

A descent method for a reformulation of the second-order cone complementarity problem

Jein-Shan Chen^{a,*}, Shaohua Pan^b

^aDepartment of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

^bSchool of Mathematical Sciences, South China University of Technology, Guangzhou 510640, China

Received 27 July 2006; received in revised form 31 January 2007

Abstract

Analogous to the nonlinear complementarity problem and the semi-definite complementarity problem, a popular approach to solving the second-order cone complementarity problem (SOCCP) is to reformulate it as an unconstrained minimization of a certain merit function over \mathbb{R}^n . In this paper, we present a descent method for solving the unconstrained minimization reformulation of the SOCCP which is based on the Fischer–Burmeister merit function (FBMF) associated with second-order cone [J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, *Math. Programming* 104 (2005) 293–327], and prove its global convergence. Particularly, we compare the numerical performance of the method for the symmetric affine SOCCP generated randomly with the FBMF approach [J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, *Math. Programming* 104 (2005) 293–327]. The comparison results indicate that, if a scaling strategy is imposed on the test problem, the descent method proposed is comparable with the merit function approach in the CPU time for solving test problems although the former may require more function evaluations.

© 2007 Elsevier B.V. All rights reserved.

MSC: 26B05; 90C33

Keywords: Second-order cone; Complementarity; Merit function; Descent method

1. Introduction

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

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

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

$$\mathcal{H} = \mathcal{H}^{n_1} \times \cdots \times \mathcal{H}^{n_m}, \quad (2)$$

* Corresponding author. Tel.: +886 2 29325417; fax: +886 2 29332342.

E-mail addresses: jschen@math.ntnu.edu.tw (J.-S. Chen), shhpan@scut.edu.cn (S. Pan).

¹ Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

where $m, n_1, \dots, n_m \geq 1, n_1 + \dots + n_m = n$, 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 $n_1 = \dots = n_m = 1$. 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., [7,10–12]. Unless otherwise stated, in the first three sections of 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 carried over to the case where \mathcal{K} is a product of second-order cones without difficulty).

Various methods have been proposed for solving the SOCCP. They include interior-point methods [1,24,26,28,30], reformulating SOC constraints as smooth convex constraints [31], and (noninterior) smoothing Newton methods [5,16,20]. These methods require solving a nontrivial system of linear equations at each iteration. In the recent paper [4], an alternative approach based on reformulating the SOCCP as an unconstrained smooth minimization problem was studied. In particular, they were finding a smooth function $\psi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ such that

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

We call such a ψ a *merit function*. Then 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 (5). For this approach to be effective, the choice of ψ is crucial. In the case of NCP, a popular choice is

$$\psi_{\text{FB}}(a, b) = \frac{1}{2} \sum_{i=1}^n \phi_{\text{FB}}(a_i, b_i)^2$$

for all $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$ and $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$, where ϕ_{FB} is the well-known Fischer–Burmeister NCP-function [13,14] defined by

$$\phi_{\text{FB}}(a_i, b_i) = \sqrt{a_i^2 + b_i^2} - a_i - b_i.$$

It has been shown that ψ_{FB} is smooth (even though ϕ_{FB} is not differentiable) and is a merit function for NCP [8,21,22]. These two functions can be extended to the case of SOCCP via Jordan algebra shown as below. 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* associated with \mathcal{K}^n as

$$x \circ y := (\langle x, y \rangle, y_1x_2 + x_1y_2).$$

The identity element under this product is $e := (1, 0, \dots, 0)^T \in \mathbb{R}^n$. 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, if $x \in \mathcal{K}^n$, then there exists a unique vector in \mathcal{K}^n , denoted by $x^{1/2}$, such that $(x^{1/2})^2 = x^{1/2} \circ x^{1/2} = x$. Then,

$$\phi_{\text{FB}}(x, y) := (x^2 + y^2)^{1/2} - x - y \tag{6}$$

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 $\langle x, y \rangle = 0, x \in \mathcal{K}^n, y \in \mathcal{K}^n$. Hence, $\psi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ given by

$$\psi_{\text{FB}}(x, y) := \frac{1}{2} \|\phi_{\text{FB}}(x, y)\|^2 \tag{7}$$

is a merit function for SOCCP because ψ_{FB} satisfies (4) as well. Therefore, the SOCCP is equivalent to the global minimization problem:

$$\min_{\zeta \in \mathbb{R}^n} f_{\text{FB}}(\zeta) := \psi_{\text{FB}}(F(\zeta), \zeta). \tag{8}$$

It was also shown in the paper [4] that, like the NCP case, ψ_{FB} is smooth and, when ∇F is positive semi-definite, every stationary point of (8) solves the SOCCP. For semi-definite complementarity problem (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 aforementioned features hold for the following analog of the SDCP merit function studied in [32]:

$$\psi_{\text{YF}}(x, y) := \psi_0(\langle x, y \rangle) + \psi_{\text{FB}}(x, y), \tag{9}$$

where $\psi_0 : \mathbb{R} \rightarrow [0, \infty)$ is any smooth function satisfying

$$\psi_0(t) = 0 \quad \forall t \leq 0 \quad \text{and} \quad \psi'_0(t) > 0 \quad \forall t > 0. \tag{10}$$

In [32], $\psi_0(t) = \frac{1}{4}(\max\{0, t\})^4$ was considered. In fact, the function ψ_{YF} , which was recently studied in [4], is also a SOCCP version merit function that enjoys favorable properties as what ψ_{FB} has. Moreover, ψ_{YF} possesses properties of bounded level sets and error bound.

