An SQP-Type Algorithm for Nonlinear Second-Order Cone Programs [1]

Hirokazu Kato[2] and Masao Fukushima[2,3]

March 20, 2006

Abstract We propose an SQP-type algorithm for solving nonlinear second-order cone programming (NSOCP) problems. At every iteration, the algorithm solves a convex SOCP subproblem in which the constraints involve linear approximations of the constraint functions in the original problem and the objective function is a convex quadratic function. Those subproblems can be transformed into linear SOCP problems, for which efficient interior point solvers are available. We establish global convergence and local quadratic convergence of the algorithm under appropriate assumptions. We report numerical results to examine the effectiveness of the algorithm.

Key Words Nonlinear second-order cone programming problem, SQP method, global convergence, quadratic convergence

1 Introduction

Linear second-order cone programming (LSOCP) [1, 9, 10] and linear semidefinite programming (LSDP) [12, 14] have extensively been studied in the last decade, since they have desirable theoretical properties as well as many important applications [2, 9, 14]. For solving those problems, efficient interior point algorithms have been proposed and the software implementing those algorithms has been developed. On the other hand, nonlinear programming (NLP) has long been studied and a number of effective methods such as sequential quadratic programming methods (SQP) [3, 7] and interior point methods [15, 16] have been proposed. However, the study of nonlinear second-order cone programming (NSOCP) and nonlinear semidefinite programming (NSDP), which are natural extensions of LSOCP and LSDP, respectively, are much more recent and still in its preliminary phase. Optimality conditions for NSOCP and NSDP are studied in [4, 5]. An interior point method has been proposed for NSOCP in [17]. Globally convergent SQP-type method and successive linearization method have been developed for NSDP in [6] and [8], respectively.

In this paper, we propose an SQP-type algorithm for NSOCP. At every iteration, the algorithm solves a subproblem in which the constraints involve linear approximations of the constraints in the original problem and the objective function is a convex quadratic function. The subproblem can be transformed into an LSOCP problem, for which efficient interior point methods [13] are available. To ensure global convergence, the algorithm employs a line search strategy with the l_1-penalty function.

The organization of this paper is as follows: In Section 2, we formulate the nonlinear second-order cone programming problem. In Subsection 3.1, we describe the algorithm for NSOCP. In Subsection 3.2, we show global convergence of the algorithm. In Subsection 3.3, we consider the local convergence behavior of the algorithm. In Section 4, we report some numerical results.

---

[2]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan.

[3]Corresponding author. Email: fuku@i.kyoto-u.ac.jp

For vector $x \in \Re^n$, x_0 denotes the first component and $\bar{x}$ is the subvector consisting of the remaining components, that is, $x = \begin{pmatrix} x_0 \\ \bar{x} \end{pmatrix}$. The second-order cone of dimension n is defined by $K^n := \{x \in \Re^{n+1} \mid x_0 \geq \|\bar{x}\|\}$. For simplicity, $(x^T, y^T)^T$ is written as $(x, y)^T$. The Euclidean norm of vector x is denoted $\|x\| := \sqrt{x^T x}$.

2   Nonlinear Second-Order Cone Program

We are interested in the following nonlinear second-order cone program (NSOCP):

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \in K, \end{aligned} \tag{1}$$

where $f : \Re^n \to \Re$, $g : \Re^n \to \Re^m$ and $h : \Re^n \to \Re^l$ are twice continuously differentiable functions, K is the Cartesian product of second-order cones given by $K := K^{l_1} \times K^{l_2} \times \cdots \times K^{l_s}$, and $l := l_1 + \cdots + l_s$. Throughout the paper, we denote $h(x) = (h_1(x), \cdots, h_s(x))^T$ and $h_i(x) = (h_{i0}(x), \bar{h}_i(x))^T \in \Re^{l_i}$ $(i = 1, \cdots, s)$.

The Karush-Kuhn-Tucker (KKT) conditions for NSOCP(1) are given by

$$\begin{aligned} \nabla f(x^*) - \nabla g(x^*)\zeta^* - \nabla h(x^*)\eta^* &= 0, \\ g(x^*) &= 0, \\ h_i(x^*) \in K^{l_i}, \quad \eta_i^* &\in K^{l_i}, \\ h_i(x^*)^T \eta_i^* = 0, \quad i &= 1, \cdots, s, \end{aligned} \tag{2}$$

where $\zeta^* \in \Re^m$ and $\eta_i^* \in \Re^{l_i}(i = 1, \cdots, s)$ are Lagrange multiplier vectors. The KKT conditions are necessary optimality conditions under certain constraint qualifications [5]. We call a vector x^* a stationary point of problem (1) if there exist Lagrange multipliers ζ^* and η^* satisfying the KKT conditions (2). In this paper, we assume that there exists a triple (x^*, ζ^*, η^*) satisfying the KKT conditions (2) of problem (1). Such (x^*, ζ^*, η^*) is called a KKT triple of (1).

3   SQP-Type Algorithm for NSOCP

3.1   Algorithm

The algorithm solves the following subproblem at every iteration:

$$\begin{aligned} \min \quad & \nabla f(x^k)^T \Delta x + \tfrac{1}{2}\Delta x^T M_k \Delta x \\ \text{s.t.} \quad & g(x^k) + \nabla g(x^k)^T \Delta x = 0, \\ & h(x^k) + \nabla h(x^k)^T \Delta x \in K, \end{aligned} \tag{3}$$

where x^k is a current iterate and M_k is a symmetric positive definite matrix approximating the Hessian of Lagrangian function of problem (1) in some sense. The subproblem (3) is a convex programming problem. Therefore, under certain constraint qualifications, a vector Δx is an optimal solution of (3) if and only if there exist Lagrange multiplier vectors λ and μ satisfying the following KKT conditions for (3):

$$\nabla f(x^k) + M_k \Delta x - \nabla g(x^k)\lambda - \nabla h(x^k)\mu = 0,$$

$$g(x^k) + \nabla g(x^k)^T \Delta x = 0, \tag{4}$$
$$h_i(x^k) + \nabla h_i(x^k)^T \Delta x \in K^{l_i}, \quad \mu_i \in K^{l_i},$$
$$(h_i(x^k) + \nabla h_i(x^k)^T \Delta x)^T \mu_i = 0, \quad i = 1, \cdots, s.$$

Additionally, the subproblem (3) can be transformed into a linear second-order cone programming problem, for which efficient interior point methods are available [13].

