Bilinearity and Complementarity in Robust Control

Mehran Mesbahi* Michael G. Safonov† George P. Papavassilopoulos‡

Abstract

We present an overview of the key developments in the methodological, structural, and computational aspects of the Bilinear Matrix Inequality (BMI) feasibility problem. In this direction, we present the connections of the BMI with robust control theory, its geometric properties, including interpretations of the BMI as a rank constrained Linear Matrix Inequality (LMI), as an Extreme Form Problem (EFP), and as a Semi-Definite Complementarity Problem (SDCP). Computational implications and algorithms are also discussed.

1 Introduction

The simultaneous appearance of the (unknown) variables x and y in the matrix inequality:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j F_{ij} > 0, \tag{1.1}$$

for a given set of symmetric matrices $F_{ij} \in \mathbf{R}^{p\times p}$ $(i = 1, \ldots n; j = 1, \ldots m)$, not only provides an unexpectedly powerful formulation for a wide range of robust control problems, but also introduces new and elegant structural and computational questions. A possible initial attempt to rename the products $x_i y_j$ as z_{ij} in (1.1) and rewriting it in terms of z_{ij}'s as a linear matrix inequality (LMI) [6],

$$\sum_{i,j} z_{ij} F_{ij} > 0, \tag{1.2}$$

introduces yet another twist to this problem, since the unknown variables z_{ij}'s are now constrained to be related in a rather peculiar manner, for example,

$$z_{ij} z_{(i+1)(j+1)} = x_i y_i x_{i+1} y_{i+1} = z_{i(j+1)} z_{(i+1)j}.$$

Since at the present time this approach seems to present neither aesthetic insights, nor suggest a computational approach for finding the variables x and y in (1.1) (or prove the non-existence of them), we abandon our initial temptation and decide to treat the problem

*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109. The research of the author was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

†Department of EE—Systems, University of Southern California, CA 90089; research was supported in part by AFOSR Grant F49620-95-I-0095.

‡Department of EE—Systems, University of Southern California, CA 90089; research was supported in part by the National Science Foundation grant CCR-9222734.

with complete regard to its most distinguished property: *bilinearity*; we shall refer to this problem as the Bilinear Matrix Inequality (BMI).

What makes the BMI an extremely important problem in robust control, however, is far beyond an initial mathematical curiosity to make the LMI "bilinear." The BMI is introduced in response to some of the most important issues facing the field of robust control, and in particular those which *can not* be addressed in the LMI framework [20], [54].

The chapter presents various facets of the research performed by the authors and their colleagues on the BMI problem. The topics to be covered are categorized as,

1. methodological,
2. structural, and
3. computational.

Under the first category, we present the very important results pertaining to the formulation of the robust synthesis problems as a BMI. In particular, we cover reformulating the μ/k_m synthesis [13], [51], along with *specifications on the controller order and structure*, to a BMI. Our presentation in this part is based on the results reported in Safonov et al. [54], and influenced by the dissertations of Goh [19] and Ly [38].

We then proceed to present some of the structural aspects of the BMI. It turns out that the investigations into the geometry of the solution set of the BMI lead to some very interesting non-convex and convex programming problems over cones. In particular, we discuss the results pertaining to the equivalence of the BMI with examining the magnitude of the diameter of a certain convex set [55]. The results connecting the BMI to a class of cone optimization problems, namely, the Semi-Definite Complementarity Problems [28], [39], [41], [42], will also be given particular attention. The computational methods for solving the BMI are briefly reviewed in the final section.

§1.1 introduces the relevant preliminaries.

1.1 Preliminaries

1.1.1 Convex Analysis. Let $\mathcal{H}$ be a finite dimensional Hilbert space equipped with the inner product $<\cdot,\cdot>$: $\mathcal{H} \times \mathcal{H} \to \mathbf{R}$ (e.g., the n-dimensional Euclidean space or the space of $n \times n$ matrices, with the appropriate notion of an inner product defined on them). A set $\mathcal{K} \subseteq \mathcal{H}$ is a cone if for all $\alpha \geq 0$, $\alpha\mathcal{K} \subseteq \mathcal{K}$. $\mathcal{K}$ is a convex cone, if $\mathcal{K}$ is a cone and it is convex, i.e., for all $\alpha \in [0,1]$, $\alpha\mathcal{K} + (1-\alpha)\mathcal{K} \subseteq \mathcal{K}$, or equivalently, if $\mathcal{K}$ is a cone and $\mathcal{K} + \mathcal{K} \subseteq \mathcal{K}$. A convex cone $\mathcal{K}$ is called pointed if $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$, and solid if it has a non-empty interior. An extreme form (or an extreme ray) of a convex cone $\mathcal{K}$ is a subset $E = \{\alpha x : \alpha \geq 0\}$ of $\mathcal{K}$, such that if $x = \alpha y + (1-\alpha)z$, for $0 < \alpha < 1$, and $y, z \in \mathcal{K}$, one can conclude that $y, z \in E$ [23], [24]. The set of extreme forms (rays) of the cone $\mathcal{K}$ is denoted by $\dot{\mathcal{K}}$. The dual cone of a set $S \subseteq \mathcal{H}$, denoted by S^*, is defined to be,

$$S^* = \{y \in \mathcal{H} : <x, y> \geq 0; \ \ \forall x \in S\}.$$

If S is a pointed closed convex cone, then the interior of its dual cone, $\operatorname{int} S^*$, is given by,

$$\operatorname{int} S^* = \{y \in \mathcal{H} : <x, y> > 0; \ \ \forall x \in S, x \neq 0\}.$$

A closed convex cone in a finite dimensional Hilbert space is pointed if and only if its dual is solid [4]. In particular, $\mathbf{SR}_+^{p \times p}$ is a pointed, closed, and self-dual cone — the cone which induces the ordering used in the LMI and the BMI formulations. It can easily be shown that S^* is always a convex set, and that if $S_1 \subseteq S_2$, then $S_2^* \subseteq S_1^*$. In addition, $S = (S^*)^*$ if and only if S is a closed convex cone. Given a pointed closed convex cone $\mathcal{K} \subseteq \mathcal{H}$, a linear map $M : \mathcal{H} \to \mathcal{H}$ is called $\mathcal{K}$-positive if for all $0 \neq X \in \mathcal{K}$, $M(X) \in \operatorname{int} \mathcal{K}^*$. Furthermore, a linear map $M : \mathcal{H} \to \mathcal{H}$ is called $\mathcal{K}$-copositive if $< X, M(X) > \geq 0$, for all $X \in \mathcal{K}$ [4], [22], [28].

We are now ready to formulate the cone problems that are considered in §3.2. The Cone-LP is formulated as follows: Given a cone $\mathcal{K} \subseteq \mathcal{H}$, a linear map $M : \mathcal{H} \to \mathcal{H}$, and the elements Q and C in $\mathcal{H}$, find $Z \in \mathcal{H}$ (if it exists) as a solution to:

$$\min < C, Z > \tag{1.3}$$
$$Z \in \mathcal{K}, \tag{1.4}$$
$$Q + M(Z) \in \mathcal{K}^*. \tag{1.5}$$

Similarly, the Cone-LCP is formulated as follows: Given a cone $\mathcal{K} \subseteq \mathcal{H}$, a linear map $M : \mathcal{H} \to \mathcal{H}$, and $Q \in \mathcal{H}$, find $Z \in \mathcal{H}$ (if it exists) such that:

$$Z \in \mathcal{K} \tag{1.6}$$
$$Q + M(Z) \in \mathcal{K}^*, \tag{1.7}$$
$$< Z, Q + M(Z) >= 0. \tag{1.8}$$

The above instances of the Cone-LP and the Cone-LCP shall be referred to as the Cone-LP$_{\mathcal{K}}(C, Q, M)$ and Cone-LCP$_{\mathcal{K}}(Q, M)$. When $\mathcal{K}$ is the nonnegative orthant in the n-dimensional Euclidean space, the Cone-LP$_{\mathcal{K}}(C, Q, M)$ (1.3)–(1.5) and the Cone-LCP$_{\mathcal{K}}(Q, M)$ (1.6)–(1.8), are equivalent to the familiar Linear Programming and the Linear Complementarity problems [8], [9].

A problem which serves as a bridge between the BMI and the Cone-LP/LCPs is the Extreme Form Problem (EFP): Given a pointed closed convex cone $\mathcal{K} \subseteq \mathcal{H}$, a linear map $M : \mathcal{H} \to \mathcal{H}$, find $X \in \mathcal{H}$ (if it exists), such that,

$$X \in \mathcal{K} \tag{1.9}$$
$$M(X) \in \text{ int } \mathcal{K}^*, \tag{1.10}$$
$$X \text{ is an extreme form of } \mathcal{K}. \tag{1.11}$$

The above instance of the EFP is referred to as the EFP$_{\mathcal{K}}(M)$. When $\mathcal{K}$ is the nonnegative orthant in the n-dimensional Euclidean space, the EFP is a trivial problem. It should be noted that the solution set of an EFP is generally nonconvex; given the two extreme forms of $\mathcal{K}$ that solve the EFP$_{\mathcal{K}}(M)$, a strict convex combination of them is not an extreme form of $\mathcal{K}$. It is also important to note that the EFP requires $M(X)$ to lie in the *interior* of the dual cone. This is in light of the fact that for certain important classes of linear maps M, including the map that is encountered in the context of the BMIs, $M(X)$ is known to lie in, but possibly on the *boundary* of, the dual cone, for all the extreme forms X of the cone $\mathcal{K}$. As it is shown in §3.2, the EFP formulation, in-spite of its nice geometrical interpretation, includes as its special case, the BMI problem. Since the extreme forms of the matrix classes which are considered in the chapter can be characterized by their rank, the EFP formulation of the BMI also translates directly to a rank minimization problem over a convex set of matrices [39] (refer to §3.2).

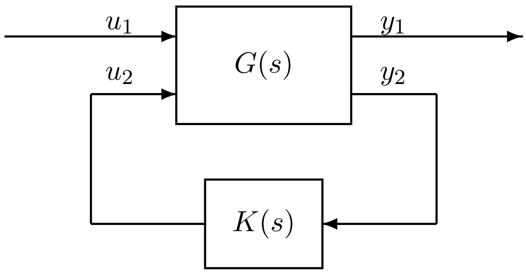

FIG. 2. $T(s) := \mathcal{F}_L\{G(s), K(s)\}$

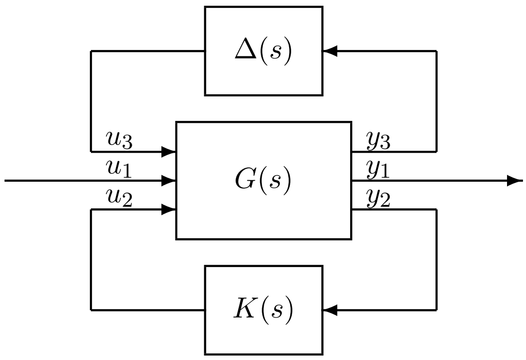

FIG. 3. $T_\Delta := \mathcal{F}_U\{\Delta(s),\ \mathcal{F}_L\{G(s), K(s)\}\}$

1.1.2 Control Theory. We use r_i and s_i to denote the dimensions of ith input and the ith output of the finite dimensional linear time invariant (LTI) plant $G(s)$, respectively; the order of $G(s)$ is denoted by n; similarly q and q_M designate the order of the controller and the multiplier to be synthesized.

