A SURVEY ON SOC COMPLEMENTARITY FUNCTIONS AND SOLUTION METHODS FOR SOCPS AND SOCCPS

Jein-Shan Chen* and Shaohua Pan†

This paper is dedicated to the memory of Professor Paul Tseng.

Abstract: This paper makes a survey on SOC complementarity functions and related solution methods for the second-order cone programming (SOCP) and second-order cone complementarity problem (SOCCP). Specifically, we discuss the properties of four classes of popular merit functions, and study the theoretical results of associated merit function methods and numerical behaviors in the solution of convex SOCPs. Then, we present suitable nonsinguarity conditions for the B-subdifferentials of the natural residual (NR) and Fischer-Burmeister (FB) nonsmooth system reformulations at a (locally) optimal solution, and test the numerical behavior of a globally convergent FB semismooth Newton method. Finally, we survey the properties of smoothing functions of the NR and FB SOC complementarity functions, and provide numerical comparisons of the smoothing Newton methods based on them. The theoretical results and numerical experience of this paper provide a comprehensive view on the development of this field in the past ten years.

Key words: *second-order cone, complementarity functions, merit functions, smoothing function, nonsmooth Newton methods, smoothing Newton methods*

Mathematics Subject Classification: *26B05, 26B35, 90C33, 65K05*

1 Introduction

The second-order cone (SOC) in $\mathbb{R}^n$ $(n \geq 1)$, also called the Lorentz cone, is defined as

$$\mathcal{K}^n := \left\{ (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_1 \geq \|x_2\| \right\},$$

where $\|\cdot\|$ denotes the Euclidean norm. If $n = 1$, then $\mathcal{K}^n$ is the set of nonnegative reals $\mathbb{R}_+$. We are interested in optimization and complementarity problems whose constraints involve the direct product of SOCs. In particular, we are interested in the SOC complementarity system which is to find vectors $x, y \in \mathbb{R}^n$ and $\zeta \in \mathbb{R}^l$ satisfying

$$x \in \mathcal{K}, \quad y \in \mathcal{K}, \quad \langle x, y \rangle = 0, \quad E(x, y, \zeta) = 0, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $E : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n \times \mathbb{R}^l$ is a continuously differentiable mapping, and $\mathcal{K}$ is the direct product of SOCs given by

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m} \tag{1.2}$$

with $m, n_1, \ldots, n_m \geq 1$, and $n_1 + \cdots + n_m = n$. Throughout this paper, corresponding to the structure of $\mathcal{K}$, we write $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m)$ with $x_i, y_i \in \mathbb{R}^{n_i}$.

A special case of (1.1) is the generalized second-order cone complementarity problem (SOCCP) which, for given two continuously differentiable mappings $F = (F_1, \ldots, F_m)$ and $G = (G_1, \ldots, G_m)$ with $F_i, G_i \colon \mathbb{R}^n \to \mathbb{R}^{n_i}$, is to *find a vector $\zeta \in \mathbb{R}^n$ such that*

$$F(\zeta) \in \mathcal{K}, \quad G(\zeta) \in \mathcal{K}, \quad \langle F(\zeta), G(\zeta) \rangle = 0. \tag{1.3}$$

When G becomes an identity one, (1.3) reduces to finding a vector $\zeta \in \mathbb{R}^n$ such that

$$\zeta \in \mathcal{K}, \quad F(\zeta) \in \mathcal{K}, \quad \langle \zeta, F(\zeta) \rangle = 0, \tag{1.4}$$

which is a direct extension of the NCPs studied well in the past 30 years (see [22, 24]). Another special case of (1.1) is the KKT conditions of the second-order cone programming

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & Ax = b, \quad x \in \mathcal{K} \end{aligned} \tag{1.5}$$

where $f \colon \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable function, A is an $m \times n$ matrix with full row rank, and $b \in \mathbb{R}^m$. When f is linear, (1.5) becomes the standard linear SOCP that has wide applications in engineering design, control, finance, management science, and so on; see [1, 35] and the references therein. In addition, system (1.1) arises directly from some engineering and practical problems; for example, the three-dimensional frictional contact problems [34] and the robust Nash equilibria [28].

During the past ten years, there appeared active research for SOCPs and SOCCPs, and various methods had been proposed which include the interior-point methods [1, 35, 41, 63, 51], the smoothing Newton methods [18, 25, 27], the semismooth Newton methods [32, 43], and the merit function methods [10, 12]. Among others, the last three kinds of methods are typically developed by an SOC complementarity function. Recall that a mapping $\phi \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is an *SOC complementarity function* associated with $\mathcal{K}^n$ if

$$\phi(x, y) = 0 \iff x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y \rangle = 0. \tag{1.6}$$

However, there are lack of comprehensive studies for the properties of SOC complementarity functions and the numerical behavior of related solution methods. In this work, we give a survey for popular SOC complementarity functions and the related merit function methods, semismooth Newton methods and smoothing Newton methods.

The squared norm of SOC complementarity functions gives a merit function associated with $\mathcal{K}^n$, where $\psi \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is called a merit function associated with $\mathcal{K}^n$ if

$$\psi(x, y) = 0 \iff x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y \rangle = 0. \tag{1.7}$$

Apart from this, there are other ways to construct merit functions; for example, the LT merit function in Subsection 3.3. Here we are interested in those smooth ψ so that the SOCCP (1.3) can be reformulated as an unconstrained smooth minimization problem

$$\min_{\zeta \in \mathbb{R}^n} \ \Psi(\zeta) := \sum_{i=1}^{m} \psi\left(F_i(\zeta), G_i(\zeta)\right), \tag{1.8}$$

in the sense that ζ^* is a solution to (1.3) if and only if it solves (1.8) with zero optimal value. This is the so-called merit function approach. Note that with a smooth merit function ψ, system (1.1) can also be reformulated as a smooth minization problem

$$\min_{(x,y,\zeta)\in\mathbb{R}^{2n+l}} \|E(x,y,\zeta)\|^2 + \psi(x,y),$$

but the reformulation is not effective for the solution of (1.1) due to the conflict between the feasibility and the decrease of complementarity gap involved in the objective. So, in this paper we consider the merit function methods for the SOCCP (1.3). In Section 3, we survey and compare the properties of four classes of popular smooth merit functions associated with $\mathcal{K}^n$. In Section 4, we focus on the theoretical results of corresponding merit function methods, and their numerical performance in the solution of linear SOCPs from DIMACS [52] and nonlinear convex SOCPs generated randomly.

With an SOC complementarity function ϕ associated with $\mathcal{K}^n$, we can rewrite (1.1) as

$$\Phi(z) = \Phi(x,y,\zeta) := \begin{pmatrix} E(x,y,\zeta) \\ \phi(x_1,y_1) \\ \vdots \\ \phi(x_m,y_m) \end{pmatrix} = 0, \tag{1.9}$$

By [22, Prop. 9.1.1], system (1.9) is effective only for those nondifferentiable but (strongly) semismooth ϕ. Two popular such ϕ are the vector-valued natural residual (NR) function $\phi_{\mathrm{NR}}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and Fischer-Burmeister (FB) function $\phi_{\mathrm{FB}}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$:

$$\phi_{\mathrm{NR}}(x,y) := x - (x-y)_+ \tag{1.10}$$

and

$$\phi_{\mathrm{FB}}(x,y) := (x+y) - (x^2+y^2)^{1/2}, \tag{1.11}$$

where $(\cdot)_+$ denotes the Euclidean projection onto $\mathcal{K}^n$, x^2 means the Jordan product of x and itself, and $x^{1/2}$ with $x \in \mathcal{K}^n$ is the unique square root of x such that $x^{1/2} \circ x^{1/2} = x$. The two nondifferentiable functions are strongly semismooth, where the proof for ϕ_{NR} can be found in [18, Prop. 4.3], [9, Prop. 7] or [27, Prop. 4.5], and the proof for ϕ_{FB} is given by Sun and Sun [58] and Chen [11] by using different techniques. In Section 5, we review the nonsingularity conditions for the B-subdifferentials of Φ at a solution of (1.1) without strict complementarity, and test the behavior of a global FB nonsmooth Newton method.

Let $\theta\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be a continuously differentiable on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++}$ with $\theta(\cdot,\cdot,0) \equiv \phi(\cdot,\cdot)$ for $\phi = \phi_{\mathrm{NR}}$ or ϕ_{FB}. Then (1.1) is also equivalent to the augmented system

$$\Theta(\omega) = \Theta(\varepsilon,x,y,\zeta) := \begin{pmatrix} \varepsilon \\ E(x,y,\zeta) \\ \theta(x_1,y_1,\varepsilon) \\ \vdots \\ \theta(x_m,y_m,\varepsilon) \end{pmatrix} = 0, \tag{1.12}$$

which is continuously differentiable in $\mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$. In the past several years, some smoothing Newton methods have been proposed for (1.1) by solving a sequence of smooth systems or a single augmented system (see, e.g., [25, 18, 27]), but there is no comprehensive study for their numerical performance. Motivated by the efficiency of the smoothing Newton

method [54], we in Section 6 apply it for the system (1.12) involving the CHKS smoothing function and the squared smoothing function of ϕ_{NR}, and the FB smoothing function, respectively, and compare their numerical behaviors. Similar to the NR and FB nonsmooth Newton methods, the locally superlinear (quadratic) convergence of these smoothing methods does not require the strict complementarity of solutions. So, these nonsmooth and smoothing methods are superior to interior point methods in theory since singular Jacobians will occur to the latter if strict complementarity is not satisfied.

Throughout this paper, I means an identity matrix of appropriate dimension, $\mathbb{R}^n$ ($n \geq 1$) denotes the space of n-dimensional real column vectors, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is identified with $\mathbb{R}^{n_1 + \cdots + n_m}$. For a given set S, we denote $\mathrm{int}(S)$ and $\mathrm{bd}(S)$ by the interior and boundary of S, respectively. For any $x \in \mathbb{R}^n$, we write $x \succeq_{\mathcal{K}^n} 0$ (respectively, $x \succ_{\mathcal{K}^n} 0$) to mean $x \in \mathcal{K}^n$ (respectively, $x \in \mathrm{int}(\mathcal{K}^n)$). For any differentiable $F \colon \mathbb{R}^n \to \mathbb{R}^l$, we denote $F'(x) \in \mathbb{R}^{l \times n}$ by the Jacobian of F at x, and $\nabla F(x)$ by the transposed Jacobian of F at x. A square matrix $B \in \mathbb{R}^{n \times n}$ is said to be positive definite if $\langle u, Bu \rangle > 0$ for all nonzero $u \in \mathbb{R}^n$, and B is said to be positive semidefinite if $\langle u, Bu \rangle \geq 0$ for all $u \in \mathbb{R}^n$.

2 Preliminaries

This section recalls some background materials that are needed in the subsequent sections. For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, their Jordan product [23] is defined by

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

The Jordan product, unlike scalar or matrix multiplication, is not associative, which is a main source on complication in the analysis of SOCCP. The identity element under this product is $e := (1, 0, \ldots, 0)^T \in \mathbb{R}^n$. For any given $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the matrix

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

will be used, which can be viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ given by $L_x y = x \circ y$.

For each $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, let $\lambda_1(x), \lambda_2(x)$ and $u_x^{(1)}, u_x^{(2)}$ be the spectral values and the corresponding spectral vectors of x, respectively, given by

$$\lambda_i(x) := x_1 + (-1)^i \|x_2\| \quad \text{and} \quad u_x^{(i)} := \frac{1}{2}\Big(1, \ (-1)^i \bar{x}_2\Big), \quad i = 1, 2$$

with $\bar{x}_2 = x_2 / \|x_2\|$ if $x_2 \neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{n-1}$ satisfying $\|\bar{x}_2\| = 1$. Then x admits a spectral factorization [23] associated with $\mathcal{K}^n$ in the form of

$$x = \lambda_1(x) u_x^{(1)} + \lambda_2(x) u_x^{(2)}.$$

When $x_2 \neq 0$, the spectral factorization is unique. The following lemma states the relation between the spectral factorization of x and the eigenvalue decomposition of L_x.

Lemma 2.1 ([25]). *For any given $x \in \mathbb{R}^n$, let $\lambda_1(x), \lambda_2(x)$ be the spectral values of x, and $u_x^{(1)}, u_x^{(2)}$ be the associated spectral vectors. Then, L_x has the eigenvalue decomposition*

$$L_x = U(x) \mathrm{diag}\left(\lambda_2(x), x_1, \cdots, x_1, \lambda_1(x)\right) U(x)^T$$

where

$$U(x) = \left(\sqrt{2}u_x^{(2)}, u_x^{(3)}, \cdots, u_x^{(n)}, \sqrt{2}u_x^{(1)}\right) \in \mathbb{R}^{n \times n}$$

is an orthogonal matrix, and $u_x^{(i)}$ for $i = 3, \ldots, n$ have the form of $(0, \bar{u}_i)$ with $\bar{u}_3, \ldots, \bar{u}_n$ being any unit vectors in $\mathbb{R}^{n-1}$ that span the linear subspace orthogonal to x_2.

By using Lemma 2.1, it is not hard to calculate the inverse of L_x whenever it exists:

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \dfrac{\det(x)}{x_1}I + \dfrac{1}{x_1}x_2x_2^T \end{bmatrix} \tag{2.1}$$

where $\det(x) := x_1^2 - \|x_2\|^2$ denotes the determinant of x.

By the spectral factorization above, for any given scalar function $g \colon \mathbb{R} \to \mathbb{J} \subseteq \mathbb{R}$, we may define the associated vector-valued function $g^{\mathrm{soc}} \colon \mathbb{R}^n \to S \subseteq \mathbb{R}^n$ by

$$g^{\mathrm{soc}}(x) := g(\lambda_1(x))u_x^{(1)} + g(\lambda_2(x))u_x^{(2)}. \tag{2.2}$$

For example, taking $g(t) = \sqrt{t}$ for $t \geq 0$, we have that $g^{\mathrm{soc}}(x) = x^{1/2}$ with $x \in \mathcal{K}^n$. The vector-valued g^{soc} inherits many desirable properties from g (see [9]). The following lemma provides the formulas to compute the Jacobian of g^{soc} and its inverse.

Lemma 2.2. *Let $g \colon \mathbb{R} \to \mathbb{J} \subseteq \mathbb{R}$ be a given scalar function, and $g^{\mathrm{soc}} \colon \mathbb{R}^n \to S \subseteq \mathbb{R}^n$ be defined by (2.2). If g is differentiable on $\mathrm{int}(\mathbb{J})$, then g^{soc} is differentiable in $\mathrm{int}(S)$ with*

$$\nabla g^{\mathrm{soc}}(x) = \begin{cases} g'(x_1)I & \text{if } x_2 = 0, \\ \begin{bmatrix} b(x) & c(x)\dfrac{x_2^T}{\|x_2\|} \\ c(x)\dfrac{x_2}{\|x_2\|} & a(x)I + (b(x) - a(x))\dfrac{x_2x_2^T}{\|x_2\|^2} \end{bmatrix} & \text{if } x_2 \neq 0 \end{cases}$$

for any $x = (x_1, x_2) \in \mathrm{int}(S)$, where

$$a(x) = \frac{g(\lambda_2(x)) - g(\lambda_1(x))}{\lambda_2(x) - \lambda_1(x)}, \ b(x) = \frac{g'(\lambda_2(x)) + g'(\lambda_1(x))}{2}, \ c(x) = \frac{g'(\lambda_2(x)) - g'(\lambda_1(x))}{2}.$$

If $\nabla g^{\mathrm{soc}}(\cdot)$ is invertible at $x \in \mathrm{int}(S)$, then letting $d(x) = b^2(x) - c^2(x)$, we have that

$$(\nabla g^{\mathrm{soc}}(x))^{-1} = \begin{cases} (g'(x_1))^{-1}I & \text{if } x_2 = 0, \\ \begin{bmatrix} \dfrac{b(x)}{d(x)} & -\dfrac{c(x)}{d(x)}\dfrac{x_2^T}{\|x_2\|} \\ -\dfrac{c(x)}{d(x)}\dfrac{x_2}{\|x_2\|} & \dfrac{1}{a(x)}I + \left(\dfrac{b(x)}{d(x)} - \dfrac{1}{a(x)}\right)\dfrac{x_2x_2^T}{\|x_2\|^2} \end{bmatrix} & \text{if } x_2 \neq 0. \end{cases}$$

Proof. The first part is direct by Prop. 5.2 of [25] or Prop. 5 of [9]. For the second part, it suffices to calculate the inverse of $\nabla g^{\mathrm{soc}}(x)$ when $x_2 \neq 0$. By the expression of ∇g^{soc}, it is easy to verify that $b(x) + c(x)$ and $b(x) - c(x)$ are the eigenvalues of $\nabla g^{\mathrm{soc}}(x)$ with $(1, \frac{x_2}{\|x_2\|})$ and $(1, -\frac{x_2}{\|x_2\|})$ being the corresponding eigenvectors, and $a(x)$ is the eigenvalue of multiplicity $n - 2$ with corresponding eigenvectors of the form $(0, \bar{v}_i)$, where $\bar{v}_1, \ldots, \bar{v}_{n-2}$ are any unit vectors in $\mathbb{R}^{n-1}$ that span the subspace orthogonal to x_2. By this, using an elementary calculation yields the formula of $(\nabla g^{\mathrm{soc}}(x))^{-1}$. $\qquad\square$

We next recall some joint properties of two mappings which are the direct extensions of the uniform Cartesian P-property [17], the uniform Jordan P-property [60], the weak coerciveness [67], and the R_0-property [6], respectively.

Definition 2.3. The mappings $F = (F_1, \ldots, F_m)$ and $G = (G_1, \ldots, G_m)$ are said to have

(i) **joint uniform Cartesian P-property** if there exists a constant $\varrho > 0$ such that, for every $\zeta, \xi \in \mathbb{R}^n$, there exists an index $\nu \in \{1, 2, \ldots, m\}$ such that

$$\langle F_\nu(\zeta) - F_\nu(\xi)), G_\nu(\zeta) - G_\nu(\xi) \rangle \geq \varrho \|\zeta - \xi\|^2.$$

(ii) **joint uniform Jordan P-property** if there exists a constant $\varrho > 0$ such that, for every $\zeta, \xi \in \mathbb{R}^n$,

$$\lambda_2 \left[(F(\zeta) - F(\xi)) \circ (G(\zeta) - G(\xi)) \right] \geq \varrho \|\zeta - \xi\|^2.$$

(iii) **joint Cartesian weak coerciveness** if there is an element $\xi \in \mathbb{R}^n$ such that

$$\lim_{\|\zeta\| \to \infty} \max_{1 \leq i \leq m} \frac{\langle G_i(\zeta) - G_i(\xi), F_i(\zeta) \rangle}{\|\zeta - \xi\|} = +\infty.$$

(iv) **joint Cartesian strong coerciveness** if the last equation holds for all $\xi \in \mathbb{R}^n$.

(v) **joint Cartesian R_0^w-property** if, for any sequence $\{\zeta^k\} \subseteq \mathbb{R}^n$ satisfying

$$\|\zeta^k\| \to +\infty, \ \limsup_{k \to \infty} \|(F(\zeta^k))_-\| < +\infty, \ \limsup_{k \to \infty} \|(G(\zeta^k))_-\| < +\infty,$$

there holds that

$$\limsup_{k \to \infty} \max_{1 \leq i \leq m} \langle F_i(\zeta^k), G_i(\zeta^k) \rangle = +\infty.$$

It is easy to see that the joint uniform Cartesian P-property implies the joint Cartesian strong coerciveness. From the arguments in [47], it follows that the joint uniform Cartesian P-property implies the joint uniform Jordan P-property, and the joint Cartesian weak coerciveness with respect to an element ξ with $G(\xi) \in \mathcal{K}$ implies the joint Cartesian R_0^w-property. Now we are not clear whether the joint uniform Jordan P-property implies the joint Cartesian weak coerciveness. Note that the above several properties do not imply the joint monotonicity of F and G, but the joint monotonicity of F and G with some additional conditions may imply their joint Cartesian R_0^w-property; see the remarks after Prop. 4.2. The following definition recalls the concept of linear growth of a mapping, which is weaker than the global Lipschitz continuity.

Definition 2.4. A mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to have **linear growth** if there exists a constant $C > 0$ such that $\|F(\zeta)\| \leq \|F(0)\| + C\|\zeta\|$ for any $\zeta \in \mathbb{R}^n$.

We next introduce the Cartesian (strict) column monotonicity of matrices M and N, which is weaker than the (strict) column monotonicity introduced in [22, page 1014] and [37, page 222]. Particularly, when N is invertible, this property reduces to the Cartesian P_0 (P)-property of the matrix $-N^{-1}M$ introduced by Chen and Qi [17].

