMIs (i.e. not AMIs) for which only one minimum of $\Lambda(x,y)$ exists.

## 2.3 Connection between Biaffine and Bilinear Cases

We now examine the relationship between the bilinear and biaffine matrix inequalities.

Given a bilinear matrix function $F(x,y) := \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{i,j}$, a biaffine matrix function $\tilde{F}(x,y) := \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{i,j} + \sum_{i=1}^{n_x} x_i F_{i,0} + \sum_{j=1}^{n_y} y_j F_{0,j} + F_{0,0}$ is easily generated, where $F_{0,0}$, $F_{0,j}$ and $F_{i,0}$ are all zero matrices. Clearly then $F(x,y) > 0 \Leftrightarrow \tilde{F}(x,y)$.

On the other hand, given the biaffine matrix function of (1), we may define

$$\tilde{F}(\tilde{x},\tilde{y}) := \sum_{i=0}^{n_x}\sum_{j=0}^{n_y} \tilde{x}_i \tilde{y}_j F_{i,j} \tag{15}$$

It then holds that

**Proposition 2.5** *If there exist $\tilde{x} \in \mathbf{R}^{n_x+1}$ and $\tilde{y} \in \mathbf{R}^{n_y+1}$ such that $\tilde{F}(\tilde{x},\tilde{y}) > 0$ and $\tilde{x}_0\tilde{y}_0 \geq 0$, then there exist $x^* \in \mathbf{R}^{n_x}$ and $y^* \in \mathbf{R}^{n_y}$ such that $F(x^*,y^*) > 0$.*

Also note that it is possible to convert biaffine problems into bilinear problems by using the embedding suggested in [35].

# 3 A Geometrical Interpretation of the BMI Feasibility Problem

Given a set of symmetric matrices $F_{i,j}$ defining a BMI, from (1) and (2), define the *matrix numerical range*:

$$\mathcal{W}(F_{i,j}) := \left\{ W \in \mathbf{R}^{(n_x+1)\times(n_y+1)} : W(z) = \begin{bmatrix} z^T F_{0,0} z & \cdots & z^T F_{0,n_y} z \\ \vdots & \ddots & \vdots \\ z^T F_{n_x,0} z & \cdots & z^T F_{n_x,n_y} z \end{bmatrix}, \|z\|_2 = 1 \right\} \tag{16}$$

The concept of a matrix numerical range introduced here is an obvious extension of the concept of an *m-dimensional field of values* [19], or an *m-form numerical range* [11]. We may convert the problem of finding the feasible set of a BMI into a matrix numerical range problem as follows:

For any fixed $z \in \mathbf{R}^m$ with $\|z\|_2 = 1$,

$$\begin{aligned} z^T F(x,y) z &= z^T F_{0,0} z + \sum_{i=1}^{n_x} x_i z^T F_{i,0} z + \sum_{j=1}^{n_y} y_j z^T F_{0,j} z \\ &\quad + \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j z^T F_{i,j} z \\ &= \begin{bmatrix} 1 & x^T \end{bmatrix} W(z) \begin{bmatrix} 1 \\ y \end{bmatrix} \end{aligned} \tag{17}$$

where $W(z)$ is as given in (16), so that clearly $W(z) \in \mathcal{W}(F_{i,j})$. It follows that

$$\begin{aligned} \underline{\lambda}\{F(x,y)\} &\equiv \min_{\|z\|_2=1} z^T F(x,y) z \\ &\equiv \min_{W\in\mathcal{W}(F_{i,j})} \begin{bmatrix} 1 & x^T \end{bmatrix} W \begin{bmatrix} 1 \\ y \end{bmatrix} \\ &\equiv \min_{W\in\mathcal{W}(F_{i,j})} \operatorname{tr}\left\{ \begin{bmatrix} 1 \\ y \end{bmatrix} \begin{bmatrix} 1 & x^T \end{bmatrix} W \right\} \\ &\equiv \min_{W\in\mathcal{W}(F_{i,j})} \left\langle \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 & y^T \end{bmatrix}, W \right\rangle \end{aligned} \tag{18}$$

