A new merit function and its related properties for the second-order cone complementarity problem

Jein-Shan Chen [1]

Department of Mathematics

National Taiwan Normal University

Taipei, Taiwan 11677

June 21, 2005

(revised October 6, 2005)

(second revised November 15, 2005)

Abstract Recently, J.-S. Chen and P. Tseng extended two merit functions for the nonlinear complementarity problem (NCP) and the semidefinite complementarity problem (SDCP) to the second-order cone commplementarity problem (SOCCP) and showed several favorable properties. In this paper, we extend a merit function for the NCP studied by Yamada, Yamashita, and Fukushima to the SOCCP and show that the SOCCP is equivalent to an unconstrained smooth minimization via this new merit function. Furthermore, we study conditions under which the new merit function provides a global error bound which plays an important role in analyzing the convergence rate of iterative methods for solving the SOCCP; and conditions under which the new merit function has bounded level sets which ensures that the sequence generated by a descent method has at least one accumulation point.

Key words. Second-order cone, complementarity, merit function, error bound, bounded level sets

1 Introduction

In this paper, we consider the natural extension of nonlinear complementarity problem (NCP), the second-order cone complementarity problem (SOCCP), that is, finding $\zeta \in \mathbb{R}^n$ satisfying

$$\langle F(\zeta), \zeta \rangle = 0, \quad F(\zeta) \in \mathcal{K}, \quad \zeta \in \mathcal{K}, \tag{1}$$

---

[1] E-mail: jschen@math.ntnu.edu.tw, TEL: 886-2-29325417, FAX: 886-2-29332342.

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, $F : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth (i.e., continuously differentiable) mapping, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOC), also called Lorentz cones [9]. In other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_m}, \tag{2}$$

where $m, n_1, \ldots, n_m \geq 1$, $n_1 + \cdots + n_m = n$, and

$$\mathcal{K}^{n_i} := \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n_i-1} \mid \|x_2\| \leq x_1\}, \tag{3}$$

with $\| \cdot \|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative reals $\mathbb{R}_+$. A special case of (2) is $\mathcal{K} = \mathbb{R}_+^n$, the nonnegative orthant in $\mathbb{R}^n$, which corresponds to $m = n$ and $n_1 = \cdots = n_m = 1$. If $\mathcal{K} = \mathbb{R}_+^n$, then (1) reduces to the nonlinear complementarity problem (NCP). The NCP plays a fundamental role in optimization theory and has many applications in engineering and economics; see, e.g., [7, 10, 11, 12]. Throughout this paper, we assume $\mathcal{K} = \mathcal{K}^n$ for simplicity, i.e., $\mathcal{K}$ is a single second-order cone (all the analysis can be easily carried over to the general case where $\mathcal{K}$ has the direct product structure (2)).

There have been proposed various methods for solving SOCCP. They include interior-point methods [1, 20, 21, 22, 25], and (non-interior) smoothing Newton methods [6, 16, 17]. In the recent paper [3], an alternative approach based on reformulating SOCCP as an unconstrained smooth minimization problem was studied. In particular, they were finding a smooth function $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ such that

$$\psi(x, y) = 0 \quad \Longleftrightarrow \quad x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y \rangle = 0. \tag{4}$$

We call such a ψ a *merit function*. Then, the SOCCP can be expressed as an unconstrained smooth (global) minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} \ \psi(F(\zeta), \zeta). \tag{5}$$

Various gradient methods such as conjugate gradient methods and quasi-Newton methods [2, 15] can be applied to slove (5). There have some advantages for this approach as explained in [3]. For this approach to be effective, the choice of ψ is crucial. In the case of NCP, corresponding to (1) when $\mathcal{K} = \mathbb{R}_+^n$, a popular choice is

$$\psi(x, y) \ = \ \frac{1}{2} \sum_{i=1}^n \phi(x_i, y_i)^2$$

for all $x = (x_1, ..., x_n)^T \in \mathbb{R}^n$, where ϕ is the well-known Fischer-Burmeister (FB) NCP-function [13, 14] defined by

$$\phi(x_i, y_i) = \sqrt{x_i^2 + y_i^2} - x_i - y_i.$$

It has been shown that ψ is smooth (even though ϕ is not differentiable) and satisfies (4), see [8, 18, 19]. In the case of SOCCP, for any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their *Jordan product* [3, 9] associated with $\mathcal{K}^n$ as

$$x \circ y := (\langle x, y \rangle,\ y_1 x_2 + x_1 y_2). \tag{6}$$

The identity element under this product is $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual componentwise addition of vectors. It is known that $x^2 \in \mathcal{K}^n$ for all $x \in \mathbb{R}^n$. Moreover, for $x \in \mathcal{K}^n$, there exists a unique vector in $\mathcal{K}^n$, denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Thus, the Fischer-Burmeister function associated with second-order cone (SOC),

$$\phi_{\text{FB}}(x, y) := (x^2 + y^2)^{1/2} - x - y, \tag{7}$$

is well-defined for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ and maps $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. It was shown in [16] that $\phi_{\text{FB}}(x, y) = 0$ if and only if $x \in \mathcal{K}^n$, $y \in \mathcal{K}^n$, $\langle x, y \rangle = 0$. Hence, $\psi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ given by