In this paper, we focus on the following equivalent reformulation of SOCCP, which arises via the merit function ψ_{YF} defined as in (9) and (10):

$$\min_{\zeta \in \mathbb{R}^n} f_{\text{YF}}(\zeta) := \psi_{\text{YF}}(F(\zeta), \zeta). \tag{11}$$

We are motivated by the work [32] showing a descent method for the SDCP. Thus, the main purpose of the paper is to explore the extension to SOCCP. In other words, we wish to adopt the algorithm therein to solve the equivalent reformulation (11) of the SOCCP and prove its global convergence (see Section 3). In particular, we also compare the numerical performance of the descent algorithm for the symmetric affine SOCCPs generated randomly with the Fischer–Burmeister merit function (FBMF) approach [4]. Here it is worth pointing out that the proposed algorithm does not work for the other reformulation (8). The reason is that $f_{\text{FB}}(\zeta)$ lacks property of bounded level sets and does not provide error bound due to the absence of the term ψ_0 .

Some words about our notation. Throughout this paper, \mathbb{R}^n denotes the space of n -dimensional real column vectors. For any differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x . For any differentiable mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\nabla F(x)$ is an $n \times m$ matrix which denotes the transposed Jacobian of F at x .

2. Preliminaries

As mentioned in the Introduction, ψ_{YF} satisfies (4), so the SOCCP can be recast as an equivalent global minimization (11). It was shown in [4] that the function f_{YF} is smooth, has bounded level sets, and provides error bound for the unconstrained minimization reformulation. Moreover, every stationary point of problem (11) is a solution of the SOCCP. In this section, we review some basic concepts and properties that will be used for proving the convergence results of the descent algorithm later. Since the work of [4] already includes as special cases the following lemmas, we here omit the proofs.

Lemma 2.1 (Chen and Tseng [4, Proposition 3.2]). *Let $\phi_{\text{FB}}, \psi_{\text{FB}}$ be given by (6) and (7), respectively, and ψ_{YF} be given by (9) and (10). Then ψ_{FB} and ψ_{YF} are both smooth on $\mathbb{R}^n \times \mathbb{R}^n$.*

Lemma 2.2 (Chen and Tseng [4, Proposition 4.2]). *Let ψ_{YF} be given by (9) and (10) and $f_{\text{YF}}(\zeta)$ be defined as (11). Then, for every $\zeta \in \mathbb{R}^n$ such that $\nabla F(\zeta)$ is positive semi-definite, either $f_{\text{YF}}(\zeta) = 0$ or $\nabla f_{\text{YF}}(\zeta) \neq 0$ with $\langle d(\zeta), \nabla f_{\text{YF}}(\zeta) \rangle < 0$, where*

$$d(\zeta) := -(\psi'_0(\langle F(\zeta), \zeta \rangle)\zeta + \nabla_x \psi_{\text{FB}}(F(\zeta), \zeta)). \tag{12}$$

In what follows, we say that F is *monotone* if

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

and 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 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 if $\nabla F(\zeta)$ is positive definite for all $\zeta \in \mathbb{R}^n$.

Lemma 2.3 (Chen and Tseng [4, Proposition 5.2]). *Suppose that F is a differentiable and monotone mapping from \mathbb{R}^n to \mathbb{R}^n . Suppose also that the SOCCP (1) is strictly feasible, i.e., there exists $\tilde{\zeta} \in \mathbb{R}^n$ such that $F(\tilde{\zeta}), \tilde{\zeta} \in \text{int}(\mathcal{H}^n)$. Then the level set*

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

is nonempty and bounded for all $\gamma \geq 0$, where f_{YF} is given by (11).

Remark 2.1. It is known that Lemma 2.3 is also true if the conditions of monotonicity and strict feasibility are replaced by strong monotonicity.

We next recall some basic results about the spectral factorization associated with \mathcal{H}^n . Any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, admits a spectral factorization of the form

$$x = \lambda_1(x) \cdot u_x^{(1)} + \lambda_2(x) \cdot u_x^{(2)}, \tag{13}$$

where $\lambda_i(x)$ and $u_x^{(i)}$ for $i = 1, 2$ are the spectral values and the associated spectral vectors of x given by

$$\lambda_i(x) = x_1 + (-1)^i \|x_2\|, \\ u_x^{(i)} = \begin{cases} \frac{1}{2} \begin{pmatrix} 1, (-1)^i \frac{x_2}{\|x_2\|} \end{pmatrix} & \text{if } x_2 \neq 0, \\ \frac{1}{2} (1, (-1)^i w_2) & \text{if } x_2 = 0 \end{cases} \tag{14}$$

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 set $\{u_x^{(1)}, u_x^{(2)}\}$ is called a *Jordan frame* and has the following properties.

Property 2.1. *For any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$ with the spectral values $\lambda_1(x), \lambda_2(x)$ and spectral vectors $u_x^{(1)}, u_x^{(2)}$ given as in (14), we have*

(a) $u_x^{(1)}$ and $u_x^{(2)}$ are orthogonal under Jordan product and have length $1/\sqrt{2}$, i.e.,

$$u_x^{(1)} \circ u_x^{(2)} = 0, \quad \|u_x^{(1)}\| = \|u_x^{(2)}\| = \frac{1}{\sqrt{2}}.$$

(b) $u_x^{(i)}$ and $u_x^{(i)}$ are idempotent under Jordan product, i.e., $u_x^{(i)} \circ u_x^{(i)} = u_x^{(i)}$ for $i = 1, 2$.