We may then give an equivalent formulation of the BMI Problem:

**Theorem 3.1** *The BMI (2) has a solution if there exists a dyad $X \in \mathbf{R}^{(n_x+1)\times(n_y+1)}$ such that*

$$\langle X, W\rangle > 0 \tag{19}$$

*for all $W \in \mathcal{W}(F_{i,j})$, and the $(1,1)$ element of $X$ is non-negative.*

*Conversely, a dyad $X \in \mathbf{R}^{(n_x+1)\times(n_y+1)}$ such that (19) holds for all $W \in \mathcal{W}(F_{i,j})$ exists if the BMI (2) has a solution.*

*Proof:* The proof for the converse part is trivial. Suppose $(\tilde{x},\tilde{y})$ is in the feasible set $\mathcal{S}$ of the BMI (1). Then clearly, $X := \begin{bmatrix} 1 \\ \tilde{x} \end{bmatrix} \begin{bmatrix} 1 & \tilde{y}^T \end{bmatrix}$ gives (19).

On the other hand, suppose there exists a dyad $X$ such that (19) holds and the $(1,1)$ element of $X$ is positive. Since (19) holds for $X$ if and only if it holds for every $\kappa X$, $\kappa > 0$, we may normalize the $(1,1)$ element of $X$, and the factorization of $X$ to $\begin{bmatrix} 1 \\ \tilde{x} \end{bmatrix} \begin{bmatrix} 1 & \tilde{y}^T \end{bmatrix}$, where $(\tilde{x},\tilde{y})$ is in the feasible set $\mathcal{S}$ trivially follows.

Suppose the $(1,1)$ element of $X$ is 0. Then if we write $X = \bar{x}\bar{y}^T$ where $\bar{x} := \begin{bmatrix} x_0 \\ x \end{bmatrix}$ and $\bar{y} := \begin{bmatrix} y_0 & y^T \end{bmatrix}$, it holds that either or both of $x_0$ and $y_0$ are zero. Assume $x_0 = 0, y_0 < 0$, then (19) implies

$$\begin{aligned} & \textstyle\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i y_j F_{i,j} + \sum_{i=1}^{n_x} x_i y_0 F_{i,0} > 0 \\ \Leftrightarrow\ & \textstyle\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} x_i \frac{1}{y_0} y_j F_{i,j} + \sum_{i=1}^{n_x} x_i F_{i,0} < 0 \end{aligned}$$

so that there exists a $\kappa < 0$ such that $F(\kappa x, \frac{1}{y_0} y) > 0$. The proof for the cases $y_0 > 0$ and $(x_0, y_0) = (0,0)$ are similar. *Q.E.D.*

We also have the following corollary:

**Corollary 3.1** *If the BMI (1) is linear in either $x$ or $y$ or both, then the BMI (1) has a solution if and only if there exists a dyad $X \in \mathbf{R}^{(n_x+1)\times(n_y+1)}$ such that (19) holds for all $W \in \mathcal{W}(F_{i,j})$.*

For the rest of the discussion in this section, we will assume that the BMI (1) is bilinear.

## 3.1 Finding the Minimum Norm Element

Now, clearly, a necessary condition for the *dyad* $X := x_o y_o^T$ to exist such that (19) holds is for co $\mathcal{W}(F_{i,j})$ to exclude the origin. However, the condition is not sufficient, because a dyad $x_o y_o^T$ fulfilling condition (19) may not exist even if co $\mathcal{W}(F_{i,j})$ excludes the origin.

We introduce the following terminology:

$$\Pr(\mathcal{W}(F_{i,j})) := \arg \min_{W \in \operatorname{co}\mathcal{W}(F_{i,j})} \|W\|_F \tag{20}$$

i.e., $\Pr(\mathcal{W}(F_{i,j}))$ is the minimum norm element of co $\mathcal{W}(F_{i,j})$.

