COMPUTATIONAL & APPLIED MATHEMATICS

Volume 30, N. 3, pp. 569–588, 2011

ISSN 0101-8205
www.scielo.br/cam

A smoothing Newton-type method for second-order cone programming problems based on a new smoothing Fischer-Burmeister function

LIANG FANG[1]* and ZENGZHE FENG[2]

[1]College of Mathematics and System Science, Taishan University, 271021, Tai'an, P.R. China

[2]College of Information Engineering, Taishan Medical University, 271016, Tai'an, P.R. China

E-mails: fangliang3@163.com / fengzengzhe@163.com

Abstract. A new smoothing function of the well known Fischer-Burmeister function is given. Based on this new function, a smoothing Newton-type method is proposed for solving second-order cone programming. At each iteration, the proposed algorithm solves only one system of linear equations and performs only one line search. This algorithm can start from an arbitrary point and it is Q-quadratically convergent under a mild assumption. Numerical results demonstrate the effectiveness of the algorithm.

Mathematical subject classification: 90C25, 90C30, 90C51, 65K05, 65Y20.

Key words: second-order cone programming, smoothing method, interior-point method, Q-quadratic convergence, central path, strong semismoothness.

1 Introduction

The second order cone (SOC) in $\mathcal{R}^n (n \geq 2)$, also called the Lorentz cone or the ice-cream cone, is defined as

$$\mathcal{Q}_n = \left\{(x_1; x_2) | \, x_1 \in \mathcal{R}, x_2 \in \mathcal{R}^{n-1} \text{ and } x_1 \geq \|x_2\|)\right\},$$

#CAM-227/10. Received: 20/VI/10. Accepted: 03/VIII/10.
*Corresponding author.

here and below, $\|\cdot\|$ refers to the standard Euclidean norm, n is the dimension of $\mathcal{Q}_n$, and for convenience, we write $x = (x_1; x_2)$ instead of $(x_1, x_2^T)^T$. It is easy to verify that the SOC $\mathcal{Q}_n$ is self-dual, that is

$$\mathcal{Q}_n = \mathcal{Q}_n^* = \left\{s \in \mathcal{R}^n : s^T x \geq 0, \text{ for all } x \in \mathcal{Q}_n\right\}.$$

We may often drop the subscripts if the dimension is evident from the context.

For any $x = (x_1; x_2), y = (y_1; y_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$, their Jordan product is defined as [5]

$$x \circ y = \left(x^T y; x_1 y_2 + y_1 x_2\right).$$

Second-order cone programming (SOCP) problems are to minimize a linear function over the intersection of an affine space with the Cartesian product of a finite number of SOCs. The study of SOCP is vast important as it covers linear programming, convex quadratic programming, quadratically constraint convex quadratic optimization as well as other problems from a wide range of applications in many fields, such as engineering, control, optimal control and design, machine learning, robust optimization and combinatorial optimization and so on [13, 24, 4, 23, 29, 22, 18, 10, 11].

Without loss of generality, we consider the SOCP problem with a single SOC

$$\text{(PSOCP)} \qquad \min\left\{\langle c, x\rangle : Ax = b, x \in \mathcal{Q}\right\} \tag{1}$$

and its dual problem

$$\text{(DSOCP)} \qquad \max\left\{\langle b, y\rangle : A^T y + s = c, s \in \mathcal{Q}, y \in \mathcal{R}^m\right\}, \tag{2}$$

where $c \in \mathcal{R}^n$, $A \in \mathcal{R}^{m\times n}$ and $b \in \mathcal{R}^m$, with an inner product $\langle\cdot,\cdot\rangle$, are given data. $x \in \mathcal{Q}$ is variable and the set $\mathcal{Q}$ is an SOC of dimension n. Note that our analysis can be easily extended to the general case with Cartesian product of SOCs.

We call $x \in \mathcal{Q}$ primal feasible if $Ax = b$. Similarly $(y, s) \in \mathcal{R}^m \times \mathcal{Q}$ is called dual feasible if $A^T y + s = c$. For a given primal-dual feasible point $(x, y, s) \in \mathcal{Q} \times \mathcal{R}^m \times \mathcal{Q}$, $\langle x, s\rangle$ is called the duality gap due to the well known weak dual theorem, i.e., $\langle x, s\rangle \geq 0$, which follows that

$$\langle c, x\rangle - \langle b, y\rangle = \langle A^T y + s, x\rangle - \langle Ax, y\rangle = \langle x, s\rangle \geq 0.$$

Let us note that $\langle x, s\rangle = 0$ is sufficient for optimality of primal and dual feasible $(x, y, s) \in \mathcal{Q} \times \mathcal{R}^m \times \mathcal{Q}$.

Throughout the paper, we make the following Assumption:

Assumption 2.1. Both (PSOCP) and its dual (DSOCP) are strictly feasible.

It is well known that under the Assumption 2.1, the SOCP is equivalent to its *optimality conditions*:

$$\begin{aligned} &Ax = b, \\ &A^T y + s = c, \\ &x \circ s = 0,\ \ x, s \in \mathcal{Q},\ y \in \mathcal{R}^m, \end{aligned} \tag{3}$$

where $\langle x, s\rangle = 0$ is usually referred to as the complementary condition.

There are an extensive literature focusing on interior-point methods (IPMs) for (PSOCP) and (DSOCP) (see, e.g., [1, 17, 6, 23, 16, 11] and references therein). IPMs typically deal with the following perturbation of the optimality conditions:

$$\begin{aligned} &Ax = b, \\ &A^T y + s = c, \\ &x \circ s = \mu e,\ \ x, s \in \mathcal{Q},\ \ y \in \mathcal{R}^m, \end{aligned} \tag{4}$$

where $\mu > 0$ and $e = (1; 0) \in \mathcal{R} \times \mathcal{R}^{n-1}$ is identity element. This set of conditions are called the *central path conditions* as they define a trajectory approaching the solution set as $\mu \downarrow 0$. Conventional IPMs usually apply a Newton-type method to the equations in (4) with a suitable line search dealing with constraints $x \in \mathcal{Q}$ and $s \in \mathcal{Q}$ explicitly.

