

Article

A Lifting-Penalty Method for Quadratic Programming with a Quadratic Matrix Inequality Constraint

Wei Liu ^{1,2}, Li Yang ³ and Bo Yu ^{1,*}¹ School of Mathematical Sciences, Dalian University of Technology, Dalian 116025, China; liuwei@bnu.edu.cn² School of Applied Mathematics, Beijing Normal University, Zhuhai 519087, China³ School of Mathematics and Physics Science, Dalian University of Technology, Panjin 124221, China; yangli96@dlut.edu.cn

* Correspondence: yubo@dlut.edu.cn

Received: 23 December 2019; Accepted: 15 January 2020; Published: 21 January 2020



Abstract: In this paper, a lifting-penalty method for solving the quadratic programming with a quadratic matrix inequality constraint is proposed. Additional variables are introduced to represent the quadratic terms. The quadratic programming is reformulated as a minimization problem having a linear objective function, linear conic constraints and a quadratic equality constraint. A majorization–minimization method is used to solve instead a l_1 penalty reformulation of the minimization problem. The subproblems arising in the method can be solved by using the current semidefinite programming software packages. Global convergence of the method is proven under some suitable assumptions. Some examples and numerical results are given to show that the proposed method is feasible and efficient.

Keywords: penalty method; majorization–minimization method; quadratic matrix inequality

1. Introduction

In this paper, we consider the quadratic programming with a quadratic matrix inequality constraint of the following form:

$$\begin{aligned} \min_x \quad & x^T Q x + q^T x, \\ \text{s.t.} \quad & \sum_{i,j=1}^d x_i x_j A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00} \succeq 0, \\ & \mathcal{A}x - b \in K, \end{aligned} \quad (1)$$

where $q, x \in \mathbb{R}^n$, $Q, A_{ij} \in \mathbb{S}^m$, $\mathcal{A}$ is a linear operator, the cone K is a product of semidefinite cones, second-order cones, nonnegative orthants and Euclidean spaces, $\mathbb{S}^m$ is the space of $m \times m$ symmetric matrices, and $A \succeq 0$ indicates that A is positive semidefinite.

Results of [1–6] have shown that problem (1) has important applications in many areas, including control theory and robust optimization. Moreover, it contains a wide range of optimization problems, the well-known linear and quadratic programming problems [7–14] and bilinear matrix inequality (BMI) feasibility problem [15] are its special case. The quadratic programming [16,17] and the problem of checking the solvability of a general BMI [18] are NP-hard. The problem (1) is not so easy to be solved computationally. It can be categorized as a special case of more general nonlinear semidefinite programming (NSDP). The first- and second-order optimality conditions and numerical methods for NSDP problems have been developed; see, for example, [19–25].

Nonsmooth optimization methods suited for eigenvalue optimization were used to compute locally optimal solutions for the BMI problem in [26]. A linearized convex-concave semidefinite programming algorithm for solving the following problem:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & A(x) - B(x) \preceq 0, \\ & x \in \mathcal{X}, \end{aligned} \quad (2)$$

was proposed in [1], where $f(x)$ is convex, $\mathcal{X}$ is a nonempty and closed convex set, and $A(x)$ and $B(x)$ are positive semidefinite convex. Based on the fact that the bilinear form $X^T Y + Y^T X$ can be decomposed as a difference between two positive semidefinite convex mappings, the method was applied to BMI optimization formulations of the static state/output feedback controller design problems. Following the same line of the work in [1], an iterative procedure to search a local optimum of more general nonconvex problem was developed in [27], based on a positive semidefinite convex overestimate of a positive semidefinite nonconvex matrix mapping. These local methods may not be able to obtain a global optimum.

It is easy to see that the problem (1) is equivalent to rank one constrained optimization problem, by introducing additional variable y_{ij} that represents the quadratic term $x_i x_j$. Rank-constrained optimization problems are considered in [28,29]. In [30], polynomial-time checkable sufficient conditions, which guarantees that the semidefinite relaxations of quadratically constrained quadratic programs are exact, are given.

To attain global optimal solution, optimization approaches have been proposed. Based on the generalized Benders decomposition, a global approach is proposed for problems with BMI constraints in [31]. A slight general formulation of the BMI feasibility problem known as the following BMI eigenvalue problem (BMIEP):

$$\begin{aligned} \min_{x,y,\lambda} \quad & \lambda \\ \text{s.t.} \quad & \lambda I - A(x,y) \succeq 0, \\ & l \leq x \leq u, \end{aligned}$$

was dealt in [32], where $A(x,y) = \sum_{ij} x_i x_j A_{ij} + \sum_i x_i A_{i0} + \sum_j y_j A_{j0} + A_{00}$. They proposed robust Branch-and-Cut algorithms, which improves the first implemented branch-and-bound algorithm [33] for the BMIEP. Other branch-and-bound algorithms can refer to [34,35]. Though the approaches based on the generalized Benders decomposition and branch-and-bound algorithms are global methods, it is in general impractical to solve large-scale problems. A solution method of reduction of variables is proposed for BMI problems in system and control designs in [36]. The proposed method consists of a principle of variable classification, a procedure for problem transformation and a hybrid algorithm. The proposed method can address feasibility, single-objective, and multiobjective problems with BMI constraints. However, it can fail in two circumstances.