Given the convex set co $\mathcal{W}(F_{i,j})$, it is fairly easy to find $\Pr(\mathcal{W}(F_{i,j}))$ using the algorithm of [16], with its various refinements [4, 39, 18], provided that given any $W_0 \in \mathcal{W}(F_{i,j})$, a $V \in \mathcal{W}(F_{i,j})$ such that $\langle W_0, V\rangle = \min_{W\in\mathcal{W}(F_{i,j})} \langle W_0, W\rangle$ can be found. This is achieved as follows:

First note that given $W_0 \in \mathcal{W}(F_{i,j})$,

$$\begin{aligned} \min_{W\in\mathcal{W}(F_{i,j})} \langle W_0, W\rangle &= \min_{\|z\|_2=1} \left\langle W_0, \left[z^T F_{i,j} z\right]_{i,j} \right\rangle \\ &= \underline{\lambda}\left\{ \sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_{i,j}^0 F_{i,j} \right\} \end{aligned} \tag{21}$$

which is of course easily computed. The minimizing $V \in \mathcal{W}(F_{i,j})$ is then obtained from setting $z$ to be the normalized eigenvector corresponding to $\underline{\lambda}\left\{\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} w_{i,j}^0 F_{i,j}\right\}$. Note that $w_{i,j}^0$, $i = 1,\dots,n_x$, $j = 1,\dots,n_y$ are the entries of $W_0$. The connection to the upper bound (12) should be noted.

The real question then is as follows: Given that there exists a known $\bar{W} := \Pr(\mathcal{W}(F_{i,j}))$ of rank greater than 1, such that $\langle \bar{W}, W\rangle > 0$, for all $W \in \mathcal{W}(F_{i,j})$, can we find a dyad $xy^T$ such that $\langle xy^T, W\rangle > 0$, $\forall\, W \in \mathcal{W}(F_{i,j})$ ?

## 3.2 Finding the Dyad

Given a matrix $H \in \mathbf{R}^{n_x \times n_y}$, the "best" approximation of a matrix $H$ with a dyad is the one given by the singular vectors of $H$ corresponding to its largest singular value.

**Proposition 3.1** *Given $H \in \mathbf{R}^{p\times q}$ of rank $\rho$, with non-zero singular values $\sigma_1 > \sigma_2 > \dots > \sigma_\rho > 0$ and with corresponding left and right orthonormal singular vector sets $u_1, u_2, \dots, u_p$, and $v_1, v_2, \dots, v_q$, it holds that*

$$\sigma_1 u_1 v_1^T = \arg \min_{x\in\mathbf{R}^p, y\in\mathbf{R}^q} \|H - xy^T\|_F \tag{22}$$

*and the minimum value is $\sqrt{\|H\|_F^2 - \sigma_1^2}$. Further, for any $\gamma > 0$,*

$$\gamma u_1 v_1^T = \arg \max_{x\in\mathbf{R}^p, y\in\mathbf{R}^q} \frac{\langle H, xy^T\rangle}{\|H\|_F \|xy^T\|_F} \tag{23}$$

*and the maximum value is $\frac{\sigma_1}{\|H\|_F}$.*

As we have shown above, given a matrix $H \neq 0$, finding a dyad $xy^T$ so that $\langle xy^T, H\rangle > 0$ is easy. Suppose now we have a set of matrices $Q := \{H_1, \dots, H_N\}$. The problem

$$\max_{(x,y)} \min_k \langle xy^T, H_k\rangle \tag{24}$$

is considerably more difficult. In fact, [36] shows that the problem is equivalent to the problem of determining whether the diameter of an intersection of ellipsoids centered about the origin is greater than 2.

Choose $\rho > \max_{H_i\in Q} \bar{\sigma}(H_i)$ so that $\begin{bmatrix} I & \frac{1}{\rho}H_i \\ \frac{1}{\rho}H_i^T & I \end{bmatrix} > 0$ for all $i = 1, \dots, N$. Define ellipsoids

$$C_i := \left\{ w \in \mathbf{R}^{n_x+n_y} : w^T \begin{bmatrix} I & \frac{1}{\rho}H_i \\ \frac{1}{\rho}H_i^T & I \end{bmatrix} w \leq 1 \right\} \tag{25}$$

The following result then holds:

