J. Math. Anal. Appl. 366 (2010) 164–180

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

A proximal gradient descent method for the extended second-order cone linear complementarity problem

Shaohua Pan [a,1], Jein-Shan Chen [b,*,2]

[a] School of Mathematical Sciences, South China University of Technology, Guangzhou 510640, China
[b] Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

ARTICLE INFO

Article history:
Received 9 February 2009
Available online 14 January 2010
Submitted by J.A. Filar

Keywords:
Extended second-order cone linear complementarity problems
Optimization reformulations
Proximal gradient method
Descent
Linear convergence rate

ABSTRACT

We consider an extended second-order cone linear complementarity problem (SOCLCP), including the generalized SOCLCP, the horizontal SOCLCP, the vertical SOCLCP, and the mixed SOCLCP as special cases. In this paper, we present some simple second-order cone constrained and unconstrained reformulation problems, and under mild conditions prove the equivalence between the stationary points of these optimization problems and the solutions of the extended SOCLCP. Particularly, we develop a proximal gradient descent method for solving the second-order cone constrained problems. This method is very simple and at each iteration makes only one Euclidean projection onto second-order cones. We establish global convergence and, under a local Lipschitzian error bound assumption, linear rate of convergence. Numerical comparisons are made with the limited-memory BFGS method for the unconstrained reformulations, which verify the effectiveness of the proposed method.

1. Introduction

We consider an extended second-order cone linear complementarity problem (SOCLCP) which is to find a pair of vectors $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}^p$ such that

$$\begin{cases} Mx - Ny + Pz \in \Omega, \\ x \in \mathcal{K}, \ \ y \in \mathcal{K}, \ \ \langle x, y \rangle = 0 \end{cases} \tag{1}$$

where M and N are $m \times n$ real matrices, P is an $m \times p$ real matrix, Ω is defined by

$$\Omega := \left\{ u \in \mathbb{R}^m \,\middle|\, Eu - r \in \mathcal{E} \right\} \tag{2}$$

with $E \in \mathbb{R}^{l \times m}$, $r \in \mathbb{R}^l$, and $\mathcal{E} \subseteq \mathbb{R}^l$ being a closed convex cone, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs), also called Lorentz cones. In other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_q} \tag{3}$$

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

1 The author's work is supported by National Young Natural Science Foundation (No. 10901058) and Guangdong Natural Science Foundation (No. 9251802902000001).
2 Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2010.01.011

where $q, n_1, \ldots, n_q \geqslant 1$, $n_1 + n_2 + \cdots + n_q = n$, and

$$\mathcal{K}^{n_i} := \left\{(x_1, x_2) \in \mathbb{R} \times \mathbb{R}^{n_i - 1} \;\middle|\; x_1 \geqslant \|x_2\|\right\}.$$

Throughout this paper, we assume that the SOCLCP (1) is feasible, i.e.,

$$\left\{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \;\middle|\; Mx - Ny + Pz \in \Omega,\ x \in \mathcal{K},\ y \in \mathcal{K}\right\} \neq \emptyset. \tag{4}$$

The SOCLCP (1) is a generalization of the extended linear complementarity problem that is known to have wide applications in linear and quadratic programming problems, bimatrix game problems, market and network equilibrium problems [2,20,24]. As will be illustrated in Section 2, the problem includes many special types of SOCLCPs, such as the generalized SOCLCP, the horizontal SOCLCP, the vertical SOCLCP, and the mixed SOCLCP, which can all be rewritten as (1) with $\mathcal{E} = \{0\}$. In view of the work in [16], we conjecture that the special cases of (1) with $\mathcal{E}$ being an SOC or nonnegative orthant cone will arise from some engineering and practical problems directly.

In recent ten years, there has been active interest in reformulating a nonpolyhedral symmetric cone complementarity problem as an optimization problem with suitable merit functions. For example, Tseng [25] first considered such reformulations for the semidefinite complementarity problem, Chen and Tseng [5] studied the Fischer–Burmeister unconstrained minimization reformulation for the second-order cone complementarity problem (SOCCP), Andreani et al. [1] proposed box-constrained minimization reformulations for a generalization of the SOCCP, and Kong et al. [17] studied the implicit Lagrangian reformulation for the general symmetric cone complementarity problem.

Motivated by Solodov's work [24] for the extended linear complementarity problem, in this paper we propose some simple SOC constrained reformulations and unconstrained reformulations for (1), and under mild conditions establish the equivalence between the stationary points of these optimization problems and the solutions of (1). Moreover, for these simple SOC constrained reformulation problems, we develop a proximal gradient descent method. The method has very small computation work at each iteration, and makes one Euclidean projection onto SOCs to generate a feasible descent direction. As will be demonstrated in Section 4, the method can be subsumed into the framework proposed in [8,26] for minimizing a sum of a smooth function and a convex separable function. Nevertheless, the analysis of its global convergence, and linear rate of convergence under a local Lipschitzian error bound, will become much simpler now.

In addition, for the proximal gradient descent method, we report numerical experience for solving (1) with $P = 0$ and $\mathcal{E}$ being the Cartesian product of SOCs or nonnegative orthant cone, and numerical comparisons with the limited-memory BFGS method [4] for the unconstrained minimization reformulation based on the Fischer–Burmeister merit function. The comparison results show that among the proposed constrained reformulations (see Section 3), the one based on the logarithmic function ψ_4 has better performance than those derived from the entropy function ψ_3 and the quadratic functions ψ_2 and ψ_5, and now Algorithm 4.1 has comparable performance with the limited-memory BFGS method when l is not close to m. To our best knowledge, there are no papers to discuss numerical performance of such equivalent constrained reformulation problems.

This paper is organized as follows. Section 2 reviews some background materials about SOCs and Jordan product, and illustrates that (1) includes many special SOCLCPs. In Section 3, we present some simple SOC constrained reformulations and unconstrained reformulations, and establish the equivalence between the stationary points of these optimization problems and the solutions of (1) under some mild conditions. In Section 4, a proximal gradient algorithm is developed for solving the equivalent SOC constrained reformulation problems, and the linear convergence of the algorithm is also established. In Section 5, numerical results are reported for the special cases of (1) in which $P = 0$, and $\mathcal{E}$ is the Cartesian product of SOCs or the nonnegative orthant cones $\mathbb{R}^l_+$.

Throughout this paper, I represents an identity matrix of suitable dimension, $\|\cdot\|$ denotes the Euclidean norm, $\mathrm{int}(\mathcal{K}^n)$ means the interior of $\mathcal{K}^n$, and $\mathbb{R}^n$ denotes the space of n-dimensional real column vectors, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_q}$ is identified with $\mathbb{R}^{n_1 + \cdots + n_q}$. For any $x, y \in \mathbb{R}^n$, we write $x \succeq_{\mathcal{K}} y$ and $x \succ_{\mathcal{K}} y$ to mean $x - y \in \mathcal{K}$ and $x - y \in \mathrm{int}(\mathcal{K})$, respectively. For any closed convex cone $\mathcal{E}$, the notation $[x]^+_{\mathcal{E}}$ means the minimum Euclidean norm projection of x onto $\mathcal{E}$, and $\mathcal{E}^\circ$ denotes the polar cone of $\mathcal{E}$, defined by

$$\mathcal{E}^\circ := \left\{v \in \mathbb{R}^l \;\middle|\; \langle v, u\rangle \leqslant 0 \text{ for all } u \in \mathcal{E}\right\}.$$

In addition, we denote $0^+\Omega$ by the recession cone of Ω, and from [23] it follows that

$$0^+\Omega = \left\{d \in \mathbb{R}^m \;\middle|\; Ed \in \mathcal{E}\right\},$$

and therefore the polar cone of $0^+\Omega$ is given by

$$\begin{aligned}\left(0^+\Omega\right)^\circ &= \left\{v \in \mathbb{R}^m \;\middle|\; \langle v, d\rangle \leqslant 0 \text{ for all } d \in 0^+\Omega\right\}\\ &= \left\{v \in \mathbb{R}^m \;\middle|\; v = E^T w \text{ for some } w \in \mathcal{E}^\circ\right\}.\end{aligned}$$

We recall that a square matrix Q is said to be *copositive* on $\mathcal{E}$ if $\langle Qv, v\rangle \geqslant 0$ for all $v \in \mathcal{E}$, and *strictly copositive* if the latter inequality is strict for all $0 \neq v \in \mathcal{E}$. A pair of matrices $M, N \in \mathbb{R}^{m\times n}$ is said to be *X-row-block-sufficient* with respect to (w.r.t.) Ω if

$$\left.\begin{aligned}&\left\langle \left(M^T \upsilon\right)_i, \left(N^T \upsilon\right)_i\right\rangle \leqslant 0, \quad i = 1, \ldots, q\\ &\upsilon \in \left(0^+\Omega\right)^\circ\end{aligned}\right\} \implies \left\langle \left(M^T \upsilon\right)_i, \left(N^T \upsilon\right)_i\right\rangle = 0, \quad i = 1, 2, \ldots, q.$$

When $q=n$, this reduces to the definition of X-row-sufficiency w.r.t. Ω in [13]. Clearly, the copositiveness of MN^T on $(0^+\Omega)^\circ$ implies X-row-block-sufficiency of M and N.

2. Preliminaries and examples

We start with the definition of Jordan product [7] associated with the cone $\mathcal{K}^n$. For any $x=(x_1,x_2),\ y=(y_1,y_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$, the Jordan product of x and y is defined as

$$x\circ y:=\big(\langle x,y\rangle,\ y_1x_2+x_1y_2\big). \tag{5}$$

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 $e:=(1,0,\ldots,0)^T\in\mathbb{R}^n$. Given a vector $x=(x_1,x_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$, let

$$L_x:=\begin{bmatrix} x_1 & x_2^T\\ x_2 & x_1I\end{bmatrix},$$

which can be viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. It is not hard to verify that $L_xy=x\circ y$ and $L_{x+y}=L_x+L_y$ for any $x,y\in\mathbb{R}^n$, L_x is positive semidefinite if and only if $x\in\mathcal{K}^n$, and L_x is positive definite if and only if $x\in\operatorname{int}(\mathcal{K}^n)$. Also, if L_x is invertible,

$$L_x^{-1}=\frac{1}{\det(x)}\begin{bmatrix} x_1 & -x_2^T\\ -x_2 & \frac{\det(x)}{x_1}I+\frac{1}{x_1}x_2x_2^T\end{bmatrix} \tag{6}$$

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

We recall from [7,9] that each $x=(x_1,x_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$ admits a spectral factorization associated with $\mathcal{K}^n$ in the form of

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

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

$$\lambda_i(x):=x_1+(-1)^i\|x_2\|,\qquad u_x^{(i)}:=\frac{1}{2}\big(1,\ (-1)^i\bar{x}_2\big) \tag{7}$$

with $\bar{x}_2=\frac{x_2}{\|x_2\|}$ if $x_2\neq 0$, and otherwise $\bar{x}_2$ being any vector in $\mathbb{R}^{n-1}$ satisfying $\|\bar{x}_2\|=1$. If $x_2\neq 0$, the factorization is unique. By the spectral factorization, we readily have

$$[x]^+_{\mathcal{K}^n}=\max\big\{0,\lambda_1(x)\big\}u_x^{(1)}+\max\big\{0,\lambda_2(x)\big\}u_x^{(2)}. \tag{8}$$

Next we review some properties of the Fischer–Burmeister (FB) merit function studied by [5] for the second-order complementarity problem. The merit function is defined as

$$\psi_{\mathrm{FB}}(x,y):=\frac{1}{2}\big\|\phi_{\mathrm{FB}}(x,y)\big\|^2 \tag{9}$$

where $\phi_{\mathrm{FB}}:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ is the vector-valued FB function given by

$$\phi_{\mathrm{FB}}(x,y)=\big(x^2+y^2\big)^{1/2}-(x+y).$$

Lemma 2.1. *(See [5].) Let $\psi_{\mathrm{FB}}:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be given by (9). Then, for any $x,y\in\mathbb{R}^n$,*

(a) *$\psi_{\mathrm{FB}}(x,y)\geqslant 0$, and $\psi_{\mathrm{FB}}(x,y)=0\Longleftrightarrow x\in\mathcal{K}^n,\ y\in\mathcal{K}^n,\ \langle x,y\rangle=0$.*
(b) *ψ_{FB} is continuously differentiable. Moreover, $\nabla_x\psi_{\mathrm{FB}}(0,0)=\nabla_y\psi_{\mathrm{FB}}(0,0)=0$, and*

$$\nabla_x\psi_{\mathrm{FB}}(x,y)=\big(L_xL^{-1}_{(x^2+y^2)^{1/2}}-I\big)\phi_{\mathrm{FB}}(x,y),$$
$$\nabla_y\psi_{\mathrm{FB}}(x,y)=\big(L_yL^{-1}_{(x^2+y^2)^{1/2}}-I\big)\phi_{\mathrm{FB}}(x,y)$$

if $x^2+y^2\in\operatorname{int}(\mathcal{K}^n)$, and if $x^2+y^2\notin\operatorname{int}(\mathcal{K}^n)$ and $(x,y)\neq(0,0)$,

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

(c) $\langle \nabla_x \psi_{\rm FB}(x,y), \nabla_y \psi_{\rm FB}(x,y)\rangle \geqslant 0$, *and the equality holds if and only if* $\psi_{\rm FB}(x,y)=0$.
(d) $\langle x, \nabla_x \psi_{\rm FB}(x,y)\rangle + \langle y, \nabla_y \psi_{\rm FB}(x,y)\rangle = 2\psi_{\rm FB}(x,y)$.
(e) $\psi_{\rm FB}(x,y)=0 \Longleftrightarrow \nabla\psi_{\rm FB}(x,y)=0 \Longleftrightarrow \nabla_x\psi_{\rm FB}(x,y)=0 \Longleftrightarrow \nabla_y\psi_{\rm FB}(x,y)=0$.