$$\psi_{\text{FB}}(x, y) := \frac{1}{2} \|\phi_{\text{FB}}(x, y)\|^2, \tag{8}$$

is a merit function for SOCCP. It was also shown in the paper [3] that, like the NCP case, ψ_{FB} is smooth and, when ∇F is positive semi-definite, every stationary point of (5) solves SOCCP. For SDCP, which is a natural extension of NCP where $\mathbb{R}^n_+$ is replaced by the cone of positive semi-definite matrices $\mathcal{S}^n_+$ and the partial order $\leq$ is also changed by $\preceq_{\mathcal{S}^n_+}$ (a partial order associated with $\mathcal{S}^n_+$ where $A \preceq_{\mathcal{S}^n_+} B$ means $B - A \in \mathcal{S}^n_+$) accordingly, the above features hold for the following analog of the SDCP merit function studied by Yamashita and Fukushima [27]:

$$\psi_{\text{YF}}(x, y) := \psi_1(\langle x, y \rangle) + \psi_{\text{FB}}(x, y), \tag{9}$$

where $\psi_1 : \mathbb{R} \to \mathbb{R}_+$ is any smooth function satisfying

$$\psi_1(t) = 0 \ \ \forall t \leq 0 \qquad \text{and} \qquad \psi_1'(t) > 0 \ \ \forall t > 0. \tag{10}$$

In [27], $\psi_1(t) = \frac{1}{4}(\max\{0, t\})^4$ was considered. In fact, the function ψ_{YF}, which was recently studied in [3], is also a SOCCP version merit function that enjoys favorable properties as what ψ_{FB} has and possesses additional properties including bounded level sets and error bound.

In this paper, we make a slight modification of ψ_{YF}, for which ψ_1 is replaced by the mapping $\psi_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ that is given by

$$\psi_0(x, y) := \frac{1}{2} \|(x \circ y)_+\|^2, \tag{11}$$

where $(\cdot)_+$ denotes the orthogonal projection onto $\mathcal{K}^n$. If we observe closely, we may see there is some relation between ψ_0 and ψ_1 : both are smooth functions. Moreover, if we

let $\hat{\psi}_1 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be $\hat{\psi}_1(x, y) := \psi_1(x, y)$, then the graphs of ψ_0 and $\hat{\psi}_1$ share similar features. In other words, our new merit function $\psi_{\text{new}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is defined as

$$\psi_{\text{new}}(x, y) := \alpha\, \psi_0(x, y) + \psi_{\text{FB}}(x, y), \tag{12}$$

where $\alpha > 0$. When $\alpha = 0$, ψ_{new} reduces to ψ_{FB} which is the squared norm of Fischer-Burmeister function (8) studied in [3]. Thus, this new merit function can be viewed as the extension of the squared norm of Fischer-Burmeister function. We will show that the SOCCP is equivalent to the following global minimization via the new merit function ψ_{new}:

$$\min_{\zeta \in \mathbb{R}^n} f(\zeta) \quad \text{where} \quad f(\zeta) := \psi_{\text{new}}(F(\zeta), \zeta). \tag{13}$$

Indeed, this new merit function ψ_{new} was studied by K. Yamada, N. Yamashita, and M. Fukushima in [26] for the NCP case. We are motivated by their work and wish to explore its extension to the SOCCP. Analogous to the additional properties that ψ_{YF} (given as (9)-(10)) possesses and as will be seen in Sec. 4, if F is strongly monotone [7] then f provides a global error bound which plays an important role in analyzing the convergence rate of some iterative methods for solving the SOOCP; and if F is monotone and a strictly feasible solution exists then f has bounded level sets which will ensure that the sequence generated by a descent algorithm has at least one accumulation point. All these properties will make it possible to construct a descent algorithm for solving the equivalent unconstrained reformulation of the SOCCP. In contrast, the merit function induced by ψ_{FB} lacks these properties. In addition, we will show that ψ_{new} is continuously differentiable and its gradient has a computable formula. All the aforementioned features are significant reasons for choosing and studying this new merit function ψ_{new}.

It is known that SOCCP can be reduced to an SDCP by observing that, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we have $x \in \mathcal{K}^n$ if and only if

$$L_x := \left[\begin{array}{cc} x_1 & x_2^T \\ x_2 & x_1 I \end{array} \right]$$

is positive semi-definite (also see [16, p. 437] and [23]). However, this reduction increases the problem dimension from n to $n(n+1)/2$ and it is not known whether this increase can be mitigated by exploiting the special "arrow" structure of L_x.

Throughout this paper, $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping $F = (F_1, ..., F_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \ \cdots \ \nabla F_m(x)]$ is a n by m matrix which denotes the transpose Jacobian of F at x. Also, we let $\mathcal{C}^* := \{y \mid \langle x, y \rangle \geq 0 \ \forall x \in \mathcal{C}\}$ be the dual cone of $\mathcal{C}$ which is any closed convex cone.