Comparing conditions (2) and (4) yields the next proposition. The proof is straightforward and hence is omitted.

Proposition 1 *Under certain constraint qualifications, $\Delta x = 0$ is an optimal solution of subproblem (3) if and only if x^k is a stationary point of NSOCP (1) .*

This proposition allows us to deduce that the algorithm is globally convergent if $\{M_k\}$ is bounded and $\lim_{k \to \infty} \|\Delta x^k\| = 0$, where Δx^k is the solution of subproblem (3). Note that a subproblem (3) may be infeasible, even if the original NSOCP (1) is feasible. In SQP methods for nonlinear programming problems, some remedies to avoid this difficulty have been proposed [3]. In this paper, we simply assume that subproblem (3) is always feasible and hence has a unique optimal solution Δx^k.

The algorithm uses the l_1-penalty function as a merit function to determine step sizes:

$$P_\alpha(x) := f(x) + \alpha \Big(\sum_{j=1}^m |g_j(x)| + \sum_{i=1}^s \max\{0, -(h_{i0}(x) - \|\bar{h}_i(x)\|)\} \Big), \tag{5}$$

where $\alpha > 0$ is a penalty parameter.

The algorithm is formally stated as follows:

Algorithm 1

Step 0 *Choose $x^0 \in \Re^n$, $\alpha_0 > 0$, $\beta \in (0, 1)$, $\sigma \in (0, 1)$, $\tau > 0$, and set $k := 0$.*

Step 1 *Choose a symmetric positive definite matrix $M_k \in \Re^{n \times n}$. Find the solution Δx^k and the corresponding Lagrange multipliers λ^k and μ^k satisfying the KKT conditions (4) of subproblem (3). If $\|\Delta x^k\| = 0$, then STOP. Otherwise, go to Step 2.*

Step 2 *Set the penalty parameter as follows: If $\alpha_k \geq \max \left\{ \max_{1 \leq j \leq m} |\lambda_j^k|, \max_{1 \leq i \leq s} \mu_{i0}^k \right\}$, then $\alpha_{k+1} := \alpha_k$; otherwise, $\alpha_{k+1} := \max \left\{ \alpha_k, \max_{1 \leq j \leq m} |\lambda_j^k|, \max_{1 \leq i \leq s} \mu_{i0}^k \right\} + \tau$.*

Step 3 *Compute the smallest nonnegative integer r satisfying*

$$P_{\alpha_{k+1}}(x^k) - P_{\alpha_{k+1}}(x^k + (\beta)^r \Delta x^k) \geq \sigma(\beta)^r \Delta x^{kT} M_k \Delta x^k, \tag{6}$$

and set the step size $t_k := (\beta)^r$.

Step 4 *Set $x^{k+1} := x^k + t_k \Delta x^k$, $k := k + 1$, and go to Step 1.*

3.2 Global Convergence

In this subsection, we show that Algorithm 1 has a global convergence property. For simplicity, we assume $s := 1$. The arguments in what follows apply in a similar manner to the case where $s > 1$.

When $s = 1$, the KKT conditions (4) reduce to

$$\nabla f(x^k) + M_k \Delta x - \nabla g(x^k)\lambda - \nabla h(x^k)\mu = 0,$$
$$g(x^k) + \nabla g(x^k)^T \Delta x = 0, \tag{7}$$
$$h(x^k) + \nabla h(x^k)^T \Delta x \in K^l, \quad \mu \in K^l,$$
$$(h(x^k) + \nabla h(x^k)^T \Delta x)^T \mu = 0$$

and the penalty function used as a merit function is given by

$$P_\alpha(x) = f(x) + \alpha\Big(\sum_{j=1}^m |g_j(x)| + \max\{0, -(h_0(x) - \|\bar{h}(x)\|)\}\Big), \tag{8}$$

where $h(x) := (h_0(x), \bar{h}(x))^T$ with $h_0 : \Re^n \to \Re$ and $\bar{h} : \Re^n \to \Re^{l-1}$.

To prove global convergence of Algorithm 1, we make the next two assumptions.
(A.1) At every iteration, subproblem (3) has a KKT triple $(\Delta x^k, \lambda^k, \mu^k)$.
(A.2) The generated sequence $\{(x^k, \lambda^k, \mu^k)\}$ is bounded.

Assumption (A.1) implies that subproblem (3) has a unique optimal solution since M_k is a positive definite matrix. Below, we will show that the optimal solution Δx^k of subproblem (3) affords a descent direction of the penalty function P_{α_k} at x^k, provided $\alpha_k \geq \max\Big\{ \max_{1\leq j\leq m} |\lambda_j^k|, \mu_0^k\Big\}$. Thus we can determine the step size t_k in Step 3, and Algorithm 1 is well defined. Assumption (A.2) is standard in SQP methods for nonlinear programming.

In what follows, we denote

$$\varphi(x) := \max\{0, -(h_0(x) - \|\bar{h}(x)\|)\},$$

$$\psi(x) := \sum_{j=1}^m |g_j(x)|.$$

The next lemma gives a formula for the directional derivative of φ.

Lemma 1 *The directional derivative $\varphi'(x; \Delta x)$ of φ at x along the direction $\Delta x = (\Delta x_0, \Delta \bar{x})^T$ is given by*

$$\varphi'(x; \Delta x) = \begin{cases} -\nabla h_0(x)^T \Delta x + \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|}\Delta x & \left(\begin{array}{l} h_0(x) < \|\bar{h}(x)\|, \bar{h}(x) \neq 0 \text{ or} \\ h_0(x) = \|\bar{h}(x)\| \neq 0 \text{ and} \\ \nabla h_0(x)^T \Delta x < \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|}\Delta x \end{array}\right) \\[2em] -\nabla h_0(x)^T \Delta x + \|\nabla \bar{h}(x)^T \Delta x\| & \left(\begin{array}{l} h_0(x) < \|\bar{h}(x)\|, \bar{h}(x) = 0 \text{ or} \\ h_0(x) = \bar{h}(x) = 0 \text{ and} \\ \nabla h_0(x)^T \Delta x < \|\nabla \bar{h}(x)^T \Delta x\| \end{array}\right) \\[2em] 0 & \text{(otherwise)}. \end{cases}$$