To close this section, we present some special examples of the extended SOCLCP (1).

2.1. The generalized SOCLCP

Given matrices $A, B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{m\times p}$, and a vector $b \in \mathbb{R}^m$, the generalized SOCLCP is to find $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}^p$ such that

$$Ax + By + Cz = b, \quad z \in \widehat{\mathcal{K}},\ x \in \mathcal{K},\ y \in \mathcal{K},\ \langle x, y\rangle = 0 \tag{10}$$

where $\widehat{\mathcal{K}} \subset \mathbb{R}^p$ is the Cartesian product of SOCs. Clearly, when $\mathcal{K}$ and $\widehat{\mathcal{K}}$ degenerate into $\mathbb{R}^n_+$ and $\mathbb{R}^p_+$, respectively, (10) becomes the generalized LCP of [28]. Letting

$$M = [A\ C], \qquad N = -[B\ C], \qquad x' = \begin{pmatrix} x \\ 0 \end{pmatrix}, \qquad y' = \begin{pmatrix} y \\ z \end{pmatrix},$$

we can rewrite (10) as (1) with $P = 0$, $E = I$, $r = b$ and $\mathcal{E} = \{0\}$, i.e.,

$$Mx' - Ny' \in \Omega, \quad x' \in \mathcal{K} \times \widehat{\mathcal{K}},\ y' \in \mathcal{K} \times \widehat{\mathcal{K}},\ \langle x', y'\rangle = 0.$$

2.2. The horizontal SOCLCP

Given matrices $A, B \in \mathbb{R}^{m\times n}$ and a vector $b \in \mathbb{R}^m$, the *horizontal* SOCLCP is to find $x, y \in \mathbb{R}^n$ such that

$$Ax - By = b, \quad x \in \mathcal{K},\ y \in \mathcal{K},\ \langle x, y\rangle = 0. \tag{11}$$

If $m = n$ and $A = I$, this reduces to the standard SOCLCP [14]; whereas if $\mathcal{K} = \mathbb{R}^n_+$, it reduces to the horizontal linear complementarity problem [10,27]. Obviously, (11) is an extended SOCLCP with $M = A$, $N = B$, $P = 0$ and $E = I$, $r = b$, $\mathcal{E} = \{0\}$.

2.3. The vertical SOCLCP

Given matrices $A, B \in \mathbb{R}^{n\times p}$ and vectors $c, d \in \mathbb{R}^n$, the *vertical* SOCLCP is to find $z \in \mathbb{R}^p$ such that

$$Az + c \in \mathcal{K}, \qquad Bz + d \in \mathcal{K}, \qquad \langle Az + c, Bz + d\rangle = 0. \tag{12}$$

When $\mathcal{K} = \mathbb{R}^n_+$, this reduces to the vertical linear complementarity problem [12]. Letting $x = Az + c$ and $y = Bz + d$, we can reformulate the vertical SOCLCP as (1) with

$$M = \begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad N = \begin{bmatrix} 0 \\ -I \end{bmatrix}, \qquad P = -\begin{bmatrix} A \\ B \end{bmatrix}, \qquad E = I, \qquad r = \begin{bmatrix} c \\ d \end{bmatrix}, \qquad \mathcal{E} = \{0\}.$$

2.4. The mixed SOCLCP

Given $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times p}$ and $D \in \mathbb{R}^{n\times n}$, and vectors $c \in \mathbb{R}^m$, $d \in \mathbb{R}^n$, the mixed SOCLCP is to find $z \in \mathbb{R}^p$, $y \in \mathbb{R}^n$ such that

$$Az + By + c = 0, \quad y \in \mathcal{K}, \quad Cz + Dy + d \in \mathcal{K}, \quad \langle y, Cz + Dy + d\rangle = 0. \tag{13}$$

When $\mathcal{K} = \mathbb{R}^n_+$, this reduces to the mixed linear complementarity problem [11]. Letting $x = Cz + Dy + d$, this problem can be rewritten as (1) with

$$M = \begin{bmatrix} 0 \\ -I \end{bmatrix}, \qquad N = -\begin{bmatrix} B \\ D \end{bmatrix}, \qquad P = \begin{bmatrix} A \\ C \end{bmatrix}, \qquad E = I, \qquad r = \begin{bmatrix} -c \\ -d \end{bmatrix}, \qquad \mathcal{E} = \{0\}.$$

3. Constrained and unconstrained reformulations

In this section, we give some simple SOC constrained reformulations and unconstrained reformulations for the SOCLCP (1), and then under some mild assumptions establish the equivalence between the stationary points of these problems and the solutions of (1). In the sequel, we write $x = (x_1, \ldots, x_q)$, $y = (y_1, \ldots, y_q) \in \mathbb{R}^n$ with $x_i, y_i \in \mathbb{R}^{n_i}$, and let

$$\nabla_x \psi(x,y) := \big(\nabla_{x_1}\psi(x,y), \nabla_{x_2}\psi(x,y), \ldots, \nabla_{x_q}\psi(x,y)\big),$$
$$\nabla_y \psi(x,y) := \big(\nabla_{y_1}\psi(x,y), \nabla_{y_2}\psi(x,y), \ldots, \nabla_{y_q}\psi(x,y)\big).$$

From [15, p. 121], $x \in \mathcal{E}$ if and only if (iff for short) $[x]^+_{\mathcal{E}^\circ} = 0$. This means that

$$Mx - Ny + Pz \in \Omega \quad \Longleftrightarrow \quad \big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ} = 0, \tag{14}$$

and finding (x, y, z) so that $Mx - Ny + Pz \in \Omega$ is equivalent to seeking a global minimum of $\|[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ}\|^2$ with zero optimal value. If $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies

(i) $\psi(x, y) \geqslant 0$ for all $x, y \in \mathcal{K}$, and $\psi(x, y) = 0 \Longleftrightarrow \langle x, y\rangle = 0$,

then the SOCLCP (1) can be reformulated as an SOC constrained problem

$$\begin{aligned} &\min \frac{1}{2}\big\|\big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\|^2 + \gamma\psi(x, y) \\ &\text{s.t.} \quad x \in \mathcal{K}, \ \ y \in \mathcal{K} \end{aligned} \tag{15}$$

where $\gamma > 0$ is a constant to balance the feasibility and the complementarity in (1). If ψ is a merit function for the complementarity condition involved in (1), i.e.,

(I) $\psi(x, y) \geqslant 0$ for all $x, y \in \mathbb{R}^n$, and $\psi(x, y) = 0 \Longleftrightarrow x \in \mathcal{K},\ y \in \mathcal{K},\ \langle x, y\rangle = 0$,

then (1) can be reformulated as an unconstrained minimization problem

$$\min_{(x,y,z)\in\mathbb{R}^{2n+p}} \frac{1}{2}\big\|\big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\|^2 + \gamma\psi(x, y). \tag{16}$$

There are many functions satisfying the requirement in (i). A direct choice for ψ is

$$\psi(x, y) = \sum_{i=1}^{q} h\big(\langle x_i, y_i\rangle\big),$$

with $h : \mathbb{R} \to \mathbb{R}$ satisfying $h(t) \geqslant 0$ for all $t \geqslant 0$ and $h(t) = 0$ iff $t = 0$; for example,

$$\text{the linear function} \quad \psi_1(x, y) := \langle x, y\rangle = \sum_{i=1}^{q} x_i^T y_i,$$

$$\text{the quadratic function} \quad \psi_2(x, y) := \frac{1}{2}\sum_{i=1}^{q}\big(x_i^T y_i\big)^2,$$

$$\text{the entropy function} \quad \psi_3(x, y) := \sum_{i=1}^{q}\big[\big(1 + x_i^T y_i\big)\ln\big(1 + x_i^T y_i\big) - x_i^T y_i\big],$$

$$\text{the logarithmic function} \quad \psi_4(x, y) := \sum_{i=1}^{q}\ln\big[1 + \big(x_i^T y_i\big)^2\big].$$

Noting that $x, y \in \mathcal{K}$ and $\langle x, y\rangle = 0$ iff $x, y \in \mathcal{K}$ and $x \circ y = 0$, another choice for ψ is

$$\psi(x, y) = g(x \circ y)$$

with $g : \mathbb{R}^n \to \mathbb{R}_+$ satisfying $g(u) = 0$ iff $u = 0$. For example, taking $g(u) = \frac{1}{2}\|u\|^2$,

$$\psi_5(x, y) = \frac{1}{2}\|x \circ y\|^2 = \sum_{i=1}^{q}\frac{1}{2}\|x_i \circ y_i\|^2.$$

This function will become ψ_2 used in [24] when $\mathcal{K}$ degenerates to $\mathbb{R}^n_+$ and the Jordan product becomes the componentwise product of vectors. In addition, we may choose ψ as a merit function for the complementarity condition in (1), such as the FB merit function

$$\psi_{\mathrm{FB}}(x, y) = \sum_{i=1}^{q}\psi_{\mathrm{FB}}(x_i, y_i)$$

where $\psi_{\mathrm{FB}}(x_i, y_i)$ is defined as in (9), the regularized FB merit function

$$\psi_{\mathrm{YF}}(x, y) := \sum_{i=1}^{q}\bigg[\frac{1}{2}\big(\max\{0, x_i^T y_i\}\big)^2 + \psi_{\mathrm{FB}}(x_i, y_i)\bigg], \tag{17}$$

and the implicit Lagrangian function defined by

$$\psi_\alpha(x,y) := \langle x,y\rangle + \frac{1}{2\alpha}\left(\left\|[x-\alpha y]_{\mathcal{K}}^+\right\|^2 - \|x\|^2 + \left\|[y-\alpha x]_{\mathcal{K}}^+\right\|^2 - \|y\|^2\right) \quad \alpha > 1.$$

To establish the equivalence between the stationary point set of (15) and the solution set of (1), we require that ψ also possesses (some of) the following favorable properties:

(ii) ψ is continuously differentiable everywhere in $\mathcal{K}\times\mathcal{K}$.
(iii) For all $x, y\in\mathcal{K}$, $\langle\nabla_{x_i}\psi(x,y), \nabla_{y_i}\psi(x,y)\rangle \geqslant 0$, $i=1,2,\ldots,q$.
(iv) $\langle\nabla_x\psi(x,y), \nabla_y\psi(x,y)\rangle = 0$ and $x,y\in\mathcal{K} \Longrightarrow \psi(x,y)=0$.
(v) If there exist vectors $w,s\in\mathcal{K}$ such that $\langle w,x\rangle = 0$ and $\langle s,y\rangle = 0$ for $x,y\in\mathcal{K}$, then $\langle w_i, \nabla_{y_i}\psi(x,y)\rangle = 0$ and $\langle s_i, \nabla_{x_i}\psi(x,y)\rangle = 0$ for all $i=1,2,\ldots,q$.
(vi) $\psi(x,y)=0$ and $x,y\in\mathcal{K} \Longrightarrow \nabla_x\psi(x,y) = \nabla_y\psi(x,y) = 0$.
(vii) $\langle x, \nabla_x\psi(x,y)\rangle + \langle y, \nabla_y\psi(x,y)\rangle \geqslant 0$ for all $x,y\in\mathcal{K}$, and the equality holds iff $\langle x,y\rangle = 0$ or $x\circ y = 0$.