**Proposition 3.2 ([36])** *There exists $(x, y)$ such that*

$$\langle xy^T, H_i\rangle > 0, \quad \forall i = 1, \dots, N \tag{26}$$

*if and only if there exists a vector $w$ such that $\|w\|_2 > 1$ and $w \in \bigcap_{i=1}^N C_i$.*

We may therefore conclude that given the set of matrices $Q$, the task of finding a dyad which separates the matrices from the origin is equivalent to the optimization problem

$$\max \|w\|_2^2 \qquad \text{subject to} \quad w \in \bigcap_{i=1}^N C_i \tag{27}$$

a *maximization* of a convex function over a convex set $\bigcap_{i=1}^N C_i$; a difficult global optimization problem although various algorithms exist, see [21].

# 4 Local Optimization Approaches

While it is clear from the preceding sections that the use of standard nonlinear programming techniques on the function $\Lambda(x, y) := -\underline{\lambda}\{F(x, y)\}$, even if it were differentiable, is going to be suboptimal due to the multi-extremal or *non*local-global nature of $\Lambda(x, y)$, it remains true that the development of local procedures constitutes an important first step towards establishing viable global optimization approaches.

## 4.1 The BMI as a Double LMI

As is clear, if $x$ is fixed, finding a $y$ such that (2) holds is a LMI feasibilty problem, which has been extensively studied and for which various reliable and efficient algorithms exist, as we have already noted. We may therefore propose the following algorithm:

**Algorithm 4.1 Alternating Minimization of $\Lambda(x, y)$**

- **given:** $k := 0$, *Arbitrary* $x^{(0)} \in \mathbf{R}^{n_x}$.
- **repeat:**

  *R1. Set* $k := k + 1$.

  *R2. Find* $y^{(k)} := \arg\min \left\{\Lambda(x^{(k-1)}, y) : y \in \mathbf{R}^{n_y}\right\}$.

  *R3. Find* $x^{(k)} := \arg\min \left\{\Lambda(x, y^{(k)}) : x \in \mathbf{R}^{n_x}\right\}$.

- **until:** $\Lambda(x^{(k-1)}, y^{(k)}) < 0$ **or** $\Lambda(x^{(k)}, y^{(k)}) < 0$.

Clearly, the subproblems R2 and R3 are unconstrained convex AMI minimization problems. Taken together, the two AMI subproblems may be regarded to be analogous to a cyclic coordinate descent algorithm in coordinates $x$ and $y$.

Unfortunately, since $\Lambda(x, y)$ is nonsmooth, $\bar{x}$ minimizing $\Lambda(x, \bar{y})$ and $\bar{y}$ minimizing $\Lambda(x, \bar{y})$ separately do not imply that $\Lambda(x, y)$ is even stationary at $(\bar{x}, \bar{y})$. As an example consider the continuous but non-differentiable function

$$f(x, y) := \max\left\{y - 2x,\ x - 2y,\ \frac{1}{4}(x^2 + y^2 - 16)\right\} \tag{28}$$

which has a minimum at $(2, 2)$. Note that $f(1, y)$ is minimized at $y = 1$ and $f(x, 1)$ is minimized at $x = 1$, but $(1, 1)$ clearly is not a minimum. It therefore holds that Algorithm 4.1 will not work in general.

## 4.2 Subgradient Descent Methods

In theory, one may use standard nonsmooth optimization methods [34, 7, 26] specialized for eigenvalue type problems [33, 32] to obtain local minima of $\Lambda(x, y)$. First note that:

**Proposition 4.1** *$\Lambda(x, y)$ is locally Lipschitz continuous.*

This follows from (5).