Definition 2.5. The matrices $M, N \in \mathbb{R}^{n \times n}$ are said to be

(i) **Cartesian column monotone** if for any $u, v \in \mathbb{R}^n$ with $u \neq 0, v \neq 0$,

$$Mu + Nv = 0 \implies \exists \nu \in \{1, \ldots, m\} \text{ s.t. } u_\nu \neq 0 \text{ and } \langle u_\nu, v_\nu \rangle \geq 0.$$

(ii) **Cartesian strictly column monotone** if for any $u, v \in \mathbb{R}^n$ with $(u, v) \neq (0, 0)$,

$$Mu + Nv = 0 \implies \exists \nu \in \{1, \ldots, m\} \text{ s.t. } \langle u_\nu, v_\nu \rangle > 0.$$

To close this section, we recall the concept of B-subdifferential for a locally Lipschitz continuous mapping. If $H \colon \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous, then the set

$$\partial_B H(z) := \left\{ V \in \mathbb{R}^{m \times n} \mid \exists \{z^k\} \subseteq D_H : z^k \to z, \ H'(z^k) \to V \right\}$$

is nonempty and called the *B-subdifferential* [55] of H at z, where $D_H \subseteq \mathbb{R}^n$ is the set of points at which H is differentiable. The convex hull of $\partial_B H(z)$ is called the generalized Jacobian of Clarke [20], i.e. $\partial H(z) = \operatorname{conv} \partial_B H(z)$. We assume that the reader is familiar with the concept of (strong) semismoothness, and refer to [49, 55, 56] for the details.

Unless otherwise stated, in the rest of this paper, we assume that $F = (F_1, \ldots, F_m)$ and $G = (G_1, \ldots, G_m)$ with $F_i, G_i \colon \mathbb{R}^n \to \mathbb{R}^{n_i}$ are continuously differentiable. For a given $x \in \mathbb{R}^l$ for some $l \geq 2$, we write $x = (x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$, where x_1 is the first component of the vector x and x_2 consists of the remaining $l - 1$ components of x.

3 Merit Functions Associated with $\mathcal{K}^n$

This section reviews four classes of smooth merit functions associated with $\mathcal{K}^n$ and their properties related to the merit function approach. The nondifferentiable NR function

$$\psi_{\mathrm{NR}}(x, y) := \| x - (x - y)_+ \|^2 \quad \forall x, y \in \mathbb{R}^n \tag{3.1}$$

is needed, which plays a crucial role in error bound estimations of other merit functions.

3.1 Implicit Lagrangian Function

The implicit Lagrangian $\psi_{\mathrm{MS}} \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, parameterized by $\alpha > 1$, is defined as

$$\begin{aligned}
\psi_{\mathrm{MS}}(x, y) &:= \max_{u, v \in \mathcal{K}^n} \left\{ \langle x, y - v \rangle - \langle y, u \rangle - \frac{1}{2\alpha}(\| x - u \|^2 + \| y - v \|^2) \right\} \\
&= \langle x, y \rangle + \frac{1}{2\alpha} \left(\| (x - \alpha y)_+ \|^2 - \| x \|^2 + \| (y - \alpha x)_+ \|^2 - \| y \|^2 \right).
\end{aligned} \tag{3.2}$$

The function is introduced by Mangasarian and Solodov [38] for NCPs, and extended to semidefinite complementarity problems (SDCPs) by Tseng [61] and general symmetric cone complementarity problems (SCCPs) by Kong et al. [33]. By Theorem 3.2(b) of [33], ψ_{MS} is a merit function induced by the trace of the SOC complementarity function

$$\phi_{\mathrm{MS}}(x, y) := x \circ y + \frac{1}{2\alpha} \left[(x - \alpha y)_+^2 - x^2 + (y - \alpha x)_+^2 - y^2 \right] \quad \forall x, y \in \mathbb{R}^n, \alpha > 1. \tag{3.3}$$

The following results are extensions of known results, particularly [62, 65, 39], for NCPs.

Lemma 3.1. *For any fixed $\alpha > 1$ and all $x, y \in \mathbb{R}^n$, we have the following results.*

(a) $\psi_{\mathrm{MS}}(x, y) = 0 \iff x \in \mathcal{K}^n, \ y \in \mathcal{K}^n, \ \langle x, y \rangle = 0 \iff \phi_{\mathrm{MS}}(x, y) = 0.$

(b) ϕ_{MS} *and* ψ_{MS} *are continuously differentiable everywhere, with*

$$\begin{aligned}
\nabla_x \psi_{\mathrm{MS}}(x, y) &= y + \alpha^{-1} \left((x - \alpha y)_+ - x \right) - (y - \alpha x)_+, \\
\nabla_y \psi_{\mathrm{MS}}(x, y) &= x + \alpha^{-1} \left((y - \alpha x)_+ - y \right) - (x - \alpha y)_+.
\end{aligned}$$

(c) *The gradient function $\nabla\psi_{\mathrm{MS}}$ is globally Lipschitz continuous.*

(d) $\langle x, \nabla_x\psi_{\mathrm{MS}}(x,y)\rangle + \langle y, \nabla_y\psi_{\mathrm{MS}}(x,y)\rangle = 2\psi_{\mathrm{MS}}(x,y).$

(e) $\langle \nabla_x\psi_{\mathrm{MS}}(x,y), \nabla_y\psi_{\mathrm{MS}}(x,y)\rangle \geq 0.$

(f) $\psi_{\mathrm{MS}}(x,y) = 0$ *if and only if* $\nabla_x\psi_{\mathrm{MS}}(x,y) = 0$ *and* $\nabla_y\psi_{\mathrm{MS}}(x,y) = 0.$

(g) $(\alpha-1)\|\phi_{\mathrm{NR}}(x,y)\|^2 \geq \psi_{\mathrm{MS}}(x,y) \geq (1-\alpha^{-1})\|\phi_{\mathrm{NR}}(x,y)\|^2.$

(h) $\alpha^{-1}(\alpha-1)^2\psi_{\mathrm{MS}}(x,y) \leq \|\nabla_x\psi_{\mathrm{MS}}(x,y) + \nabla_y\psi_{\mathrm{MS}}(x,y)\|^2 \leq 2\alpha(\alpha-1)\psi_{\mathrm{MS}}(x,y).$

Proof. The proofs of parts (a)–(b) and (e)–(f) are given in [33]. Parts (c)–(d) are direct by the expressions of ψ_{MS} and $\nabla\psi_{\mathrm{MS}}$. Part (g) is a direct application of [62, Prop. 2.2] with $\tilde{\pi} = -\psi_{\mathrm{MS}}$. Part (h) is easily shown by [50, Theorem 4.2] and (b) and (g). $\qquad\square$

Analogous to the NCPs and SDCPs, the implicit Lagrangian has the most favorable properties among all projection merit functions. So, we do not review others in this class.

$\boxed{3.2}$ Fischer-Burmeister (FB) Merit Function

From [25], ϕ_{FB} in (1.11) is an SOC complementarity function, and whence its squared norm

$$\psi_{\mathrm{FB}}(x,y) := \frac{1}{2}\|\phi_{\mathrm{FB}}(x,y)\|^2. \tag{3.4}$$

is a merit function associated with $\mathcal{K}^n$. The function ψ_{FB} was shown to be continuously differentiable everywhere with globally Lipschitz continuous gradient [10, 16], although ϕ_{FB} itself is not differentiable. Recently, we extend these favorable properties of ψ_{FB} to the following one-parametric class of merit functions (see [14, 15]):

$$\psi_\tau(x,y) := \frac{1}{2}\|\phi_\tau(x,y)\|^2, \tag{3.5}$$

where $\tau \in (0,4)$ is an arbitrary fixed parameter and $\phi_\tau \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$\phi_\tau(x,y) := (x+y) - \left[(x-y)^2 + \tau(x \circ y)\right]^{1/2}. \tag{3.6}$$

Clearly, when $\tau = 2$, ψ_τ becomes the FB merit function ψ_{FB}. The one-parametric class of functions was originally proposed by Kanzow and Kleinmichel [31] for NCPs, and was proved to share all desirable properties of the FB NCP function. The following lemma summarizes those properties of ψ_τ used in the merit function approach.

Lemma 3.2. *For any fixed $\tau \in (0,4)$ and all $x, y \in \mathbb{R}^n$, we have the following results.*

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

(b) ψ_τ *is continuously differentiable everywhere with* $\nabla_x\psi_\tau(0,0) = \nabla_y\psi_\tau(0,0) = 0.$ *Also, if* $w = (x-y)^2 + \tau(x \circ y) \in \mathrm{int}(\mathcal{K}^n)$, *then*

$$
\begin{aligned}
\nabla_x\psi_\tau(x,y) &= \left(I - L_{x+\frac{\tau-2}{2}y}L_{\sqrt{w}}^{-1}\right)\phi_\tau(x,y), \\
\nabla_y\psi_\tau(x,y) &= \left(I - L_{y+\frac{\tau-2}{2}x}L_{\sqrt{w}}^{-1}\right)\phi_\tau(x,y);
\end{aligned}
$$

and if $(x - y)^2 + \tau(x \circ y) \in \mathrm{bd}(\mathcal{K}^n)$ *and* $(x, y) \neq (0, 0)$,

$$\nabla_x \psi_\tau(x, y) = \left[1 - \frac{x_1 + \frac{\tau-2}{2} y_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2) x_1 y_1}} \right] \phi_\tau(x, y),$$

$$\nabla_y \psi_\tau(x, y) = \left[1 - \frac{y_1 + \frac{\tau-2}{2} x_1}{\sqrt{x_1^2 + y_1^2 + (\tau - 2) x_1 y_1}} \right] \phi_\tau(x, y).$$

(c) *The gradient function $\nabla \psi_\tau$ is globally Lipschitz continuous.*

(d) $\langle x, \nabla_x \psi_\tau(x, y) \rangle + \langle y, \nabla_y \psi_\tau(x, y) \rangle = 2 \psi_\tau(x, y).$

(e) $\langle \nabla_x \psi_\tau(x, y), \nabla_y \psi_\tau(x, y) \rangle \geq 0$, *with equality holding if and only if $\psi_\tau(x, y) = 0$.*

(f) $\psi_\tau(x, y) = 0 \iff \nabla_x \psi_\tau(x, y) = 0 \iff \nabla_y \psi_\tau(x, y) = 0.$

(g) *There exist constant $c_1 > 0$ and $c_2 > 0$ independent on x, y such that*

$$c_1 \|\phi_{\mathrm{NR}}(x, y)\| \leq \|\phi_\tau(x, y)\| \leq c_2 \|\phi_{\mathrm{NR}}(x, y)\|.$$

(h) *There exist constants $C_1 > 0$ and $C_2 > 0$ only dependent on n, τ such that*

$$C_1 \|\phi_\tau(x, y)\| \leq \|\nabla_x \psi_\tau(x, y) + \nabla_y \psi_\tau(x, y)\| \leq C_2 \|\phi_\tau(x, y)\|.$$

Proof. The proofs of parts (a)–(b) and (d)–(e) can be found in [14]. Part (c) is proved in [15, Theorem 3.1]. Part (f) follows by parts (a), (b) and (e). Parts (g) and (h) are established in [3]. □

Comparing Lemma 3.2 with Lemma 3.1, we see that the functions ψ_{FB} and ψ_{MS} share with similar favorable properties, but the properties (e)–(f) of ψ_{FB} are stronger than those of ψ_{MS}, which make ψ_{FB} require a weaker stationary point condition; see Prop. 4.1.

It should be pointed out that the squared norms of Evtushenko and Purtov [21] SOC complementarity functions $\phi_\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\phi_\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, defined as

$$\phi_\alpha(x, y) := -(x \circ y) + \frac{1}{2\alpha}(x + y)_-^2 \qquad 0 < \alpha \leq 1,$$

$$\phi_\beta(x, y) := -(x \circ y) + \frac{1}{2\beta}\left((x)_-^2 + (y)_-^2\right) \quad 0 < \beta < 1, \qquad (3.7)$$

also provide the smooth merit functions ψ_α and ψ_β associated with $\mathcal{K}^n$. But, since they do not enjoy the property (e) of ψ_τ or the weaker property (e) of ψ_{MS}, it is hard to find the conditions to guarantee that every stationary point of Ψ_α and Ψ_β is a solution of SOCCPs (see the proof of Prop. 4.1). In addition, unlike in the setting of NCPs, the squared norm of penalized FB SOC complementarity function is not smooth even nondifferentiable. So, this paper does not include these functions.

3.3 Luo and Tseng (LT) Merit Function

The third class of smooth merit functions is an extension of the class of functions introduced by Luo and Tseng [37] for NCPs, and subsequently extended to SDCPs in [61, 66]. In the setting of SOCs, this class of functions is defined as

$$\psi_{\mathrm{LT}}(x, y) := \psi_0\left(\langle x, y \rangle\right) + \widehat{\psi}(x, y), \quad \forall x, y \in \mathbb{R}^n \qquad (3.8)$$

where $\psi_0 : \mathbb{R} \to \mathbb{R}_+$ is an arbitrary smooth function satisfying

$$\psi_0(0) = 0, \quad \psi_0'(t) = 0 \quad \forall t \leq 0, \text{ and } \psi_0'(t) > 0 \quad \forall t > 0 \tag{3.9}$$

and $\widehat{\psi} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is an arbitrary smooth function such that

$$\widehat{\psi}(x,y) = 0, \ \langle x,y \rangle \leq 0 \iff x \in \mathcal{K}^n, \ y \in \mathcal{K}^n, \ \langle x,y \rangle = 0. \tag{3.10}$$

The requirment for ψ_0 is a little different from the original LT merit functions [37]. There are many functions satisfying (3.9) such as the polynomial function $q^{-1} \max(0,t)^q$ $(q \geq 2)$, the exponential function $\exp(\max(0,t)^2) - 1$, and logarithmic function $\ln(1 + \max(0,t)^2)$. In addition, there are many choices for $\widehat{\psi}$ such as ψ_{MS}, ψ_τ and the following

$$\widehat{\psi}_1(x,y) := \frac{1}{2} \left(\|(x)_-\|^2 + \|(y)_-\|^2 \right) \quad \text{and} \quad \widehat{\psi}_2(x,y) := \frac{1}{2} \|\phi_{\mathrm{FB}}(x,y)_+\|^2. \tag{3.11}$$

In this paper, we are particularly interested in three subclasses of ψ_{LT} with $\widehat{\psi}$ chosen as ψ_{FB}, $\widehat{\psi}_1$ and $\widehat{\psi}_2$. Among others, ψ_{LT} with $\widehat{\psi} = \psi_{\mathrm{FB}}$ is an analog of the merit function studied by Yamashita and Fukushima [66] for SDCPs. In view of this, we write ψ_{LT} with $\widehat{\psi} = \psi_{\mathrm{FB}}$ as ψ_{YF}. We also write ψ_{LT} with $\widehat{\psi} = \widehat{\psi}_1$ and $\widehat{\psi}_2$ as ψ_{LT1} and ψ_{LT2}, respectively.

Lemma 3.3. *Let ψ be one of the functions $\psi_{\mathrm{YF}}, \psi_{\mathrm{LT1}}$ and ψ_{LT2}. Then, for all $x, y \in \mathbb{R}^n$,*

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

(b) *ψ is continuously differentiable everywhere. Furthermore,*

$$\begin{aligned}
\nabla_x \psi_{\mathrm{YF}}(x,y) &= \psi_0'(\langle x,y \rangle) y + \nabla_x \psi_{\mathrm{FB}}(x,y), \\
\nabla_y \psi_{\mathrm{YF}}(x,y) &= \psi_0'(\langle x,y \rangle) x + \nabla_y \psi_{\mathrm{FB}}(x,y),
\end{aligned}$$

where $\nabla_x \psi_{\mathrm{FB}}$ and $\nabla_y \psi_{\mathrm{FB}}$ are given by Lemma 3.2(c) with $\tau = 2$;

$$\nabla_x \psi_{\mathrm{LT1}}(x,y) = \psi_0'(\langle x,y \rangle) y + (x)_-, \quad \nabla_y \psi_{\mathrm{LT1}}(x,y) = \psi_0'(\langle x,y \rangle) x + (y)_-;$$

when $\psi = \psi_{\mathrm{LT2}}$, $\nabla_x \psi_{\mathrm{LT2}}(0,0) = \nabla_y \psi_{\mathrm{LT2}}(0,0) = 0$, and if $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$,

$$\begin{aligned}
\nabla_x \psi_{\mathrm{LT2}}(x,y) &= \psi_0'(\langle x,y \rangle) y + \left(I - L_x L_{(x^2+y^2)^{1/2}}^{-1} \right) \phi_{\mathrm{FB}}(x,y)_+, \\
\nabla_y \psi_{\mathrm{LT2}}(x,y) &= \psi_0'(\langle x,y \rangle) x + \left(I - L_y L_{(x^2+y^2)^{1/2}}^{-1} \right) \phi_{\mathrm{FB}}(x,y)_+,
\end{aligned}$$

and if $x^2 + y^2 \in \mathrm{bd}^+(\mathcal{K}^n)$,

$$\begin{aligned}
\nabla_x \psi_{\mathrm{LT2}}(x,y) &= \psi_0'(\langle x,y \rangle) y + \left(1 - \frac{x_1}{\sqrt{x_1^2 + y_1^2}} \right) \phi_{\mathrm{FB}}(x,y)_+, \\
\nabla_y \psi_{\mathrm{LT2}}(x,y) &= \psi_0'(\langle x,y \rangle) x + \left(1 - \frac{y_1}{\sqrt{x_1^2 + y_1^2}} \right) \phi_{\mathrm{FB}}(x,y)_+.
\end{aligned}$$

(c) *The gradient $\nabla \psi$ is globally Lipschitz continuous on any bounded set of $\mathbb{R}^n \times \mathbb{R}^n$.*

(d) $\langle x, \nabla_x \psi(x,y) \rangle + \langle y, \nabla_y \psi(x,y) \rangle \geq 2\psi_0'(\langle x,y \rangle) \langle x,y \rangle + 2\psi(x,y) \geq 2\widehat{\psi}(x,y).$

(e) $\langle \nabla_x \psi(x,y), \nabla_y \psi(x,y) \rangle \geq 0$, *and when* $\psi = \psi_{\mathrm{YF}}$ *and* ψ_{LT2}, $\langle \nabla_x \psi(x,y), \nabla_y \psi(x,y) \rangle = 0$ *if and only if* $\psi(x,y) = 0$.

(f) *When* $\psi = \psi_{\mathrm{YF}}$ *and* ψ_{LT2}, $\psi(x,y) = 0 \Longleftrightarrow \nabla_x \psi(x,y) = 0 \Longleftrightarrow \nabla_y \psi(x,y) = 0$; *and when* $\psi = \psi_{\mathrm{LT1}}$, $\psi(x,y) = 0 \Longleftrightarrow \nabla_x \psi(x,y) = 0$ *and* $\nabla_y \psi(x,y) = 0$.

(g) *If* ψ_0 *is convex and nondecreasing in* $\mathbb{R}$, *then* ψ_{LT1} *is a convex function over* $\mathbb{R}^n \times \mathbb{R}^n$.

Proof. When $\psi = \psi_{\mathrm{YF}}$, from the definition of ψ_{YF} and Lemma 3.2(a) and (c)–(d), we readily get parts (a)–(c); parts (d)–(e) are easily verified by using part (b), Lemma 3.2(e) with $\tau = 2$ and equation (3.9). When $\psi = \psi_{\mathrm{LT1}}$ and ψ_{LT2}, parts (a)–(b) and (d)–(e) are established in Prop. 3.1 and Prop. 3.2 of [12] except the smoothness of ψ_{LT2}, which is implied by Lemma 1 of appendix. Part (c) is immediate by using the expressions of $\nabla \psi_{\mathrm{LT1}}$ and $\nabla \psi_{\mathrm{LT2}}$ and noting that $\nabla \widehat{\psi}$ is globally Lipschitz continuous on $\mathbb{R}^n \times \mathbb{R}^n$.

When $\psi = \psi_{\mathrm{YF}}$ and ψ_{LT2}, part (f) follows by parts (b) and (e), and when $\psi = \psi_{\mathrm{LT1}}$, part (f) follows by parts (b) and (d). By Prop. 3.1(b) of [12], $\widehat{\psi}_1$ is convex over $\mathbb{R}^n \times \mathbb{R}^n$. Since ψ_0 is convex and nondecreasing in $\mathbb{R}$, it is easy to verify that $\psi_0(\langle x,y \rangle)$ is also convex over $\mathbb{R}^n \times \mathbb{R}^n$. So, we obtain the result of part (g). $\qquad\square$

Comparing Lemma 3.3 with Lemmas 3.1 and 3.2, we observe that ψ_{MS} and ψ_τ have two remarkable advantages over the LT class of merit functions: one is the positive homogeneity of ψ_{MS} and ψ_τ, which makes the corresponding merit functions for SOCCPs overcome the bad-scaling of problems; the other is that their gradients have the same growth as the merit function itself, which is the key to establish convergence rate of some descent algorithms. It should be pointed out that although the LT class of merit functions does not possess the property (g) of ψ_{MS} and ψ_{FB}, the corresponding merit functions for the SOCCPs may provide a global error bound under a weaker condition (see Prop. 4.3), and moreover, Lemma 3.5 below shows that they have faster growth than ψ_{MS} and ψ_{FB}.

3.4   A Variant of LT Merit Function

A variant of the LT merit functions is the function $\widehat{\psi}_{\mathrm{LT}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ defined by

$$\widehat{\psi}_{\mathrm{LT}}(x,y) := \psi_0(\|(x \circ y)_+\|^2) + \widehat{\psi}(x,y) \quad \forall x,y \in \mathbb{R}^n, \tag{3.12}$$

where ψ_0 satisfies the first and the third properties of (3.9) and $\widehat{\psi}$ satisfies (3.10). This class of merit functions was considered by Chen [12]. In this work we are interested in $\widehat{\psi}_{\mathrm{LT}}$ with $\widehat{\psi} = \psi_{\mathrm{FB}}, \widehat{\psi}_1$ and $\widehat{\psi}_2$, and write them as $\widehat{\psi}_{\mathrm{YF}}, \widehat{\psi}_{\mathrm{LT1}}$ and $\widehat{\psi}_{\mathrm{LT2}}$, successively.