The spectral factorization (13)–(14) of x as well as x^2 and $x^{1/2}$ has various interesting properties; see [16]. For instance, for any $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, with spectral values $\lambda_1(x), \lambda_2(x)$ and spectral vectors $u_x^{(1)}, u_x^{(2)}$, the following results hold:

- (1) $x^2 = \lambda_1^2(x)u_x^{(1)} + \lambda_2^2(x)u_x^{(2)} \in \mathcal{H}^n$.
- (2) If $x \in \mathcal{H}^n$, then $0 \leq \lambda_1(x) \leq \lambda_2(x)$ and $x^{1/2} = \sqrt{\lambda_1(x)}u_x^{(1)} + \sqrt{\lambda_2(x)}u_x^{(2)}$.

To close this section, we present a property of ψ_{FB} associated with the spectral value.

Lemma 2.4 (Cheng and Tseng [4, Lemma 9(a)]). *For any $\{(x^k, y^k)\}_{k=1}^\infty \subseteq \mathbb{R}^n \times \mathbb{R}^n$, let $\lambda_1(x^k) \leq \lambda_2(x^k)$ and $\mu_1(y^k) \leq \mu_2(y^k)$ denote the spectral values of x^k and y^k , respectively. Then, if $\lambda_1(x^k) \rightarrow -\infty$ or $\mu_1(y^k) \rightarrow -\infty$, we have $\psi_{\text{FB}}(x^k, y^k) \rightarrow \infty$.*

3. Main results

In this section, we propose a descent method for solving the unconstrained minimization reformulation (11) of the SOCCP and prove its global convergence. The proposed method uses $d(\zeta)$ defined as (12) as its direction. Now let us describe the algorithm.

Algorithm 3.1. (Step 0) Choose $\zeta^0 \in \mathbb{R}^n$, $\varepsilon \geq 0$, $\sigma \in (0, 1/2)$, $\beta \in (0, 1)$ and set $k := 0$.

(Step 1) If $f_{\text{YF}}(\zeta^k) \leq \varepsilon$, then stop.

(Step 2) Compute $d(\zeta^k) := -(\psi'_0(\langle F(\zeta^k), \zeta^k \rangle)_{\zeta^k} + \nabla_x \psi_{\text{FB}}(F(\zeta^k), \zeta^k))$.

(Step 3) Find a step-size $t_k := \beta^{m_k}$, where m_k is the smallest nonnegative integer m satisfying Armijo’s rule:

$$f_{\text{YF}}(\zeta^k + \beta^m d(\zeta^k)) \leq (1 - \sigma \beta^{2m}) f_{\text{YF}}(\zeta^k). \tag{15}$$

(Step 4) Set $\zeta^{k+1} := \zeta^k + t_k d(\zeta^k)$, $k := k + 1$ and go to Step 1.

Note that the above algorithm is ∇F -free, i.e., there is no need to compute the Jacobian matrix of F , and moreover, the computation work in each iteration is very small, i.e., only several vector multiplications. In fact, this type of algorithm was also studied for the NCP (see [17]) and the SDCP (see [32]) and the most remarkable feature of this type of algorithm is that not only the step-size but also the search direction itself is adjusted via Armijo’s rule. In practical experience, σ is usually chosen close to zero, and β is usually chosen in $(\frac{1}{10}, \frac{1}{2})$ depending on the confidence we have on the quality of the initial step-size (see [2]).

Next, we prove the global convergence of Algorithm 3.1. Without any loss of generality, we suppose $\varepsilon = 0$ so that the algorithm generates an infinite sequence $\{\zeta^k\}$.

Proposition 3.1. *Suppose that F is monotone and the SOCCP (1) is strictly feasible. Then the sequence $\{\zeta^k\}$ generated by Algorithm 3.1 has at least one accumulation point, and any accumulation point is a solution of the SOCCP (1).*

Proof. The proof is standard and can be found in [2]. For completeness, we here present its proof by the following three steps.

- (i) First, we show that, whenever ζ^k is not a solution, there exists a nonnegative integer m_k in Step 3 of Algorithm 3.1. Suppose not, then for any positive integer m , we have

$$f_{\text{YF}}(\zeta^k + \beta^m d(\zeta^k)) - f_{\text{YF}}(\zeta^k) > -\sigma \beta^{2m} f_{\text{YF}}(\zeta^k).$$

Dividing by β^m on both sides and letting $m \rightarrow \infty$ yields

$$\langle \nabla f_{\text{YF}}(\zeta^k), d(\zeta^k) \rangle \geq 0. \tag{16}$$

Since F is monotone which is equivalent to $\nabla F(\zeta)$ is positive semi-definite, the inequality (16) contradicts Lemma 2.2. Hence, we can find an integer m_k in Step 3.

