A Sequential Parametric Convex Approximation Method for Solving Bilinear Matrix Inequalities ⋆

Donghwan Lee [a], Jianghai Hu [a]

[a] *Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, USA*

Abstract

The goal of this paper is to develop an algorithm for solving optimization problems subject to bilinear matrix inequalities (BMIs), which are widely known to be of great significance in engineering, especially in control theory and its applications. Motivated by the convex-concave programming and path-following approaches, we propose a sequential convex optimization algorithm subject to a linear matrix inequality (LMI) constraint which approximates the BMI constraint. The feasible region of the LMIs is a convex inner approximation of that of the BMI constraints around the current iteration point. The approximations depend on variables that can be adjusted through the iterative convex subproblems. Its convergence property is also provided. In particular, it is proved that if all the feasible points satisfy the Mangasarian-Fromovitz constraint qualification, then there exists a subsequence of the subproblem solutions that converges to a stationary point of the BMI problem. Finally, an example of the static output-feedback controller design problem is provided for comparative analysis.

Key words: Cicero; Catiline; orations.

1 Introduction

This paper is devoted to a class of optimization problems minimizing a convex objective function subject to bilinear matrix inequality (BMI) constraints, called the BMI problem (BMIP) throughout the paper. This problem arises in many engineering applications, for example, the linear consensus protocol design [1], the resource allocation problem in wireless networks [2], the structured control design [3], and the switched controller design [4].

In particular, the BMIP arises frequently in the study of control systems. For example, the structured controller design problem (SCDP), which includes the reduced-order controller [5], static output-feedback controller (zero-order controller) [3], decentralized controller [6], and distributed controller design problems, is one of the most well-known problems in the systems and control theory that can be reduced to BMIPs. In [5], the SCDP was formulated as a concave programming, minimizing a concave objective function with linear matrix inequality (LMI) constraints, and solved by means of the the Frank and Wolfe feasible direction algorithm. In [7], the SCDP was converted into an equality constrained optimization, and then solved by using the augmented Lagrange multiplier method [8]. The freely available package, HIFOO [9], applied the Broyden-Fletcher-Goldfarb-Shanno algorithm [10] with penalty terms to the SCDP. In [11], a simple path-following method was suggested. It solves iteratively linearized semidefinite programming (SDP) subproblems, and can be easily applied to various BMIPs related to control design problems, for instance, the switched control design [4]. In [12], approaches based on the branch and bound method, interior-point method, and block coordinate descent method were considered.

In addition, it is of interest to note that some BMIPs can be converted to and solved as rank constrained optimization problems. For example, the SCDP was formulated as rank constrained LMI feasibility problems, and solved by using a Newton-like method [13], an alternating projection algorithm [14], and the cone complementarity linearization method [15]. Especially, the Newton-like method [13] was used for the LMIRank package.

Recently, the idea of the DC (difference of convex functions) programming [16–18] for the general non-convex programming has been extended to the optimizations with BMI constraints in [19]. The DC programming

⋆ This paper was not presented at any IFAC meeting.

Email addresses: lee1923@purdue.edu (Donghwan Lee), jianghai@purdue.edu (Jianghai Hu).

method is known as an effective technique for solving a class of non-convex optimizations, where the non-convex objective function or the non-convex functions in the inequality constraints are expressed as a difference of convex functions (or convex-concave decomposition). Then, the concave function is linearized to obtain a convex constraint whose feasible set is an inner convex approximation of the non-convex feasible set of the original inequality constraint.

On the other hand, there have been several researches on the more general nonlinear SDP (NLSDP) problems. Theoretical foundations on the first- and second-order necessary and sufficient conditions for the NLSDPs were investigated in [20]. In [21], the NLSDP was formulated as a maximum eigenvalue minimization problem, and the subgradient method was suggested. As a generalization of the sequential quadratic programming method for nonlinear programs, the sequential SDP was considered in [22] and [23], where the local quadratic convergence of the sequential SDP was proved under the second-order sufficient condition. In addition, a globally convergent sequential SDP algorithm was proposed in [24] based on improved merit functions. In [25], a modified augmented Lagrangian method was developed and proved to be suitable for the large-scale SDP and NLSDP. The algorithm was applied for the BMIPs in [26] as well.

The goal of this paper is to develop an algorithm to solve the BMIPs. The main contributions of this paper consist of the proposition of a sequential convex optimization algorithm and the proof of its convergence. In particular, it is proved that if all the feasible points satisfy the Mangasarian-Fromovitz constraint qualification [20], then there exists a subsequence of the sequential solutions of the convex subproblems that converges to a stationary point of the BMIPs.

The main idea of the paper was motivated by several approaches, such as the path-following method [11], the DC programming [19], the inner approximation algorithm [27], and the sequential parametric convex approximation (SPCA) method in [28]. Although these previous approaches have demonstrated their effectiveness through many applications, we found that depending on the BMIPs, alternative methods may have better performance. Even though it is a difficult task to provide theoretical comparisons, we present numerical examples that illustrate advantages of the proposed algorithm over the existing ones. In the proposed approach, instead of exploiting the DC decomposition in [19], an algebraic matrix inequality is used to obtain convex over approximations of the bilinear terms on the cone of positive semidefinite matrices. A significant difference of the proposed method compared to the DC programming is that auxiliary matrix variables are introduced in the approximations and sequentially adjusted by solving the subproblems so as to accelerate the convergence. Through numerical experiments on the SCDP, especially the static output-feedback control design problem [3], it is demonstrated that, in some cases, the proposed method outperforms the DC programming approach at the expense of higher computational efforts. On the other hand, the proposed method preserves some favorable properties of the DC programming approach. First, it is relatively easy to implement by using standard SDP solvers and to be modified to solve different BMIPs. In addition, at every iteration, a solution to the subproblem is guaranteed to be feasible, and no step size rule is required to ensure the feasibility and the convergence.

The proposed method can be categorized into the broader classes of nonlinear programming, such as the inner approximation algorithm [27] and the SPCA method [28]. In the sense that the bilinear terms are preserved as over approximations instead of ignoring them, it can be interpreted as an extension of the path-following algorithm [11] as well.

2 Preliminaries

In this paper, we follow the notation used in [19]. Let $\mathbb{S}^p$ be the set of symmetric matrices of size $p \times p$, $\mathbb{S}_+^p$ and $\mathbb{S}_{++}^p$ be the set of symmetric positive semidefinite and positive definite matrices, respectively. For given matrices X and Y in $\mathbb{S}^p$, the relation $X \succeq Y$ (respectively, $X \preceq Y$) means $X - Y \in \mathbb{S}_+^p$ (respectively, $Y - X \in \mathbb{S}_+^p$) and $X \succ Y$ (respectively, $X \prec Y$) is $X - Y \in \mathbb{S}_{++}^p$ (respectively, $Y - X \in \mathbb{S}_{++}^p$). The quantity $X \circ Y := \operatorname{trace}(X^T Y)$ is an inner product of two matrices X and Y defined on $\mathbb{S}^p$, where $\operatorname{trace}(Z)$ is the trace of matrix Z. In addition, the following standard notation will be used: $\operatorname{He}\{A\} := A^T + A$; I_n and I: $n \times n$ identity matrix and identity matrix of appropriate dimensions; $||\cdot||$: Euclidean norm of a vector or spectral norm of a matrix; s.t.: abbreviation of "subject to."

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be strongly convex [29, chapter 9.1.2] with parameter $\rho > 0$ if $f(\cdot) - (1/2)\rho \left\|\cdot\right\|^2$ is convex. We define the derivative of a matrix-valued mapping $\mathcal{F}$ at $z_0 \in \mathbb{R}^n$ as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^{p\times p}$ defined by

$$D_z\mathcal{F}[z]|_{z=z_0}[d] := \sum_{i=1}^{n} d_i \left.\frac{\partial \mathcal{F}[z]}{\partial z_i}\right|_{z=z_0}, \quad \forall d \in \mathbb{R}^n.$$

Let $\mathcal{A} : \mathbb{R}^n \to \mathbb{S}^p$ be a linear mapping defined as $\mathcal{A}[x] := \sum_{i=}^{n} x_i \mathcal{A}_i$, where $\mathcal{A}_i \in \mathbb{S}^p$ for $i \in \{1, \ldots, n\}$. The adjoint operator of $\mathcal{A}$, $\mathcal{A}^*$, is defined as $\mathcal{A}^* Z := [\mathcal{A}_1 \circ Z, \mathcal{A}_2 \circ Z, \ldots, \mathcal{A}_n \circ Z]^T$ for any $Z \in \mathbb{S}^p$. The concept of the convexity for the matrix-valued mapping is defined as follows.

Definition 1 ([20]) *The matrix-valued mapping $\mathcal{A}$: $\mathbb{R}^n \to \mathbb{S}^p$ is said to be positive semidefinite convex (psd-*

convex) on a convex subset $C \subseteq \mathbb{R}^n$ if for all $t \in [0, 1]$, and $x, y \in C$, one has $\mathcal{A}[tx + (1-t)y] \preceq t\mathcal{A}[x] + (1-t)\mathcal{A}[y]$. The matrix-valued mapping $\mathcal{A}$ is said to be positive semidefinite concave (psd-concave) on $C \subseteq \mathbb{R}^n$ if $-\mathcal{A}$ is psd-convex.