The aim of this paper is to present a lifting-penalty method for the solution of problem (1). We reformulate a quadratic matrix inequality constraint as linear matrix inequality constraints and a single quadratic equality constraint, but instead of a rank one constraint or a quadratic matrix equality constraint. Then, we use a majorization–minimization method to solve instead a l_1 penalty reformulation of the minimization problem. For the fixed penalty problem, global convergence to a Karush–Kuhn–Tucker (KKT) point of the majorization–minimization method is proven. The organization of this paper is as follows. In Section 2, a lifting-penalty method, which can obtain an ϵ -optimal solution of the problem (1), and its global convergence are given. In Section 3, numerical results are reported to show that the proposed method is efficient.

Throughout this paper, we use the notations in Table 1.

Table 1. Mathematical symbols and their meaning.

Symbol	Meaning
$\phi'(x)$	The derivative of ϕ with respect to x
$\nabla\phi(x)$	The gradient of ϕ at x
$\mathbb{S}_+^m$	The set of symmetric positive semidefinite matrices of dimension $m \times m$
$\mathbb{S}_{++}^m$	The set of symmetric positive definite matrices of dimension $m \times m$
I	The identity matrix of appropriate dimension
$A \succeq B$	$A - B$ belongs to $\mathbb{S}_+^m$
$A \succ B$	$A - B$ belongs to $\mathbb{S}_{++}^m$
$A \bullet B$	The inner product of A and B
$\text{tr}(A)$	The trace (sum of diagonal elements) of the matrix A
$\lambda_{\min}(A)$	The minimum eigenvalue of the matrix A
$N_{\Omega}(x)$	The normal cone to the set Ω at the point $x \in \Omega$

2. The Lifting-Penalty Method

This section concerns a lifting-penalty method for solving the problem (1). Its motivation is simple to describe: Rather than solving the original problem (1) directly, we reformulate it as a minimization problem having a linear objective function, linear conic constraints and a quadratic equality constraint by lifting x to a symmetric matrix Y , and then we use l_1 penalty method to solve the reformulation of (1). For a given value of the penalty parameter, the majorization–minimization (MM) method is used to solve the penalty problem. Global convergence of the MM method is proven under strict feasibility and boundedness of the original problem (1). An ϵ -optimal solution to problem (1) is obtained by solving the penalized problem as long as the penalty parameter is chosen appropriately.

Letting $Y = xx^T$, problem (1) can be stated equivalently as the following minimization problem:

$$\begin{aligned}
 \min_{x,Y} \quad & f(x, Y) = \text{tr}(QY) + q^T x \\
 \text{s.t.} \quad & \sum_{i,j=1}^d y_{ij} A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00} \succeq 0, \\
 & Y = xx^T, \\
 & \mathcal{A}x - b \in K.
 \end{aligned} \tag{3}$$

By Schur complement Theorem and the equivalence between the equality $Y = xx^T$ and the following system

$$\begin{aligned}
 Y - xx^T &\succeq 0, \\
 \text{tr}(Y - xx^T) &= \text{tr}(Y) - x^T x = 0,
 \end{aligned} \tag{4}$$

we can state problem (3) equivalently as follows:

$$\begin{aligned}
 \min_{x,Y} \quad & f(x, Y) \\
 \text{s.t.} \quad & \sum_{i,j=1}^d y_{ij} A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00} \succeq 0, \\
 & \begin{pmatrix} 1 & x^T \\ x & Y \end{pmatrix} \succeq 0, \\
 & \text{tr}(Y) - x^T x = 0, \\
 & \mathcal{A}x - b \in K.
 \end{aligned} \tag{5}$$

We assume throughout this paper that the feasible set of (1) is bounded. Hence, assume without loss of generality that $x^T x \leq \kappa/2$. By using a penalty function similar to the l_1 penalty function defined in [37], we consider the following penalized problem:

$$\begin{aligned}
\min_{x,Y} \quad & f_\mu(x,Y) = f(x,Y) + \mu|\text{tr}(Y) - x^T x| = f(x,Y) + \mu(\text{tr}(Y) - x^T x) \\
\text{s.t.} \quad & \sum_{i,j=1}^d y_{ij} A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00} \succeq 0, \\
& \begin{pmatrix} 1 & x^T \\ x & Y \end{pmatrix} \succeq 0, \quad \text{tr}(Y) \leq \kappa, \\
& \mathcal{A}x - b \in K,
\end{aligned} \tag{6}$$

where $\mu > 0$ is the penalty parameter. To simplify the notation, we denote $(b, -\kappa)$ and $K \times \mathbb{R}_+$ by $\tilde{b}$ and $\tilde{K}$, respectively, and define the linear operator $\tilde{\mathcal{A}}: (x, Y) \rightarrow (\mathcal{A}x, -\text{tr}(Y))$. Hence, we can rewrite constraints $\text{tr}(Y) \leq \kappa$ and $\mathcal{A}x - b \in K$ more compactly as $\tilde{\mathcal{A}}(x, Y) - \tilde{b} \in \tilde{K}$.

Let Ω denote the feasible set for problem (5). Unfortunately, strict feasibility of Ω (i.e., Slater's constraint qualification) fails. Therefore, we cannot establish similar results to the exactness (Theorem 4.4 in [37]) of the l_1 penalty function. However, when the penalty parameter μ is large enough, it can be proven that an ϵ -optimal solution to problem (1) can be obtained by solving the penalized problem as long as the penalty parameter is sufficiently large.