Lemma 3.4. *Let* ψ *be one of the functions* $\widehat{\psi}_{\mathrm{YF}}, \widehat{\psi}_{\mathrm{LT1}}$ *and* $\widehat{\psi}_{\mathrm{LT2}}$. *Then, for all* $x,y \in \mathbb{R}^n$,

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

(b) ψ *is continuously differentiable everywhere, with*

$$\begin{aligned} \nabla_x \psi(x,y) &= 2\psi_0' \left(\|(x \circ y)_+\|^2 \right) L_y(x \circ y)_+ + \nabla_x \widehat{\psi}(x,y), \\ \nabla_y \psi(x,y) &= 2\psi_0' \left(\|(x \circ y)_+\|^2 \right) L_x(x \circ y)_+ + \nabla_y \widehat{\psi}(x,y), \end{aligned}$$

where $\nabla_x \widehat{\psi}(x,y)$ *and* $\nabla_y \widehat{\psi}(x,y)$ *are same as in Lemma 3.3.*

(c) *The gradient* $\nabla \psi$ *is globally Lipschitz continuous on any bounded set of* $\mathbb{R}^n \times \mathbb{R}^n$.

(d) $\langle x, \nabla_x \psi(x,y) \rangle + \langle y, \nabla_y \psi(x,y) \rangle = 4\psi_0' \left(\|(x \circ y)_+\|^2 \right) \|(x \circ y)_+\|^2 + 2\widehat{\psi}(x,y).$

(e) $\psi(x,y) = 0 \iff \nabla_x \psi(x,y) = 0$ *and* $\nabla_y \psi(x,y) = 0.$

Proof. The proofs are same as that of Lemma 3.3, and we omit them. $\qquad\square$

For the class of merit functions $\widehat{\psi}_{\mathrm{LT}}$, it is difficult to establish the following inequality

$$\langle \nabla_x \widehat{\psi}_{\mathrm{LT}}(x,y), \nabla_y \widehat{\psi}_{\mathrm{LT}}(x,y) \rangle \geq 0 \quad \forall x, y \in \mathbb{R}^n$$

although numerical simulations show that they possess the property. The main difficulty is to estimate the terms $\langle L_y(x \circ y)_+, \nabla_y \widehat{\psi}(x,y) \rangle$ and $\langle L_x(x \circ y)_+, \nabla_x \widehat{\psi}(x,y) \rangle.$

To close this section, we characterize the growth of the above merit functions via a lemma, whose proof is direct by the arguments of [47, Sec. 4] and the remarks after it.

Lemma 3.5. *If the sequence* $\{(x^k, y^k)\} \subseteq \mathbb{R}^n \times \mathbb{R}^n$ *satisfies one of the conditions:*

(i) $\liminf_{k\to\infty} \lambda_1(x^k) = -\infty$ *or* $\liminf_{k\to\infty} \lambda_1(y^k) \to -\infty;$

(ii) $\{\lambda_1(x^k)\}$ *and* $\{\lambda_1(y^k)\}$ *are bounded below,* $\lambda_2(x^k), \lambda_2(y^k) \to +\infty,$ *and* $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \nrightarrow 0;$

(iii) $\{\lambda_1(x^k)\}$ *and* $\{\lambda_1(y^k)\}$ *are bounded below, and* $\limsup_{k\to\infty} \langle x^k, y^k \rangle = +\infty,$

then $\limsup_{k\to\infty} \psi(x^k, y^k) \to \infty$ *for* $\psi = \psi_{\mathrm{YF}}, \psi_{\mathrm{LT1}}, \psi_{\mathrm{LT2}}, \widehat{\psi}_{\mathrm{YF}}, \widehat{\psi}_{\mathrm{LT1}}$ *and* $\widehat{\psi}_{\mathrm{LT2}}$. *If* $\{(x^k, y^k)\}$ *satisfies (i) or (ii), then* $\limsup_{k\to\infty} \psi(x^k, y^k) \to \infty$ *with* $\psi = \psi_{\mathrm{NR}}, \psi_{\mathrm{MS}}, \psi_\tau.$

The condition (ii) of Lemma 3.5 implies the condition (iii) since, when $\{\lambda_1(x^k)\}$ and $\{\lambda_1(y^k)\}$ are bounded below and $\lambda_2(x^k), \lambda_2(y^k) \to +\infty$, there must exist a vector $d \in \mathbb{R}^n$ such that $x^k - d \in \mathcal{K}^n$ and $y^k - d \in \mathcal{K}^n$, which along with $\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \nrightarrow 0$ yields that $\frac{\langle x^k, y^k \rangle}{\|x^k\|\|y^k\|} \to c > 0$ (taking a subsequence if necessary), and $\limsup_{k\to\infty} \langle x^k, y^k \rangle = +\infty$ then follows. Hence, ψ_{LT} and its variant $\widehat{\psi}_{\mathrm{LT}}$ have faster growth than ψ_τ and ψ_{MS}.

4 Merit Function Approach and Applications

This section is devoted to the merit function methods for the generalized SOCCP (1.3), which yields a solution of (1.3) by solving an unconstrained minimization (1.8) with ψ being one of the merit functions introduced in last section. Throughout this section, we assume that $\mathcal{K}$ has the Cartesian structure of (1.2), and for any $\zeta \in \mathbb{R}^n$, write

$$\nabla_x \psi(F(\zeta), G(\zeta)) = \left(\nabla_{x_1} \psi(F_1(\zeta), G_1(\zeta)), \dots, \nabla_{x_m} \psi(F_m(\zeta), G_m(\zeta)) \right),$$
$$\nabla_y \psi(F(\zeta), G(\zeta)) = \left(\nabla_{y_1} \psi(F_1(\zeta), G_1(\zeta)), \dots, \nabla_{y_m} \psi(F_m(\zeta), G_m(\zeta)) \right).$$

When applying effective gradient-type methods for the problem (1.8), we expect only a stationary point due to the nonconvexity of merit functions. Thus, it is necessary to know what conditions can guarantee every stationary point of Ψ to be a solution of (1.3). The following proposition provides a suitable condition for the first three classes of functions.

Proposition 4.1. *Let* Ψ *be given by (1.8) with* ψ *being one of the previous merit functions.*

(a) *When $\psi = \psi_\tau, \psi_{\mathrm{YF}}$ and ψ_{LT2}, every stationary point of Ψ is a solution of (1.3) if $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ are Cartesian column monotone for any $\zeta \in \mathbb{R}^n$.*

(b) *When $\psi = \psi_{\mathrm{MS}}$ or ψ_{LT1}, every stationary point of Ψ is a solution of (1.3) if $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ are Cartesian strictly column monotone for any $\zeta \in \mathbb{R}^n$.*

Proof. Since F and G are continuously differentiable, by Lemmas 3.1–3.3(b), the function Ψ is continuously differentiable with

$$\nabla \Psi(\zeta) = \nabla F(\zeta) \nabla_x \psi(F(\zeta), G(\zeta)) + \nabla G(\zeta) \nabla_y \psi(F(\zeta), G(\zeta)). \tag{4.1}$$

Let $\zeta \in \mathbb{R}^n$ be an arbitrary but fixed stationary point of the function Ψ. Then,

$$\nabla F(\zeta) \nabla_x \psi(F(\zeta), G(\zeta)) + \nabla G(\zeta) \nabla_y \psi(F(\zeta), G(\zeta)) = 0. \tag{4.2}$$

Suppose that ζ is not a solution of (1.3). When $\psi = \psi_\tau, \psi_{\mathrm{YF}}$ and ψ_{LT2}, we must have $\nabla_x \psi(F(\zeta), G(\zeta)) \neq 0$ and $\nabla_y \psi(F(\zeta), G(\zeta)) \neq 0$ by Lemma 3.2– 3.3(f). Since $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ are Cartesian column monotone, equality (4.2) implies that there exists an index $\nu \in \{1, \ldots, m\}$ such that $\nabla_{x_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)) \neq 0$ and

$$\langle \nabla_{x_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)), \nabla_{y_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)) \rangle \leq 0.$$

Along with Lemma 3.2–3.3(e), we have $\psi(F_\nu(\zeta), G_\nu(\zeta)) = 0$. This, by Lemma 3.2–3.3(f), implies $\nabla_{x_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)) = 0$, and then we get a contradiction. When $\psi = \psi_{\mathrm{MS}}$ or ψ_{LT1}, by Lemma 3.1 and 3.3(f) we have $(\nabla_x \psi(F(\zeta), G(\zeta)), \nabla_y \psi(F(\zeta), G(\zeta)) \neq (0, 0)$. Since $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ are Cartesian strictly column monotone, (4.2) implies that there exists an index $\nu \in \{1, \ldots, m\}$ such that $\nabla_{x_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)) \neq 0$ and

$$\langle \nabla_{x_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)), \nabla_{y_\nu} \psi(F_\nu(\zeta), G_\nu(\zeta)) \rangle < 0,$$

which is impossible by Lemma 3.1 and 3.3(e). The proof is completed. $\qquad \square$

When $\nabla G(\zeta)$ is invertible, since the Cartesian (strict) column monotonicity of $\nabla F(\zeta)$ and $-\nabla G(\zeta)$ is equivalent to the Cartesian (P) P_0-property of $\nabla G(\zeta)^{-1} \nabla F(\zeta)$, Prop. 4.1 extends the results of [10, Prop. 3] and [12, Prop. 3.3] for ψ_{YF} and ψ_{LT2}, respectively, as well as recovers the result of [44, Prop. 5.1]. When G is an identity mapping, in view of Lemmas 3.1– 3.3(e), using the same arguments as in [33, Theorem 5.3] can prove that every regular stationary point of Ψ is a solution of (1.3). From [33], we know that the regularity is weaker than the Cartesian P-property of ∇F, but it is not clear whether it is weaker than the Cartesian P_0-property of ∇F.

The property that $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y) \rangle \geq 0$ for all $x, y \in \mathbb{R}^n$ plays a crucial role in the proof of Prop. 4.1. For the variant of LT functions $\widehat{\Psi}_{\mathrm{LT}}$, since the desirable property is not established, we can not provide suitable stationary point conditions for $\widehat{\Psi}_{\mathrm{LT}}$.

When solving the minimization problem (1.8), to guarantee that the iterative sequence generated has a limit point, it is necessary to require that Ψ has bounded level sets, which is implied by the coerciveness of Ψ, i.e., $\limsup_{\|\zeta\| \to \infty} \Psi(\zeta) = +\infty$. The following proposition provides the weakest coerciveness conditions for the previous merit functions.

Proposition 4.2. *Suppose the mappings F and G satisfy one of the following conditions:*

(C.1) *F and G have the joint uniform Jordan P-property and the linear growth.*

(C.2) *F and G have the joint Cartesian weak coerciveness and the linear growth.*

(C.3) *F and G are the joint Cartesian R_0^w-functions.*

Then $\Psi_{\mathrm{NR}}, \Psi_{\mathrm{MS}}$ and Ψ_τ are coercive under (C.1) or (C.2); $\Psi_{\mathrm{LT1}}, \Psi_{\mathrm{LT2}}, \widehat{\Psi}_{\mathrm{LT1}}$ and $\widehat{\Psi}_{\mathrm{LT2}}$ are coercive under (C.3); and Ψ_{YF} and $\widehat{\Psi}_{\mathrm{YF}}$ are coercive under one of (C.1)–(C.3).

Proof. The arguments are similar to those of [47], and we here omit them. □

From the remarks after Def. 2.3, we know that the functions $\Psi_{\mathrm{NR}}, \Psi_{\mathrm{MS}}$ and Ψ_τ require a stronger coerciveness condition than the functions $\Psi_{\mathrm{YF}}, \Psi_{\mathrm{LT1}}, \Psi_{\mathrm{LT2}}$ and $\widehat{\Psi}_{\mathrm{YF}}, \widehat{\psi}_{\mathrm{LT1}}, \widehat{\Psi}_{\mathrm{LT2}}$. It is worthwhile to point out, when F and G are jointly monotone mappings satisfying $\lim_{\|\zeta\|\to\infty} \|F(\zeta)\| + \|G(\zeta)\| = \infty$ and the generalized SOCCP (1.3) is strictly feasible, i.e., there exists $\hat{\zeta} \in \mathbb{R}^n$ such that $F(\hat{\zeta}), G(\hat{\zeta}) \in \mathrm{int}(\mathcal{K})$, using the same arguments as those of [47, Prop. 3.4] yields that F and G have the joint Cartesian R_0^w-property. Thus, Prop. 4.2 extends the results of [10, Prop. 6] and [12, Prop. 4.3]. In addition, this proposition also extends the results obtained in [45, 44] for the boundedness of level sets of $\Psi_{\mathrm{NR}}, \Psi_{\mathrm{MS}}, \Psi_\tau$.

Next we review the global error bound results for the previous merit functions. These results play a key role in establishing the convergence rate of merit function approaches.

Proposition 4.3. *Let ζ^* be a solution of the SOCCP (1.3) and Ψ be given by (1.8) with ψ being one of the merit functions introduced in last section. Then,*

(a) *when $\psi = \psi_{\mathrm{YF}}, \psi_{\mathrm{LT1}}, \psi_{\mathrm{LT2}}$ and $\widehat{\psi}_{\mathrm{YF}}, \widehat{\psi}_{\mathrm{LT1}}, \widehat{\psi}_{\mathrm{LT2}}$, if F and G have the joint uniform Cartesian P-property, then there exists a constant $\kappa > 0$ such that for any $\zeta \in \mathbb{R}^n$,*

$$\kappa\|\zeta - \zeta^*\|^2 \leq \left[\psi_0^{-1}(\Psi(\zeta))\right]^{1/2} + \Psi(\zeta)^{1/2}; \tag{4.3}$$

(b) *when $\psi = \psi_{\mathrm{NR}}, \psi_{\mathrm{MS}}$ and ψ_{FB}, if F and G are Lipschitz continuous and have joint uniform Cartesian P-property, then there exist constants $\kappa_1, \kappa_2 > 0$ such that*

$$\kappa_1 \Psi(\zeta) \leq \|\zeta - \zeta^*\|^2 \leq \kappa_2 \Psi(\zeta), \quad \forall \zeta \in \mathbb{R}^n. \tag{4.4}$$

Proof. (a) From Def. 2.3(i), there exist an index $\nu \in \{1, \ldots, m\}$ and $\varrho > 0$ such that

$$\varrho\|\zeta - \zeta^*\|^2 \leq \langle F_i(\zeta) - F_i(\zeta^*), G_i(\zeta) - G_i(\zeta^*)\rangle \quad \forall \zeta \in \mathbb{R}^n.$$

Using the same arguments as Prop. 5 of [10], we obtain that for any $\zeta \in \mathbb{R}^n$,

$$\hat{\kappa}\|\zeta - \zeta^*\|^2 \leq \langle F_i(\zeta), G_i(\zeta)\rangle + \|(F_i(\zeta))_-\| + \|(G_i(\zeta))_-\|$$

where $\hat{\kappa} := \dfrac{\varrho}{\max\{1, \|F_i(\zeta^*)\|, \|G_i(\zeta^*)\|\}}$. Note that, for any $\zeta \in \mathbb{R}^n$,

$$\langle F_i(\zeta), G_i(\zeta)\rangle \leq (F_i(\zeta)^T G_i(\zeta))_+ \text{ or } \sqrt{2}\|(F_i(\zeta) \circ G_i(\zeta)_+\|.$$

From the increasing of $\psi_0^{-1}(t)$ on $[0, +\infty)$ and the nonnegativity of ψ_{FB} and $\widehat{\psi}$, we get

$$(F_i(\zeta)^T G_i(\zeta))_+ \text{ or } \|(F_i(\zeta) \circ G_i(\zeta))_+\| \leq \left[\psi_0^{-1}(\psi(F_i(\zeta), G_i(\zeta)))\right]^{1/2} \leq \left[\psi_0^{-1}(\Psi(\zeta))\right]^{1/2}.$$

In addition, using Lemma 8 of [10], it is easy to verify that

$$\begin{aligned}
\|(F_i(\zeta))_-\| + \|(G_i(\zeta))_-\| &\leq 2\sqrt{2}\psi_{\mathrm{FB}}(F_i(\zeta), G_i(\zeta))^{1/2} \text{ or } 2\widehat{\psi}(F_i(\zeta), G_i(\zeta))^{1/2} \\
&\leq 2\sqrt{2}\Psi(\zeta)^{1/2}.
\end{aligned}$$

Combining the last two equations, we readily obtain the result in (4.3) with $\kappa = \hat{\kappa}/(2\sqrt{2})$. (b) When $\psi = \psi_{\mathrm{NR}}$ and ψ_{MS}, the result is established in Theorem 6.3 of [33]. When $\psi = \psi_{\mathrm{FB}}$, the result is direct by Lemma 3.2(g) and the result for Ψ_{NR}. $\qquad \square$

The constant κ in Prop. 4.3 is dependent on ζ^* and ϱ, whereas κ_1, κ_2 are also dependent on the Lipschitz constants of F and G besides ζ^* and ϱ. This proposition shows that the LT class of merit functions Ψ_{LT} and its variant $\widehat{\Psi}_{\mathrm{LT}}$ may provide an upper global error estimation under a weaker condition than the functions $\Psi_{\mathrm{NR}}, \Psi_{\mathrm{MS}}$ and Ψ_{FB}, although the latter also provides a lower global error estimation. Prop. 4.3(a) extends the results of [10, Prop. 5] and [12, Prop. 4.1]. From this proposition, we readily recover the error bound results of these merit functions in the setting of NCPs [30, 37].

Up to now, we have established the theoretical foundations of the merit function methods when applying existing gradient-type descent methods, for example the BFGS method, for solving (1.8). Except these existing minimization methods, we may develop new descent algorithms for the unconstrained reformulation (1.8) in view of the following proposition. These methods have some attractive features for some special classes of SOCCPs.

Proposition 4.4. *Let ψ be one of the first three classes of merit functions in last section. Suppose that $\nabla G(\zeta)$ is invertible for every $\zeta \in \mathbb{R}^n$, and $\bar{\zeta}$ is not a solution of (1.3). Let*

$$d(\zeta, \beta) := -\left[\nabla G(\zeta)^{-1}\right]^T \left[\nabla_x \psi(F(\zeta), G(\zeta)) + \beta \nabla_y \psi(F(\zeta), G(\zeta))\right] \quad \forall \zeta \in \mathbb{R}^n. \qquad (4.5)$$

(a) *Then, when $\psi = \psi_\tau, \psi_{\mathrm{YF}}$ and ψ_{LT2}, $\langle d(\bar{\zeta}, \beta), \nabla \Psi(\bar{\zeta}) \rangle < 0$ for sufficiently small $\beta > 0$, provided that $\nabla G(\bar{\zeta})^{-1} \nabla F(\bar{\zeta})$ is positive semidefinite.*

(b) *When $\psi = \psi_{\mathrm{MS}}$ and ψ_{LT1}, $\langle d(\bar{\zeta}, \beta), \nabla \Psi(\bar{\zeta}) \rangle < 0$ for sufficiently small $\beta > 0$, provided that $\nabla G(\bar{\zeta})^{-1} \nabla F(\bar{\zeta})$ is positive definite.*

Proof. Using the definition of $d(\zeta, \beta)$ and formula (4.1), we calculate that

$$\begin{aligned}
\langle d(\zeta, \beta), \nabla \Psi(\zeta) \rangle &= -\langle \nabla_x \psi(F(\zeta), G(\zeta)), \ \nabla G(\zeta)^{-1} \nabla F(\zeta) \nabla_x \psi(F(\zeta), G(\zeta)) \rangle \\
&\quad -\langle \nabla_x \psi(F(\zeta), G(\zeta)), \nabla_y \psi(F(\zeta), G(\zeta)) \rangle \\
&\quad -\beta \langle \nabla_x \psi(F(\zeta), G(\zeta)), \ \nabla G(\zeta)^{-1} \nabla F(\zeta) \nabla_y \psi(F(\zeta), G(\zeta)) \rangle \\
&\quad -\beta \|\nabla_y \psi(F(\zeta), G(\zeta))\|^2.
\end{aligned}$$

Together with Lemmas 3.1–3.3(e)–(f), we obtain the desired results. $\qquad \square$

Motivated by Prop. 4.4 and the descent algorithms for NCPs (see, e.g., [65, 39, 68]), we next utilize the direction $d(\zeta, \beta)$ to design an algorithm for the generalized SOCCPs that involve an invertible ∇G and a positive definite (or positive semidefinite) $\nabla G^{-1} \nabla F$.

Algorithm 4.1

Step 0. Let ψ be from the first three classes of merit functions of last section. Choose a point $\zeta^0 \in \mathbb{R}^n$, $\sigma \in (0, 1/2)$ and $\gamma, \beta \in (0, 1)$ with $\gamma < \beta$. Set $k := 0$.

Step 1. If $\Psi(\zeta^k) = 0$, then stop and ζ^k is a solution of the SOCCP.