Consider the matrix-valued mapping $\mathcal{F} : \mathbb{R}^n \to \mathbb{S}^p$

$$\mathcal{F}(z) = C + \mathcal{L}[z] + \mathrm{He}\{\mathcal{A}[x]\mathcal{B}[y]\}, \tag{1}$$

where $z = [x^T, y^T]^T \in \mathbb{R}^n$, $x \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}^{n_y}$ are variables, $\mathcal{L} : \mathbb{R}^n \to \mathbb{S}^p$, $\mathcal{A} : \mathbb{R}^{n_x} \to \mathbb{R}^{p\times q}$, and $\mathcal{B} : \mathbb{R}^{n_y} \to \mathbb{R}^{q\times p}$ are linear mappings. In this paper, we consider a class of optimization problems with a BMI constraint (BMIP) of the form

$$\min_z \quad f(z) \quad \text{s.t.} \quad \mathcal{F}[z] \preceq 0, \quad z \in \Omega, \tag{2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and $\Omega \subset \mathbb{R}^n$ is a nonempty, closed, and convex set. The algorithm proposed in this paper can be directly extended to the optimization problems with multiple BMI constraints. Thus, the general case will not be considered here for notational simplicity and to save space.

To proceed, denote by $\mathcal{D} := \{z \in \Omega : \mathcal{F}[z] \preceq 0\}$ the feasible set of (2), and define $\mathcal{D}_0 := \{z \in \mathrm{ri}(\Omega) : \mathcal{F}[z] \prec 0\}$, where $\mathrm{ri}(\Omega)$ is the set of relative interior points of the convex set Ω [29, chapter 2.1.3]. Throughout the paper, the following assumptions apply.

Assumption 1 *$\mathcal{D}_0$ is nonempty.*

Assumption 2 *f is bounded from below on $\mathcal{D}$, and is differentiable.*

If we define the Lagrange function $L(z, \Lambda) = f(x) + \Lambda \circ \mathcal{F}[z]$, where $\Lambda \in \mathbb{S}^n$ is the Lagrange multiplier, the generalized KKT condition [30,31, Theorem 12.1] of (2), which includes an abstract set constraint $z \in \Omega$, can be written as

$$\begin{cases} 0 \in \nabla f(z) + D_z\mathcal{F}[z]^*\Lambda + \mathcal{N}_\Omega(z) \\ \mathcal{F}[z] \preceq 0, \quad \Lambda \succeq 0, \quad \mathcal{F}[z] \circ \Lambda = 0 \end{cases} \tag{3}$$

where $\mathcal{N}_\Omega(z)$ is the normal cone of Ω at z [30, Definition 12.7] defined by

$$\mathcal{N}_\Omega(z) := \begin{cases} \{w \in \mathbb{R}^n : w^T(z-y) \geq 0, \forall y \in \Omega\}, & \text{if } z \in \Omega \\ \emptyset, & \text{otherwise} \end{cases}$$

Note that the KKT condition in (3) is based on the KKT condition in [29, cahpter 11.6] for optimizations subject to generalized inequality constraints. The definitions of the KKT point and the stationary point are introduced below.

Definition 2 ([19, 30]) *A pair $(z, \Lambda) = (z^*, \Lambda^*)$ satisfying (3) is called a KKT point, z^* is called a stationary point, and Λ^* is called the corresponding multiplier of (2).*

The existence of the Lagrange multipliers is not always guaranteed. However, under certain conditions, called the constraint qualifications (CQs), the Lagrange multipliers exist. One of the useful CQs is the Mangasarian-Fromovitz constraint qualification (MFCQ) [30,32]. An MFCQ for optimizations subject to nonlinear SDPs was given in [20]. The MFCQ for the optimizations involving abstract convex set constraints is provided in [32, Proposition 3.3.12]. Combining the two results, we can obtain the following MFCQ for optimizations with abstract set and nonlinear SDP constraints.

Definition 3 (MFCQ [20, 32]) *We say that the MFCQ holds at a feasible point $z_0 \in \Omega$ of (2) if there exists a vector $z \in \Omega$ such that*

$$\mathcal{F}[z_0] + D_z\mathcal{F}[z]|_{z=z_0}[z - z_0] \prec 0.$$

Under the MFCQ, the first-order necessary condition for the optimality can be obtained.

Lemma 1 (First-order necessary condition [20]) *Let $z_0 \in \mathbb{R}^n$ be a locally optimal solution of the problem (2), and suppose that the MFCQ holds at z_0. Then, the set of the Lagrange multiplier $\Lambda \in \mathbb{S}^n$ satisfying the KKT condition (3) is nonempty and bounded.*

3 A sequential parametric convex approximation method

In this section, we propose an algorithm for solving the BMIP (2). The proposed method can be also viewed as an extension of the sequential parametric convex approximation method [28] to the BMIP (2).

To explain the proposed approach, we reformulate the matrix $\mathcal{F}[z]$ in (2) by

$$\begin{aligned} &\mathcal{F}[z] \\ &= C + \mathcal{L}[z] + \mathrm{He}\{\mathcal{A}[x]\mathcal{B}[y]\} \\ &= C + \mathcal{L}[z - z_k + z_k] \\ &+ \mathrm{He}\{\mathcal{A}[x - x_k + x_k]\mathcal{B}[y - y_k + y_k]\} \\ &= \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z - z_k] \\ &+ \mathrm{He}\{\mathcal{A}[x - x_k]\mathcal{B}[y - y_k]\}, \end{aligned} \tag{4}$$

where $z_k := [x_k^T, y_k^T]^T$, and

$$\begin{aligned} D_z\mathcal{F}[z]|_{z=z_k}[z - z_k] \\ = \mathcal{L}[z - z_k] + \mathrm{He}\{\mathcal{A}[x_k]\mathcal{B}[y - y_k]\} \\ + \mathrm{He}\{\mathcal{A}[x - x_k]\mathcal{B}[y_k]\}. \end{aligned}$$

If the last term in (4) is ignored, then the linearization $\mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z-z_k]$ of $\mathcal{F}[z]$ around z_k is obtained. The path-following algorithm in [11] for the BMIP (2) is to solve (2) with the BMI constraint replaced with the linearized constraint. Instead of dropping the bilinear term $\text{He}\{\mathcal{A}[x-x_k]\mathcal{B}[y-y_k]\}$ in (4), we can obtain an over estimation of $\mathcal{F}[z]$ over the cone of positive semidefinite matrices by using a matrix inequality, which is often used in the control theory literature, e.g., [33, Proposition 2.1 and Proposition 2.2], and [34].

Lemma 2 *Let A and E be real matrices of appropriate dimensions. Then, for any $S \in \mathbb{S}^n_{++}$,*

$$\text{He}\{DE\} \preceq DSD^T + E^T S^{-1} E.$$

PROOF. For any $S \in \mathbb{S}^n_{++}$, the inequality is obtained by expanding $(D^T - S^{-1}E)^T S (D^T - S^{-1}E) \succeq 0$. □

Using Lemma 2, an upper bound on $\text{He}\{\mathcal{A}[x-x_k]\mathcal{B}[y-y_k]\}$ in (4) is obtained as

$$\begin{aligned}&\text{He}\{\mathcal{A}[x-x_k]\mathcal{B}[y-y_k]\}\\ &\quad\preceq \mathcal{A}[x-x_k]S\mathcal{A}[x-x_k]^T\\ &\quad + \mathcal{B}[y-y_k]^T S^{-1}\mathcal{B}[y-y_k],\end{aligned}$$

where $S \in \mathbb{S}^n_{++}$. Thus, for any given $z_k \in \Omega$, an over estimation of $\mathcal{F}[z]$ is given by

$$\begin{aligned}\mathcal{F}[z] &\preceq \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z-z_k]\\ &\quad + \mathcal{A}[x-x_k]S\mathcal{A}[x-x_k]^T\\ &\quad + \mathcal{B}[y-y_k]^T S^{-1}\mathcal{B}[y-y_k]\\ &\quad =: \mathcal{K}[z;\, z_k,\, S]. \end{aligned} \tag{5}$$

The mapping $\mathcal{K}[z;\, z_k,\, S]$ in (5) has the following properties.

Proposition 1 *For any given $z_k \in \Omega$, $S \in \mathbb{S}^n_{++}$, the matrix-valued mapping $\mathcal{K}[\cdot;\, z_k,\, S]$ in (5) satisfies the following properties:*

(1) $\mathcal{F}[z] \preceq \mathcal{K}[z;\, z_k,\, S], \quad \forall z \in \mathbb{R}^n$;
(2) $\mathcal{F}[z_k] = \mathcal{K}[z_k;\, z_k,\, S]$;
(3) $D_z\mathcal{K}[z;\, z_k,\, S]|_{z=z_k}[u] = D_z\mathcal{F}[z]|_{z=z_k}[u], \quad \forall u \in \mathbb{R}^n$.