Let Ω_c denote the feasible set for (6), (x_f, Y_f) be a feasible solution of (5), (x_μ^*, Y_μ^*) and (x^*, Y^*) be optimal solutions of (6) and (5), respectively, and (x_c^*, Y_c^*) be an optimal solution of the following problem:

$$\min_{x,Y} \{f(x,Y) : (x,Y) \in \Omega_c\}. \tag{7}$$

Theorem 1. Let $\epsilon > 0$ be a given constant, and assume that $\mu \geq (f(x_f, Y_f) - f(x_c^*, Y_c^*)) / \epsilon$, then

$$\begin{aligned}
& \text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^* \leq \epsilon, \\
& f(x_\mu^*, Y_\mu^*) \leq f(x^*, Y^*) - \mu(\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^*) \leq f(x^*, Y^*).
\end{aligned}$$

Proof. By noting that (x_f, Y_f) is a feasible solution of (5), we have that $\text{tr}(Y_f) - x_f^T x_f = 0$. Therefore,

$$\begin{aligned}
f(x_f, Y_f) &= f(x_f, Y_f) + \mu(\text{tr}(Y_f) - x_f^T x_f) = f_\mu(x_f, Y_f) \\
&\geq f_\mu(x_\mu^*, Y_\mu^*) = f(x_\mu^*, Y_\mu^*) + \mu(\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^*) \\
&\geq f(x_c^*, Y_c^*) + \mu(\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^*).
\end{aligned} \tag{8}$$

It follows from (8) and the assumption on μ that

$$\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^* \leq (f(x_f, Y_f) - f(x_c^*, Y_c^*)) / \mu \leq \epsilon.$$

Because (x^*, Y^*) is an optimal solution of (5), by replacing (x_f, Y_f) in (8) by (x^*, Y^*) , we have that

$$f(x^*, Y^*) \geq f_\mu(x_\mu^*, Y_\mu^*) = f(x_\mu^*, Y_\mu^*) + \mu(\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^*).$$

Hence,

$$f(x_\mu^*, Y_\mu^*) \leq f(x^*, Y^*) - \mu(\text{tr}(Y_\mu^*) - x_\mu^{*T} x_\mu^*) \leq f(x^*, Y^*).$$

□

Based on Theorem 1, we give the following Algorithm 1 for finding an ϵ -optimal solution for problem (1).

Algorithm 1: The lifting-penalty method for solving (1)

- Step 1.** Choose an initial point $(x_f, Y_f) \in \Omega$, $\mu_1 > 0$, $s_1 > s_2 > 0$, $\epsilon_1 > 0$. Set $(x^1, Y^1) = (x_f, Y_f)$ and $j = 1$.
- Step 2.** Taking (x^j, Y^j) to be the initial point, compute an optimal solution (x^{j+1}, Y^{j+1}) of (6) with $\mu = \mu_j$.
- Step 3.** Set $\mu^{j+1} = \begin{cases} s_1 \mu_j, & \text{tr}(Y^j) - x^{jT} x^j > 0.1; \\ s_2 \mu_j, & \text{otherwise.} \end{cases}$
- Step 4.** If $\text{tr}(Y^{j+1}) - x^{(j+1)T} x^{j+1} \leq \epsilon_1$, stop; else, set $j = j + 1$, and go to Step 2.

In Step 1, to obtain a feasible point of the problem (5), we use the MM method to solve the following problem:

$$\begin{aligned} \min_{x, Y} \quad & \text{tr}(Y) - x^T x \\ \text{s.t.} \quad & (x, Y) \in \Omega_c. \end{aligned} \quad (9)$$

More detailed discussion on the MM method can be found in [38–41]. In Step 2, we use the MM method to solve the problem (6) with $\mu = \mu_j$. Let

$$\hat{f}_\mu(x, Y, z) = f(x, Y) + \mu \text{tr}(Y) - \mu z^T z - 2\mu z^T (x - z),$$

where $z \in \mathbb{R}^n$. From the definition of $\hat{f}_\mu(x, Y, z)$, we have that

$$\begin{aligned} \hat{f}_\mu(x, Y, z) &\geq f_\mu(x, Y), \text{ for any } z \in \mathbb{R}^n, \\ \hat{f}_\mu(x, Y, x) &= f_\mu(x, Y). \end{aligned}$$

That is, $\hat{f}_\mu(x, Y, z)$ is the majorization function of $f_\mu(x, Y)$ at z . Therefore, we can apply the following Algorithm 2 to the problem (6) for $\mu = \mu_j$.

Algorithm 2: The MM method for (6) with a fixed value of the parameter μ

- Step 1.** Choose an initial point $(x^1, Y^1) \in \Omega_c$, $\epsilon_2 > 0$; Set $k = 1$.
- Step 2.** Taking (x^k, Y^k) to be the initial point, compute an optimal solution (x^{k+1}, Y^{k+1}) of the following problem:

$$\begin{aligned} \min_{x, Y} \quad & \hat{f}_\mu(x, Y, x_k) \\ \text{s.t.} \quad & (x, Y) \in \Omega_c. \end{aligned} \quad (10)$$

- Step 3.** If $\|(x^{k+1}, Y^{k+1}) - (x^k, Y^k)\| \leq \epsilon_2$, stop; else, set $k = k + 1$, and go to Step 2.

In Algorithm 2, linear conic programming problem (10) can be solved by software packages SeDuMi [42], SDPT3 [43] or SDPNAL [44]. The MM method has global convergence properties, which states that the cluster points of $\{(x^k, Y^k)\}$ are all KKT points of (6) for any given μ . To prove it, we need a constraint qualification. We first give the following Lemma.

Lemma 1. *If strict feasibility of the problem (1) holds, then strict feasibility of the problem (6) is also ensured.*

Proof. If the problem (1) is strictly feasible, then there exists a point $\tilde{x} \in \mathbb{R}^n$ such that

$$\sum_{i,j=1}^d \tilde{x}_i \tilde{x}_j A_{ij} + \sum_{i=1}^n \tilde{x}_i A_{i0} + A_{00} \succ 0 \text{ and } \mathcal{A}\tilde{x} - b \in \text{ri } K, \quad (11)$$

where $\text{ri } K$ is the relative interior of the cone K .