Step 2. Let l_k be the smallest nonnegative integer l satisfying

$$\begin{aligned}
\Psi(\zeta^k &+ \gamma^l d(\zeta^k, \beta^l)) - \Psi(\zeta^k) \\
&\leq -\sigma \gamma^{2l} \|\nabla_x \psi(F(\zeta^k), G(\zeta^k)) + \nabla_y \psi(F(\zeta^k), G(\zeta^k))\|^2 \qquad (4.6)
\end{aligned}$$

where $d(\zeta, \beta)$ is defined as in (4.5), and set

$$d^k(\beta^{l_k}) := d(\zeta^k, \beta^{l_k}) \;\; \text{and} \;\; \zeta^{k+1} := \zeta^k + \gamma^{l_k} d^k(\beta^{l_k}).$$

Step 3. Let $k := k + 1$, and then go to Step 1.

When applying Algorithm 4.1 for solving the SOCCP (1.4), the computation of the search direction $d(\zeta, \beta)$ and the stepsize does not require the gradients of F. This makes the method suitable for large-scale problems, as well as applications where the derivatives of $F(\cdot)$ are not available or are costly to compute.

By Prop. 4.4, when $\nabla G^{-1} \nabla F$ is positive definite (semidefinite), the direction $d(\zeta, \beta)$ is descent for sufficiently small $\beta > 0$. To achieve this goal, we start Algorithm 4.1 with some reasonable small β and adapt iteratively by decreasing it if the linesearch step fails or the algorithm does not appear to make predicted progress. Such adaptive choice for the constant β is also adopted in descent algorithms for NCPs; see [39, 68].

For Algorithm 4.1, by Lemmas 3.1–3.3(e)–(f), using the same arguments as those of [46, Theorem 4.1] yields the following global convergence results.

Theorem 4.5. *Suppose that $\nabla G(\zeta)$ is invertible for every $\zeta \in \mathbb{R}^n$, and $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ is positive semidefinite for every $\zeta \in \mathbb{R}^n$ when $\psi = \psi_\tau, \psi_{\mathrm{YF}}$ and ψ_{LT2}, and $\nabla G(\zeta)^{-1} \nabla F(\zeta)$ is positive definite for every $\zeta \in \mathbb{R}^n$ when $\psi = \psi_{\mathrm{MS}}$ and ψ_{LT1}. Then, Algorithm 4.1 is well defined for any initial point ζ^0, and any accumulation point of the sequence $\{\zeta^k\}$ generated by Algorithm 4.1 is a solution of the SOCCP (1.3).*

When F and G satisfy the corresponding conditions of Prop. 4.2, the sequence $\{\zeta^k\}$ generated by Algorithm 4.1 has always an accumulation point. Together with Theorem 4.5, clearly, when G is an identity mapping and F is strongly monotone and Lipschitz continuous, the sequence $\{\zeta^k\}$ converges to the unique solution of (1.3). Particularly, for $\psi = \psi_{\mathrm{MS}}$ and ψ_{FB}, using the properties (e) and (g)–(h) of Lemmas 3.1 and 3.2, and the same arguments as those of [46, Theorem 5.1], we may establish the R-linear rate of convergence of Algorithm 4.1 under strong monotonicity.

Theorem 4.6. *Let $\Psi = \Psi_{\mathrm{MS}}$ or Ψ_{FB}. Suppose that $\nabla G(\zeta)$ is invertible for any $\zeta \in \mathbb{R}^n$, and F and G are jointly strongly monotone and have linear growth. If ∇F and ∇G are Lipschitz continuous on the set $L_\Psi(\zeta) := \{\zeta \in \mathbb{R}^n : \Psi(\zeta) \leq \Psi(\zeta^0)\}$, then the sequence $\{\Psi(\zeta^k)\}$ converges to zero Q-linearly and $\{\zeta^k\}$ converges R-linearly to the solution of (1.3).*

When $\psi = \psi_{\mathrm{YF}}$ and ψ_{LT}, we can not establish the convergence rate of Algorithm 4.1 since it is not clear whether the inequality $\|\nabla_x \psi(x,y) + \nabla_y \psi(x,y)\|^2 \geq c_1 \psi(x,y)$ for some $c_1 > 0$ holds or not, although they have more desirable error bound results.

The discussions above show that Ψ_{FB} seems to possess more desirable properties than $\Psi_{\mathrm{MS}}, \Psi_{\mathrm{YF}}, \Psi_{\mathrm{LT1}}$ and Ψ_{LT2}, although for some properties the LT merit functions $\Psi_{\mathrm{YF}}, \Psi_{\mathrm{LT1}}$ and Ψ_{LT2} need weaker conditions; for example, the coercive property and the global error bound property. In the next subsections, we compare their numerical performance by applying these merit function methods for convex SOCPs and SOCCPs.

4.1 Applications in the Solution of Convex SOCPs

We employ the merit function methods to solve the SOCP (1.5) with a twice continuously differentiable convex f. From [10], it follows that the KKT optimality conditions of (1.5)

can be reformulated as the generalized SOCCP (1.3) with

$$F(\zeta) = \hat{x} + (I - A^T(AA)^{-1})\zeta \ \text{ and } \ G(\zeta) = \nabla f(F(\zeta)) - A^T(AA^T)^{-1}A\zeta$$

where $\hat{x} \in \mathbb{R}^n$ satisfies $Ax = b$. We apply the BFGS method for the minimization reformulation (1.8) with ψ chosen as one of the previous merit functions.

In the BFGS method, we revert to the steepest descent direction $-\nabla\Psi(\zeta)$ whenever $p^T q \leq 10^{-5}\|p\|\|q\|$, where $p := \zeta - \zeta^{\text{old}}$ and $q := \nabla\Psi(\zeta) - \nabla\Psi(\zeta^{\text{old}})$. We adopt a nonmonotone line search as described in [26] to seek a suitable stepsize, i.e., we compute the smallest nonnegative integer l_k such that

$$\Psi(\zeta^k + \rho^{l_k} d^k) \leq \mathcal{W}_k + \sigma\rho^{l_k}\nabla\Psi(\zeta^k)^T d^k$$

where d^k denotes the direction in the kth iteration, ρ and σ are parameters in $(0,1)$, and $\mathcal{W}_k = \max_{j=k-m_k,\cdots,k}\Psi(\zeta^j)$ and where, for a given nonnegative integer $\hat{m}$ and s, we set

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

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

$$\rho = 0.5, \ \ \sigma = 10^{-4}, \ \ \hat{m} = 5 \ \text{ and } \ s = 5.$$

Since the SOCP instances to solve have a large n, we employ a limited-memory BFGS method [4] (L-BFGS, for short) with 5 limited-memory vector-updates to solve (1.8). In the L-BFGS method, we adopt the choice of $\gamma = p^T q/q^T q$ recommended by [42, page 226] for the scaling matrix $H^0 = \gamma I$, where p and q are same as above.

All tests were done at a PC of Pentium 4 with 2.8GHz CPU and 512MB memory. The computer codes were written in Matlab 6.5. During the testing, we computed the vector $\hat{x}$ in F as a solution of $\min_d \|Ad - b\|$ using Matlab's least square solver. We evaluated F and G with the Cholesky factorization of AA^T, which is efficient when A is sparse. In particular, given such a factorization $RR^T = AA^T$, we compute $x = F(\zeta)$ and $y = G(\zeta)$ for each ζ via two matrix-vector multiplications and two forward/backward solvers:

$$Ru = A\zeta, \ \ \ R^T v = u, \ \ \ w = A^T v, \ \ \ x = d + \zeta - w \ \ \ y = \nabla f(x) - w.$$

We started the BFGS method with $\zeta^0 = 0$, and terminated the iteration whenever one of the following conditions was satisfied: **(1)** $\Psi(\zeta) \leq \epsilon_1$ and $|F(\zeta)^T G(\zeta)| \leq \epsilon_2$; **(2)** The step-length is less than 10^{-12}; **(3)** The number of iteration is over $k_{\max}$. The parameter α in Ψ_{MS} is always chosen as 50. For ψ_0 in Ψ_{LT} and $\widehat{\Psi}_{\text{LT}}$, we select $\psi_0(t) = \max(0, t)^2/2$.

We tested two groups of instances for convex SOCPs. The first one is composed of four standard linear SOCPs from the DIMACS Implementation Challenge library [52]. We solved them by use of the L-BFGS method with $\epsilon_1 = 10^{-6}$, $\epsilon_2 = 10^{-4}$ and $k_{\max} = 10^5$. Numerical results are reported in Table 1, in which **NF** denotes the function evaluations for solving each problem, **Iter** means the number of iterations needed for each problem, $\Psi(\zeta^k)$ denotes the merit function value at the final iteration, and for Ψ_{LT1} and Ψ_{LT2}, the results on the second line are obtained by choosing $\psi_0(t) = \log(1 + \max(0, t)^2)$.

From Table 1, we see that all merit function methods can not yield desired result for the difficult "nb_L1" within 10^5 iterations, the implicit Lagrangian function method can not

Table 1: **Numerical results for linear SOCPs from DIMACS Library**

Name	nb			nb_L1			nb_L2			nb_L2_bessel		
	NF	Iter	$\Psi(\zeta^k)$	NF	Iter	$\Psi(\zeta^k)$	NF	Iter	$\Psi(\zeta^k)$	NF	Iter	$\Psi(\zeta^k)$
$\Psi_{\rm FB}$	3340	2776	8.19e–7	14200	10000	4.52e–3	634	429	1.22e–9	190	143	7.34e–8
$\Psi_{\rm MS}$	9908	8619	4.98e–5*	*	*	*	843	603	9.60e–9	189	162	8.84e–7
$\Psi_{\rm YF}$	2029	1615	9.58e–7	15257	10000	3.02e–3	519	282	6.87e–10	249	183	7.85e–7
$\widehat{\Psi}_{\rm YF}$	4421	3688	7.81e–7	11671	10000	5.29e–3	634	458	6.21e–10	109	96	1.77e–7
$\Psi_{\rm LT1}$	2147	1484	2.17e–8	15046	10000	1.73e–3	1590	1189	9.79e–7	921	664	8.90e–7
	760	559	8.08e–8	16539	10000	2.03e–3	1852	1468	9.98e–7	664	537	9.25e–7
$\Psi_{\rm LT2}$	1882	1391	2.15e–8	18414	10000	6.03e–4	1850	1341	9.73e–7	640	531	8.66e–7
	924	746	7.62e–8	21050	10000	7.82e–4	1488	1115	9.98e–7	792	628	9.79e–7
$\widehat{\Psi}_{\rm LT1}$	12269	10000	1.33e–9	10319	10000	2.50e–3	1647	1277	1.64e–9	1002	905	1.09e–10
$\widehat{\Psi}_{\rm LT2}$	12376	10000	1.57e–9	10191	10000	2.38e–3	1480	1147	1.46e–9	1029	918	2.66e–11

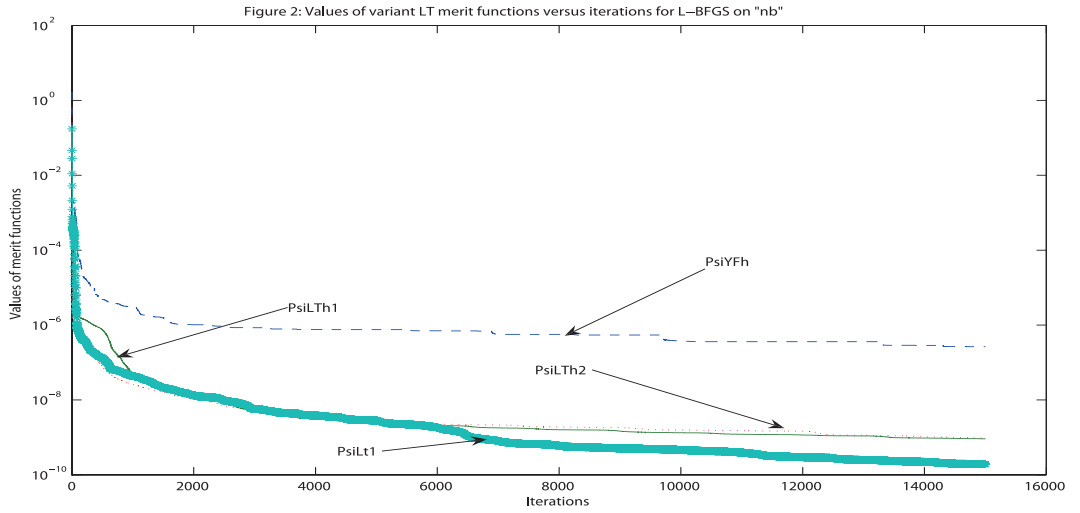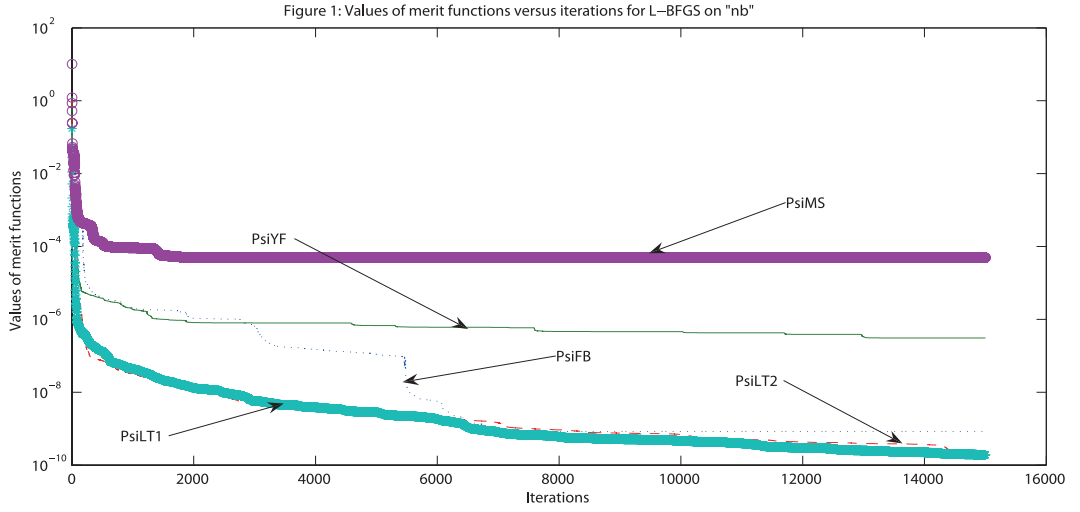*The notation $*$ means that the merit function method fails due to too small stepsize.*

yield the desired result for "nb" due to too small stepsize, and $\widehat{\Psi}_{\rm LT1}$ and $\widehat{\Psi}_{\rm LT2}$ can not yield the desired result for "nb" within 10^5 iterations. Table 1 shows that the LT merit functions $\Psi_{\rm LT1}$ and $\Psi_{\rm LT2}$ have similar performance for these problems, and their variants $\widehat{\Psi}_{\rm LT1}$ and $\widehat{\Psi}_{\rm LT2}$ also have similar performance, but the LT merit functions $\Psi_{\rm YF}, \Psi_{\rm LT1}, \Psi_{\rm LT2}$ are superior to their variants $\widehat{\Psi}_{\rm YF}, \widehat{\Psi}_{\rm LT1}, \widehat{\Psi}_{\rm LT2}$. Figures 1 and 2 below also show this fact. In addition, Figures 1 and 2 indicate that among all merit functions, $\Psi_{\rm MS}, \Psi_{\rm YF}$ and $\widehat{\Psi}_{\rm YF}$ have the worst convergence rate, and $\Psi_{\rm LT1}, \Psi_{\rm LT2}$ have the best convergence rate, although Table 1 shows that $\widehat{\Psi}_{\rm YF}$ has better performance than $\widehat{\Psi}_{\rm LT1}$ and $\widehat{\Psi}_{\rm LT2}$. We observe from Table 1 and Figure 1 that $\Psi_{\rm FB}$ has comparable performance with $\Psi_{\rm LT1}$ and $\Psi_{\rm LT2}$, and moreover has much better performance than $\Psi_{\rm MS}$ and $\Psi_{\rm YF}$ for these test problems.

For the LT merit functions $\Psi_{\rm LT}$ and its variant $\widehat{\Psi}_{\rm LT}$, we also tested their performance with ψ_0 chosen as $\log(1 + \max(0, t)^2)$, which has slower growth than the quadratic function $\max(0, t)^2/2$. We found that such a choice of ψ_0 does not improve numerical performance of $\Psi_{\rm YF}$ and the variant of LT merit functions, and it gives some improvements for $\Psi_{\rm LT1}$ and $\Psi_{\rm LT2}$ (see the result on the second line for $\Psi_{\rm LT1}$ and $\Psi_{\rm LT2}$ in Table 1).

The second group of problems consists of the nonlinear convex SOCP (1.5) generated randomly with $f(x) = \frac{1}{2}x^T Q x + c^T x$, where Q is an $n \times n$ symmetric positive semidefinite matrix, and $c \in \mathbb{R}^n$. We generate such instances with $m = 120, n = 2602$ and

$$\mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^1 \times \underbrace{\mathcal{K}^5 \times \cdots \times \mathcal{K}^5}_{200} \times \mathcal{K}^{800} \times \mathcal{K}^{800}.$$
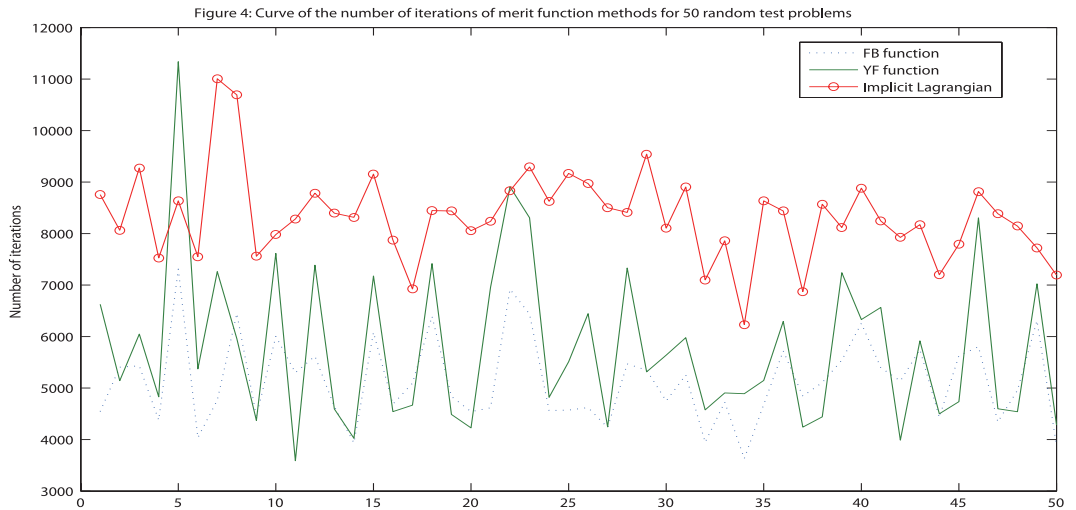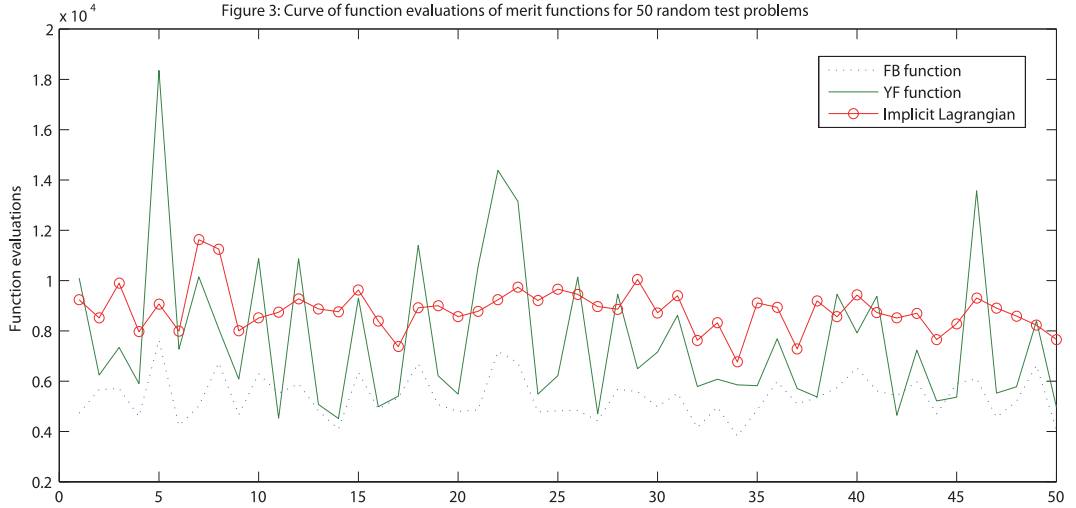
During the tests, the vector c is chosen randomly from the interval $[-1, 1]$, Q is given by

Figure 1: Values of merit functions versus iterations for L–BFGS on "nb"

Figure 2: Values of variant LT merit functions versus iterations for L–BFGS on "nb"

$Z^T Z$ with Z being an $n \times n$ matrix generated by MATLAB's "sprand" with %1 nonzero density, the matrix $A \in \mathbb{R}^{m \times n}$ is generated by MATLAB's "sprandn" with %1 nonzero density, and the vector $b \in \mathbb{R}^m$ is generated in the following way: choose $w \in \text{int}(\mathcal{K})$ and then let $b = Aw$, which guarantees the feasibility of the SOCPs generated. We observed that this class of problems has a bad scaling, and their optimal objective values attain 10^4.