Let $1 \leq \bar{k}(x, y) \leq m$ denote the multiplicity of the minimum eigenvalue of $F(x, y)$. Then $\Lambda(x, y)$ is differentiable at all $(x, y)$ such that $\bar{k}(x, y) = 1$. Let $U(x, y)$ be any orthonormal matrix which spans the null space of $\underline{\lambda}\{F(x, y)\}\, I - F(x, y)$. Then $U(x, y)z$ is an orthogonal eigenvector of $F(x, y)$ corresponding to $\underline{\lambda}\{F(x, y)\}$ for any $z \in \mathbf{SR}^{\bar{k}}$. It then holds that the generalized gradient (or the subdifferential) of $\Lambda(x, y)$ is:

$$\partial\Lambda(x, y) = \operatorname{co}\left\{\begin{bmatrix} v^x \\ v^y \end{bmatrix} \in \mathbf{R}^{n_x+n_y} : v_i^x := -z^T U^T \frac{\partial}{\partial x_i} F(x, y)\, Uz,\ v_j^y := -z^T U^T \frac{\partial}{\partial y_j} F(x, y)\, Uz,\ z \in \mathbf{SR}^{\bar{k}}\right\} \tag{29}$$

where we have written $U(x,y) = U$ for brevity. The computation of $\frac{\partial}{\partial x_i}F(x,y)$ is trivial, e.g.,

$$\frac{\partial}{\partial x_i}F(x,y) = F_{i,0} + \sum_{j=1}^{n_y} y_j F_{i,j}$$

Further, if $(\bar{x}, \bar{y})$ gives a local minimum of $\Lambda(x,y)$, then

$$0 \in \partial\Lambda(\bar{x},\bar{y}) \tag{30}$$

Unfortunately, the practical implementation of subgradient descent methods remains problematic. Refer to [33, 32] for more details.

## 4.3 Interior Point Methods

As we have previously noted, interior point methods form the core of much of the progress achieved with LMI problems. However, the application of interior point methods to nonconvex problems remains rare, although we note that the proofs for convergence to *local* minima for the interior point methods given in [12], e.g. Theorems 8 and 32, do not require convexity.

Given the biaffine matrix function $F(x,y)$ of (1), introduce

$$\phi_\alpha(x,y) := \begin{cases} -\ln\,\det[\alpha I_m + F(x,y)], & \underline{\lambda}\{\alpha I_m + F(x,y)\} > 0 \\ \infty, & \text{otherwise} \end{cases} \tag{31}$$

Here, note that $\phi_\alpha(x,y)$ is convex in $(x,\alpha)$, and also convex in $(y,\alpha)$ and smooth. Clearly, $\phi_\alpha(x,y,\alpha) < \infty \Leftrightarrow \alpha - \Lambda(x,y) > 0$. For some $0 < B < \infty$, define $\mathcal{B} := \left\{(x,y) \in \mathbf{R}^{n_x \times n_y} : \|x\|_\infty \le B,\ \|y\|_\infty \le B,\ B < \infty\right\}$. Then $\phi_\alpha(x,y)$, and $\Lambda(x,y)$ are both bounded below over $\mathcal{B}$.

The derivatives of $\phi_\alpha(x,y)$ are straightforward extensions of the derivatives of the barrier functions of [37, 23] and are as given below, with $G := G(x,y,\alpha) := [\alpha I + F(x,y)]^{-1}$:

$$\frac{\partial}{\partial x_i}\phi_\alpha(x,y) = -\operatorname{tr}\{GF_{i,0}\} - \sum_{j=1}^{n_y} y_j \operatorname{tr}\{GF_{i,j}\} \tag{32}$$

$$\begin{aligned}\frac{\partial^2}{\partial x_i \partial x_k}\phi_\alpha(x,y) &= \operatorname{tr}\{GF_{i,0}GF_{k,0}\} + \sum_{j=1}^{n_y} y_j \operatorname{tr}\{GF_{i,j}GF_{k,0}\} \\ &\quad + \sum_{l=1}^{n_y} y_l \operatorname{tr}\{GF_{i,0}GF_{k,l}\} + \sum_{j=1}^{n_y}\sum_{l=1}^{n_y} y_j y_l \operatorname{tr}\{GF_{i,j}GF_{k,l}\}\end{aligned} \tag{33}$$

$$\begin{aligned}\frac{\partial^2}{\partial x_i \partial y_l}\phi_\alpha(x,y) &= \operatorname{tr}\{GF_{i,0}GF_{0,l}\} \\ &\quad + \sum_{j=1}^{n_y} y_j \operatorname{tr}\{GF_{i,j}GF_{0,l}\} + \sum_{k=1}^{n_x} x_k \operatorname{tr}\{GF_{i,0}GF_{k,l}\} \\ &\quad + \sum_{j=1}^{n_y}\sum_{k=1}^{n_x} y_j x_k \operatorname{tr}\{GF_{i,j}GF_{k,l}\} - \operatorname{tr}\{GF_{i,l}\}\end{aligned} \tag{34}$$