Taking $\bar{Y} = \bar{x}\bar{x}^T$, we know that $(\bar{x}, \bar{Y})$ satisfies

$$\begin{aligned} G(\bar{x}, \bar{Y}) &= \sum_{i,j=1}^d \bar{y}_{ij} A_{ij} + \sum_{i=1}^n \bar{x}_i A_{i0} + A_{00} \succ 0, \\ \tilde{A}(\bar{x}, \bar{Y}) - \tilde{b} &\in \text{ri } \tilde{K}, \quad \begin{pmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{Y} \end{pmatrix} \succeq 0, \quad \begin{pmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{Y} \end{pmatrix} \not\succeq 0. \end{aligned} \quad (12)$$

Denote the vector of all eigenvalues of $Z \in \mathbb{S}^m$ arranged in non-increasing order by $\lambda(Z)$. By the definition of $G(\bar{x}, \bar{Y})$ and Mirsky's Theorem ([45], Cor 4.12), we have that for any

$$\varepsilon \in (0, \lambda_{\min}(G(\bar{x}, \bar{Y}))/2 \| \lambda(\sum_{i=1}^d A_{ii}) \|_{\infty}),$$

$$\begin{aligned} &\| \lambda(G(\bar{x}, \bar{Y} + \varepsilon I)) - \lambda(G(\bar{x}, \bar{Y})) \|_{\infty} = \| \lambda(G(\bar{x}, \bar{Y}) + \varepsilon \sum_{i=1}^d A_{ii}) - \lambda(G(\bar{x}, \bar{Y})) \|_{\infty} \\ &\leq \| \lambda(\varepsilon \sum_{i=1}^d A_{ii}) \|_{\infty} = \varepsilon \| \lambda(\sum_{i=1}^d A_{ii}) \|_{\infty} \leq \lambda_{\min}(G(\bar{x}, \bar{Y}))/2. \end{aligned}$$

Therefore, $\lambda(G(\bar{x}, \bar{Y} + \varepsilon I)) > 0$. Take $\tilde{Y} = \bar{Y} + \varepsilon I$, we have that

$$G(\bar{x}, \tilde{Y}) \succ 0, \quad \tilde{Y} - \bar{x}\bar{x}^T = \varepsilon I \succ 0. \quad (13)$$

From Schur complement Theorem and (13), we know that $(\bar{x}, \tilde{Y})$ is a strictly feasible point for (6). $\square$

Lemma 2. *If the feasible set of the problem (1) is nonempty and bounded, then, for any given $\mu > 0$, Ω_c and the solution sets of the problems (6) and (10) are nonempty and compact.*

Proof. Because the feasible set of the problem (1) is a nonempty subset of Ω_c , Ω_c is nonempty. It is trivial to show that Ω_c is closed. We will prove Ω_c is bounded. For any $(x, Y) \in \Omega_c$, $Y \succeq 0$ together with $\text{tr}(Y) \leq \kappa$ implies that Y is bounded, and $x^T x \leq \text{tr}(Y) \leq \kappa$ implies that x is bounded. As a result, Ω_c is nonempty and compact.

From the continuity of $f_{\mu}(\cdot)$ and $\hat{f}_{\mu}(\cdot)$, we know that the solution sets of the problems (6) and (10) are also nonempty and compact. $\square$

By introducing the product topology on $\mathbb{R}^n \times \mathbb{S}^m$ with the induced inner product

$$\langle (x_1, Y_1), (x_2, Y_2) \rangle = \langle x_1, x_2 \rangle + \langle Y_1, Y_2 \rangle = x_1^T x_2 + \text{tr}(Y_1 Y_2), \quad (14)$$

for any fixed z and any direction $(d, D) \in \mathbb{R}^n \times \mathbb{S}^m$, we calculate the derivative of $\hat{f}_{\mu}(x, Y, z)$ with respect to (x, Y) as follows:

$$\hat{f}'_{\mu}(x, Y, z)(d, D) = \langle q - 2\mu z, d \rangle + \langle Q + \mu I, D \rangle. \quad (15)$$

The gradient $\nabla \hat{f}_{\mu}(x, Y, z)$ of $\hat{f}_{\mu}(x, Y, z)$ with respect to (x, Y) , therefore, may be interpreted as the pair of matrices:

$$\nabla \hat{f}_{\mu}(x, Y, z) = (q - 2\mu z, Q + \mu I).$$

Theorem 2. *Suppose that Slater's constraint qualification for problem (1) holds, and the feasible set of (1) is bounded. For any given $\mu > 0$, let $\{(x^k, Y^k)\}$ be the sequence generated by Algorithm 2, then $\lim_{k \rightarrow \infty} f_{\mu}(x^k, Y^k) = \inf_k f_{\mu}(x^k, Y^k)$, and any cluster point of $\{(x^k, Y^k)\}$ is a KKT point of (6).*

Proof. From the boundedness of the feasible set of problem (1) and Lemma 2, we know that problem (10) has a nonempty and compact solution set and there exists a point $(x^{k+1}, Y^{k+1}) \in \arg \min_{(x, Y) \in \Omega_c} \{\hat{f}_{\mu}(x, Y, x^k)\}$.