2 Preliminaries

In this section, we review some definitions and preliminary results developed by the author and his co-author in [3, 4] that will be used in the subsequent analysis. First, we recall from [16] that each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ admits a spectral factorization, associated with $\mathcal{K}^n$, of the form

$$x = \lambda_1 u^{(1)} + \lambda_2 u^{(2)},$$

where λ_1, λ_2 and $u^{(1)}, u^{(2)}$ are the spectral values and the associated spectral vectors of x given by

$$
\begin{aligned}
\lambda_i &= x_1 + (-1)^i \|x_2\|, \\
u^{(i)} &= \begin{cases} \frac{1}{2}\left(1, \ (-1)^i \dfrac{x_2}{\|x_2\|}\right) & \text{if } x_2 \neq 0; \\ \frac{1}{2}\left(1, \ (-1)^i w_2\right) & \text{if } x_2 = 0, \end{cases}
\end{aligned}
$$

for $i = 1, 2$, with w_2 being any vector in $\mathbb{R}^{n-1}$ satisfying $\|w_2\| = 1$. If $x_2 \neq 0$, the factorization is unique. The above spectral factorization of x, as well as x^2 and $x^{1/2}$ and the matrix L_x, have various interesting properties; see [16]. For instances, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, with spectral values λ_1, λ_2 and spectral vectors $u^{(1)}, u^{(2)}$, the following results hold : (1) $x^2 = \lambda_1^2 u^{(1)} + \lambda_2^2 u^{(2)} \in \mathcal{K}^n$. (2) If $x \in \mathcal{K}^n$, then $0 \leq \lambda_1 \leq \lambda_2$ and $x^{1/2} = \sqrt{\lambda_1}\, u^{(1)} + \sqrt{\lambda_2}\, u^{(2)}$. (3) If $x \in \text{int}(\mathcal{K}^n)$, then $0 < \lambda_1 \leq \lambda_2$, and L_x is invertible with

$$
L_x^{-1} = \frac{1}{x_1^2 - \|x_2\|^2} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \dfrac{x_1^2 - \|x_2\|^2}{x_1} I + \dfrac{1}{x_1} x_2 x_2^T \end{bmatrix}.
$$

In general, we have $x \circ y = L_x y$ for all $y \in \mathbb{R}^n$, and $L_x \succ 0$ if and only if $x \in \text{int}(\mathcal{K}^n)$.

We now recall definitions of monotonicity of a mapping which is needed for the assumptions of our main results later. We say that F is *monotone* if

$$\langle F(\zeta) - F(\xi), \zeta - \xi \rangle \geq 0 \quad \forall \zeta, \ \xi \in \mathbb{R}^n.$$

Similarly, F is *strongly monotone* if there exists $\rho > 0$ such that

$$\langle F(\zeta) - F(\xi), \zeta - \xi \rangle \geq \rho \|\zeta - \xi\|^2 \quad \forall \zeta, \ \xi \in \mathbb{R}^n.$$

It is well known that, when F is continuously differentiable, F is monotone if and only if $\nabla F(\zeta)$ is positive semi-definite for all $\zeta \in \mathbb{R}^n$ while F is strongly monotone if and only is $\nabla F(\zeta)$ is positive definite for all $\zeta \in \mathbb{R}^n$. For more details about monotonicity, please refer to [7].

The next useful lemma, describing special properties of x, y with $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, was used to prove Lemma 2.2 and will be also used to prove Prop. 3.1.

Lemma 2.1 *For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, we have*

$$
\begin{array}{rcl}
x_1^2 & = & \|x_2\|^2, \\
y_1^2 & = & \|y_2\|^2, \\
x_1 y_1 & = & x_2^T y_2, \\
x_1 y_2 & = & y_1 x_2.
\end{array}
$$

Proof. See [3, Lemma 3.2]. $\square$

Lemma 2.2 *Let ϕ_{FB} and ψ_{FB} be defined as in (7) and (8), respectively. Then the following holds.*

(a) $\psi_{\text{FB}}(x, y) = 0 \iff x \in \mathcal{K}^n, y \in \mathcal{K}^n, x \circ y = 0 \iff x \in \mathcal{K}^n, y \in \mathcal{K}^n, \langle x, y \rangle = 0.$

(b) ψ_{FB} *is continuously differentiable at every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. Moreover, $\nabla_x \psi_{\text{FB}}(0, 0) = \nabla_y \psi_{\text{FB}}(0, 0) = 0$. If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, then*

$$
\begin{array}{rcl}
\nabla_x \psi_{\text{FB}}(x, y) & = & \left(L_x L_{(x^2+y^2)^{1/2}}^{-1} - I \right) \phi_{\text{FB}}(x, y), \\
\nabla_y \psi_{\text{FB}}(x, y) & = & \left(L_y L_{(x^2+y^2)^{1/2}}^{-1} - I \right) \phi_{\text{FB}}(x, y).
\end{array}
\tag{14}
$$

If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, then $x_1^2 + y_1^2 \neq 0$ and

$$
\begin{array}{rcl}
\nabla_x \psi_{\text{FB}}(x, y) & = & \left(\dfrac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\text{FB}}(x, y), \\[4mm]
\nabla_y \psi_{\text{FB}}(x, y) & = & \left(\dfrac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\text{FB}}(x, y).
\end{array}
\tag{15}
$$

Proof. These results come from Prop. 3.1, Prop. 3.2, and Lemma 3.1 of [3]. $\square$

In what follows, for each $x \in \mathbb{R}^n$, $(x)_+$ denotes the nearest-point (in the Euclidean norm) projection of x onto $\mathcal{K}^n$. The following lemmas are crucial to our properties of error bound and bounded level sets in Sec. 4. They are results in [3] by the author and his co-author, the reader can find the proofs therein.