Proof   We show this lemma by cases.
(i) If $h_0(x) < \|\bar{h}(x)\|$, then

$$\begin{aligned} \varphi'(x; \Delta x) &= \lim_{t\searrow 0} \frac{1}{t}(-h_0(x + t\Delta x) + \|\bar{h}(x + t\Delta x)\| + h_0(x) - \|\bar{h}(x)\|) \\[1em] &= -\nabla h_0(x)^T \Delta x + \lim_{t\searrow 0} \frac{1}{t}(\|\bar{h}(x + t\Delta x)\| - \|\bar{h}(x)\|) \\[1em] &= \begin{cases} -\nabla h_0(x)^T \Delta x + \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|}\Delta x & (\bar{h}(x) \neq 0) \\ -\nabla h_0(x)^T \Delta x + \|\nabla \bar{h}(x)^T \Delta x\| & (\bar{h}(x) = 0). \end{cases} \end{aligned}$$

4

(ii) If $h_0(x) = \|\bar{h}(x)\| = 0$, then

$$
\begin{aligned}
\varphi'(x; \Delta x) &= \lim_{t \searrow 0} \frac{1}{t} \max\left\{0, -(h_0(x + t\Delta x) - \|\bar{h}(x + t\Delta x)\|)\right\} \\
&= \begin{cases} -\nabla h_0(x)^T \Delta x + \|\nabla \bar{h}(x)^T \Delta x\| & \left(\nabla h_0(x)^T \Delta x < \|\nabla \bar{h}(x)^T \Delta x\|\right) \\ 0 & \left(\nabla h_0(x)^T \Delta x \geq \|\nabla \bar{h}(x)^T \Delta x\|\right). \end{cases}
\end{aligned}
$$

(iii) If $h_0(x) = \|\bar{h}(x)\| \neq 0$, then

$$
\begin{aligned}
\varphi'(x; \Delta x) &= \lim_{t \searrow 0} \frac{1}{t} \max\left\{0, (-h_0(x + t\Delta x) + \|\bar{h}(x + t\Delta x)\| + h_0(x) - \|\bar{h}(x)\|)\right\} \\
&= \max\left\{0, -\nabla h_0(x)^T \Delta x + \lim_{t \searrow 0} \frac{1}{t}(\|\bar{h}(x + t\Delta x)\| - \|\bar{h}(x)\|\right\} \\
&= \begin{cases} -\nabla h_0(x)^T \Delta x + \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|} \Delta x & \left(\nabla h_0(x)^T \Delta x < \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|} \Delta x\right) \\ 0 & \left(\nabla h_0(x)^T \Delta x \geq \frac{(\nabla \bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|} \Delta x\right). \end{cases}
\end{aligned}
$$

(iv) If $h_0(x) > \|\bar{h}(x)\|$, then $\varphi'(x; \Delta x) = 0$. ■

In the next lemma, using the directional derivative $\varphi'(x; \Delta x)$ given in Lemma 1, we derive an inequality that is used to prove global convergence of the algorithm.

Lemma 2 *Let $(\Delta x^k, \lambda^k, \mu^k)$ be a KKT triple of subproblem (3). If $\alpha > \mu_0^k$, then the directional derivative $\varphi'(x^k; \Delta x^k)$ of φ at x^k along the direction Δx^k satisfies the inequality*

$$
-\mu^{kT} h(x^k) + \alpha \varphi'(x^k; \Delta x^k) \leq 0.
$$

Proof Using the formula of $\varphi'(x; \Delta x)$ given in Lemma 1, we show the desired inequality by cases.
(i) If $h_0(x^k) < \|\bar{h}(x^k)\|$ and $\bar{h}(x^k) \neq 0$, then we have

$$
\begin{aligned}
&-\mu^{kT} h(x^k) + \alpha \varphi'(x^k; \Delta x^k) \\
&= -\mu^{kT} h(x^k) + \alpha\left(-\nabla h_0(x^k)^T \Delta x^k + \frac{(\nabla \bar{h}(x^k)\bar{h}(x^k))^T}{\|\bar{h}(x^k)\|} \Delta x^k\right) \\
&\leq -\mu^{kT} h(x^k) + \alpha\left(h_0(x^k) - \|\bar{h}(x^k) + \nabla \bar{h}(x^k)^T \Delta x^k\| + \frac{(\nabla \bar{h}(x^k)\bar{h}(x^k))^T}{\|\bar{h}(x^k)\|} \Delta x^k\right) \\
&= -\mu^{kT} h(x^k) + \alpha(h_0(x^k) - \|\bar{h}(x^k) + \nabla \bar{h}(x^k)^T \Delta x^k\| \\
&\quad + \frac{\bar{h}(x^k)^T(\bar{h}(x^k) + \nabla \bar{h}(x^k)^T \Delta x^k)}{\|\bar{h}(x^k)\|} - \|\bar{h}(x^k)\|) \\
&\leq (\alpha - \mu_0^k)h_0(x^k) - \bar{\mu}^{kT}\bar{h}(x^k) - \alpha\|\bar{h}(x^k)\| \\
&\leq (\alpha - \mu_0^k)h_0(x^k) - (\alpha - \|\bar{\mu}^k\|)\|\bar{h}(x^k)\| \\
&\leq -(\mu_0^k - \|\bar{\mu}^k\|)\|\bar{h}(x^k)\| \\
&\leq 0,
\end{aligned}
$$

where the first inequality holds by $h(x^k) + \nabla h(x^k)^T \Delta x^k \in K^l$ in the KKT conditions (7), the second and the third inequalities follow from Cauchy-Schwarz inequality, and the fourth and the last inequalities follow from $\alpha \geq \mu_0^k \geq \|\bar{\mu}^k\|$ and $h_0(x^k) < \|\bar{h}(x^k)\|$, $\mu^k \in K^l$, respectively.

(ii) If $h_0(x^k) < \|\bar{h}(x^k)\|$ and $\bar{h}(x^k) = 0$, then we have

$$
\begin{aligned}
-\mu^{kT}h(x^k) + \alpha\varphi'(x^k; \Delta x^k) &= -\mu_0^{kT}h_0(x^k) + \alpha(-\nabla h_0(x^k)^T\Delta x^k + \|\nabla\bar{h}(x^k)^T\Delta x^k\|) \\
&\leq (\alpha - \mu_0^k)h_0(x^k) \\
&\leq 0,
\end{aligned}
$$

where the first inequality follows from $h(x^k) + \nabla h(x^k)^T\Delta x^k \in K^l$ and the last inequality holds by $\alpha \geq \mu_0^k$.