Recently smoothing Newton methods [2, 14, 25, 19, 20, 7, 15, 8, 12] have attracted a lot of attention partially due to their superior numerical performances. However, some algorithms [2, 19] depend on the assumptions of uniform nonsingularity and strict complementarity. Without the uniform nonsingularity assumption, the algorithm given in [27] usually needs to solve two linear systems of equations and to perform at least two line searches at each iteration. Lastly, Qi, Sun and Zhou [20] proposed a class of new smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities under a nonsingularity assumption. The method in [20] was shown to be

locally superlinearly/quadratically convergent without strict complementarity. Moreover, the smoothing methods available are mostly for solving the complementarity problems [2, 3, 19, 20, 7, 8], but there is little work on smoothing methods for the SOCP.

Under certain assumptions, IPMs and smoothing methods are globally convergent in the sense that every limit point of the generated sequence is a solution of optimality conditions (3). However, with the exception of the infeasible IPMs [2, 21, 20], they need a feasible starting point.

Fukushima, Luo and Tseng studied Lipschitzian and differential properties of several typical smoothing functions for second-order cone complementarity problems. They derived computable formulas for their Jacobians, which provide a theoretical foundation for constructing corresponding non-interior point methods. The purpose of this paper is just to present such a non-interior point method for problem (PSOCP), which employs a new smoothing function to characterize the central path conditions. We stress on the demonstration of the global convergence and locally quadratic convergence of the proposed algorithm.

The new smoothing algorithm to be discussed here is based on perturbed optimality conditions (4) and the main difference from IPMs is that we reformulate (4) as a smoothing linear system of equations. It is shown that our algorithm has the following good properties:

(i) The algorithm can start from an arbitrary initial point;

(ii) The algorithm needs to solve only one linear system of equations and perform only one line search at each iteration;

(iii) The algorithm is globally and locally Q-quadratically convergent under mild assumption, without strict complementarity. The result is stronger than the corresponding results for IPMs.

The following notations and terminologies are used throughout the paper. We use "," for adjoining vectors and matrices in a row and ";" for adjoining them in a column. $\mathcal{R}^n$ ($n \geq 1$) denotes the space of n-dimensional real column vectors, and $\mathcal{R}^n \times \mathcal{R}^m$ is identified with $\mathcal{R}^{n+m}$. Denote $x^2 = x \circ x$. For any $x, y \in \mathcal{R}^n$, we write $x \succcurlyeq_Q y$ or $x \preccurlyeq_Q y$ (respectively, $x \succ_Q y$ or $x \prec_Q y$) if

$x - y \in \mathcal{Q}$ or $y - x \in \mathcal{Q}$ (respectively, $x - y \in \text{int}\mathcal{Q}$ or $y - x \in \text{int}\mathcal{Q}$, where $\text{int}\mathcal{Q}$ denotes the interior of $\mathcal{Q}$). $\mathcal{R}_+(\mathcal{R}_{++})$ denotes the set of nonnegative (positive) reals. For $x \in \mathcal{R}^n$ with eigenvalues λ_1 and λ_2, we can define the Frobenius norm

$$\|x\|_F := \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{\text{tr}(x^2)}.$$

Since both eigenvalues of e are equal to one, $\|e\|_F = \sqrt{2}$. For any $x, y \in \mathcal{R}^n$, the Euclidean inner product and norm are denoted by $\langle x, y\rangle = x^T y$ and $\|x\| = \sqrt{x^T x}$ respectively.

The paper is organized as follows. In Section 2, we give the equivalent formulation of the perturbed optimality conditions and some preliminaries. A smoothing function associated with the SOC $\mathcal{Q}$ and its properties are given in Section 3. In Section 4, we describe our algorithm. The convergence of the new algorithm is analyzed in Section 5. Numerical results are shown in Section 6.

2 Preliminaries

For any vector $x = (x_1; x_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$, we define its spectral decomposition associated with SOC $\mathcal{Q}$ as

$$x = \lambda_1 u_1 + \lambda_2 u_2, \tag{5}$$

where the spectral values λ_i and the associated spectral vectors u_i of x are given by

$$\lambda_i = x_1 + (-1)^{i+1}\|x_2\|, \tag{6}$$

$$u_i = \begin{cases} \dfrac{1}{2}\left(1; (-1)^{i+1}\dfrac{x_2}{\|x_2\|}\right), & x_2 \neq 0; \\ \dfrac{1}{2}(1; (-1)^{i+1}\omega), & x_2 = 0, \end{cases} \tag{7}$$

for $i = 1, 2$, with any $\omega \in \mathcal{R}^{n-1}$ such that $\|\omega\| = 1$. If $x_2 \neq 0$, then the decomposition (5) is unique. Some interesting properties of λ_1, λ_2 and u_1, u_2 are summarized below.

Property 2.1. For any $x = (x_1; x_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$, the spectral values λ_1, λ_2 and spectral vectors u_1, u_2 as given by (6) and (7), have the following properties:

(i) $u_1 + u_2 = e$.

(ii) u_1 and u_2 are idempotent under the Jordan product, i.e., $u_i^2 = u_i$, $i = 1, 2$.

(iii) u_1 and u_2 are orthogonal under the Jordan product, i.e., $u_1 \circ u_2 = 0$, and have length $\sqrt{2}/2$.

(iv) λ_1, λ_2 are nonnegative (respectively, positive) if and only if $x \in \mathcal{Q}$ (respectively, $x \in \text{int}\mathcal{Q}$).

