

Global Optimization for the Biaffine Matrix Inequality Problem

KEAT-CHOON GOH, MICHAEL G. SAFONOV and
GEORGE P. PAPAVASSILOPOULOS

*Electrical Engineering - Systems, University of Southern California,
Los Angeles, CA 90089–2563, U.S.A. (Email: goh@bode.usc.edu)*

(Received: 31 January 1995; accepted: 1 June 1995)

Abstract. It has recently been shown that an extremely wide array of robust controller design problems may be reduced to the problem of finding a feasible point under a Biaffine Matrix Inequality (BMI) constraint. The BMI feasibility problem is the bilinear version of the Linear (Affine) Matrix Inequality (LMI) feasibility problem, and may also be viewed as a bilinear extension to the Semidefinite Programming (SDP) problem. The BMI problem may be approached as a biconvex global optimization problem of minimizing the maximum eigenvalue of a biaffine combination of symmetric matrices. This paper presents a branch and bound global optimization algorithm for the BMI. A simple numerical example is included. The robust control problem, i.e., the synthesis of a controller for a dynamic physical system which guarantees stability and performance in the face of significant modelling error and worst-case disturbance inputs, is frequently encountered in a variety of complex engineering applications including the design of aircraft, satellites, chemical plants, and other precision positioning and tracking systems.

Key words: Branch and bound, bilinear matrix inequalities, linear matrix inequalities, robust control

1. Introduction

This paper discusses the Biaffine Matrix Inequality (BMI) feasibility problem introduced in [35] and presents a global optimization algorithm for its solution. The BMI feasibility problem is of great significance in the field of robust control theory because it captures the essence of the robust control synthesis problem. Finding efficient ways to solve the BMI feasibility problem is therefore of great interest, and will find many applications in the control of space structures, high performance fighter aircraft, distillation columns, robots, CD players, disk drives and other areas where guaranteed worst case performance is required.

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 $F : \mathbf{R}^{n_x} \times \mathbf{R}^{n_y} \rightarrow \mathbf{R}^{m \times m}$,

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

The *BMI Feasibility Problem* is to find, if it exists, $(x, y) \in \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$ such that

$$F(x, y) < 0. \tag{2}$$

Notation will be standard. In particular, \mathbf{R} denotes the set of real numbers, and \mathbf{R}^n the n -dimensional real vector space. For any vector $z \in \mathbf{R}^n$, $\|z\|_\infty := \max_{i \in \{1, \dots, n\}} |z_i|$, and for any $(x, y) \in \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$, $\|(x, y)\|_\infty := \left\| \begin{bmatrix} x^T \\ y^T \end{bmatrix}^T \right\|_\infty$. The space of $m \times m$ real matrices will be denoted $\mathbf{R}^{m \times m}$. For symmetric matrices A and B , $\bar{\lambda}\{A\}$ and $\underline{\lambda}\{A\}$ refer to the greatest (most positive) and smallest (most negative) eigenvalues of A , and $A > 0$ means $\underline{\lambda}\{A\} > 0$, $A > B$ means $A - B > 0$. For scalars \underline{b} , $\bar{b} \in \mathbf{R}$ with $\underline{b} \leq \bar{b}$, $[\underline{b}, \bar{b}]$ denotes a closed interval $\subset \mathbf{R}$.

This paper will mainly be concerned with the problem of finding a global solution to the following minimax problem. Given $F(x, y)$ of (1), define

$$\Lambda(x, y) := \bar{\lambda}\{F(x, y)\}. \tag{3}$$

The *BMI Eigenvalue Minimization Problem* is

$$\min_{(x, y)} \Lambda(x, y). \tag{4}$$

Clearly, there exists a solution (x, y) to the BMI feasibility problem (2) if and only if $0 > \min_{(x, y)} \Lambda(x, y)$.

The Linear or Affine Matrix Inequality (LMI/AMI) or Semidefinite Programming (SDP) problem of, e.g. [3, 29, 8, 38], characterized by the problem

$$\min_x \bar{\lambda} \left\{ F_0 + \sum_{i=1}^n x_i F_i \right\}, \tag{5}$$

where $F_i = F_i^T$, is a special case of the BMI problem which emerges if either one of x or y in (4) is fixed. The LMI/SDP problem is convex, and relatively efficient polynomial time algorithms for its solution exist [17, 37].

It may be shown that the function $\Lambda(x, y)$ is convex in x for fixed y , and convex in y for fixed x , but not convex in (x, y) in general. Further, the function is not local-global in general; i.e. the function $\Lambda(x, y)$ may have local minima which are not global minima. While it is fairly straightforward to find at least one local minimum of the $\Lambda(x, y)$ over some compact domain $M \subset \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$, the complete solution of the minimization problem, i.e., for some $\epsilon > 0$, find any (\bar{x}, \bar{y}) such that $\Lambda(\bar{x}, \bar{y}) \leq \Lambda(x, y) + \epsilon$ for all $(x, y) \in M \subset \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$, is a *global* optimization problem. Hence, unlike the LMI or SDP problems, the general BMI problem is computationally “hard” [18, 31, 28, 36], and it is unlikely that computational algorithms with polynomial time worst case performance bounds exist for the BMI problem.

While the BMI feasibility problem is difficult, it has several agreeable features which may be exploited.

- Since the problem is bilinear, an obvious approach to the minimization problem (4) is to minimize $\Lambda(x, y)$ alternatingly over x with y fixed, and vice versa.
- The close relation of the BMI problem to the LMI problem allows various adaptations of LMI approaches for the efficient computation of local minima with respect to (x, y) jointly.
- The bilinear form of the semidefiniteness constraint suggests an LMI relaxation of the problem for calculating lower bounds, and for obtaining good starting points for local optimization.

The contribution of this paper is to utilize the above characteristics to present a branch and bound global optimization algorithm for the BMI which finds an ϵ -global minimum of $\Lambda(x, y)$ over any fixed compact hyper-rectangle within a finite number of iterations. Branch and bound algorithms do not have polynomial time worst case performance bounds, but are nevertheless effective for small problems, and can be useful for finding improved suboptimal solutions for large problems. An earlier version of this work is given in [22].