(iii) If $h_0(x^k) = 0$ and $\bar{h}(x^k) = 0$, then $\nabla h_0(x^k)^T\Delta x^k \geq \|\nabla\bar{h}(x^k)^T\Delta x^k\|$ implies $h_0(x^k) + \nabla h_0(x^k)^T\Delta x^k \geq \|\bar{h}(x^k) + \nabla\bar{h}(x^k)^T\Delta x^k\|$, which in turn implies $\varphi'(x^k; \Delta x^k) = 0$ by the formula shown in Lemma 1. Therefore we obtain

$$
-\mu^{kT}h(x^k) + \alpha\varphi'(x^k; \Delta x^k) = -\mu^{kT}h(x^k) = 0.
$$

(iv) Suppose $h_0(x^k) = \|\bar{h}(x^k)\| \neq 0$. If $\nabla h_0(x)^T\Delta x < \frac{(\nabla\bar{h}(x)\bar{h}(x))^T}{\|\bar{h}(x)\|}\Delta x$, then similarly to case (i), we have

$$
\begin{aligned}
&-\mu^{kT}h(x^k) + \alpha\varphi'(x^k; \Delta x^k) \\
&= -\mu^{kT}h(x^k) + \alpha\left(-\nabla h_0(x^k)^T\Delta x^k + \frac{(\nabla\bar{h}(x^k)\bar{h}(x^k))^T}{\|\bar{h}(x^k)\|}\Delta x^k\right) \\
&\leq -\mu^{kT}h(x^k) + \alpha\left(h_0(x^k) - \|\bar{h}(x^k) + \nabla\bar{h}(x^k)^T\Delta x^k\| + \frac{(\nabla\bar{h}(x^k)\bar{h}(x^k))^T}{\|\bar{h}(x^k)\|}\Delta x^k\right) \\
&= -\mu^{kT}h(x^k) + \alpha(h_0(x^k) - \|\bar{h}(x^k) + \nabla\bar{h}(x^k)^T\Delta x^k\| \\
&\quad + \frac{\bar{h}(x^k)^T(\bar{h}(x^k) + \nabla\bar{h}(x^k)^T\Delta x^k)}{\|\bar{h}(x^k)\|} - \|\bar{h}(x^k)\|) \\
&\leq (\alpha - \mu_0^k)h_0(x^k) - \bar{\mu}^{kT}\bar{h}(x^k) - \alpha\|\bar{h}(x^k)\| \\
&\leq (\alpha - \mu_0^k)h_0(x^k) - (\alpha - \|\bar{\mu}^k\|)\|\bar{h}(x^k)\| \\
&= -(\mu_0^k - \|\bar{\mu}^k\|)\|\bar{h}(x^k)\| \\
&\leq 0.
\end{aligned}
$$

Otherwise, $\varphi'(x^k, \Delta x^k) = 0$, so it follows from $\mu^k \in K^l$ and Cauchy-Schwarz inequality that

$$
\begin{aligned}
-\mu^{kT}h(x^k) + \alpha\varphi'(x^k; \Delta x^k) &= -\mu^{kT}h(x^k) \\
&= -\mu_0^k h_0(x^k) - \bar{\mu}^{kT}\bar{h}(x^k) \\
&\leq -(\mu_0^k - \|\bar{\mu}^k\|)\|\bar{h}(x^k)\| \\
&\leq 0.
\end{aligned}
$$

(v) If $h_0(x^k) > \|\bar{h}(x^k)\|$, then it follows from $\mu^k \in K^l$ and Cauchy-Schwarz inequality that

$$
\begin{aligned}
-\mu^{kT}h(x^k) + \alpha\varphi'(x^k; \Delta x^k) &= -\mu^{kT}h(x^k) \\
&= -\mu_0^k h_0(x^k) - \bar{\mu}^{kT}\bar{h}(x^k) \\
&\leq -(h_0(x^k) - \|\bar{h}(x^k)\|)\|\bar{\mu}^k\| \\
&\leq 0.
\end{aligned}
$$

$\blacksquare$

The next lemma gives an equality regarding the directional derivative $\psi'(x; \Delta x)$ of the function ψ. The proof is not difficult and hence is omitted.

Lemma 3 *Let Δx^k be the optimal solution of subproblem (3). Then the directional derivative $\psi'(x^k; \Delta x^k)$ of ψ at x^k along the direction Δx^k satisfies the equality*

$$\psi'(x^k; \Delta x^k) = -\sum_{j=1}^{m} |g_j(x^k)|.$$

From the above lemmas, we obtain the following lemma.

Lemma 4 *Let $(\Delta x^k, \lambda^k, \mu^k)$ be a KKT triple of subproblem (3). If $\alpha \geq \max \left\{ \max_{1 \leq j \leq m} |\lambda_j^k|, \mu_0^k \right\}$, then the directional derivative $P'_\alpha(x^k; \Delta x^k)$ of the penalty function P_α at x^k along the direction Δx^k satisfies the inequality*

$$P'_\alpha(x^k; \Delta x^k) \leq -\Delta x^{kT} M_k \Delta x^k.$$

Proof   By the KKT conditions (7) and Lemma 3, we have

$$
\begin{aligned}
P'&(x^k; \Delta x^k) \\
&= \nabla f(x^k)^T \Delta x^k + \alpha(\psi'(x^k; \Delta x^k) + \varphi'(x^k, \Delta x^k)) \\
&= -\Delta x^{kT} M_k \Delta x^k + \lambda^{kT} \nabla g(x^k)^T \Delta x^k + \mu^{kT} \nabla h(x^k)^T \Delta x^k + \alpha(\psi'(x^k; \Delta x^k) + \varphi'(x^k, \Delta x^k)) \\
&= -\Delta x^{kT} M_k \Delta x^k - \lambda^{kT} g(x^k) - \mu^{kT} h(x^k) + \alpha(-\sum_{j=1}^{m} |g_j(x^k)| + \varphi'(x^k, \Delta x^k)) \\
&\leq -\Delta x^{kT} M_k \Delta x^k - \sum_{j=1}^{m} (\lambda_j^k + \alpha)|g_j(x^k)| - \mu^{kT} h(x^k) + \alpha\varphi'(x^k, \Delta x^k)).
\end{aligned}
$$