Proposition 3.1. *The function ψ_1 satisfies all the properties except* (vi)*, the functions ψ_2–ψ_5 satisfy all the properties* (i)–(vii)*, ψ_{FB} and ψ_{YF} satisfy all the properties except* (v)*, and ψ_α satisfies all the properties except* (iv) *and* (v)*.*

Proof. It is easily seen that ψ_1 satisfies all the properties except (vi). For ψ_2 and ψ_4, it is easy to check that they satisfy (i)–(vii). For ψ_3, using the properties of $(1+t)\ln(1+t)-t$ and noting that $\nabla_{x_i}\psi_3(x,y) = \ln(1+x_i^T y_i)y_i$ and $\nabla_{y_i}\psi_3(x,y) = \ln(1+x_i^T y_i)x_i$, we can verify that ψ_3 satisfies (i)–(vi). Also, from $\ln(1+t)\leqslant t$ for all $t\geqslant 0$, it follows that

$$\langle x, \nabla_x\psi_3(x,y)\rangle + \langle y, \nabla_y\psi_3(x,y)\rangle \geqslant 2\psi_3(x,y),$$

which together with (i) implies that ψ_3 satisfies (vii). Clearly, ψ_5 satisfies the properties (i), (ii) and (vi), and it suffices to check that it satisfies (iii)–(v) and (vii). Since

$$\langle x, \nabla_x\psi_5(x,y)\rangle + \langle y, \nabla_y\psi_5(x,y)\rangle = \langle x, y\circ(x\circ y)\rangle + \langle y, x\circ(x\circ y)\rangle = 2\|x\circ y\|^2,$$

it follows that ψ_5 satisfies (vii). If there exist $w=(w_1,\ldots,w_q)$, $s=(s_1,\ldots,s_q)\in\mathcal{K}$ such that $\langle w,x\rangle=0$ and $\langle s,y\rangle = 0$ for $x,y\in\mathcal{K}$, then we must have $w_i\circ x_i = 0$ and $s_i\circ y_i = 0$ for all $i=1,2,\ldots,q$. Consequently, ψ_5 satisfies (v) since for all $i=1,2,\ldots,q$,

$$\langle w_i, \nabla_{y_i}\psi_5(x,y)\rangle = \langle w_i, x_i\circ(x_i\circ y_i)\rangle = \langle w_i\circ x_i, x_i\circ y_i\rangle = 0,$$
$$\langle s_i, \nabla_{x_i}\psi_5(x,y)\rangle = \langle s_i, y_i\circ(x_i\circ y_i)\rangle = \langle s_i\circ y_i, x_i\circ y_i\rangle = 0.$$

In addition, for all $x,y\in\mathcal{K}$, we can compute that for all $i=1,2,\ldots,q$,

$$\langle \nabla_{x_i}\psi_5(x,y), \nabla_{y_i}\psi_5(x,y)\rangle = \langle x_i\circ(x_i\circ y_i), y_i\circ(x_i\circ y_i)\rangle.$$

From Lemma 1 of Appendix A, it follows that ψ_5 satisfies the property (iii), and moreover,

$$\langle \nabla_{x_i}\psi_5(x,y), \nabla_{y_i}\psi_5(x,y)\rangle = 0 \quad\Longrightarrow\quad x_i^T y_i = 0 \quad \text{for } i=1,2,\ldots,q.$$

This implies that ψ_5 also satisfies the property (iv).

From Lemma 2.1, ψ_{FB} satisfies all the properties except (v). By the expression of ψ_{YF} and Lemma 2.1, it is easy to check that ψ_{YF} satisfies all the properties except (v). From Lemma 4.2, Theorem 4.3, and Prop. 4.4(1) of [17], ψ_α satisfies (i)–(iii) and (vi). In addition, using the gradient formulas of ψ_α, $\langle x, \nabla_x\psi_\alpha(x,y)\rangle + \langle y, \nabla_y\psi_\alpha(x,y)\rangle = 2\psi_\alpha(x,y)$. This together with (i) implies that ψ_α also satisfies the property (vii). $\square$

Proposition 3.1 shows that ψ_2–ψ_5 share with the same favorable properties. But, it should be noted that their growth in the cone $\mathcal{K}\times\mathcal{K}$ is different. It is easy to verify that

$$\psi_5(x,y) \gg \psi_2(x,y) \gg \psi_3(x,y) \gg \psi_1(x,y) \gg \psi_4(x,y) \quad \forall x,y\in\mathcal{K}\times\mathcal{K}, \tag{18}$$

where $\psi\gg\phi$ means ψ has faster growth than ϕ.

Theorem 3.1. *Suppose that one of the following conditions is satisfied*:

(a) *ψ satisfies* (i)–(vi)*, and M and N are X-row-block-sufficient w.r.t. Ω.*
(b) *ψ satisfies* (i)–(vi)*, and MN^T is copositive on $(0^+\Omega)^\circ$.*
(c) *ψ satisfies* (i)–(iii) *and* (v)–(vii)*, and MN^T is strictly copositive on $(0^+\Omega)^\circ$.*
(d) *ψ satisfies* (i)–(ii) *and* (vi)–(vii)*, and $M^T v\in\mathcal{K}$ and $-N^T v\in\mathcal{K}$ for all $v\in(0^+\Omega)^\circ$.*
(e) *ψ satisfies* (i) *and* (vii)*, and $0\in\Omega$.*

Then, (x,y,z) is a stationary point of (15) *iff it solves* (1)*.*

Proof. Suppose that (x, y, z) is a solution of (1). Then, (x, y, z) is feasible for (15) and the corresponding objective value is zero. This means that (x, y, z) is a solution of (15). Notice that the constraints of (15) are convex and satisfy the Slater constraint qualification, and hence (x, y, z) is a stationary point of (15).

Let (x, y, z) be a stationary point of (15). Then, there exist $w, s \in \mathbb{R}^n$ such that

$$M^T v + \gamma \nabla_x \psi(x, y) - w = 0, \tag{19}$$

$$-N^T v + \gamma \nabla_y \psi(x, y) - s = 0, \qquad P^T v = 0, \tag{20}$$

$$w \in \mathcal{K}, \quad x \in \mathcal{K}, \quad \langle w, x\rangle = 0, \tag{21}$$

$$s \in \mathcal{K}, \quad y \in \mathcal{K}, \quad \langle s, y\rangle = 0 \tag{22}$$

where $v := E^T[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ}$. Since $[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ} \in \mathcal{E}^\circ$, we have $v \in (0^+\Omega)^\circ$. We next show that (x, y, z) solves (1) under the given assumptions.

(a) From Eqs. (19) and (20), it follows that for all $i = 1, 2, \dots, q$,

$$\big(M^T v\big)_i = w_i - \gamma \nabla_{x_i} \psi(x, y), \qquad \big(N^T v\big)_i = -s_i + \gamma \nabla_{y_i} \psi(x, y),$$

where $w_i, s_i \in \mathbb{R}^{n_i}$ are ith subvectors of w and s, respectively. By this, we have

$$\begin{aligned}
\big\langle \big(M^T v\big)_i, \big(N^T v\big)_i\big\rangle &= -\gamma^2\big\langle \nabla_{x_i}\psi(x, y), \nabla_{y_i}\psi(x, y)\big\rangle - \langle w_i, s_i\rangle \\
&\quad + \gamma\big\langle s_i, \nabla_{x_i}\psi(x, y)\big\rangle + \gamma\big\langle w_i, \nabla_{y_i}\psi(x, y)\big\rangle \\
&= -\gamma^2\big\langle \nabla_{x_i}\psi(x, y), \nabla_{y_i}\psi(x, y)\big\rangle - \langle w_i, s_i\rangle \\
&\leqslant -\gamma^2\big\langle \nabla_{x_i}\psi(x, y), \nabla_{y_i}\psi(x, y)\big\rangle \\
&\leqslant 0 \quad \text{for all } i = 1, 2, \dots, q
\end{aligned}$$

where the second equality is due to (v), and the first inequality is using $w_i, s_i \in \mathcal{K}^{n_i}$, and the last inequality is from (iii). Since $v \in (0^+\Omega)^\circ$, and M and N are X-row-block-sufficient w.r.t. Ω, we have $\langle (M^T v)_i, (N^T v)_i\rangle = 0$ for all $i = 1, 2, \dots, q$. Combining with the last inequality yields $\langle \nabla_{x_i}\psi(x, y), \nabla_{y_i}\psi(x, y)\rangle = 0$ for all $i = 1, 2, \dots, q$, which means that $\langle \nabla_x\psi(x, y), \nabla_y\psi(x, y)\rangle = 0$. This along with (iv) and $x, y \in \mathcal{K}$ gives $\psi(x, y) = 0$, and consequently, $\langle x, y\rangle = 0$. Since $x \in \mathcal{K}$ and $y \in \mathcal{K}$, to prove that (x, y, z) is a solution of (1), it remains to prove that $Mx - Ny + Pz \in \Omega$. Since $\psi(x, y) = 0$, from (vi) we get $\nabla_x\psi(x, y) = \nabla_y\psi(x, y) = 0$. Thus, (19)–(22) reduce to the KKT conditions of

$$\begin{aligned}
&\min \ \frac{1}{2}\big\|\big[E\big(Mx' - Ny' + Pz'\big) - r\big]^+_{\mathcal{E}^\circ}\big\|^2 \\
&\text{s.t.} \ \ x' \in \mathcal{K}, \ y' \in \mathcal{K}.
\end{aligned}$$

Since this is a convex program, (x, y, z) is its solution. Noting that the convex program has a zero optimal value by the assumption (4), we have $Mx - Ny + Pz \in \Omega$ from (14).

(b) The result follows by part (a) and the fact that copositiveness of MN^T implies X-row-block-sufficiency w.r.t. Ω of M and N.

(c) Using Eqs. (19)–(20) and the properties (iii) and (v), we have

$$\begin{aligned}
\big\langle MN^T v, v\big\rangle &= -\gamma^2\big\langle \nabla_x\psi(x, y), \nabla_y\psi(x, y)\big\rangle - \langle w, s\rangle \\
&\quad - \gamma\big\langle w, \nabla_y\psi(x, y)\big\rangle - \gamma\big\langle s, \nabla_x\psi(x, y)\big\rangle \\
&= -\gamma^2\big\langle \nabla_x\psi(x, y), \nabla_y\psi(x, y)\big\rangle - \langle w, s\rangle \\
&\leqslant 0.
\end{aligned}$$

This is equivalent to saying that $-\langle MN^T v, v\rangle \geqslant 0$. Since $v \in (0^+\Omega)^\circ$ and MN^T is strictly copositive on $(0^+\Omega)^\circ$, we then have $v = 0$. Combining it with (19)–(22) yields that

$$0 = \langle w, x\rangle + \langle s, y\rangle = \gamma\big\langle x, \nabla_x\psi(x, y)\big\rangle + \gamma\big\langle y, \nabla_y\psi(x, y)\big\rangle,$$

and $\langle x, y\rangle = 0$ holds by (vii). The proof of $Mx - Ny + Pz \in \Omega$ is the same as before by (vi).

(d) From Eqs. (19)–(22), it follows that

$$\begin{aligned}
0 &= \langle x, w\rangle + \langle y, s\rangle \\
&= \big\langle x, \gamma\nabla_x\psi(x, y)\big\rangle + \big\langle y, \gamma\nabla_y\psi(x, y)\big\rangle + \big\langle x, M^T v\big\rangle - \big\langle y, N^T v\big\rangle \\
&\geqslant \big\langle x, M^T v\big\rangle - \big\langle y, N^T v\big\rangle \geqslant 0
\end{aligned} \tag{23}$$

where the first inequality is due to (vii) and the second is from the assumption that $M^T v \in \mathcal{K}$ and $-N^T v \in \mathcal{K}$. This means that $\langle x, \nabla_x \psi(x, y)\rangle + \langle y, \nabla_y \psi(x, y)\rangle = 0$ since each term on the left hand side of the first inequality is nonnegative. From (vii), it then follows that $\langle x, y\rangle = 0$. By (vi), the proof that $Mx - Ny + Pz \in \Omega$ follows as before.