Consider the LTI plant $G(s)$ partitioned as,

$$\begin{aligned} G(s) &= \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \\ &:= \mathcal{F}_U\left\{\frac{1}{s}I, \mathbf{S}_G\right\}, \end{aligned} \tag{1.12}$$

where

$$\mathbf{S}_G := \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \tag{1.13}$$

and $A \in \mathbf{R}^{n\times n}$, $D_{11} \in \mathbf{R}^{s_1\times r_1}$, $D_{22} \in \mathbf{R}^{s_2\times r_2}$ and $G_{ij}(s) := D_{ij} + C_i(Is - A)^{-1}B_j$.

With respect to this particular partition of the transfer matrix G, the feedback connection shown in Figure 1 has the transfer function

$$T(s) \;:=\; \mathcal{F}_U\left\{\frac{1}{s}I, \left[\begin{array}{c|c} A_t & B_t \\ \hline C_t & D_t \end{array}\right]\right\} \;:=\; \mathcal{F}_L\{G(s), K(s)\}, \tag{1.14}$$

where $K(s)$ denotes a controller of order q given by

$$K(s) \;:=\; \mathcal{F}_U\left\{\frac{1}{s}I, \mathbf{S}_K\right\} \tag{1.15}$$

and

$$\mathbf{S}_K \ := \ \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]. \tag{1.16}$$

The inclusion of the uncertainty in this framework is now really an extra bonus; see Figure 3. The resulting transfer function from u_1 to y_1 is simply,

$$T_\Delta := \ \mathcal{F}_U\{\Delta(s), T(s)\}$$

provided that all the appropriate inverses exist.

Motivated by the way that uncertainty manifests itself in the generalized plants (those which include the nominal plant, the sensors, and the actuators), one is lead to consider uncertainties of the form:

$$\Delta = \mathrm{diag}[\delta_1 I_{n_1}, \ldots, \delta_L I_{n_L}, \Delta_{L+1}, \ldots, \Delta_{L+F}] \tag{1.17}$$

for some prescribed positive integers L, F, and $n_1, \ldots, n_L$, such that $\|\Delta\|_\infty \leq 1$. The δ_i's are real or complex valued, and in general correspond to parametric uncertainties; Δ_i's on the other hand are time invariant operators used to account for such things as unmodeled dynamics. Given that Δ has arisen from our modeling technique and/or our inadequate knowledge about the nature of things, and that G is a physical reality, K is our only hope to make the system in Figure 2 operate in a way that is desirable to us: the controller K is chosen to provide internal stability and (external) performance for all possible perturbations Δ; performance in robust control setting is often taken as the ability of the system to reject the disturbances lumped in the term u_1, necessitating not only $\|T\|_\infty < 1$ (as the result of the small gain condition, the so-called bounded realness of the operator T [56], [61]) but also requiring $\|T_\Delta\|_\infty$ to be as small as possible.

Let us shift our attention back to the unperturbed configuration of Figure 2. It is well-known that the bilinear sector transform $\widetilde{T}(s) = \ \mathrm{sect}\{T(s)\} := (I - T(s))(I + T(s))^{-1}$ maps bounded real systems $\|T(s)\|_\infty < 1$ into positive real systems $\mathbf{Herm}(\widetilde{T}(j\omega)) > 0$ and vice versa. A routine calculation reveals that the state space matrices of $\widetilde{T}(s) = \ \mathrm{sect}\{T(s)\}$ and $T(s)$ are related by

$$\mathbf{S}_{\widetilde{T}} := \left[\begin{array}{c|c} \tilde{A}_t & \tilde{B}_t \\ \hline \tilde{C}_t & \tilde{D}_t \end{array}\right] = \ \mathcal{F}_L\{S, \mathbf{S}_T\}\,, \tag{1.18}$$

where

$$S := \left[\begin{array}{cc|cc} 0 & 0 & I & 0 \\ 0 & I & 0 & -\sqrt{2}I \\ \hline I & 0 & 0 & 0 \\ 0 & \sqrt{2}I & 0 & -I \end{array}\right]; \tag{1.19}$$

the corresponding bilinearly transformed open-loop plant has the state-space system matrix

$$\mathbf{S}_{\tilde{G}} := \left[\begin{array}{c|cc} \tilde{A} & \tilde{B}_1 & \tilde{B}_2 \\ \hline \tilde{C}_1 & \tilde{D}_{11} & \tilde{D}_{12} \\ \tilde{C}_2 & \tilde{D}_{21} & \tilde{D}_{22} \end{array}\right] = \ \mathcal{F}\{S, \mathbf{S}_G\}\,. \tag{1.20}$$

It is known that we may linearly parameterize the set of realizable $\mathbf{S}_{\widetilde{T}}$ matrices as follows [45], [46]:

$$\mathbf{S}_{\widetilde{T}} = R + U\mathbf{S}_Q V, \tag{1.21}$$

where $\mathbf{S}_Q$ is the state space system matrix of

$$Q(s) \quad := \quad K(s)\left(I - \tilde{D}_{22}K(s)\right)^{-1}, \tag{1.22}$$

and

$$\left[\begin{array}{c|c} R & U \\ \hline V & \tilde{D}_{22} \end{array}\right] := \left[\begin{array}{cc|c} \tilde{A} & \tilde{B}_1 & \tilde{B}_2 \\ \tilde{C}_1 & \tilde{D}_{11} & \tilde{D}_{12} \\ \hline \tilde{C}_2 & \tilde{D}_{21} & \tilde{D}_{22} \end{array}\right]. \tag{1.23}$$

Thus we have all the necessary machinery to translate between positive real and bounded real conditions. The principle tool in translating frequency domain results related to positive real conditions in terms of the corresponding matrix inequalities is the following important lemma.

LEMMA 1.1 (GENERALIZED STRONG POSITIVE REAL LMI: [1], [54]). *Let $G(s) = C(Is - A)^{-1}B + D$ be a minimal state-space realization. Then for some $\epsilon > 0$*

$$\mathbf{Herm}(G(j\omega)) > \epsilon I \quad \forall \omega,$$

if and only if there exists $P = P^T$ such that

$$\mathbf{Herm}(\begin{bmatrix} -P & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}) > 0. \tag{1.24}$$

Moreover, $G(s)$ is stable if and only if $P > 0$.

2 Methodological Aspects of the BMI

Mathematics often takes a precise problem formulation as its starting point, theoretical research in engineering more often than not takes it close to its ending. It is common to see statements like "since this problem is equivalent to solving a system of linear equations, we consider the problem solved," in the engineering literature. But when should we consider an engineering problem *solved*? Certainly, one could respond to this question by saying that an engineering problem is solved when the corresponding mathematical problem is solved; however, this correspondence is never one to one and, moreover, there are still many questions as to when a mathematical problem, specially in optimization, is declared to be solved. Let us then take the following as our starting point:

> An engineering problem is solved when a computationally reasonable method (on a reasonable model of computation) exists for at least one of its equivalent mathematical formulations.

Given that the statement above is a reasonable approximation to our actual motivation in control research, we strive for a problem formulation which leads to an efficient computational method for the solution of a control problem. On the other hand, when formulating an engineering problem, we are inclined to formulate it in such way that the mathematical formulation is directly linked to our engineering intuition and judgment. The frequency domain techniques in control system design is a prime example of such preferences. To make matters even more interesting, we often search for a formulation which not only captures our engineering considerations, but that which has a nice mathematical representation. Alas often these considerations cannot be satisfied at the same time, although exceptions exist.[1]

[1] Linear Programming, the Simplex Algorithm, and the corresponding economic interpretation, are beautiful examples of such an interaction — even though finding a polynomial time algorithm for Linear

The engineering problem considered in this chapter is the synthesis of controllers for systems whose models are not precisely known. We adopt the framework of the robust control theory, and in particular that of μ/k_m analysis and synthesis. Roughly speaking, with reference to Figure 1, μ/k_m synthesis is the problem of finding the controller K such that $\mathcal{F}_U\{\Delta, T\}$ is stable for all admissible Δ's. Since one would like to exploit the structure of the uncertainty Δ (of the form 1.17) in order to establish stability, we are led beyond requiring $\|T\|_\infty < 1$ to such conditions as solving [62],

$$\inf_{D(s),K(s)} \|D(s)\ \mathcal{F}_L\{G(s), K(s)\}\ D(s)^{-1}\|_\infty, \tag{2.25}$$

where D is a stable, minimum phase scaling matrix such that $D(s)\Delta(s) = \Delta(s)D(s)$; the infimum is guaranteed to provide a lower bound (upper bound) for the multivariable stability margin k_M [49], [50] (resp. μ [11]), where for an asymptotically stable transfer matrix $T(s)$,

$$k_M := \inf_{w\in\mathbf{R}} \inf_{\Delta} \{k : \det(I - kT\Delta) = 0)\} \tag{2.26}$$

and

$$\mu := 1/k_M. \tag{2.27}$$

The D-K iteration is a solution method for the μ/k_m synthesis which proceeds by alternating between solving an $\mathcal{H}^\infty$optimization problem by fixing $D(s)$ in (2.25), followed by a convex optimization with K fixed [2], [7], [51]. An improvement in the conservativeness of the D-K iteration is known as the D, G-K iteration [15], [47]. Both design techniques can be enhanced by considering the μ/k_m synthesis in the positive real framework, resulting in what is known as the M-K iteration [53]. The M-K iteration approach has a very nice property of being able to bypass the curve fitting step in the D-K iteration; this latter approach in fact forms the basis for the proof of Theorem 2.1 (below).

2.1 Limitations of the LMI Approach

It remains true however that much difficulty remains in the robust synthesis of practical, non-conservative controllers. At least three very important classes of robust control design issues have not been found to be readily transformable into the LMI framework, which has offered a very promising direction to study many robust analysis problems. These are (1) μ/k_m-synthesis via *dynamical* scalings/multipliers, (2) *fixed-order* control synthesis and (3) *decentralized* controller design (i.e., synthesis of controllers with "block diagonal" or other specified structure).

- Consider the robust control problem of μ/k_m-synthesis. Current μ/k_m techniques [2],[7], remain inherently conservative, i.e. sub-optimal, because the D-K iteration approach of alternately synthesizing first a controller $K(s)$ and then diagonal scalings $D(s)$ is in no way guaranteed to achieve a globally *or, it turns out, even a locally* optimal solution. The D-K iteration can, in theory, get stuck at points which are local minima with respect to D alone and with respect to K alone, but are not a true local minima with respect to *joint* variations in D and K.

Programming which corresponds to a certain pivoting strategy of the Simplex method is still an open problem.

- One of the main objections to the modern control synthesis theories, such as LQG, $\mathcal{H}^\infty$ and μ/k_m, is that the resultant controllers are typically of relatively high order — at least as high as the original plant and often much higher in the very usual case where the plant must be augmented by dynamical scalings, multipliers or weights in order to achieve the desired performance.