Then Lemma 2 together with the inequality $\alpha \geq \max \left\{ \max_{1 \leq j \leq m} |\lambda_j^k|, \mu_0^k \right\}$ yields the desired inequality. ∎

When $\Delta x^k \neq 0$, by Lemma 4 and the positive definiteness of the matrix M_k, we have

$$
\begin{aligned}
P_\alpha&(x^k) - P_\alpha(x^k + t\Delta x^k) - \sigma t \Delta x^{kT} M_k \Delta x^k \\
&= -t P'_\alpha(x^k; \Delta x^k) + o(t) - \sigma t \Delta x^{kT} M_k \Delta x^k \\
&\geq (1 - \sigma) t \Delta x^{kT} M_k \Delta x^k + o(t) \\
&> 0
\end{aligned}
$$

for any sufficiently small $t > 0$. This ensures that we can always determine the step size t_k in Step 3 of Algorithm 1.

Now we are ready to establish global convergence of Algorithm 1.

Theorem 1   *Suppose that Assumptions (A.1) and (A.2) hold. Let $\{(x^k, \lambda^k, \mu^k)\}$ be a sequence generated by Algorithm 1, and (x^*, λ^*, μ^*) be any accumulation point. Assume that there exist some positive scalars γ and Γ such that*

$$\gamma \|z\|^2 \leq z^T M_k z \leq \Gamma \|z\|^2, \quad \forall z \in \Re^n, \quad \forall k \in \{0, 1, 2, \cdots\}.$$

Then, (x^*, λ^*, μ^*) satisfies the KKT conditions (2) of NSOCP (1)

Proof   Since $\{M_k\}$ is bounded, we only need to show $\lim_{k \to \infty} \|\Delta x^k\| = 0$ from Proposition 1. First note that, from (A.2) and the way of updating the penalty parameter, α_k stays constant $\bar{\alpha}$ eventually.

Consequently, $\{P_{\bar{\alpha}}(x^k)\}$ is monotonically nonincreasing for sufficiently large k. Meanwhile, by (6) and the positive definiteness of M_k, we have

$$P_{\bar{\alpha}}(x^k) - P_{\bar{\alpha}}(x^{k+1}) \geq \sigma t_k \Delta x^{kT} M_k \Delta x^k > 0.$$

Since $\{P_{\bar{\alpha}}(x^k)\}$ is bounded below by (A.2), we have

$$\lim_{k \to \infty} P_{\bar{\alpha}}(x^k) - P_{\bar{\alpha}}(x^{k+1}) = 0.$$

Therefore, it holds that

$$\lim_{k \to \infty} t_k \Delta x^{kT} M_k \Delta x^k = 0.$$

Moreover, it follows from the given assumption that

$$t_k \Delta x^{kT} M_k \Delta x^k \geq t_k \gamma \|\Delta x^k\|^2.$$

Hence, we have $\lim\limits_{k \to \infty} t_k \|\Delta x^k\|^2 = 0$. It clearly holds that $\lim\limits_{k' \to \infty} \|\Delta x^{k'}\| = 0$ for any subsequence $\{\Delta x^{k'}\}$ such that $\liminf\limits_{k' \to \infty} t_{k'} > 0$. Let us consider an arbitrary subsequence $\{t_{k'}\}$ such that $\lim\limits_{k' \to \infty} t_{k'} = 0$. Then, by the Armijo rule in Step 3, we have

$$P_{\bar{\alpha}}(x^{k'}) - P_{\bar{\alpha}}(x^{k'} + \bar{t}_{k'} \Delta x^{k'}) < \sigma \bar{t}_{k'} \Delta x^{k'T} M_{k'} \Delta x^{k'},$$

where $\bar{t}_{k'} := \frac{t_{k'}}{\beta}$. On the other hand, since $P'_{\bar{\alpha}}(x^{k'}; \Delta x^{k'}) \leq -\Delta x^{k'T} M_{k'} \Delta x^{k'}$ by Lemma 4, it follows that

$$P_{\bar{\alpha}}(x^{k'}) - P_{\bar{\alpha}}(x^{k'} + \bar{t}_{k'} \Delta x^{k'}) = -\bar{t}_{k'} P'_{\bar{\alpha}}(x^{k'}; \Delta x^{k'}) + o(\bar{t}_{k'}) \geq \bar{t}_{k'} \Delta x^{k'} M_{k'} \Delta x^{k'} + o(\bar{t}_{k'}).$$

Combining the above inequalities yields $\bar{t}_{k'} \Delta x^{k'} M_{k'} \Delta x^{k'} + o(\bar{t}_{k'}) < \sigma \bar{t}_{k'} \Delta x^{k'} M_{k'} \Delta x^{k'}$, i.e.,

$$0 > (1 - \sigma) \bar{t}_{k'} \Delta x^{k'} M_{k'} \Delta x^{k'} + o(\bar{t}_{k'}) > (1 - \sigma) \bar{t}_{k'} \gamma \|\Delta x^{k'}\|^2 + o(\bar{t}_{k'}).$$

Thus we obtain

$$(1 - \sigma) \gamma \|\Delta x^{k'}\|^2 + \frac{o(\bar{t}_{k'})}{\bar{t}_{k'}} < 0,$$

which yields $\limsup\limits_{k' \to \infty} \|\Delta x^{k'}\| \leq 0$. Consequently, we have $\lim\limits_{k \to \infty} \|\Delta x^k\| = 0$. ∎

3.3 Local Convergence

In this subsection, we consider local behavior of a sequence generated by Algorithm 1. For that purpose, we make use of the results for generalized equations [11].