In view of the fact that the variant of LT merit function has worse performance than the corresponding LT merit function, we did not apply them for this class of problems. We employed the above L-BFGS method with $\epsilon_1 = 10^{-8}$, $\epsilon_2 = 10^{-4}$ and $k_{\max} = 10^5$ to solve the unconstrained minimization reformulations based on the FB merit function, the implicit Lagrangian function and the LT merit functions, for 50 test problems generated randomly as above. We tested that among the three LT merit functions, Ψ_{LT1} and Ψ_{LT2} fail to this class of test problems; for example, for the first test problem, the function Ψ_{LT2} has the value 3.85×10^{-3} in the 2×10^5th iteration, and the function Ψ_{LT1} has the value 1.37×10^{-2} in the 2×10^5th iteration. Figures 3 and 4 below depict the curves of function evaluations and

the number of iterations, respectively, of the FB function method, the implicit Lagrangian function method and the YF function method. From the two figures, we see that the FB merit function method requires the lest function evaluations and iterations, and the YF merit function method requires less function evaluations and iterations for most test problems than the implicit Lagrangian function method.

Figure 3: Curve of function evaluations of merit functions for 50 random test problems

Figure 4: Curve of the number of iterations of merit function methods for 50 random test problems

To sum up, for the linear SOCPs in the DIMACS Implementation Challenge library, the FB merit function has comparable performance with the LT merit functions, and for the nonlinear SOCPs with bad scaling, it has much better performance. Among all merit functions, Ψ_{LT1} and Ψ_{LT2} have best convergence rate, and Ψ_{MS}, Ψ_{YF} and $\widehat{\Psi}_{\mathrm{YF}}$ have the worse convergence rate. The LT merit functions are always superior to their variants. In addition, the implicit Lagrangian with $\alpha = 50$ has worse performance than the YF merit function and the FB merit function for both groups of test problems.

5 Semismooth Newton Methods and Applications

This section concentrates on equation reformulation methods based on system (1.9) with ϕ chosen as ϕ_{NR} and ϕ_{FB}, respectively. Since the corresponding Φ is not differentiable, the nonsmooth Newton method [49, 55, 56], i.e. the following iterative method

$$z^{k+1} := z^k - (W^k)^{-1}\Phi(z^k), \quad W^k \in \partial_B \Phi(z^k), \quad k = 0, 1, 2, \dots, \tag{5.1}$$

can be applied for (1.9) to get a solution of (1.1), where $\partial_B \Phi(z^k)$ denotes the B-subdifferential of Φ at z^k whose existence is guaranteed by the global Lipschitz continuity of ϕ_{NR} and ϕ_{FB} (see [9, 58]). From the results of [55, 56], in order to establish fast local convergence of the method (5.1), on one hand, Φ is required to be sufficiently 'smooth' such as (strongly) semismooth; and on the other hand, it satisfies a local nonsingularity condition. In view of the strong semismoothness of ϕ_{NR} and ϕ_{FB}, we only need to check the nonsingularity of the B-subdifferential of Φ at an arbitrary solution of (1.1).

5.1 Nonsingularity Conditions

The following two lemmas review the B-subdifferentials of ϕ_{NR} and ϕ_{FB} at a general point, respectively, where Lemma 5.1 is a direct consequence of [53, Lemma 14], and Lemma 5.2 is implied by Prop. 4.2–4.3.

Lemma 5.1. *For any given* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, *with* $z = x - y$ *it holds that*

$$\partial_B \phi_{\mathrm{NR}}(x, y) = \left\{ \begin{bmatrix} I - V & V \end{bmatrix} \mid V \in \partial_B(z)_+ \right\},$$

where the B-subdifferential $\partial_B(z)_+$ *of the projector operator is characterized as follows:*

(a) *If* $z_1 > \|z_2\|$, *then* $\partial_B(z)_+ = \{I\}$; *if* $z_1 < -\|z_2\|$, *then* $\partial_B(z)_+ = \{0\}$.

(c) *If* $|z_1| < \|z_2\|$, *then* $\partial_B(z)_+ = \left\{ \dfrac{1}{2} \begin{pmatrix} 1 & \bar{z}_2^T \\ \bar{z}_2 & I + \dfrac{z_1}{\|z_2\|}(I - \bar{z}_2 \bar{z}_2^T) \end{pmatrix} \right\}$ *with* $\bar{z}_2 \equiv \dfrac{z_2}{\|z_2\|}$.

(d) *If* $z_1 = \|z_2\| \neq 0$, *then* $\partial_B(z)_+ = \left\{ I, \dfrac{1}{2} \begin{pmatrix} 1 & \bar{z}_2^T \\ \bar{z}_2 & 2I - \bar{z}_2 \bar{z}_2^T \end{pmatrix} \right\}$ *with* $\bar{z}_2 \equiv \dfrac{z_2}{\|z_2\|}$.

(e) *If* $-z_1 = \|z_2\| \neq 0$, *then* $\partial_B(z)_+ = \left\{ 0, \dfrac{1}{2} \begin{pmatrix} 1 & \bar{z}_2^T \\ \bar{z}_2 & \bar{z}_2 \bar{z}_2^T \end{pmatrix} \right\}$ *with* $\bar{z}_2 \equiv \dfrac{z_2}{\|z_2\|}$.

(f) *If* $z = 0$, *then*

$$\partial_B(z)_+ = \left\{ 0, I \right\} \bigcup \left\{ \frac{1}{2} \begin{pmatrix} 1 & \bar{z}_2^T \\ \bar{z}_2 & (w_0 + 1)I - w_0 \bar{z}_2 \bar{z}_2^T \end{pmatrix} \;\middle|\; |w_0| \leq 1 \text{ and } \|\bar{z}_2\| = 1 \right\}.$$

Particularly, V *and* $I - V$ *are positive semidefinite symmetric matrices.*

Lemma 5.2. *For any given* $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, *the following statements hold.*

(a) *If* $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, *then* ϕ_{FB} *is continuously differentiable at* (x, y) *with*

$$\partial_B \phi_{\mathrm{FB}}(x, y) = \left\{ \begin{bmatrix} I - L_{(x^2+y^2)^{1/2}}^{-1} L_x & I - L_{(x^2+y^2)^{1/2}}^{-1} L_y \end{bmatrix} \right\}.$$

(b) *If $x^2 + y^2 \in \mathrm{bd}^+(\mathcal{K}^n)$, then*

$$\partial_B \phi_{\mathrm{FB}}(x,y) = \left\{ \left[I - X - \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} u^T \quad I - Y - \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix} v^T \right] \right|$$

$$u = (u_1, u_2) = \left(\frac{\zeta_2^T y_2 + \zeta_1 x_1}{\sqrt{x_1^2 + y_1^2}}, \ \frac{-y_1 \zeta_2 - \zeta_1 x_2}{\sqrt{x_1^2 + y_1^2}} \right),$$

$$v = (v_1, v_2) = \left(\frac{-\zeta_2^T x_2 + \zeta_1 y_1}{\sqrt{x_1^2 + y_1^2}}, \ \frac{x_1 \zeta_2 - \zeta_1 y_2}{\sqrt{x_1^2 + y_1^2}} \right)$$

$$\text{for some } \zeta = (\zeta_1, \zeta_2) \text{ satisfying } \zeta_1^2 + \|\zeta_2\|^2 = 1 \Big\},$$

where $\bar{w}_2 = w_2/\|w_2\|$ with $w_2 = 2(x_1 x_2 + y_1 y_2)$, and

$$X = \frac{1}{2\sqrt{x_1^2 + y_1^2}} \begin{pmatrix} x_1 & x_2^T \\ x_2 & 2x_1 I - \bar{w}_2 x_2^T \end{pmatrix},$$

$$Y = \frac{1}{2\sqrt{x_1^2 + y_1^2}} \begin{pmatrix} y_1 & y_2^T \\ y_2 & 2y_1 I - \bar{w}_2 y_2^T \end{pmatrix}.$$

(c) *If $(x,y) = (0,0)$, then $\partial_B \phi_{\mathrm{FB}}(x,y) \subseteq \{ [I - L_g \ \ I - L_h] \mid g^2 + h^2 = e \} \cup S$, where*

$$S \equiv \left\{ \left[I - \frac{1}{2} p u^T - \frac{1}{2} q \xi^T - H L_s \quad I - \frac{1}{2} p v^T - \frac{1}{2} q \eta^T - H L_\omega \right] \right|$$

$$p = \begin{pmatrix} 1 \\ -\bar{w}_2 \end{pmatrix}, \ q = \begin{pmatrix} 1 \\ \bar{w}_2 \end{pmatrix}, \ H = \begin{pmatrix} 0 & 0 \\ 0 & I - \bar{w}_2 \bar{w}_2^T \end{pmatrix} \text{ for } \|\bar{w}_2\| = 1;$$

$$u, v \in \mathbb{R}^n \text{ satisfy } p^T u = 2u_1, q^T u = 0, p^T v = 0, q^T v = 2v_1 \text{ and}$$

$$|u_1| \le \|u_2\| \le 1, \ |v_1| \le \|v_2\| \le 1,$$

$$(u_1 - v_1) \le \|u_2 - v_2\|, \quad (u_1 + v_1) \le \|u_2 + v_2\|, \tag{5.2}$$

$$(u_1 - v_1)^2 + \|u_2 + v_2\|^2 \le 2, \quad (u_1 + v_1)^2 + \|u_2 - v_2\|^2 \le 2;$$

$$\xi, \eta \in \mathbb{R}^n \text{ satisfy } p^T u = 0, q^T u = 2u_1, p^T v = 2v_1, q^T v = 0 \text{ and } (5.2);$$

$$s = \sigma u + (1-\sigma)\xi, \ \omega = \sigma v + (1-\sigma)\eta \text{ for } \sigma \in [0, 1/2] \Big\}.$$

Comparing Lemma 5.2 with Lemma 5.1, we observe that the B-subdifferential of ϕ_{FB} at a general point has a more complicated structure than that of ϕ_{NR}, and moreover, the elements in $\partial_B \phi_{\mathrm{FB}}$ do not possess the desired properties of the elements in $\partial_B \phi_{\mathrm{NR}}$. But, as will be shown in the sequel, the nonsingularity of B-subdifferential of Φ_{FB} at a solution of (1.1) can be established under the same conditions as used for the B-subdifferential of Φ_{NR}.

Let (x^*, y^*, ζ^*) be a solution of (1.1). Define the index sets associated with x^* and y^*:

$$\begin{aligned}
J_{I0} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* \in \mathrm{int}(\mathcal{K}^{n_j}), \ y_j^* = 0 \right\}, \\
J_{BB} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* \in \mathrm{bd}^+(\mathcal{K}^{n_j}), \ y_j^* \in \mathrm{bd}^+(\mathcal{K}^{n_j}) \right\}, \\
J_{0I} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* = 0, \ y_j^* \in \mathrm{int}(\mathcal{K}^{n_j}) \right\}, \\
J_{B0} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* \in \mathrm{bd}^+(\mathcal{K}^{n_j}), \ y_j^* = 0 \right\}, \\
J_{0B} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* = 0, \ y_j^* \in \mathrm{bd}^+(\mathcal{K}^{n_j}) \right\}, \\
J_{00} &:= \left\{ j \in \{1, \dots, m\} \mid x_j^* = 0, \ y_j^* = 0 \right\}.
\end{aligned} \tag{5.3}$$

From [1], these index sets form a partition of $\{1, \ldots, m\}$. We next review two important properties for the elements of B-subdifferentials of ϕ_{NR} and ϕ_{FB} at (x^*, y^*).

Lemma 5.3. *Let (x^*, y^*, ζ^*) be an arbitrary solution of (1.1). Let $[U_j \; V_j] \in \partial_B \phi(x_j^*, y_j^*)$ for $j = 1, \ldots, m$ with $\phi = \phi_{\mathrm{NR}}$ or ϕ_{FB}. Then, for any $(\triangle u)_j, (\triangle v)_j \in \mathbb{R}^{n_j}$,*

$$U_j(\triangle u)_j + V_j(\triangle v)_j = 0 \implies \langle (\triangle u)_j, (\triangle v)_j \rangle \le 0 \quad \text{for } j = 1, 2, \ldots, m \qquad (5.4)$$

and

$$U_j(\triangle u)_j + V_j(\triangle v)_j = 0 \implies \begin{cases} (\triangle v)_j = 0 & \text{if } j \in J_{I0}, \\ (\triangle u)_j = 0 & \text{if } j \in J_{0I}, \\ \langle (\triangle u)_j, y_j^* \rangle = 0, \; \langle (\triangle v)_j, x_j^* \rangle = 0 & \text{if } j \in J_{BB}, \\ (\triangle u)_j = \mathbb{R}(y_{j1}^*, -y_{j2}^*) & \text{if } j \in J_{0B}, \\ (\triangle v)_j = \mathbb{R}(x_{j1}^*, -x_{j2}^*) & \text{if } j \in J_{B0}. \end{cases} \qquad (5.5)$$

Particularly, for all $j \in J_{BB}$, the following implication also holds:

$$U_j(\triangle u)_j + V_j(\triangle v)_j = 0 \implies \langle (\triangle u)_j, (\triangle v)_j \rangle = \frac{y_{j1}^*}{x_{j1}^*} \left((\triangle u)_{j1}^2 - \|(\triangle u)_{j2}\|^2 \right). \qquad (5.6)$$

Proof. For $\phi = \phi_{\mathrm{NR}}$, the implication (5.4) is direct by [40, Prop. 1], and the implications (5.5) and (5.6) are implied by [64, Prop. 3.1]. For $\phi = \phi_{\mathrm{FB}}$, the implication (5.4) is given by [48, Prop. 4.2], and the implications (5.5) and (5.6) are given by [48, Prop. 4.1]. $\qquad \square$

Lemma 5.4. *Let (x^*, y^*, ζ^*) be an arbitrary solution of (1.1). Let $[U_j \; V_j] \in \partial_B \phi(x_j^*, y_j^*)$ for $j = 1, \ldots, m$ with $\phi = \phi_{\mathrm{NR}}$ or ϕ_{FB}. Then, for any $(\triangle u)_j, (\triangle v)_j \in \mathbb{R}^{n_j}$,*

$$U \triangle u + V \triangle v = 0 \implies \langle \triangle u, \triangle v \rangle \le \sum_{j=1}^m \Upsilon_{x_j^*}(y_j^*, (\triangle u)_j)$$

where for any given $\varpi \in \mathbb{R}^n$, $\Upsilon_{\varpi_j} : \mathbb{R}^{n_j} \times \mathbb{R}^{n_j} \to \mathbb{R}$ is the linear-quadratic function

$$\Upsilon_{\varpi_j}(\xi_j, \eta_j) := \begin{cases} \dfrac{\xi_{j1}}{\varpi_{j1}} \eta_j^T \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} \eta_j & \text{if } j \in J_{BB} \cup J_{B0}, \\ 0 & \text{otherwise.} \end{cases} \qquad (5.7)$$

Proof. The proof is seen in [64, Prop. 3.1] for $\phi = \phi_{\mathrm{NR}}$, and [48, Prop. 4.3] for $\phi = \phi_{\mathrm{FB}}$. $\quad \square$

Now we are ready to establish the nonsingularity conditions for the B-subdifferential $\partial_B \Phi(z^*)$ with $z^* = (x^*, y^*, \zeta^*)$ and Φ given by the generalized SOCCP (1.3). To the end, we also need to recall the tangent cone $\mathcal{T}_{\mathcal{K}}(y^*)$ of $\mathcal{K}$ at x^* and the critical cone $\mathcal{C}(y^* - x^*; \mathcal{K})$ of $\mathcal{K}$ at $y^* - x^*$. By [2, Lemma 25], the tangent cone $\mathcal{T}_{\mathcal{K}}(y^*)$ takes the form of

$$\mathcal{T}_{\mathcal{K}}(y^*) := \Big\{ d \in \mathbb{R}^n \mid d_j \in \mathcal{K}^{n_j} \text{ for } j \in J_{I0} \cup J_{B0} \cup J_{00};$$
$$d_j^T(y_{j1}^*, -y_{j2}^*) \ge 0 \text{ for } j \in J_{BB} \cup J_{0B} \Big\}. \qquad (5.8)$$

Noting that $\mathcal{C}(y^* - x^*; \mathcal{K}) = \mathcal{T}_{\mathcal{K}}(y^*) \cap (x^*)^\perp$, we have from [2, Corollary 26] that

$$\mathcal{C}(y^* - x^*; \mathcal{K}) = \Big\{ d \in \mathbb{R}^n \,\Big|\, d_j = 0 \text{ for } j \in J_{I0}, \; d_j = \mathbb{R}_+(x_{j1}^*, -x_{j2}^*) \text{ for } j \in J_{B0},$$
$$\langle d_j, x_j^* \rangle = 0 \text{ for } j \in J_{BB}, \; d_j \in \mathcal{K}^{n_j} \text{ for } j \in J_{00},$$
$$d_j^T(y_{j1}^*, -y_{j2}^*) \ge 0 \text{ for } j \in J_{0B} \Big\}.$$

Consequently, the affine hull of $\mathcal{C}(y^*-x^*;\mathcal{K})$, denoted by aff $(\mathcal{C}(y^*-x^*;\mathcal{K}))$, has the form

$$\text{aff}\,(\mathcal{C}(y^*-x^*;\mathcal{K})) = \Big\{ d \in \mathbb{R}^n \,\Big|\, d_j = 0 \text{ for } j \in J_{I0},\ d_j = \mathbb{R}(x_{j1}^*, -x_{j2}^*) \text{ for } j \in J_{B0}, \quad (5.9)$$

$$\langle d_j, x_j^* \rangle = 0 \text{ for } j \in J_{BB} \Big\}.$$

Theorem 5.5. *Let ζ^* be a solution of the SOCCP (1.3) with $x^* = F(\zeta^*)$ and $y^* = G(\zeta^*)$. Suppose that $G'(\zeta^*)$ is invertible with $M = F'(\zeta^*)(G'(\zeta^*))^{-1}$, and*

$$\langle d, H(\zeta^*)d \rangle > 0 \quad \forall d \in \text{aff}\,(\mathcal{C}(y^*-x^*;\mathcal{K})) \setminus \{0\} \tag{5.10}$$

where $H(\zeta^*) = \sum_{j=1}^m H_j(\zeta^*) \in \mathbb{R}^{n\times n}$ with $H_j(\zeta^*) \in \mathbb{R}^{n\times n}$ defined by

$$H_j(\zeta^*) := \begin{cases} -\dfrac{y_{j1}^*}{x_{j1}^*} M_j^T \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} M_j & \text{if } j \in J_{BB} \cup J_{B0} \\ 0 & \text{otherwise.} \end{cases}$$

Then, any element in $\partial_B \Phi(x^*, y^*, \zeta^*)$ is nonsingular.

Proof. Let W be an arbitrary element in $\partial_B \Phi(x^*, y^*, \zeta^*)$. To prove that W is nonsingular, we let $(\triangle x, \triangle y, \triangle \zeta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ be such that $W(\triangle x, \triangle y, \triangle \zeta) = 0$. Since now

$$E(x, y, \zeta) = \begin{pmatrix} F(\zeta) - x \\ G(\zeta) - y \end{pmatrix},$$

by the definition of Φ and [5, Lemma 1], there exists a $[U\ \ V] \in \partial_B \phi(x^*, y^*)$ such that

$$W(\triangle x, \triangle y, \triangle \zeta) = \begin{bmatrix} F'(\zeta)\triangle\zeta - \triangle x \\ G'(\zeta)\triangle\zeta - \triangle y \\ U\triangle x + V\triangle y \end{bmatrix} = 0 \tag{5.11}$$

where $U = \text{diag}(U_1, \ldots, U_m)$ and $V = \text{diag}(V_1, \ldots, V_m)$ with $[U_j\ \ V_j] \in \partial_B \phi(x_j^*, y_j^*)$ for $j = 1, \ldots, m$. Note that the third equation of (5.11) is equivalent to

$$U_j(\triangle x)_j + V_j(\triangle y)_j = 0, \quad j = 1, 2, \ldots, m.$$

From the implication (5.5) in Lemma 5.3, it then follows that

$$(\triangle y)_j = 0 \text{ for } j \in J_{I0}; \ (\triangle y)_j = \mathbb{R}(x_{j1}^*, -x_{j2}^*) \text{ for } j \in J_{B0}; \ \langle (\triangle y)_j, x_j^* \rangle = 0 \text{ for } j \in J_{BB}.$$

Comparing with the definition of aff $(\mathcal{C}(y^*-x^*;\mathcal{K}))$ in (5.9), we have

$$\triangle y \in \text{aff}\,(\mathcal{C}(y^*-x^*;\mathcal{K})). \tag{5.12}$$

This, together with the third equation of (5.11), (5.4) and Lemma 5.4, implies that

$$
\begin{aligned}
0 \geq \langle \triangle y, \triangle x \rangle &\geq -\sum_{j=1}^m \Upsilon_{x_j^*}(y_j^*, (\triangle x)_j) \\
&= -\sum_{j \in J_{BB} \cup J_{B0}} \frac{y_{j1}^*}{x_{j1}^*}(\triangle y)^T M_j^T \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} M_j \triangle y \\
&= \langle \triangle y, H(\zeta^*) \triangle y \rangle
\end{aligned}
$$

where the first equality is using $\triangle x = M\triangle y$. This by (5.10) implies $\triangle y = 0$, and $\triangle \zeta = 0$ and $\triangle x = 0$ then follow successively from the second and first equations of (5.11). $\square$

For the SOCCP (1.4), the conditions of Theorem 5.5 are equivalent to requring that

$$- \left(F_j'(\zeta^*)d \right)^T \begin{pmatrix} 1 & 0 \\ 0 & -I \end{pmatrix} \left(F_j'(\zeta^*)d \right) > 0 \quad \forall d \in \text{aff}\left(\mathcal{C}(y^* - x^*; \mathcal{K}) \right) \setminus \{0\}.$$

When (1.1) becomes the KKT optimality conditions of the nonlinear SOCP (1.5), the non-singularity of $\partial_B \Phi(z^*)$ can be established under the strong second-order sufficient condition [2] and constraint nondegeneracy, which for the nonlinear SOCP (1.5) are as follows.