- (ii) Secondly, we show that the sequence $\{\zeta^k\}$ generated by the algorithm has at least one accumulation point. By the descent property of Algorithm 3.1, the sequence $\{f_{\text{YF}}(\zeta^k)\}_{k \in \mathbb{N}}$ is decreasing. Hence by Lemma 2.3, we have that $\{\zeta^k\}$ is bounded, and consequently has at least one accumulation point.
- (iii) Finally, we prove that any accumulation point of $\{\zeta^k\}$ is a solution of the SOCCP (1). Let ζ^* be an arbitrary accumulation point of $\{\zeta^k\}_{k \in \mathbb{N}}$. In other words, there is a subsequence $\{\zeta^k\}_{k \in K}$ converging to ζ^* , where K is a subset of \mathbb{N} . We know $d(\cdot)$ is continuous (since ψ_0 and ψ_{FB} are smooth) which implies $\{d(\zeta^k)\}_{k \in K}$ converges to $d(\zeta^*)$. Next, we need to discuss two cases. First, we consider the case where there exists a constant $\bar{\beta}$ such that $\beta^{m_k} \geq \bar{\beta} > 0$ for all $k \in K$. Then, from (15), we have

$$f_{\text{YF}}(\zeta^{k+1}) \leq (1 - \sigma \bar{\beta}^2) f_{\text{YF}}(\zeta^k)$$

for all $k \in K$ and the entire sequence $\{f_{\text{YF}}(\zeta^k)\}_{k \in K}$ is decreasing. Thus, we obtain $f_{\text{YF}}(\zeta^*) = 0$ (by taking the limit) which says ζ^* is a solution of the SOCCP (1). Now, we consider the other case where there exists a further

subsequence such that $\beta^{m_k} \rightarrow 0$. Note that by Armijo’s rule (15) in Step 3, we have

$$f_{\text{YF}}(\zeta^k + \beta^{m_k-1}d(\zeta^k)) - f_{\text{YF}}(\zeta^k) > -\sigma\beta^{2(m_k-1)}f_{\text{YF}}(\zeta^k).$$

Dividing by β^{m_k-1} both sides and passing the limit on the further subsequence, we obtain

$$\langle \nabla f_{\text{YF}}(\zeta^*), d(\zeta^*) \rangle \geq 0,$$

which yields that ζ^* is a solution of the SOCCP (1) by Lemma 2.2. \square

Proposition 3.2. *Let F be a continuously differentiable and strongly monotone mapping. Then the sequence $\{\zeta^k\}$ generated by Algorithm 3.1 converges to the unique solution of the SOCCP (1).*

Proof. The proof is routine (see [7]); however, we present it for completeness. We know that the property of bounded level sets is also held when F is strongly monotone, so following the same arguments as in the proof of Proposition 3.1, we again obtain that $\{\zeta^k\}$ has at least one accumulation point and any accumulation point is a solution of the SOCCP (1).

On the other hand, the strong monotonicity of F implies that the SOCCP (1) has at most one solution. To see this, say there are two solutions $\zeta^*, \zeta^* \in \mathbb{R}^n$ such that

$$\begin{cases} \langle F(\zeta^*), \zeta^* \rangle = 0, \\ F(\zeta^*) \in \mathcal{H}^n, \quad \zeta^* \in \mathcal{H}^n \end{cases} \quad \text{and} \quad \begin{cases} \langle F(\zeta^*), \zeta^* \rangle = 0, \\ F(\zeta^*) \in \mathcal{H}^n, \quad \zeta^* \in \mathcal{H}^n. \end{cases}$$

Since F is strongly monotone, we have $\langle F(\zeta^*) - F(\zeta^*), \zeta^* - \zeta^* \rangle > 0$. However,

$$\begin{aligned} \langle F(\zeta^*) - F(\zeta^*), \zeta^* - \zeta^* \rangle &= \langle F(\zeta^*), \zeta^* \rangle + \langle F(\zeta^*), \zeta^* \rangle - \langle F(\zeta^*), \zeta^* \rangle - \langle F(\zeta^*), \zeta^* \rangle \\ &= -\langle F(\zeta^*), \zeta^* \rangle - \langle F(\zeta^*), \zeta^* \rangle \\ &\leq 0, \end{aligned}$$

where the inequality is due to $F(\zeta^*), \zeta^*, F(\zeta^*), \zeta^*$ are all in \mathcal{H} . Hence, it is a contradiction and therefore there is at most one solution for the SOCCP (1).

From all the above, it says there is a unique solution ζ^* , so the entire sequence $\{\zeta^k\}$ must converge to ζ^* . \square

Propositions 3.1 and 3.2 may not be so surprising since they seem as expected. Nonetheless, we do not take them for granted before we prove them even though we think they should be true. Now, the results of Propositions 3.1 and 3.2 do fill up the gap in the literature. We notice that Lemma 2.3 plays an important role in the proofs for them. In fact, the assumption of strict feasibility is necessary for Lemma 2.3 to be held. For example, when $F(\zeta) \equiv 0$, every $\zeta \in \mathcal{H}^n$ is a solution of SOCCP (1) and hence the solution set is unbounded. In the following, we continue a further study of considering another (weaker) condition to replace this kind of *strict* condition by F being a R_{01} -function (will be defined in Definition 3.1) that is a new concept recently developed for linear and nonlinear transformations on Euclidean Algebra [18,25,29].