First note that the KKT conditions of NSOCP (1) can be rewritten as the generalized equation

$$0 \in F(y) + \partial \delta_C(y), \tag{9}$$

where F is a vector-valued function and $\partial \delta_C(y)$ is the normal cone of a closed convex set C at y, which is defined by

$$\partial \delta_C(y) := \begin{cases} \emptyset & \text{if} \quad y \notin C \\ \{w \mid w^T(c - y) \leq 0 \quad \forall c \in C\} & \text{if} \quad y \in C. \end{cases}$$

Indeed, by defining the Lagrangian of NSOCP (1) by

$$L(x, \zeta, \eta) := f(x) - g(x)^T \zeta - h(x)^T \eta,$$

the KKT conditions (2) are represented as

$$0 \in \nabla_x L(x, \zeta, \eta) + \partial \delta_{\Re^n}(x),$$
$$0 \in \nabla_\zeta L(x, \zeta, \eta) + \partial \delta_{\Re^m}(\zeta),$$
$$0 \in \nabla_\eta L(x, \zeta, \eta) + \partial \delta_{K^*}(\eta),$$

where $K^* := \{\eta \in \Re^l \mid \eta^T \xi \geq 0 \ \forall \xi \in K\}$ is the dual cone of K. Since $\partial \delta_{\Re^n}(x) = \{0\}, \partial \delta_{\Re^m}(\zeta) = \{0\}$ and $K^* = K$, we can rewrite the KKT conditions (2) as the generalized equation (9) with $y := (x, \zeta, \eta)^T$, $C := \Re^n \times \Re^m \times K$ and

$$F(y) := \begin{bmatrix} \nabla_x L(x, \zeta, \eta) \\ \nabla_\zeta L(x, \zeta, \eta) \\ \nabla_\eta L(x, \zeta, \eta) \end{bmatrix}. \tag{10}$$

On the other hand, if we choose $M_k := \nabla_{xx}^2 L(x^k, \lambda^k, \mu^k)$, we can express the KKT conditions of subproblem (3) as

$$0 \in \nabla_x L(x^k, \lambda, \mu) + \nabla_{xx}^2 L(x^k, \lambda^k, \mu^k)\Delta x + \partial \delta_{\Re^n}(x),$$
$$0 \in \nabla_\zeta L(x^k, \lambda, \mu) + \nabla_{\zeta x}^2 L(x^k, \lambda^k, \mu^k)\Delta x + \partial \delta_{\Re^m}(\lambda),$$
$$0 \in \nabla_\eta L(x^k, \lambda, \mu) + \nabla_{\eta x}^2 L(x^k, \lambda^k, \mu^k)\Delta x + \partial \delta_K(\mu),$$

which is precisely the generalized equation

$$0 \in F(z^k) + F'(z^k)(z - z^k) + \partial \delta_C(z), \tag{11}$$

where $z^k = (x^k, \lambda^k, \mu^k), z = (x^k + \Delta x, \lambda, \mu)$ and F is defined by (10). This can be regarded as the application of Newton's method for the generalized equation (9). Thus, a sequence $\{z^k\}$ generated by (11) is expected to converge fast to a solution of (10). To be more precise, we use the notion of a regular solution [11].

Definition 1 *Let y^* be a solution of the generalized equation (9) and F be Fréchet differentiable at y^*. Define the set-valued mapping T by $T(y) := F(y^*) + F'(y^*)(y - y^*) + \partial \delta_C(y)$. If there exist neighborhoods U of 0 and V of y^* such that the mapping $T^{-1} \cap V$ is single-valued and Lipschitzian on U, then y^* is called a regular solution of the generalized equation (9).*

We suppose that F is Fréchet differentiable with Lipschitz constant L and the generalized equation (11) at $k = 0$

$$0 \in F(z^0) + F'(z^0)(z - z^0) + \partial \delta_C(z)$$

has a regular solution with Lipschitz constant Λ. Then (11) has a regular solution at every iteration and the following inequality holds for a sequence $\{z^k\}$ generated by (11), provided z^0 is sufficiently close to a regular solution y^* of the generalized equation (9) (see [11]):

$$\|z^k - y^*\| \leq (2^{(l+n+m)}\Lambda L)^{-1}(2\Lambda L\|z^0 - z^1\|)^{(2^k)},$$

which means that the sequence $\{z^k\}$ converges R-quadratically to y^*.

Next we consider the relation between the regularity of a solution and the second-order conditions for NSOCP (1). We recall the notion of nondegeneracy in second-order cone programming [4].

Definition 2 *For each $i = 1, \cdots, s$, given a vector $\hat{w}_i \in K^{l_i}$, define the function ϕ_i as follows:*
(i) if $\hat{w}_i = 0$, then $\phi_i : \Re^{l_i} \to \Re^{l_i}$ and $\phi_i(w_i) := w_i$;
(ii) if $\hat{w}_{i0} > \|\bar{\hat{w}}_i\|$, then $\phi_i : \Re^{l_i} \to \Re^0$ and $\phi_i(w_i) := 0$;
(iii) if $\hat{w}_{i0} = \|\bar{\hat{w}}_i\| \neq 0$, then $\phi_i : \Re^{l_i} \to \Re^1$ and $\phi_i(w_i) := \|\bar{w}_i\| - w_{i0}$.
Let x be a feasible solution of NSOCP (1). If the matrix

$$(\nabla g(x), \nabla h_1(x)\nabla\phi_1(h_1(x)), \cdots, \nabla h_s(x)\nabla\phi_s(h_s(x)))$$

has full column rank, then x is said to be nondegenerate. Here, $\nabla h_i(x)\nabla\phi_i(h_i(x)) = \nabla h_i(x)$ if $h_i(x) = 0$, $\nabla h_i(x)\nabla\phi_i(h_i(x)) = -\nabla h_{i0}(x) + \frac{\nabla \bar{h}_i(x)\bar{h}_i(x)}{\|\bar{h}_i(x)\|}$ if $h_{i0}(x) = \|\bar{h}_i(x)\| \neq 0$, and $\nabla h_i(x)\nabla\phi_i(h_i(x))$ is vacuous if $h_{i0}(x) > \|\bar{h}_i(x)\|$.