By noting that $f_\mu(\cdot)$ is continuous and $\{(x^k, Y^k)\}$ is bounded, we know that $\inf_k \{f_\mu(x^k, Y^k)\}$ is finite. It follows from definitions of $f_\mu(x, Y)$ and $\hat{f}_\mu(x, Y, z)$ that

$$f_\mu(x^{k+1}, Y^{k+1}) \leq \hat{f}_\mu(x^{k+1}, Y^{k+1}, x^k) \leq \hat{f}_\mu(x^k, Y^k, x^k) = f_\mu(x^k, Y^k).$$

That is, $\{f_\mu(x^k, Y^k)\}$ is a monotonically decreasing sequence. Hence,

$$\lim_{k \rightarrow \infty} f_\mu(x^k, Y^k) = \inf_k f_\mu(x^k, Y^k).$$

Because (x^{k+1}, Y^{k+1}) is an optimal solution of (10), we have that

$$0 \in (q - 2\mu x^k, Q + \mu I) + N_{\Omega_c}(x^{k+1}, Y^{k+1}). \quad (16)$$

We first consider the case $(x^{k+1}, Y^{k+1}) = (x^k, Y^k)$ for some integer $k \geq 0$. When $(x^{k+1}, Y^{k+1}) = (x^k, Y^k)$, it follows from (16) that

$$0 \in (q - 2\mu x^{k+1}, Q + \mu I) + N_{\Omega_c}(x^{k+1}, Y^{k+1}). \quad (17)$$

Denote $G_1(x, Y) = \sum_{i,j=1}^d y_{ij} A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00}$, $G_2(x, Y) = \begin{pmatrix} 1 & x^T \\ x & Y \end{pmatrix}$, $G_3(x, Y) = \tilde{A}(x, Y) - \tilde{b}$ and $G_i^k = G_i(x^k, Y^k)$. From Lemma 1, we know that Slater's constraint qualification for (6) holds. This, together with convexity, implies that

$$N_{\Omega_c}(x^{k+1}, Y^{k+1}) = (G_1^{k+1})'^* N_{\mathbb{S}_+^m}(G_1^{k+1}) + (G_2^{k+1})'^* N_{\mathbb{S}_+^{n+1}}(G_2^{k+1}) + \tilde{A}^* N_{\tilde{K}}(G_3^{k+1}), \quad (18)$$

where $(G_i^{k+1})'^*$ and $\tilde{A}^*$ are the adjoint operators of $(G_i^{k+1})'$ and $\tilde{A}$, respectively. It follows from (17) and (18) that (x^{k+1}, Y^{k+1}) is a KKT point of problem (6).

Next, we assume that $(x^{k+1}, Y^{k+1}) \neq (x^k, Y^k)$ for all $k \geq 0$. Because $\{(x^k, Y^k)\}$ is bounded, we know that $\{(x^k, Y^k)\}$ has cluster points and any cluster point $(\bar{x}, \bar{Y}) \in \Omega_c$. By the equality $\lim_{k \rightarrow \infty} f_\mu(x^k, Y^k) = \inf_k f_\mu(x^k, Y^k)$, we have that $f_\mu(\bar{x}, \bar{Y}) = \inf_k f_\mu(x^k, Y^k)$. Now, we prove that $(\bar{x}, \bar{Y})$ is a KKT point of (6). For any $k \geq 0$, we have that

$$\begin{aligned} & f_\mu(x^{k+1}, Y^{k+1}) - f_\mu(x^k, Y^k) \\ &= \text{tr}((Q + \mu I)(Y^{k+1} - Y^k)) + q^T(x^{k+1} - x^k) - \mu(x^{k+1})^T x^{k+1} + \mu x^{kT} x^k \\ &= \text{tr}((Q + \mu I)(Y^{k+1} - Y^k)) + (q - 2\mu x^k)^T(x^{k+1} - x^k) + 2\mu x^{kT}(x^{k+1} - x^k) - \mu(x^{k+1})^T x^{k+1} + \mu x^{kT} x^k \\ &= \langle (q - 2\mu x^k, Q + \mu I), (x^{k+1} - x^k, Y^{k+1} - Y^k) \rangle - \mu(x^{k+1} - x^k)^T(x^{k+1} - x^k). \end{aligned} \quad (19)$$

It follows from (16) that $-(q - 2\mu x^k, Q + \mu I) \in N_{\Omega_c}(x^{k+1}, Y^{k+1})$, which implies that

$$\langle (q - 2\mu x^k, Q + \mu I), (x^{k+1} - x^k, Y^{k+1} - Y^k) \rangle = \langle -(q - 2\mu x^k, Q + \mu I), (x^k - x^{k+1}, Y^k - Y^{k+1}) \rangle \leq 0. \quad (20)$$

From (19) and (20), we have that

$$f_\mu(x^{k+1}, Y^{k+1}) - f_\mu(x^k, Y^k) \leq -\mu \|x^{k+1} - x^k\|^2, \quad (21)$$

which implies that

$$\mu \lim_{l \rightarrow \infty} \sum_{k=1}^l \|x^{k+1} - x^k\|^2 \leq f_\mu(x^1, Y^1) - \lim_{l \rightarrow \infty} f_\mu(x^{l+1}, Y^{l+1}) < +\infty.$$

Hence, $\lim_{k \rightarrow \infty} (x^{k+1} - x^k) = 0$. Let k_j satisfy that $x^{k_j} \rightarrow \bar{x}$ and $Y^{k_j} \rightarrow \bar{Y}$, then

$$\lim_{j \rightarrow \infty} x^{k_j} = \lim_{j \rightarrow \infty} x^{k_j-1} = \bar{x}. \quad (22)$$