Lemma 2.3 *Let $\mathcal{C}$ be any closed convex cone in $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, let $x_{\mathcal{C}}^+$ and $x_{\mathcal{C}}^-$ denote the nearest-point (in the Euclidean norm) projection of x onto $\mathcal{C}$ and $-\mathcal{C}^*$, respectively. The following results hold.*

(a) *For any $x \in \mathbb{R}^n$, we have $x = x_{\mathcal{C}}^+ + x_{\mathcal{C}}^-$ and $\|x\|^2 = \|x_{\mathcal{C}}^+\|^2 + \|x_{\mathcal{C}}^-\|^2$.*

6

(b) *For any $x \in \mathbb{R}^n$ and $y \in \mathcal{C}$, we have $\langle x, y \rangle \leq \langle x_{\mathcal{C}}^+, y \rangle$.*

Lemma 2.4 *Let ϕ_{FB}, ψ_{FB} be given by (7) and (8), respectively. For any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we have*

$$4\psi_{\mathrm{FB}}(x, y) \geq 2\left\|\phi_{\mathrm{FB}}(x, y)_+\right\|^2 \geq \left\|(-x)_+\right\|^2 + \left\|(-y)_+\right\|^2.$$

Lemma 2.5 *Let ϕ_{FB}, ψ_{FB} be given by (7) and (8), respectively. For any $\{(x^k, y^k)\}_{k=1}^\infty \subseteq \mathbb{R}^n \times \mathbb{R}^n$, let $\lambda_1^k \leq \lambda_2^k$ and $\mu_1^k \leq \mu_2^k$ denote the spectral values of x^k and y^k, respectively. Then the following results hold.*

(a) *If $\lambda_1^k \to -\infty$ or $\mu_1^k \to -\infty$, then $\psi(x^k, y^k) \to \infty$.*

(b) *Suppose that $\{\lambda_1^k\}$ and $\{\mu_1^k\}$ are bounded below. If $\lambda_2^k \to \infty$ or $\mu_2^k \to \infty$, then $\langle x, x^k \rangle + \langle y, y^k \rangle \to \infty$ for any $x, y \in \mathrm{int}(\mathcal{K}^n)$.*

3 A new merit function and its properties

In this section, we study the new merit function ψ_{new} given as (11)-(12), i.e.,

$$\psi_{\mathrm{new}}(x, y) := \alpha\,\psi_0(x, y) + \psi_{\mathrm{FB}}(x, y),$$

where $\alpha > 0$ and

$$\psi_0(x, y) := \frac{1}{2}\|(x \circ y)_+\|^2, \qquad \psi_{\mathrm{FB}}(x, y) := \frac{1}{2}\|(x^2 + y^2)^{1/2} - x - y\|^2.$$

As we will see, ψ_{new} has several favorable properties which are parallel to what the function ψ_{FB} has. One important property that we will prove is that the SOCCP is indeed equivalent to the reformulation (13) :

$$\min_{\zeta \in \mathbb{R}^n} f(\zeta) \quad \text{where} \quad f(\zeta) := \psi_{\mathrm{new}}(F(\zeta), \zeta).$$

Other properties which will be shown are that the function f is smooth (Prop. 3.2) and has bounded level sets (Prop. 4.2) as well as providing an error bound (Prop. 4.1).

Lemma 3.1 *Let $\psi_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ be given by (11). Then ψ_0 is continuously differentiable and*

$$\begin{aligned} \nabla_x \psi_0(x, y) &= L_y \cdot (x \circ y)_+, \\ \nabla_y \psi_0(x, y) &= L_x \cdot (x \circ y)_+. \end{aligned} \tag{16}$$

Proof. For any $z \in \mathbb{R}^n$, we can factor z as $z = \lambda_1 u^{(1)} + \lambda_2 u^{(2)}$. Then let $g : \mathbb{R}^n \to \mathbb{R}^n$ be defined as

$$g(z) := \frac{1}{2}((z)_+)^2 = \hat{g}(\lambda_1)u^{(1)} + \hat{g}(\lambda_2)u^{(2)},$$

where $\hat{g} : \mathbb{R} \to \mathbb{R}$ is given by $\hat{g}(\lambda) := \frac{1}{2}(\max(0, \lambda))^2$. From the continuous differentiability of $\hat{g}$ and Prop. 5.2 of [5], the vector-valued function g is also continuously differentiable. Hence, the first component $g_1(z) := \frac{1}{2}\|(z)_+\|^2$ of $g(z)$ is continuously differentiable as well. By an easy computation, we have $\nabla g_1(z) = (z)_+$. Now, let