PROOF. The statement 1) was proved in (5), and the statement 2) can be proved by setting $z = z_k$ in (5). To prove 3), note that $D_z\mathcal{F}[z]|_{z=z_k}[u]$ is given by

$$D_z\mathcal{F}[z]|_{z=z_k}[u] = \mathcal{L}[u] + \text{He}\{\mathcal{A}[x_k]\mathcal{B}[w]\} + \text{He}\{\mathcal{A}[v]\mathcal{B}[y_k]\}.$$

where $v \in \mathbb{R}^{n_x}$, $w \in \mathbb{R}^{n_y}$ are appropriately partitioned vectors such that $u = [v,\, w]^T$. In addition, $D_\xi\mathcal{K}(\xi;\, z_k,\, S)|_{\xi=z}[u]$ is obtained as

$$\begin{aligned}&D_\xi\mathcal{K}(\xi;\, z_k,\, S)|_{\xi=z}[u]\\ &\quad = D_\xi\mathcal{F}(z)|_{\xi=z}[u]\\ &\quad + \mathcal{A}[x-x_k]S\mathcal{A}[v]^T + \mathcal{A}[v]S\mathcal{A}[x-x_k]^T\\ &\quad + \mathcal{B}[w]^T S^{-1}\mathcal{B}[y-y_k] + \mathcal{B}[y-y_k]^T S^{-1}\mathcal{B}[w].\end{aligned}$$

Plugging $z = z_k = [x_k^T,\, y_k^T]^T$ into the above equality yields $D_\xi\mathcal{K}(\xi;\, z_k,\, S)|_{\xi=z_k}[u] = D_\xi F(z)|_{\xi=z_k}[u]$, concluding the proof. □

It is worth mentioning that the above properties of the approximation $\mathcal{K}[z;\, z_k,\, S]$ are equivalent to that given in [27, page 682]. Therefore, the set of $z \in \Omega$ such that $\mathcal{K}[z;\, z_k,\, S] \preceq 0$ is an inner approximation of the feasible set of (2) around z_k. Instead of solving (2), we can solve the following approximate problem for a fixed $S \in \mathbb{S}^n_{++}$:

$$\begin{cases}\min\limits_z f(z) \quad \text{s.t.}\\ \mathcal{K}[z;\, z_k,\, S] \preceq 0, \quad z \in \Omega\end{cases} \tag{6}$$

It is easy to prove that $\mathcal{K}[z;\, z_k,\, S]$ is psd-convex in $z \in \mathbb{R}^n$. Therefore, the problem (6) is a convex optimization. The path-following approach in [11] replaces the whole BMI constraints with their linearizations, and the DC programming approach in [19] replaces only the concave terms in the BMIs with their linearizations while preserving the affine and convex terms. In the proposed approach, we replace the bilinear terms with their convex quadratic over approximations while preserving the affine terms. If an optimal solution x^* to the above problem is close to x_k, then the constraint (6) approximates the constraint of (2) because of the statement (2) of Proposition 1. Using the Schur complement, (6) is converted to the equivalent form

$$\min_{z\in\Omega} f(z) \quad \text{s.t.}$$
$$\begin{bmatrix}\mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z-z_k] & * & *\\ \mathcal{A}[z-z_k]^T & -S^{-1} & *\\ \mathcal{B}[z-z_k] & 0 & -S\end{bmatrix} \preceq 0. \tag{7}$$

The inequality (7) is an LMI constraint, and the above convex optimization can be solved using standard convex optimization techniques [29]. If the set Ω consists of LMI constraints, then it can be solved by using SDP solvers. By repeatedly solving the problem and using the current optimal value x^* for the next point x_{k+1}, we can obtain an iterative convex optimization algorithm for the BMIP.

Although it is difficult to compare the tightness of the proposed approximation with that of the DC programming, then the quality of the approximation can be improved since the gap between the over approximation and the original bilinear term in Lemma 2 can be reduced by an appropriate choice of S. However, the determination of S in (7) is a non-convex problem due to the inverse of S in (7). One possible way to determine $S \succ 0$ through the convex program is to linearize the inverse $-S^{-1}$. A key observation is that $f(S) = S^{-1}$ is psd-convex on $\mathbb{S}^n_{++}$ by [19, Lemma 3.1], and hence, $g(S) = -S^{-1}$ is psd-concave. By linearizing $-S^{-1}$ around S_k, we have the over approximation $-S^{-1} \preceq -2S_k^{-1} + S_k^{-1} S S_k^{-1}$ around S_k. For completeness and easy reference, it is formally stated in the following lemma.

Lemma 3 *Suppose that $\mathcal{S} : \mathbb{R}^n \to \mathbb{S}^p$ is a linear mapping defined as $\mathcal{S}[x] = \sum_{i=1}^n x_i \mathcal{S}_i$, where $\mathcal{S}_i \in \mathbb{S}^p$ for $i \in \{1, \ldots, n\}$. If $\mathcal{S}[x] \succ 0$ and $\mathcal{S}[y] \succ 0$, then $-\mathcal{S}[y]^{-1} \preceq -2\mathcal{S}[x]^{-1} + \mathcal{S}[x]^{-1}\mathcal{S}[y]\mathcal{S}[x]^{-1}$ holds.*

PROOF. From the formulation of the derivative of the matrix inverse [35], we have

$$\begin{aligned} D_z(-\mathcal{S}[z]^{-1})\Big|_{z=x}[d] &= -\sum_{i=1}^n d_i \frac{\partial(\mathcal{S}[z]^{-1})}{\partial z_i}\Bigg|_{z=x} \\ &= -\sum_{i=1}^n d_i \left\{ -\mathcal{S}[z]^{-1} \frac{\partial \mathcal{S}[z]}{\partial z_i} \mathcal{S}[z]^{-1} \right\}\Bigg|_{z=x} \\ &= \mathcal{S}[x]^{-1}(D_z\mathcal{S}[z]|_{z=x}[d])\mathcal{S}[x]^{-1}. \end{aligned}$$

Using this result, if $\mathcal{S}[x]$ and $\mathcal{S}[y]$ are invertible, then the linearization of $-\mathcal{S}[y]^{-1}$ around x is

$$-\mathcal{S}[x]^{-1} + \mathcal{S}[x]^{-1}(D_z\mathcal{S}[z]|_{z=x}[y-x])\mathcal{S}[x]^{-1}.$$

Since the mapping $g(S) = -S^{-1}$ is psd-concave on $\mathbb{S}^n_{++}$, for $\mathcal{S}[x] \succ 0$ and $\mathcal{S}[y] \succ 0$, we have

$$\begin{aligned} -\mathcal{S}[y]^{-1} &\preceq -\mathcal{S}[x]^{-1} + \mathcal{S}[x]^{-1}(D_z\mathcal{S}[z]|_{z=x}[y-x])\mathcal{S}[x]^{-1} \\ &= -2\mathcal{S}[x]^{-1} + \mathcal{S}[x]^{-1}\mathcal{S}[y]\mathcal{S}[x]^{-1}. \end{aligned}$$

This completes the proof. □

Algorithm 1 A sequential parametric convex approximation algorithm for BMIP

Initialize $z_0 \in \mathcal{D}_0$ and set $k \leftarrow 0$, $S_k = I_n$;
repeat
Solve

$$(z_{k+1}, S_{k+1}) = \operatorname*{arg\,min}_{z\in\mathbb{R}^n,\, S\in S^n} f_\rho(z; z_k) \quad \text{s.t.} \tag{9}$$

$$\begin{bmatrix} \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z-z_k] & * & * \\ S_k\mathcal{A}[z-z_k]^T & -2S_k+S & * \\ \mathcal{B}[z-z_k] & 0 & -S \end{bmatrix} \preceq 0,$$

$$c_1 I \preceq S \preceq c_2 I, \quad -2S_k + S \preceq -c_3 I \quad z \in \Omega, \tag{10}$$

where $\rho > 0$, $c_2 > c_1 > 0$, $c_3 > 0$, $f_\rho(z; z_k) := f(z) + \frac{\rho}{2}\|z - z_k\|^2$.
$k \leftarrow k+1$;
until a certain stopping criterion is satisfied.

By replacing $-S^{-1}$ in (7) with $-2S_k^{-1} + S_k^{-1}SS_k^{-1}$ and multiplying (7) from the left and right by the block diagonal matrix $\mathrm{diag}(I,\, S_k,\, I)$, we obtain the following convex optimization:

$$\min_{z\in\Omega,\, S\in\mathbb{S}^n} f(z) \quad \text{s.t.} \tag{8}$$

$$\begin{bmatrix} \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z-z_k] & * & * \\ S_k\mathcal{A}[z-z_k]^T & -2S_k+S & * \\ \mathcal{B}[z-z_k] & 0 & -S \end{bmatrix} \preceq 0,$$

$$S \succ 0, \quad 2S_k - S \succ 0.$$

By sequentially solving the above convex program and using the current optimal point for the next point x_{k+1} and S_{k+1}, Algorithm 1 shown at the top of the next page is obtained.