2. Robust Control Motivation

The robust control motivation for the study of the BMI problem is now briefly outlined. Robust control deals with the analysis and synthesis of dynamic physical systems which are subject to significant unknown but bounded modelling uncertainties and inaccuracy, and disturbance inputs, see, e.g. [11]. The physical systems to be controlled may include satellites, fighter aircraft, distillation columns, disk drives etc. The robust control synthesis problem is to design a dynamical controller which guarantees worst-case stability and performance.

It is now clear that while the robust *analysis* problem may be recast as an LMI problem, e.g. [14, 34], the general problem of *synthesizing* a robust controller is a BMI problem [20, 21]. In fact, it can be shown that a wide array of difficult control problems, such as μ/K_m -synthesis, decentralized control, constrained order \mathcal{H}^∞ synthesis, the synthesis of one controller for multiple plants etc. reduce to BMI problems.

It may be noted that global optimization techniques have previously been applied to robustness analysis [10, 5, 30, 6], and another robust control problem which may benefit from a global optimization approach is the LMI with rank constraints problem of, e.g. [13, 9].

As an example, the μ/K_m -synthesis problem [33, 12], which is in many senses, the key problem in robust control, one needs to *simultaneously* find a suitable *multiplier* and controller pair. In [35, 21] it is shown that the problem is equivalent to the BMI problem of solving for the real matrices T, S, P, W and Q , where T, S, P are symmetric, such that the following matrix inequalities hold,

$$\text{herm} \left\{ \begin{bmatrix} T & 0 \\ 0 & W \end{bmatrix} (R_{MT} + U_{MT} Q V_{MT}) \right\} > 0,$$

$$P > 0, \quad -\text{herm}\{P(R_{G_A} + U_{G_A} Q V_{G_A})\} > 0,$$

$$\text{herm}\left\{\begin{bmatrix} S & 0 \\ 0 & W \end{bmatrix} M\right\} > 0.$$

where for any square matrix X , $\text{herm}\{X\} := \frac{1}{2}(X + X^T)$, and $M, R_{MT}, U_{MT}, V_{MT}, R_{G_A}, U_{G_A}, V_{G_A}$ are known/prescribed matrices. Local minimization procedures coupled with the ability to make good initial guesses are often sufficient to give significantly improved results compared with other currently available approaches.

3. Preliminaries

Several important facts about the BMI problem are listed. See [20] for further details. Firstly, note that for any two symmetric matrices A, B , $\bar{\lambda}\{A\} + \underline{\lambda}\{B\} \leq \bar{\lambda}\{A + B\} \leq \bar{\lambda}\{A\} + \bar{\lambda}\{B\}$, e.g. [24, Theorem 4.3.1]. It then follows that $\Lambda(x, y)$ is continuous in (x, y) . Further, $\Lambda(x, y)$ is nondifferentiable, but expressions for its subdifferential at any given (x, y) may be obtained.

A function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ is said to be Lipschitzian on some compact $M \subset \mathbf{R}^n$ if there exists an L such that $|f(x) - f(y)| \leq L\|x - y\|, \forall x, y \in M$, where $\|\cdot\|$ denotes any norm on \mathbf{R}^n . Since the subdifferential of $\Lambda(x, y)$ of (4) is uniformly bounded over any compact $M \subset \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$, it follows that $\Lambda(x, y)$ is Lipschitzian on any compact M .

For the rest of this paper, restrict the domain of $\Lambda(x, y)$ of (4) to some closed bounded hyper-rectangle $X \times Y \subset \mathbf{R}^{n_x} \times \mathbf{R}^{n_y}$ where

$$X := [\underline{b}_{x_1}, \bar{b}_{x_1}] \times \dots \times [\underline{b}_{x_{n_x}}, \bar{b}_{x_{n_x}}], \tag{6}$$

$$Y := [\underline{b}_{y_1}, \bar{b}_{y_1}] \times \dots \times [\underline{b}_{y_{n_y}}, \bar{b}_{y_{n_y}}], \tag{7}$$

for some bounds $-\infty < \underline{b}_{x_i} \leq \bar{b}_{x_i} < \infty, i = 1 \dots, n_x$, and $-\infty < \underline{b}_{y_j} \leq \bar{b}_{y_j} < \infty, j = 1 \dots, n_y$. Consider any closed hyper-rectangle $\mathcal{Q} \subset X \times Y$ of the form

$$\mathcal{Q} := [\underline{a}_{x_1}, \bar{a}_{x_1}] \times \dots \times [\underline{a}_{x_{n_x}}, \bar{a}_{x_{n_x}}] \times [\underline{a}_{y_1}, \bar{a}_{y_1}] \times \dots \times [\underline{a}_{y_{n_y}}, \bar{a}_{y_{n_y}}], \tag{8}$$

where \underline{a}_{x_i} and \bar{a}_{x_i} are the lower and upper bounds on the variable x_i for the rectangle \mathcal{Q} , and $\underline{b}_{x_i} \leq \underline{a}_{x_i} \leq \bar{a}_{x_i} \leq \bar{b}_{x_i}$. Similarly for y_j variables. Note that if $x_i \in [\underline{a}_{x_i}, \bar{a}_{x_i}]$ and $y_j \in [\underline{a}_{y_j}, \bar{a}_{y_j}]$, then $x_i y_j \in [\underline{a}_{w_{i,j}}, \bar{a}_{w_{i,j}}]$ where

$$\underline{a}_{w_{i,j}} := \min \left\{ \underline{a}_{x_i} \underline{a}_{y_j}, \underline{a}_{x_i} \bar{a}_{y_j}, \bar{a}_{x_i} \underline{a}_{y_j}, \bar{a}_{x_i} \bar{a}_{y_j} \right\}, \tag{9}$$

$$\bar{a}_{w_{i,j}} := \max \left\{ \underline{a}_{x_i} \underline{a}_{y_j}, \underline{a}_{x_i} \bar{a}_{y_j}, \bar{a}_{x_i} \underline{a}_{y_j}, \bar{a}_{x_i} \bar{a}_{y_j} \right\}. \tag{10}$$

Given any \mathcal{Q} , define the hyper-rectangle $\mathcal{W}(\mathcal{Q}) \subset \mathbf{R}^{n_x \times n_y}$ as follows

$$\mathcal{W}(\mathcal{Q}) := \left\{ W \in \mathbf{R}^{n_x \times n_y} : w_{i,j} \in [\underline{a}_{w_{i,j}}, \bar{a}_{w_{i,j}}] \right\}. \tag{11}$$

Define the *affine* matrix function of x, y and W ,

$$F_L(x, y, W) := F_{0,0} + \sum_{i=1}^{n_x} x_i F_{i,0} + \sum_{j=1}^{n_y} y_j F_{0,j} + \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} w_{i,j} F_{i,j}, \quad (12)$$

where $W \in \mathcal{W}(\mathcal{Q})$. Define also

$$\Lambda_L(x, y, W) := \bar{\lambda}\{F_L(x, y, W)\}. \quad (13)$$

Note that $\Lambda_L(x, y, W)$ is a convex function of (x, y, W) and its global minimum over $\mathcal{W}(\mathcal{Q})$ may be computed reasonably efficiently using standard LMI solvers [37, 17].