(e) Using the equality (23) and noting that $P^T v = 0$, we have

$$\begin{aligned}
0 &= \bigl\langle x, \gamma \nabla_x \psi(x, y)\bigr\rangle + \bigl\langle y, \gamma \nabla_y \psi(x, y)\bigr\rangle + \langle Mx - Ny + Pz, v\rangle \\
&= \bigl\langle x, \gamma \nabla_x \psi(x, y)\bigr\rangle + \bigl\langle y, \gamma \nabla_y \psi(x, y)\bigr\rangle \\
&\quad + \bigl\langle E(Mx - Ny + Pz) - r + r,\ \bigl[E(Mx - Ny + Pz) - r\bigr]^+_{\mathcal{E}^\circ}\bigr\rangle \\
&= \gamma\bigl[\bigl\langle x, \nabla_x \psi(x, y)\bigr\rangle + \bigl\langle y, \nabla_y \psi(x, y)\bigr\rangle\bigr] + \bigl\|\bigl[E(Mx - Ny + Pz) - r\bigr]^+_{\mathcal{E}^\circ}\bigr\|^2 \\
&\quad + \bigl\langle r, \bigl[E(Mx - Ny + Pz) - r\bigr]^+_{\mathcal{E}^\circ}\bigr\rangle.
\end{aligned}$$

Notice that $-r \in \mathcal{E}$ since $0 \in \Omega$, and $[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ} \in \mathcal{E}^\circ$. Hence, we have $\langle -r, [E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ}\rangle \leqslant 0$ by the definition of the polar cone. This shows that the last term on the right hand side of the last equality is nonnegative, whereas the first term is also nonnegative by (vii). Thus, from the last equality it follows that $\langle x, y\rangle = 0$ and $[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ} = 0$. Together with (14), $x, y \in \mathcal{K}$, and the property (i), it follows that (x, y, z) solves the SOCLCP (1). □

By Proposition 3.1 and Theorem 3.1, when ψ is chosen as one of ψ_2–ψ_5, the stationary point set of (15) coincides with the solution set of (1) under any of the assumptions of Theorem 3.1; when $\psi = \psi_{\mathrm{FB}}$, ψ_{YF} or ψ_α, the two sets are equivalent only under the assumptions (d) and (e); whereas when $\psi = \psi_1$, the equivalence holds only under the assumption (e). This means that the constrained reformulations associated with ψ_2–ψ_5 are superior to those with other functions.

Next we establish the equivalence between the stationary points of (16) and the solutions of (1). We require that ψ satisfies (some of) the following properties except (I):

(II) ψ is continuously differentiable in $\mathbb{R}^n \times \mathbb{R}^n$.
(III) For all $x, y \in \mathcal{K}$, $\langle \nabla_{x_i} \psi(x, y), \nabla_{y_i} \psi(x, y)\rangle \geqslant 0$, $i = 1, 2, \ldots, q$.
(IV) $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y)\rangle = 0 \Longrightarrow \psi(x, y) = 0$.
(V) $\psi(x, y) = 0 \Longrightarrow \nabla_x \psi(x, y) = 0, \nabla_y \psi(x, y) = 0$.
(VI) $\langle x, \nabla_x \psi(x, y)\rangle + \langle y, \nabla_y \psi(x, y)\rangle = c\psi(x, y)$, where $c > 0$.

Theorem 3.2. *Suppose that one of the following conditions is satisfied*:

(a) ψ *satisfies* (I)–(V), *and* M *and* N *are* X*-row-block-sufficient with respect to* Ω.
(b) ψ *satisfies* (I)–(V), *and* MN^T *is copositive on* $(0^+\Omega)^\circ$.
(c) ψ *satisfies* (I)–(III) *and* (VI), *and* MN^T *is strictly copositive on* $(0^+\Omega)^\circ$.
(d) ψ *satisfies* (I) *and* (VI), *and* $0 \in \Omega$.

Then (x, y, z) *is a stationary point of* (16) *if and only if it solves* (1).

Proof. Suppose that (x, y, z) is a solution of (1). Then, (x, y, z) is a solution of (16) since the objective value of (16) at this point is zero. Consequently, (x, y, z) is a stationary point of (16). Next, let (x, y, z) be a stationary point of (16). Then,

$$M^T v + \gamma \nabla_x \psi(x, y) = 0, \qquad -N^T v + \gamma \nabla_y \psi(x, y) = 0, \qquad P^T v = 0, \tag{24}$$

where $v := E^T[E(Mx - Ny + Pz) - r]^+_{\mathcal{E}^\circ}$. From the first two equalities, we have

$$\bigl(M^T v\bigr)_i = -\gamma \nabla_{x_i} \psi(x, y), \qquad \bigl(N^T v\bigr)_i = \gamma \nabla_{y_i} \psi(x, y), \quad i = 1, 2, \ldots, q. \tag{25}$$

(a) Using Eq. (25) and the property (III), it follows that

$$\bigl\langle \bigl(M^T v\bigr)_i, \bigl(N^T v\bigr)_i\bigr\rangle = -\gamma^2\bigl\langle \nabla_{x_i} \psi(x, y), \nabla_{y_i} \psi(x, y)\bigr\rangle \leqslant 0, \quad i = 1, 2, \ldots, q.$$

This, by the given assumption, implies that $\langle (M^T v)_i, (N^T v)_i\rangle = 0$ for all $i = 1, 2, \ldots, q$. Consequently, $\langle \nabla_x \psi(x, y), \nabla_y \psi(x, y)\rangle = 0$. Combining with the properties (IV) and (I), we have $x \in \mathcal{K}$, $y \in \mathcal{K}$ and $\langle x, y\rangle = 0$. To prove that (x, y, z) is a solution of (1), it remains to argue that $Mx - Ny + Pz \in \Omega$. Since $\psi(x, y) = 0$, by (V), (24) reduces to

$$M^T v = 0, \qquad N^T v = 0, \qquad P^T v = 0.$$

This means that (x, y, z) is a stationary point of the following convex program

$$\min_{(x,y,z) \in \mathbb{R}^{2n+p}} \frac{1}{2}\bigl\|\bigl[E(Mx - Ny + Pz) - r\bigr]^+_{\mathcal{E}^\circ}\bigr\|^2,$$

which has a zero optimal value by the assumption (4). Thus, $Mx - Ny + Pz \in \Omega$ by (14).

(b) The result is direct by part (a).

(c) From (25) and (III), $\langle MN^T v, v\rangle = -\gamma^2\langle \nabla_x\psi(x,y), \nabla_y\psi(x,y)\rangle \leqslant 0$. This, by the strict copositivity of MN^T on $(0^+\Omega)^\circ$, implies $v = 0$. Substituting $v = 0$ into (24), we have $\nabla_x\psi(x,y) = \nabla_y\psi(x,y) = 0$. From (VI) and (I), we get $x \in \mathcal{K}$, $y \in \mathcal{K}$ and $\langle x, y\rangle = 0$. Using the same arguments as before leads to $Mx - Ny + Pz \in \Omega$.

(d) From (24), clearly, $\langle Mx - Ny + Pz, v\rangle = -\gamma\langle x, \nabla_x\psi(x,y)\rangle - \gamma\langle y, \nabla_x\psi(x,y)\rangle$. Hence,

$$\begin{aligned}
0 &= \gamma\big\langle x, \nabla_x\psi(x,y)\big\rangle + \gamma\big\langle y, \nabla_x\psi(x,y)\big\rangle + \langle Mx - Ny + Pz, v\rangle \\
&= \gamma\big\langle x, \nabla_x\psi(x,y)\big\rangle + \gamma\big\langle y, \nabla_x\psi(x,y)\big\rangle \\
&\quad + \big\langle E(Mx - Ny + Pz) - r + r,\ \big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\rangle \\
&= \gamma\big\langle x, \nabla_x\psi(x,y)\big\rangle + \gamma\big\langle y, \nabla_x\psi(x,y)\big\rangle \\
&\quad + \big\|\big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\|^2 + \big\langle r, \big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\rangle.
\end{aligned}$$

By (I) and (VI), using the same arguments as in Theorem 3.1(e) yields the result. □

From the proof of Proposition 3.1, ψ_{FB} and ψ_{YF} satisfy (I)–(VI), whereas ψ_α satisfies all the properties except (IV). Thus, by Theorem 3.2, the stationary point set of (16) with $\psi = \psi_{\mathrm{FB}}$ and ψ_{YF} coincides with the solution set of (1) under any of the assumptions of Theorem 3.2; whereas for $\psi = \psi_\alpha$ their equivalence holds only under (c) or (d).

For convenience, from now to the end of Section 4, we assume $\mathcal{K} = \mathcal{K}^n$, and all analysis can be carried over to the case where $\mathcal{K}$ has the structure as in (3). Next, we study two important properties of the objective function of (15) with ψ being one of ψ_2–ψ_5, ψ_{FB}, ψ_{YF} and ψ_α. Let $w := (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$, and write

$$f(w) := \frac{1}{2}\big\|\big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\|^2 + \gamma\psi(x,y). \tag{26}$$

In addition, we denote the feasible set of (15) by $S := \{w = (x, y, z) \mid x \succcurlyeq_{\mathcal{K}} 0,\ y \succcurlyeq_{\mathcal{K}} 0\}$.

Proposition 3.2. *Let f be given by* (26)*. Then, f is smooth, and its gradient function $\nabla f(w)$ is Lipschitz continuous on any bounded set $S_1 \subseteq S$ when ψ is chosen as one of ψ_2–ψ_5 and ψ_{YF}, and when ψ is chosen as ψ_{FB} or ψ_α, it is globally Lipschitz continuous.*

Proof. The smoothness of f is due to the smoothness of $\|[\cdot]^+_{\mathcal{E}^\circ}\|^2$ and property (ii). Let

$$f_1(w) := \frac{1}{2}\big\|\big[E(Mx - Ny + Pz) - r\big]^+_{\mathcal{E}^\circ}\big\|^2.$$

From the nonexpansive property of the projection operator [3, Prop. 2.1.3], $\nabla f_1(w)$ and $\nabla\psi_\alpha$ are Lipschitz continuous. In addition, $\nabla\psi_{\mathrm{FB}}$ is Lipschitz continuous by [6]. Thus, $\nabla f(w)$ with $\psi = \psi_{\mathrm{FB}}$ or ψ_α is Lipschitz continuous. To prove that $\nabla f(w)$ with ψ being one of ψ_2–ψ_5 and ψ_{YF} is Lipschitz continuous on the bounded set S_1, it suffices to show that $\nabla_x\psi$ is Lipschitz continuous on S_1 due to the Lipschitz continuity of ∇f_1 and the symmetry between $\nabla_x\psi(x,y)$ and $\nabla_y\psi(x,y)$. For any $(x,y), (a,b) \in S_1$, we have

$$\begin{aligned}
\big\|\nabla_x\psi_2(x,y) - \nabla_x\psi_2(a,b)\big\| &= \big\|\max\{0, x^Ty\}y - \max\{0, a^Tb\}b\big\| \\
&\leqslant \max\{0, x^Ty\}\|y - b\| + \big|\max\{0, x^Ty\} - \max\{0, a^Tb\}\big|\|b\| \\
&\leqslant \max\{0, x^Ty\}\|y - b\| + \big|x^Ty - a^Tb\big|\|b\| \\
&\leqslant \max\{0, x^Ty\}\|y - b\| + \big(\|x\|\|y - b\| + \|x - a\|\|b\|\big)\|b\| \\
&\leqslant C_2\big(\|x - a\| + \|y - b\|\big)
\end{aligned}$$

where $C_2 > 0$ is a constant, and the last inequality is due to the boundedness of S_1;