Remark 1 *Due to some technical reasons, the optimization* (8) *is modified in the subproblem of Algorithm 1. In particular, the constraint $c_1 I \preceq S \preceq c_2 I$ in* (10) *ensures that each S_k is nonsingular and the sequence $\{S_k\}_{k\geq 0}$ is bounded. Moreover, $-2S_k + S \preceq -c_3 I$ is included in the algorithm to guarantee that $-2S_k + S_{k+1}$ is nonsingular for all $k \geq 0$. The term $\frac{\rho}{2}\|z - z_k\|^2$ in the objective function in* (9) *is a regularization term to guarantee that the value of the function f is strictly descent at each iteration, e.g., see the statement (1) of Lemma 6.*

Similarly to the DC programming in [19], a favorable feature of Algorithm 1 is that the optimal solution of the subproblem at each iteration is a feasible point of the original problem (2).

Proposition 2 *Let $\{(z_k,\, S_k)\}_{k\geq 0}$ be a sequence of optimal solutions generated by Algorithm 1. For every $k \geq 0$, z_k is a feasible solution to* (2)*, i.e., $z_k \in \Omega$, $\mathcal{F}[z_k] \preceq 0$.*

PROOF. The proof will be completed by the induction argument. From Algorithm 1, the feasibility of the initial point is guaranteed, i.e., $z_0 \in \Omega$, $\mathcal{F}[z_0] \preceq 0$. Suppose that z_k is a feasible point of (2). Then, it is guaranteed that the subproblem in Algorithm 1 is always feasible for any $S_k \in \mathbb{S}^n_{++}$ because of the trivial feasible point $(z, S) = (z_k, S_k)$. Let (z_{k+1}, S_{k+1}) be an optimal solution to the subproblem in Algorithm 1. By plugging $(z, S) = (z_{k+1}, S_{k+1})$ into the constraint of (9), and applying the Schur complement, we get

$$\begin{bmatrix} \begin{pmatrix} \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z_{k+1} - z_k] \\ +\mathcal{B}[z_{k+1} - z_k]^T S_{k+1}^{-1} \mathcal{B}[z_{k+1} - z_k] \end{pmatrix} & * \\ S_k \mathcal{A}[z_{k+1} - z_k]^T & -2S_k + S_{k+1} \end{bmatrix} \preceq 0.$$

Multiplying the above matrix by $[I, \mathcal{A}[z_{k+1} - z_k]]$ from the left and by its transpose from the right, we have

$$\begin{aligned} &\mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z_{k+1} - z_k] \\ &\quad + \mathcal{A}[z_{k+1} - z_k]^T S_{k+1} \mathcal{A}[z_{k+1} - z_k]^T \\ &\quad + \mathcal{B}[z_{k+1} - z_k]^T S_{k+1}^{-1} \mathcal{B}[z_{k+1} - z_k] \\ &\quad = \mathcal{K}[z_{k+1}; z_k, S_{k+1}] \\ &\quad \preceq 0. \end{aligned}$$

By the statement (1) of Proposition 1, this implies that $z_{k+1} \in \Omega$ satisfies $\mathcal{F}[z_{k+1}] \preceq 0$. Therefore, z_{k+1} is a feasible point of (2). This completes the proof. □

Remark 2 *Both the DC programming approach in [19] and Algorithm 1 use over approximations of the BMI constraints over the cone of positive semidefinite matrices. Intuitively, the difference between the two approaches can be explained using the following simple example. Consider the BMI constraint* $\Gamma(X, Y) + X^T Y + Y^T X \preceq 0$*, where the matrices* X *and* Y *are the decision variables, and the matrix-valued mapping* $\Gamma(X, Y)$ *is linear in* (X, Y)*. According to [19, Lemma 3.1], the bilinear term* $X^T Y + Y^T X$ *can be represented as a difference of psd-convex and psd-concave terms (psd-convex-concave mapping):* $X^T Y + Y^T X = X^T X + Y^T Y - (X - Y)^T (X - Y)$*. In fact, a more general psd-convex-concave mapping can be derived by introducing the auxiliary matrix* $S \succ 0$ *as follows:* $X^T Y + Y^T X = X^T S X + Y^T S^{-1} Y - (X - S^{-1} Y)^T S (X - S^{-1} Y)$*. Note that the last term* $-(X - S^{-1}Y)^T S (X - S^{-1} Y)$ *is concave in* X *and* Y*. If we set* S *to be a constant, i.e.* $S = I$*, and linearize the last term* $(X - S^{-1}Y)^T S (X - S^{-1}Y)$ *with respect to* (X, Y) *at the point* $(X, Y) = (X_k, Y_k)$*, then the over approximation of the DC programming method is obtained. Instead, if we drop the last term and linearize* S^{-1} *at* $S = S_k$*, then the over approximation of the proposed Algorithm 1 is obtained. From the interpretation, it is not easy to claim which approximation is better than the other. Since the last term is entirely dropped in Algorithm 1, it can be seen as a less accurate one in general. However, the auxiliary matrix* S *can be adjusted as a decision variable of the convex subproblem, the over approximation can be tightened at each iteration.*

4 Convergence analysis

In this section, a convergence analysis of Algorithm 1 will be provided. First of all, the convex subproblem of Algorithm 1 is converted to a simpler yet equivalent form.

Proposition 3 *Suppose that* $\{(z_k, S_k)\}_{k \geq 0}$ *is a sequence of the optimal solutions generated by Algorithm 1. Then, for each* $k \geq 0$*,* z_{k+1} *is an optimal solution to the following optimization problem:*

$$\min_{z \in \mathcal{D}(z_k, S_k, S_{k+1})} f_\rho(z; z_k) \tag{11}$$

where

$$\mathcal{D}(z_k, S_k, S_{k+1}) := \{z \in \Omega : \mathcal{H}[z; z_k, S_k, S_{k+1}] \preceq 0\}, \tag{12}$$

$$\begin{aligned} &f_\rho(z; z_k) := f(z) + \frac{\rho}{2}\|z - z_k\|^2, \\ &\mathcal{H}[z; z_k, S_k, S_{k+1}] := \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z - z_k] \\ &\qquad + \mathcal{A}[z - z_k] S_k (2S_k - S_{k+1})^{-1} S_k \mathcal{A}[z - z_k]^T \\ &\qquad + \mathcal{B}[z - z_k]^T S_{k+1}^{-1} \mathcal{B}[z - z_k]. \end{aligned} \tag{13}$$

PROOF. If we plug $S = S_{k+1}$ into the constraint of (9), then z_{k+1} is the unique optimal solution to

$$\begin{aligned} &\min_{z \in \Omega} f_\rho(z; z_k) \quad \text{s.t.} \\ &\begin{bmatrix} \mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[z - z_k] & * & * \\ S_k \mathcal{A}[z - z_k]^T & -2S_k + S_{k+1} & * \\ \mathcal{B}[z - z_k] & 0 & -S_{k+1} \end{bmatrix} \\ &\preceq 0, \end{aligned} \tag{14}$$

where the uniqueness follows from the fact that $f_\rho(z; z_k)$ is strongly convex with the parameter $\rho > 0$ and the feasible set is convex. Since $S_{k+1} \succ 0$ and $S_{k+1} - 2S_k \succ 0$, we can apply the Schur complement twice to (14) to obtain $\mathcal{H}[z; z_k, S_k, S_{k+1}] \preceq 0$. Note that from the standard Schur complement argument, $\mathcal{H}[z; z_k, S_k, S_{k+1}] \preceq 0$ and the LMI constraint in (9) are equivalent. Therefore, the problem (9) can be equivalently converted to (11). □

In the sequel, the convergence analysis of Algorithm 1 will be carried out based on the optimization (11). Therefore, it is of great interest to understand the relations among the mappings $\mathcal{H}[\cdot;\, z_k,\, S_k,\, S_{k+1}]$, $\mathcal{K}[\cdot;\, z_k,\, S_k]$, and $\mathcal{F}[\cdot]$.

Lemma 4 *For any given $z_k \in \Omega$, $S_k \in \mathbb{S}^n_{++}$, $S_{k+1} \in \mathbb{S}^n_{++}$, the matrix-valued mapping $\mathcal{H}[\cdot;\, z_k,\, S_k,\, S_{k+1}]$ satisfies the following properties*

(1) $\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}] = \mathcal{K}[z;\, z_k,\, S_k]$ *if* $S_{k+1} = S_k$;
(2) $\mathcal{F}[z_k] = \mathcal{H}[z_k;\, ,\, z_k,\, S_k,\, S_{k+1}]$;
(3) $D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_k}[u] = D_z\mathcal{F}[z]|_{z=z_k}[u]$ *for* $u \in \mathbb{R}^n$;
(4)

$$\begin{aligned} &D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}[u] \\ &\quad = D_z\mathcal{F}[z]|_{z=z_k}[u] \\ &\quad + \mathrm{He}\{\mathcal{A}[v]S_k(2S_k - S_{k+1})^{-1}S_k\mathcal{A}[x_{k+1} - x_k]^T\} \\ &\quad + \mathrm{He}\{\mathcal{B}[w]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]\}, \end{aligned}$$

where $u = [v^T,\, w^T]^T$ *and* $z_k = [x_k^T,\, y_k^T]^T$;
(5) $\mathcal{K}[z,\, z_k,\, S_{k+1}] \preceq \mathcal{H}[z,\, z_k,\, S_k,\, S_{k+1}], \quad \forall z \in \mathbb{R}^n$;
(6) $\mathcal{F}[z] \preceq \mathcal{H}[z,\, z_k,\, S_k,\, S_{k+1}], \quad \forall z \in \mathbb{R}^n$.