$$z(x, y) := x \circ y = (\langle x, y \rangle , \; x_1 y_2 + y_1 x_2),$$

then we have $\psi_0(x, y) = g_1(z(x, y))$. Applying the chain rule, we obtain

$$\nabla_x \psi_0 = \nabla_x z \cdot \nabla g_1(z) = L_y \cdot (x \circ y)_+,$$
$$\nabla_y \psi_0 = \nabla_y z \cdot \nabla g_1(z) = L_x \cdot (x \circ y)_+,$$

where

$$\nabla_x z(x, y) = \begin{bmatrix} y_1 & y_2^T \\ y_2 & y_1 I \end{bmatrix} = L_y \quad \text{and} \quad \nabla_y z(x, y) = \begin{bmatrix} x_1 & x_2^T \\ x_2 & x_1 I \end{bmatrix} = L_x.$$

Thus, the proof is completed. □

Proposition 3.1 *Let* $\psi_{\mathrm{new}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ *be defined as in (11)-(12). Then the following results hold.*

(a) $\psi_{\mathrm{new}}(x, y) \geq 0$ *for all* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$.

(b) $\psi_{\mathrm{new}}(x, y) = 0 \iff x \in \mathcal{K}^n, \; y \in \mathcal{K}^n, \; x \circ y = 0 \iff x \in \mathcal{K}^n, \; y \in \mathcal{K}^n, \; \langle x, y \rangle = 0.$

(c) ψ_{new} *is continuously differentiable at every* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. *Moreover,* $\nabla_x \psi_{\mathrm{new}}(0, 0) = \nabla_y \psi_{\mathrm{new}}(0, 0) = 0$. *If* $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, *then*

$$\begin{aligned}
\nabla_x \psi_{\mathrm{new}}(x, y) &= \alpha \; L_y \cdot (x \circ y)_+ + \left(L_x L_{(x^2+y^2)^{1/2}}^{-1} - I \right) \phi_{\mathrm{FB}}(x, y), \\
\nabla_y \psi_{\mathrm{new}}(x, y) &= \alpha \; L_x \cdot (x \circ y)_+ + \left(L_y L_{(x^2+y^2)^{1/2}}^{-1} - I \right) \phi_{\mathrm{FB}}(x, y).
\end{aligned} \tag{17}$$

If $(x, y) \neq (0, 0)$ *and* $x^2 + y^2 \notin \mathrm{int}(\mathcal{K}^n)$, *then* $x_1^2 + y_1^2 \neq 0$ *and*

$$\begin{aligned}
\nabla_x \psi_{\mathrm{new}}(x, y) &= 2\alpha \; |x_1| \cdot (y)_+^2 + \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\mathrm{FB}}(x, y), \\
\nabla_y \psi_{\mathrm{new}}(x, y) &= 2\alpha \; |y_1| \cdot (x)_+^2 + \left(\frac{y_1}{\sqrt{x_1^2 + y_1^2}} - 1 \right) \phi_{\mathrm{FB}}(x, y).
\end{aligned} \tag{18}$$

Proof. (a) It is clear by definition.

(b) We only need to prove the first equivalence since the second one is a known result in [16]. Suppose $\psi_{\text{new}}(x, y) = 0$, it yields $\psi_{\text{FB}}(x, y) = 0$. Thus, the desirable result follows by Lemma 2.2(a). On the other hand, $x \in \mathcal{K}^n$, $y \in \mathcal{K}^n$, $x \circ y = 0$ imply $\psi_{\text{FB}}(x, y) = 0$; and $\psi_0(x, y) = 0$ from $x \circ y = 0$. Therefore, $\psi_{\text{new}}(x, y) = 0$.

(c) If $(x, y) = (0, 0)$, it is easy to know $\nabla_x \psi_0(0, 0) = \nabla_y \psi_0(0, 0) = 0$ by Lemma 3.1. Hence $\nabla_x \psi_{\text{new}}(0, 0) = \nabla_y \psi_{\text{new}}(0, 0) = 0$. If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, then the results follow by Lemma 2.2(b) and Lemma 3.1. If $(x, y) \neq (0, 0)$ and $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$, then by applying Lemma 2.1, we have

$$
\begin{aligned}
x \circ y &= (\langle x, y \rangle \,,\, x_1 y_2 + y_1 x_2) \\
&= (x_1 y_1 + x_2^T y_2 \,,\, x_1 y_2 + y_1 x_2) \\
&= (2x_1 y_1 \,,\, 2x_1 y_2) \\
&= 2x_1 y.
\end{aligned}
$$

Therefore,

$$
L_y \cdot (x \circ y)_+ = L_y \cdot (2x_1 y)_+ = 2|x_1| \, L_y \cdot (y)_+ = 2|x_1| \cdot (y \circ (y)_+) = 2|x_1| \cdot (y)_+^2,
$$

where the last equality is due to

$$
y \circ y_+ = [(y)_+ + (y)_-] \circ (y)_+ = (y)_+^2 + (y)_- \circ (y)_+ = (y)_+^2.
$$

Similarly, we have $L_x \cdot (x \circ y)_+ = 2|y_1| \cdot (x)_+^2$. This together with Lemma 2.2 lead to the desired results. $\square$