$$\begin{aligned}
\big\|\nabla_x\psi_4(x,y) - \nabla_x\psi_4(a,b)\big\| &= \left\|\frac{2x^Ty}{1 + (x^Ty)^2}y - \frac{2a^Tb}{1 + (a^Tb)^2}b\right\| \\
&\leqslant \frac{2x^Ty\|y - b\|}{1 + (x^Ty)^2} + \left\|\frac{2x^Ty}{1 + (x^Ty)^2} - \frac{2a^Tb}{1 + (a^Tb)^2}\right\|\|b\| \\
&\leqslant \|y - b\| + \frac{2|x^Ty - a^Tb|\cdot|1 - x^Tya^Tb|}{[1 + (x^Ty)^2][1 + (a^Tb)^2]}\|b\| \\
&\leqslant \|y - b\| + 6\big|x^Ty - a^Tb\big|\|b\| \\
&\leqslant \|y - b\| + 6\big(\|x\|\|y - b\| + \|x - a\|\|b\|\big)\|b\| \\
&\leqslant C_4\big(\|x - a\| + \|y - b\|\big), \quad C_4 > 0 \text{ is a constant;}
\end{aligned}$$

$$
\begin{aligned}
\left\|\nabla_x\psi_3(x,y)-\nabla_x\psi_3(a,b)\right\| &= \left\|\ln\left(1+x^Ty\right)y-\ln\left(1+a^Tb\right)b\right\| \\
&\leqslant \ln\left(1+x^Ty\right)\|y-b\|+\left|\ln\left(1+x^Ty\right)-\ln\left(1+a^Tb\right)\right|\|b\| \\
&\leqslant \ln\left(1+x^Ty\right)\|y-b\|+\frac{|x^Ty-a^Tb|}{1+a^Tb}\|b\| \\
&\leqslant \ln\left(1+x^Ty\right)\|y-b\|+\left(\|x\|\|y-b\|+\|x-a\|\|b\|\right)\|b\| \\
&\leqslant C_3\left(\|x-a\|+\|y-b\|\right)
\end{aligned}
$$

where $C_3>0$ is a constant, and the second inequality uses $\ln t\leqslant t-1$ $(t>0)$;

$$
\begin{aligned}
\left\|\nabla_x\psi_5(x,y)-\nabla_y\psi_5(a,b)\right\| &= \left\|y\circ(y\circ x)-b\circ(a\circ b)\right\| \\
&\leqslant 3\|x\circ y\|\|y-b\|+3\|y\circ x-a\circ b\|\|b\| \\
&\leqslant 3\|x\circ y\|\|y-b\|+9\left(\|y-b\|\|x\|+\|x-a\|\|b\|\right)\|b\| \\
&\leqslant C_5\left(\|x-a\|+\|y-b\|\right)
\end{aligned}
$$

where $C_5>0$ is a constant, and the first inequality and the second one use $\|x\circ y\|\leqslant 3\|x\|\|y\|$. The above inequalities show that $\nabla_x\psi(x,y)$ with ψ being one of ψ_2–ψ_5 and ψ_{YF} is Lipschitz continuous on the bounded set S_1. $\square$

The following proposition provides a condition to guarantee that the level sets of f

$$L_f(c):=\left\{w=(x,y,z)\in S\;\middle|\; f(w)\leqslant c\right\}$$

are bounded for all $c\geqslant 0$. The property is very important since it ensures that the feasible descent sequence of f always has an accumulation point.

Proposition 3.3. *The level sets $L_f(c)$ are bounded for all $c\geqslant 0$, if for any $w=(x,y,z)\in S$ satisfying $\|w\|=1$ and $\langle x,y\rangle=0$, there holds that $Mx-Ny+Pz\notin 0^+\Omega$.*

Proof. Assume on the contrary there exists an unbounded sequence $\{w^k=(x^k,y^k,z^k)\}\subset L_f(c)$ for some $c\geqslant 0$. Then $f(w^k)\leqslant c$ for all k. Since $\{w^k\}$ is unbounded, there exists a subsequence $\{w^k\}_{k\in K_1}$ satisfying $\|w^k\|\to+\infty$. By passing to a subsequence if necessary, we assume that $\{w^k/\|w^k\|\}_{k\in K_1}\to w^*=(x^*,y^*,z^*)$. Then, $w^*\in S$ and $\|w^*\|=1$. If $\langle x^*,y^*\rangle=0$, by the given assumption $Mx^*-Ny^*+Pz^*\notin 0^+\Omega$, and so

$$\left\|\left[E\left(Mx^*-Ny^*+Pz^*\right)\right]^+_{\mathcal{E}^\circ}\right\|\neq 0.$$

Noting that $[\beta x]^+_{\mathcal{E}^\circ}=\beta[x]^+_{\mathcal{E}^\circ}$ for any $\beta\geqslant 0$, the last equation is equivalent to

$$\lim_{k\to\infty,\,k\in K_1}\frac{\|[E(Mx^k-Ny^k+Pz^k)-r]^+_{\mathcal{E}^\circ}\|}{\|w^k\|}\neq 0,$$

which implies $\lim_{k\to\infty,\,k\in K_1} f_1(w^k)=\infty$. Combining with the nonnegativity of ψ, we have $\lim_{k\to\infty} f(w^k)=+\infty$, a contradiction to the fact $f(w^k)\leqslant c$ for all k.

If $\langle x^*,y^*\rangle\neq 0$, then $\langle x^*,y^*\rangle>0$, which implies that $\lim_{k\to\infty,\,k\in K_1}\langle x^k,y^k\rangle=+\infty$. Since $\max\{0,t\}^2$, $\ln(1+t^2)$, and $(1+t)\ln(1+t)-t$ are increasing on $[0,+\infty)$, we have $\lim_{k\to\infty,\,k\in K_1}\psi(x^k,y^k)=+\infty$ when ψ is chosen as one of ψ_2–ψ_4 and ψ_{YF}. Since $\|x\circ y\|\geqslant|\langle x,y\rangle|$, we also have $\lim_{k\to\infty,\,k\in K_1}\psi_5(x^k,y^k)=+\infty$. In addition, $\langle x^*,y^*\rangle\neq 0$ implies that $x^*\circ y^*\neq 0$ since $x^*\in\mathcal{K}$ and $y^*\in\mathcal{K}$. Therefore, $(x^k/\|w^k\|)\circ(y^k/\|w^k\|)\nrightarrow 0$. Using Lemma 5.2(b) of [21] and Proposition 4.2(ii) of [22], $\lim_{k\to\infty,\,k\in K_1}\psi_{\mathrm{FB}}(x^k,y^k)=+\infty$ and $\lim_{k\to\infty,\,k\in K_1}\psi_\alpha(x^k,y^k)=+\infty$. Thus, we prove $\lim_{k\to\infty,\,k\in K_1}\psi(x^k,y^k)\to\infty$, and hence $\lim_{k\to\infty,\,k\in K_1} f(w^k)=+\infty$ when ψ is chosen as one of ψ_2–ψ_5 or the functions ψ_{FB}, ψ_{YF} and ψ_α. This gives a contradiction to the fact that $f(w^k)\leqslant c$ for all k. $\square$

4. The solution of SOC constrained problem

In this section, we develop a proximal gradient descent method for solving the equivalent SOC constrained reformulation problem (15). This method will generate a direction $d=(d_x,d_y,d_z)\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^p$ at a given point $w=(x,y,z)\in S$ by solving

$$
\begin{aligned}
&\min\ \left\langle\nabla f(w),d'\right\rangle+\frac{1}{2}\rho\left\|d'\right\|^2 \\
&\text{s.t.}\quad x+d'_x\succcurlyeq_{\mathcal{K}} 0 \\
&\qquad\ \ y+d'_y\succcurlyeq_{\mathcal{K}} 0
\end{aligned}
\tag{27}
$$

where $d' = (d'_x, d'_y, d'_z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$, and $\rho > 0$ is an arbitrary constant, and then seek a stepsize $\alpha > 0$ along the direction d. The idea of this method is not new. In fact, the subproblem (27) can be regarded as a special case of the ones used by [8,26] for minimizing the sum of a smooth function and a separable convex function since, using the indicator function $\delta(\cdot \mid \mathcal{K})$, it can be rewritten as

$$\min \langle \nabla f(w), d' \rangle + \frac{1}{2}\rho \|d'\|^2 + P(w + d') \tag{28}$$

where $P(w + d') := \delta(x + d'_x \mid \mathcal{K}) + \delta(y + d'_y \mid \mathcal{K})$ is convex, and separable when $q > 1$.

Before describing our algorithm, we present two technical lemmas, where Lemma 4.1 implies that the nonzero solution of (27) at $w \in S$ must be a feasible descent direction of f at this point, and Lemma 4.2 provides an alternative characterization for the stationary points of (15) which will be used as the termination condition of our algorithm.

Lemma 4.1. *Let $w = (x, y, z)$ be any given point in S and $d = (d_x, d_y, d_z)$ be the solution of* (27)*. Then, for any $\alpha \in [0, 1]$, $w + \alpha d \in S$ and*

$$\langle \nabla f(w), d \rangle \leqslant -\frac{1}{2}\rho \|d\|^2. \tag{29}$$

Proof. Since $w = (x, y, z) \in S$, we have $x \succcurlyeq_{\mathcal{K}} 0$ and $y \succcurlyeq_{\mathcal{K}} 0$. Notice that $x + d_x \succcurlyeq_{\mathcal{K}} 0$ and $y + d_y \succcurlyeq_{\mathcal{K}} 0$. Hence, for any $\alpha \in [0, 1]$, $x + \alpha d_x = (1 - \alpha)x + \alpha(x + d_x) \in \mathcal{K}$ and $y + \alpha d_y = (1 - \alpha)y + \alpha(y + d_y) \in \mathcal{K}$. This means that $w + \alpha d \in S$. Noting that $d' = 0$ is a feasible solution of (27) since $x, y \succcurlyeq_{\mathcal{K}} 0$, whereas d is the optimal solution, we have $\langle \nabla f(w), d \rangle + \frac{1}{2}\rho \|d\|^2 \leqslant 0$, which implies the desired result in (29). $\square$

Lemma 4.2. *Let $w = (x, y, z)$ be any given point in S and $d = (d_x, d_y, d_z)$ be the solution of* (27)*. Then, w is a stationary point of* (15) *iff $d = 0$.*

Proof. Suppose that w is a stationary point of (15). Then there exist ξ and η such that

$$\begin{aligned} &\nabla_x f(w) - \xi = 0, \qquad \nabla_y f(w) - \eta = 0, \qquad \nabla_z f(w) = 0, \\ &\langle x, \xi \rangle = 0, \qquad \langle y, \eta \rangle = 0, \qquad \xi \succcurlyeq_{\mathcal{K}} 0, \qquad \eta \succcurlyeq_{\mathcal{K}} 0. \end{aligned} \tag{30}$$

The last two equations imply that for any $w' = (x', y', z')$ with $x', y' \in \mathcal{K}$,

$$\langle \nabla f(w), w' - w \rangle = \langle \xi, x' - x \rangle + \langle \eta, y' - y \rangle = \langle \xi, x' \rangle + \langle \eta, y' \rangle \geqslant 0. \tag{31}$$

If $d \neq 0$, then from (29) d is a feasible descent direction of f at the feasible point w, which contradicts (31). We next consider the sufficiency. Since (27) is a convex program whose constraints satisfy the Slater constraint qualification, there exist ζ and ν such that

$$\begin{aligned} &\nabla_x f(w) + d_x - \zeta = 0, \qquad \nabla_y f(w) + d_y - \nu = 0, \qquad \nabla_z f(w) = 0, \\ &\langle x + d_x, \zeta \rangle = 0, \qquad \langle y + d_y, \nu \rangle = 0, \quad \zeta \succcurlyeq_{\mathcal{K}} 0, \ \nu \succcurlyeq_{\mathcal{K}} 0. \end{aligned}$$

When $d = (d_x, d_y, d_z) = 0$, these conditions are the same as those in (30), i.e., the KKT conditions of the problem (15). Consequently, w is a stationary point of (15). $\square$

Algorithm 4.1 *(The PGD method).*

Step 0. Choose $w^0 = (x^0, y^0, z^0) \in S$, $\beta \in (0, 1)$, $\sigma \in (0, 1)$, and $\epsilon > 0$. Set $k := 0$.
Step 1. Choose $\rho_k > 0$, and solve (27) with $w = w^k$ to get its solution $d^k = (d^k_x, d^k_y, d^k_z)$.
Step 2. If $\|d^k\| \leqslant \epsilon$, then stop; and otherwise go to Step 3.
Step 3. Let α^k be the largest element of $\{1, \beta, \beta^2, \ldots\}$ satisfying

$$f(w^k + \alpha^k d^k) \leqslant f(w^k) + \sigma \alpha^k \langle \nabla f(w^k), d^k \rangle. \tag{32}$$

Step 4. Set $w^{k+1} := w^k + \alpha^k d^k$ and $k := k + 1$, and go to Step 1.