The following modified method of centers [22, 12, 5] type algorithm [1] may then be defined:

**Algorithm 4.2 BMI Method of Centers Algorithm**

- **given:** $k := 0$, $\theta \in (0,1)$.
  *Arbitrary* $(x^{(0)}, y^{(0)}) \in \mathcal{B}$ *and* $\alpha^{(0)} > \Lambda(x^{(0)}, y^{(0)})$.
- **repeat:**

  *R1. Find* $(x^{(k+1)}, y^{(k+1)}) := \arg\min \phi_{\alpha^{(k)}}(x,y),\ (x,y) \in \mathcal{B}$.

  *R2. Set* $\alpha^{(k+1)} := \theta\alpha^{(k)} + (1-\theta)\Lambda(x^{(k+1)}, y^{(k+1)})$.

  *R3. Set* $k := k+1$.
- **until:** $\alpha^{(k)}$ *converges or* $\alpha^{(k)} < 0$.

[1] Since $\phi_\alpha(x,y)$ is not convex, the term "centers" is used very loosely.

Note that if $\theta = 0$, then we get the "pure" method of centers algorithm as dicussed in [22, 12]. Now, notice that since $\alpha^{(k)} := \theta\alpha^{(k-1)} + (1-\theta)\Lambda(x^{(k)}, y^{(k)})$, and $\alpha^{(k-1)} > \Lambda(x^{(k)}, y^{(k)})$ always, we have

$$\alpha^{(k)} - \alpha^{(k-1)} = -(1-\theta)\left(\alpha^{(k-1)} - \Lambda(x^{(k)}, y^{(k)})\right) < 0 \tag{35}$$

so that $\alpha^{(k)}$ is strictly monotone decreasing. Since $\Lambda(x,y)$ is bounded below, there must exist $\bar{\alpha} > -\infty$ such that $\alpha^{(k)} \searrow \bar{\alpha}$. then $\alpha^{(k-1)} - \Lambda(x^{(k)}, y^{(k)}) \searrow 0$ and $\alpha^{(k)} - \Lambda(x^{(k)}, y^{(k)}) \searrow 0$. It therefore holds that Algorithm 4.2 approaches the "pure" method of centers as $k \to \infty$. From [12, Theorem 32], $\alpha^{(k)}$ then converges to a *local* minimum of $\Lambda(x,y)$ if the minimization of Step R1 is local.

We recognise that if we can obtain *global* solutions to the minimization of $\phi_\alpha(x,y)$ for each fixed $\alpha$, then Algorithm 4.2 will indeed solve the BMI problem (4). We note however that because $\phi_\alpha(x,y)$ is not continuous in general, global optimization of $\phi_\alpha(x,y)$ is difficult.

A barrier method type [12, 23] algorithm with $\phi_\alpha(x,y)$ replaced by

$$\psi_{\mu^{(k)}}(x,y,\alpha) := \alpha + \mu^{(k)}\phi_\alpha(x,y) \tag{36}$$

may also be used, where $\mu^{(k)} \searrow 0$, and at each iteration $\psi_{\mu^{(k)}}(x,y,\alpha)$ is minimized over $(x,y,\alpha)$. Note that convergence to a local minimum is then guaranteed by [12, Theorem 8].

We will now discuss the application of gradient descent methods for the local minimizations required by Algorithm 4.2. We recognise that since $\phi_\alpha(x,y)$ is self-concordant [29] with respect to $x$ if $y$ is fixed and vice-versa, if $\phi_\alpha(x,y)$ is to be minimized with respect to $x$ and $y$ alternatingly with fixed $y$ and $x$ respectively, a very effective approach would be to use the Newton's method of [29], which would give polynomial time convergence for each of the alternating minimization steps. Furthermore, since $\phi_\alpha(x,y)$ is *smooth* for the region under consideration, convergence to a stationary point of $\phi_\alpha(x,y)$ is guaranteed. Note that even if a stationary point of $\phi_\alpha(x,y)$ is reached, the possibility that the point is a saddle point, and not a local minimum, needs to be checked.