The global minimum of $\Lambda_L(x, y, W)$ over $\mathcal{W}(\mathcal{Q})$ gives a lower bound to $\Lambda(x, y)$ over \mathcal{Q} . This follows immediately from the fact that if $(x, y) \in \mathcal{Q}$, then the dyad $xy^T \in \mathcal{W}(\mathcal{Q})$.

PROPOSITION 1. *It holds that*

$$\min_{(x,y) \in \mathcal{Q}} \bar{\lambda}\{F(x, y)\} \geq \min_{(x,y) \in \mathcal{Q}, W \in \mathcal{W}(\mathcal{Q})} \bar{\lambda}\{F_L(x, y, W)\}. \quad (14)$$

Define the size of a hyper-rectangle $M \in \mathbf{R}^p$, for any integer p , by

$$\text{size}\{M\} := \max \left\{ \|z^{(1)} - z^{(2)}\|_\infty : z^{(1)}, z^{(2)} \in M \right\}. \quad (15)$$

For $\mathcal{Q} \in X \times Y$ of the form (8), and $\mathcal{W}(\mathcal{Q})$ of (11), it follows that

$$\begin{aligned} \text{size}\{\mathcal{Q}\} &= \max \left\{ \max_i (\bar{b}_{x_i} - \underline{b}_{x_i}), \max_j (\bar{b}_{y_j} - \underline{b}_{y_j}) \right\}, \\ \text{size}\{\mathcal{W}(\mathcal{Q})\} &= \max_{i,j} (\bar{a}_{w_{i,j}} - \underline{a}_{w_{i,j}}), \end{aligned}$$

i.e., the lengths of their longest sides.

LEMMA 2. *Suppose $B := \max_{\mathcal{Q}} \|(x, y)\|_\infty$. For any $\delta > 0$, if $\text{size}\{\mathcal{Q}\} \leq \delta$, then $\text{size}\{\mathcal{W}(\mathcal{Q})\} \leq \delta(2B + \delta)$.*

Proof. If $\text{size}\{\mathcal{Q}\} \leq \delta$, assume without loss of generality that for some i and some j , $x_i \in [\underline{a}_{x_i}, \underline{a}_{x_i} + \delta]$, $y_j \in [\underline{a}_{y_j}, \underline{a}_{y_j} + \delta]$. It then follows that if $w_{i,j} \in [\underline{a}_{w_{i,j}}, \bar{a}_{w_{i,j}}]$, then $\underline{a}_{w_{i,j}} = \min \left\{ \underline{a}_{x_i} \underline{a}_{y_j}, \underline{a}_{x_i} \underline{a}_{y_j} + \underline{a}_{x_i} \delta, \underline{a}_{x_i} \underline{a}_{y_j} + \underline{a}_{y_j} \delta, \underline{a}_{x_i} \underline{a}_{y_j} + \underline{a}_{x_i} \delta + \underline{a}_{y_j} \delta + \delta^2 \right\}$, and similarly for $\bar{a}_{w_{i,j}}$ (with a max instead of a min). Clearly,

$$\bar{a}_{w_{i,j}} - \underline{a}_{w_{i,j}} = \max \left\{ |\underline{a}_{x_i} \delta|, |\underline{a}_{y_j} \delta|, |\underline{a}_{x_i} \delta + \underline{a}_{y_j} \delta + \delta^2| \right\}.$$

The lemma follows by noting that, necessarily, $|\underline{a}_{x_i}| < B$, $|\underline{a}_{y_j}| < B$.

4. Local Minimization Approaches

This section will briefly discuss various approaches to find a local minimum for the function $\Lambda(x, y)$. Local minimization algorithms give upper bounds for problem (4), and are sometimes sufficient for practical purposes.

The minimization (4) is very closely related to the LMI eigenvalue minimization problem and a possible approach is to use currently available LMI algorithms to alternately minimize $\Lambda(x, y)$ with respect to x with y fixed and vice versa. See Algorithm 1 as follows.

ALGORITHM 1. Alternating LMI for BMI.

Fix $\epsilon > 0$. Select some $(x^{(0)}, y^{(0)})$. $k := 0$.
 Repeat, {
 R1. $x^{(k+1)} := \arg \min \Lambda(x, y^{(k)})$.
 R2. $y^{(k+1)} := \arg \min \Lambda(x^{(k+1)}, y)$.
 R3. $k := k + 1$.
 } until $\Lambda(x^{(k-1)}, y^{(k-1)}) - \Lambda(x^{(k)}, y^{(k)}) < \epsilon$.

For a number of reasons, Algorithm 1 is not guaranteed to converge to a local minimum of $\Lambda(x, y)$, see, e.g. [39, 20]. As an example, consider the BMI problem where $\Lambda(x, y) = \max \{y - 2x, x - 2y, xy - 6\}$. Clearly, $(x, y) = (2, 2)$ gives a global minimum. Note that $\Lambda(1, y)$ is minimized by $y = 1$, and $\Lambda(x, 1)$ is minimized by $x = 1$, but the non-differentiable point $(1, 1)$ is not a local minimum of $\Lambda(x, y)$.

A simple approach for the BMI problem which converges to a local minimum (under certain reasonable assumptions), based on the Method of Centers of [26, 15, 27, 7], is now presented. Given the biaffine matrix function $F(x, y)$ of (1), introduce

$$\phi_\alpha(x, y) := \begin{cases} -\log \det[\alpha I - F(x, y)], & \underline{\lambda}\{\alpha I - F(x, y)\} > 0, \\ \infty, & \text{otherwise.} \end{cases} \quad (16)$$

Note that the barrier function $\phi_\alpha(x, y)$ is convex in (x, α) , and also convex in (y, α) . Furthermore, for all (x, y) such that $\underline{\lambda}\{\alpha I - F(x, y)\} > 0$, $\phi_\alpha(x, y)$ is smooth and at least twice differentiable, with derivatives given in [20]. A method of centers algorithm, Algorithm 2, for the BMI follows.