From Lemmas 4.1 and 4.2, we see that Algorithm 4.1 is well defined and generates a feasible sequence $\{w^k\} = \{(x^k, y^k, z^k)\}$ such that the objective value sequence $\{f(w^k)\}$ is monotonically decreasing. In each iteration, the main work of Algorithm 4.1 is to solve the subproblem (27) with $w = w^k$, which is equivalent to making one Euclidean projection on the closed convex cone S since the subproblem can be rewritten as

$$\min_{w \in S} \langle \nabla f(w^k), w \rangle + \frac{1}{2}\rho_k \|w - w^k\|^2. \tag{33}$$

By a simple computation, $d^k = (d_x^k, d_y^k, d_z^k)$ has the following explicit expression

$$d_x^k = \left[x^k - \rho_k^{-1}\nabla_x f\left(w^k\right)\right]_{\mathcal{K}}^{+} - x^k,$$
$$d_y^k = \left[y^k - \rho_k^{-1}\nabla_y f\left(w^k\right)\right]_{\mathcal{K}}^{+} - y^k,$$
$$d_z^k = -\rho_k^{-1}\nabla_z f\left(w^k\right).$$

It is worthwhile to mention that the solution of the more general subproblem

$$\min \left\langle \nabla f\left(w^k\right), d'\right\rangle + \frac{1}{2}\left(d'\right)^T H^k d'$$
$$\text{s.t.}\quad x^k + d'_x \succcurlyeq_{\mathcal{K}} 0, \quad y^k + d'_y \succcurlyeq_{\mathcal{K}} 0$$

as used in [26] with H^k being a $(2n+p)\times(2n+p)$ symmetric positive definite matrix to approximate the Hessian of f at w^k, is equivalent to the solution of a SOCLCP instead of the scaled projection onto $\mathcal{K}$. This is different from the nonnegative orthant cone case.

Now we concentrate on the convergence of Algorithm 4.1. We first establish the global convergence under the assumption that the parameter ρ_k is uniformly bounded.

Theorem 4.1. *Let $w^k = \{(x^k, y^k, z^k)\}$ be a sequence generated by Algorithm 4.1 with $0 < \rho_1 \leqslant \rho_k \leqslant \rho_2$ for all k. Then each cluster point of $\{w^k\}$ is a stationary point of (15).*

Proof. Let $\{w^k\}_K$ be a subsequence of $\{w^k\}$ converging to some $\hat{w}$. Then $\hat{w} \in S$ since $\{w^k\} \subseteq S$ and S is closed. Also, since f is continuous, we have $\lim_{k\to\infty, k\in K} f(w^k) = f(\hat{w})$. This means that the sequence $\{f(w^k)\}_K$ is convergent and $\{f(w^{k+1}) - f(w^k)\}_K \to 0$.

Case 1: $\liminf_{k\in K, k\to\infty} \alpha^k > 0$. In this case, by Step 3 of Algorithm 4.1 and Lemma 4.1,

$$f\left(w^{k+1}\right) - f\left(w^k\right) \leqslant \sigma\alpha^k\left\langle\nabla f\left(w^k\right), d^k\right\rangle \leqslant -\frac{1}{2}\sigma\rho_1\alpha^k\left\|d^k\right\|^2 \quad \forall k.$$

Taking the limit $k \to \infty$ with $k \in K$ on the both sides and using $\{f(w^{k+1}) - f(w^k)\}_K \to 0$, we get $\{\alpha^k\|d^k\|^2\}_{k\in K} \to 0$, which implies $\{d^k\}_{k\in K} \to \hat{d} = 0$ since $\liminf_{k\in K, k\to\infty} \alpha^k > 0$. On the other hand, for any $d' = (d'_x, d'_y, d'_z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$ satisfying $\hat{x} + d'_x \succcurlyeq_{\mathcal{K}} 0$ and $\hat{y} + d'_y \succcurlyeq_{\mathcal{K}} 0$, we have $x^k + d'_x \succcurlyeq_{\mathcal{K}} 0$, $y^k + d'_y \succcurlyeq_{\mathcal{K}} 0$ for sufficiently large k, and moreover,

$$\left\langle\nabla f\left(w^k\right), d^k\right\rangle + \frac{1}{2}\rho_k\left\|d^k\right\|^2 \leqslant \left\langle\nabla f\left(w^k\right), d'\right\rangle + \frac{1}{2}\rho_k\left\|d'\right\|^2 \leqslant \left\langle\nabla f\left(w^k\right), d'\right\rangle + \frac{1}{2}\rho_2\left\|d'\right\|^2.$$

Taking the limit $k \to \infty$ with $k \in K$ yields $0 \leqslant \langle\nabla f(\hat{w}), d'\rangle + \frac{1}{2}\rho_2\|d'\|^2$. This means that $\hat{d} = 0$ is a solution of the subproblem (27) with $w = \hat{w}$ and $\rho = \rho_2$. From Lemma 4.2, it then follows that $\hat{w}$ is a stationary point of (15).

Case 2: $\lim_{k\in K, k\to\infty} \alpha^k = 0$. Suppose that $\{d^k\}_K \nrightarrow 0$. By passing to a subsequence if necessary, we can assume that for some $\delta > 0$, $\|d^k\| \geqslant \delta$ for all $k \in K$. Since α^k is chosen by the Armijo rule, we have $f(w^k + (\alpha^k/\beta)d^k) - f(w^k) > \sigma(\alpha^k/\beta)\langle\nabla f(w^k), d^k\rangle$ for any $k \in K$. Dividing both sides by $\|d^k\|$, this inequality becomes

$$\frac{f(w^k + \hat{\alpha}^k\hat{d}^k) - f(w^k)}{\hat{\alpha}^k} > \sigma\left\langle\nabla f\left(w^k\right), \hat{d}^k\right\rangle \quad \forall k \in K, \tag{34}$$

where $\hat{d}^k = d^k/\|d^k\|$ and $\hat{\alpha}^k = \alpha^k\|d^k\|/\beta$. Since $\{\hat{d}^k\}_K$ is bounded, we assume $\{\hat{d}^k\}_K \to \hat{d}$ (by passing to a subsequence if necessary). From Case 1, we know that $\{\alpha^k\|d^k\|^2\}_{k\in K} \to 0$, which, by $\|d^k\| \geqslant \delta$ for all $k \in K$, implies $\{\hat{\alpha}^k\}_K \to 0$. Taking the limit $k \to \infty$ with $k \in K$ in the inequality (34), we obtain

$$(1-\sigma)\left\langle\nabla f(\hat{w}), \hat{d}\right\rangle \geqslant 0. \tag{35}$$

On the other hand, using Lemma 4.1 and noting that $\|d^k\| \geqslant \delta$ for all $k \in K$, we have

$$\left\langle\nabla f\left(w^k\right), \hat{d}^k\right\rangle \leqslant -\frac{1}{2}\rho_k\left\|d^k\right\| \leqslant -\frac{1}{2}\rho_1\delta \quad \forall k \in K.$$

Taking the limit $k \to \infty$ with $k \in K$ in the inequality yields $\langle\nabla f(\hat{w}), \hat{d}\rangle \leqslant -\frac{1}{2}\rho_1\delta$, which clearly contradicts (35). So, $\{d^k\}_K \to 0$. Using the same arguments as in Case 1, we have that $\hat{w}$ is a stationary point of (15). $\square$

Notice that the sequence $\{w^k = (x^k, y^k, z^k)\}$ generated by Algorithm 4.1 is contained in the level set $L_f(f(w^0))$. Therefore, $\{w^k\}$ always has a cluster point, provided that the matrices M, N and P satisfy the assumption of Proposition 3.3.

Next we concentrate on analyzing the linear rate of convergence of Algorithm 4.1. The following technical lemma will be used in the subsequent analysis.

Lemma 4.3. *For any $w = (x, y, z) \in S$, let $d = (d_x, d_y, d_z)$ be the solution of* (27) *with $\rho > 0$. Then, for any $\bar{w} = (\bar{x}, \bar{y}, \bar{z}) \in S$ and $w' = w + \alpha d$ with $\alpha \in [0, 1]$, we have*

$$\big\langle \nabla f(w), w' - \bar{w}\big\rangle \leqslant -\big\langle \nabla f(w), d\big\rangle + \rho\|d\|\|\bar{w} - w\|. \tag{36}$$

Proof. By the definition of d, it is not hard to verify that d is also a solution of

$$\begin{aligned} &\min_u \big\langle \nabla f(w) + \rho d, u\big\rangle \\ &\text{s.t.}\;\; x + u_x \succcurlyeq_{\mathcal{K}} 0 \\ &\qquad y + u_y \succcurlyeq_{\mathcal{K}} 0 \end{aligned} \tag{37}$$

where $u = (u_x, u_y, u_z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$. Since $\bar{w} - w$ is a feasible solution of (37), we get $\langle \nabla f(w) + \rho d, d\rangle \leqslant \langle \nabla f(w) + \rho d, \bar{w} - w\rangle$, which implies that

$$\big\langle \nabla f(w), w - \bar{w}\big\rangle \leqslant -\big\langle \nabla f(w) + \rho d, d\big\rangle + \rho\langle d, \bar{w} - w\rangle.$$

Using this inequality and the definition of w', we have

$$\begin{aligned} \big\langle \nabla f(w), w' - \bar{w}\big\rangle &= \big\langle \nabla f(w), w - \bar{w}\big\rangle + \alpha\big\langle \nabla f(w), d\big\rangle \\ &\leqslant -(1-\alpha)\big\langle \nabla f(w), d\big\rangle - \rho\|d\|^2 + \rho\langle d, \bar{w} - w\rangle \\ &\leqslant -\big\langle \nabla f(w), d\big\rangle + \rho\|d\|\|\bar{w} - w\|, \end{aligned}$$

where the last inequality is since $\langle \nabla f(w), d\rangle \leqslant 0$ and $\alpha \in [0, 1]$. □

Similar to [26], we also need a local Lipschitzian error bound assumption on the distance to the set of stationary point of (15), denoted by $\widehat{S}$. Such assumption was often used to establish the rate of convergence for iterative methods, such as gradient projection and coordinate descent methods of constrained smooth optimization; see [18,19].

Assumption 1. $\widehat{S} \neq \emptyset$ and, there exist $\tau > 0$ and $\epsilon > 0$ such that $\operatorname{dist}(w, \widehat{S}) \leqslant \tau\|d\|$ whenever $w \in L_f(f(w^0))$ and $\|d\| \leqslant \epsilon$, where d is the solution of (27) with $\rho > 0$.

Theorem 4.2. *Let $\{w^k\}$ and $\{d^k\}$ be generated by Algorithm* 4.1 *with $0 < \rho_1 \leqslant \rho_k \leqslant \rho_2$ for all k. If M, N and P satisfy the assumption of Proposition* 3.3 *and Assumption* 1 *holds, then $\{f(w^k)\}$ converges at least Q-linearly and $\{w^k\}$ converges at least R-linearly.*

Proof. The proof is similar to that of [26, Theorem 2], but the arguments here are much simpler. First, f is Lipschitz continuous over any bounded set S_2 due to its smoothness. This implies that there exists a scalar $\delta > 0$ such that

$$\|w - w'\| > \delta \quad \text{whenever } w, w' \in S_2,\; f(w) \neq f(w'). \tag{38}$$

By Proposition 3.3, $L_f(f(w^0))$ is bounded. Since $\{w^k\} \subset L_f(f(w^0))$ by the construction of Algorithm 4.1, $\{w^k\}$ is bounded. By passing to a subsequence if necessary, we can assume that $\{w^k\}$ converges to some $\hat{w}$. Using the same arguments as in Theorem 4.1, we then have $\{d^k\} \to 0$. From Assumption 1 and $\{w^k\} \subset L_f(f(w^0))$, it follows that

$$\|w^k - \bar{w}^k\| \leqslant \tau\|d^k\| \quad \forall k \geqslant \text{some } \bar{k}, \tag{39}$$

where $\tau > 0$ and $\bar{w}^k \in \widehat{S}$ satisfies $\|w^k - \bar{w}^k\| = \operatorname{dist}(w^k, \widehat{S})$. Noting that $\{w^k\}$ is bounded, the inequality (39) implies that $\{\bar{w}^k\}_{k\geqslant\bar{k}}$ is bounded. By (38), there exist an index $\hat{k} \geqslant \bar{k}$ and a scalar $\bar{\upsilon}$ such that $f(\bar{w}^k) = \bar{\upsilon}$ for all $k \geqslant \hat{k}$.