PROOF. The statements 1)-4) can be proved using direct calculations, and are thus omitted for brevity. To prove the statement (5), observe that from Lemma 3, we obtain

$$\begin{aligned} &-S_{k+1}^{-1} \preceq -2S_k^{-1} + S_k^{-1}S_{k+1}S_k^{-1} \\ &\quad \Leftrightarrow 2S_k^{-1} - S_k^{-1}S_{k+1}S_k^{-1} \preceq S_{k+1}^{-1}. \end{aligned}$$

Pre- and post-multiplying the above inequality by S_k results in $2S_k - S_{k+1} \preceq S_k S_{k+1}^{-1} S_k$. Taking the inverse of the left- and right-hand sides of the last inequality, one gets $S_k^{-1}S_{k+1}S_k^{-1} \preceq (2S_k - S_{k+1})^{-1}$. Applying the inequality to the definition of $\mathcal{H}[z,\, z_k,\, S_k,\, S_{k+1}]$ proves the statement (5). The statement (6) follows by combining the statement (5) and the relation (1) in Proposition 1. This completes the proof. □

The above properties coincide with the properties given in [27, page 682]. Therefore, the proposed approach can be interpreted as a SPCA method for the BMIP. The following lemma introduces the KKT condition for (11).

Lemma 5 *Let*

$$z_{k+1} = \operatorname*{arg\,min}_{z \in \mathcal{D}(z_k,\, S_k,\, S_{k+1})} f_\rho(z;\, z_k),$$

and suppose that $z_{k+1} \in \mathcal{D}(z_k,\, S_k,\, S_{k+1})$ satisfies the MFCQ for the problem (11), i.e., there exists $\xi \in \Omega$ such that

$$\begin{aligned} &\mathcal{H}[z_{k+1},\, z_k,\, S_k,\, S_{k+1}] \\ &\quad + D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}(\xi - z_{k+1}) \prec 0. \end{aligned}$$

Then, there exists a Lagrange multiplier Λ_{k+1} satisfying the KKT condition

$$\begin{aligned} &0 \in \nabla f(z_{k+1}) + \rho(z_{k+1} - z_k) \\ &\quad + \left[D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}\right]^* \Lambda_{k+1} + \mathcal{N}_\Omega(z_{k+1}), \end{aligned} \tag{15}$$

$$\mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] \preceq 0, \quad \Lambda_{k+1} \succeq 0, \tag{16}$$

$$\Lambda_{k+1} \circ \mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] = 0. \tag{17}$$

PROOF. By assumption, since z_{k+1} is an optimal solution to (11), and $z_{k+1} \in \mathcal{D}(z_k,\, S_k,\, S_{k+1})$ satisfies the MFCQ, by Lemma 1, the set of the Lagrange multiplier Λ_{k+1} solving the KKT (15)-(17) is nonempty and bounded [20, page 306]. This completes the proof. □

The following result proves that the sequence of the objective functions $\{f_\rho(z_{k+1};\, ,\, z_k)\}_{k \geq 0}$ is monotonically decreasing.

Lemma 6 *Suppose that $\{(z_k,\, S_k)\}_{k \geq 0}$ is a sequence of solutions generated by Algorithm 1. The, the following statements are true:*

(1) $f(z_{k+1}) \leq f(z_k) - \frac{\rho}{2}\|z_k - z_{k+1}\|^2$;
(2) *If there exists $z \in \Omega$ such that $\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}] \prec 0$, then*

$$\begin{aligned} &f(z_{k+1}) - f(z_k) \\ &\quad \leq -\rho\|z_{k+1} - z_k\|^2 \\ &\quad - \Lambda_{k+1} \circ [\mathcal{A}[x_{k+1} - x_k]S_k(2S_k - S_{k+1})^{-1} \\ &\quad \times S_k\mathcal{A}[x_{k+1} - x_k]^T] \\ &\quad - \Lambda_{k+1} \circ [\mathcal{B}[y_{k+1} - y_k]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]], \end{aligned} \tag{18}$$

where Λ_{k+1} is a Lagrange multiplier satisfying the KKT condition in (15)-(17).

PROOF. (1) First of all, from the statement 2) of Lemma 4 and Proposition 2, we have $\mathcal{F}(z_k) = \mathcal{H}(z_k,\, z_k,\, S_k,\, S_{k+1}) \preceq 0$ and $z_k \in \Omega$. By the definition of $\mathcal{D}(z_k,\, S_k,\, S_{k+1})$ in (12), one concludes $z_k \in \mathcal{D}(z_k,\, S_k,\, S_{k+1})$. In addition, since z_{k+1} is an optimal solution of (11), we have

$$\begin{aligned} &f(z_{k+1}) + \frac{\rho}{2}\|z_{k+1} - z_k\|^2 \\ &\quad = \min_{z \in \mathcal{D}(z_k,\, S_k,\, S_{k+1})} f_\rho(z;\, z_k) \end{aligned}$$

$$\leq f(z_k) + \frac{\rho}{2}\|z_k - z_k\|^2$$
$$= f(z_k),$$

and the desired result follows.

(2) The proof follows similar lines of the proof of [19, Lemma 4.2]. From the assumption of the convexity of $f(z)$, $f(z_k) \geq f(z) + \nabla f(z)^T(z_k - z)$, it follows from the first inclusion in the KKT condition (15) that

$$\begin{aligned} & f(z_k) - f(z_{k+1}) + ([D_z\mathcal{H}]^*\Lambda_{k+1})^T(z_k - z_{k+1}) \\ & \quad \geq (\nabla f(z_{k+1})^T + [D_z\mathcal{H}]^*\Lambda_{k+1})(z_k - z_{k+1}) \\ & \quad \geq \rho(z_{k+1} - z_k)^T(z_{k+1} - z_k), \end{aligned} \tag{19}$$

where the superscript $*$ is the adjoint operator, we use the shorthand notation

$$D_z\mathcal{H} := D_z\mathcal{H}(z;\, z_k,\, S_k,\, S_{k+1})|_{z=z_{k+1}},$$

and the last inequality follows from the KKT condition (15)

$$-\nabla f(z_{k+1}) - \rho(z_{k+1} - z_k) - [D_z\mathcal{H}]^*\Lambda_{k+1} \in \mathcal{N}_\Omega(z).$$

On the other hand, by plugging $u = z_k - z_{k+1}$ into the statement 4) of Lemma 4, we have

$$\begin{aligned} & D_z\mathcal{H}[z_k - z_{k+1}] \\ & \quad = -\mathcal{H}[z_{k+1},\, z_k,\, S_k,\, S_{k+1}] + \mathcal{F}[z_k] \\ & \quad - \mathcal{A}[x_{k+1} - x_k]S_k(2S_k - S_{k+1})^{-1}S_k\mathcal{A}[x_{k+1} - x_k]^T \\ & \quad - \mathcal{B}[y_{k+1} - y_k]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]. \end{aligned}$$

Thus,

$$\begin{aligned} & ([D_z\mathcal{H}]^*\Lambda_{k+1})^T(z_k - z_{k+1}) \\ & \quad = \Lambda_{k+1} \circ \{D_z\mathcal{H}[z_k - z_{k+1}]\} \\ & \quad = \Lambda_{k+1} \circ \{-\mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] + \mathcal{F}[z_k] \\ & \quad - \mathcal{A}[x_{k+1} - z_k]S_k(2S_k - S_{k+1})^{-1}S_k A(x_{k+1} - x_k)^T \\ & \quad - \mathcal{B}[y_{k+1} - y_k]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]\} \\ & \quad \leq -\Lambda_{k+1} \circ \{\mathcal{A}[x_{k+1} - x_k]S_k(2S_k - S_{k+1})^{-1}S_k \\ & \quad \times \mathcal{A}[x_{k+1} - x_k]^T\} \\ & \quad - \Lambda_{k+1} \circ \{\mathcal{B}[y_{k+1} - y_k]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]\}, \end{aligned}$$

where the last inequality follows from the KKT condition (17), i.e., $\Lambda_{k+1} \circ \mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] = 0$, and from the fact that $\Lambda_{k+1} \succeq 0$ and $\mathcal{F}[z_k] \preceq 0$. Thus, (19) can be represented by

$$\begin{aligned} & f(z_k) - f(z_{k+1}) + ([D_z\mathcal{H}]^*\Lambda_{k+1})^T(z_k - z_{k+1}) \\ & \quad = f(z_k) - f(z_{k+1}) \\ & \quad + \Lambda_{k+1} \circ D_z\mathcal{H}(z_k - z_{k+1}) \\ & \quad \leq f(z_k) - f(z_{k+1}) \\ & \quad - \Lambda_{k+1} \circ \{\mathcal{A}[x_{k+1} - x_k]S_k \\ & \quad \times (2S_k - S_{k+1})^{-1}S_k\mathcal{A}[x_{k+1} - x_k]^T\} \\ & \quad - \Lambda_{k+1} \circ \{\mathcal{B}[y_{k+1} - y_k]^T S_{k+1}^{-1}\mathcal{B}[y_{k+1} - y_k]\}. \end{aligned}$$

Combining the last inequality with (19), the proof is completed. □

Based on Lemma 6, it can be proved that the sequence of the the optimal objective function value of the subproblem of Algorithm 1 is nonincreasing, and thus, converges.