ALGORITHM 2. Method of Centers for BMI [23].

Fix $\epsilon > 0$, $\delta > 0$ and $\theta \in (0, 1)$.

Choose some $(x^{(0)}, y^{(0)})$. Set $\alpha^{(0)} := \delta + \Lambda(x^{(0)}, y^{(0)})$, $k := 0$.
 Repeat, {
 R1. $(x^{(k+1)}, y^{(k+1)}) := \arg \text{local min } \phi_{\alpha^{(k)}}(x, y)$.
 R2. $\alpha^{(k+1)} := (1 - \theta)\Lambda(x^{(k+1)}, y^{(k+1)}) + \theta\alpha^{(k)}$.

R3. $k := k + 1$.
 } until $\Lambda(x^{(k-1)}, y^{(k-1)}) - \Lambda(x^{(k)}, y^{(k)}) < \epsilon$.

In Algorithm 2, the minimization in Step R1 is a local minimization with initial point $(x^{(k)}, y^{(k)})$. It may be shown that $\alpha^{(k)} > \alpha^{(k+1)}$ for all k , and under minor assumptions on the uniformity of the local minimization, $(x^{(k)}, y^{(k)})$ will converge to a local minimum of $\Lambda(x, y)$.

5. BMI Branch and Bound

A branch and bound global optimization algorithm which finds an ϵ -global minimum of $\Lambda(x, y)$ over the compact hyper-rectangle $X \times Y$ in a finite number of iterations is now presented. An ϵ -global minimum of $\Lambda(x, y)$ is any (\bar{x}, \bar{y}) such that $\Lambda(\bar{x}, \bar{y}) \leq \Lambda(x, y) + \epsilon$ for all $(x, y) \in X \times Y$. Much of the following notation and terminology is from [5, 4] and the global optimization text [25].

Before presenting a branch and bound algorithm for the BMI, it should be noted that it is, in theory, possible to establish the global optimum to any given tolerance by *exhaustively gridding* the entire domain. This follows from the fact that $\Lambda(x, y)$ is Lipschitz over any bounded domain. However, since local optimization algorithms are available for $\Lambda(x, y)$, a branch and bound approach will generally be very much more efficient; provided that given any set $Q \subset X \times Y$, one can calculate a good lower bound for $\Lambda(x, y)$ over Q .

5.1. UPPER AND LOWER BOUNDS

The objective is to minimize $\Lambda(x, y)$ over the domain $X \times Y$. The basic requirement for a branch and bound algorithm for globally minimizing $\Lambda(x, y)$ is for the existence of two functions, Φ_L and Φ_U , on the family of hyper-rectangles of the form (8), which $X \times Y$ will be partitioned into, such that the following conditions, C1 and C2, hold.

C1. For every hyper-rectangle $Q \subset X \times Y$, $\Phi_L(Q)$ gives a lower bound, and $\Phi_U(Q)$ an upper bound, on $\min \Lambda(x, y)$ over Q , i.e.,

$$\Phi_L(Q) \leq \min_{(x,y) \in Q} \Lambda(x, y) \leq \Phi_U(Q). \tag{17}$$

C2. As $\text{size}\{Q\} \searrow 0$, $\Phi_U(Q) - \Phi_L(Q) \searrow 0$ uniformly, i.e.,

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that,} \\ \forall Q \subset X \times Y, \text{size}\{Q\} \leq \delta \Rightarrow \Phi_U(Q) - \Phi_L(Q) < \epsilon. \tag{18}$$

Define the following functions on the family of hyper-rectangles of the form $Q \subset X \times Y$,

$$\Phi_U(Q) := \text{local } \min_{(x,y) \in Q} \Lambda(x, y), \tag{19}$$

$$\Phi_L(Q) := \min_{(x,y) \in Q, W \in \mathcal{W}(Q)} \Lambda_L(x, y, W). \tag{20}$$

The function $\Phi_L(\mathcal{Q})$ is well defined, by the convexity of $\Lambda_L(x, y, W)$ over $(x, y) \in \mathcal{Q}$, $W \in \mathcal{W}(\mathcal{Q})$, and may be computed reasonably efficiently in polynomial time. However, the function $\Phi_U(\mathcal{Q})$ of (19) requires some further clarification.

Fix \mathcal{Q} and suppose $(x^*, y^*, W^*) := \arg \min_{(x,y) \in \mathcal{Q}, W \in \mathcal{W}(\mathcal{Q})} \Lambda_L(x, y, W)$. For the purposes of the proof that the algorithm to be presented will converge in finite time, all that is required is for $\Phi_U(\mathcal{Q}) \leq \Lambda(x^*, y^*)$. This will be assumed to hold, and will certainly hold if $\Phi_U(\mathcal{Q})$ is computed using any reasonable local minimization algorithm with initial points (x^*, y^*) .

ASSUMPTION 1. Fix \mathcal{Q} . If $(x^*, y^*, W^*) := \arg \min_{(x,y) \in \mathcal{Q}, W \in \mathcal{W}(\mathcal{Q})} \Lambda_L(x, y, W)$, then $\Phi_U(\mathcal{Q}) \leq \Lambda(x^*, y^*)$.

Of course, in general, it would be expected that $\Phi_U(\mathcal{Q})$ be significantly less than $\Lambda(x^*, y^*)$. Clearly, for any \mathcal{Q} , a reasonable and sound way to compute $\Phi_U(\mathcal{Q})$ will be to use either Algorithm 1 or Algorithm 2, or, first Algorithm 1, and then followed by Algorithm 2.

THEOREM 3. The functions Φ_U and Φ_L as defined in (19) and (20), on the family of sets $\{\mathcal{Q} : \mathcal{Q} \subset X \times Y\}$, fulfil conditions **C1** and **C2**.

Proof. Firstly, Proposition 1 implies that Φ_U and Φ_L satisfy condition **C1**.