Given an element $x = (x_1; x_2) \in \mathcal{R}^n$, we define the arrow-shaped matrix

$$L_x = \begin{pmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{pmatrix},$$

where I represents the $(n-1) \times (n-1)$ identity matrix. It is easy to verify that $x \circ s = L_x \circ s = L_s \circ x = L_x L_s e$ for any $s \in \mathcal{R}^n$. Moreover, L_x is symmetric positive definite if and only if $x \in \text{int}\mathcal{Q}$, i.e., $x \succ_{\mathcal{Q}} 0$.

3 A smoothing function associated with the SOC $\mathcal{Q}$ and its properties

First, let us introduce a smoothing function. In [7], it has been shown that the vector valued Fischer-Burmeister function $\phi_{FB}(x, s) : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}^n$ defined by

$$\phi_{FB}(x, s) = x + s - \sqrt{x^2 + s^2} \tag{8}$$

satisfies the following important property

$$\phi_{FB}(x, s) = 0 \Longleftrightarrow x \in \mathcal{Q}, s \in \mathcal{Q}, x \circ s = 0. \tag{9}$$

The Fischer-Burmeister function has many interesting properties. However, it is typically nonsmooth, because it is not derivable at $(0; 0) \in \mathcal{R} \times \mathcal{R}^{n-1}$, which limits its practical applications. Recently, some smoothing methods are presented, such as the method using Chen-Harker-Kanzow-Smale smoothing function (see [9, 28] and its references therein).

We now smoothing the function ϕ_{FB}, so that we get a characterization of the central path conditions (4). By smoothing the symmetrically perturbed ϕ_{FB}, we obtain the new vector-valued function $\Phi : \mathcal{D} \to \mathcal{R}^n$, defined by

$$\Phi(\mu, x, s) = \frac{1}{2\mu}\left[\mu^2 e - (x - \mu e)^2 - (s - \mu e)^2\right], \tag{10}$$

where $\mathcal{D} = \{(\mu, x, s) \in \mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^n : \mu > \|\sqrt{x^2 + s^2}\|_F\}$.

Proposition 3.1. *For any* $(\mu_1, x, s),\ (\mu_2, x, s) \in \mathcal{D}$,

$$\|\Phi(\mu_1, x, s) - \Phi(\mu_2, x, s)\|_F < |\mu_1 - \mu_2|.$$

Proof. For any $(\mu_1, x, s),\ (\mu_2, x, s) \in \mathcal{D}$, without loss of generality, we assume $\mu_1 \geq \mu_2 > \sqrt{x^2 + s^2}$. Thus, we have

$$\begin{aligned}
&\|\Phi(\mu_1, x, s) - \Phi(\mu_2, x, s)\|_F \\
&= \| - \frac{x^2 + s^2}{2\mu_1} - \frac{\mu_1}{2}e + \frac{x^2 + s^2}{2\mu_2} + \frac{\mu_2}{2}e\|_F \\
&= \|\frac{(\mu_1 - \mu_2)(x^2 + s^2)}{2\mu_1\mu_2} - \frac{\mu_1 - \mu_2}{2}e\|_F \\
&\leq \|\frac{(\mu_1 - \mu_2)(x^2 + s^2)}{2\mu_1\mu_2}\|_F + |\frac{\mu_1 - \mu_2}{2}| \\
&= \frac{|\mu_1 - \mu_2|}{2\mu_1\mu_2}\|x^2 + s^2\|_F + \frac{|\mu_1 - \mu_2|}{2} \\
&\leq \frac{|\mu_1 - \mu_2|}{2\mu_1\mu_2}\|\sqrt{x^2 + s^2}\|_F\|\sqrt{x^2 + s^2}\|_F + \frac{|\mu_1 - \mu_2|}{2} \\
&< \frac{|\mu_1 - \mu_2|}{2\mu_1\mu_2}\mu_1\mu_2 + \frac{|\mu_1 - \mu_2|}{2} \\
&= |\mu_1 - \mu_2|,
\end{aligned}$$

which completes the proof. □

As we will show, the function $\Phi(\mu, x, s)$ have many good properties that can be used to characterize the central path conditions (4). $\Phi(\mu, x, s)$ is smooth

for any $(\mu, x, s) \in \mathcal{D}$. This property plays an important role in the analysis of the quadratic convergence of our smoothing Newton method. Semismoothness is a generalization concept of smoothness, which was originally introduced in [15] and then extended by L. Qi in 1993. Semismooth functions include smooth functions, piecewise smooth functions, convex and concave functions, etc. The composition of (strongly) semismooth functions is still a (strongly) semismooth function.

Definition 3.1. *Suppose that $H : \mathcal{R}^n \to \mathcal{R}^m$ is a locally Lipschitz continuous function. H is said to be semismooth at $x \in \mathcal{R}^n$ if H is directionally differentiable at x and for any $V \in \partial H(x + \triangle x)$*

$$H(x + \triangle x) - H(x) - V(\triangle x) = o(\|\triangle x\|).$$

H is said to be p-order $(0 < p < \infty)$ semismooth at x if H is semismooth at x and

$$H(x + \triangle x) - H(x) - V(\triangle x) = O(\|\triangle x\|^{1+p}).$$

In particular, H is called strongly semismooth at x if H is 1-order semismooth at x.

The following concept of a smoothing function of a nondifferentiable function was introduced by Hayashi, Yamashita and Fukushima [8].

Definition 3.2. *A function $H : \mathcal{R}^n \to \mathcal{R}^m$ is said to be a semismooth (respectively, p-order semismooth) function if it is semismooth (respectively, p-order semismooth) everywhere in $\mathcal{R}^n$.*

In fact, the function $\Phi(\mu, x, s)$ given by (10) is a smoothing function of $\phi_{FB}(x, s)$. Thus, we can solve a family of smoothing subproblems $\Phi(\mu, x, s) = 0$ for $\mu > 0$ and obtain a solution of $\Phi_{FB}(x, s) = 0$ by letting $\mu \downarrow 0$.