Because (x^{k_j}, Y^{k_j}) is an optimal solution, there exists $D^{k_j} \in N_{\Omega_c}(x^{k_j}, Y^{k_j})$ such that

$$D^{k_j} = -(q - 2\mu x^{k_j-1}, Q + \mu I)$$

for any $k_j \geq 1$. Hence, for any $(\tilde{x}, \tilde{Y}) \in \Omega_c$,

$$\langle -(q - 2\mu x^{k_j-1}, Q + \mu I), (\tilde{x} - x^{k_j}, \tilde{Y} - Y^{k_j}) \rangle = \langle D^{k_j}, (\tilde{x} - x^{k_j}, \tilde{Y} - Y^{k_j}) \rangle \leq 0,$$

which, from the convergence of the subsequence $\{(x^{k_j}, Y^{k_j})\}$ and (22), gives rise to

$$\langle -(q - 2\mu \bar{x}, Q + \mu I), (\tilde{x} - \bar{x}, \tilde{Y} - \bar{Y}) \rangle \leq 0,$$

which implies that $-(q - 2\mu \bar{x}, Q + \mu I) \in N_{\Omega_c}(\bar{x}, \bar{Y})$. That is, $0 \in \nabla f_{\mu}(\bar{x}, \bar{Y}) + N_{\Omega_c}(\bar{x}, \bar{Y})$, which, together with Slater's constraint qualification, implies that $(\bar{x}, \bar{Y})$ is a KKT point of problem (6). $\square$

3. Numerical Experiments

In this section, the examples and some preliminary numerical results taken by the lifting-penalty method (LPM) and a modified augmented Lagrangian method (MALM) [46] are given below.

All numerical experiments are done by running MATLAB[®] 2016 on a notebook PC Intel[®] Core[™] i7-4810MQ CPU with 2.8 GHz and 16 GB RAM. The linear conic optimization problems in our method are solved by a SDPT3 solver. The optimization subproblems in MALM are solved by the subroutine fmincon. The parameters in the algorithms are set as follows:

$$\begin{aligned} \text{LPM : } & \mu_1 = 0.25, s_1 = 4, s_2 = 1.4, \epsilon_1 = \epsilon_2 = 10^{-6}; \\ \text{MALM : } & \tau = 0.25, \gamma = 10, c_1 = 1. \end{aligned}$$

Example 1 (HE1 in D in Section V, [3]). Find matrix K such that $A + BKC$ is Hurwitz, i.e., eigenvalues of matrix $A + BKC$ all belong to the left half plane $\mathcal{D} = \{t \in \mathbb{C} : t + t^* < 0\}$ of the complex plane, where t^* is the conjugate of t ,

$$A = \begin{pmatrix} -0.0366 & 0.0271 & 0.0188 & -0.4555 \\ 0.0482 & -1.01 & 0.0024 & -4.0208 \\ 0.1002 & 0.3681 & -0.7070 & 1.42 \\ 0 & 0 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 0.4422 & 0.1761 \\ 3.5446 & -7.5922 \\ -5.52 & 4.49 \\ 0 & 0 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}.$$

The problem amounts to solving a nonconvex feasibility problem $S(K) \succ 0$ (see [3,47,48]). We solved it by solving the following non-strict optimization problem:

$$\min_{K, \lambda} \lambda \quad \text{s.t. } \lambda I + S(K) \succeq 0, \quad -50 \leq K_{ij} \leq 50, \quad -1 \leq \lambda \leq 0.$$

From Figure 2 in [3], we know that the problem is nonconvex. For the starting point $K^{(0)} = (0, 0)^T$, our method ended with the final solution $K = (-0.4293, 0.5514)^T$.

Example 2. Let $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$,

$$\begin{aligned} \max_{x, \lambda} \quad & \lambda, \\ \text{s.t.} \quad & -\lambda I + \sum_{i,j=1}^d x_i x_j A_{ij} + \sum_{i=1}^n x_i A_{i0} + A_{00} \succeq 0, \\ & x_i \in [-1, 1], \quad i = 1, \dots, n, \end{aligned}$$

where $A_{ij} \in \mathbb{S}^m$ is generated by the Matlab function `sprandn(m,m,r)`. We symmetrized the matrices by copying the upper triangular part to the lower one after creation.

In our experiment, 20 problem instances are randomly generated for each value of (n, m, d) , and the additional variable Y is a $d \times d$ symmetric matrix, $r = 0.2$. Table 2 lists n, m, d and the average CPU time in seconds, respectively.

Table 2. Numerical results of Example 2.

			Time in Seconds			Time in Seconds			Time in Seconds	
			LPM	MALM		LPM	MALM		LPM	MALM
30	10	6	1.71	4.07	10	4.62	6.86	15	5.97	11.91
40			2.19	6.22		2.99	8.47		11.28	13.23
50			2.10	6.46		9.12	9.42		12.29	13.91
30	15	6	1.91	4.42	10	3.46	6.46	15	13.90	10.04
40			1.92	6.75		8.31	9.36		12.32	14.89
50			2.25	7.75		6.78	10.32		13.63	17.07
30	20	6	1.85	12.93	10	2.42	17.79	15	27.19	34.87
40			2.12	18.71		3.32	28.41		24.83	43.37
50			2.51	23.90		9.02	35.78		28.70	51.18
30	30	6	2.21	14.65	10	2.89	23.25	15	26.64	39.39
40			2.47	20.68		3.14	30.03		42.43	46.73
50			2.57	28.45		5.00	37.70		46.62	55.09

From the results listed in Table 2, we find that our method requires less time than the LPM for most instances. Moreover, consuming time seems to be more sensitive to d , compared with n and m . Preliminary computational experiences show that our method is competitive with the LPM.