Proposition 4 *If $\{z_k\}_{k\geq 0}$ is the sequence generated by* (9)*, then $\{f(z_k)\}_{k\geq 0}$ converges.*

PROOF. By Lemma 6, the sequence $\{f(z_k)\}_{k\geq 0}$ is nonincreasing. From Assumption 2, since f is bounded from below on $\mathcal{D}$, $\{f(z_k)\}_{k\geq 0}$ is bounded from below as well, thus converges. □

In what follows, we propose the main convergence result, which claims that if every feasible point of the BMIP (2) satisfies the MFCQ and the sequence of the optimal solutions $\{z_k\}_{k\geq 0}$ of the subproblem of Algorithm 1 is bounded, then there exists a limit point of $\{z_k\}_{k\geq 0}$ that is a stationary point of (2).

Proposition 5 *Let $\{(z_k,\, S_k)\}_{k\geq 0}$ be the sequence of solutions generated by* (9)*. If every feasible point of* (2) *satisfies the MFCQ and $\{z_k\}_{k\geq 0}$ is bounded, then there exists a limit point of $\{z_k\}_{k\geq 0}$ that is a stationary point of* (2)*.*

PROOF. Lemma 6 implies

$$\frac{\rho}{2}\|(z_{k+1} - z_k)\|^2 \leq f(z_k) - f(z_{k+1}), \quad \forall k \geq 0.$$

By Proposition 4, $\{f(z_k)\}_{k\geq 0}$ converges, and thus, the above inequality ensures

$$\lim_{k\to\infty} \|z_{k+1} - z_k\|^2 = 0. \tag{20}$$

Since $\{z_k\}_{k\geq 0}$ is bounded by assumption, there exists a limit point $\bar{z} \in \Omega$ of $\{z_k\}_{k\geq 0}$, i.e., there exists a subset K of the set of nonnegative integers such that $\lim_{k\to\infty,\, k\in K} z_k = \bar{z}$. By (20), it also follows that $\lim_{k\to\infty,\, k\in K} z_{k+1} = z^*$. Using the definition of $\mathcal{H}[\cdot;\, z_k,\, S_k,\, S_{k+1}]$ in (13) and the feasibility of z_{k+1} in (11), we have

$$\begin{aligned}\mathcal{F}[z_k] &+ D_z\mathcal{F}[z]|_{z=z_k}[z_{k+1}-z_k] \\ &\preceq \mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] \\ &\preceq 0.\end{aligned}$$

By taking the limit $k\to\infty$, $k\in K$ in the above inequality, using the continuity of $\mathcal{F}[\cdot]$, $\mathcal{H}[\cdot;\,\cdot,\, S_k,\, S_{k+1}]$, and since the positive semidefinite cone is closed, we have $\mathcal{F}[\bar{z}]\preceq 0$.

Therefore, every limit point $\bar{z}$ of $\{z_k\}_{k\geq 0}$ is feasible. By assumption, $\bar{z}\in\Omega$ satisfies the MFCQ, i.e., there exists $\xi\in\Omega$ such that

$$\mathcal{F}[\bar{z}] + D_z\mathcal{F}[z]|_{z=\bar{z}}[\xi-\bar{z}] \prec 0.$$

Therefore, there exists $k_0\geq 0$ such that

$$\mathcal{F}[z_k] + D_z\mathcal{F}[z]|_{z=z_k}[\xi-z_k] \prec 0, \quad \forall k\geq k_0,\, k\in K.$$

On the other hand, by using the statements (2) and (3) of Lemma 4 and (20), we have

$$\begin{aligned}\lim_{k\to\infty,\, k\in K} &(\mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] \\ &+ D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}[\xi-z_{k+1}]) \\ &= \mathcal{F}[\bar{z}] + D_z\mathcal{F}[z]|_{z=\bar{z}}[\xi-\bar{z}].\end{aligned}$$

Therefore, there exists $k_1\geq k_0$ such that

$$\begin{aligned}&\mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] \\ &\quad + D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}[\xi-z_{k+1}] \prec 0, \\ &\quad \forall k\geq k_1,\, k\in K.\end{aligned}$$

This implies that for all $k\geq k_1$, $k\in K$, the MFCQ holds at z_{k+1} for the optimization problem (11). By Lemma 5, there exists Λ_{k+1} solving the KKT (15)-(17) for all $k\geq k_1$, $k\in K$.

Next, we will prove that $\{\Lambda_{k+1}\}_{k\geq k_1,\, k\in K}$ is bounded. Assume by contradiction that $\{\Lambda_{k+1}\}_{k\geq k_1,\, k\in K}$ is not bounded. If we define $\Phi_{k+1} := \Lambda_{k+1}/\|\Lambda_{k+1}\|$, then $\{\Phi_{k+1}\}_{k\geq k_1,\, k\in K}$ is bounded and, from the KKT condition (15), satisfies

$$\begin{aligned}0 \in &\frac{\nabla f(z_{k+1}) + \rho(z_{k+1}-z_k)}{\|\Lambda_{k+1}\|} \\ &+ \left[D_z\mathcal{H}[z;\, z_k,\, S_k,\, S_{k+1}]|_{z=z_{k+1}}\right]^*\Phi_{k+1} \\ &+ \mathcal{N}_\Omega(z_{k+1}),\end{aligned} \tag{21}$$

and there exists a subset $R\subseteq K$ such that

$$\lim_{k\to\infty,\, k\in R}\Phi_{k+1} = \bar{\Phi}, \quad \lim_{k\to\infty,\, k\in R} z_k = \bar{z}, \quad \lim_{k\to\infty,\, k\in R} z_{k+1} = \bar{z}.$$

By taking the limit $k\to\infty$, $k\in R$ in (21), we have

$$0 \in [D_z\mathcal{F}(z)|_{z=\bar{z}}]^*\bar{\Phi} + \mathcal{N}_\Omega(\bar{z}), \tag{22}$$

which, by the definition of the normal cone $\mathcal{N}_\Omega(\bar{z})$ at $\bar{z}$, implies

$$\bar{\Phi}\circ(D_z\mathcal{F}[z]|_{z=\bar{z}}(y-\bar{z})) \geq 0, \quad \forall y\in\Omega. \tag{23}$$

By assumption, every feasible point of (2) satisfies the MFCQ, so there exists $\xi\in\Omega$ such that $D_z\mathcal{F}[z]|_{z=\bar{z}}[\xi-\bar{z}]\prec 0$, implying that for any $\bar{\Phi}\succeq 0$, $\bar{\Phi}\circ(D_z\mathcal{F}[z]|_{z=\bar{z}}[\xi-\bar{z}]) < 0$. However, this contradicts with (23). Therefore, $\{\Lambda_{k+1}\}_{k\geq k_1,\, k\in K}$ is bounded, and there exists $R\subseteq K$ such that $\lim_{k\to\infty,\, k\geq k_1,\, k\in R}\Lambda_{k+1} = \bar{\Lambda}$. Therefore, it can be proved that

$$\begin{aligned}\lim_{k\to\infty,\, k\geq k_1 k\in R} &\Lambda_{k+1}\circ\mathcal{H}[z_{k+1};\, z_k,\, S_k,\, S_{k+1}] \\ &= \bar{\Lambda}\circ\mathcal{F}[\bar{z}] = 0\end{aligned} \tag{24}$$

and

$$0 \in \nabla f(\bar{z}) + [D_z\mathcal{F}[z]|_{z=\bar{z}}]^*\bar{\Lambda} + \mathcal{N}_\Omega(\bar{z}),$$

where the equality in (24) is derived from the statement 2) of Lemma 4. Therefore, the limit point $(\bar{z},\, \bar{\Lambda})$ is a KKT point. This completes the proof. □

Remark 3 *The proofs of Proposition 4 and Proposition 5 are based on the proofs of [36, Corollary 2.3] and [36, Proposition 3.2], where non-convex optimization problems are solved by the SPCA of the non-convex inequality constraints. In particular, [36, Proposition 3.2] uses the linear independence constraint qualification (LICQ) [30, Definition 12.4] for the convergence, while in Proposition 4 of this paper, the MFCQ is used for the proof of the convergence. To prove the boundedness of the Lagrange multipliers, we apply the idea of the proof of [37, Theorem 11] and the proof of [32, proposition 3.3.5] to prove the Fritz John necessary condition for the optimality of the general nonlinear programming. In the proof of Proposition 5, we used the non-existence of a non-zero* $\bar{\Phi}\succeq 0$ *such that* (22) *holds to prove the boundedness of the sequence of Lagrange multipliers. The non-existence condition is called the pseudonormality in [32], which is a generalization of the constraint* *qualifications and ensures the existence of the Lagrange multipliers. Finally, proofs similar to the proof of Proposition 5 also appear in [18, Theorem 2].*