- Control system designs for very large or complex systems must often be implemented in a *decentralized* fashion; that is, local loops are decoupled and closed separately with little or no direct communication among local controllers. The synthesis of optimally robust decentralized systems has obvious benefits. However, even the synthesis of optimal decentralized $\mathcal{H}^\infty$ and LQG control systems has remained beyond the scope of the existing theories.

Neither the LMI framework nor other existing theories has yet proved to be sufficiently flexible to handle problems in the foregoing classes. The purpose of this section is to demonstrate that the BMI framework is sufficiently flexible to simultaneously accommodate all three of the foregoing types of control design specifications, in addition to handling all those which the LMI handles. In particular we provide a proof of the following theorem:

THEOREM 2.1 ([54]). *The μ/k_m synthesis problem can be formulated as the following matrix inequality: Given real matrices $R_{aug}, U_{aug}, V_{aug}$, and $R_{\tilde{A}_t}, U_{\tilde{A}_t}, V_{\tilde{A}_t}$, find matrices $Z, \mathbf{S}_Q$ such that*

$$\mathbf{Herm}\left(Z \left[\begin{array}{c|c} \begin{bmatrix} I_{n+q} & 0 \\ 0 & R_{\tilde{A}_t} + U_{\tilde{A}_t}\mathbf{S}_Q V_{\tilde{A}_t} \end{bmatrix} & 0 \\ \hline 0 & R_{aug} + U_{aug}\mathbf{S}_Q V_{aug} \end{array}\right]\right) > 0 \qquad (2.28)$$

where

$$Z := \left[\begin{array}{c|c} \begin{bmatrix} X & 0 \\ 0 & -X \end{bmatrix} & 0 \\ \hline 0 & \begin{bmatrix} -P & 0 & 0 \\ 0 & W & 0 \\ 0 & 0 & W \end{bmatrix} \end{array}\right] \qquad (2.29)$$

and $P = P^T$ and $X = X^T \in \mathbf{R}^{(n+q)\times(n+q)}$, and W is constrained to have the structure consistent with the uncertainty structure of the underlying μ/k_m problem (1.17).

The rest of this section is devoted to the proof of Theorem 2.1 and one of its important consequences in the design of the decentralized controllers.

2.2 Proof of Theorem 2.1

We first make a note of the following: any constraint of the form $x_i = 1$, $y_i = 1$ for an instance of the BMI can be transformed to an instance with no such constraints [54].

Proof. [54] In the μ/k_m-synthesis procedure outlined in [53], the μ/k_m problem is shown to be solved if there exists a suitably structured block-diagonal rational multiplier matrix $M(s)$ such that

$$\mathbf{Herm}(M(s)\ \operatorname{sect}\{T(s)\}) > 0,\ \forall s = j\omega. \qquad (2.30)$$

No constraint is placed on the stability of $M(s)$, but it is required to be uniformly bounded on the $j\omega$-axis and to satisfy the generalized strong positive real condition, i.e., for some $\epsilon > 0$,

$$\mathbf{Herm}(M(s)) > \epsilon I,\ \forall s = j\omega. \qquad (2.31)$$

It is supposed in [36] (without loss of generality if $N = \infty$) that $M(s)$ is a weighted sum of certain suitably structured block-diagonal transfer function matrices $M_i(s)$,

$$M(s) := \sum_{i=1}^{N} W_i\, M_i(s) \equiv W\tilde{M}(s), \tag{2.32}$$

where

$$\tilde{M}(s) := \begin{bmatrix} M_1(s) \\ \vdots \\ M_N(s) \end{bmatrix} \tag{2.33}$$

and $W := [W_1, \ldots, W_N] \in \mathbf{R}^{s_1 \times q s_1}$. In most cases $W_i \in \mathbf{R}$ but, more generally as when there are repeated uncertainty blocks, W_i may be block diagonal matrices of a certain specified form. The specific details of the structure constraints on W_i and $M_i(s)$ required for the various types of real and/or complex uncertainty block structures are described in [36] and [52].

One may form the "augmented" closed-loop system

$$T_{\text{aug}}(s) := \begin{bmatrix} M(s) & \text{sect}\{T(s)\} & 0 \\ & 0 & M(s) \end{bmatrix}; \tag{2.34}$$

then, $T_{\text{aug}}(s)$ has a state-space realization of the form (1.21)

$$\mathbf{S}_{T_{\text{aug}}} := \left[\begin{array}{c|c} A_{\text{aug}} & B_{\text{aug}} \\ \hline C_{\text{aug}} & D_{\text{aug}} \end{array}\right] = R_{\text{aug}} + U_{\text{aug}}\mathbf{S}_Q V_{\text{aug}}. \tag{2.35}$$

In view of the form of (2.34), A_{aug} naturally assumes the form

$$A_{\text{aug}} = \begin{bmatrix} \tilde{A}_t & 0 & 0 \\ * & A_{\tilde{M}} & 0 \\ 0 & 0 & A_{\tilde{M}} \end{bmatrix}, \tag{2.36}$$

where $A_{\tilde{M}}$ denotes the A-matrix of $\tilde{M}(s)$ and $\tilde{A}_t \in \mathbf{R}^{(n+q)\times(n+q)}$ is the A-matrix of the bilinear sector transformed closed-loop plant $\text{sect}\{T(s)\}$. Under (2.30)–(2.31), the stability of $\text{sect}\{T(s)\}$ is equivalent to the stability of the original untransformed closed-loop plant $T(s)$ [53]. In view of (2.36) and (2.35), one has

$$\tilde{A}_t = R_{\tilde{A}_t} + U_{\tilde{A}_t}\mathbf{S}_Q V_{\tilde{A}_t} := \mathbf{E}_{n+q}(R_{\text{aug}} + U_{\text{aug}}\mathbf{S}_Q V_{\text{aug}})\mathbf{E}_{n+q}^T, \tag{2.37}$$

where,

$$\mathbf{E}_{n+q} := [I_{n+q}, 0, \ldots, 0]\,. \tag{2.38}$$

Theorem 2.1 is now proved by applying Lemma 1.1 for checking the positive realness of T_{aug}.

We note that the matrix P in the above theorem is the solution to the LMI which results from the application of Lemma 1.1. Since we do not require A_{aug} to be stable, no definiteness condition is imposed on P. Instead, stability of the closed-loop $\tilde{A}_t$ (and hence $T(s)$) is ensured by testing existence of $X = X^T > 0$ such that $\mathbf{Herm}(-X\tilde{A}_t) > 0$ (e.g., [44][page 63]); this is the role of the matrix X in the BMI (2.28).

Clearly (2.28) is a BMI feasibility problem. It is jointly linear in the parameters of the matrices W, X, P. It is affine in the controller parameter matrix $\mathbf{S}_Q$. Some special cases of this problem have been found by Packard et al. [45], [46] to be reducible to LMI's via the Parrott theorem. These include full-state feedback $\mathcal{H}^\infty$ and full-order $\mathcal{H}^\infty$ with constant diagonal scalings ($M(s) =$ "Constant Matrix"), as well as certain simultaneous stabilization and related gain scheduling problems. But optimal solution to even the fixed order ($q < n$) $\mathcal{H}^\infty$ problem (i.e., $M(s) = I$) has remained previously elusive despite some determined efforts. The foregoing BMI formulation (2.28) provides a simple formulation of these and the related synthesis problems.

Interestingly, the BMI formulation is flexible enough to accommodate constraints on controller structure as well. To simplify matters, we make the following assumption (refer to (1.13)):

$$D_{22} = 0.$$

Hence by (1.22),

$$K(s) = Q(s). \tag{2.39}$$

Note that no significant loss of generality results from this assumption since it can always be made to hold via a singular perturbation of the plant (given that our constraints do not require infinite bandwidth controller).

With $D_{22} = 0$, it is clear that the controller $K(s)$ inherits the same block structure as $Q(s)$. In particular, if the state-space matrices A_Q, B_Q, C_Q, D_Q are constrained to have a block-diagonal structure, then $K(s)$ will be block-diagonal too; i.e., $K(s)$ will be a "decentralized" controller.

2.3 Why is the BMI Formulation Important?

We conclude this section with a recapitulation of the reasons why the BMI formulation is important:

- The main attraction of the BMI formulation is its simplicity and generality. It allows the controller and multiplier/scaling optimization in μ/k_m-synthesis to be formulated as a single finite dimensional optimization over the controller parameters $\mathbf{S}_Q$ and the multiplier/scaling and Lyapunov parameters W, X, P. Consequently, application of nonlinear programming techniques to the BMI at least assures convergence to a joint local optimum in D and K — the D-K iteration of μ/k_m-synthesis cannot make this claim. The BMI formulation also eliminates the curve fitting step of the traditional D-K iteration approaches to μ/k_m.

- A broad spectrum of robust control synthesis problems can be formulated within the BMI framework. These include order-constrained controller μ/k_m-synthesis with specifications requiring such additional properties as decentralized control. Gain scheduling and simultaneous μ/k_m-synthesis for several plants also fall in this framework.

We note however that the BMI formulation is not without its drawbacks. One major concern is that biconvex optimization problems are in general difficult to solve. Moreover, much available structure is hidden in the BMI formulation. For example, we know that the BMI for the full order output feedback $\mathcal{H}^\infty$ synthesis may be reduced to an LMI. It seems unlikely however that order constrained or decentralized control problems will admit an LMI embedding. It will be interesting to see just how broad a class of BMI's will admit an embedding within the LMI framework.

3 Structural Aspects of the BMI

Optimization often provides a very convenient framework for constructing computational procedures for a wide range of problems. Applying a standard trick that changes a feasibility problem to an optimization one results in rewriting the BMI as,

$$\inf \alpha \tag{3.40}$$

subject to

$$\alpha I + \sum_i \sum_j x_i y_j F_{ij} \geq 0; \tag{3.41}$$

the BMI has a feasible point if and only if the value of the infimum is negative. At this point we can theoretically apply some general purpose *global* optimization technique to (3.40)–(3.41). Our intuition however suggests that the geometrical properties of the BMI should be useful in the construction of the computational procedures. Our goal is to gain as much insight into the geometrical and analytic properties of the BMI to the extend that our choice of the algorithm for its solution is judicious and transcends beyond a rather blind application of some global optimization technique.

There are at least three issues which have to be addressed in connection with the BMI and the global optimization methods:

1. What are the geometric interpretations of the BMI?
2. What are the specific properties of the global optimization problem which arises from the BMI, and whether these properties can be used to devise more efficient algorithms for the BMI?
3. Which instances of the BMI can be solved efficiently? Moreover, are there instances for which certain "structural" properties can be established?

All of the above issues can be addressed by studying the BMI *on its own*. We believe however that many important structural and computational issues of the BMI can be studied by establishing a connection between the BMI and problems which are more well-understood in optimization theory. This section is devoted to such investigations. In the first part, we present the result showing that the BMI can be formulated as a convex *maximization* problem. The second part establishes the connection between the BMI and two optimization problems over cones, namely the Extreme Form (EFP) and the Semi-Definite Complementarity Problems (SDCP).