Proposition 3.2 *Let f be defined as (11)-(13). Then f is smooth with $f(\zeta) \geq 0$ for all $\zeta \in \mathbb{R}^n$ and $f(\zeta) = 0$ if and only if ζ solves the SOCCP. Moreover, suppose that the SOCCP has at least one solution. Then, ζ is a global minimization of f if and only if ζ solves the SOCCP.*

Proof. The results follow by Prop. 3.1 and definition of f. $\square$

4   Error bound and bounded level sets

The error bound is an important concept that indicates how close an arbitrary point is to the solution set of SOCCP. Thus, an error bound may be used to provide stopping criterion for an iterative method. As below, we establish a proposition about the error bound of f given as (11)-(13). We need the next technical lemma to prove the error bound property.

Lemma 4.1 *Let* $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ *and* $y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$. *Then, we have*

$$\langle x, y \rangle \leq \sqrt{2} \, \|(x \circ y)_+\|. \tag{19}$$

Proof. First, we observe the fact that

$$
\begin{aligned}
x \in \mathcal{K}^n &\iff (x)_+ = x, \\
x \in -\mathcal{K}^n &\iff (x)_+ = 0, \\
x \notin \mathcal{K}^n \cup -\mathcal{K}^n &\iff (x)_+ = \lambda_2 u^{(2)},
\end{aligned}
$$

where λ_2 is the bigger spectral value of x with the corresponding spectral vector $u^{(2)}$ defined as in Sec. 2. Hence, we have three cases.

Case(1): If $x \circ y \in \mathcal{K}^n$, then $(x \circ y)_+ = x \circ y$. By definition of Jordan product of x and y as (6), i.e., $x \circ y = (\langle x, y \rangle, \ x_1 y_2 + y_1 x_2)$. It is clear that $\|(x \circ y)_+\| \geq \langle x, y \rangle$ and hence (19) holds.

Case(2): If $x \circ y \in -\mathcal{K}^n$, then $(x \circ y)_+ = 0$. Since $x \circ y \in -\mathcal{K}^n$, by definition of Jordan product again, we have $\langle x, y \rangle \leq 0$. Hence, it is true that $\sqrt{2}\|(x \circ y)_+\| \geq \langle x, y \rangle$.

Case(3): If $x \circ y \notin \mathcal{K}^n \cup -\mathcal{K}^n$, then $(x \circ y)_+ = \lambda_2 u^{(2)}$ where

$$
\begin{aligned}
\lambda_2 &= \langle x, y \rangle + \|x_1 y_2 + y_1 x_2\|, \\
u^{(2)} &= \frac{1}{2} \left(1, \ \frac{x_1 y_2 + y_1 x_2}{\|x_1 y_2 + y_1 x_2\|} \right).
\end{aligned}
$$

If $\langle x, y \rangle \leq 0$, then (19) is trivial. Thus, we can assume $\langle x, y \rangle > 0$. In fact, the desired inequality (19) follows from the below.

$$
\begin{aligned}
\|(x \circ y)_+\|^2 &= \frac{1}{2} \lambda_2^2 \\
&= \frac{1}{2} \left(\langle x, y \rangle^2 + 2 \langle x, y \rangle \cdot \|x_1 y_2 + y_1 x_2\| + \|x_1 y_2 + y_1 x_2\|^2 \right) \\
&\geq \frac{1}{2} \langle x, y \rangle^2,
\end{aligned}
$$

where the first equality is by $\|u^{(2)}\| = 1/\sqrt{2}$. $\quad \square$

Proposition 4.1 *Suppose that F is strongly monotone mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. Also, suppose that SOCCP has a solution ζ^*. Then there exists a scalar $\tau > 0$ such that*

$$\tau \|\zeta - \zeta^*\|^2 \leq \|(F(\zeta) \circ \zeta)_+\| + \|(-F(\zeta))_+\| + \|(-\zeta)_+\| \quad \forall \zeta \in \mathbb{R}^n. \tag{20}$$

Moreover,

$$\tau \|\zeta - \zeta^*\|^2 \leq \sqrt{2} \left(\frac{1}{\alpha} + 2 \right) f(\zeta)^{1/2} \quad \forall \zeta \in \mathbb{R}^n, \tag{21}$$

where $\alpha > 0$, and f is given by (11)-(13) .

Proof. Since F is strongly monotone, there exists a scalar $\rho > 0$ such that, for any $\zeta \in \mathbb{R}^n$,