Remark 4 *If Assumption 2 does not hold, i.e., f is not bounded from below on $\mathcal{D}$, then this means that in some cases, $\lim_{k\to\infty} f(x_k) = -\infty$. In general, the solution with a too large $|f(x_k)|$ may not be useful for practical purposes. Similarly, if $\{x_k\}_{k\geq 0}$ is unbounded, the final solution chosen from the sequence will be meaningless. In these cases, we can modify the objective function f and the set $\mathcal{D}$ to guarantee the boundedness of $\{x_k\}_{k\geq 0}$ and $\{f(x_k)\}_{k\geq 0}$. For instance, the constraint $\|x\|_\infty \leq b$ with a real number $b > 0$ can be included in $\mathcal{D}$ to ensure that $\{x_k\}_{k\geq 0}$ is bounded, where $\|\cdot\|_\infty$ is the ∞-norm.*

5 Example

All numerical examples were solved by MATLAB R2008a running on a Windows 10 PC with Intel Core i5-4210 2.6G Hz CPU, 4 GB RAM. The convex optimization problems were solved with SeDuMi [38] and Yalmip [39]. In this section, the static output feedback (SOF) controller design problem [3] will be considered to illustrate the proposed method. The spectral abscissa optimization problem for the SOF controller design, addressed also in [19, page 1383], is a well-known BMIP of the form (2) in the systems and control literature, e.g., [11, 15, 38]. Consider the continuous-time linear time-invariant system $\dot{x}(t) = Ax(t) + Bu(t)$, $y(t) = Cx(t)$, $x(0) = \mathbb{R}^{n_x}$, where $(A, B, C) \in \mathbb{R}^{n_x\times n_x} \times \mathbb{R}^{n_x\times n_u} \times \mathbb{R}^{n_y\times n_x}$, $t \geq 0$, $x(t) \in \mathbb{R}^{n_x}$ is the state, $u(t) \in \mathbb{R}^{n_u}$ is the control input, and $y(t) \in \mathbb{R}^{n_y}$ is the measured output. The goal is to find the gain matrix $F \in \mathbb{R}^{n_u\times n_y}$ so that the system can be stabilized by the SOF controller $u(t) = Fy(t)$. Specifically, the design problem can be formulated as the non-convex optimization problem

$$\inf_{F\in\mathbb{R}^{n_u\times n_y}} f(F),$$

where $f(F) := \max\{\mathrm{Re}\{\lambda\} : \lambda \in \Lambda(A+BFC)\}$, $\mathrm{Re}\{\lambda\}$ denotes the real part of $\lambda \in \mathbb{C}$, and $\Lambda(A+BFC)$ is the spectrum of $A+BFC$. The problem can be expressed as the BMIP (2)

$$\begin{aligned}&\inf_{P,\,F,\,\alpha} \alpha \quad \text{s.t.} \\ &(A+BFC)^T P + P(A+BFC) - 2\alpha P \preceq 0, \\ &(P,\,F,\,\alpha) \in \Omega,\end{aligned} \tag{25}$$

where

$$\begin{aligned}\Omega = \{(P,\,F,\,\alpha) \in \mathbb{R}^{n_x\times n_x} \times \mathbb{R}^{n_x\times n_u} \times \mathbb{R} :& \\ P = P^T,\, P - 10^{-6} I \succeq 0\}.&\end{aligned} \tag{26}$$

To apply Algorithm 1, we replace $(F,\,P,\,\alpha)$ in (25) with $(F_k + \Delta F,\, P_k + \Delta P,\, \alpha_k + \Delta\alpha)$ to have

$$\begin{aligned}&\mathrm{He}\{(P_k + \Delta P)(A + B(F_k + \Delta F)C)\} \\ &\quad - 2(\alpha_k + \Delta\alpha)(P_k + \Delta P) \preceq 0,\end{aligned}$$

where $\Delta F = F - F_k$, $\Delta P = P - P_k$, and $\Delta\alpha = \alpha - \alpha_k$. By expanding and rearranging the terms in the last matrix inequality, we obtain

$$\begin{aligned}&\mathrm{He}\{PA + PBF_kC + P_kB\Delta FC\} \\ &\quad - 2(\alpha P_k + \alpha_k \Delta P) \\ &\quad + \mathrm{He}\{\Delta P(B\Delta FC - \Delta\alpha I_n)\} \preceq 0.\end{aligned} \tag{27}$$

Noting that the last term in the left-hand side of the above inequality is bilinear, the LMI constraint (10) is obtained as (28) at the top of page 12. Then, the subproblem (9) of Algorithm 1 is obtained as follows:

$$\begin{aligned}&(P_{k+1},\, F_{k+1},\, \alpha_{k+1},\, S_{k+1}) \\ &:= \mathop{\arg\inf}_{(P,\,F,\,\alpha)\in\Omega,\, S\in\mathbb{S}^n} \Big\{\alpha + 0.005\,\|\Delta P\|_F^2 \\ &\qquad\qquad + 0.005\,\|\Delta F\|_F^2 + 0.005\,\|\Delta\alpha\|_F^2\Big\} \\ &\text{s.t.}\quad \text{LMI (28) holds, and} \\ &10^{-6} I \preceq S \preceq 10^4 I, \quad -2S_k + S \preceq -10^{-6} I,\end{aligned}$$

where $\|\cdot\|_F$ is the Frobenius norm.

There are several remarks on the implementation of Algorithm 1.

(1) We use the same method of [19] to determine an initial feasible point of (25), i.e., $F_0 = 0$, $\alpha_0 = -0.5f(0)$, and P_0 is chosen as a solution to the LMIs $P_0 \succeq 10^{-6} I$, $A^T P_0 + P_0 A + 2\alpha_0 P_0 \preceq 0$;
(2) Stopping criterion: As in [19], Algorithm 1 is terminated if one of the following conditions is satisfied:
  (a) $\|x_{k+1} - x_k\|_\infty / (\|x_k\|_\infty + 1) \leq 10^{-3}$;
  (b) the maximum number of iterations, $K_{\max} = 100$, reaches;
  (c) the objective function is not significantly improved after two successive iterations, i.e., $|f(F_k) - f(F_{k-1})| \leq 10^{-4}(1 + |f(F_{k-1})|)$ and $|f(F_{k+1}) - f(F_k)| \leq 10^{-4}(1 + |f(F_k)|)$.
(3) Ω is convex and closed;
(4) The MFCQ is satisfied for every feasible point $(P_k,\, F_k,\, \alpha_k) \in \Omega$ of (25). To prove this, note that the MFCQ at $(P_k,\, F_k,\, \alpha_k)$ is the existence of $(P,\, F,\, \alpha) \in \Omega$ such that (27) is satisfied. If $(P_k,\, F_k,\, \alpha_k) \in \Omega$ is feasible for (25), i.e., $(A + BF_kC)^T P_k + P_k(A + BF_kC) - 2\alpha_k P_k = \mathrm{He}\{P_kA + P_kBF_kC\} - 2\alpha_k P_k \preceq 0$, then it can be proved that $P = P_k$, $F = F_k$, and any $\alpha_k < \alpha$ satisfy the condition for the MFCQ.

$$
\begin{bmatrix} \text{He}\{PA + PBF_kC + P_kB\Delta FC\} - 2(\alpha P_k + \alpha_k\Delta P) & * & * \\ S_k\Delta P & S - 2\text{S}_k & * \\ B\Delta FC - \Delta\alpha I_n & 0 & -S \end{bmatrix} \preceq 0, \tag{28}
$$

In this section, Algorithm 1 is compared with the DC programming [19], the LMIRank [13] (a MATLAB toolbox for solving rank constrained LMI feasibility problems), and the HIFOO [9] (an open-source Matlab package for structured controller design). In addition, we use the system data (A, B, C) extracted from the publicly available database COMPleib library [40]. We set the initial controller for the HIFOO, LMIRank, and DC programming to that used by Algorithm 1. For the HIFOO and LMIRank, the default options are used. Since LMIRank can only deal with the feasibility problem of the rank constrained LMIs, a simple bisection algorithm is applied to solve the optimization problem.

The comparison results of the optimal objective function value f^* of (25) computed using different approaches are summarized in Table 1 with corresponding computational times, where for each system data (A, B, C), the smallest estimated optimal objective function value is highlighted in blue.

The results suggest that the computational time of Algorithm 1 is comparable with that of the DC programming, whereas its performance is in general better than that of the DC programming in terms of the estimated optimal objective function value.

Moreover, the LMIRank outperforms Algorithm 1 in terms of the estimated optimal objective function value. However, the computational time of the LMIRank is higher than that of Algorithm 1.