The two approaches discussed in this section, beside providing very nice geometrical insights into the structure of the solution set of the BMI, also suggest computational procedures for its solution, a subject which we shall comment on in §4.

3.1 Concave Programming Approach

It is shown that a BMI has a non-empty solution set if and only if the diameter of a certain convex set is greater than two. The convex set in question is simply the intersection of ellipsoids centered at the origin. In this avenue we end up addressing the first two issues raised above about the relationship between the BMI and the global optimization methods. Specifically, we prove the following theorem.

Theorem 3.1 ([55]). *The BMI (1.1) is feasible if and only if the diameter of a suitably constructed convex set* $\mathbf{C}$ *is strictly greater than* 2*; i.e.,*

$$\max_{w_1, w_2 \in \mathbf{C}} \|w_1 - w_2\| > 2. \tag{3.42}$$

The implications of Theorem 3.1 are twofold. First, it opens up a wide range of possibilities for the application of the concave minimization algorithms for the solution of the BMI. At the same time, the concave minimization result indicates, in a rather transparent way, why it is not a good idea to spend research time looking for a *polynomial time* BMI solvers.[2] This is due to the fact that concave minimization belongs to a class of computational problems for which the existence of a polynomial time algorithm is highly unlikely (concave minimization is NP-hard [59]). The NP-hardness of the BMI was also explicitly proved in [58].

The rest of the section is devoted to the proof of the Theorem 3.1 and some of its implications [55].

We begin by noting that the BMI problem, that of finding the vectors x and y such that $\sum_{ij} x_i y_j F_{ij} > 0$ (1.1) is equivalent to finding the corresponding vectors such that $\sum_{ij} x_i y_j F_{ij} < 0$, which can be written as,

$$\min_{\|z\|=1, z\in\mathbf{R}^p} x^T G(z) y < 0 \tag{3.43}$$

where

$$[G(z)]_{i,j} = z^T F_{ij} z \in \mathbf{R}^{n\times m}. \tag{3.44}$$

Let ρ be a real positive number such that

$$\rho > \max_{\|z\|=1} \ \sigma_{\max}(G(z)) \tag{3.45}$$

and let $\mathbf{C} \subseteq \mathbf{R}^{n+m}$ be the convex set

$$\mathbf{C} := \{w : w^T Q(z) w \leq 1, \ z \in \mathbf{R}^p, \|z\| = 1\} \tag{3.46}$$

where

$$Q(z) = \begin{bmatrix} I & \frac{1}{\rho}G(z) \\ \frac{1}{\rho}G^T(z) & I \end{bmatrix}. \tag{3.47}$$

Notice that by (3.45), one has that the matrix $Q(z) > 0$, for all $\|z\| = 1$. It follows that the set $\mathbf{C}$ is the intersection of ellipsoids in $\mathbf{R}^{n+m}$.

Suppose that $\mathbf{C}$ has diameter strictly greater than 2. $\mathbf{C}$ is the intersection of an infinite number of ellipsoids and thus its maximum diameter is achieved at some $\bar{w}_1 = (\bar{x}, \bar{y})$ on the boundary of $\mathbf{C}$ and, by symmetry, also at $\bar{w}_2 = -\bar{w}_1$. It thus holds that,

$$\bar{w}_1^T Q(z) \bar{w}_1 \leq 1, \quad \forall z, \ \|z\| = 1 \tag{3.48}$$

where

$$\bar{w}_1 := \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \tag{3.49}$$

and

$$\|\bar{w}_1\| = \|\bar{x}\|^2 + \|\bar{y}\|^2 > 1. \tag{3.50}$$

Inequality (3.48) is equivalent to

$$\|\bar{x}\|^2 + \|\bar{y}\|^2 + \frac{2}{\rho}\bar{x}^T G(z)\bar{y} \leq 1, \ \forall z, \ \|z\| = 1 \tag{3.51}$$

[2] It is generally accepted in computer science that polynomial running time of an algorithm is equivalent to efficiency.

from which it follows that

$$\frac{2}{\rho}\bar{x}^T G(z)\bar{y} \leq 1 - (\|\bar{x}\|^2 + \|\bar{y}\|^2) < 0. \tag{3.52}$$

Thus

$$\bar{x}^T G(z)\bar{y} < 0, \text{ uniformly in } z, \ \|z\| = 1. \tag{3.53}$$

and consequently $-\bar{x}, \bar{y}$ satisfy (1.1).

On the other hand, suppose $(-\bar{x}, \bar{y})$ satisfy (1.1). Without loss of generality we may assume $\|\bar{x}\|^2 + \|\bar{y}\|^2 = 1$. It then holds that

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}^T Q \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \|\bar{x}\|^2 + \|\bar{y}\|^2 + \frac{2}{\rho}\bar{x}^T G(z)\bar{y} \tag{3.54}$$

$$= 1 + \frac{2}{\rho}\bar{x}^T G(z)\bar{y} < 1, \tag{3.55}$$

uniformly in $z, \ \|z\| = 1$. Thus the radius of the set $\mathbf{C}$ is strictly greater than $(\|\bar{x}\|^2 + \|\bar{y}\|^2) = 1$. Since $\mathbf{C}$ is hermitian and centered about the origin, the diameter is precisely twice the radius. Hence, the diameter of $\mathbf{C}$ is strictly greater than 2.

The ρ's used in (3.47) may all be the same or they may be chosen to depend on z. All we need is to guarantee that

$$Q(z) = \begin{bmatrix} I & \frac{1}{\rho}G(z) \\ \frac{1}{\rho}G^T(z) & I \end{bmatrix} > 0 \tag{3.56}$$

and thus we can take ρ to be z-dependent, e.g.,

$$\rho = \rho(z) > \sigma_{\max}(G(z)). \tag{3.57}$$

Finding a different ρ for each z is a laborious task. A single constant ρ that will satisfy (3.45) for all z can easily be computed via the matrix inequality

$$\sigma_{\max}(\bar{Q}) \leq \sigma_{\max}(\text{abs}(Q(z))) \leq \sigma_{\max}(Q) \tag{3.58}$$

where ij-th entry of the $n \times m$ matrix $\bar{Q}$ is given by

$$[\bar{Q}]_{ij} = \sigma_{\max}(F_{ij}). \tag{3.59}$$

In view of Theorem 3.1, the following statement is obvious: Consider the optimization problem,

$$\max_{x,y} \ \|x\|^2 + \|y\|^2 \tag{3.60}$$

subject to

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} I & \frac{1}{\rho}G(z) \\ \frac{1}{\rho}G^T(z) & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq 1 \quad \forall z \in \mathbf{R}^p, \ \|z\| = 1. \tag{3.61}$$

Then, there exists a feasible pair (x, y) (1.1) if and only if the global optimum of (3.60)–(3.61) is strictly greater than 1; note that since all points of the form $(x, y) = (\hat{x}, 0)$ or $(x, y) = (0, \hat{y})$ with $\|\hat{x}\| = \|\hat{y}\| = 1$ satisfy (3.61), the optimum of (3.60) is always greater than or equal to 1. Actually, instead of solving the problem (3.60) for its global optimum, all we need is a point (x, y) where $J(x, y) > 1$. Note that (3.60)–(3.61) is a nonlinear programming problem where a convex function is to be maximized subject to an infinite number of quadratic constraints (parameterized by z) all of which are ellipsoids centered at the origin.

We note that the BMI can be formulated as follows: Find an $n \times m$ matrix N of rank 1 (i.e., $N = xy^T$) such that for all z, with $\|z\| = 1$,

$$\mathbf{Tr}(G(z)N) = x^T G(z) y < 0; \tag{3.62}$$

that is, in the Hilbert space of $n \times m$ real matrices find a hyperplane that strictly separates the origin from the set $\mathbf{W} = \{G(z) \mid \|z\| = 1\}$ and in addition it should hold that the matrix N that defines the perpendicular to this hyperplane is of rank one. In the absence of the restriction rank$(N) = 1$, the problem may be interpreted as the LMI,

$$\epsilon_* = \min_{\epsilon, N} \epsilon \tag{3.63}$$

subject to

$$\mathbf{Tr}(NG(z)) \leq \epsilon, \qquad \forall z,\ \|z\| = 1 \tag{3.64}$$

$$|[N]_{ij}| \leq 1, \qquad \forall i = 1, \ldots, n; j = 1, \ldots, m; \tag{3.65}$$

a solution exists if and only if the minimal cost $\epsilon_* < 0$. If the rank of N is restricted to be less than $\min\{m, n\}$, the LMI formulation fails. We notice that the (3.63)–(3.65) formulation of the BMI comes remarkably close to an LMI. At the same time however, this problem formulation signifies the role that rank restricted LMIs play in the control synthesis problems in a rather transparent way.

We shift our attention now to the cone programming formulations of the BMI.

3.2 Cone Programming Approach

We present an approach for solving the BMI based on its connections with certain problems defined over matrix cones. These problems are, among others, the cone generalization of the linear programming (LP) and the linear complementarity problem (LCP) (referred to as the Cone-LP and the Cone-LCP, respectively). Specifically, we show that solving a given BMI is equivalent to examining the solution set of a suitably constructed Cone-LP or Cone-LCP. In this direction we end up addressing all of the issues raised above pertaining to the BMIs and global optimization techniques.

The main results proved in this section are the following.

THEOREM 3.2 ([39]). *The BMI is an instance of the EFP.*

THEOREM 3.3 ([39]). *The BMI has a solution if and only if a suitably constructed SDCP with a copositive linear map has a rank one solution.*

The emphasis of this section is on the matrix theoretic aspects of the cone problems which the BMI leads to. In this direction, we spent quite an effort to understand the geometry of a certain classes of matrix cones, and subsequently to establish their relevant properties, helpful in our understanding of the BMI.

The rest of this section is devoted to the proofs of the above two theorems.

3.2.1 Few Initial Steps. Let us rewrite the BMI (1.1) as:

$$\sum_i x_i \sum_j y_j F_{ij} = \sum_i x_i F_i^y > 0 \tag{3.66}$$

where,

$$F_i^y = \sum_j y_j F_{ij} \in \mathbf{SR}^{p\times p}.$$

As it becomes apparent by the subsequent developments, it is convenient to assume that $m = p$, and when necessary, that y_j's $(1 \le j \le m)$, are nonnegative. The first assumption is made to avoid defining inner products between matrix classes of different dimensions. The second assumption is made in §3.2.2 to facilitate the formulation of the problems in terms of dual cones. In §3.2.3 shall drop the non-negativity assumption on the vector y for the Cone-LCP formulation. These assumptions are warranted for the following reason. First, note that if we define $F_j^x = \sum_i x_i F_{ij} \in \mathbf{SR}^{p\times p}$ $(1 \le j \le m)$, then $\sum_i x_i F_i^y = \sum_j y_j F_j^x$. But the last sum is a *linear* inequality in F_j^x's. Thus, as it is customary in the Linear Programming, one can assume that $m \le p$ and that y_j's are positive (by an appropriate augmentation). Now we would only need to define $F_{ij} \equiv 0$ $(1 \le i \le n; m < j \le p)$, for the assumption $m = p$ to be justified.