It is showed in [4] that when a local optimal solution x^* of NSOCP(1) is nondegenerate, (x^*, ζ^*, η^*) is a regular solution of the generalized equation representing the KKT conditions (2) of NSOCP (1) if and only if (x^*, ζ^*, η^*) satisfies the following second-order condition:

$$d^T\Big(\nabla^2_{xx}L(x^*,\zeta^*,\eta^*) + \sum_{i=1}^s H_i(x^*,\zeta^*,\eta_i^*)\Big)d > 0 \quad \forall d \in C_0(x^*) \cap C_{K^{l_1}}(x^*) \cap \cdots \cap C_{K^{l_s}}(x^*), \, d \neq 0, \quad (12)$$

where

$$C_0(x^*) = \{d \in \Re^n \mid \nabla g(x^*)^T d = 0\}$$

and for $i = 1, \cdots, s$

$$C_{K^{l_i}}(x^*) = \left\{d \in \Re^n \left| \begin{array}{ll} \nabla h_i(x^*)^T d = 0 & \text{if } \eta_{i0}^* > \bar{\eta}_i^*, \\ \nabla h_i(x^*)^T d \in \text{span}\{R_i\eta_i^*\} & \text{if } \eta_{i0}^* = \|\bar{\eta}_i^*\| \neq 0, h_i(x^*) = 0, \\ d^T\nabla h_i(x^*)\eta_i^* = 0 & \text{if } \eta_{i0}^* = \|\bar{\eta}_i^*\| \neq 0, h_{i0}(x^*) = \|\bar{h}_i(x^*)\| \neq 0, \\ \text{no condition} & \text{otherwise} \end{array}\right.\right\},$$

$$H_i(x^*,\zeta^*,\eta_i^*) = \left\{\begin{array}{ll} -\frac{\eta_{i0}^*}{h_{i0}(x^*)}\nabla h_i(x^*)R_i\nabla h_i(x^*)^T & \text{if } h_{i0}(x^*) = \|\bar{h}_i(x^*)\| \neq 0, \\ 0 & \text{otherwise} \end{array}\right.$$

with $R_i := \begin{pmatrix} 1 & 0^T \\ 0 & -I_{l_i-1} \end{pmatrix} \in \Re^{l_i \times l_i}$. Summarizing the above arguments, we have the next theorem about the local behavior of a sequence $\{(x^k, \zeta^k, \eta^k)\}$ generated by Algorithm 1.

Theorem 2 *Suppose $M_k = \nabla^2_{xx}L(x^k, \zeta^k, \eta^k)$ and $t_k = 1$ for all $k > \bar{k}$, where $\bar{k}$ is a positive integer. If, for some $k > \bar{k}$, (x^k, ζ^k, η^k) is sufficiently close to a KKT triple (x^*, ζ^*, η^*) of NSOCP (1) such that x^* is nondegenerate and (x^*, ζ^*, η^*) satisfies the second-order condition (12), then the sequence $\{(x^k, \zeta^k, \eta^k)\}$ generated by Algorithm 1 converges R-quadratically to (x^*, ζ^*, η^*). In particular, $\{x^k\}$ converges R-quadratically to x^*.*

4   Numerical Experiments

We implemented Algorithm 1 in MATLAB (Version 6.5) using the SDPT3-Solver (Version 3.0) [13] to solve the subproblems by transforming them into LSOCPs. In Algorithm 1, we set the parameters as $\alpha_0 = 1, \tau = 0.01, \sigma = 0.2, \beta = 0.95$ and use the stopping criterion $\|\Delta x^k\| < 10^{-4}$.

Experiment 1. First, we consider the following convex NSOCP:

$$\min \quad x^T C x + \sum_{i=1}^{n} (d_i x_i^4 + f_i x_i)$$
$$\text{s.t.} \quad Ax + \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \in K := K^{l_1} \times \cdots \times K^{l_s}, \tag{13}$$

where $d_i, f_i \, (i = 1, \cdots, n)$ are scalars, $b_j \, (j = 1, \cdots, s)$ are l_j-dimensional vectors with $l_1 + \cdots + l_s = n$, C is an $n \times n$ symmetric positive semidefinite matrix, and A is an $n \times n$ matrix. We generate ten problem instances for each of $n = 10, 30, 50$. We determine the constants as follows: d_i and $f_i \, (i = 1, \cdots, n)$ are randomly chosen from the intervals [0,1] and [−1, 1], respectively, and C is given by $C := Z^T Z$, where Z is an $n \times n$ matrix whose elements are randomly chosen from the interval [0,1]. The elements of the matrix A are chosen from the interval [0,2]. Vectors $b_j \in \Re^{l_j} \, (j = 1, \cdots, s)$ are determined as $b_{j0} = 1, \bar{b}_j = 0$. Then, problem (13) is always feasible, since $x = 0$ satisfies the constraints.

Each problem instance is solved by Algorithm 1 using an initial point whose elements are randomly generated from the interval $[-1, 1]$. The following two updating formulas for matrices M_k are tested. Both update formulas ensure the positive definiteness of M_k for all k.

Modified Newton formula. At iteration k, if the Hessian $\nabla_{xx}^2 L(x^k, \mu^{k-1})$ of the Lagrangian is a positive definite matrix, then set $M_k = \nabla_{xx}^2 L(x^k, \mu^{k-1})$; otherwise, set $M_k = \nabla_{xx}^2 L(x^k, \mu^{k-1}) + (|\xi_k| + 0.1)I$, where ξ_k is the minimum eigenvalue of $\nabla_{xx}^2 L(x^k, \mu^{k-1})$. At the first iteration, M_0 is set to be the identity matrix I.

Quasi-Newton formula. We set $M_0 = I$ and subsequently update M_k by

$$M_{k+1} = M_k - \frac{M_k v^k v^{kT} M_k}{v^{kT} M_k v^k} + \frac{u^k u^{kT}}{v^{kT} u^k},$$

where $v^k = x^{k+1} - x^k$, $w^k = \nabla_x L(x^{k+1}, \lambda^k, \mu^k) - \nabla_x L(x^k, \lambda^k, \mu^k)$, $u^k = \theta_k w^k + (1 - \theta_k) M_k v^k$, and θ_k is determined by

$$\theta_k = \begin{cases} 1 & \text{if } v^{kT} w^k \geq 0.2 v^{kT} M_k v^k \\ \frac{0.8 v^{kT} M_k v^k}{v^{kT}(M_k v^k - w^k)} & \text{otherwise.} \end{cases}$$

This is a modified BFGS update suggested in the SQP method for NLP [3].