Definition 5.6. Let x^* be a stationary point of the SOCP (1.5) such that the multiplier set $\Lambda(x^*) = \{(\zeta^*, y^*)\}$. We say the strong second-order sufficient condition holds at x^* if

$$\langle d, \nabla^2 f(x)d \rangle > 0 \quad \forall d \in \text{aff}(\mathcal{C}(x^* - y^*; \mathcal{K})) \setminus \{0\}. \tag{5.13}$$

where $\text{aff}(\mathcal{C}(x^* - y^*; \mathcal{K}))$ is the affine hull of the critical cone $\mathcal{C}(x^* - y^*; \mathcal{K})$ given by

$$\text{aff}(\mathcal{C}(x^* - y^*; \mathcal{K})) = \Big\{ d \in \mathbb{R}^n \ \Big| \ d_j = 0 \text{ for } j \in J_{0I}, \ \langle d_j, y_j^* \rangle = 0 \text{ for } j \in J_{BB},$$
$$d_j \in \mathbb{R}(y_{j1}^*, -y_{j2}^*) \text{ for } j \in J_{0B}; \ Ad = 0 \Big\}.$$

Definition 5.7. We say that a feasible point $\hat{x}$ to (1.5) is constraint nondegenerate if

$$\begin{pmatrix} A \\ I \end{pmatrix} \mathbb{R}^n + \begin{pmatrix} \{0\} \\ \text{lin}\left(\mathcal{T}_\mathcal{K}(\hat{x})\right) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^m \\ \mathbb{R}^n \end{pmatrix}. \tag{5.14}$$

Theorem 5.8. *Let $z^* = (x^*, y^*, \zeta^*)$ be a KKT point of the nonlinear SOCP (1.5). Suppose that the strong second-order sufficient condition (5.13) holds at x^* and x^* is constraint nondegenerate, then any element in $\partial_B \Phi(z^*)$ is nonsingular.*

Proof. The proof is found in [64, Theorem 3.1] and [48, Theorem 4.1]. □

By [64] the strong second-order sufficient condition and constraint nondegeneracy are aslo sufficient for the nonsingularity of $\partial_B \Phi_{\text{NR}}(z^*)$, but now it is not clear whether they are sufficient for the nonsingularity of $\partial_B \Phi_{\text{FB}}(z^*)$. Observe that Kanzow et al. [32] also present a nonsingularity condition without strict complementartiy for $\partial_B \Phi_{\text{NR}}(z^*)$ by using the algebraic technique. However, their condition is more complicated than the strong second-order sufficient condition and constraint nondegeneracy.

5.2 FB Nonsmooth Newton Method

The last subsection discusses the fast local convergence of the NR and FB nonsmooth Newton methods. Note that the NR nonsmooth Newton method is hard to be globalized due to the nondifferentiability of natural residual merit function. On the contrast, the FB nonsmooth Newton method is easily globalized by the smoothness of FB merit function. In this subsection, we present a global convergent FB semismooth method for system (1.1).

Let $\Phi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$ be the operator defined by (1.9) with $\phi = \phi_{\text{FB}}$, and $\Psi_{\text{FB}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}_+$ be the natural merit function for the system $\Phi_{\text{FB}}(z) = 0$, i.e., $\Psi_{\text{FB}}(z) = \|\Phi_{\text{FB}}(z)\|^2$ for any $w \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$. The following algorithm is similar to the global convergent semismooth Newton methods developed in [36] for NCPs.

Algorithm 5.1 (**FB semismooth Newton method**)

Step 0. Choose constants $\rho > 0, p > 2, \delta \in (0,1)$ and $\sigma \in (0,1/2)$. Choose a starting point $z^0 = (x^0, y^0, \zeta^0) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$. Set $k := 0$.

Step 1. If $\|\Phi_{\mathrm{FB}}(z^k)\| = 0$, then stop.

Step 2. Select an element $W^k \in \partial_B \Phi_{\mathrm{FB}}(z^k)$, and find the solution d^k of the system

$$W^k d = -\Phi_{\mathrm{FB}}(z^k). \tag{5.15}$$

If the solution d^k does not satisfy the condition that

$$\nabla \Psi_{\mathrm{FB}}(z^k)^T d^k \leq -\rho \|d^k\|^p, \tag{5.16}$$

then set $d^k = -\nabla \Psi_{\mathrm{FB}}(z^k)$.

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$\Psi_{\mathrm{FB}}(z^k + \delta^l d^k) \leq \Psi_{\mathrm{FB}}(z^k) + \sigma \delta^l \nabla \Psi_{\mathrm{FB}}(z^k).$$

Step 4. Define $z^{k+1} := z^k + \delta^{l_k} d^k$. Let $k := k + 1$, and then go to Step 1.

As remarked in [36], Algorithm 5.1 is virtually indistinguishable from a global Newton algorithm for the solution of an differentiable system of equations, except the selection of W^k in Step 2. From the definition of Φ_{FB}, any element W^k in $\partial_B \Phi_{\mathrm{FB}}(z^k)$ has the form

$$W^k = \begin{bmatrix} E_x'(x^k, y^k, \zeta^k) & E_y'(x^k, y^k, \zeta^k) & E_\zeta'(x^k, y^k, \zeta^k) \\ U^k & V^k & 0 \end{bmatrix},$$

where $U^k = \mathrm{diag}(U_1^k, \ldots, U_m^k)$ and $V^k = \mathrm{diag}(V_1^k, \ldots, V_m^k)$ with $[U_j^k \ \ V_j^k] \in \partial_B \phi_{\mathrm{FB}}(x_j^k, y_j^k)$. Hence, the selection of W^k from $\partial_B \Phi_{\mathrm{FB}}(z^k)$ reduces to the selection of $[U_j^k \ \ V_j^k]$ from $\partial_B \phi_{\mathrm{FB}}(x_j^k, y_j^k)$. We next give a simple way to evaluate an element $[U_j^k \ \ V_j^k] \in \partial \phi_{\mathrm{FB}}(x_j^k, y_j^k)$, which is verified to be effective by the subsequent numerical experiments.

Procedure to evaluate an element $[U_j \ \ V_j] \in \partial \phi_{\mathrm{FB}}(x_j, y_j)$

Step 1: If $x_j, y_j \in \mathbb{R}$, then go to Step 2; otherwise go to Step 3.

Step 2: If $x_j = y_j = 0$, then set $U_j = V_j = 1/2$. Otherwise, let

$$U_j = 1 - x_j / \sqrt{x_j^2 + y_j^2} \ \ \text{and} \ \ V_j = 1 - y_j / \sqrt{x_j^2 + y_j^2}.$$

Step 3: Let $w_j = x_j^2 + y_j^2$. If $w_j \in \mathrm{int}(\mathcal{K}^{n_j})$, set

$$U_j = I - L_{\sqrt{w_j}}^{-1} L_{x_j} \ \ \text{and} \ \ V_i = I - L_{\sqrt{w_j}}^{-1} L_{y_j}.$$

If $w_j \in \mathrm{bd}^+(\mathcal{K}^{n_j})$, set $\bar{w}_{j2} = \frac{w_{j2}}{\|w_{j2}\|}$, $u_j = v_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\bar{w}_{j2} \end{pmatrix}$, and

$$U_j = I - \frac{1}{2\sqrt{x_{j1}^2 + y_{j1}^2}} \begin{bmatrix} x_{j1} & x_{j2}^T \\ x_{j2} & 2x_{j1}I - \bar{w}_{j2}x_{j2}^T \end{bmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_{j2} \end{pmatrix} u_j^T,$$

$$V_j = I - \frac{1}{2\sqrt{x_{j1}^2 + y_{j1}^2}} \begin{bmatrix} y_{j1} & y_{j2}^T \\ y_{j2} & 2y_{j1}I - \bar{w}_{j2}y_{j2}^T \end{bmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -\bar{w}_{j2} \end{pmatrix} v_j^T.$$

If $(x_j, y_j) = (0,0)$, let $\tilde{g}_j = (1,1,\ldots,1)^T, \tilde{h}_j = (1,-1,\ldots,-1)^T \in \mathbb{R}^{n_j}$, and set

$$g_j = \frac{\tilde{g}_j}{\sqrt{\|\tilde{g}_j\|^2 + \|\tilde{g}_j\|^2}} \quad \text{and} \quad g_j = \frac{\tilde{h}_j}{\sqrt{\|\tilde{g}_j\|^2 + \|\tilde{g}_j\|^2}};$$

and let $\bar{w}_{j2} = \frac{w_{j2}}{\|w_{j2}\|}$ with $w_{j2} = (1,1,\ldots,1) \in \mathbb{R}^{n_j - 1}$ and calculate

$$\xi_j = \eta_j = \begin{pmatrix} 1 \\ \bar{w}_{j2} \end{pmatrix}, \ u_j = v_j = \begin{pmatrix} 1 \\ -\bar{w}_{j2} \end{pmatrix}, \ s_j = \omega_j = u_j,$$

$$\tilde{U}_j = I - \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_{j2} \end{pmatrix}\xi_j^T - \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_{j2} \end{pmatrix}u_j^T - \begin{bmatrix} 0 & 0 \\ 0 & (I - \bar{w}_{j2}\bar{w}_{j2}^T) \end{bmatrix}L_{s_j},$$

$$\tilde{V}_j = I - \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_{j2} \end{pmatrix}\eta_j^T - \frac{1}{2}\begin{pmatrix} 1 \\ \bar{w}_{j2} \end{pmatrix}v_j^T - \begin{bmatrix} 0 & 0 \\ 0 & (I - \bar{w}_{j2}\bar{w}_{j2}^T) \end{bmatrix}L_{\omega_j}.$$

Finally, take $U_j = \frac{1}{2}(L_{g_j} + \tilde{U}_j)$ and $V_i = \frac{1}{2}(L_{h_j} + \tilde{V}_j)$.

For Algorithm 5.1, using the similar arguments as in [36] and the properties of FB merit functions in Sections 3–4, it is not hard to establish its global convergence. Now we take a closer look at the numerical performance of Algorithm 5.1. We implemented a nonmontone line search version of Algorithm 5.1, i.e., in Step 3 we compute the smallest nonnegative integer l_k satisfying

$$\Psi_{\text{FB}}(z^k + \delta^{l_k}d^k) \leq \mathcal{W}_k - \sigma\delta^{l_k}\Psi_{\text{FB}}(z^k)$$

where $\mathcal{W}_k = \max_{j=k-m_k,\cdots,k}\Psi_{\text{FB}}(z^j)$ and where, for a given nonnegative integer $\hat{m}$ and s,

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

Throughout the tests, we used $\hat{m} = 5, s = 5$ and the parameters below for Algorithm 5.1:

$$\rho = 10^{-8}, \quad p = 2.1, \quad \delta = 0.5, \quad \sigma = 0.5 \times 10^{-4}. \tag{5.17}$$

We terminated the algorithm once $\Psi_{\text{FB}}(z^k) \leq \epsilon$ and $k \leq 150$. The linear system (5.15) in Algorithm 5.1 is solved by "\" of Matlab, i.e., $d^k = -W^k\backslash\Phi_{\text{FB}}(z^k)$. All tests were done in Matlab 6.5 on a PC of Pentium 4 with 2.8GHz CPU and 512MB memory.

Table 2: **Numerical results of Algorithm 5.1 for linear SOCPs**

	(m,n)	**NF**	**Iter**	**Ngra**	$\Psi_{\text{FB}}(z^k)$	**Time**	**Optval**
nb	$(123, 2383)$	89	38	3	7.48e–9	128.4	-0.05070337
nb_L1	$(915, 3176)$	268	149	1	6.74e–9	489.1	-13.01227034
nb_L2	$(123, 4195)$	64	14	0	2.10e–9	1360.8	-1.62897295
nb_L2_bessel	$(123, 2641)$	24	15	0	3.60e–10	50.6	-0.10256981

We tested two groups of convex SOCP instances. The first one is composed of four standard linear SOCPs from the DIMACS Implementation Challenge library [52]. We solved them by Algorithm 5.1 with the error tolerance ϵ chosen as 1.0×10^{-8}. Numerical results are reported in Table 2 where **NF** and **Iter** have the same meaning as in Table 1, **Ngra** means the number of negative gradient steps adopted by Algorithm 5.1 for solving each problem, **Time** denotes the CPU time in seconds for solving each problem, and **Optval** denotes the optimal objective values of linear SOCPs. We note that Algorithm 5.1 yields solutions with desirable accuracy to four test problems within 150 iterations.

The second group of test instances is composed of nonlinear convex SOCPs in (1.5) generated randomly with $f(x)$ same as that of Subsection 4.1, but $\mathcal{K}$ has the structure

$$\mathcal{K} = \underbrace{\mathcal{K}^1 \times \cdots \times \mathcal{K}^1}_{50} \times \underbrace{\mathcal{K}^5 \times \cdots \times \mathcal{K}^5}_{200} \times \mathcal{K}^{800} \times \mathcal{K}^{800},$$

and $m = 120, n = 2650$. The matrices $Q \in \mathbb{R}^{n \times n}$ and $A \in \mathbb{R}^{m \times n}$ and the vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are generated in the same way as in Subsection 4.1. We applied the above nonmonotone line search version of Algorithm 5.1 with the parameters in (5.17) and tolerance $\epsilon = 1.0 \times 10^{-9}$ for 50 test instances generated randomly. The computational results are as follows: Algorithm 5.1 solves all problems successfully with given accuracy, the average function evaluations and the average iterations needed are **11** and **10**, respectively, the average CPU time used by each problem is **857.3**(s), and all problems do not make use of the negative gradient steps. As pointed out in Subsection 4.1, this class of problems has a bad scaling, and the optimal value attains the order of 10^3, but numerical results show that the FB semismooth Newton method does not suffer from this.

6 Smoothing Newton Methods and Applications

This section focuses on equation reformulation methods based on smoothing functions of SOC complementarity functions ϕ_{NR} and ϕ_{FB}, that is, the smoothing Newton methods based on the augmented system (1.12) involving the Chen-Mangasarian (CM) smoothing functions of ϕ_{NR}, the squared smoothing function of ϕ_{NR}, and the smoothing function of ϕ_{FB}. In particular, we provide numerical comparisons for their behaviors in solving linear SOCPs from DIMACS and nonlinear convex SOCPs generated randomly.

6.1 Smoothing Newton Methods

Let $g : \mathbb{R} \to \mathbb{R}_+$ be an arbitrary continuously differentiable convex function satisfying

$$\lim_{t \to -\infty} g(t) = 0, \quad \lim_{t \to +\infty} (g(t) - t) = 0 \ \text{ and } \ g'(t) \in (0,1) \ \text{ for all } t \in \mathbb{R}. \tag{6.1}$$

The CM family of smoothing functions of ϕ_{NR} associated with $\mathcal{K}^n$ [25] is defined as

$$\varphi_{\mathrm{CM}}(x, y, \varepsilon) := x - \varepsilon g^{\mathrm{soc}}\left(\frac{x-y}{\varepsilon}\right) \quad \forall x, y \in \mathbb{R}^n, \ \varepsilon > 0 \tag{6.2}$$

which is a natural generalization of CM family of smoothing functions [8] for the NCPs. Two most popular smoothing functions among the CM family are the CHKS smoothing function [7, 29, 57] and the log-exponential smoothing function [8], respectively, with

$$g(t) := \frac{\sqrt{t^2 + 4} + t}{2} \ \text{ and } \ g(t) := \ln\left(\exp(t) + 1\right). \tag{6.3}$$

Another common smoothing function of ϕ_{NR} is the squared smoothing function given by

$$\varphi_{\text{SQ}}(x, y, \varepsilon) := \frac{1}{2} \left[(x + y) - \left((x - y)^2 + 4\varepsilon^2 e \right)^{1/2} \right] \quad \forall x, y \in \mathbb{R}^n, \ \varepsilon > 0. \tag{6.4}$$

This function was employed to develop a smoothing Newton method in [18], where numerical comparisons with the interior point method SDPT3 for linear SOCPs indicate that the former is very promising. The FB smoothing function associated with $\mathcal{K}^n$ is

$$\varphi_{\text{FB}}(x, y, \varepsilon) := (x + y) - (x^2 + y^2 + 2\varepsilon^2 e)^{1/2} \quad \forall x, y \in \mathbb{R}^n, \ \varepsilon > 0. \tag{6.5}$$

The following proposition shows that φ_{CM}, φ_{SQ} and φ_{FB} defined as above are the uniform smooth approximation of the corresponding SOC complementarity function, and characterizes the properties of Jacobians of these smoothing functions.

Proposition 6.1. *Let φ be one of $\varphi_{\text{CM}}, \varphi_{\text{SQ}}$ and φ_{FB} defined as above. Then,*

(a) *φ is positively homogeneous, i.e., $\varphi(tx, ty, t\varepsilon) = t\varphi(x, y, \varepsilon)$ for all $x, y \in \mathbb{R}^n$, $\varepsilon, t > 0$.*

(b) *For any $x, y \in \mathbb{R}^n$ and $\varepsilon_2 > \varepsilon_1 > 0$, there hold that*

$$\kappa(\varepsilon_2 - \varepsilon_1)e \ \succeq_{\mathcal{K}^n} \ \varphi(x, y, \varepsilon_2) - \varphi(x, y, \varepsilon_1) \ \succ_{\mathcal{K}^n} \ 0,$$
$$\kappa\varepsilon_1 e \ \succeq_{\mathcal{K}^n} \ \varphi(x, y, 0) - \varphi(x, y, \varepsilon_1) \ \succ_{\mathcal{K}^n} \ 0.$$

where $\varphi(x, y, 0) = \lim_{\varepsilon \downarrow 0} \varphi(x, y, \varepsilon)$, and $\kappa = g(0)$ if $\varphi = \varphi_{\text{CM}}$, and otherwise $\kappa = \sqrt{2}$.

(c) *φ is continuously differentiable everywhere in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++}$. Furthermore,*

$$\varphi'_{\text{CM}}(x, y, \varepsilon) = \left[I - \nabla g^{\text{soc}}(z) \ \ \nabla g^{\text{soc}}(z) \ \ \nabla g^{\text{soc}}(z)^T z - g^{\text{soc}}(z) \right] \ \text{with} \ z = \frac{x - y}{\varepsilon}$$

where $\nabla g^{\text{soc}}(z)$ has the same expression as given in Lemma 2.2, and

$$\varphi'_{\text{SQ}}(x, y, \varepsilon) = \frac{1}{2} \left[I - L_z^{-1} L_{x-y} \ \ I + L_z^{-1} L_{x-y} \ \ -4\varepsilon L_z^{-1} e \right] \ \text{with} \ z = ((x - y)^2 + 4\varepsilon^2 e)^{1/2};$$

$$\varphi'_{\text{FB}}(x, y, \varepsilon) = \left[I - L_z^{-1} L_x \ \ I - L_z^{-1} L_y \ \ -2\varepsilon L_z^{-1} e \right] \ \text{with} \ z = (x^2 + y^2 + 2\varepsilon^2 e)^{1/2}.$$

(d) *The partial Jacobians φ'_x and φ'_y of φ are nonsingular in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++}$.*

(e) *The matrices $(\varphi'_x)^{-1}\varphi'_y$ and $(\varphi'_y)^{-1}\varphi'_x$ are positive definite in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_{++}$.*

Proof. Part (a) is direct by the expression of φ. When $\varphi = \varphi_{\text{CM}}$ and φ_{FB}, part (b) is proved in [25, Prop. 5.1]. Using similar arguments, we can prove that part (b) holds for $\varphi = \varphi_{\text{SQ}}$. Part (c) is immediate by [9, Prop. 5]. When $\varphi = \varphi_{\text{CM}}$, the proof of part (d) can be found in [25, Prop. 6.1], and when $\varphi = \varphi_{\text{SQ}}$ and φ_{FB}, part (d) is direct by the expressions of φ'_x and φ'_y. When $\varphi = \varphi_{\text{SQ}}$ and φ_{FB}, the proof of part (e) can be found in Prop. 6.1 and 6.2 of [25], respectively; and when $\varphi = \varphi_{\text{SQ}}$, since $z^2 \succ_{\mathcal{K}^n} (x - y)^2$ and $z \succ_{\mathcal{K}^n} 0$, we have from Prop. 3.4 of [25] that $L_z^2 - L_{x-y}^2$ is positive definite, and part (e) follows by noting that $L_z^2 - L_{x-y}^2 = (L_z - L_{x-y})(L_z + L_{x-y}) + (L_z + L_{x-y})(L_z - L_{x-y})$. $\square$

With $\varphi = \varphi_{\text{CM}}, \varphi_{\text{SQ}}$ and φ_{FB} above, we define the function $\theta \colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ by

$$\theta(x, y, \varepsilon) := \begin{cases} \varphi(x, y, |\varepsilon|) & \text{if } \varepsilon \neq 0, \\ \phi(x, y) & \text{if } \varepsilon = 0. \end{cases} \tag{6.6}$$

The following proposition summarizes some favorable properties of the function θ.