Recall that the Gordan's theorem of alternative [8], [57], relates the solvability of the following two systems of linear inequalities: Given $A \in \mathbf{R}^{m\times n}$, the system $Ax > 0$ has a solution if and only if the system $y^T A = 0, y \ge 0, y \ne 0$, has no solution. This theorem can be generalized for the linear inequalities over matrix cones as follows:

PROPOSITION 3.1. *Given the symmetric matrices $A_i's \in \mathbf{SR}^{p\times p}$ $(1 \le i \le n)$, the system $\sum_{i=1}^n x_i A_i > 0$, has a solution if and only if the system,*

$$\mathbf{Tr}(A_i Z) = 0,\ (1 \le i \le n), \quad Z \ge 0, \quad Z \ne 0,$$

has no solution.

¿From the Gordan's theorem of alternative over the cone of symmetric positive semi-definite matrices, one concludes that the BMI (1.1) does not have a solution if and only if,

$$(\forall y \ge 0)\ (\exists Z \ge 0, Z \ne 0): \quad \mathbf{Tr}(F_i^y Z) = 0. \tag{3.67}$$

Therefore, the BMI (1.1) has a solution if and only if,

$$(\exists y \ge 0)\ (\forall Z \ge 0, Z \ne 0): \quad \sum_i \mathbf{Tr}(F_i^y Z)^2 > 0. \tag{3.68}$$

Now let,

$$F_i = [\overbrace{\mathbf{vec}\ F_{i1}\, 0 \ldots 0}^{p}, \overbrace{0\, \mathbf{vec}\ F_{i2} \ldots 0}^{p}, \ldots, \overbrace{0\, 0 \ldots \mathbf{vec}\ F_{ip}}^{p}] \in \mathbf{R}^{p^2\times p^2} \tag{3.69}$$

and Y=diag$(y) \in \mathbf{SR}^{p\times p}$. Since $(F_i^y)^T = F_i^y = \sum_j y_j F_{ij}$ (recalling that F_{ij}'s are symmetric matrices),

$$\mathbf{vec}\ (F_i^y)^T = F_i\, \mathbf{vec}\ Y; \tag{3.70}$$

thereby,

$$\begin{aligned} \mathbf{Tr}(F_i^y Z) &= (\mathbf{vec}\ (F_i^y)^T)^T (\mathbf{vec}\ Z) \\ &= (F_i \mathbf{vec}\ Y)^T (\mathbf{vec}\ Z) \\ &= (\mathbf{vec}\ Y)^T F_i^T (\mathbf{vec}\ Z). \end{aligned} \tag{3.71}$$

Combining (3.68) and (3.71) we conclude that (1.1) has a solution if and only if there exists $Y \ge 0$, Y=diag(y), for some $y \ge 0$, such that for all $Z \ge 0, Z \ne 0$,

$$(\mathbf{vec}\ Z)^T \{\sum_i F_i (\mathbf{vec}\ Y)(\mathbf{vec}\ Y)^T F_i^T\} (\mathbf{vec}\ Z) > 0. \tag{3.72}$$

Let $X = (\mathbf{vec}\, Y)(\mathbf{vec}\, Y)^T$, $Y \in \mathbf{R}^{p\times p}$, and,

$$M(X) = \sum_i F_i X F_i^T. \tag{3.73}$$

We observe the following:

1. According to (3.69), for all $p \times p$ skew-symmetric matrices Z, $F_i\,\mathbf{vec}\; Z = 0$ $(i = 1, \ldots, n)$. Consequently, if there exists a matrix $Y \in \mathbf{R}^{p\times p}$ such that (3.72) holds, then one can assume that Y is *symmetric*, since *the skew-symmetric part of* Y *does not contribute to the left hand side of the inequality (3.72)*: Let $Y = Y_1 + Y_2$, with Y_1 and Y_2 being the symmetric and skew-symmetric part of Y, respectively. Then for $1 \le i \le n$, $F_i\,\mathbf{vec}\; Y = F_i(\mathbf{vec}\; Y_1 + \mathbf{vec}\; Y_2) = F_i\,\mathbf{vec}\; Y_1$.

2. According to (3.70), for all matrices $Y \in \mathbf{R}^{p\times p}$, $F_i\,\mathbf{vec}\; Y = \mathbf{vec}\; W_i$, for some $W_i \in \mathbf{SR}^{p\times p}$. Therefore if $X = (\mathbf{vec}\, Y)(\mathbf{vec}\, Y)^T$, then $M(X)$ can be represented by,

$$M(X) = \sum_i (\mathbf{vec}\; W_i)(\mathbf{vec}\; W_i)^T. \tag{3.74}$$

REMARK 3.1. *Suppose that the vector y is not required to be nonnegative in the above analysis. It is clear that the above steps are still valid with the obvious modifications, and that the end result would read as follows: The BMI has a solution if and only if there exists a diagonal matrix Y, such that for all $Z \ge 0, Z \ne 0$, the inequality (3.72) holds. We shall use this observation later in §3.2.2.*

The inequality (3.72) can be interpreted as requiring $M(X)$ to belong to a certain matrix class. The matrices in this class are symmetric (given that X is symmetric) and have quadratic forms which are positive over the **vec** form of the non-zero matrices in $\mathbf{SR}_+^{p\times p}$. This observation justifies the introduction of the following matrix classes.

Denote by $\mathcal{PSD}$, the class of $p^2 \times p^2$ symmetric positive semi-definite matrices, i.e., matrices for the which the quadratic form is nonnegative over the **vec** form of the $p \times p$ matrices ($\mathbf{R}^{p\times p}$),

$$\mathcal{PSD} = \{A \in \mathbf{SR}^{p^2\times p^2} : (\mathbf{vec}\; Z)^T A\, (\mathbf{vec}\; Z) \ge 0; Z \in \mathbf{R}^{p\times p}\}. \tag{3.75}$$

Let $\mathcal{PSD}_0$ denote the subset of $p^2 \times p^2$ symmetric matrices with quadratic forms nonnegative over the **vec** form of the symmetric $p \times p$ matrices ($\mathbf{SR}^{p\times p}$), and with the **vec** form of the skew-symmetric matrices ($\mathbf{SKR}^{p\times p}$) in their null space, i.e.,

$$\begin{aligned} \mathcal{PSD}_0 = \{A \in \mathbf{SR}^{p^2\times p^2} \;\; : \;\; & (\mathbf{vec}\; Z)^T A\, (\mathbf{vec}\; Z) \ge 0; Z \in \mathbf{SR}^{p\times p}; \\ & A(\mathbf{vec}\; W) = 0;\ \text{for all } W \in \mathbf{SKR}^{p\times p}\}. \end{aligned} \tag{3.76}$$

Clearly both $\mathcal{PSD}$ (3.75) and $\mathcal{PSD}_0$ (3.76) are closed convex cones. Moreover, it can be shown that certain essential features of the $\mathcal{PSD}$ cone can be generalized for the class of $\mathcal{PSD}_0$ matrices, including rank one decomposition property, self-duality, and the unity rank of the extreme forms [24], [39].

Let $\mathcal{C}$ denote the class of symmetric PSD-copositive matrices,

$$\mathcal{C} = \{A \in \mathbf{SR}^{p^2\times p^2} : (\mathbf{vec}\; Z)^T A\, (\mathbf{vec}\; Z) \ge 0;\ Z \ge 0\}. \tag{3.77}$$

A particular subset of $\mathcal{C}$ which be useful in our cone formulations is a subset of matrices in $\mathcal{C}$ which have the $p \times p$ skew-symmetric matrices in their null space, i.e.,

$$\mathcal{C}_0 := \{A \in \mathcal{C} : A\,\mathbf{vec}\; W = 0;\ W \in \mathbf{SKR}^{p\times p}\}. \tag{3.78}$$

Finally, let $\mathcal{B}$ denote the class of symmetric PSD-completely positive matrices,

$$\mathcal{B} = \{A \in \mathbf{SR}^{p^2 \times p^2} : A = \sum_{i=1}^{t} (\mathbf{vec}\ Z_i)(\mathbf{vec}\ Z_i)^T ; Z_i \geq 0, t \geq 1\}. \tag{3.79}$$

One can establish the following results.

LEMMA 3.1 ([39]). *The matrix classes $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{C}_0$ are closed convex cones in $\mathbf{SR}^{p^2\times p^2}$. Moreover, $\mathcal{B}^* = \mathcal{C}$, $\mathcal{C}^* = \mathcal{B}$, $\mathcal{C}$ is solid, and $\mathcal{B}$ is pointed.*

LEMMA 3.2 ([39]). *The extreme forms of $\mathcal{B}$ are matrices* $(\mathbf{vec}\ Z)(\mathbf{vec}\ Z)^T$, $Z \geq 0$.
Note that $\mathcal{B} \subseteq \mathcal{PSD}_0 \subseteq \mathcal{PSD} \subseteq \mathcal{C}$, and $\mathcal{PSD}_0 \subseteq \mathcal{C}_0 \subseteq \mathcal{C}$.

It is now observed that Equation (3.72) states whether a *nonlinear* combination of matrices $F_i's \in \mathbf{R}^{p^2\times p^2}$ $(1 \leq i \leq n)$, belongs to the *interior* of the cone of PSD-copositive matrices $\mathcal{C}$. In fact, due the particular form of the linear map M (3.74), $M(X)$ is required to be in $\mathcal{C}_1 := \mathcal{C}_0 \cap \operatorname{int}\ \mathcal{C}$.

3.2.2 EFP Formulation of the BMI. We present the proof of Theorem 3.2 in this section, i.e., the reformulation of the BMI as the $\mathrm{EFP}_{\mathcal{B}}(M)$, where M is defined by (3.73), and $\mathcal{B}$ is the cone of PSD-completely positive matrices.

Proof. [39] If the BMI has a solution X, then there exists $Y = \operatorname{diag}(y)$, $y \geq 0$, $X = (\mathbf{vec}\ Y)(\mathbf{vec}\ Y)^T$, such that $M(X) \in \operatorname{int}\ \mathcal{C}$, and hence, the $\mathrm{EFP}_{\mathcal{B}}(M)$ has a solution.

Conversely, suppose that the $\mathrm{EFP}_{\mathcal{B}}(M)$ has a solution X. Then there exists $V \geq 0$, such that $X = (\mathbf{vec} V)(\mathbf{vec}\ V)^T$ and $M(X) \in \operatorname{int}\ \mathcal{C}$, i.e., for all $Z \geq 0, Z \neq 0$,

$$(\mathbf{vec}\ Z)^T \{\sum_i F_i (\mathbf{vec}\ V)(\mathbf{vec} V)^T F_i^T\} (\mathbf{vec}\ Z) > 0.$$

Let $V = T^T Y T$ be such that Y is diagonal, T is nonsingular, and $T^T = T^{-1}$. Then $\mathbf{vec}\ Y = (T \otimes T)\mathbf{vec}\ V$.