Definition 3.3 [8]**.** *For a nondifferentiable function $g : \mathcal{R}^n \to \mathcal{R}^m$, we consider a function $g_\mu : \mathcal{R}^n \to \mathcal{R}^m$ with a parameter $\mu > 0$ that has the following properties:*

(i) *g_μ is differentiable for any $\mu > 0$;*

(ii) $\lim_{\mu \downarrow 0} g_\mu(x) = g(x)$ *for any* $x \in \mathcal{R}^n$.

Such a function g_μ is called a smoothing function of g.

Theorem 3.1. *For any* $x, s \in \mathcal{R}^n$, *let* $w := \sqrt{x^2 + s^2}$ *whose spectral decomposition associated with SOC* $\mathcal{Q}$ *is* $w = \lambda_1 u_1 + \lambda_2 u_2$. *If* $\mu > \|\sqrt{x^2 + s^2}\|_F$, *then the following results hold:*

$$\text{(i) } w \succcurlyeq 0; \quad \text{(ii) } \mu e \succ w; \quad \text{(iii) } \mu^2 e \succ w^2.$$

Proof. Assume the spectral decomposition of w associated with SOC $\mathcal{Q}$ is $w = \lambda_1 u_1 + \lambda_2 u_2$. From $\mu > \|\sqrt{x^2 + s^2}\|_F$ we have

$$\mu > \|\sqrt{x^2 + s^2}\|_F = \|\lambda_1 u_1 + \lambda_2 u_2\|_F = \sqrt{\lambda_1^2 + \lambda_2^2} \geq \max\{\lambda_1, \lambda_2\}. \tag{11}$$

Thus, we have $0 \leq \lambda_i < \mu, i = 1, 2$, which means that (i) holds. By

$$0 \leq \lambda_i < \mu, i = 1, 2, \ \mu e - w = (\mu - \lambda_1) u_1 + (\mu - \lambda_2) u_2 \succ 0,$$

which yields (ii), and $\mu^2 e - w^2 = (\mu^2 - \lambda_1^2) u_1 + (\mu^2 - \lambda_2^2) u_2 \succ 0$, which gives (iii). □

Now, we give the main properties of $\Phi(\mu, x, s)$:

Theorem 3.2. (i) $\Phi(\mu, x, s)$ *is globally Lipschitz continuous for any* $(\mu, x, s) \in \mathcal{D}$. *Moreover,* $\Phi(\mu, x, s)$ *is continuously differentiable at any* $(\mu, x, s) \in \mathcal{D}$ *with its Jacobian*

$$\Phi'(\mu, x, s) = \begin{pmatrix} \frac{1}{2\mu^2}\left(x^2 + s^2 - \mu^2 e\right) \\ I - \frac{1}{\mu} L_x \\ I - \frac{1}{\mu} L_s \end{pmatrix}. \tag{12}$$

(ii) $\lim_{\mu \downarrow 0} \Phi(\mu, x, s) = \phi_{FB}(x, s)$ *for any* $(x, s) \in \mathcal{R}^n \times \mathcal{R}^n$. *Thus,* $\Phi(\mu, x, s)$ *is a smoothing function of* $\Phi_{FB}(x, s)$.

Proof. (i) It is not difficult to show that $\Phi(\mu, x, s)$ is globally Lipschitz continuous, and continuously differentiable at any $(\mu, x, s) \in \mathcal{D}$. Now we prove (12). For any $(\mu, x, s) \in \mathcal{D}$, from (10) and by simply calculation, we have

$$\begin{aligned}
\Phi_{\mu}^{'}(\mu, x, s) &= \frac{1}{2\mu^2}(x^2 + s^2 - \mu^2 e), \\
\Phi_{x}^{'}(\mu, x, s) &= I - \frac{1}{\mu} L_x, \\
\Phi_{s}^{'}(\mu, x, s) &= I - \frac{1}{\mu} L_s,
\end{aligned}$$

which yield (12).

Next, we prove (ii). For any $x = (x_1; x_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$ and $s = (s_1; s_2) \in \mathcal{R} \times \mathcal{R}^{n-1}$, denote $w = \sqrt{x^2 + s^2}$ whose spectral decomposition associated with SOC $\mathcal{Q}$ is $w = \lambda_1 u_1 + \lambda_2 u_2$. From Theorem 3.1, we have $w \succcurlyeq 0$, and $0 \leq \lambda_i \leq \mu, i = 1, 2$. Therefore, we have

$$\begin{aligned}
&\|\Phi(\mu, x, s) - \phi_{FB}(x, s)\|_F \\
&= \left\| \frac{1}{2\mu}\left[\mu^2 e - (x - \mu e)^2 - (s - \mu e)^2\right] - x - s + \sqrt{x^2 + s^2} \right\|_F \\
&= \left\| \frac{1}{2\mu}(x^2 + s^2 + \mu^2 e) - \sqrt{x^2 + s^2} \right\|_F \\
&= \left\| \frac{1}{2\mu}(\sqrt{x^2 + s^2} - \mu e)^2 \right\|_F \\
&= \frac{1}{2\mu} \left\| [(\lambda_1 - \mu)u_1 + (\lambda_2 - \mu)u_2]^2 \right\|_F \\
&= \frac{1}{4\mu}\left[(\lambda_1 - \mu)^2 + (\lambda_2 - \mu)^2\right] \\
&\leq \frac{1}{4\mu}\left(\mu^2 + \mu^2\right) \\
&= \frac{1}{2}\mu \to 0, \text{ when } \mu \downarrow 0.
\end{aligned}$$

Thus, we have $\lim_{\mu \downarrow 0} \Phi(\mu, x, s) = \phi_{FB}(x, s)$. Therefore, it follows from (i) and Definition 3.3 that $\Phi(\mu, x, s)$ is a smoothing function of $\phi_{FB}(x, s)$. □

4 Description of the algorithm

Based on the smoothing function (10) introduced in the previous section, the aim of this section is to propose the smoothing Newton-type algorithm for the SOCP and show the well-definedness of it under suitable assumptions.