Definition 3.1. For a mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, it is called a

(a) R_{01} -function if for any sequence $\{\zeta^k\}$ such that

$$\|\zeta^k\| \rightarrow \infty, \quad \frac{(-\zeta^k)_+}{\|\zeta^k\|} \rightarrow 0, \quad \frac{(-F(\zeta^k))_+}{\|\zeta^k\|} \rightarrow 0, \tag{17}$$

we have

$$\liminf_{k \rightarrow \infty} \frac{\langle \zeta^k, F(\zeta^k) \rangle}{\|\zeta^k\|^2} > 0;$$

(b) R_{02} -function if for any sequence $\{\zeta^k\}$ such that (17), we have

$$\liminf_{k \rightarrow \infty} \frac{\omega(\zeta^k \circ F(\zeta^k))}{\|\zeta^k\|^2} > 0.$$

The above concepts are extensions of the ones defined for NCP and for SDCP. It is also known that every R_{01} -function is R_{02} -function [23, Lemma 4]; and if F has the uniform Jordan P -property (see [18,25,29]), then F is R_{02} -function [23, Lemma 5]. However, it is not clear whether uniform P -property (see [18,25,29]) implies R_{02} -function or not. With this new concept, Lemma 2.3 and Proposition 3.1 can be improved as Lemma 3.1 and Proposition 3.3, respectively. These results are significant not only because they are brand-new but also because there is no need of the assumption of strict feasibility therein.

Lemma 3.1. *Let f_{YF} be given as in (11). Suppose that F is a R_{01} -function. Then the level set*

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

is bounded for all $\gamma \geq 0$.

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 from Lemma 2.4 that $f_{YF}(\zeta^k) \rightarrow \infty$, which contradicts $\{\zeta^k\} \subset \mathcal{L}(\gamma)$. Therefore, $\{(-\zeta^k)_+\}$ and $\{(-F(\zeta^k))_+\}$ are bounded above, which says the conditions in (17) are satisfied. Then, by the assumption of R_{01} -function, we have

$$\liminf_{k \rightarrow \infty} \frac{\langle \zeta^k, F(\zeta^k) \rangle}{\|\zeta^k\|^2} > 0.$$

This yields $\langle \zeta^k, F(\zeta^k) \rangle \rightarrow \infty$, and hence $f_{YF}(\zeta^k) \rightarrow \infty$ by definition of f_{YF} given as in (11). Thus, it is a contradiction to $\{\zeta^k\} \subset \mathcal{L}(\gamma)$. \square

Proposition 3.3. *Let F be a continuously differentiable mapping. Suppose that F is R_{01} -function. Then the sequence $\{\zeta^k\}$ generated by Algorithm 3.1 has at least one accumulation point, and any accumulation point is a solution of the SOCCP (1).*

Proof. By applying Lemma 3.1 and following the same arguments as in Proposition 3.1, the desired results hold. We omit it. \square

From [23,29], the condition of R_{01} -function is weaker than strong monotonicity, and it is also weaker than monotonicity plus strict feasibility in certain sense. However, it is not clear yet whether R_{01} -function can be replaced by R_{02} -function in our brand-new results.

4. Numerical results

In this section, we report our computational experience with solving the symmetric affine SOCCPs generated randomly by the proposed algorithm, and compare the numerical performance with the FBMF approach [4]. Unless otherwise stated, the function f_{YF} in Algorithm 3.1 is always defined as in (11), where ψ_{YF} is defined by (9) and (10) with $\psi_0(t) = \frac{1}{2}(\max\{0, t\})^2$.

The symmetric affine SOCCP is stated as follows: finding $\zeta \in \mathbb{R}^n$ such that

$$\langle F(\zeta), \zeta \rangle = 0, \quad \zeta \in \mathcal{H}, \quad F(\zeta) = M\zeta + q \in \mathcal{H}, \tag{18}$$

where $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$ are a given symmetric positive semidefinite matrix and a vector, respectively. In our experiments, the matrix M and the vector q are generated by the following procedure. Elements of q were chosen randomly from the interval $[-1, 1]$ and the matrix M was obtained by setting $M = NN^T$, where N is a square matrix

whose nonzero elements are chosen randomly from the interval $[-1, 1]$. In this procedure, the number of nonzero elements of N is determined so that the nonzero density of M can be approximately estimated.

All experiments were done at a PC with 2.8 GHz CPU and 512 MB memory. The computer codes were all written in Matlab 6.1. To improve the numerical behavior of Algorithm 3.1, we replaced the standard Armijo rule by the nonmonotone line search as described in [19], i.e., we computed the smallest nonnegative integer m such that

$$f_{\text{YF}}(\zeta^k + \beta^m d(\zeta^k)) \leq \mathcal{W}_k - \sigma \beta^{2m} f_{\text{YF}}(\zeta^k), \tag{19}$$

where \mathcal{W}_k is given by

$$\mathcal{W}_k = \max_{j=k-m_k, \dots, k} f_{\text{YF}}(\zeta^j)$$

and where, for given nonnegative integers \hat{m} and s , we set

$$m_k = \begin{cases} 0 & \text{if } k \leq s, \\ \min\{m_{k-1} + 1, \hat{m}\} & \text{otherwise.} \end{cases} \tag{20}$$

Throughout the experiments, unless otherwise stated, we used the following parameters:

$$\hat{m} = 5, \quad s = 5, \quad \beta = 0.3 \quad \text{and} \quad \sigma = 1.0e - 4. \tag{21}$$

For the FBMF approach [4], we chose a limited-memory BFGS algorithm with five limited-memory vector-updates [3] to solve the unconstrained minimization reformulation (8) for the SOCCP (1). For the scaling matrix $H^0 = \gamma I$ in the BFGS algorithm, we adopted the choice of $\gamma = p^T q / p^T q$ recommended by [27, p. 226], where $p := \zeta - \zeta^{\text{old}}$ and $q := \nabla f_{\text{FB}}(\zeta) - \nabla f_{\text{FB}}(\zeta^{\text{old}})$. To ensure convergence, we revert to the steepest descent direction $-\nabla f_{\text{FB}}(\zeta)$ whenever the current direction Δ fails to satisfy the sufficient descent condition

$$\nabla f_{\text{FB}}(\zeta)^T \Delta \leq -10^{-4} \|\nabla f_{\text{FB}}(\zeta)\| \|\Delta\|.$$

In addition, we also employed the same nonmonotone line search as above to seek a suitable step-length, except that the parameter β is chosen as 0.2.