$$
\begin{aligned}
& \rho\|\zeta - \zeta^*\|^2 \\
\leq\ & \langle F(\zeta) - F(\zeta^*),\ \zeta - \zeta^* \rangle \\
=\ & \langle F(\zeta), \zeta \rangle + \langle -F(\zeta), \zeta^* \rangle + \langle F(\zeta^*), -\zeta \rangle \\
\leq\ & \langle F(\zeta), \zeta \rangle + \langle (-F(\zeta))_+, \zeta^* \rangle + \langle F(\zeta^*), (-\zeta)_+ \rangle \\
\leq\ & \langle F(\zeta), \zeta \rangle + \|(-F(\zeta))_+\|\ \|\zeta^*\| + \|F(\zeta^*)\|\ \|(-\zeta)_+\| \\
\leq\ & \sqrt{2}\|(F(\zeta) \circ \zeta)_+\| + \|(-F(\zeta))_+\|\ \|\zeta^*\| + \|F(\zeta^*)\|\ \|(-\zeta)_+\| \\
\leq\ & \max\{\sqrt{2}, \|F(\zeta^*)\|, \|\zeta^*\|\} \Big(\|(F(\zeta) \circ \zeta)_+\| + \|(-F(\zeta))_+\| + \|(-\zeta)_+\| \Big),
\end{aligned}
$$

where the second inequality uses Lemma 2.3(b) while the fourth inequality is from (19). Then, setting $\tau := \dfrac{\rho}{\max\{\sqrt{2}, \|F(\zeta^*)\|, \|\zeta^*\|\}}$ yields (20).

Moreover, we have

$$
\|(F(\zeta) \circ \zeta)_+\| = \sqrt{2}\ \psi_0(F(\zeta), \zeta)^{1/2} \leq \frac{\sqrt{2}}{\alpha}\ f(\zeta)^{1/2},
$$

and

$$
\begin{aligned}
\|(-F(\zeta))_+\| + \|(-\zeta)_+\| \ &\leq\ \sqrt{2} \left(\|(-F(\zeta))_+\|^2 + \|(-\zeta)_+\|^2 \right)^{1/2} \\
&\leq\ 2\sqrt{2}\ \psi_{\text{FB}}(F(\zeta), \zeta)^{1/2} \\
&\leq\ 2\sqrt{2}\ f(\zeta)^{1/2},
\end{aligned}
$$

where the second inequality is true by Lemma 2.4. Thus,

$$
\|(F(\zeta) \circ \zeta)_+\| + \|(-F(\zeta))_+\| + \|(-\zeta)_+\| \leq \sqrt{2} \left(\frac{1}{\alpha} + 2 \right) f(\zeta)^{1/2}.
$$

This together with (20) yield (21). $\qquad\square$

The boundedness of level sets of a merit function is also important since it ensures that the sequence generated by a descent method has at least one accumulation. The following proposition gives conditions under which f has bounded level sets.

Proposition 4.2 *Suppose that F is a monotone mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ and that SOCCP is strictly feasible, i.e., there exists $\hat{\zeta} \in \mathbb{R}^n$ such that $F(\hat{\zeta}), \hat{\zeta} \in \text{int}(\mathcal{K}^n)$. Then the level set*

$$
\mathcal{L}(\gamma) := \{\zeta \in \mathbb{R}^n \mid f(\zeta) \leq \gamma\}
$$

is bounded for all $\gamma \geq 0$, where f is given by (11)-(13) with $\alpha > 0$.

Proof. We will prove this result by contradiction. Suppose there exists an unbounded sequence $\{\zeta^k\} \subset \mathcal{L}(\gamma)$ for some $\gamma \geq 0$. It can be seen that the sequence of the smaller spectral values of $\{\zeta^k\}$ and $\{F(\zeta^k)\}$ are bounded below. In fact, if not, it follows form Lemma 2.5(a) that $f(\zeta^k) \to \infty$, which contradicts $\{\zeta^k\} \subset \mathcal{L}(\gamma)$. Therefore, the unboundedness of $\{\zeta^k\}$ leads to that the sequence of the bigger spectral values of $\{\zeta^k\}$ tends to infinity. Now, let $\hat{\zeta}$ be a strictly feasibile solution of the SOCCP. Since F is monotone, we have

$$\langle F(\zeta^k) - F(\hat{\zeta}) \, , \, \zeta^k - \hat{\zeta} \rangle \geq 0,$$

which yields

$$\langle F(\zeta^k), \hat{\zeta} \rangle + \langle F(\hat{\zeta}), \zeta^k \rangle \leq \langle F(\zeta^k), \zeta^k \rangle + \langle F(\hat{\zeta}), \hat{\zeta} \rangle. \tag{22}$$

Then, by Lemma 2.5(b) and $F(\hat{\zeta}), \hat{\zeta} \in \text{int}(\mathcal{K}^n)$, we obtain $\langle F(\zeta^k), \hat{\zeta} \rangle + \langle F(\hat{\zeta}), \zeta^k \rangle \to \infty$, which together with (22) lead to $\langle F(\zeta^k), \zeta^k \rangle \to \infty$. Thus, by Lemma 4.1 and (12)-(13), we have

$$\|(F(\zeta^k) \circ \zeta^k)_+\| \to \infty \quad \Longrightarrow \quad \psi_{\text{new}}(F(\zeta^k), \zeta^k) \to \infty \quad \Longrightarrow \quad f(\zeta^k) \to \infty.$$

But, this contradicts $\{\zeta^k\} \subset \mathcal{L}(\gamma)$. Therefore, we complete the proof. $\quad\square$