Let $z := (\mu, x, c - A^T y) \in \mathcal{D}$. By using the smoothing function (10), we define the function $G(\mu, x, y) : \mathcal{R} \times \mathcal{R}^n \times \mathcal{R}^m \to \mathcal{R} \times \mathcal{R}^m \times \mathcal{R}^n$ by

$$G(z) := \begin{pmatrix} \mu \\ Ax - b \\ \Phi(\mu, x, c - A^T y) \end{pmatrix}. \tag{13}$$

In view of (9) and (13), $z^* := (\mu^*, x^*, y^*)$ is a solution of the system $G(z) = 0$ if and only if $(x^*, y^*, c - A^T y^*)$ solves the optimality conditions (3).

It is well-known that problems (PSOCP) and (DSOCP) are equivalent to (13) [1, 20]. Therefore, z^* is a solution of $G(z) = 0$ if and only if $(x^*, y^*, c - A^T y^*)$ is the optimal solution of (PSOCP) and (DSOCP). Then we can apply Newton's method to the nonlinear system of equations $G(z) = 0$.

Let $\gamma \in (0, 1)$ and define the function $\beta : \mathcal{R}^{n+m+1} \to \mathcal{R}_+$ by

$$\beta(z) := \gamma \min\left\{1, \ \|G(z)\|^2\right\}. \tag{14}$$

Now we are in the position to give a formal description of our algorithm.

Algorithm 4.1. (*A smoothing Newton-type method for* SOCP).

Step 0. Choose constants $\delta \in (0, 1), \sigma \in (0, 1)$, and $\mu_0 \in \mathcal{R}_{++}$, and let $\bar{z} := (\mu_0, 0, 0) \in \mathcal{D}$. Let $(x_0, y_0) \in \mathcal{R}^n \times \mathcal{R}^m$ be arbitrary initial point and $z_0 := (\mu_0, x_0, y_0)$. Choose $\gamma \in (0, 1)$ such that $\gamma\mu_0 < 1/2$. Set $k := 0$.

Step 1. If $G(z_k) = 0$, then stop. Else, let $\beta_k := \beta(z_k)$.

Step 2. Compute $\Delta z_k := (\Delta\mu_k, \Delta x_k, \Delta y_k)$ by solving the following system of linear equations

$$G(z_k) + G'(z_k)\Delta z_k = \beta_k \bar{z}. \tag{15}$$

Step 3. Let ν_k be the smallest nonnegative integer ν such that

$$\|G(z_k + \delta^\nu \Delta z_k)\|^2 \leq \left[1 - \sigma(1 - 2\gamma\mu_0)\delta^\nu\right] \|G(z_k)\|^2. \tag{16}$$

Let $\lambda_k := \delta^{\nu_k}$.

Step 4. Set $z_{k+1} := z_k + \lambda_k \Delta z_k$ and $k := k + 1$. Go to step 1.

In order to analyze our algorithm, we study the Lipschitzian, smoothness and differential properties of the function $G(z)$ given by (13). Moreover, we derive the computable formula for the Jacobian of the function $G(z)$ and give the condition for the Jacobian to be nonsingular.

Throughout the rest of this paper, we make the following assumption:

Assumption 4.1. The matrix A has full row rank, i.e., all the row vectors of A are linearly independent.

Lemma 4.1 [7]. *For any $x, s \in \mathcal{R}^n$ and any $\omega \succ_Q 0$, we have*

$$\omega^2 \succ_Q x^2+s^2 \Rightarrow L_\omega - L_x \succ 0,\ L_\omega - L_s \succ 0,\ (L_\omega - L_x)(L_\omega - L_s) \succ 0. \tag{17}$$

Moreover, (17) *remains true when "$\succ$" is replaced by "$\succcurlyeq$" everywhere.*

Theorem 4.1. *Let $z := (\mu, x, y) \in \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{D}$ be defined by* (13). *Then the following results hold.*

(i) *G is globally Lipschitz continuous, and continuously differentiable at any $z := (\mu, x, y) \in \mathcal{D}$ with its Jacobian*

$$G'(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & A & 0 \\ I - \frac{1}{2\mu^2}\left[x^2 + (c - A^T y)^2 - \mu^2 e\right] & I - \frac{1}{\mu} L_x & -\left(I - \frac{1}{\mu} L_{(c - A^T y)}\right) A^T \end{pmatrix}. \tag{18}$$

(ii) *Under Assumption* 4.1, *$G'(z)$ is nonsingular for any $\mu > 0$.*

Proof. By Theorem 3.1, we can easily show that (i) holds. Now we prove (ii). For any fixed $\mu > 0$, let

$$\Delta z := (\Delta\mu, \Delta x, \Delta y) \in \mathcal{R} \times \mathcal{R}^n \times \mathcal{R}^m.$$

It is sufficient to prove that the linear system of equations

$$G'(z)\Delta z = 0 \tag{19}$$

has only zero solution, i.e., $\Delta\mu = 0$, $\Delta x = 0$ and $\Delta y = 0$. By (18) and (19), we have

$$\Delta\mu = 0, \tag{20}$$

$$A\Delta x = 0, \tag{21}$$

$$\left(I - \frac{1}{\mu}L_x\right)\Delta x - \left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)A^T\Delta y = 0. \tag{22}$$

Since

$$e^2 \succ \left(\frac{x}{\mu}\right)^2 + \left(\frac{c - A^Ty}{\mu}\right)^2,$$

It follows from Lemma 4.1 that

$$I - \frac{1}{\mu}L_x \succ 0,\ I - \frac{1}{\mu}L_{(c-A^Ty)} \succ 0,\ \left(I - \frac{1}{\mu}L_x\right)\left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right) \succ 0. \tag{23}$$

Premultiplying (23) by $\Delta x^T\left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)^{-1}$ and taking into account $A\Delta x = 0$, we have

$$\Delta x^T\left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)^{-1}\left(I - \frac{1}{\mu}L_x\right)\Delta x = 0. \tag{24}$$

Denote

$$\Delta\tilde{x} = \left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)^{-1}\Delta x. \tag{25}$$

From (24), we obtain

$$\Delta\tilde{x}^T\left(I - \frac{1}{\mu}L_x\right)\left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)\Delta\tilde{x} = 0. \tag{26}$$

By (23), $\left(I - \frac{1}{\mu}L_x\right)\left(I - \frac{1}{\mu}L_{(c-A^Ty)}\right)$ is positive definite. Therefore, (26) yields $\Delta\tilde{x} = 0$, and it follows from (25) that $\Delta x = 0$. Since A has full row rank, (21) implies $\Delta y = 0$. Thus the linear system of equations (19) has only zero solution, and hence $G'(z)$ is nonsingular. Thus, the proof is completed. □

By Theorem 4.1, we can show that Algorithm 4.1 is well-defined.