Table 1: Computational results for the convex NSOCP (13)

		modified Newton			quasi-Newton		
n	K	k_{ave}	k_{min}	k_{max}	k_{ave}	k_{min}	k_{max}
10	$K^5 \times K^5$	12.11	7	19	22.89	15	31
30	$K^5 \times K^5 \times K^{20}$	13.03	8	25	31.54	22	52
50	$K^5 \times K^5 \times K^{20} \times K^{20}$	13.97	8	29	38.86	25	68

In our experiments with the modified Newton formula, we observed that M_k was chosen to be $\nabla_{xx}^2 L(x^k, \mu^{k-1})$ and the step size was equal to 1 in the final stage of the iterations. In the case of the quasi-Newton formula, the step size was also equal to 1 in the final stage of the iterations. Table 1 shows the average k_{ave}, the minimum k_{min}, and the maximum k_{max} numbers of iterations for ten runs, along with the problem size and the Cartesian structure of the second-order cone K of each

test problem. We find that, for the convex NSCOP (13), the modified Newton formula results in faster convergence than the quasi-Newton formula. This suggests that the convexity of the objective function and the linearity of the constraint functions can be better exploited in the modified Newton formula, since it uses the Hessian matrices of those functions in a direct manner.

Experiment 2. Next, we consider the following nonconvex NSOCP:

$$
\begin{aligned}
\min \quad & x^T C x + \sum_{i=1}^{n} (d_i x_i^4 + e_i x_i^3 + f_i x_i) \\
\text{s.t.} \quad & \begin{pmatrix} a_1(e^{x_1} - 1) \\ a_2(e^{x_2} - 1) \\ \vdots \\ a_n(e^{x_n} - 1) \end{pmatrix} + \begin{pmatrix} \hat{a}_1 x_1 x_2 \\ \hat{a}_2 x_2 x_3 \\ \vdots \\ \hat{a}_n x_n x_1 \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \in K := K^{l_1} \times \cdots, \times K^{l_s},
\end{aligned}
\tag{14}
$$

where the constants are similar to those in the previous test problem, except that C is an $n \times n$ indefinite matrix. We generate ten problem instances for each of $n = 10, 30, 50$. We determine the constants as follows: $a_i, \hat{a}_i, e_i, f_i$ $(i = 1, \cdots, n)$ and the elements of C are randomly chosen from the interval $[-1, 1]$, and $d_i (i = 1, \cdots, n)$ are randomly chosen from the interval $[0, 1]$. Vectors $b_j \in \Re^{l_j} (j = 1, \cdots, s)$ are determined as $b_{j0} = 1, \bar{b}_j = 0$ similarly to the case of problem (13). Note that the objective function and the constraint functions are in general nonconvex.

As in the previous experiment, each problem instance is solved by Algorithm 1 using an initial point whose elements are randomly generated from the interval $[-1, 1]$. We test the two formulas for updating matrices M_k, the modified Newton formula and the quasi-Newton formula. The results are shown in Table 2.

Table 2: Computational results for the nonconvex NSOCP (14)

n	K	modified Newton			quasi-Newton		
		k_{ave}	k_{min}	k_{max}	k_{ave}	k_{min}	k_{max}
10	$K^5 \times K^5$	24.31	11	116	24.96	12	56
30	$K^5 \times K^5 \times K^{20}$	59.44	19	183	39.75	25	91
50	$K^5 \times K^5 \times K^{20} \times K^{20}$	68.64	20	180	50.22	31	97

Because of the lack of convexity in the objective and constraint functions, the Hessian of the Lagrangian is not likely to be positive definite even if x^k is close to a stationary point of the problem. Thus, the matrices M_k determined by the modified Newton formula may substantially differ from $\nabla_{xx}^2 L(x^k, \mu^{k-1})$. We have observed that the algorithm with the modified Newton formula performs somewhat inefficiently compared with the previous experiment, although it exhibits fast local convergence when $\nabla_{xx}^2 L(x^k, \mu^{k-1})$ becomes positive definite near a solution. In fact, Table 2 suggests that the quasi-Newton formula works more effectively especially when $\nabla_{xx}^2 L(x^k, \mu^{k-1})$ is indefinite.

References

[1] F. Alizadeh and D. Goldfarb: *Second-order cone programming.* Mathematical Programming, Vol. 95, 2003, pp. 3–51.

[2] A. Ben-Tal and A. Nemirovski: *Robust convex optimization.* Mathematics of Operations Research, Vol. 23, 1998, pp. 769–805.

[3] P.T. Boggs and J.W. Tolle: *Sequential quadratic programming.* Acta Numerica. Vol. 4, 1995, pp. 1–51.

[4] J.F. Bonnans and H. Ramírez: *Perturbation analysis of second-order-cone programming problems.* Mathematical Programming, Vol. 104, 2005, pp. 205–227.

[5] J.F. Bonnans and A. Shapiro: *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York, 2000.

[6] R. Correa and H. Ramírez: *A global algorithm for nonlinear semidefinite programming.* SIAM Journal on Optimization, Vol. 15, 2004, pp. 303–318.

[7] S.P. Han: *A globally convergent method for nonlinear programming.* Journal of Optimization Theory and Applications, Vol. 22, 1977, pp. 297–309.

[8] C. Kanzow, C. Nagel, H. Kato and M. Fukushima: *Successive linearization methods for nonlinear semidefinite programs.* Computational Optimization and Applications, Vol. 31, 2005, pp. 251–273.

[9] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret: *Applications of second-order cone programming.* Linear Algebra and Its Applications, Vol. 284, 1998, pp. 193–228.

[10] R.D.C. Monteiro and T. Tsuchiya: *Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions.* Mathematical Programming, Vol. 88, 2000, pp. 61–83.

[11] S.M. Robinson: *Generalized equations.* in A. Bachem et al. (eds.) *Mathematical Programming: The State of the Art,* Springer-Verlag, Berlin, 1983, pp. 346–367.

[12] M.J. Todd: *Semidefinite optimization.* Acta Numerica, Vol. 10, 2001, pp. 515–560.

[13] R.H. Tütüncü, K.C. Toh and M.J. Todd: *Solving semidefinite-quadratic-linear programs using SDPT3.* Mathematical Programming, Vol. 95, 2003, pp. 189–217.

[14] L. Vandenberghe and S. Boyd: *Semidefinite programming.* SIAM Review, Vol. 38, 1996, pp. 49–95.

[15] R.J. Vanderbei and D.F. Shanno: *An interior-point algorithm for nonconvex nonlinear programming.* Computational Optimization and Applications, Vol. 13, 1999, pp. 231–252.

[16] H. Yamashita and H. Yabe: *An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization.* SIAM Journal on Optimization, Vol. 14, 2003, pp. 479–499.

[17] H. Yamashita and H. Yabe: *A primal-dual interior point method for nonlinear optimization over second order cones.* Technical Report, Mathematical Systems Inc., Tokyo, May 2005 (revised February 2006).