From the inequality $\bar{\lambda}\{A + B\} \leq \bar{\lambda}\{A\} + \bar{\lambda}\{B\}$ for hermitian matrices, and defining $v_{i,j} := v_{i,j}(x_i, y_j, w_{i,j}) = x_i y_j - w_{i,j}$,

$$\begin{aligned} \Lambda(x, y) &\leq \Lambda_L(x, y, W) + \bar{\lambda} \left\{ \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} (x_i y_j - w_{i,j}) F_{i,j} \right\} \\ \Rightarrow \Lambda(x, y) - \Lambda_L(x, y, W) &\leq \bar{\lambda} \left\{ \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} v_{i,j} F_{i,j} \right\} \\ \Rightarrow \Lambda(x^*, y^*) - \Lambda_L(x^*, y^*, W^*) &\leq \bar{\lambda} \left\{ \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} v_{i,j}^* F_{i,j} \right\}, \end{aligned}$$

where $(x^*, y^*, W^*) := \arg \min_{(x,y) \in \mathcal{Q}, W \in \mathcal{W}(\mathcal{Q})} \Lambda_L(x, y, W)$, and $v_{i,j}^* := v_{i,j}(x_i^*, y_j^*, w_{i,j}^*)$.

From Lemma 2, for scalar $B := \max_{X \times Y} \|(x, y)\|_\infty$,

$$\begin{aligned} \text{size}\{\mathcal{Q}\} \leq \delta &\Rightarrow \text{size}\{\mathcal{W}(\mathcal{Q})\} \leq \delta(2B + \delta) \\ \Rightarrow |v_{i,j}| &\leq \delta(2B + \delta), \forall i \in \{1, \dots, n_x\}, \forall j \in \{1, \dots, n_y\}. \end{aligned}$$

Define $\Delta(\mathcal{Q}) := \Phi_U(\mathcal{Q}) - \Phi_L(\mathcal{Q})$. Clearly, $\Delta(\mathcal{Q}) \geq 0$, and by Assumption 5.1, $\Delta(\mathcal{Q}) \leq \Lambda(x^*, y^*) - \Lambda_L(x^*, y^*, W^*)$. It follows that

$$\Delta(\mathcal{Q}) \leq \bar{\lambda} \left\{ \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} v_{i,j}^* F_{i,j} \right\} \leq \max_{|v_{i,j}| \leq \delta(2B + \delta)} \bar{\lambda} \left\{ \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} v_{i,j} F_{i,j} \right\},$$

so that condition **C2** follows.

5.2. A BMI BRANCH AND BOUND ALGORITHM

Given the functions $\Phi_U(Q)$, $\Phi_L(Q)$ and Theorem 3, it is straightforward to adapt the branch and bound algorithm given in [5, 4] to yield Algorithm 3. See also [25].

ALGORITHM 3. Branch and Bound Algorithm for BMI

Fix $\epsilon > 0$. Set $k := 0$, $Q_0 := X \times Y$, $S_0 := \{Q_0\}$.

$L_0 := \Phi_L(Q_0)$, $U_0 := \Phi_U(Q_0)$.

Repeat {

R1. Select \bar{Q} from S_k such that

$L_k = \Phi_L(\bar{Q})$.

R2. Split \bar{Q} along its longest edge into \bar{Q}_1 and \bar{Q}_2 .

R3. For $i = 1, 2$, {
 Compute $\Phi_L(\bar{Q}_i)$.

If $\Phi_L(\bar{Q}_i) \leq U_k$, {
 Compute $\Phi_U(\bar{Q}_i)$.

}

R4. $S_{k+1} := \{S_k - \{\bar{Q}\}\} \cup \{\bar{Q}_1, \bar{Q}_2\}$.

R5. $U_{k+1} := \min \{U_k, \min_{Q \in S_{k+1}} \Phi_U(Q)\}$. $L_{k+1} := \min_{Q \in S_{k+1}} \Phi_L(Q)$.

R6. $S_{k+1} = S_{k+1} - \{Q : \Phi_L(Q) > U_{k+1}\}$.

R7. $k := k + 1$.
 } until $U_k - L_k < \epsilon$.

In Algorithm 3, S_k is the collection of hyper-rectangles $\{Q_1, Q_2, \dots, Q_{\bar{k}}\}$ after k iterations, where $\bar{k} \leq k$. At the $(k + 1)$ th iteration, \bar{Q} , the hyper-rectangle in S_k with the smallest lower bound, $\Phi_L(Q)$ is identified, and split along its longest side into hyper-rectangles \bar{Q}_1 , and \bar{Q}_2 (Step R2).

In Step R3, the lower bounds corresponding to \bar{Q}_1 , and \bar{Q}_2 , $\Phi_L(\bar{Q}_1)$ and $\Phi_L(\bar{Q}_2)$, are calculated. For $i = 1, 2$, if $\Phi_L(\bar{Q}_i) \leq U_k$, $\Phi_U(\bar{Q}_i)$ is calculated.

The new partition set S_{k+1} is formed from S_k by deleting \bar{Q} and adding \bar{Q}_1 and \bar{Q}_2 . The new estimates for the upper bound, U_{k+1} , and the lower bound, L_{k+1} , are then calculated. R6 is an optional pruning step. The procedure then repeats, unless it terminates under the condition $U_k - L_k < \epsilon$.

Note that necessarily, at each iteration k ,

- for $i = 1, 2$, $\Phi_L(\bar{Q}_i) \geq \Phi_L(\bar{Q}) = L_k$ in Step R3. Further, if (x^*, y^*, W^*) is the unique solution of $\min_{(x,y) \in \bar{Q}, W \in \mathcal{W}(\bar{Q})} \Lambda_L(x, y, W)$, then $\Phi_L(\bar{Q}_i) = \Phi_L(\bar{Q})$ if and only if $(x^*, y^*, W^*) \in \bar{Q}_i$.

– for $i = 1, 2$, $\text{size}\{\bar{Q}_i\} \leq \text{size}\{\bar{Q}\}$, and the “volume” of \bar{Q}_i is half that of \bar{Q} .

Further, if \bar{Q} has a unique longest side, then for $i = 1, 2$, $\text{size}\{\bar{Q}_i\} < \text{size}\{\bar{Q}\}$. Broadly speaking, it can be seen that the lower bound L_k will generally increase every iteration, and further, U_k is nonincreasing. In fact, for a fixed $\epsilon > 0$, Algorithm 3 terminates within a finite number of iterations.