Theorem 4.2. *Suppose that Assumption* 4.1 *holds. If* $\mu_k > 0$, *then Algorithm* 4.1 *is well-defined for any* $k \geq 0$.

Proof. Since A has full row rank, it follows from Theorem 4.1 that $G'(z_k)$ is nonsingular for any $\mu_k > 0$. Therefore, Step 2 is well-defined at the kth iteration. Then by following the proof of Lemma 5 in [20], we can show the well-definedness of Step 3. The proof is completed. □

5 Convergence analysis

In this section, we analyze the global and local convergence properties of Algorithm 4.1. It is shown that any accumulation point of the iteration sequence is a solution of the system $G(z) = 0$. If the accumulation point z^* satisfies a nonsingularity assumption, then the iteration sequence converges to z^* locally Q-quadratically without any strict complementarity assumption. To show the global convergence of Algorithm 4.1, we need the following Lemma (see [20], Proposition 6).

Lemma 5.1. *Suppose that Assumption* 4.1 *holds. For any* $\tilde{z} := (\tilde{u}, \tilde{x}, \tilde{y}) \in \mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^m$, *and* $G'(\tilde{z})$ *is nonsingular, then there exists a closed neighborhood* $\mathcal{N}(\tilde{z})$ *of* $\tilde{z}$ *and a positive number* $\bar{\alpha} \in (0, 1]$ *such that for any* $z = (u, x, y) \in \mathcal{N}(\tilde{z})$ *and all* $\alpha \in [0, \bar{\alpha}]$, *we have* $u \in \mathcal{R}_{++}$, $G'(z)$ *is invertible and*

$$\|G(z + \alpha\Delta z)\|^2 \leq \left[1 - \sigma(1 - 2\gamma\mu_0)\alpha\right] \|G(z)\|^2. \tag{27}$$

Theorem 5.1. *Suppose that Assumption* 4.1 *holds and that* $\{z_k\}$ *is the iteration sequence generated by Algorithm* 4.1. *Then the following results hold.*

(i) $\mu_k \in \mathcal{R}_{++}$ *and* $z_k \in \Theta$ *for any* $k \geq 0$, *where*

$$\Theta = \{z = (\mu, x, y) \in \mathcal{D} : \mu \geq \beta(z)\mu_0\}. \tag{28}$$

(ii) *Any accumulation point* $z^* := (\mu^*, x^*, y^*)$ *of* $\{z_k\}$ *is a solution of* $G(z) = 0$.

Proof. Suppose that $\mu_k > 0$. It follows from (15) and Step 4 that

$$\Delta\mu_k = -\mu_k + \beta_k\mu_0, \tag{29}$$

$$\mu_{k+1} = \mu_k + \lambda_k\Delta\mu_k. \tag{30}$$

Substituting (29) into (30), we have

$$\mu_{k+1} = \mu_k - \lambda_k \mu_k + \lambda_k \beta_k \mu_0 = (1 - \lambda_k)\mu_k + \lambda_k \beta_k \mu_0 > 0, \tag{31}$$

which, together with $\mu_0 > 0$ and $\lambda_k = \delta^{\nu_k} \in (0, 1)$ implies that $\mu_k \in \mathcal{R}_{++}$ for any $k \geq 0$.

Now we prove $z_k \in \Theta$ for any $k \geq 0$ by induction. Since

$$\beta_0 = \beta(z_0) = \gamma \min\left\{1, \ \|G(z_0)\|^2\right\} \leq \gamma \in (0, 1),$$

it is easy to see that $z_0 \in \Theta$. Suppose that $z_k \in \Theta$, then

$$\mu_k \geq \beta_k \mu_0. \tag{32}$$

We consider the following two cases:

Case (I): If $\|G(z_k)\| > 1$, then

$$\beta_k = \gamma. \tag{33}$$

Since $\beta_{k+1} = \gamma \min\{1, \ \|G(z_{k+1})\|^2\} \leq \gamma$, it follows from (16), (31), (32) and (33) that

$$\mu_{k+1} - \beta_{k+1}\mu_0 \geq (1 - \lambda_k)\beta_k \mu_0 + \lambda_k \beta_k \mu_0 - \gamma \mu_0 = \beta_k \mu_0 - \gamma \mu_0 = 0. \tag{34}$$

Case (II): If $\|G(z_k)\| \leq 1$, then

$$\beta_k = \gamma \|G(z_k)\|^2. \tag{35}$$

By (16), we have $\|G(z_{k+1})\| \leq \|G(z_k)\| \leq 1$. From (31), (35), and taking into account $\beta_{k+1} = \gamma \|G(z_{k+1})\|^2$, we have

$$\begin{aligned} \mu_{k+1} - \beta_{k+1}\mu_0 &= (1 - \lambda_k)\mu_k + \lambda_k \beta_k \mu_0 - \gamma \mu_0 \|G(z_{k+1})\|^2 \\ &\geq (1 - \lambda_k)\beta_k \mu_0 + \lambda_k \beta_k \mu_0 - \gamma \mu_0 \|G(z_k)\|^2 \\ &= \beta_k \mu_0 - \gamma \mu_0 \|G(z_k)\|^2 \\ &= 0. \end{aligned} \tag{36}$$

Combining (34) and (36) yields that $z_k \in \Theta$ for any $k \geq 0$.