Proposition 6.2. *Let θ be defined as in (6.6) with $\varphi = \varphi_{\mathrm{CM}}, \varphi_{\mathrm{SQ}}$ or φ_{FB}. Then,*

(a) *θ is continuously differentiable at any (x, y, ε) with $\varepsilon \neq 0$. In particular, in this case,*

$$\|\theta'(x, y, \varepsilon)\| \leq C,$$

where $C > 0$ is a constant independent on x, y and ε.

(b) *θ is globally Lipschitz continuous and directionally differentiable everywhere.*

(c) *θ is a strongly semismooth function if $\varphi = \varphi_{\mathrm{CM}}$ with g given by (6.3), φ_{SQ} or φ_{FB}.*

Proof. (a) The first part is a direct consequence of Prop. 6.1(c). For the second part, when $\varphi = \varphi_{\mathrm{CM}}$, by the properties of g in (6.1) and the expression of $\nabla g^{\mathrm{soc}}(z)$ in Lemma 2.2, it is easy to verify that the boundedness of θ'; when $\varphi = \varphi_{\mathrm{SQ}}$ and φ_{FB}, using the same arguments as in those of [10, Lemma 4] can show that θ' is bounded.

(b) Using part (a) and Prop. 6.1(b) and noting that ϕ_{NR} and ϕ_{FB} are globally Lipschitz continuous and directionally differentiable everywhere, we readily get the result.

(c) When $\varphi = \varphi_{\mathrm{CM}}$ with g given by (6.3), letting $h : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$h(t, \varepsilon) := \frac{\sqrt{t^2 + 4\varepsilon^2} + t}{2} \quad \text{and} \quad h(t, \varepsilon) := \varepsilon \ln\left(1 + \exp(-t/\varepsilon)\right) \quad \forall t, \varepsilon \in \mathbb{R},$$

it is not hard to see that for any $x, y \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}$,

$$\theta(x, y, \varepsilon) = x - \left[h(\lambda_1(z), \varepsilon)u_z^{(1)} + h(\lambda_2(z), \varepsilon)u_z^{(2)}\right]$$

where $\lambda_1(z)u_z^{(1)} + \lambda_2(z)u_z^{(2)}$ is the spectral decomposition of $z = x - y$. From Prop. 1 and Prop. 2 of [54], the above h are strongly semismooth functions. Hence, θ is strongly semismooth everywhere in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ by [9, Prop. 7]. When $\varphi = \varphi_{\mathrm{SQ}}$ and φ_{FB}, the result is implied by [18, Theorem 4.2] and [58, Theorem 3.2], respectively. $\square$

Unless otherwise stated, $\theta(x, y, \varepsilon)$ in the rest of this section is the function associated with $\mathcal{K}$, i.e. $\theta(x, y, \varepsilon) = (\theta(x_1, y_1, \varepsilon), \ldots, \theta(x_m, y_m, \varepsilon))$ with $\theta(x_i, y_i, \varepsilon)$ defined as in (6.6). Let $\Theta : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n$ be the operator defined by (1.12) with such θ. The following proposition shows that Θ is continuously differentiable at any $\omega \in \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$, and has nonsingular Jacobians under some mild assumptions.

Proposition 6.3. *Let Θ be defined by (1.12) with θ given as in (6.6). Then,*

(a) *the operator Θ is continuously differentiable at any $\omega \in \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$ and*

$$\Theta'(\omega) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & E_x'(x, y, \zeta) & E_y'(x, y, \zeta) & E_\zeta'(x, y, \zeta) \\ \theta_\varepsilon'(x, y, \varepsilon) & D_x(x, y, \varepsilon) & D_y(x, y, \varepsilon) & 0 \end{bmatrix}$$

where

$$\begin{aligned} D_x(x, y, \varepsilon) &:= \mathrm{diag}\left(\theta_{x_1}'(x_1, y_1, \varepsilon), \cdots, \theta_{x_m}'(x_m, y_m, \varepsilon)\right), \\ D_y(x, y, \varepsilon) &:= \mathrm{diag}\left(\theta_{y_1}'(x_1, y_1, \varepsilon), \cdots, \theta_{y_m}'(x_m, y_m, \varepsilon)\right). \end{aligned} \tag{6.7}$$

(b) *$\Theta'(\omega)$ is nonsingular provided that $\mathrm{rank}\, E_\zeta'(x, y, \zeta) = l$ and for any $u \neq 0, v \neq 0$,*

$$E'(x, y, \zeta)(u, v, s) = 0 \Rightarrow \exists \nu \in \{1, \ldots, m\} \text{ s.t. } u_\nu \neq 0 \text{ and } \langle u_\nu, v_\nu \rangle \geq 0. \tag{6.8}$$

Proof. Part (a) is direct by Prop. 6.2(a) and the definition of Θ. We next prove part (b). By the expression of $\Theta'(\omega)$, it suffices to prove that the following system

$$E'_x(x,y,\zeta)u + E'_y(x,y,\zeta)v + E'_\zeta(x,y,\zeta)s = 0$$
$$D_x(x,y,\zeta)u + D_y(x,y,\zeta)v = 0 \tag{6.9}$$

has only zero solutions. If one of u and v is zero, then we must have $u = 0$ and $v = 0$ from the second equation since $D_x(x,y,\zeta)$ and $D_y(x,y,\zeta)$ are nonsingular by Prop. 6.1(c). Together with the first equation and the assumption of rank $E'_\zeta(x,y,\zeta) = l$, we get $s = 0$. Thus, we prove that $u = 0, v = 0, s = 0$ under this condition. If $u \neq 0$ and $v \neq 0$, then the first equation of (6.9) and the given assumption imply that there exists a $\nu \in \{1,\ldots,m\}$ such that $u_\nu \neq 0$ and $\langle u_\nu, v_\nu \rangle \geq 0$. Note that the second equation of (6.9) is equivalent to

$$\varphi'_{x_i}(x_i,y_i,\varepsilon)u_i + \varphi'_{y_i}(x_i,y_i,\varepsilon)v_i = 0 \quad \text{for all } i = 1,2,\ldots,m.$$

For $i = \nu$, since $\varphi'_{y_\nu}(x_\nu,y_\nu,\varepsilon)$ is nonsingular by Prop. 6.1(d), it follows that

$$u_\nu^T \left[\varphi'_{y_\nu}(x_\nu,y_\nu,\varepsilon)\right]^{-1} \left[\varphi'_{x_\nu}(x_\nu,y_\nu,\varepsilon)\right] u_\nu + \langle u_\nu, v_\nu \rangle = 0.$$

From Prop. 6.1(e), the matrix $\left[\varphi'_{y_\nu}(x_\nu,y_\nu,\varepsilon)\right]^{-1} \left[\varphi'_{x_\nu}(x_\nu,y_\nu,\varepsilon)\right]$ is positive definite, and from the last equality and $\langle u_\nu, v_\nu \rangle \geq 0$, it then follows that $u_\nu = 0$. Thus, we obtain a contradiction. The proof is completed. $\square$

The condition in Prop. 6.3(b) is weaker than the conditions (6.2)–(6.3) of [25]. When $l = 0$, the condition of Prop. 6.3(b) is equivalent to saying that E'_x and E'_y are Cartesian column monotone, whereas the condition (6.3) of [25] is equivalent to saying that E'_x and E'_y are column monotone. For the SOCCP (1.4), the condition in Prop. 6.3(b) is equivalent to requiring the Cartesian P_0-property of F', whereas the condition (6.3) of [25] is equivalent to requiring the positive semidefiniteness of F'. Recently, for the SOCCP (1.4), Chua et al. [19] establish the nonsingularity of $\Theta'(\omega)$ with $\varphi = \varphi_{\text{CM}}$ and φ_{SQ} under the uniform nonsingularity of F, which is another nonmonotone property of F. Now it is not clear whether this condition is weaker than the Cartesian P_0-property of F' for differentiable F.

Let $\Xi: \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}_+$ be the natural merit function of $\Theta(w) = 0$, i.e.,

$$\Xi(\omega) = \|\Theta(w)\|^2 \quad \forall \omega \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l. \tag{6.10}$$

The following proposition provides the coerciveness conditions of Ξ for the SOCCP (1.3).

Proposition 6.4. *Let Θ and Ξ be defined by (1.12) and (6.10), respectively. Suppose that*

$$E(x,y,\zeta) \equiv \left(\begin{array}{c} F(\zeta) - x \\ G(\zeta) - y \end{array} \right)$$

with F and G being continuous. Then Ξ is coercive under (C.1) or (C.2) of Prop. 4.2.

Proof. We first prove the result under the condition (C.1) of Prop. 4.2. Suppose on the contrary that there exist a constant $\gamma \geq 0$ and a sequence $\{\omega^k\}$ with $\|\omega^k\| \to \infty$ such that $\Xi(\omega^k) \leq \gamma$. Since $\{\varepsilon^k\}$ is bounded by $\Xi(\omega^k) \leq \gamma$, we must have $\|(x^k, y^k, \zeta^k)\| \to \infty$. Observe that $\|\zeta^k\| \to \infty$ necessarily holds. If not, using the continuity of F and G, and $\Xi(\omega^k) \leq \gamma$, we deduce that $\{x^k\}$ and $\{y^k\}$ are bounded, which contradicts the fact that $\|(x^k, y^k, \zeta^k)\| \to \infty$. From the uniform Jordan P-property and the linear growth of F and

G, it then follows that $\|F(\zeta^k)\|, \|G(\zeta^k)\| \to \infty$. If not, we assume without loss of generality that $\{F(\zeta^k)\}$ is bounded. Define the bounded sequence $\{\xi^k\}$ by

$$\xi_i^k = \begin{cases} \zeta_i^k & \text{if } i \in J \\ 0 & \text{otherwise} \end{cases}$$

where $J := \left\{ i \in \{1, \ldots, m\} \mid \|\zeta_i^k\| \text{ is unbounded} \right\} \neq \emptyset$. Using the boundedness of $\{F(\zeta^k)\}$ and $\{F(\xi^k)\}$, and the linear growth of G, we obtain that

$$\lambda_2 \left[(F(\zeta^k) - F(\xi^k)) \circ (G(\zeta^k) - G(\xi^k)) \right] \leq C_1 \|\zeta^k\| + C_2$$

for some $C_1, C_2 > 0$, which contradicts the uniform Jordan P-property of F and G. Thus,

$$\|\zeta^k\| \to +\infty, \quad \|F(\zeta^k)\| \to \infty, \quad \|G(\zeta^k)\| \to \infty, \quad \|x^k\| \to \infty, \quad \|y^k\| \to \infty.$$

By Prop. 6.1 and $\Xi(\omega^k) \leq \gamma$, we have that $\{\phi(x^k, y^k)\}$ is bounded with $\phi = \phi_{\text{NR}}$ or ϕ_{FB}. This together with Lemma 3.5 and the last equation implies that $\{\lambda_1(x^k)\}$ and $\{\lambda_1(y^k)\}$ are bounded below, but $\lambda_2(x^k), \lambda_2(y^k) \to +\infty$. From the proof of [47, Prop. 4.2(a)], we know that the uniform Jordan P-property and the linear growth of F and G implies that $\lim_{k\to\infty} \frac{F(\zeta^k)}{\|F(\zeta^k)\|} \circ \frac{G(\zeta^k)}{\|G(\zeta^k)\|} \neq 0$. Using the boundedness of $\{F(\zeta^k) - x^k\}$ and $\{G(\zeta^k) - y^k\}$, it is easy to argue that $\lim_{k\to\infty} \frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|} \neq 0$. From Lemma 3.5, it then follows that $\{\phi(x^k, y^k)\}$ is unbounded, which is impossible.

When the condition (C.2) of Prop. 4.2 holds, using the similar arguments as above and those of [47, Prop. 4.2(d)], we can prove the desired result. □

Note that, when $E(x, y, \zeta) \equiv \begin{pmatrix} Ax - b \\ A^T \zeta + y - c \end{pmatrix}$ with $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$, the row full rank of A can not guarantee the coerciveness of Ξ. For example, let

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 2 \end{pmatrix}, \quad x^k = \begin{pmatrix} k \\ k-1 \\ 0 \end{pmatrix}, \quad y^k = 0, \quad \varepsilon^k \in [0, 1],$$

and $\{\zeta^k\}$ is an arbitrary bounded sequence in $\mathbb{R}^2$. Clearly, A has full row rank. Since $x^k \in \mathcal{K}^3$ and $y^k = 0$, we have that $\{\phi(x^k, y^k)\}$ is bounded with $\phi = \phi_{\text{NR}}$ or ϕ_{FB}, and so is $\{\theta(x^k, y^k, \varepsilon^k)\}$. In addition, it is easy to verify that $\{Ax^k - b\}$ and $\{A^T \zeta^k + y^k - c\}$ are bounded. This means that $\{\Xi(\omega^k)\}$ is bounded, but $\|\omega^k\| \to \infty$, i.e., Ξ is not coercive. This partly interprets why in Subsection 6.2 using the smoothing method below to solve some linear SOCPs requires much more iterations than the interior point methods.

Motivated by the efficiency of the smoothing Newton method [54], we next apply this method for solving the SOC complementarity system (1.1), i.e., we want to obtain a solution of (1.1) by solving a single augmented smooth system $\Theta(\omega) = 0$. Choose $\bar{\varepsilon} > 0$ and $\gamma \in (0, 1)$ such that $\bar{\varepsilon}\gamma < 1$, and let $\bar{\omega} = (\bar{\varepsilon}, 0) \in \mathbb{R}_{++} \times \mathbb{R}^{2n+l}$. Define

$$\beta(\omega) := \gamma \min \{1, \Xi(\omega)\}.$$

The smoothing Newton method [54] for solving system (1.1) is described as follows.

Algorithm 6.1 (Smoothing Newton method)

Step 0. Select a smoothing function φ of ϕ_{NR} or ϕ_{FB}. Choose constants $\delta \in (0, 1)$ and $\sigma \in (0, 1/2)$, and a point $(x^0, \zeta^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n$. Let $\varepsilon^0 = \bar{\varepsilon}$ and $k := 0$.

Step 1. If $\Theta(\omega^k) = 0$, then stop. Otherwise, let $\beta_k := \beta(\omega^k)$.

Step 2. Compute the direction $d\omega^k := (d\varepsilon^k, dx^k, dy^k, d\zeta^k) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$ by

$$\Theta(\omega^k) + \Theta'(\omega^k)d\omega = \beta_k\bar{\omega}. \tag{6.11}$$

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$\Xi(\omega^k + \delta^l d\omega^k) \leq \left[1 - 2\sigma(1 - \gamma\|\bar{w}\|)\delta^l\right] \Xi(\omega^k). \tag{6.12}$$

Step 4. Define $\omega^{k+1} := \omega^k + \delta^{l_k} d\omega^k$. Let $k := k + 1$, and then go to Step 1.

From Prop. 6.3(a), it follows that the mapping $\Theta(\cdot)$ is continuously differentiable at any $\omega^k \in \mathbb{R}_{++} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$, and if for each $\varepsilon > 0$ and $(x, y, \zeta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l$, the mapping E satisfies the condition of Prop. 6.3(b), then $\Theta'(\omega)$ is nonsingular. As remarked after Prop. 6.3, there are many types of SOCCPs, even nonmonotone SOCCPs, such that E satisfies this condition. The main computation work of Algorithm 6.1 is to calculate the direction $d\omega^k$ by (6.11). In Subsection 6.2, we analyze that for the standard linear SOCPs, the calculation of $d\omega^k$ needs only one factorization of an $m \times m$ positive definite matrix, whereas for nonlinear SOCPs, it requires one factorization of an $n \times n$ positive definite matrix and one factorization of an $m \times m$ positive definite matrix.

Note that $\bar{\varepsilon} > 0$ and the starting $\omega^0 = (\varepsilon^0, x^0, y^0, \zeta^0)$ belongs to the following set

$$\Omega := \left\{\omega = (\varepsilon, x, y, \zeta) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^l \mid \varepsilon \geq \beta(\omega)\bar{\varepsilon}\right\}.$$

Therefore, using the same arguments as those of [54], we can prove that Algorithm 6.1 generates an infinite sequence $\{\omega^k\}$ with $\varepsilon^k \in \mathbb{R}_{++}$ and $\omega^k \in \Omega$, provided that for each k with $\varepsilon^k > 0$ and $\omega^k \in \Omega$, the Jacobian $\Theta'(\omega^k)$ is invertible. Particularly, we have the following global convergence result, whose proof is similar to that of [54, Theorem 4].

Theorem 6.5. *Suppose that for all $\omega = (\varepsilon, x, y, \zeta) \in \Omega$, rank $E'_\zeta(x, y, \zeta) = l$ and the implication (6.8) holds for any $u \neq 0, v \neq 0$. Then, an infinite sequence $\{\omega^k\}$ is generated by Algorithm 6.1 and each accumulation point ω^* of $\{\omega^k\}$ is a solution of $\Theta(w) = 0$.*

When applying Algorithm 6.1 for the SOCCP (1.3), the sequence generated is bounded if the mappings F and G satisfy one of the conditions of Prop. 4.2. For the linear SOCPs, the counterexample after Prop. 6.4 shows that the full row rank of A can not guarantee the boundedness of level sets of the merit function Ξ, although numerical results in Subsection 6.2 demonstrate that the sequence generated by Algorithm 6.1 is generally bounded for this class of problems.

Assume that φ in Algorithm 6.1 is chosen as φ_{CM} with g given by (6.3), φ_{SQ} or φ_{FB}. By Prop. 6.2(c), the operator Θ is semismooth everywhere whenever E is semismooth, and strongly semismooth everywhere whenever E is strongly semismooth. Therefore, under the nonsingularity assumption of $\partial_B\Theta(\omega^*)$, using the similar arguments to those of [54], we can establish the following local superlinear (or quadratic) convergence results.

Theorem 6.6. *Suppose that for all $\omega = (\varepsilon, x, y, \zeta) \in \Omega$, rank $E'_\zeta(x, y, \zeta) = l$ and the implication (6.8) holds for any $u \neq 0, v \neq 0$, and ω^* is an accumulation point of the infinite sequence $\{\omega^k\}$ generated by Algorithm 6.1 with $\varphi = \varphi_{\mathrm{CM}}$ for g given by (6.3), φ_{SQ} or φ_{FB}.*

If E is semismooth at ω^* and all $V \in \partial_B \Theta(\omega^*)$ are nonsingular, then the whole sequence $\{\omega^k\}$ converges to ω^*, and

$$\|\omega^{k+1} - \omega^*\| = o(\|\omega^k - \omega^*\|) \quad \text{and} \quad \varepsilon^{k+1} = o(\varepsilon^k).$$

If, in addition, E is strongly semismooth at ω^*, then

$$\|\omega^{k+1} - \omega^*\| = O(\|\omega^k - \omega^*\|^2) \quad \text{and} \quad \varepsilon^{k+1} = O((\varepsilon^k)^2).$$

Using the same arguments as in [64] and [48], it is not hard to verify that the conditions of Theorems 5.5 and 5.8 may guarantee the nonsingularity of $\partial_B \Theta(\omega^*)$. Thus, analogous to the NR and FB semismooth Newton methods, the local superlinear (or quadratic) convergence of the smoothing Newton methods with the CHKS smoothing function, the log-exponential smoothing function, the squared smoothing function and the FB smoothing function do not require the strict complementarity of solutions.

6.2 Applications of Smoothing Newton Methods

In what follows, we use Algorithm 6.1 with φ chosen as the CHKS smoothing function, the squared smoothing function, and the FB smoothing function, respectively, to solve the SOCP (1.5) with a twice continuously differentiable convex f. Unless otherwise stated, φ_{CM} appearing in the subsection denotes the CHKS smoothing function.

First, let us take a closer look at the calculation of Newton direction $d\omega^k$ for this class of problems. Let θ be defined by (6.6) with φ being one of $\varphi_{\text{CM}}, \varphi_{\text{SQ}}$ and φ_{FB}. It is easy to see that the KKT optimality conditions of (1.5) can be reformulated as $\Theta(\omega) = 0$ with

$$E(x, y, \zeta) \equiv \begin{pmatrix} Ax - b \\ A^T \zeta + y - \nabla f(x) \end{pmatrix}.$$

By Prop. 6.3(a) and the expression of E, equation (6.11) can be equivalently written as

$$
\begin{aligned}
d\varepsilon &= \beta_k \bar{\varepsilon} - \varepsilon^k \\
A dx &= b - Ax^k \\
-\nabla^2 f(x^k) dx + dy + A^T d\zeta &= \nabla f(x^k) - A^T \zeta^k - y^k \\
D_x^k dx + D_y^k dy &= -\theta'_\varepsilon(x^k, y^k, \varepsilon^k) d\varepsilon - \varphi(x^k, y^k, \varepsilon^k)
\end{aligned}
\tag{6.13}
$$

where $D_x^k = D_x(x^k, y^k, \zeta^k)$ and $D_y^k = D_y(x^k, y^k, \zeta^k)$ with $D_x(\cdot, \cdot, \cdot)$ and $D_y(\cdot, \cdot, \cdot)$ defined by (6.7). Since D_x^k and D_y^k are nonsingular by Prop. 6.1(d) and the definition of θ, we obtain from the last two equations of (6.13) that

$$\left[(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k) \right] dx = A^T d\zeta + \text{Res}^k \tag{6.14}$$

where

$$\text{Res}^k = (-\nabla f(x^k) + A^T \zeta^k + y^k) - (D_y^k)^{-1} \left[\theta'_\varepsilon(x^k, y^k, \varepsilon^k) d\varepsilon + \varphi(x^k, y^k, \varepsilon^k) \right].$$

Since $(D_y^k)^{-1} D_x^k$ is positive definite by the definition of θ and Prop. 6.1(e), and $\nabla^2 f(x^k)$ is symmetric positive semidefinite by the convexity of f, it follows that $(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k)$ is nonsingular. Using (6.14) and the second equation of (6.13) then yields that

$$d\zeta = (\Sigma^k)^{-1} \left[(b - Ax^k) - A \left[(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k) \right]^{-1} \text{Res}^k \right], \tag{6.15}$$

where

$$\Sigma^k = A \left[(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k) \right]^{-1} A^T.$$

Substituting $d\zeta$ into (6.14) yields dx, and dy follows from the third equation of (6.13).