THEOREM 4. *For a fixed $\epsilon > 0$, Algorithm 3 terminates within a finite number of iterations.*

Proof. The branch and bound algorithm used is standard, and the general proof may be found in, e.g., [25, Theorem IV.3], which gives convergence of the bounds U_k and L_k for the case $\epsilon = 0$. It therefore suffices to prove that the general requirements for [25, Theorem IV.3] are fulfilled. The termination of Algorithm 3 within a finite number of iterations then follows since $\epsilon > 0$.

Firstly, since any hyper-rectangle \bar{Q} has only $n_x + n_y$ sides, given any $\delta > 0$, there exists a positive integer N such that the size $\{Q^N\} < \delta$, where $\{Q^j\}$ is a sequence of hyper-rectangles successively split along their longest sides and $Q^1 := \bar{Q}$. It follows that since Φ_U and Φ_L fulfil conditions **C1** and **C2**, the bounding operation of Algorithm 3 is *consistent* [25, Definition IV.4].

Further, since the set to be split in Step R2, \bar{Q} , is selected to be such that $\Phi_L(\bar{Q}) = L_k$, the selection operation is *bound improving* [25, Definition IV.6]. The conditions for [25, Theorem IV.3] are therefore fulfilled.

An example of the performance of Algorithm 3 is given in the next section. It should also be noted that if only the BMI feasibility problem is of interest, i.e., if it is desired only to find (x, y) such that $\Lambda(x, y) < 0$, then the algorithm is terminated if either $U_k < 0$ or $L_k \geq 0$.

5.3. ACCELERATING ALGORITHM CONVERGENCE

Algorithm 3 may be modified several ways to accelerate its convergence, improve on its efficiency and to reduce its memory requirements. The pruning of S_k to remove hyper-rectangles for which the global minimum cannot occur, has already been mentioned. Pruning is not strictly necessary; however, it reduces the number of hyper-rectangles under consideration, and therefore the memory requirements of Algorithm 3.

Another refinement is possible in Step R3. If (x^*, y^*, W^*) is a solution of $\min_{(x,y) \in \bar{Q}, W \in \mathcal{W}(\bar{Q})} \Lambda_L(x, y, W)$, and $(x^*, y^*, W^*) \in \bar{Q}_i$ for some $i = 1, 2$, then $\Phi_L(\bar{Q}_i) = \Phi_L(\bar{Q})$. Hence the computation of $\Phi_L(\bar{Q}_i)$ for that i may be bypassed.

Also, it has been observed that if (x^*, y^*) is the solution to local $\min_{(x,y) \in \bar{Q}} \Lambda(x, y)$, and $(x^*, y^*) \in \bar{Q}_i$ for some $i = 1, 2$, then one may set $\Phi_U(\bar{Q}_i) = \Phi_U(\bar{Q})$ for that i , and bypass the local BMI solver for \bar{Q}_i with minimal effect on the accuracy of the upper bound. Hence the computation of $\Phi_U(\bar{Q}_i)$ for that i may be bypassed.

Finally, it is noted that the computation of $\Phi_U(\cdot)$ is potentially expensive. The authors' computational experience indicates that $\Phi_U(\cdot)$ is most efficiently computed by running Algorithm 1 first, and then using the output of Algorithm 1 as the initial point for a few iterations of Algorithm 2.

6. Discussion

6.1. FURTHER REMARKS ON ALGORITHM 3

The key to Algorithm 3 is the development of the lower bound function Φ_L , which exploits the *bilinearity* of the problem. However, if the geometry of the BMI problem is such that the lower bound of (14) is conservative, then the convergence of Algorithm 3 will be slow.

Even though convergence to a global minimum is guaranteed in Algorithm 3, it may not be *practical* to apply the algorithm to obtain complete solutions to even moderate sized BMI problems due to the fact that in the worst-case, the number of iterations required by Algorithm 3 to provide a guaranteed ϵ -global optimum will increase exponentially with respect to the dimensions of the BMI problem, i.e., n_x and n_y . See, e.g. [18].

However, it should be noted that Algorithm 3 need not be run to termination, and termination after some fixed number of iterations may be sufficient to produce significant improvement over local optimization results. Also, as previously noted, if only the *feasibility* problem is of interest, then early termination is also possible.

It may also be noted that the authors' computational experience indicates that the global minimum is often found after only a small number of iterations, and the remainder of the algorithm is devoted to tightening the lower bound, L_k .

The geometrical/variational aspects of the BMI feasibility problem are discussed in [20]. In particular, the BMI problem was shown to be equivalent to that of finding a hyperplane separating a given matrix numerical range (field of values) and the origin, subject to the constraint that the hyperplane being generated by a dyad. The connection between that viewpoint and the lower bound of Proposition 1 is obvious.

6.2. OTHER GLOBAL APPROACHES

There are several references in the mathematical programming literature to branch and bound methods for biconvex and bilinear problems, e.g. [2, 1, 25], and references therein. Also available is the Benders decomposition (see [19] and references therein) which leads to the primal-relaxed dual algorithm of [16]. The algorithm of [16] is of particular interest. However, it requires closed form formulae for the gradients of the function to be minimized, which is unavailable in the BMI context.

Another obvious approach to the BMI global optimization problem is to use multistart methods, e.g. [32]. Various “intelligent” optimization methods may also be used. However, unless the underlying structure of the BMI problem is exploited, it is doubtful whether these methods will offer any improvement over Algorithm 3. Finally, one may also attempt global optimization over y alone, i.e., on $\tilde{\Lambda}(y) := \min_x \Lambda(x, y)$; since for fixed y , $\min_x \Lambda(x, y)$ may be calculated relatively efficiently.