4. Conclusions

In this paper, a lifting-penalty method for solving a quadratic optimization problem involving a quadratic matrix inequality constraint is introduced. By introducing additional variables, we reformulate a quadratic matrix inequality constraint as linear matrix inequality constraints and a single quadratic equality constraint but instead of a rank one constraint or a quadratic matrix equality constraint. Its global convergence result has been given under mild assumptions. Then, the method was applied to a feasibility problem and a problem of maximizing the smallest eigenvalue of a symmetric matrix. The numerical results show that the proposed method is competitive with the LPM. Note, however, that linear conic optimization subproblems arising in our method have the same feasible set, so the development of an efficient method for solving a family of linear conic optimization problems with the same feasible set is our future work.

Author Contributions: B.Y. supervised the research and helped W.L. at every step, especially framework building, analysis of the results, and writing of the manuscript. W.L. contributed the idea, framework building, implementation of the results, and writing of the manuscript. L.Y. helped with analyses of the introduction, results, and literature review. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the National Natural Science Foundation of China (11301050, 11171051, 11971092), and the Fundamental Research Funds for the Central Universities (DUT17RC(4)38).

Acknowledgments: The authors thank the anonymous referees, whose comments and suggestions led to an improved version of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dinh, Q.T.; Gumussoy, S.; Michiels, W.; Diehl, M. Combining Convex–Concave Decompositions and Linearization Approaches for Solving BMIs, with Application to Static Output Feedback. *IEEE Trans. Autom. Control* **2012**, *57*, 1377–1390. [\[CrossRef\]](#)
2. Goh, K.C.; Safonov, M.; Ly, J. Robust synthesis via bilinear matrix inequalities. *Int. J. Robust Nonlinear Control* **1996**, *6*, 1079–1095. [\[CrossRef\]](#)
3. Henrion, D.; Löfberg, J.; Kočvara, M.; Stingl, M. Solving polynomial static output feedback problems with PENBMI. In Proceedings of the 44th IEEE Conference on Decision and Control, Sevilla, Spain, 15 December 2005; Volume 1, pp. 7581–7586.
4. VanAntwerp, J.G.; Braatz, R.D. A tutorial on linear and bilinear matrix inequalities. *J. Process. Control* **2000**, *10*, 363–385. [\[CrossRef\]](#)
5. Lim, Y.; Oh, K.; Ahn, H. Stability and Stabilization of Fractional-Order Linear Systems Subject to Input Saturation. *IEEE Trans. Autom. Control* **2013**, *58*, 1062–1067. [\[CrossRef\]](#)
6. Fazelnia, G.; Madani, R.; Kalbat, A.; Lavaei, J. Convex Relaxation for Optimal Distributed Control Problems. *IEEE Trans. Autom. Control* **2017**, *62*, 206–221. [\[CrossRef\]](#)
7. Dikin, I. Iterative solutions of problems of linear and quadratic programming. *Sov. Math. Dokl.* **1967**, *8*, 674–675.
8. Ye, Y.; Tse, E. An extension of Karmarkar’s projective algorithm for convex quadratic programming. *Math. Program.* **1989**, *44*, 157–179, doi:10.1007/BF01587086. [\[CrossRef\]](#)
9. Ye, Y. On affine scaling algorithms for nonconvex quadratic programming. *Math. Program.* **1992**, *56*, 285–300. [\[CrossRef\]](#)
10. Zorkaltsev, V.; Mokryi, I. Interior Point Algorithms in Linear Optimization. *J. Appl. Ind. Math.* **2018**, *12*, 191–199. [\[CrossRef\]](#)
11. Anstreicher, K.; Wolkowicz, H. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.* **2000**, *22*, 41–55. [\[CrossRef\]](#)
12. Beck, A. Quadratic matrix programming. *SIAM J. Optim.* **2007**, *17*, 1224–1238. [\[CrossRef\]](#)
13. Ding, Y.; Ge, D.; Wolkowicz, H. On equivalence of semidefinite relaxations for quadratic matrix programming. *Math. Oper. Res.* **2011**, *36*, 88–104. [\[CrossRef\]](#)
14. Leibfritz, F.; Mostafa, E.M.E. An interior point constrained trust region method for a special class of nonlinear semidefinite programming problems. *SIAM J. Optim.* **2002**, *12*, 1048–1074. [\[CrossRef\]](#)
15. Mesbahi, M.; Papavassilopoulos, G.P. A cone programming approach to the bilinear matrix inequality problem and its geometry. *Math. Program.* **1997**, *77*, 247–272. [\[CrossRef\]](#)
16. Pardalos, P.M.; Vavasis, S.A. Quadratic programming with one negative eigenvalue is NP-hard. *J. Glob. Optim.* **1991**, *1*, 15–22. [\[CrossRef\]](#)
17. Sahni, S. Computationally related problems. *SIAM J. Comput.* **1974**, *3*, 262–279. [\[CrossRef\]](#)
18. Toker, O.; Özbay, H. On the NP-hardness of solving Bilinear Matrix Inequalities and simultaneous stabilization with static output feedback. In Proceedings of the 1995 American Control Conference (ACC’95), Seattle, WA, USA, 21–23 June 1995; pp. 2525–2526.
19. Fares, B.; Noll, D.; Apkarian, P. Robust control via sequential semidefinite programming. *SIAM J. Control Optim.* **2002**, *40*, 1791–1820. [\[CrossRef\]](#)
20. Gómez, W.; Ramírez, H. A filter algorithm for nonlinear semidefinite programming. *Comput. Appl. Math.* **2010**, *29*, 297–328.
21. Kanzow, C.; Nagel, C.; Kato, H.; Fukushima, M. Successive linearization methods for nonlinear semidefinite programs. *Comput. Optim. Appl.* **2005**, *31*, 251–273. [\[CrossRef\]](#)
22. Shapiro, A. First, and second order analysis of nonlinear semidefinite programs. *Math. Program.* **1997**, *77*, 301–320. [\[CrossRef\]](#)
23. Sun, D.; Sun, J.; Zhang, L. The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Math. Program.* **2008**, *114*, 349–391. [\[CrossRef\]](#)