5   Concluding Remarks

In this paper, we have proposed a new merit function for the SOCCP. We also have shown that the function enjoys some favorable properties. In particular, it provides a global error bound under strong monotonicity of F only, and it has bounded level sets when F is monotone and the SOCCP is strictly feasible. With these properties, it is possible to construct a descent algorithm, as done in [4, 26, 27], for solving the unconstrained minimization reformulation (13) of the SOCCP and investigate its global convergence. We leave it as one of the future topics. In fact, there have already had systematic study on merit functions for NCP and SDCP cases, however, not much for the SOCCP case. Therefore, it would be worth of considering other merit functions for SOCCP and build a systematic study accordingly. On the other hand, recently there have definitions of P-properties for the nonlinear transformations on Euclidean Jordan algebras (see [24] for details). In particular, they proved the following implications.

$$\text{strongly monotone} \Longrightarrow \text{uniform Jordan } P\text{-property} \Longrightarrow \text{uniform } P\text{-property} \Longrightarrow$$
$$P\text{-property}$$

Therefore, another interesting future topic is to see whether the assumptions used in Prop. 4.1 and Prop. 4.2 can be weaken to P-properties.

References

[1] F. ALIZADEH AND S. SCHMIETA, *Symmetric cones, potential reduction methods, and word-by-word extensions*, in Handbook of Semidefinite Programming, edited by H. Wolkowicz, R. Saigal, and L. Vandenberghe, Kluwer Academic Publishers, Boston, pp. 195–233, 2000.

[2] D. P. BERTSEKAS, *Nonlinear Programming*, 2nd edition, Athena Scientific, Belmont, 1999.

[3] J.-S. CHEN AND P. TSENG, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, to appear in Mathematical Programming, (2005).

[4] J.-S. CHEN, *A descent method for solving reformulation of the second-order cone complementarity problem*, submitted, (2004).

[5] J.-S. CHEN, X. CHEN, AND P. TSENG, *Analysis of nonsmooth vector-valued functions associated with second-order cones*, Mathematical Programming, vol. 101 (2004), pp. 95-117.

[6] X.-D. CHEN, D. SUN, AND J. SUN, *Complementarity functions and numerical experiments for second-order cone complementarity problems,* Computational Optimization and Applications, vol. 25 (2003), pp. 39-56.

[7] F. FACCHINEI AND J.-S. PANG, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York, 2003.

[8] F. FACCHINEI AND J. SOARES, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM Journal on Optimization, vol. 7 (1997), pp. 225-247.

[9] U. FARAUT AND A. KORÁNYI, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[10] M. C. FERRIS AND J.-S. PANG, *Engineering and economic applications of complementarity problems*, SIAM Review, vol. 39 (1997), pp. 669-713.

[11] M. C. FERRIS AND J.-S. PANG, editors, *Complementarity and Variational Problems: State of the Art*, SIAM Publications, Philadelphia, 1996.

[12] M. C. FERRIS, O. L. MANGASARIAN, AND J.-S. PANG, editors, *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.

[13] A. Fischer, *A special Newton-type optimization methods*, Optimization, vol. 24 (1992), pp. 269-284.

[14] A. Fischer, *Solution of the monotone complementarity problem with locally Lipschitzian functions*, Mathematical Programming, vol. 76 (1997), pp. 513-532.

[15] R. Fletcher, *Practical Methods of Optimization*, 2nd edition, Wiley-Interscience, Chichester, 1987.

[16] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM Journal on Optimization, vol. 12 (2002), pp. 436-460.

[17] S. Hayashi, N. Yamashita, and M. Fukushima, *On the coerciveness of merit functions for the second-order cone complementarity problem*, Report, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan, June (2001).

[18] C. Kanzow, *An unconstrained optimization technique for large scale linearly constrained convex minimization problems*, Computing, vol. 53 (1994), pp. 101-117.

[19] C. Kanzow, *Nonlinear complementarity as unconstrained optimization*, Journal of Optimization Theory and Applications, vol. 88 (1996), pp. 139-155.

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

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

[22] S. Schmieta and F. Alizadeh, *Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones*, Mathematics of Operations Research, vol. 26 (2001), pp. 543-564.

[23] C.-K. Sim and G. Zhao, *A note on treating second order cone problem as a special case of semidefinite problem*, Report, Department of Mathematics, National University of Singapore, Singapore, May 2002; to appear in Mathematical Programming.

[24] J. Tao and M. S. Gowda, *Some P-properties for the nonlinear transformations on Euclidean Jordan Algebra*, Technical Report, Department of Mathematics and Satistics, University of Maryland, 2004.

[25] T. Tsuchiya, *A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming*, Optimization Methods and Software, vol. 11 (1999), pp. 141–182.

[26] K. YAMADA, N. YAMASHITA, AND M. FUKUSHIMA, *A new derivative-free descent method for the nonlinear complementarity problems*, in Nonlinear Optimization and Related Topics edited by G.D. Pillo and F. Giannessi, Kluwer Academic Publishers, Netherlands, (2000), pp. 463–489.

[27] N. YAMASHITA AND M. FUKUSHIMA, *A new merit function and a descent method for semidefinite complementarity problems*, in Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods edited by M. Fukushima and L. Qi, Kluwer Academic Publishers, Boston, (1999), pp. 405–420.