Descent along *joint* $(x,y)$ directions may also be attempted, although since the Hessian of $\phi_\alpha(x,y)$ may not be positive definite, pure Newton descent cannot be used. However, our computational experience indicates that joint descent is more efficient than alternating descent. The computation of the Hessian of $\phi_\alpha(x,y)$ with respect to $(x,y)$ is expensive, and can be circumvented by the use of the very standard Broyden-Fletcher-Goldfarb-Shanno quasi-Newton method [27, 17].

Our computational experience with the Algorithm 4.2 is favorable in general. The main difficulty that arises seems to be the problem of ill conditioning as $\alpha^{(k)} - \Lambda(x^{(k)}, y^{(k)}) \searrow 0$, which is an inherent characteristic of interior point problems, e.g. [23].

# 5 Biconvex Global Optimization

We now briefly discuss possible global optimization approaches for the problem (4). Examples of the application of global optimization techniques to robust control problems include [9, 3, 30].

We first note that, in general, global optimization problems are NP-hard. This, of course, does not mean that the problems are unsolvable, only that they will require algorithms which will not have worst case polynomial time bounds, unlike, say, LMI or linear programming problems.

With regard to our particular problem of finding a global minimum to a biconvex function, the references [1, 21] may be consulted for surveys of various approaches to the bilinear and biconvex programming problems.

Branch and bound methods seem to be a relatively promising approach for exploiting biconvexity and the preliminary study of [1] is worth further investigation in this respect. We also have the Benders decomposition approach of [15] and the references therein, which leads to the primal-relaxed dual algorithm of [13]. Multistart methods may also be used.

## 6 Summary and Conclusion

This paper presents a preliminary study of the BMI feasibility problem (2) and its associated nonsmooth biconvex optimization problem (4). Trivial cases and upper and lower bounds are examined. A geometric perspective is presented. Extensions of LMI type minimum eigenvalue maximization approaches to the BMI case are discussed. The main difficulty in approaching BMI problems seems to be its nonsmooth biconvex nature. In this respect, the key seems to be to construct a global optimization approach which fully exploits the biconvexity of the problem.

## References


[1] F.A. Al-Khayyal. Jointly constrained bilinear programs and related problems: An overview. *Computers and Mathmatics with Applications*, 19(11):53–62, 1990.

[2] J.C. Allwright. On maximizing the minimum eigenvalue of a linear combination of symmetric matrices. *SIAM J. Matrix Analysis & Applications*, 10(3):347–382, 1989.

[3] V. Balakrishnan, S. Boyd, and S. Balemi. Computing the minimum stability degree of parameter-dependent linear systems. In S.P. Bhattacharyya and L.H. Keel, editors, *Control of Uncertain Dynamic Systems*, pages 359–378. CRC Press, Inc., Boca Raton, FL., 1991.

[4] R.O. Barr. An efficient computational procedure for a generalized quadratic programming problem. *SIAM J. Control and Optimization*, 7(3):415–429, 1969.

[5] S.P. Boyd and L. El Ghaoui. Method of centers for minimizing generalized eigenvalues. *Linear Algebra and its Applications*, 188–189:63–111, 1993.

[6] S.P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*. Final Draft, December 1993.

[7] F. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley Interscience, New York, NY, 1983.

[8] J. Cullum, W.E. Donath, and P. Wolfe. The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. *Mathematical Programming Study*, 3:35–55, 1975.

[9] R.E. De Gaston and M.G. Safonov. Exact calculation of the multivariable stability margin. *IEEE Trans. on Automatic Control*, AC-33:156–171, 1988.

[10] J.C. Doyle. Structured uncertainty in control system design. *IEE Proceedings*, 129-D(6):242–250, 1982.

[11] M.K.H. Fan and A. L. Tits. *m*-form numerical range and the computation of the structured singular value. *IEEE Trans. on Automatic Control*, AC-33(3):284–289, March 1988.