Now, fixing any $k \geqslant \hat{k}$ and using the Mean Value Theorem, it follows that

$$\begin{aligned} f(w^{k+1}) - \bar{\upsilon} &= \nabla f(\tilde{w}^k)^T(w^{k+1} - \bar{w}^k) \\ &= \big(\nabla f(\tilde{w}^k) - \nabla f(w^k)\big)^T(w^{k+1} - \bar{w}^k) + \nabla f(w^k)^T(w^{k+1} - \bar{w}^k) \end{aligned}$$

where $\tilde{w}^k$ is a point lying on the segment joining w^{k+1} with $\bar{w}^k$. Since $\{w^k\}$ and $\{\bar{w}^k\}_{k\geqslant\bar{k}}$ are bounded, the sequence $\{\tilde{w}^k\}_{k\geqslant\hat{k}}$ is also bounded. Using Proposition 3.2, we have

$$\|\nabla f(\tilde{w}^k) - \nabla f(w^k)\| \leqslant L\|\tilde{w}^k - w^k\| \leqslant L\big[\|\bar{w}^k - w^{k+1}\| + \|\bar{w}^k - w^k\|\big] \quad \forall k \geqslant \hat{k},$$

for some constant $L > 0$. Combining the last two equations, we obtain

$$
\begin{aligned}
f(w^{k+1}) - \bar{\upsilon} &\leqslant L\|w^{k+1} - \bar{w}^k\|^2 + L\|w^{k+1} - \bar{w}^k\|\|w^k - \bar{w}^k\| + \nabla f(w^k)^T(w^{k+1} - \bar{w}^k) \\
&\leqslant L\|w^{k+1} - \bar{w}^k\|^2 + L\|w^{k+1} - \bar{w}^k\|\|w^k - \bar{w}^k\| - \langle \nabla f(w^k), d^k\rangle + \rho_2\|d^k\|\|\bar{w}^k - w^k\| \\
&\leqslant L(\alpha^k\|d^k\| + \|w^k - \bar{w}^k\|)^2 + L(\alpha^k\|d^k\|\|w^k - \bar{w}^k\| + \|w^k - \bar{w}^k\|^2) \\
&\quad - \langle \nabla f(w^k), d^k\rangle + \rho_2\|d^k\|\|\bar{w}^k - w^k\| \\
&\leqslant C_1\|d^k\|^2 - \langle \nabla f(w^k), d^k\rangle \quad \text{for all } k \geqslant \hat{k},
\end{aligned}
\tag{40}
$$

where the second step uses Lemma 4.3 with $\bar{w} = \bar{w}^k$ and $w' = w^{k+1}$, the third step is from $\|w^{k+1} - \bar{w}^k\| \leqslant \|w^{k+1} - w^k\| + \|w^k - \bar{w}^k\|$ and $\|w^{k+1} - w^k\| = \alpha^k d^k$, the fourth step is due to (39), and C_1 is a constant depending on L, τ, ρ_2 only. In addition, using (29),

$$\|d^k\|^2 \leqslant -2\rho_1^{-1}\langle \nabla f(w^k), d^k\rangle.$$

This means that the right hand side of (40) is bounded above by

$$-C_2\langle \nabla f(w^k), d^k\rangle \quad \text{for all } k \geqslant \hat{k},$$

where $C_2 > 0$ is depending on L, τ, ρ_1, ρ_2 only. Then, by Step 3 of Algorithm 4.1,

$$f(w^{k+1}) - \bar{\upsilon} \leqslant C_3\big(f(w^k) - f(w^{k+1})\big) \quad \forall k \geqslant \hat{k}, \tag{41}$$

where $C_3 = C_2/(\sigma\alpha^k)$. On the other hand, for all $k \geqslant \hat{k}$, we have

$$
\begin{aligned}
\bar{\upsilon} - f(w^k) &= \nabla f(\hat{w}^k)^T(\bar{w}^k - w^k) \\
&\leqslant \nabla f(\hat{w}^k)^T(\bar{w}^k - w^k) + \nabla f(\bar{w}^k)^T(w^k - \bar{w}^k) \\
&= \big(\nabla f(\bar{w}^k) - \nabla f(\hat{w}^k)\big)^T(w^k - \bar{w}^k) \\
&\leqslant L\|w^k - \bar{w}^k\|^2
\end{aligned}
\tag{42}
$$

which the first step uses the Mean Value Theorem with $\hat{w}^k$ being a point on the segment joining w^k with $\bar{w}^k$, the second step follows since $\nabla f(\bar{w}^k)^T(w^k - \bar{w}^k) \geqslant 0$, implied by $\bar{w}^k \in \widehat{S}$, and the last step is due to the Lipschitz continuity of ∇f in any bounded set. Combining with $\{w^k - \bar{w}^k\} \to 0$, the inequality (42) implies

$$\liminf_{k\to\infty} f(w^k) \geqslant \bar{\upsilon}. \tag{43}$$

From Eqs. (41) and (43), it then follows that

$$0 \leqslant f(w^{k+1}) - \bar{\upsilon} \leqslant \frac{C_3}{1 + C_3}\big(f(w^k) - \bar{\upsilon}\big) \quad \forall k \geqslant \hat{k}.$$

This shows that $\{f(w^k)\}$ converges at least Q-linearly.

Using Step 3 of Algorithm 4.1, the inequality (29) and $w^{k+1} - w^k = \alpha^k d^k$, we have

$$f(w^{k+1}) - f(w^k) \leqslant -\frac{1}{2}\sigma\alpha^k\rho_k\|d^k\|^2 \leqslant -\frac{1}{2}\sigma\rho_1(\alpha^k)^{-1}\|w^{k+1} - w^k\|^2$$

which in turn implies that $\|w^{k+1} - w^k\| \leqslant \sqrt{2\alpha^k(f(w^k) - f(w^{k+1}))(\sigma\rho_1)^{-1}}$ for any k. Since $\{f(w^k) - f(w^{k+1})\}$ converges to 0 at least Q-linearly and $\sup_k \alpha^k \leqslant 1$, this implies that the sequence $\{w^k\}$ converges at least R-linearly. □

Assumption 1 seems to be a little stronger, and now we do not know what properties of f can guarantee it to hold, except the strong convexity of f over the set S, which is clearly not satisfied by f with ψ being any of functions introduced in Section 3. In addition, since the feasible set of (15) is nonpolyhedral, the results obtained in [18,19] cannot be utilized. We will leave this problem as our future research topic.

5. Numerical experience

In this section, we test the numerical performance of Algorithm 4.1 with solving the SOC constrained reformulations for two classes of special extended SOCLCPs, where $P = 0$ and $\mathcal{E}$ is the nonnegative orthant cone $\mathbb{R}^l_+$ or the Cartesian product of second-order cones $\widetilde{\mathcal{K}} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_p}$ with $n_1, \ldots, n_p \geqslant 1$ and $n_1 + \cdots + n_p = l$.

All test problems were generated randomly. The matrices M, N and E were generated by "sprandn" of Matlab with approximately 0.01 mn normally distributed nonzero entries. The vector r was obtained from $r = E(Mu - Nv)$, where $u = (u_1, \ldots, u_q) \in \mathcal{K}$ and $v = (v_1, \ldots, v_q) \in \mathcal{K}$ with $u_i, v_i \in \mathcal{K}^{n_i}$ were generated in the following way: let the elements of u_i

Table 1
Numerical results of Algorithm 4.1 for the SOCLCP with $\mathcal{E} = \widetilde{\mathcal{K}}$.

No.	ψ_2				ψ_4				ψ_3			
	It	**NF**	**Obj**	**Cpu**	**It**	**NF**	**Obj**	**Cpu**	**It**	**NF**	**Obj**	**Cpu**
1	893	1298	1.22e–5	302.5	68	175	7.12e–7	48.0	622	836	1.07e–5	225.3
2	681	1013	1.08e–5	245.4	36	135	2.22e–11	36.6	763	1082	1.15e–5	297.1
3	650	933	1.05e–5	216.3	55	169	3.12e–7	46.9	201	380	6.15e–6	107.3
4	436	582	8.83e–6	27.9	33	154	0	42.3	369	515	8.16e–6	141.5
5	714	1095	1.07e–5	258.2	64	163	5.06e–7	45.6	592	816	9.78e–6	225.4
6	673	947	1.06e–5	223.7	74	157	1.05e–6	43.9	90	219	3.82e–6	61.7
7	609	853	9.89e–6	201.1	61	144	4.57e–7	38.8	69	210	7.66e–7	58.9
8	683	1027	1.06e–5	245.1	44	125	1.02e–7	34.7	562	745	9.39e–6	200.6
9	695	1012	1.10e–5	232.0	36	129	8.97e–9	35.8	600	804	1.01e–5	232.0
10	729	1066	1.07e–5	251.3	37	158	3.11e–10	43.5	208	426	5.99e–6	116.2

Table 2
Numerical results of Algorithm 4.1 and L-BFGS for $\mathcal{E} = \widetilde{\mathcal{K}}$.

No.	Algorithm 4.1					L-BFGS				
	Iter	**NF**	**Obj**	**Gap**	**Cpu**	**Iter**	**NF**	**Obj**	**Gap**	**Cpu**
1	90	231	2.95e–6	5.43e–6	288.7	99	127	9.34e–10	1.55e–6	152.5
2	80	209	1.51e–6	3.89e–6	264.8	92	120	1.38e–9	8.92e–6	138.0
3	85	209	2.10e–6	4.58e–6	255.2	105	131	2.37e–9	7.03e–6	157.9
4	79	222	1.52e–6	3.90e–6	275.3	136	183	9.50e–9	9.96e–6	206.7
5	93	259	3.75e–6	6.12e–6	323.1	86	120	2.99e–9	2.53e–6	129.5
6	49	195	1.99e–10	4.54e–8	251.6	121	160	1.36e–9	8.35e–6	187.0
7	77	206	1.22e–6	3.49e–6	262.0	73	98	2.20e–9	5.83e–6	112.3
8	324	492	6.64e–6	8.13e–6	625.3	109	141	7.33e–9	2.06e–6	163.3
9	90	235	3.01e–6	5.49e–6	297.4	142	172	7.20e–9	5.27e–6	208.7
10	87	227	2.42e–6	4.92e–6	287.1	105	130	9.68e–10	1.06e–5	159.3

be chosen randomly from a normal distribution with mean -1 and variance 4, and then set $u_{i1} = \|u_{i2}\|$; let the elements of v_i be chosen randomly from a normal distribution with mean 0 and variance 1, and then set $v_{i1} = \|v_{i2}\|$, where u_{i1} and v_{i1} are the first elements of u_i and v_i, respectively. Such way guarantees that the assumption in (4) holds. We chose $n_1 = \cdots = n_q$ and $n_1 = \cdots = n_p$ to construct $\mathcal{K}$ and $\widetilde{\mathcal{K}}$, respectively.

All experiments were done with a PC of 2.8 GHz CPU and 512 MB memory, and the computer codes were all written in Matlab 7.0. During our tests, we adopted $\gamma = 10^5$ for the reformulation problems, and chose the following parameters for Algorithm 4.1:

$$\epsilon = 10^{-5}, \qquad \beta = 0.5, \qquad \sigma = 0.1, \qquad \rho_{k+1} = \min\{1.05\rho_k, 10^3\} \quad \text{with } \rho_0 = 10.$$

The starting point (x^0, y^0) of Algorithm 4.1 was chosen as $x^0 = (x_1^0, \ldots, x_q^0)$ and $y^0 = (y_1^0, \ldots, y_q^0)$ with $x_i^0 = (10, \omega_i/\|\omega_i\|)$ and $y_i^0 = (10, \eta_i/\|\eta_i\|)$, where $\omega_i, \eta_i \in \mathbb{R}^{n_i-1}$ for all $i = 1, 2, \ldots, m$ were generated by Matlab's **rand.m**.