During the experiments, we started Algorithm 3.1 and the FBMF approach with the starting point $\zeta^0 = 0.001(1, 1, \dots, 1)^T$ and terminated the iterate once one of the following conditions is satisfied:

- (1) $\max\{\Psi(\zeta), |F(\zeta)^T \zeta|\} \leq 10^{-4}$, where Ψ represents f_{YF} or f_{FB} .
- (2) The number of iteration is over 50 000.
- (3) The step-length is lower than 10^{-16} .

We have done the following three groups of experiments.

Experiment A: Testing the influence of the scaling strategy on Algorithm 3.1 and the FBMF method. Note that, when the matrix M and the vector q in (18) are replaced by

$$\bar{M} = \frac{M}{w} \quad \text{and} \quad \bar{q} = \frac{q}{w}, \tag{22}$$

where $w \geq 1$ is a constant, the optimal solution of problem (18) does not change. Hence, in this experiment, we generated 10 test problems with sparsity 0.5% and 10% and $m = 10, n_1 = n_2 = \dots = n_m = 100$, and then solved each problem and their different scaled formulations with Algorithm 3.1 and the FBMF approach. Numerical results are summarized in Tables 1 and 2, where **NO.** represents the number of problem, **Den** denotes the approximate sparsity of M , **Nf** and **Time**, respectively, denote the total number of function evaluations and the CPU time for solving each problem.

From Tables 1 and 2, we see that, when $w > 1$, i.e., imposing the scaling strategy on the original problems, Algorithm 3.1 and the FBMF approach require much less function evaluations. Therefore, the scaling strategy in (22) can greatly improve the numerical performance of Algorithm 3.1 and the merit function approach. In particular, for those problems to which Algorithm 3.1 fails due to too small step-length, using the scaling strategy can yield satisfying solutions. This implies that Algorithm 3.1 has more dependence on the scaling strategy than the MF approach.

Experiment B: Testing Algorithm 3.1 and the FBMF approach on the affine SOCCP (18) with various degree of sparsity. In this experiment, we generated 10 test problems with $m = 1$ and $n = 1000$ for each nonzero density 0.1%,

Table 1
Numerical results of Algorithm 3.1 for the scaled problems

NO.	Den (%)	w = 1		w = 10		w = 50		w = 100	
		Nf	Time	Nf	Time	Nf	Time	Nf	Time
1	0.5	21 693	69.68	6716	21.71	4201	15.31	5608	21.89
2	0.5	55 916	175.9	26 234	85.81	17 836	68.15	24 073	98.15
3	0.5	11 897	44.57	989	3.82	803	3.40	1168	5.34
4	0.5	14 860	53.68	998	4.04	776	3.28	1047	5.07
5	0.5	13 260	48.53	553	2.01	553	2.32	733	3.20
6	10	–	–	2238	89.67	237	10.09	99	4.46
7	10	–	–	2518	95.54	264	10.64	114	5.21
8	10	–	–	8592	344.4	228	10.26	162	7.23
9	10	–	–	–	–	273	12.78	81	6.18
10	10	–	–	1982	82.60	239	10.98	125	5.56

“–” means that the iteration was stopped since the step-length was less than 10^{-16} .

Table 2
Numerical results of the FBMF method for the scaled problems

NO.	Den (%)	w = 1		w = 10		w = 50		w = 100	
		Nf	Time	Nf	Time	Nf	Time	Nf	Time
1	0.5	8135	56.96	4346	30.01	3076	22.14	4649	30.56
2	0.5	9086	57.56	14 020	91.06	16 619	117.5	21 284	149.9
3	0.5	611	3.95	531	3.70	812	5.89	976	6.82
4	0.5	1030	7.65	677	5.17	493	4.09	970	6.96
5	0.5	769	5.56	403	2.98	583	4.09	591	4.32
6	10	6682	488.0	807	64.15	132	9.65	100	7.40
7	10	4668	337.7	737	56.85	247	19.21	185	16.37
8	10	5639	431.1	812	63.82	131	10.12	114	9.20
9	10	4616	347.4	723	57.21	112	9.20	81	6.18
10	10	5818	452.6	702	59.12	220	17.04	96	7.59

Table 3
Numerical results for the affine SOCCP with sparsity 0.1%

NO.	Algorithm 3.1		MF method		NO.	Algorithm 3.1		MF method	
	Nf	Time	Nf	Time		Nf	Time	Nf	Time
1	597	0.76	369	1.10	2	*	*	*	*
3	539	0.85	325	0.98	4	*	*	*	*
5	*	*	*	*	6	*	*	*	*
7	254	0.34	127	0.33	8	*	*	*	*
9	*	*	*	*	10	799	0.95	143	0.28

“*” means that the iteration was stopped since the number of iteration was over 50 000.