We see from (6.15) that, if $\nabla^2 f(x^k) \neq 0$, then calculating $d\omega^k$ requires a factorization of $n \times n$ positive definite matrix $[(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k)]$ and a factorization of $m \times m$ positive definite matrix Σ^k. Note that, when $\varphi = \varphi_{\mathrm{CM}}$, the ith block of the block diagonal matrix $(D_y^k)^{-1} D_x^k$ can be achieved by calculating $(\nabla g^{\mathrm{soc}}(z_i))^{-1} - I$ via the formula in Lemma 2.2; and when $\varphi = \varphi_{\mathrm{SQ}}$ and φ_{FB}, this can be achieved by calculating $L_{z_i+(x_i-y_i)}^{-1} L_{z_i-(x_i-y_i)}$ and $L_{z_i-y_i}^{-1} L_{z_i-x_i}$, respectively, via the formula in (2.1), where z_i has the same expression as z in Prop. 6.1(b). If $\nabla^2 f(x^k) = 0$, then the computation work of $d\omega^k$ is greatly reduced. Under this case, $\Sigma^k = \sum_{i=1}^m A_i D_i^k A_i^T$ with $A_i \in \mathbb{R}^{m \times n_i}$ such that $A = [A_1 \ \cdots \ A_m]$, and

$$D_i^k = \begin{cases} (I - \nabla g^{\mathrm{soc}}(z_i))^{-1} - I & \text{if } \varphi = \varphi_{\mathrm{CM}}, \\ L_{z_i-(x_i-y_i)}^{-1} L_{z_i+x_i+y_i} & \text{if } \varphi = \varphi_{\mathrm{SQ}}, \\ L_{z_i-x_i}^{-1} L_{z_i-y_i} & \text{if } \varphi = \varphi_{\mathrm{FB}}. \end{cases}$$

So, the calculation of $d\omega^k$ only requires a factorization of $m \times m$ positive definite matrix Σ^k, where using the same technique as in Lemma 2.2 yields that $(I - \nabla g^{\mathrm{soc}}(z_i))^{-1}$ equals

$$\frac{1}{(1-b(z_i))^2 - c^2(z_i)} \begin{bmatrix} 1 - b(z_i) & c(z_i) \dfrac{z_{i2}^T}{\|z_{i2}\|} \\ c(z) \dfrac{z_{i2}}{\|z_{i2}\|} & \dfrac{(1-b(z_i))^2 - c^2(z_i)}{1 - a(z_i)} \left(I - \dfrac{z_{i2} z_{i2}^T}{\|z_{i2}\|^2} \right) + \dfrac{(1-b(z_i)) z_{i2} z_{i2}^T}{\|z_{i2}\|^2} \end{bmatrix}.$$

We implemented the nonmontone line search version of Algorithm 6.1, i.e., in Step 3 we compute the smallest nonnegative integer l_k satisfying $\omega^k + \delta^{l_k} \in \Omega$ and

$$\Xi(\omega^k + \delta^{l_k} d\omega^k) \leq \mathcal{W}_k - 2\sigma(1-\gamma)\delta^{l_k}\Xi(\omega^k)$$

where $\mathcal{W}_k = \max_{j=k-m_k,\cdots,k} \Xi(\zeta^j)$ and where, for a given nonnegative integer $\hat{m}$ and s,

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

During the testing, we used $\hat{m} = 5, s = 5$ and the following parameters of Algorithm 6.1:

$$\delta = 0.5, \quad \sigma = 0.5 \times 10^{-4}, \quad \gamma = 0.5. \tag{6.16}$$

We terminated the algorithm once $\Xi(\omega^k) \leq \epsilon$ and $k \leq 100$. Throughout the tests, the linear system of equations $\Sigma^k d\zeta = \mathrm{Res}^k$ is computed by $d\zeta = \Sigma^k \backslash \mathrm{Res}^k$, and the inverse of $[(D_y^k)^{-1} D_x^k + \nabla^2 f(x^k)]$ with $\nabla^2 f(x^k) \neq 0$ were computed by "inv" of Matlab. All tests were done in Matlab 6.5 on a PC of Pentium 4 with 2.8GHz CPU and 512MB memory.

We tested two groups of convex SOCP instances. The first one is composed of four standard linear SOCPs from the DIMACS Implementation Challenge library [52]. We solved them by Algorithm 6.1 with error tolerance ϵ chosen as 1.0×10^{-8}, and the interior point method software SeDuMi [59] with the default tolerances, respectively. We tested that $\bar{\varepsilon} = 0.5$ are favorable for the FB smoothing method and the squared smoothing method, whereas $\bar{\varepsilon} = 0.1$ is suitable for the CHKS smoothing method. Table 3 reports the results of three smoothing methods with such $\bar{\varepsilon}$, and those of the SeDuMi, where **NF**, **Iter**, **Time**

Table 3: **Numerical results of Algorithm 6.1 for linear SOCPs**

Problem	**nb**					**nb_L1**				
	NF	Iter	$\Xi(\omega^k)$	Optval	Time	NF	Iter	$\Xi(\omega^k)$	Optval	Time
$\varphi = \varphi_{\mathrm{FB}}$	51	29	5.88e–9	−0.0507032	38.6	124	45	1.52e–12	−13.0122707	92.9
$\varphi = \varphi_{\mathrm{CM}}$	228	64	8.89e–9	−0.0507044	88.3	380	61	2.35e–9	−13.0122708	134.3
$\varphi = \varphi_{\mathrm{SQ}}$	146	37	8.34e–9	−0.0507031	56.6	196	48	7.45e–10	−13.0122705	104.6
SeDuMi		21		−0.0507031	14.8		18		−13.0122706	16.7
Problem	**nb_L2**					**nb_L2_bessel**				
$\varphi = \varphi_{\mathrm{FB}}$	32	15	3.58e–9	−1.6290027	95.1	28	15	5.02e–13	−0.1025695	22.8
$\varphi = \varphi_{\mathrm{CM}}$	32	11	2.47e–10	−1.6289721	68.1	14	7	9.60e–11	−0.1025940	9.6
$\varphi = \varphi_{\mathrm{SQ}}$	27	13	2.95e–9	−1.6291627	81.6	18	11	7.43e–14	−0.1025706	15.9
SeDuMi		15		−1.6289719	30.8		16		−0.1025695	17.7

and **Optval** have the same meaning as in Table 2.

Table 3 shows that the three smoothing methods with the above $\bar{\varepsilon}$ solve all test problems with the given accuracy successfully. We note that for the easier problems "nb_L2" and "nb_L2_bessel", the FB smoothing method needs a little more iterations and function evaluations than the other two smoothing methods, but for difficult "nb" and "nb_L1", the former is superior to the latter. The squared smoothing method has better performance than the CHKS smoothing method for the two difficult problems. We also found that, when setting $\epsilon = 10^{-10}$, the FB smoothing method may yield the solutions with desirable accuracy for "nb_L1" within 100 iterations, but the CHKS smoothing method and the squared smoothing method fail to this. In addition, compared with the SeDuMi, the three smoothing methods require more iterations for those two difficult problems, but less iterations for those simple problems.

The second group of test instances is composed of nonlinear convex SOCPs in (1.5) with $f(x)$ and $\mathcal{K}$ same as in Subsection 5.2. The matrices $Q \in \mathbb{R}^{n \times n}$ and $A \in \mathbb{R}^{m \times n}$ and the vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are generated in the same way as in Subsection 5.2. We applied the above nonmonotone line search version of Algorithm 6.1 with the parameters in (6.16) and $\epsilon = 3.0 \times 10^{-9}$ for the same 50 test instances as in Subsection 5.2. Table 4 reports the results of three smoothing methods with two different $\bar{\varepsilon}$, where **SN** means the number of problems with desired accuracy, **Fail** column gives the No. of the failure problems, and

ANF columns lists the average function evaluations of the successful problems.

Table 4: **Numerical results of Algorithm 6.1 for** 50 **convex SOCPs**

Name	$\bar{\varepsilon}$	**SN**	**Fail**	**ANF**	$\bar{\varepsilon}$	**SN**	**Fail**	**ANF**
FB	0.5	47	6th, 19th, 21th	12	0.1	42	3-5th, 15th, 20th, 49th, 40-41th	10
CHKS	0.5	37	11-12th, 16th,21-22th 24th, 28th, 32th, 42-44th, 48th	10	0.1	50		8
SQ	0.5	47	16th, 19th, 24th	11	0.1	50		8

From Table 4, we observe that the value of $\bar{\varepsilon}$ has influence on the performance of three smoothing methods, and the influence is remarkable for the CHKS smoothing method. In other words, the FB smoothing method and the squared smoothing method have better robustness than the CHKS smoothing method. The FB smoothing method requires a little more function evaluations and iterations than the other two smoothing methods. Together with the results in Table 3, we conclude that the squared smoothing method is superior to the CHKS smoothing method, and for those difficult problems, it seems that the FB smoothing method has better performance than the squared smoothing method.

Comparing the results in Tables 3–4 with those of Subsection 5.2, we see that the FB semismooth method has comparable performance with the three smoothing methods in terms of functions evaluations, the number of iterations, and the accuracy, but the former requires more CPU time than the latter since the former requires one factorization of an $(2n+m) \times (2n+m)$ nonsingular matrix at each iteration. For example, for the nonlinear convex SOCP instances, the average CPU time required by the FB semismooth method for each problem is $857.3(s)$, whereas that of the smoothing methods is less than $300(s)$.

7 Conclusions

We have made a survey for the properties of SOC complementarity functions and theoretical results of related solution methods, including the merit function methods, the semismooth Newton method and the smoothing Newton methods, and pay attentions to the performance of these methods. Among the four classes of common merit functions, the FB merit function is the most desirable whether in theory or in numerical performance, although the LT merit function Ψ_{LT} and its variant $\widehat{\Psi}_{\mathrm{LT}}$ has some advantages in some aspects. Among the three popular smoothing methods, the squared smoothing method seems superior to the CHKS smoothing method, and the FB smoothing method has better performance than the squared smoothing method for those difficult problems. The global FB semismooth Newton method is comparable with the smoothing Newton methods in terms of iterations and the accuracy of solutions, despite of more CPU time.

In addition, we also observe that, compared with primal-dual interior point methods (such as the SeduMi software [59]), the semismooth Newton method and the smoothing Newton methods have worse performance for "nb_L1". One of main reasons is, as remarked after Prop. 6.4, the natural merit functions of systems $\Phi(z) = 0$ and $\Theta(\omega) = 0$ may have unbounded level sets only under the full row rank of A.

References

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

[2] J.F. Bonnans and C.H. Ramírez, Perturbation analysis of second-order cone programming problems, *Mathematical Programming* 104 (2005) 205–227.

[3] S.-J. Bi, S.-H. Pan and J.-S. Chen, The same growth of FB and NR symmetric cone complementarity functions, to appear in *Optimization Letters* 2012.

[4] R.H. Byrd, P. Lu, J. Nocedal and C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM Journal of Scientific Computing* 16 (1995) 1190–1208.

[5] Z.X. Chan and D. Sun, Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming, *SIAM Journal on Optimization* 19 (2008) 370–396.

[6] B. Chen, Error bounds for R_0-type and monotone nonlinear complementarity problems, *Journal of Optimization Theory and Applications* 108 (2001) 297–316.

[7] B. Chen and P.T. Harker, A non-interior-point continuation method for linear complementarity problems, *SIAM Journal on Matrix Analysis and Applications* 14 (1993) 1168–1190.

[8] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Computational Optimization and Applications* 5 (1996) 97–138.

[9] J.-S. Chen, X. Chen, and P. Tseng, Analysis of nonsmooth vector-valued functions associated with second-order cones, *Mathematical Programming* 101 (2004) 95–117.

[10] J.-S. Chen and P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, *Mathematical Programming* 104 (2005) 297–327.

[11] J.-S. Chen, Alternative proofs for some results of vector-valued functions associated with second-order cone, *Journal of Nonlinear and Convex Analysis* 6 (2005) 297–325.

[12] J.-S. Chen, Two classes of merit functions for the second-order cone complementarity problem, *Mathematical Methods of Operations Research* 64 (2006) 495–519.

[13] J.-S. Chen, Conditions for error bounds and bounded level sets of some merit functions for SOCCP, *Journal of Optimization Theory and Applications* 135 (2007) 459–473.

[14] J.-S. Chen and S.-H. Pan, A one-parametric class of merit functions for the second-order cone complementarity problem, *Computational Optimization and Applications* 45 (2010) 581–606.

[15] J.-S. Chen and S.-H. Pan, Lipschitz continuity of the gradient of a one-parametric class of SOC merit functions, *Optimization* 59 (2010) 661–676.

[16] J.-S. Chen, D.-F. Sun, and J. Sun, The SC^1 property of the squared norm of the SOC Fischer-Burmeister function, *Operations Research Letters* 36 (2009) 385–392.

[17] X. Chen and H. Qi, Cartesian P-proeprty and its applications to the semidefinite linear complementarity problem, *Mathematical Programming* 106 (2006) 177–201.

[18] X.-D. Chen, D. Sun, and J. Sun, Complementarity functions and numerical experiments for second-order cone complementarity problems, *Computational Optimization and Applications* 25 (2003) 39–56.

[19] C.-B. Chua, H.-L. Lin and P. Yi, Unfiorm nonsingularity and complementarity problems over symmetric cones, Optimization online, 2009.

[20] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).

[21] Y.G. Evtushenko and V.A. Purtov, Sufficient conditions for a minimum for nonlinear programming problems, *Soviet Mathematics Doklady* 30 (1984) 313–316.

[22] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II, Springer-Verlag, New York, 2003.

[23] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[24] M.C. Ferris, O.L. Mangasarian, and J.-S. Pang (eds), *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.

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

[26] L. Grippo, F. Lampariello and S. Lucidi, A nonmonotone line search technique for Newton's method, *SIAM Journal on Numerical Analysis* 23 (1986) 707–716.

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

[28] S. Hayashi, N. Yamashita, and M. Fukushima, Robust Nash equilibria and second-order cone complementarity problems, *Journal of Nonlinear and Convex Analysis* 15 (2005) 593–615.

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

[30] C. Kanzow and M. Fukushima, Equivalence of the generalized complementarity problems to differentiable unconstrained minimization, *Journal of Optimization Theory and Applications* 90 (1996) 581–613.

[31] C. Kanzow and H. Kleinmichel, A new class of semismooth Newton-type methods for nonlinear complementarity problems, *Computational Optimization and Applications* 11 (1998) 227–251.

[32] C. Kanzow, I. Ferenczi and M. Fukushima, On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity, *SIAM Journal on Optimization* 20 (2009) 297–320.

[33] L. Kong, L. Tuncel, and N. Xiu, Vector-valued Implicit Lagrangian for symmetric cone complementarity problems, *Asia-Pacific Journal of Operational Research* 26 (2009) 199–233.

[34] Y. Kanno, J.A.C. Martins and A. Pinto da Costa, Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem, *International Journal for Numerical Methods in Engineering* 65 (2006) 62–83.

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

[36] T. De. Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, *Mathematical Programming* 75 (1996) 407–439.

[37] Z.-Q. Luo and P. Tseng, A new class of merit functions for the nonlinear complementarity problem, in *Complementarity and Variational Problems: State of the Art*, M. C. Ferris and J.-S. Pang (eds), SIAM, Philadelphia, 1997, pp. 204–225.

[38] O.L. Mangasarian and M.V. Solodov, Nonlinear complentarity as unconstrained and constrained minimization, *Mathematical Programming* 62 (1993) 277–297.

[39] O.L. Mangasarian and M.V. Solodov, A linearly convergent derivative-free descent method for strongly monotone complementarity problems, *Computational Optimization and Applications* 64 (1999) 5–16.

[40] F. Meng, D. Sun and G. Zhao, Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization, *Mathematical Programming* 104 (2005) 561–581.

[41] 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* 88 (2000) 61–83.

[42] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.

[43] S.-H. Pan and J.-S. Chen, A damped Gauss-Newton method for the second-order cone complementarity problem, *Applied Mathematics and Optimization* 59 (2009) 293–318.

[44] S.-H. Pan and J.-S. Chen, A semismooth Newton method for SOCCPs based on a one-parametric class of complementarity functions, *Computational Optimization and Applications* 45 (2010) 59–88.

[45] S.-H. Pan and J.-S. Chen, Growth behavior of two classes of merit functions for symmetric cone complementarity problems, *Journal of Optimization Theory and Applications* 141 (2009) 167–191.

[46] S.-H. Pan and J.-S. Chen, A linearly convergent derivative-free descent method for the second-order cone complementarity problem, *Optimization* 59 (2010) 1173–1197.

[47] S.-H. Pan, J.-S. Chen and Z.-H. Huang, A survey on coerciveness conditions of merit functions associated with symmetric cones, manuscript, 2011.

[48] S.-J. Bi, S.-H. Pan and J.-S. Chen, Nonsingular conditions for FB system of nonsmooth SDPs, *SIAM Journal on Optimization* 21 (2011) 1392–1417.

[49] J.-S. Pang and L. Qi, Nonsmooth equations: motivation and algorithms, *SIAM Journal on Optimization* 3 (1993) 443–465.

[50] J.M. Peng, Equivalence of varational inequality problems to unconstrained minimization, *Mathematical Programming* 78 (1997) 347–355.

[51] J.M. Peng, C. Roos and T. Terlaky, A new class of polynomial primal-dual interior-point methods for second-order cone optimization based on self-regular proximities, *SIAM Journal on Optimization* 13 (2002) 179–203.

[52] G. Pataki and S. Schmieta, *The DIMACS library of semidefinite-quadratic-linear programs*, Preliminary draft, Computational Optimization Research Center, Columbia University, New York. http://dimacs.rutgers.edu/Challenges

[53] J.-S. Pang, S. Sun, and J. Sun, Semismooth homemorphisms and strong stability of semidefinite and Lorentz complementarity problems, *Mathematics of Operations Research* 28 (2003) 39–63.

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

[55] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research* 18 (1993) 227–244.

[56] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming* 58 (1993) 353–367.

[57] S. Smale, Algorithms for solving equations, in *Proceedings of the International Congresss of Mathematicians*, American Mathematical Society, Providence, RI, 1987, pp. 172–195.

[58] D. Sun and J. Sun, Strong semismoothness of Fischer-Burmeister SDC and SOC function, *Mathematical Programming* 103 (2005) 575–581.

[59] J.F. Sturm, Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones, *Optimization Methods and Software* 11 & 12 (1999) 625–653.

[60] J. Tao and M.S. Gowda, Some *P*-properties for nonlinear transformations on Euclidean Jordan algebras, *Mathematics of Operations Research* 30 (2005) 985–1004.

[61] P. Tseng, Merit functions for semi-definite complementarity problems, *Mathematical Programming* 83 (1998) 159–185.

[62] P. Tseng, N. Yamashita and M. Fukushima, Equivalence of complementarity problems to differentiable minimization: A unified approach, *SIAM Journal on Optimization* 6 (1996) 446–460.

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

[64] Y. Wang and L. W. Zhang, Properties of equation reformulation of the Karush-Kuhn-Tucker condition for nonlinear second-order cone optimization problems, *Mathematical Methods of Operations Research* 70 (2009) 195–218.

[65] N. Yamashita and M. Fukushima, On the stationary points of the Implicit Lagrangian for nonlinear complementarity problems, *Journal of Optimization Theory and Applications* 84 (1995) 653–663.

[66] 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*, M. Fukushima and L. Qi (eds), Kluwer Academic Publishers, Boston, 1999, pp. 405–420.

[67] N. Yamashita and M. Fukushima, On the level-boundedness of the natural residual function for variational inequality problems, *Pacific Journal of Optimization* 1 (2005) 625–630.

[68] K. Yamada, N. Yamashita and M. Fukushima, A new derivative-free descent method for the nonlinear complementarity problems, in *Nonlinear Optimization and Related Topics*, G.D. Pillo and F. Giannessi (eds), Kluwer Academic Publishers, Netherlands, 2000, pp. 463–489.

JEIN-SHAN CHEN
Department of Mathematics, National Taiwan Normal University
Taipei 11677, Taiwan
E-mail address: `jschen@math.ntnu.edu.tw`

SHAOHUA PAN
School of Mathematical Sciences, South China University of Technology
Guangzhou 510640, China
E-mail address: `shhpan@scut.edu.cn`