We first applied Algorithm 4.1 for solving a group of problems generated as above for $\mathcal{E} = \widetilde{\mathcal{K}}$ with $m = 2000$, $n = 2000$, $l = 1500$, $q = 50$, and $p = 50$ to test the performance of the SOC constrained reformulation problem (15) with different ψ. The numerical results corresponding to ψ_2, ψ_3 and ψ_4 are summarized in Table 1, where **It** records the number of iteration required to satisfy the termination condition, **NF** indicates the number of function evaluations of $f(w)$, **Obj** means the objective value of (15) at the final iteration, and **Cpu** denotes the CPU time in second for solving each test problem. For the function ψ_5, we cannot obtain the favorable results. From the results in Table 1 and the growth relation between ψ_2–ψ_5, we may conclude that the reformulation problem (15) has better performance if it is derived from the function $\psi(x, y)$ with slower growth over $\mathcal{K} \times \mathcal{K}$.

We also compared the numerical performance of Algorithm 4.1 for solving (15) based on ψ_4 with the limited-memory BFGS method [4] for solving (16) based on ψ_{FB}. Among others, the L-BFGS method utilized 5 limited-memory vector-update and the Armijo line search rule same as (32) except $\sigma = 10^{-4}$. We used the two methods to solve two groups of test problems with $\mathcal{E} = \widetilde{\mathcal{K}}$ and $\mathcal{E} = \mathbb{R}^l_+$. The test problems for $\mathcal{E} = \widetilde{\mathcal{K}}$ have the size of $m = 3000$, $n = 3000$, $l = 2500$, $q = 50$ and $p = 50$; whereas the test problems for $\mathcal{E} = \mathbb{R}^l_+$ have the size of $m = 2000$, $n = 2000$, $l = 1500$, $q = 50$. Algorithm 4.1 and the L-BFGS method started from the same initial point generated as above. When $\mathcal{E} = \mathbb{R}^l_+$, the parameter ρ_k was modified by $\rho_{k+1} = \min\{1.01\rho_k, 10^3\}$ with $\rho_0 = 10$. Numerical results are listed in Tables 2–3, in which **It**, **NF**, **Obj** and **Cpu** have the same meaning as in Table 1, and **Gap** means the value of $\max\{0, x^T y\}$ at the final iteration.

From Tables 2–3, we see that for most of test problems Algorithm 4.1 requires fewer iterations than the L-BFGS method, and moreover, the solutions generated have smaller **Gap**. However, Algorithm 4.1 needs more function evaluations than the L-BFGS method, and consequently a little more CPU time. Consider that Algorithm 4.1 exploits first-order information of the objective function, whereas the L-BFGS method exploits approximate second-order information of the objective function.

Table 3
Numerical results of Algorithm 4.1 and L-BFGS for $\mathcal{E} = \mathbb{R}^l_+$.

No.	Algorithm 4.1					L-BFGS				
	Iter	**NF**	**Obj**	**Gap**	**Cpu**	**Iter**	**NF**	**Obj**	**Gap**	**Cpu**
1	95	445	0	1.17e–14	115.4	214	255	1.26e–9	7.26e–6	67.9
2	88	423	1.11e–10	3.47e–10	108.3	205	255	4.97e–9	9.63e–6	69.5
3	99	376	2.81e–6	5.30e–7	98.8	272	325	8.62e–9	2.67e–6	91.5
4	133	531	4.56e–8	6.75e–7	135.2	229	270	9.80e–9	1.02e–6	76.2
5	128	596	1.41e–8	3.75e–7	154.8	221	262	6.92e–10	8.69e–7	72.3
6	97	478	2.89e–10	5.38e–8	121.5	193	236	5.13e–9	5.26e–6	65.3
7	112	552	4.44e–11	1.92e–8	142.0	271	231	8.15e–9	2.03e–6	88.2
8	124	439	4.91e–8	7.01e–7	114.8	220	269	4.93e–9	2.14e–6	72.9
9	84	408	2.22e–10	4.64e–8	105.4	307	365	1.48e–9	3.40e–6	99.3
10	155	563	7.35e–8	8.57e–7	142.1	230	273	1.19e–9	1.20e–6	74.4

This shows that Algorithm 4.1 is effective if a suitable ρ_k is selected. Notice that Algorithm 4.1 is parallelizable when $q > 1$, and therefore it is easily modified to solve the large-scale problems.

We want to point out that solving (15) with Algorithm 4.1 and solving (16) with the L-BFGS method will yield different solutions if the solution of (1) is not unique. The solution yielded by the former always lies in $\mathcal{K} \times \mathcal{K}$, whereas the one given by the latter satisfies the property approximately. In addition, we find that for the problems where $\mathcal{E} = \{0\}$, applying Algorithm 4.1 for (15) with $\psi = \psi_4$ and applying L-BFGS method for (16) cannot yield favorable numerical results, although many special SOCLCPs are reformulated as (1) with $\mathcal{E} = \{0\}$. This means that the penalized reformulations proposed are unsuitable for this class of SOCLCPs.

6. Conclusions

We proposed some SOC constrained reformulations and unconstrained reformulations for the extended SOCLCP (1), and established the equivalence between the stationary points of these optimization problems and the solutions of (1) under mild conditions. We also developed a proximal gradient descent method for solving the SOC constrained reformulation problems, and established the linear rate of convergence under a local Lipschitz error bound assumption. Numerical experiments indicated that these reformulation problems are effective for the case where $\mathcal{E}$ is nonnegative orthant cone or SOC, and the SOC constrained reformulations derived from ψ with slower growth in $\mathcal{K} \times \mathcal{K}$ have better performance. In addition, numerical comparisons with the L-BFGS method for solving (16) with $\psi = \psi_{\mathrm{FB}}$ verify the effectiveness of Algorithm 4.1 for solving (15) with $\psi = \psi_4$.

Further studies are also needed to find suitable properties of f to guarantee that the local Lipschitz error bound assumption holds. Another direction is to analyze the properties of the solution set of (1) under suitable conditions of M, N and P. We note that the SOC constrained reformulation problems and the proximal gradient descent method in this paper can be extended to general symmetric cone linear complementarity problems.

Appendix A

Lemma 1. *For any $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R} \times \mathbb{R}^{l-1}$ with $l \geqslant 1$, if $u_1 v_1 \geqslant 0$ and $\langle u, v\rangle \geqslant 0$, we have $\langle u \circ (u \circ v), v \circ (u \circ v)\rangle \geqslant 0$, and when the equality holds, $\langle u, v\rangle = 0$.*

Proof. By the definition of Jordan product, we compute that

$$\begin{aligned}\big\langle u \circ (u \circ v), v \circ (u \circ v)\big\rangle &= \big(u^T v\big)^3 + 3u^T v\big(u_1^2\|v_2\|^2 + v_1^2\|u_2\|^2\big) \\ &\quad + 6u_1^2 v_1^2 u_2^T v_2 + 5u_1 v_1\big(u_2^T v_2\big)^2 + u_1 v_1\|u_2\|^2\|v_2\|^2 \\ &\geqslant \big(u^T v\big)^3 + 3u^T v\big(u_1^2\|v_2\|^2 + v_1^2\|u_2\|^2\big) \\ &\quad + 6u_1^2 v_1^2 u_2^T v_2 + 5u_1 v_1\big(u_2^T v_2\big)^2 + u_1 v_1\big(u_2^T v_2\big)^2 \\ &= \big(u^T v\big)^3 + 3u^T v\|u_1 v_2 + v_1 u_2\|^2 \geqslant 0\end{aligned}$$

where the first inequality uses $u_1 v_1 \geqslant 0$ and $\|u_2\|^2\|v_2\|^2 \geqslant (u_2^T v_2)^2$, and the second one is due to the nonnegativity of $u^T v$. This shows that the first part holds. The second part follows directly from the above inequality. $\square$

References

[1] R. Andreani, A. Friedlander, M.P. Mello, S.A. Santos, Box-constrained minimization reformulations of complementarity problems in second-order cones, J. Global Optim. 40 (2008) 505–527.

[2] R. Andreani, J.M. Martinez, On the solution of the extended linear complementarity problem, Linear Algebra Appl. 281 (1998) 247–257.
[3] D.P. Bertsekas, Nonlinear Programming, 2nd edition, Athena Scientific, Belmont, 1999.
[4] R.H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput. 16 (1995) 1190–1208.
[5] J.-S. Chen, P. Tseng, An unconstrained smooth minimization reformulation of the second-order cone complementarity problem, Math. Program. 104 (2005) 293–327.
[6] J.-S. Chen, D. Sun, J. Sun, The SC^1 property of the squared norm of the SOC Fischer–Burmeister function, Oper. Res. Lett. 36 (2008) 385–392.
[7] J. Faraut, A. Korányi, Analysis on Symmetric Cones, Oxford Math. Monogr., Oxford University Press, New York, 1994.
[8] M. Fukushima, H. Mine, A generalized proximal point algorithm for certain non-convex minimization problems, Internat. J. Systems Sci. 12 (1981) 989–1000.
[9] M. Fukushima, Z.-Q. Luo, P. Tseng, Smoothing functions for second-order cone complementarity problems, SIAM J. Optim. 12 (2002) 436–460.
[10] M.S. Gowda, On reducing a monotone horizontal LCP to an LCP, Appl. Math. Lett. 8 (1994) 97–100.
[11] M.S. Gowda, J.-S. Pang, Stability analysis of variational inequalities and nonlinear complementarity problems, via the mixed linear complementarity problem and degree theory, Math. Oper. Res. 19 (1994) 831–879.
[12] M.S. Gowda, R. Sznajder, The generalized order linear complementarity problem, SIAM J. Matrix Anal. Appl. 15 (1994) 779–795.
[13] M.S. Gowda, On the extended linear complementarity problem, Math. Program. 72 (1996) 33–50.
[14] S. Hayashi, T. Yamaguchi, N. Yamashita, M. Fukushima, A matrix splitting method for symmetric affine second-order cone complementarity problems, J. Comput. Appl. Math. 175 (2005) 335–353.
[15] J.B. Hiriart-Urruty, C. Lemaréchal, Convex Analysis and Minimization Algorithm I, Springer, New York, 1993.
[16] Y. Kanno, J.A.C. Martins, A. Pinto da Costa, Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem, Internat. J. Numer. Methods Engrg. 65 (2006) 62–83.
[17] L. Kong, L. Tuncel, N. Xiu, Vector-valued Implicit Lagrangian for symmetric cone complementarity problems, Asia-Pac. J. Oper. Res. 26 (2009) 199–233.
[18] Z.-Q. Luo, P. Tseng, One the linear convergence of descent methods for convex essentially smooth minimization, SIAM J. Control Optim. 30 (1992) 408–425.
[19] Z.-Q. Luo, P. Tseng, Error bounds and convergence analysis of feasible descent methods: a general approach, Ann. Oper. Res. 46 (1993) 157–178.
[20] O.L. Mangasarian, J.S. Pang, The extended linear complementarity problem, SIAM J. Matrix Anal. Appl. 16 (1995) 359–368.
[21] S.-H. Pan, J.-S. Chen, A damped Gauss–Newton method for the second-order cone complementarity problem, Appl. Math. Optim. 59 (2009) 293–318.
[22] S.-H. Pan, J.-S. Chen, Growth behavior of two classes of merit functions for symmetric cone complementarity problem, J. Optim. Theory Appl. 141 (2009) 167–191.
[23] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
[24] M.V. Solodov, Some optimization reformulations for the extended linear complementarity problem, Comput. Optim. Appl. 13 (1999) 187–200.
[25] P. Tseng, Merit function for semidefinite complementarity problems, Math. Program. 83 (1998) 159–185.
[26] P. Tseng, S.-W. Yun, A coordinate gradient descent method for nonsmooth separable minimization, Math. Program. 117 (2009) 387–423.
[27] R.H. Tütüncü, M.J. Todd, Reducing horizontal linear complementarity problem, Linear Algebra Appl. 223/224 (1995) 716–720.
[28] Y. Ye, A fully polynomial-time approximation algorithm for computing a stationary point of the general linear complementarity problem, Math. Oper. Res. 18 (1993) 334–345.