24. Wolkowicz, H.; Saigal, R.; Vandenberghe, L. (Eds.) *Handbook of semidefinite programming*; International Series in Operations Research & Management Science, 27; Kluwer Academic Publishers: Boston, MA, USA, 2000.
25. Yamashita, H.; Yabe, H.; Harada, K. A primal-dual interior point method for nonlinear semidefinite programming. *Math. Program.* **2011**, *135*, 89–121. [[CrossRef](#)]
26. Thevenet, J.; Noll, D.; Apkarian, P. Nonsmooth methods for large bilinear matrix inequalities: Applications to feedback control. *Optim. Methods Softw.* **2005**, *200*, 1–24.
27. Tran Dinh, Q.; Michiels, W.; Gros, S.; Diehl, M. An inner convex approximation algorithm for BMI optimization and applications in control. In Proceedings of the 51st IEEE Conference on Decision and Control, Maui, HI, USA, 10–13 December 2012; pp. 3576–3581.
28. Kim, S.J.; Moon, Y.H. Structurally constrained and control: A rank-constrained LMI approach. *Automatica* **2006**, *42*, 1583–1588. [[CrossRef](#)]
29. Sun, C.; Dai, R. Rank-constrained optimization and its applications. *Automatica* **2017**, *82*, 128–136. [[CrossRef](#)]
30. Burer, S.; Ye, Y. Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs. *Math. Program.* **2019**. [[CrossRef](#)]
31. Beran, E.; Vandenberghe, L.; Boyd, S. A global BMI algorithm based on the generalized benders decomposition. In Proceedings of the 1997 European Control Conference (ECC), Brussels, Belgium, 1–7 July 1997; pp. 3741–3746.
32. Fukuda, M.; Kojima, M. Branch-and-Cut Algorithms for the Bilinear Matrix Inequality Eigenvalue Problem. *Comput. Optim. Appl.* **2001**, *19*, 79–105. [[CrossRef](#)]
33. Goh, K.C.; Safonov, M.G.; Papavassilopoulos, G.P. Global optimization for the biaffine matrix inequality problem. *J. Glob. Optim.* **1995**, *7*, 365–380. [[CrossRef](#)]
34. Hisaya, F.; Kohta, H. Bounds for the BMI Eigenvalue Problem. *Trans. Soc. Instrum. Control Eng.* **1997**, *33*, 616–621.
35. VanAntwerp, J.G.; Braatz, R.D.; Sahinidis, N.V. Globally optimal robust control for systems with time-varying nonlinear perturbations. *Comput. Chem. Eng.* **1997**, *21*, S125–S130. [[CrossRef](#)]
36. Chiu, W. Method of Reduction of Variables for Bilinear Matrix Inequality Problems in System and Control Designs. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 1241–1256. [[CrossRef](#)]
37. Han, S.P.; Mangasarian, O.L. Exact penalty functions in nonlinear programming. *Math. Program.* **1979**, *17*, 251–269. [[CrossRef](#)]
38. De Leeuw, J. Convergence of the majorization method for multidimensional scaling. *J. Classific.* **1988**, *5*, 163–180. [[CrossRef](#)]
39. De Leeuw, J.; Heiser, W.J. Convergence of correction matrix algorithms for multidimensional scaling. In *Geometric Representations of Relational Data*; Mathesis Press: Ann Arbor, MI, USA, 1977; pp. 735–752.
40. Figueiredo, M.A.; Bioucas-Dias, J.M.; Nowak, R.D. Majorization–minimization algorithms for wavelet-based image restoration. *IEEE Trans. Image Process.* **2007**, *16*, 2980–2991. [[CrossRef](#)] [[PubMed](#)]
41. Hunter, D.R.; Lange, K. A tutorial on MM algorithms. *Am. Stat.* **2004**, *58*, 30–37. [[CrossRef](#)]
42. Sturm, J.F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **1999**, *11/12*, 625–653. [[CrossRef](#)]
43. Tütüncü, R.H.; Toh, K.C.; Todd, M.J. Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Program.* **2003**, *95*, 189–217. [[CrossRef](#)]
44. Zhao, X.Y.; Sun, D.; Toh, K.C. A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM J. Optim.* **2010**, *20*, 1737–1765. [[CrossRef](#)]
45. Stewart, G.W.; Sun, J.G. *Matrix Perturbation Theory*; Academic Press: Cambridge, MA, USA, 1990.
46. Wu, H.; Luo, H.; Ding, X.; Chen, G. Global convergence of modified augmented Lagrangian methods for nonlinear semidefinite programming. *Comput. Optim. Appl.* **2013**, *56*, 531–558. [[CrossRef](#)]
47. Henrion, D.; Peaucelle, D.; Arzelier, D.; Šebek, M. Ellipsoidal approximation of the stability domain of a polynomial. *IEEE Trans. Automat. Control* **2003**, *48*, 2255–2259. [[CrossRef](#)]
48. Jury, E.I. *Inners and Stability of Dynamic Systems*; Wiley-Interscience (John Wiley & Sons): New York, NY, USA, 1974.