0.5%, 1%, 10%, 50% and 80%, and then solved each problem with Algorithm 3.1 and the FBMF approach. Numerical results were summarized in Tables 3 and 4, where **Nf** and **Time** are same as Experiment A, **Gap** denotes the value of $|F(\zeta)^T \zeta|$ at the final iteration, and **Scale** in Table 4 denotes the value of w in (22). In particular, the values of **Gap**, **Nf** and **Time** in Table 4 are the averages of 10 trials for each sparsity.

From Table 3, it appears that Algorithm 3.1 and the FBMF approach have similar numerical performance on those problems with sparsity 0.1%. However, from Table 4, we see that, under the scaling strategy shown, Algorithm 3.1

Table 4
Numerical results for the affine SOCCP with different sparsity

Den (%)	Scale	Algorithm 3.1			MF approach		
		Gap	Nf	Time	Gap	Nf	Time
0.5	1	9.16e – 5	1597.1	3.28	8.77e – 5	914.9	4.09
1	1	8.10e – 5	7401.4	28.76	6.51e – 5	5016.8	38.31
10	10	6.89e – 5	402.6	15.37	6.85e – 5	312.0	21.92
50	100	5.90e – 5	472.2	17.07	6.59e – 5	568.7	39.00
80	100	4.81e – 5	468.4	17.54	7.62e – 5	668.7	45.23

Table 5
Numerical results for the affine SOCCP with different \mathcal{K}

m	Scale	Algorithm 3.1			MF approach		
		Gap	Nf	Time	Gap	Nf	Time
1	10	7.95e–5	201.3	7.64	8.759e–5	167.5	10.58
10	10	8.24e–5	497.6	17.68	8.74e–5	217.1	14.03
50	10	9.34e–5	1193	49.12	9.71e–5	266.5	19.42
100	100	5.75e–5	116.4	6.62	7.86e–5	138.7	11.44
200	100	4.54e–5	129.2	9.28	7.93e–5	149.2	15.69
500	100	7.09e–5	9719.4	1115.4	7.93e–5	149.2	15.69

Table 6
Numerical results of Algorithm 3.1 for different β

NO.	$\beta = 0.3$		$\beta = 0.1$		NO.	$\beta = 0.3$		$\beta = 0.1$	
	Nf	Time	Nf	Time		Nf	Time	Nf	Time
1	75	8.42	198	21.90	2	3081	355.5	142	15.23
3	19 001	2231.4	210	22.25	4	6644	769.0	179	19.09
5	69	7.93	178	19.00	6	75	8.37	804	88.67
7	62 169	7068.4	132	14.56	8	5939	689.6	295	32.23
9	77	8.54	208	23.03	10	64	7.12	144	16.03

always needed less CPU time than the FBMF approach although the former may require more function evaluations. In addition, we also observe that the number of function evaluations required by Algorithm 3.1 will become less when the sparsity of M becomes higher.

Experiment C: Testing Algorithm 3.1 and the FBMF approach on the affine SOCCP (18) with various Cartesian structures of \mathcal{K} . To construct SOCs of various types, we chose n_i and m such that $n_1 = n_2 = \dots = n_m$ and $n_1 + \dots + n_m = 2000$. For each type of \mathcal{K} , we solved 10 test problems with nonzero density 1% by Algorithm 3.1 and the FBMF approach, respectively. Numerical results were reported in Table 5, where **Scale**, **Gap**, **Nf** and **Time** are same as Experiment A, and particularly the values of **Gap**, **Nf** and **Time** are the averages of 10 trials for each type of \mathcal{K} .

From Table 5, we see that, under the scaling strategy shown, Algorithm 3.1 is comparable with the FBMF method for the first five groups of test problems whether in the CPU time or in the number of function evaluations. For the last group of test problems, Algorithm 3.1 obviously required more CPU time and function evaluations than the FBMF approach. However, from Table 6, we see that if **Scale** is still chosen as 100 but the parameter β in the line search is chosen as 0.1 instead of 0.3, the numerical performance of Algorithm 3.1 will have a great improvement, and moreover, the CPU time and the number of function evaluations needed are comparable with those of the FBMF method.

To sum up, for the symmetric affine SOCCPs in (18), if a suitable scaling strategy and the parameter β are used, Algorithm 3.1 will be comparable with, even superior to, the FBMF method in the CPU time for solving test prob-

lems although the former may require more function evaluations. Otherwise, the FBMF approach will be superior to Algorithm 3.1 whether in the CPU time or in the number of function evaluations.

5. Final Remarks

In this paper, we investigated a descent method for the equivalent reformulation (11) of the SOCCP which was also used for the NCP and the SDCP in literature, and proved its global convergence under some mild assumptions. Numerical comparison with the FBMF approach [4] for symmetric affine SOCCPs generated randomly indicate that the descent method is comparable with, even to superior to, the FBMF approach in the CPU time if a suitable scaling strategy and the parameter β in line search are adopted. We also expect that the method can be used to deal with large SOCCPs due to very small computational work per iteration. In addition, we notice that the proposed algorithm does not work for another reformulation (8) of the SOCCP since f_{FB} lacks property of bounded level sets (Lemma 2.3) where ψ_0 plays an important role therein.

Propositions 3.1 and 3.2 are more or less an afterthought of [4], nonetheless, it does parallel the extension to the SOCCP from the NCP and SDCP cases. On the other hand, this work does a further study based on replacing the conditions of monotonicity and strict feasibility by a new (and weaker under certain sense) so-called R_{01} -function. More specifically, under the new so-called R_{01} -function condition, the level sets of f_{YF} are still bounded and the proposed descent algorithm still has global convergence. These results are significant not only because they are brand-new but also because there is no need of the assumption of strict feasibility therein.