We observe that,

$$\begin{aligned}
&(\mathbf{vec}\ Z)^T \{\sum_i F_i (\mathbf{vec}\ Y)(\mathbf{vec} Y)^T F_i^T\} (\mathbf{vec}\ Z) \\
&\quad = (\mathbf{vec}\ Y)^T \{\sum_i F_i (\mathbf{vec}\ Z)(\mathbf{vec}\ Z)^T F_i^T\} (\mathbf{vec}\ Y) \\
&\quad = (\mathbf{vec}\ V)^T (T\otimes T)^T \{\sum_i F_i (\mathbf{vec}\ Z)(\mathbf{vec}\ Z)^T F_i^T\} (T\otimes T)(\mathbf{vec}\ V) \\
&\quad = ((T\otimes T)(\mathbf{vec}\ Z))^T \{\sum_i F_i (\mathbf{vec}\ V)(\mathbf{vec}\ V)^T F_i^T\} (T\otimes T)(\mathbf{vec}\ Z) > 0
\end{aligned}$$

The last inequality follows from the fact that if $\mathbf{vec}\ W = (T\otimes T)\mathbf{vec}\ Z$, $Z \geq 0$, and T is nonsingular, then $W \geq 0$, since $W = TZT^T$.

Therefore the diagonal matrix Y, is a solution to the $\mathrm{EFP}_{\mathcal{B}}(M)$. Now define the vector $y \in \mathbf{R}_+^p$ by $y_i = Y_{ii}$ $(1 \leq i \leq p)$.

REMARK 3.2. *Suppose that the nonnegativity assumption on the vector y is dropped. It then follows that the BMI is also equivalent to finding an extreme form X of the $\mathcal{PSD}_0$ cone (3.76) such that $M(X) \in \operatorname{int} \mathcal{B}^* \equiv \operatorname{int} \mathcal{C}$, and in fact, $M(X) \in \mathcal{C}_1 := \mathcal{C}_0 \cap \operatorname{int}\ \mathcal{C}$. This observation shall be used in Section 3.2.3.*

The implication of Theorem 3.2 is that the BMI is equivalent to checking whether the image of an extreme form of the matrix cone $\mathcal{B}$ under the linear map M (which is constructed from the original data of the BMI), is in the interior of the dual cone $\mathcal{B}^*$. This equivalence thus provides a rather simple geometric interpretation of the BMI feasibility problem.

An immediate consequence of the EFP formulation is the following characterization of the BMI instances for which a solution exists.

COROLLARY 3.1 ([39]). *The BMI has a solution if the linear map M (3.73) is $\mathcal{B}$-positive.*

Proof. [39] If M is $\mathcal{B}$-positive, then every (non-zero) extreme form of $\mathcal{B}$ is mapped to the interior of $\mathcal{B}^* \equiv \mathcal{C}$. Therefore the $\mathrm{EFP}_{\mathcal{B}}(M)$, and consequently the BMI, have a solution.

3.2.3 Cone-LCP Formulation of the BMI. In this section we explore the idea of establishing a connection between the BMI and the class of linear complementarity problems over matrix cones (Cone-LCP). The EFP formulation of the BMI discussed in §3.2.2 is our main tool in this direction. In particular we shall provide a proof of Theorem 3.3.

Our motivation for the Cone-LCP approach is twofold. First, the complementarity formulation enables one to address the two important structural issues (2) and (3) mentioned in the opening paragraph of §3. The basic idea is to use the rich theory that has been developed for the linear complementarity problems over the last few decades to examine the properties of the BMI. Additionally, we shall show that the Cone-LCP that arises from the BMI can be formulated on the $\mathcal{PSD}$ cone, rather than the less (computationally) understood matrix cone $\mathcal{B}$ of §3.2.2. The complementarity problem over the $\mathcal{PSD}$ cone also has the advantage of being readily amenable to an interior point approach [30]. However, one has to note that a "general" Cone-LCP is a difficult computational problem, as is the LCP itself. One of the main advantages of the Cone-LCP formulation however, is the ability of recognizing efficiently solvable instances.

In this section we shall assume for the moment that the Cone-LCPs of the form Cone-$\mathrm{LCP}_{\mathcal{PSD}}(Q, M)$, with M being a $\mathcal{PSD}$-copositive (see §1.1), can be solved with a "reasonable" efficiency. This assumption is guided by the fact that a usual copositive LCP is a more tractable problem than a general LCP.

The starting point for the Cone-LCP approach is the remark made after the proof of Theorem 3.2: the BMI has a solution if and only if the image of an extreme form of the matrix cone $\mathcal{PSD}_0$ under the linear map M, is in the interior of $\mathcal{C}$, or in fact in $\mathcal{C}_1$. As the next proposition states, the cone $\mathcal{PSD}$ can substitute the cone $\mathcal{PSD}_0$ in the above statement.

PROPOSITION 3.2 ([39]). *There exists an extreme form of the matrix cone $\mathcal{PSD}$, X, such that $M(X) \in \operatorname{int} \mathcal{C}$, if and only if there exists an extreme form of the matrix cone $\mathcal{PSD}_0$, Y, such that $M(Y) \in \operatorname{int} \mathcal{C}$.*

Let us denote by $\bar{p} = p(p+1)/2$ the dimension of the space of symmetric $p \times p$ matrices. Before stating the main result of this section we make the following observation.

LEMMA 3.3 ([39]). *Let Y be an extreme form of the cone $\mathcal{PSD}_0$. Then there exists a symmetric $W \in \mathcal{PSD}_0$, such that $\mathbf{Tr}(YW) = 0$ and rank $(W) = \bar{p} - 1$.*

Consider the Cone-$\mathrm{LCP}_{\mathcal{PSD}}(Q, M)$ and let M be defined by the equation (3.73): Find

$X \in \mathbf{SR}^{p^2 \times p^2}$ (if it exists) such that:

$$X \in \mathcal{PSD} \tag{3.80}$$
$$Q + M(X) \in \mathcal{PSD}^* \equiv \mathcal{PSD} \tag{3.81}$$
$$\mathbf{Tr}(X(Q + M(X))) = 0 \tag{3.82}$$

THEOREM 3.4 ([39]). *The BMI has a solution if and only if there exists a matrix $Q \in \text{int}\,(-\mathcal{C})$ (or $Q \in -\mathcal{C}_1$), such that the Cone-LCP$_{\mathcal{PSD}}(Q, M)$ has a rank one solution.*

Proof. [39] Suppose that the BMI has a solution X^*, that is, $X^* = (\mathbf{vec}\ Y)(\mathbf{vec}\ Y)^T$, $Y \in \mathbf{SR}^{p\times p}$, $M(X^*) \in$ int $\mathcal{C}$, and in fact $M(X^*) \in \mathcal{C}_1$. By Lemma 3.3, there exists $W \in \mathcal{PSD}_0$, rank $W = \bar{p} - 1$, such that $\mathbf{Tr}(WX^*) = 0$. Without loss of generality, assume that $\|W\| = 1$.

Let $Q_\alpha = \alpha W - M(X^*)$. Note that since $M(X)$ is symmetric (for $X \in \mathcal{PSD}_0$), Q_α is also symmetric. Moreover, $Q_\alpha(\mathbf{vec}\ Z) = 0$, for all $Z \in \mathbf{SKR}^{p\times p}$, since both $M(X)$ and W are in the $\mathcal{PSD}_0$ (see Equation (3.74)). It suffices to show that there exists an $\alpha > 0$, such that $Q_\alpha \in \text{int}\,(-\mathcal{C})$ (or $Q \in -\mathcal{C}_1$). Since $M(X^*) \in \text{int}\,\mathcal{C}$ and $\mathcal{B}$ is closed, there exists $\beta > 0$, such that

$$\inf_{U\in\mathcal{B};\|U\|=1} \mathbf{Tr}(UM(X^*)) \geq \beta > 0.$$

Hence, for all $U \in \mathcal{B}, \|U\| = 1$,

$$\mathbf{Tr}(UQ_\alpha) = \mathbf{Tr}(U(\alpha W - M(X^*))) = \alpha(\mathbf{Tr}(UW)) - \mathbf{Tr}(UM(X^*)) \leq \alpha \mathbf{Tr}(UW) - \beta \leq \alpha - \beta.$$

Therefore choosing $\bar{\alpha} < \beta$, we see that for all $U \in \mathcal{B}, \|U\| = 1, \mathbf{Tr}(UQ_{\bar{\alpha}}) < 0$. Hence $Q_{\bar{\alpha}} \in$ int $(-\mathcal{C})$ (in fact $Q_{\bar{\alpha}} \in -\mathcal{C}_1$).

Moreover,

$$0 = \bar{\alpha}(\mathbf{Tr}(X^*W)) = \mathbf{Tr}(X^*(Q_{\bar{\alpha}} + M(X^*))).$$

By construction, $X^* \in \mathcal{PSD}$, rank $X^* = 1$, and $(Q_{\bar{\alpha}} + M(X^*)) \in \mathcal{PSD}_0 \subseteq \mathcal{PSD}$.

On the other hand suppose that there exists a rank one matrix X^* in the solution set of the Cone-LCP (Q, M), with $Q \in \text{int}\,(-\mathcal{C})$. Then there exists $Z^* \in \mathbf{R}^{p\times p}$ such that,

$$X^* = (\mathbf{vec}\ Z^*)(\mathbf{vec}\ Z^*)^T \in \mathcal{PSD}$$

Since $Q \in$ int $(-\mathcal{C})$ and $Q + M(X^*) \in \mathcal{PSD}$, using the inclusion $\mathcal{B} \subseteq \mathcal{PSD}$, one has the following:

$$\begin{aligned} \forall A \in \mathcal{B}\ (A \neq 0):\ \mathbf{Tr}(AM(X^*)) &= \mathbf{Tr}(A(Q + M(X^*) - Q)) \\ &= \mathbf{Tr}(A(Q + M(X^*))) - \mathbf{Tr}(AQ) > 0 \end{aligned}$$

The last inequality follows from the fact that, for all $A \in \mathcal{B}$ $(A \neq 0)$, $\mathbf{Tr}(AQ) < 0$. Consequently, $M(X^*) \in$ int $\mathcal{C}$.

In view of Proposition 3.2, there exists a rank one matrix,

$$Y^* = (\mathbf{vec}\ W^*)(\mathbf{vec}\ W^*)^T \in \mathcal{PSD}_0, \quad W \in \mathbf{SR}^{p\times p}$$

such that $M(Y^*) \in \text{int } \mathcal{C}$. Therefore the BMI has a solution.

The above proof can be modified in an obvious way to conclude that it is only sufficient to take $Q \in -\mathcal{C}_1$.

We shall refer to the special case of the linear complementarity problem over the positive semi-definite cone (3.80)–(3.82), as the Semi-Definite Complementarity Problem (SDCP). An immediate consequence of the above theorem is that if a matrix $Q \in$ int$(-\mathcal{C})$ cannot be found for which the corresponding SDCP has a solution, then the BMI does not have a solution.