7. A Simple BMI Example

Consider a simple BMI problem with its its corresponding LMI:

$$\begin{aligned}
 F(x, y) &:= \begin{bmatrix} -10 & -0.5 & -2 \\ -0.5 & 4.5 & 0 \\ -2 & 0 & 0 \end{bmatrix} + y \begin{bmatrix} -1.8 & -0.1 & -0.4 \\ -0.1 & 1.2 & -1 \\ -0.4 & -1 & 0 \end{bmatrix} \\
 &+ x \begin{bmatrix} 9 & 0.5 & 0 \\ 0.5 & 0 & -3 \\ 0 & -3 & -1 \end{bmatrix} + xy \begin{bmatrix} 0 & 0 & 2 \\ 0 & -5.5 & 3 \\ 2 & 3 & 0 \end{bmatrix}, \\
 F_L(x, y, w) &:= \begin{bmatrix} -10 & -0.5 & -2 \\ -0.5 & 4.5 & 0 \\ -2 & 0 & 0 \end{bmatrix} + y \begin{bmatrix} -1.8 & -0.1 & -0.4 \\ -0.1 & 1.2 & -1 \\ -0.4 & -1 & 0 \end{bmatrix} \\
 &+ x \begin{bmatrix} 9 & 0.5 & 0 \\ 0.5 & 0 & -3 \\ 0 & -3 & -1 \end{bmatrix} + w \begin{bmatrix} 0 & 0 & 2 \\ 0 & -5.5 & 3 \\ 2 & 3 & 0 \end{bmatrix}.
 \end{aligned}$$

The objective is to find the global minimum of $\bar{\lambda}\{F(x, y)\}$ for $(x, y) \in [-0.5, 2] \times [-3, 7]$. Note that $\bar{F}(1, 0, 1) = -I$, so that $\min \bar{\lambda}\{F_L(x, y, w)\} = -1$. Also note that there are *three* local minima in the domain, i.e., 3.3886, -0.4434 and -0.9565 given by the points $(0.0049, -2.0253)$, $(0.4436, 4.0174)$ and $(1.0488, 1.4179)$ respectively.

The global optimization algorithm, Algorithm 3, requires 24 iterations to reduce the difference between the minimum upper bound and the minimum lower bound, $U_k - L_k$ to within 0.5% of U_k . The ϵ -global minimum was found to be $\hat{x} = 1.0488$ and $\hat{y} = 1.4178$, and $U_{24} = \Lambda(\hat{x}, \hat{y}) = -0.9565$. The best lower bound to $\min_{(x,y)} \Lambda(x, y)$, $L_{24} = -0.9603$. Figure 1 shows the partitions generated by Algorithm 3. The endpoints of the local minimization algorithm, Algorithm 2, for each partition are also shown. The progress of Algorithm 3 is shown in Figure 2. Note that the global minimum is found fairly early, and the remainder of the iterations are devoted to tightening the lower bounds.

8. Summary and Conclusion

It has recently been shown that an extremely wide array of robust controller design problems frequently encountered in a variety of complex engineering applications may be reduced to the problem of finding a feasible point under a Baffine Matrix Inequality (BMI) constraint.

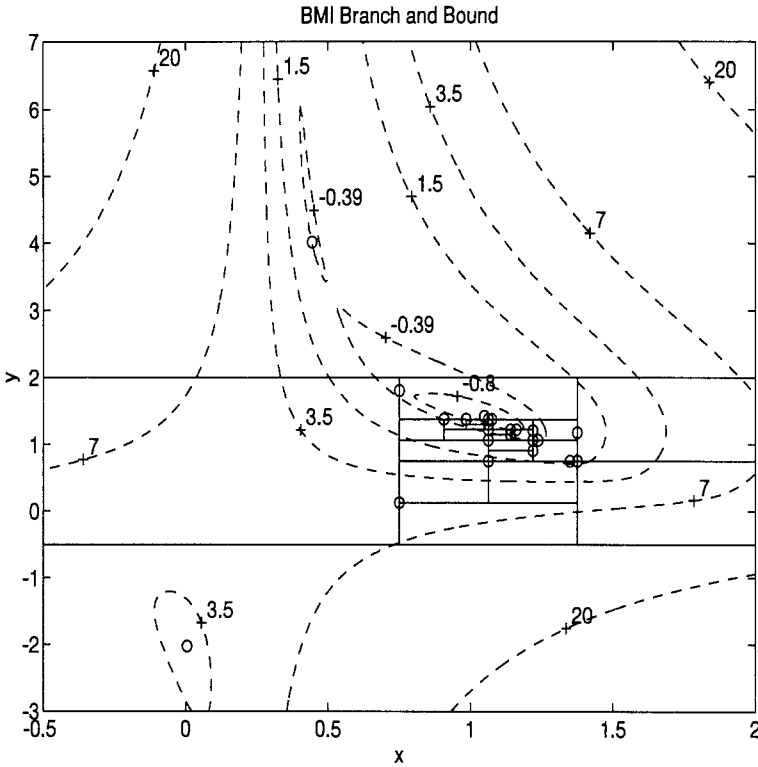


Fig. 1. Contour Plot of $\lambda\{F(x, y)\}$, with partitions generated by Algorithm 3 shown. Circles are local minimization endpoints.

The BMI feasibility problem may be viewed as a biconvex optimization problem for which a local minimum may not be a global minimum. Hence, while local optimization approaches (and even methods such as alternating LMIs) often yield sufficiently good solutions, it remains important to investigate methods for finding a solution that is guaranteed to be a global minimum within some given tolerance.

The contribution of this paper is that it provides a branch and bound global optimization algorithm for the BMI problem that is guaranteed to find a global minimum of the BMI problem to any arbitrary tolerance within a finite number of iterations. The algorithm presented exploits the bilinear structure of the problem and builds on the recent advances in solving convex LMI problems. A simple BMI example is included.

Acknowledgements

The research of K.C. Goh and M.G. Safonov is partly supported by AFOSR grant F49620-95-I-0095. The research of G.P. Papavassilopoulos is partly supported by NSF Grant CCR-9222734. The authors thank P. Gahinet, A. Nemirovskii, A.J.

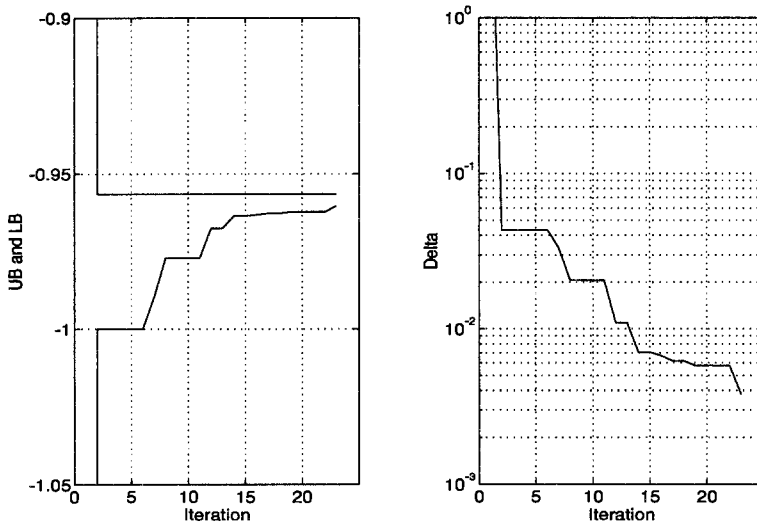


Fig. 2. Progress of the upper and lower bounds, U_k and L_k , and the logarithmic plot of the difference, $U_k - L_k$.