Now, we prove (ii). Without loss of generality, we assume that $\{z_k\}$ converges to z^* as $k \to +\infty$. Since $\{\|G(z_k)\|\}$ is monotonically decreasing and bounded from below, it follows from the continuity of $G(\cdot)$ that $\{\|G(z_k)\|\}$ converges to a nonnegative number $G(z^*)$. Then by the definition of $\beta(\cdot)$, we obtain that $\{\beta_k\}$ converges to

$$\beta^* = \gamma \min\left\{1,\ \|G(z^*)\|^2\right\}.$$

It follows from (15) and Theorem 5.1 (i) that

$$0 < \mu_{k+1} = (1-\lambda_k)\mu_k + \lambda_k\beta_k\mu_0 \le \mu_k,$$

which implies that $\{\mu_k\}$ converges to μ^*. If $\|G(z^*)\| = 0$, then we obtain the desired result. In the following, we suppose $\|G(z^*)\| > 0$. By Lemma 4.1, $0 < \beta^*\mu_0 \le \mu^*$. It follows from Theorem 4.1 that $G'(z^*)$ exists and it is invertible. Hence, by Lemma 5.1, there exists a closed neighborhood $\mathcal{N}(\tilde{z})$ of $\tilde{z}$ and a positive number $\bar{\alpha} \in (0, 1]$ such that for any $z = (\mu, x, y) \in \mathcal{N}(\tilde{z})$ and all $\alpha \in [0, \bar{\alpha}]$, we have $\mu \in \mathcal{R}_{++}$, $G'(z)$ is invertible and

$$\|G(z + \alpha\Delta z)\|^2 \le \left[1 - \sigma(1 - 2\gamma\mu_0)\alpha\right]\|G(z)\|^2. \tag{37}$$

Therefore, for a nonnegative integer $\bar{\nu}$ such that $\delta^{\bar{\nu}} \in (0, \bar{\alpha}]$, for all sufficiently large k, we have

$$\|G(z_k + \delta^{\bar{\nu}}\Delta z_k)\|^2 \le \left[1 - \sigma(1 - 2\gamma\mu_0)\delta^{\bar{\nu}}\right]\|G(z_k)\|^2.$$

For all sufficiently large k, since $\lambda_k = \delta^{\nu_k} \ge \delta^{\bar{\nu}}$, it follows from (16) that

$$\begin{aligned}\|G(z_k + \delta^{\bar{\nu}}\Delta z_k)\|^2 &\le \left[1 - \sigma(1 - 2\gamma\mu_0)\lambda_k\right]\|G(z_k)\|^2 \\ &\le \left[1 - \sigma(1 - 2\gamma\mu_0)\delta^{\bar{\nu}}\right]\|G(z_k)\|^2.\end{aligned}$$

This contradicts the fact that the sequence $\{\|G(z_k)\|\}$ converges to $\|G(z^*)\| > 0$. This completes the proof. □

To establish the locally Q-quadratic convergence of our smoothing Newton method, we need the following assumption:

Assumption 5.1. Assume that z^* satisfies the nonsingularity condition, i.e., all $V \in \partial G(z^*)$ are nonsingular.

Now we are in the position to give the rate of convergence for Algorithm 4.1.

Theorem 5.2. *Suppose that Assumption* 4.1 *holds and that* z^* *is an accumulation point of the iteration sequence* $\{z_k\}$ *generated by Algorithm* 4.1. *If Assumption* 5.1 *holds, then:*

(i) $\lambda_k \equiv 1$ *for all* z_k *sufficiently close to* z^**;*

(ii) $\{z_k\}$ *converges to* z^* *Q-quadratically, i.e.,* $\|z_{k+1} - z^*\| = O(\|z_k - z^*\|^2)$*; moreover,* $\mu_{k+1} = O(\mu_k^2)$.

Proof. By using Lemma 3.1 and Theorem 4.1, we can prove the theorem similarly as in Theorem 8 of [20]. For brevity, we omit the details here. □

6 Numerical results

In this section, we conducted some numerical experiments to evaluate the efficiency of Algorithm 4.1. All these experiments were performed on an IBM notebook computer R40e with Intel(R) Pentium(R) 4 CPU 2.00 GHz and 512 MB memory. The operating system was Windows XP SP2 and the implementations were done in MATLAB 7.0.1. For comparison purpose, we also use SDPT3 solver [12] which is an IPM for the SOCP.

For simplicity, we randomly generate six test problems with size $m = 50$ and $n = 100$. To be specific, we generate a random matrix $A \in \mathcal{R}^{m\times n}$ with full row rank and random vectors

$$x \in \text{int}\mathcal{Q},\ s \in \text{int}\mathcal{Q},\ y \in \mathcal{R}^m,\ \text{ and then let } b := Ax \text{ and } c := A^T y + s.$$

Thus the generated problems (PSOCP) and (DSOCP) have optimal solutions and their optimal values coincide, because the set of strictly feasible solutions of (PSOCP) and (DSOCP) are nonempty. Let $x_0 = e \in \mathcal{R}^n$ and $y_0 = 0 \in \mathcal{R}^m$ be initial points. The parameters used in Algorithm 4.1 were as follows:

$$\mu_0 = 0.01,\ \sigma = 0.35,\ \delta = 0.65 \text{ and } \gamma = 0.90.$$

We used $\|H(z)\| \le 10^{-5}$ as the stopping criterion.

The results in Table 1 indicate that Algorithm 4.1 performs very well. We also obtained similar results for other random examples.

	SDPT3		Algorithm 4.1	
problem	IT	CPU(s)	IT	CPU(s)
problem 1	8	0.2	6	0.09
problem 2	7	0.1	6	0.06
problem 3	9	0.2	7	0.08
problem 4	8	0.1	5	0.06
problem 5	9	0.1	7	0.06
problem 6	8	0.2	5	0.07

Table 1 – Comparison of Algorithm 4.1 and SDPT3 on SOCPs.