Corollary 3.2 ([39]). *The BMI does not have a solution if the SDCP (3.80)–(3.82) is not solvable for any $Q \in$ int $(-\mathcal{C})$ (or in fact $Q \in -\mathcal{C}_1$).*

It is noteworthy that the linear map M in the SDCP formulation, which arises in the context of the BMI, is itself copositive with respect to the matrix cone $\mathcal{PSD}$. Proposition 3.3 ([39]). *The linear map M defined by (3.73) is $\mathcal{PSD}$-copositive. Consequently if we define $M^*(X) = \sum_{i=1}^n F_i^T X F_i$, and the implication:*

$$X \in \mathcal{PSD},\ \mathbf{Tr}(XM(X)) = 0 \Rightarrow M(X) + M^*(X) = 0 \tag{3.83}$$

holds, then for all $Q \in -\mathcal{C}_1$, the Cone-LCP$_{\mathcal{PSD}}(Q, M)$ is solvable if it is feasible.

Another Cone-LCP formulation of the BMI, in addition to the one mentioned above, is to incorporate the problem of finding the matrix Q in Theorem 3.4, in setting up the corresponding Cone-LCP. For this purpose it is convenient to associate to the matrix cones $\mathcal{PSD}$, $\mathcal{B}$, and $\mathcal{C}$ (subsets of $\mathbf{SR}^{p^2\times p^2}$), the cones $\overline{\mathcal{PSD}}$, $\overline{\mathcal{B}}$, and $\overline{\mathcal{C}}$ (subsets of $\mathbf{R}^{p^4}$), which are obtained by applying the **vec** operator to these matrix cones, i.e.,

$$\overline{\mathcal{PSD}} = \{x \in \mathbf{R}^{p^4} : x = \mathbf{vec}\ A, A \in \mathcal{PSD}\}$$

and,

$$\overline{\mathcal{B}} = \{x \in \mathbf{R}^{p^4} : x = \mathbf{vec}\ A, A \in \mathcal{B}\}; \quad \overline{\mathcal{C}} = \{x \in \mathbf{R}^{p^4} : x = \mathbf{vec}\ A, A \in \mathcal{C}\}.$$

It is easy to verify that $\overline{\mathcal{PSD}}$, $\overline{\mathcal{B}}$, and $\overline{\mathcal{C}}$ are closed convex cones in $\mathbf{R}^{p^4}$.

Recall that for all $A, B \in \mathbf{SR}^{p^2\times p^2}$,

$$\mathbf{Tr}(AB) \geq 0 \Longleftrightarrow (\mathbf{vec}\ A)^T(\mathbf{vec}\ B) \geq 0.$$

Therefore, in view of the relation $\mathcal{PSD} = \mathcal{PSD}^*$, the only matrices in $\overline{\mathcal{PSD}}^*$ that are the **vec** form of a symmetric matrix are those in $\overline{\mathcal{PSD}}$.

Let $F = \sum_{i=1}^n F_i \otimes F_i \in \mathbf{R}^{p^4\times p^4}$. For the linear map M defined by (3.73) and using the property of the Kronecker products, $\mathbf{vec}\ M(X) = F\,\mathbf{vec}\ X$. Combining the above ideas with the result of Theorem 3.4, one readily obtains the following corollary.

Corollary 3.3 ([39]). *Let,*

$$\widetilde{M} = \begin{pmatrix} 0 & 0 \\ -I_{p^4\times p^4} & F \end{pmatrix}$$

and ,

$$Z = \begin{pmatrix} \mathbf{vec}\ -Q \\ \mathbf{vec}\ X \end{pmatrix} \in \overline{\mathcal{C}} \times \overline{\mathcal{PSD}}.$$

Then the BMI has a solution if and only if the homogeneous Cone-LCP$_{\overline{\mathcal{C}}\times\overline{\mathcal{PSD}}}(0,\widetilde{M})$ *has a solution of the form,*

$$\widetilde{Z} = \begin{pmatrix} \mathbf{vec}\ -\widetilde{Q} \\ \mathbf{vec}\ \widetilde{X} \end{pmatrix}$$

where $-\widetilde{Q} \in \operatorname{int}\ \mathcal{C}$*, and* $\widetilde{X}$ *has rank one.*

We note that if $Q \in -\mathcal{C}$ and $X \in \mathcal{PSD}$, then $Q + M(X)$ is automatically symmetric and therefore if $\mathbf{vec}\ (Q + M(X^*)) \in \overline{\mathcal{PSD}}^*$, then $Q + M(X^*) \in \mathcal{PSD}$.

The above corollary reduces the BMI feasibility problem to the problem of examining the solution set of a certain Cone-LCP. This can be a "tractable" problem if the solution set is finite, or if the linear map $\widetilde{M}$ enjoys certain "additional" properties. Since there are various results in the complementarity theory which pertain to the cardinality of the solution set of a Cone-LCP [28], classification of the efficiently solvable instances of the BMI can be based on these results as well.

4 Computational Aspects of the BMI

This part of the chapter is devoted to a brief overview of the computational methods for solving the BMI, those which are *directly* related with the presentation of §2 and §3. We shall first provide few words on the approaches which are *not* touched upon in this section.

Motivated by observing that the function

$$F(x, y) = \sum_{ij} x_i y_j F_{ij} \tag{4.84}$$

is convex in x for a fixed y and vice versa, Goh et al. [21] proposed a global optimization algorithm based on the branch and bound strategy for solving the BMI; in this approach, the bilinearity of the function (4.84) is successively employed in the bounding part of each iteration. The dissertation of Liu [34] discusses many interesting aspects of the *parallel* implementation of BMI solvers which are based on the branch and bound strategy (see also [35]). In [3] a global optimization technique based on the generalized Benders Decomposition which is used in bilinear and biconvex programming [16], [17], [60] is proposed. Finally we should mention the alternating LMI method [29], but this later class of algorithms is not guaranteed to find a feasible point of the BMI (1.1), even if one exists.

Back to the methods which are linked directly to the aspects of the BMI investigated in this chapter, we present a brief overview of each.

Starting with the optimization problem (3.40)–(3.41) obtained trivially from the BMI feasibility problem, one can proceed to devise a computational method based on the following strategy. Initially pick an arbitrary x_0 and y_0 and then choose $\alpha_0 > 0$ such that (3.41) is satisfied (thus we obtain a feasible point for the optimization problem); proceed by trying to reduce α_k at each step without leaving the feasible region. In the spirit of the interior point methods for solving the LMIs and Semi-Definite Programming problems, Goh et al. [18] proposed using the logarithmic barrier functional to ensure that the iterations stay inside the feasible region:

A Barrier Method for the BMI:

1) Fix $\epsilon > 0$, $\delta > 0$, $\mu_0 > 0$, and $\theta > 1$.

2) Choose some (x_0, y_0) and α_0 such that $(\alpha_0 - \delta)I + F(x_0, y_0) \geq 0$. Let $k = 0$.

3) Until $\alpha_k - \alpha_{k+1} < \epsilon$:

   3a) $(\alpha_{k+1}, x_{k+1}, y_{k+1}) = \arg\min_{\alpha,x,y}(\mu_k\alpha - \log\det(\alpha I - F(x,y))$.

   3b) $\mu_k = \theta\mu_k$.

   3c) Let $k = k + 1$. Go to 3.

□

The local minimization step (3a) is initiated from (α_k, x_k, y_k). The convergence of the above algorithm is contingent upon a wise choice of the parameter θ. Note that as $\mu_k \to \infty$, the first term of the objective functional in (3a) approaches that of minimizing the parameter α, whereas its second term ensures that this minimization is performed without leaving the feasible region of (3.40)–(3.41).

Under some mild conditions, the barrier method described above is guaranteed to converge to a local minimum of (3.40)–(3.41). The choice of the parameter θ is guided by the methods for solving LMIs and it is closely related to the self-concordant barrier parameter for the cone of positive semi-definite matrices [43].

We observe that our choice of the initial points x_0 and y_0 can be influenced by the results of the more conservative approaches, such as the alternating LMI approach or the D-K iteration. In any event the algorithm above, which is based on the BMI formulation of the μ/k_m synthesis problem along with a methodology borrowed directly from convex programming, is guaranteed to provide improvements over that of the existing methods [19].

Barrier methods are not the only variant of the interior point methods which can be used for obtaining local solutions to BMIs, although they are probably among the best well-understood. Other ipm solution methods for locally solving the BMI is presented in the dissertation of Goh [19], and in particular one which corresponds to the method of centers in the spirit of [5] and [27]. Computational examples are also provided in [19].

Motivated by the reformulation of the BMI as a concave minimization problem presented in §3.1, Safonov and Papavassilopoulos [55] proposed the following conceptual algorithm for finding a feasible point of (1.1). The approach is based on the optimization problem (3.60)–(3.61) which is obtained directly from Theorem 3.1. In order to solve (1.1), we can solve a sequence of problems of the type (3.60)–(3.61), each one having a finite number of inequalities. At the beginning of Step k, assume that the points $z^{(1)}, \ldots, z^{(k-1)}$ have been generated from an arbitrary initial guess $z^{(1)}$ with $\|z^{(1)}\| = 1$. Then, the k-th step of the algorithm is: Solve for a globally maximizing pair (x, y) in

$$J_*^{(k)} = \max J := \|x\|^2 + \|y\|^2 \tag{4.85}$$

subject to

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} I & \frac{1}{\rho^{(i)}}G(z^{(i)}) \\ \frac{1}{\rho^{(i)}}G^T(z^{(i)}) & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq 1,\ i = 1, \ldots, (k-1). \tag{4.86}$$

If the global minimum $J_*^{(k)} = 1$, stop; the problem (3.60) is infeasible. If $J_*^{(k)} > 1$, let the solution be $(x^{(k)}, y^{(k)})$ and solve

$$\max_{\|z\|=1} z^T \bar{H}(x^{(k)}, y^{(k)})z \tag{4.87}$$

where

$$\bar{H}(x^{(k)}, y^{(k)}) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i^{(k)} y_j^{(k)} H_{ij}. \tag{4.88}$$

Note that the maximal value in (4.87) is the maximal eigenvalue $\lambda_{\max}(\bar{H})$. Take $z^{(k)} \in \mathbf{R}^k$ to be a maximizing z in (4.87), i.e., $z^{(k)}$ is any unit norm eigenvector of the matrix (4.88) associated with $\lambda_{\max}(\bar{H})$. If $\lambda_{\max}(\bar{H}) < 0$, we stop and the pair $(x^{(k)}, y^{(k)})$ provide a solution of BMI (1.1); otherwise we choose $\rho^{(k)} > \lambda_{\max}(\bar{H})$ and go to Step $k+1$.

It can be shown that this process will stop in a finite number of steps if (1.1) has a solution; otherwise (1.1) is infeasible.

Note that solving (4.85) for the global maximum may be quite a time consuming problem, although there exist several algorithm for solving nonconvex maximization problems of this type [26].

Finally we observe that it is not necessary to solve (4.85) for the global maximum, but we can stop as long as a pair $(x^{(k)}, y^{(k)})$ with $\|x^{(k)}\|^2 + \|y^{(k)}\|^2 > 1$ has been generated. This may be detrimental to the speed of convergence of the algorithm but avoids spending a lot of time in finding the global maximum of (4.85). If one chooses to do this, it might be advisable now and then to solve (4.85) globally.