One future topic is to analyze the convergence rate theoretically which is more intractable. Other direction like weakening conditions which guarantees the property of bounded level sets is also interesting and worthwhile. There may be the direction as one referee pointed out which is to apply this optimization method to real-life studies, for example [6] and references therein.

Acknowledgments

The authors thank the referees for their careful reading of the paper and helpful suggestions.

References

- [1] F. Alizadeh, S. Schmieta, Symmetric cones, potential reduction methods, and word-by-word extensions, in: H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), *Handbook of Semidefinite Programming*, Kluwer Academic Publishers, Boston, 2000, pp. 195–233.
- [2] D.P. Bertsekas, *Nonlinear Programming*, second ed., Athena Scientific, Belmont, 1999.
- [3] R.H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.* 16 (1995) 1190–1208.
- [4] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, *Math. Programming* 104 (2005) 293–327.
- [5] X.-D. Chen, D. Sun, J. Sun, Complementarity functions and numerical experiments for second-order cone complementarity problems, *Comput. Optim. Appl.* 25 (2003) 39–56.
- [6] C.-T. Cheng, K.-W. Chau, Fuzzy iteration methodology for reservoir flood control operation, *J. Amer. Water Resour. Assoc.* 37 (2001) 1381–1388.
- [7] F. Facchinei, J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vols. I and II, Springer, New York, 2003.
- [8] F. Facchinei, J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, *SIAM J. Optim.* 7 (1997) 225–247.
- [9] U. Faraut, A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.
- [10] M.C. Ferris, O.L. Mangasarian, J.-S. Pang, (Eds.), *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.
- [11] M.C. Ferris, J.-S. Pang, (Eds.), *Complementarity and Variational Problems: State of the Art*, SIAM Publications, Philadelphia, PA, 1996.
- [12] M.C. Ferris, J.-S. Pang, Engineering and economic applications of complementarity problems, *SIAM Rev.* 39 (1997) 669–713.
- [13] A. Fischer, A special Newton-type optimization methods, *Optimization* 24 (1992) 269–284.
- [14] A. Fischer, Solution of the monotone complementarity problem with locally Lipschitzian functions, *Math. Programming* 76 (1997) 513–532.
- [15] R. Fletcher, *Practical Methods of Optimization*, second ed., Wiley-Interscience, Chichester, 1987.
- [16] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, *SIAM J. Optim.* 12 (2002) 436–460.
- [17] C. Geiger, C. Kanzow, On the resolution of monotone complementarity problems, *Comput. Optim. Appl.* 5 (1996) 155–173.
- [18] M.S. Gowda, R. Sznajder, J. Tao, Some P -properties for the linear transformations on Euclidean Jordan algebras, *Linear Algebra Appl.* 393 (2004) 203–232.

- [19] L. Grippo, F. Lampariello, S. Lucidi, A nonmonotone line search technique for Newton's method, *SIAM J. Numer. Anal.* 23 (1986) 707–716.
- [20] S. Hayashi, N. Yamashita, 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, 2001.
- [21] C. Kanzow, An unconstrained optimization technique for large scale linearly constrained convex minimization problems, *Computing* 53 (1994) 101–117.
- [22] C. Kanzow, Nonlinear complementarity as unconstrained optimization, *J. Optim. Theory Appl.* 88 (1996) 139–155.
- [23] Y.-J. Liu, Z.-W. Zhang, Y.-H. Wang, Some properties of a class of merit functions for symmetric cone complementarity problems, *Asia-Pacific J. Oper. Res.* 23 (2006) 473–495.
- [24] M.S. Lobo, L. Vandenbergh, S. Boyd, H. Lebret, Application of second-order cone programming, *Linear Algebra Appl.* 284 (1998) 193–228.
- [25] M. Malik, S.R. Mohan, On Q and R_0 properties of a quadratic representation in the linear complementarity problems over the second-order cone, *Linear Algebra Appl.* 397 (2005) 85–97.
- [26] R.D.C. Monteiro, T. Tsuchiya, Polynomial convergence of primal–dual algorithms for the second-order cone programs based on the MZ-family of directions, *Math. Programming* 88 (2000) 61–83.
- [27] J. Nocedal, S.J. Wright, *Numerical Optimization*, Springer, New York, 1999.
- [28] S. Schmieta, F. Alizadeh, Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones, *Math. Oper. Res.* 26 (2001) 543–564.
- [29] J. Tao, M.S. Gowda, Some P -properties for the nonlinear transformations on Euclidean Jordan Algebra, Technical Report, Department of Mathematics and Statistics, University of Maryland, 2004.
- [30] T. Tsuchiya, A convergence analysis of the scaling-invariant primal–dual path-following algorithms for second-order cone programming, *Optim. Methods Softw.* 11 (1999) 141–182.
- [31] R.J. Vanderbei, H.Y. Benson, On formulating semidefinite programming problems as smooth convex nonlinear optimization problems, ORFE 99-01, Department of Operations Research and Financial Engineering, Princeton University, Princeton, 1999.
- [32] N. Yamashita, M. Fukushima, A new merit function and a descent method for semidefinite complementarity problems, in: M. Fukushima, L. Qi (Eds.), *Reformulation—Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer Academic Publishers, Boston, 1999, pp. 405–420.