Acknowledgments. The work is supported by National Natural Science Foundation of China (10571109, 10971122), Natural Science Foundation of Shandong (Y2008A01), Scientific and Technological Project of Shandong Province (2009GG10001012), the Excellent Young Scientist Foundation of Shandong Province (BS2011SF024), and Project of Shandong Province Higher Educational Science and Technology Program (J10LA51). The authors would like to thank the anonymous referees for their valuable comments and suggestions on the paper, which have considerably improved the paper.

REFERENCES

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

[2] B. Chen and N. Xiu, *A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions*. SIAM Journal on Optimization, **9** (1999), 605–623.

[3] J. Chen, *Two classes of merit functions for the second-order cone complementarity problem*. Math. Meth. Oper. Res., **64** (2006), 495–519.

[4] R. Debnath, M. Muramatsu and H. Takahashi, *An Efficient Support Vector Machine Learning Method with Second-Order Cone Programming for Large-Scale Problems*. Applied Intelligence, **23** (2005), 219–239.

[5] J. Faraut and A. Koranyi, *Analysis on symmetric cones*. Oxford University Press, London and New York, 1994, ISBN: 0-198-53477-9.

[6] L. Faybusovich, *Euclidean Jordan Algebras and Interior-point Algorithms*. Positivity, **1** (1997), 331–357.

[7] M. Fukushima, Z. Luo and P. Tseng, *Smoothing functions for second-order-cone complimentarity problems*. SIAM J. Optim., **12**(2) (2002), 436–460.

[8] S. Hayashi, N. Yamashita and M. Fukushima, *A combined smoothing and regularized method for monotone second-order cone complementarity problems*. SIAM Journal on Optimization, **15** (2005), 593–615.

[9] K. Hotta and A. Yoshise, *Global convergence of a class of non-interior-point algorithms using Chen-Harker-Kanzow functions for nonlinear complementarity problems*. Discussion Paper Series No. 708, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305, Japan, December 1996. CMP 98:13.

[10] Y. Kanno and M. Ohsaki, *Contact Analysis of Cable Networks by Using Second-Order Cone Programming*. SIAM J. Sci. Comput., **27**(6) (2006), 2032–2052.

[11] Y.J. Kuo and H.D. Mittelmann, *Interior point methods for second-order cone programming and OR applications*. Computational Optimization and Applications, **28** (2004), 255–285.

[12] Y. Liu, L. Zhang and Y. Wang, *Analysis of a Smoothing Method for Symmetric Cone Linear Programming*. J. Appl. Math. and Computing, **22**(1-2) (2006), 133–148.

[13] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, *Applications of second-order cone programming*. Linear Algebra and its Applications, **284** (1998), 193–228.

[14] C. Ma and X. Chen, *The convergence of a one-step smoothing Newton method for P_0-NCP based on a new smoothing NCP-function*. Journal of Computational and Applied Mathematics, **216** (2008), 1–13.

[15] R. Miffin, *Semismooth and semiconvex functions in constrained optimization*. SIAM Journal on Control and Optimization, **15** (1977), 957–972.

[16] Y.E. Nesterov and M.J. Todd, *Primal-dual Interior-point Methods for Self-scaled Cones*. SIAM J. Optim., **8**(2) (1998), 324–364.

[17] J. Peng, C. Roos and T. Terlaky, *Primal-dual Interior-point Methods for Second-order Conic Optimization Based on Self-Regular Proximities*. SIAM J. Optim., **13**(1) (2002), 179–203.

[18] X. Peng and I. King, *Robust BMPM training based on second-order cone programming and its application in medical diagnosis*. Neural Networks, **21** (2008), 450–457.

[19] L. Qi and D. Sun, *Improving the convergence of non-interior point algorithm for nonlinear complementarity problems*. Mathematics of Computation, **69** (2000), 283–304.

[20] L. Qi, D. Sun and G. Zhou, *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities*. Mathematical Programming, **87** (2000), 1–35.

[21] B.K. Rangarajan, *Polynomial Convergence of Infeasible-Interior-Point Methods over Symmetric Cones*. SIAM J. Optim., **16**(4) (2006), 1211–1229.

[22] T. Sasakawa and T. Tsuchiya, *Optimal Magnetic Shield Design with Second-order Cone Programming*. SIAM J. Sci. Comput., **24**(6) (2003), 1930–1950.

[23] S.H. Schmieta and F. Alizadeh, *Extension of primal-dual interior point algorithms to symmetric cones*, Mathematical Programming, Series A **96** (2003), 409–438.

[24] P.K. Shivaswamy, C. Bhattacharyya and A.J. Smola, *Second Order Cone Programming Approaches for Handling Missing and Uncertain Data*. Journal of Machine Learning Research, **7** (2006), 1283–1314.

[25] D. Sun and J. Sun, *Strong Semismoothness of the Fischer-Burmeister SDC and SOC Complementarity Functions*. Math. Program., Ser. A **103** (2005), 575–581.

[26] K.C. Toh, R.H. Tütüncü and M.J. Todd, *SDPT3 Version 3.02-A MATLAB software for semidefinite-quadratic-linear programming*, http://www.math.nus.edu.sg/∼mattohkc/sdpt3.html, 2002.

[27] P. Tseng, *Error bounds and superlinear convergence analysis of some Newton-type methods in optimization*, in: G. Di Pillo, F. Giannessi (Eds.), Nonlinear Optimization and Related Topics, Kluwer Academic Publishers, Boston, (2000), pp. 445–462.

[28] S. Xu, *The global linear convergence and complexity of a non-interior path-following algorithm for monotone LCP based on Chen-Harker-Kanzow-Smale smooth functions*. Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, February (1997).

[29] S. Yan and Y. Ma, *Robust supergain beamforming for circular array via second-order cone programming*. Applied Acoustics, **66** (2005), 1018–1032.