Lastly we mentioned the solution method based on the SDCP approach, which is based on examining the solution set of a given SDCP. We note that often an SDCP has only a finite number of solutions — in fact, there are many classes of SDCPs for which one can guarantee even the uniqueness of the solution. This is the main advantage of the SDCP approach, since examining the feasible region defined by an LMI for the existence of a rank one matrix is itself a difficult computation problem (the rank minimization problem under LMI constraints is a powerful framework for studying many robust control synthesis problems as well [14], [40]).

The results of §3.2 provide an explicit expression for the linear maps that appear in the SDCP formulation of the BMI based on the matrices $F'_{ij}s$ $(i = 1, \ldots, n; j = 1, \ldots, m)$. The complementarity theory can be used to classify instances of the BMI which reduce to a convex optimization problem over the positive semi-definite cone. For example, as the result of Corollary 3.1 that $\mathcal{B}$-positivity guarantee the existence of a feasible point for the BMI, thus reducing the BMI to characterizing a *single* matrix. The existence of the interior point methods for monotone SDCPs also provide local and in some cases, globally convergent methods for certain classes of the BMIs.

5 Conclusion

Rooted in some of best traditions in control theory, the BMI provides a framework to address the most important issues facing the field of robust control. The BMI formulation offers direct interpretation of the robust control synthesis problem, exemplified by a search for the parameters of the controller on one hand, and the multiplier on the other — and at the same time, it has lead to elegant mathematical investigations in optimization theory, some of which were presented in this chapter. Lastly, and most importantly, the BMI formulation promises efficient, reliable, and automated procedures for the synthesis of nonconservative robust controllers, improving upon those obtained by employing the existing methods, including the D-K and the M-K iterations.

References

[1] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis - A Modern Systems Theory Approach*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[2] G. J. Balas, J. C. Doyle, K. Glover, A. Packard, and R. Smith, *μ-Analysis and Synthesis Toolbox (μ-Tools)*, Mathworks, Natick, MA, 1991.

[3] E. Beran, L. Vanderberghe, and S. P. Boyd, *A global BMI algorithm based on the generalized Benders Decomposition*, Proceedings of the European Control Conference, 1997.

[4] A. Berman, *Cones, Matrices, and Mathematical Programming*, Springer-Verlag, 1973.

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

[6] S. P. Boyd, L. EL Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

[7] R. Y. Chiang and M. G. Safonov, *Robust Control Toolbox, Ver. 2.0*, Mathworks, Natick, MA, 1992.

[8] R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, 1992.

[9] G. B. Dantzig, *Linear Programming and Extensions*, Princeton, 1963.

[10] J. David, *Algorithms and Design for Robust Controllers*, Ph.D. Dissertation, Catholic University of Leuven, 1994.

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

[12] J. C. Doyle, A. Packard, and K. Zhou, *Review of LFTs, LMIs, and μ*, Proceedings of IEEE Conference on Decision and Control, Fort Lauderdale, FL, December 1991.

[13] J. C. Doyle, *Synthesis of robust controllers and filters with structured plant uncertainty*, Proceedings of IEEE Conference on Decision and Control, Albuquerque, NM, December 1983.

[14] L. El Ghaoui and P. Gahinet, *Rank minimization under LMI constraints: a framework for output feedback problems*, Proceedings of the European Control Conference, July 1993.

[15] M. K. H. Fan, A. L. Tits, and J. C. Doyle, *Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics*, IEEE Transactions on Automatic Control, 36 (1) (1991), pp. 25–38.

[16] C. A. Floudas and V. Visweswaran, *A primal relaxed dual global optimization approach*, Journal of Optimization Theory and Applications, 10 (4) (1972), pp. 237–260.

[17] A. M. Geofrion, *Generalized Benders Decomposition*, Journal of Optimization Theory and Algorithms, 10 (4) (1972), pp. 237–260.

[18] K. C. Goh, L. Turan, M. G. Safonov, G. P. Papavassilopoulos, and J. H. Ly, *Biaffine Matrix Inequality properties and computational methods*, Proceedings of the American Control Conference, July 1994.

[19] K. C. Goh, *Robust Control Synthesis via Bilinear Matrix Inequalities*, Ph.D. Dissertation, USC, May 1995.

[20] K. C. Goh, M. G. Safonov, and J. H. Ly, *Robust synthesis via bilinear matrix inequalities*, International Journal of Robust and Nonlinear Control (to appear).

[21] K. C. Goh, M. G. Safonov, and G. P. Papavassilopoulos, *A global optimization approach for the BMI problem*, Proceedings of the IEEE Conference on Decision and Control, Lake Buena Vista, FL, December 1994.

[22] M. S. Gowda, *Complementary problems over locally compact cones*, SIAM Journal of Control and Optimization, 27 (4) (1989), pp. 836–841.

[23] M. Hall, JR., *Combinatorial Theory*, Wiley, 1987.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, 1985.

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

[26] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Second Edition, Springer-Verlag, Berlin, 1993.

[27] P. Huard, *Resolution of mathematical programming with nonlinear constraint by the method of centers*, in J. Abadie, Editor, Nonlinear Programming, North-Holland, 1967.

[28] G. Isac, *Complementarity Problems*, Springer-Verlag, 1992.

[29] T. Iwasaki and R. E. Skelton, *A complete solution to the general $\mathcal{H}^\infty$ problem, LMI existence*

conditions and state-space formulas, Proceedings of the American Control Conference, San Francisco, CA, June 1993.

[30] M. Kojima, S. Shindoh, and S. Hara, *Interior-point methods for the monotone linear complementarity problem in symmetric matrices*, Technical report, Department of Information Science, Tokyo Institute of Technology, Japan, 1994.

[31] F. C. Lee, H. Flashner, and M. G. Safonov, *Positivity embedding for noncolocated and nonsquare flexible systems*, Applied Mathematics and Computation, 70 (2 & 3) (1995), pp. 233–246.

[32] F. C. Lee, H. Flashner, and M. G. Safonov, *An LMI approach to positivity embedding*, Proceedings of the American Control Conference, Seattle, WA, June 1995.

[33] F. C. Lee, H. Flashner, and M. G. Safonov, *Positivity-based control system synthesis using alternating LMIs* Proceedings of the American Control Conference, Seattle, WA, June 1995.

[34] S-M. Liu, *Parallel Algorithms for Global Non-Convex Minimization with Applications to Robust Control Design via BMIs*, Ph.D. Dissertation, USC, August 1995.

[35] S-M. Liu and G. P. Papavassilopoulos, *Numerical experience with parallel algorithms for solving the BMI problem*, Proceedings of the IFAC World Congress, July 1996.

[36] J. H. Ly, M. G. Safonov, and R. Y. Chiang, *On computation of multivariable stability margin using generalized Popov multipliers — LMI approach*, Proceeding of the American Control Conference, July 1994.

[37] J. H. Ly, K. C. Goh, and M. G. Safonov, *LMI approach to multiplier k_m/μ-synthesis*, Proceedings of the American Control Conference, Seattle, WA, June 1995.

[38] J. H. Ly, *A Multiplier Approach to μ/k_m Synthesis*, Ph.D. Dissertation, USC, February 1995.

[39] M. Mesbahi and G. P. Papavassilopoulos, *A cone programming approach to the bilinear matrix inequality and its geometry*, Mathematical Programming, 77 (2) (1997), pp. 247–272.

[40] M. Mesbahi and G. P. Papavassilopoulos, *On the rank minimization problem over a positive semi-definite linear matrix inequality*, IEEE Transactions on Automatic Control, 42 (2) (1997), pp. 239–243.

[41] M. Mesbahi, *Matrix Inequalities, Stability Theory, Interior Point Methods, and Distributed Computation*, Ph.D. Dissertation, USC, June 1996.

[42] M. Mesbahi, G. P. Papavassilopoulos, and M. G. Safonov, *Matrix cones, complementarity problems, and the bilinear matrix inequality*, Proceedings of the IEEE Conference on Decision and Control, New Orleans, LA, December 1995.

[43] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.

[44] K. J. Narendra and J. H. Taylor, *Frequency Domain Criteria for Absolute Stability*, Academic Press, NY, 1973.

[45] A. Packard, K. Zhou, P. Pandey, and G. Becker, *A collection of robust control problems leading to LMI's*, Proceedings of IEEE Conference on Decision and Control, Brighton, England, 1991.

[46] A. Packard, K. Zhou, P. Pandey, J. Leonhardson, and G. Balas, *Optimal, constant I/O similarity scaling for full-information and state-feedback control problems*, Systems and Control Letters, 19 (1992), pp. 271–280.

[47] A. Packard and J. C. Doyle, *The complex structured singular value*, Automatica, 21 (1), (1993) pp. 71–109.

[48] R. T. Rockafellar, *Convex Analysis*, Princeton, 1970.

[49] M. G. Safonov and M. Athans, *Gain and phase margin for multiloop LQG regulators*, Proceedings of the IEEE Conference on Decision and Control, Clearwater Beach, FL., December 1976.

[50] M. G. Safonov and M. Athans, *A multiloop generalization of the circle criterion for stability margin analysis*, IEEE Transactions on Automatic Control 26 (2), (1981) pp. 415–422.

[51] M. G. Safonov, *L^∞ optimal sensitivity vs. stability margin*, Proceedings of the IEEE Conference on Decision and Control, Albuquerque, NM, December 1983.

[52] M. G. Safonov and P. H. Lee, *A multiplier method for computing real multivariable stability margin*, Proceedings of the IFAC World Congress, July 1993.

[53] M. G. Safonov and R. Y. Chiang, *Real/complex K_m-synthesis without curve fitting*, in C.T. Leondes, Editor, Control and Dynamic Systems, 56, Academic Press, New York (1993),

pp. 303–324.

[54] M. G. Safonov, K. C. Goh, and J. H. Ly, *Control system synthesis via bilinear matrix inequalities*, Proceedings of the American Control Conference Baltimore, MD, June 1994.

[55] M. G. Safonov and G. P. Papavassilopoulos, *The diameter of an intersection of ellipsoids and BMI robust synthesis*, Proceedings of the IFAC Symposium on Robust Control Design, Rio de Janeiro, Brazil, September 1994.

[56] I. W. Sandberg, *On the L_2-boundedness of solutions of nonlinear functional equations*, Bell System Technical Journal, 43 (4) Part II (1964), pp. 1581–1599.

[57] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimension*, Springer-Verlag, 1970.

[58] O. Toker and H. Özbay, *On the NP-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback*, Proceedings of the American Control Conference, 1995.

[59] S. A. Vavasis, *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.

[60] V. Visweswaran, C. A. Floudas, M. G. Ierapetritou, and E. N. Pistikipoulos, *A Decomposition-Based Optimization Approach for Solving Bilevel Linear and Quadratic Programs*, Kluwer Academic Publishers, 1996.

[61] G. Zames, *On the input-output stability of time-varying nonlinear feedback systems — parts I and II*, IEEE Transactions on Automatic Control, 15 (2 & 3) (1966) pp. 228–238 & pp. 465–467.

[62] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice-Hall, Englewood Cliffs, NJ, 1996.