Laub and M. Chilali for making available their software package LMILab for the authors' research.

References

1. F.A. Al-Khayyal. Jointly constrained bilinear programs and related problems: An overview. *Computers and Mathematics with Applications*, 19(11):53–62, 1990.
2. F.A. Al-Khayyal and J.E. Falk. Jointly constrained biconvex programming. *Mathematics of Operations Research*, 8(2):273–286, May 1983.
3. F. Alizadeh. *Combinatorial Optimization with Interior Point Methods and Semidefinite Matrices*. PhD thesis, University of Minnesota, October 1991.
4. V. Balakrishnan and S. Boyd. Global optimization in control system analysis and synthesis. In C.T. Leondes, editor, *Control and Dynamic Systems*, volume 53. Academic Press, 1992.
5. 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.
6. B.R. Barmish, C.A. Floudas, C.V. Hollot, and R. Tempo. A global linear programming solution to some open robustness problems including matrix polytope stability. *University of Wisconsin-Madison Tech. Report*, ECE-94-13, 1994.
7. S.P. Boyd and L. El Ghaoui. Method of centers for minimizing generalized eigenvalues. *Linear Algebra and its Applications*, 188–189:63–111, 1993.
8. S.P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*. S.I.A.M. Publications, Philadelphia, PA, 1994.
9. J. David. *Algorithms for Analysis and Design of Robust Controllers*. PhD thesis, Katholieke Universiteit Leuven, Belgium, 1994.
10. 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.
11. P. Dorato, editor. *Robust Control*. IEEE Press, New York, 1987.

12. J.C. Doyle. Synthesis of robust controllers and filters with structured plant uncertainty. In *Proc. IEEE Conf. on Decision and Control*, San Antonio, TX, December 1983. IEEE Press.
13. L. El Ghaoui and P. Gahinet. Rank minimization under lmi constraints: A framework for output feedback problems. In *Proceedings of 1993 European Control Conference, Groningen*, pages 1176–1179, 1993.
14. M.K.H. Fan, A.L. Tits, and J.C. Doyle. Robustness in the presence of mixed parametric uncertainty and unmodelled dynamics. *IEEE Trans. on Automatic Control*, AC-36(1):25–38, January 1991.
15. A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, Inc., New York, NY, 1968.
16. 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.
17. P. Gahinet, A. Nemirovskii, A.J. Laub, and M. Chilali. *The LMI Control Toolbox, Beta Release*. Mathworks Inc., Natick, MA, November 1994.
18. M.R. Garey and D.S. Johnson. *Computers and Intractability, A Guide to the Theory of NP-Completeness*. W.H. Freeman and Co., New York, 1979.
19. A.M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.
20. K.C. Goh. *Robust Control Synthesis via Bilinear Matrix Inequalities*. PhD thesis, University of Southern California, Los Angeles, CA., May 1995.
21. K.C. Goh, M.G. Safonov, and J.H. Ly. Robust synthesis via bilinear matrix inequalities. *International Journal of Robust and Nonlinear Control*, January 1995. Submitted to the Special Issue on Linear Matrix Inequalities in Control Theory and Applications.
22. K.C. Goh, M.G. Safonov, and G.P. Papavassilopoulos. A global optimization approach for the BMI problem. In *Proc. IEEE Conf. on Decision and Control*, pages 2009–2014, Lake Buena Vista, FL., 1994.
23. K.C. Goh, L. Turan, M.G. Safonov, G.P. Papavassilopoulos, and J.H. Ly. Bilinear Matrix Inequality properties and computational methods. In *Proc. American Control Conf.*, pages 850–855, Baltimore, MD., 1994. IEEE Press.
24. R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
25. R. Horst and H. Tuy. *Global Optimization: Deterministic Approaches*. Springer-Verlag, Berlin, second, revised edition, 1993.
26. 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.
27. 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.
28. A.S. Nemirovskii and D.B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. John Wiley and Sons, Chichester, England, 1983. English translation by E.R. Dawson.
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 μ 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. C.H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall Inc., Englewood Cliffs, NJ, 1982.
32. A.H.G. Rinnooy Kan and G.T. Timmer. Stochastic global optimization methods, Part I: Clustering methods; and Part II: Multi level methods. *Mathematical Programming*, 39:27–56, 57–78, 1987.
33. M.G. Safonov. L^∞ optimal sensitivity vs. stability margin. In *Proc. IEEE Conf. on Decision and Control*, San Antonio, TX, December 1983. IEEE Press.
34. M.G. Safonov and R.Y. Chiang. Real/Complex K_m -Synthesis without curve fitting. In C.T. Leondes, editor, *Control and Dynamic Systems*. Academic Press, 1993.
35. M.G. Safonov, K.C. Goh, and J.H. Ly. Controller synthesis via Bilinear Matrix Inequalities. In *Proc. American Control Conf.*, pages 45–49, Baltimore, MD., June 29 – July 1, 1994. IEEE Press.

36. O. Toker and H. Ozbay. On the NP-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback. In *Proc. American Control Conf.*, Seattle, WA, 1995. Submitted for consideration.
37. L. Vandenberghe and S.P. Boyd. *PDP: Software for Positive Definite Programming, User's Guide, Version 0.9*, October 1994.
38. L. Vandenberghe and S.P. Boyd. Semidefinite programming. *Technical Report, I.S.L., Stanford University*, November 1994.
39. R.E. Wendell and A.P. Hurter. Minimization of a non-separable objective function subject to disjoint constraints. *Operations Research*, 24(4):643–657, July–August 1976.
40. M.H. Wright. Interior methods for constrained optimization. *Acta Numerica*, pages 341–407, 1991.