[12] A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, Inc., New York, NY, 1968.

[13] C.A. Floudas and V. Visweswaran. A global optimization algorithm for certain classes of nonconvex NLPs, – I. Theory. *Computers and Chemical Engineering*, 14(12):1397–1417, 1990.

[14] P. Gahinet and A. Nemirovskii. General-purpose LMI solvers with benchmarks. In *Proc. IEEE Conf. on Decision and Control*, pages 3162–3165, San Antonio, TX, December 1993. IEEE Press, New York.

[15] A.M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.

[16] E.G. Gilbert. An iterative procedure for computing the minimum of a quadratic form on a convex set. *SIAM J. Control and Optimization*, 4(1):61–79, 1966.

[17] P.E. Gill, W. Murray, and M.H. Wright. *Practical Optimization*. Academic Press, London, 1981.

[18] J.E. Hauser. Proximity algorithms: Theory and implementation. *Electronics Research Laboratory Memorandum, College of Engineering, University of California, Berkeley*, UCB/ERL M86/53, 1986.

[19] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[20] R. Horst. A note on functions whose local minima are global. *Journal of Optimization Theory and Applications*, 36(3):457–463, 1982.

[21] R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer-Verlag, Berlin, second, revised edition, 1993.

[22] P. Huard. Resolution of mathematical programming with nonlinear constraints by the method of centers. In J. Abadie, editor, *Nonlinear Programming*. North-Holland Publishing Company, Amsterdam, 1967.

[23] F. Jarre. A interior point method for minimizing the maximum eigenvalue of a linear combination of matrices. *SIAM J. Control and Optimization*, 31(5):1360–1377, September 1993.

[24] N. Kamarkar. A new polynomial-time algorithm for linear programming. *Combinator.*, 4:373–395, 1984.

[25] T. Kato. *A Short Introduction to Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin-New York, 1982.

[26] C. Lemarechal. Nondifferentiable optimization. In G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, editors, *Optimization*, volume 1 of *Handbooks in Operations Research and Management Science*, pages 529–572. Elsevier Science Publishers B.V., 1989.

[27] D.G. Luenberger. *Linear and Nonlinear Programming*. John Wiley, New York, second edition edition, 1984.

[28] Yu. Nesterov and A. Nemirovskii. *Self-Concordant Functions and Polynomial-Time Methods in Convex Programming*. Central Economic and Mathematical Institute, USSR Acad. Sci., Moscow, 1989.

[29] Yu. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Publications, Philadelphia, PA, 1994.

[30] M. Newlin and P.M. Young. Mixed $\mu$ problems and branch and bound techniques. In *Proc. IEEE Conf. on Decision and Control*, pages 3175–3180, Tucson,AZ., December 16–18, 1992. IEEE Press.

[31] M.L. Overton. On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM J. Matrix Analysis & Applications*, 9(2):256–268, 1988.

[32] E. Polak. On the mathematical foundations of nondifferentiable optimization in engineering design. *SIAM Review*, 29(1):21–89, March 1987.

[33] E. Polak and Y. Wardi. Nondifferentiable optimization algorithm for designing control systems having singular value inequalities. *Automatica*, 18(3):267–283, 1982.

[34] R.T. Rockefellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[35] M.G. Safonov, K.C. Goh, and J.H. Ly. Controller synthesis via Bilinear Matrix Inequalities. In *Proc. American Control Conf.*, 1994. To be presented, July, 1994.

[36] M.G. Safonov and G.P. Papavassilopoulos. The diameter of an intersection of ellipsoids and BMI robust synthesis. In *Proceedings of the IFAC Symposium on Robust Control Design*, Rio de Janeiro, Brazil, 14–16 September, 1994. Submitted for consideration, November 1993.

[37] L. Vandenberghe and S.P. Boyd. Primal-dual potential reduction method for problems involving matrix inequalities. *Technical Report, I.S.L., Stanford University*, January 1993.

[38] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford University Press, 1965.

[39] P. Wolfe. Finding the nearest point in a polytope. *Mathematical Programming*, 11:128–149, 1976.