In addition, one can observe that the HIFOO is computationally less demanding than Algorithm 1, while Algorithm 1 can be applied to more general BMIPs of the form (2).

6 Conclusion

In this paper, we have proposed a sequential parametric convex approximation method for solving optimization problems including BMIs as constraints. It has been proved that under the MFCQ, the iterative solutions of the proposed algorithm converges to a stationary point. A numerical experiment has demonstrated the validity and potential benefits of the proposed method.

References

[1] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[2] Z. Han and K. Liu, “Noncooperative power-control game and throughput game over wireless networks,” *IEEE Transactions on Communications*, vol. 53, no. 10, pp. 1625–1629, 2005.

[3] V. L. Syrmos, C. T. Abdallah, P. Dorato, and K. Grigoriadis, “Static output feedback–a survey,” *Automatica*, vol. 33, no. 2, pp. 125–137, 1997.

[4] T. Hu, L. Ma, and Z. Lin, “Stabilization of switched systems via composite quadratic functions,” *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2571–2585, 2008.

[5] P. Apkarian and H. D. Tuan, “Robust control via concave minimization local and global algorithms,” *IEEE Transactions on Automatic Control*, vol. 45, no. 2, pp. 299–305, 2000.

[6] G. Zhai, M. Ikeda, and Y. Fujisaki, “Decentralized h_∞ controller design: a matrix inequality approach using a homotopy method,” *Automatica*, vol. 37, no. 4, pp. 565–572, 2001.

[7] B. Fares, P. Apkarian, and D. Noll, “An augmented Lagrangian method for a class of LMI-constrained problems in robust control theory,” *International Journal of Control*, vol. 74, no. 4, pp. 348–360, 2001.

[8] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 2014.

[9] S. Gumussoy, D. Henrion, M. Millstone, and M. L. Overton, “Multiobjective robust control with HIFOO 2.0,” *arXiv preprint arXiv:0905.3229*, 2009.

[10] A. S. Lewis and M. L. Overton, “Nonsmooth optimization via BFGS,” *SIAM Journal on Optimization and Control*, 2009.

[11] A. Hassibi, J. How, and S. Boyd, “A path-following method for solving BMI problems in control,” in *Proceedings of the American control conference, USA*, 1999, pp. 1385 – 1389.

[12] K.-C. Goh, M. G. Safonov, and G. P. Papavassilopoulos, “Global optimization for the biaffine matrix inequality problem,” *Journal of Global Optimization*, vol. 7, no. 4, pp. 365–380, 1995.

[13] R. Orsi, U. Helmke, and J. B. Moore, “A Newton-like method for solving rank constrained linear matrix inequalities,” *Automatica*, vol. 42, no. 11, pp. 1875–1882, 2006.

[14] K. M. Grigoriadis and R. E. Skelton, “Low-order control design for LMI problems using alternating projection methods,” *Automatica*, vol. 32, no. 8, pp. 1117–1125, 1996.

[15] L. El Ghaoui, F. Oustry, and M. AitRami, “A cone complementarity linearization algorithm for static output-feedback and related problems,” *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1171–1176, 1997.

[16] P. D. Tao and L. T. H. An, “A DC optimization algorithm for solving the trust-region subproblem,” *SIAM Journal on Optimization*, vol. 8, no. 2, pp. 476–505, 1998.

[17] P. D. Tao *et al.*, “The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems,” *Annals of Operations Research*, vol. 133, no. 1-4, pp. 23–46, 2005.

Table 1
Comparison of the values of the optimal objective function value f^* computed using Algorithm 1 and previous methods, where the smallest estimated optimal objective function value is highlighted in blue.

System	Algorithm 1		DC programming		LMIRANK		HIFOO	
	f^*	Time (s)	f^*	Time (s)	f^*	Time (s)	f^*	Time (s)
AC1	−1.9919	50.7940 s	−1.5438	32.9560 s	−11.0496	110.8050 s	−0.2064	1.1470 s
AC4	−0.0500	1.8490 s	−0.0500	1.4040 s	−0.0500	50.8950 s	−0.0500	0.3440 s
AC7	−0.0849	79.7370 s	−0.0500	29.8130 s	−0.0672	147.1430 s	−0.0322	0.4830 s
AC8	−0.4447	13.8970 s	−0.0689	10.1060 s	−0.4447	48.7270 s	−0.2107	1.2430 s
AC9	−0.4447	15.1220 s	−0.4049	61.1700 s	−0.3162	352.8110 s	−0.3951	13.3170 s
AC11	−5.5002	70.1070 s	−4.1994	58.6150 s	−4.7986	59.0010 s	−0.0003	0.3560 s
AC12	−0.0223	6.9120 s	−0.2216	33.4670 s	−113.0645	6.4250 s	−51.0269	12.4310 s
HE1	−0.2367	18.7450 s	−0.2135	33.5940 s	−0.2446	39.4360 s	−0.2457	0.5270 s
HE3	−1.2975	75.5660 s	−0.9227	42.5530 s	−2.2987	177.2510 s	−0.4716	13.0530 s
HE4	−1.3950	61.3190 s	−0.8708	41.9160 s	−1.9579	157.7620 s	−0.7338	12.7710 s
HE5	−0.0239	52.6800 s	−0.0258	38.7720 s	−0.0153	83.9090 s	−0.1830	6.5210 s
HE6	−0.0050	18.9580 s	−0.0050	10.8720 s	−0.0050	1375.9890 s	−0.0050	0.3130 s
REA1	−13.7028	53.5490 s	−3.9122	38.1450 s	−54.7418	29.6310 s	−14.8759	7.3510 s
REA2	−4.3550	21.1190 s	−2.0027	10.7710 s	−16.5533	65.1220 s	−7.0165	5.8410 s
REA3	−0.0207	2.4980 s	−0.0207	1.4900 s	−0.0207	64.8240 s	−0.0207	0.7730 s
DIS4	−9.9515	68.5790 s	−8.2552	34.0910 s	−122.2719	6.1970 s	−83.7923	0.6950 s
NN1	−3.3130	28.6730 s	−0.1354	35.3070 s	−4.3030	25.1580 s	−5.9101	1.5990 s
NN13	−2.9847	16.1450 s	−2.9594	15.2780 s	−4.4809	60.6940 s	−3.2460	3.7650 s

[18] T. D. Quoc and M. Diehl, "Sequential convex programming methods for solving nonlinear optimization problems with DC constraints," *arXiv preprint arXiv:1107.5841*, 2011.

[19] Q. T. Dinh, S. Gumussoy, W. Michiels, and M. Diehl, "Combining convex–concave decompositions and linearization approaches for solving BMIs, with application to static output feedback," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1377–1390, 2012.

[20] A. Shapiro, "First- and second-order analysis of nonlinear semidefinite programs," *Mathematical Programming*, vol. 77, no. 1, pp. 301–320, 1997.

[21] E. Polak and Y. Wardi, "Nondifferentiable optimization algorithm for designing control systems having singular value inequalities," *Automatica*, vol. 18, no. 3, pp. 267–283, 1982.

[22] B. Fares, D. Noll, and P. Apkarian, "Robust control via sequential semidefinite programming," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1791–1820, 2002.

[23] R. W. Freund, F. Jarre, and C. H. Vogelbusch, "Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling," *Mathematical Programming*, vol. 109, no. 2-3, pp. 581–611, 2007.

[24] R. Correa, "A global algorithm for nonlinear semidefinite programming," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 303–318, 2004.

[25] M. Kočvara and M. Stingl, "PENNON: A code for convex nonlinear and semidefinite programming," *Optimization methods and software*, vol. 18, no. 3, pp. 317–333, 2003.

[26] D. Henrion, J. Löfberg, M. Kočvara, and M. Stingl, "Solving polynomial static output feedback problems with PENBMI," in *Proceeding. joint IEEE Conf. Decision Control and Europ. Control Conf., Sevilla, Spain*, 2005, pp. 7581–7586.

[27] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.

[28] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, no. 1, pp. 29–51, 2010.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[30] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Science & Business Media, 2006.

[31] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, 2013.

[32] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[33] X. Li and C. E. De Souza, "Criteria for robust stability and stabilization of uncertain linear systems with state delay," *Automatica*, vol. 33, no. 9, pp. 1657–1662, 1997.

[34] Y.-Y. Cao, Y.-X. Sun, and C. Cheng, "Delay-dependent robust stabilization of uncertain systems with multiple state delays," *IEEE Transactions on Automatic Control*, vol. 43, no. 11, pp. 1608–1612, 1998.

[35] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.

[36] A. B.-T. A. Beck and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, pp. 29–51, 2010.

[37] G. Scutari, F. Facchinei, L. Lampariello, and P. Song, "Distributed methods for constrained nonconvex multi-agent pptimization-part I: theory," *arXiv preprint arXiv:1410.4754*, 2014.

[38] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[39] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD conference*, 2004, pp. 284–289.

[40] F. Leibfritz and W. Lipinski, "Description of the benchmark examples in COMPleib 1.0," *Department of Mathematics at University of Trier, Trier, Germany, Technical Report*, 2003.