Optimization
Vol. 59, No. 8, November 2010, 1173–1197

Taylor & Francis
Taylor & Francis Group

A linearly convergent derivative-free descent method for the second-order cone complementarity problem

Shaohua Pan[a] and Jein-Shan Chen[b]*

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

(*Received 25 August 2008; final version received 1 May 2009*)

We consider a class of derivative-free descent methods for solving the second-order cone complementarity problem (SOCCP). The algorithm is based on the Fischer–Burmeister (FB) unconstrained minimization reformulation of the SOCCP, and utilizes a convex combination of the negative partial gradients of the FB merit function $\psi_{\rm FB}$ as the search direction. We establish the global convergence results of the algorithm under monotonicity and the uniform Jordan P-property, and show that under strong monotonicity the merit function value sequence generated converges at a linear rate to zero. Particularly, the rate of convergence is dependent on the structure of second-order cones. Numerical comparisons are also made with the limited BFGS method used by Chen and Tseng (*An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Math. Program. 104(2005), pp. 293–327), which confirm the theoretical results and the effectiveness of the algorithm.

Keywords: second-order cone complementarity problem; Fischer–Burmeister function; descent algorithms; derivative-free methods; linear convergence

1. Introduction

We consider the conic complementarity problem of finding a vector $\zeta \in {\rm I\!R}^n$ such that

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

where $F : {\rm I\!R}^n \to {\rm I\!R}^n$ is a mapping assumed to be continuously differentiable throughout this article, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOCs). In other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \mathcal{K}^{n_2} \times \cdots \times \mathcal{K}^{n_m}, \tag{2}$$

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

$$\mathcal{K}^{n_i} := \left\{ (x_1, x_2) \in {\rm I\!R} \times {\rm I\!R}^{n_i - 1} \mid x_1 \geq \|x_2\| \right\}, \tag{3}$$

*Corresponding author. Email: jschen@math.ntnu.edu.tw.

ISSN 0233–1934 print/ISSN 1029–4945 online

DOI: 10.1080/02331930903085359
http://www.informaworld.com

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of non-negative reals IR_+. We will refer to (1)–(2) as the *second-order cone complementarity problem* (SOCCP).

As a direct extension of the non-linear complementarity problem (NCP), the SOCCP includes as a special case the Karush-Kuhn-Tucker (KKT) system of SOC programming, which has a wide range of applications in engineering design, control, finance, robust optimization and combinatorial optimization; see [1,18] and the references therein. Now there have been various methods proposed for solving the SOCCP, which include the merit function method [5], the smoothing Newton methods [6,10,12], the semismooth Newton methods [16,20], and the interior-point method [25]. We observe that the last three kinds of methods in each iteration involve the solution of a linear system of equations, which makes them unsuitable for handling large-scale SOCCPs. On the contrary, the merit function method [5], based on the Fischer–Burmeister (FB) unconstrained minimization reformulation of the SOCCP, requires much less computation work in each iteration and consequently has a certain potential for solving large-scale SOCCPs.

The FB merit function associated with the cone $\mathcal{K}^n$ is given by

$$\;_{\mathrm{FB}}(x, y) := \tfrac{1}{2}\,\|\phi_{\mathrm{FB}}(x, y)\|^2, \tag{4}$$

where $\phi_{\mathrm{FB}} : \mathrm{IR}^n \times \mathrm{IR}^n \to \mathrm{IR}^n$ is the FB function associated with $\mathcal{K}^n$, defined by

$$\phi_{\mathrm{FB}}(x, y) := (x^2 + y^2)^{1/2} - (x + y) \tag{5}$$

with $x^2 = x \circ x$ denoting the Jordan product of x and itself, $x^{1/2}$ being a vector such that $(x^{1/2})^2 = x$, and $x + y$ meaning the componentwise addition of vectors. The functions ψ_{FB} and ϕ_{FB} were studied in the papers [2,5,10,21], where ψ_{FB} was shown in [10] to satisfy

$$\;_{\mathrm{FB}}(x, y) = 0 \Longleftrightarrow x \in \mathcal{K}^n, \quad y \in \mathcal{K}^n, \quad \langle x, y\rangle = 0, \tag{6}$$

and its continuous differentiability was established by Chen and Tseng [5], and ϕ_{FB} was proved to be strongly semismooth in [21] and [2] via different ways. By equivalence (6), clearly, the SOCCP can be reformulated as an unconstrained minimization problem

$$\min_{\zeta \in \mathrm{IR}^n} \;_{\mathrm{FB}}(\zeta) := \sum_{i=1}^{m} \;_{\mathrm{FB}}(\zeta_i, F_i(\zeta)), \tag{7}$$

where $\zeta = (\zeta_1, \ldots, \zeta_m)$, $F(\zeta) = (F_1(\zeta), \ldots, F_m(\zeta))$ with $\zeta_i \in \mathrm{IR}^{n_i}$ and $F_i : \mathrm{IR}^n \to \mathrm{IR}^{n_i}$.

The merit function method in [5] was developed by applying the limited BFGS method directly for the minimization reformulation (7). In this article, we propose another merit function method based on the same reformulation, which can be viewed as an extension of the method in [23] for the NCP. Different from the limited BFGS method adopted by Chen and Tseng [5], our method does not exploit the derivative of the mapping F, but utilizes some convex combination of the negative partial gradients of ψ_{FB}, i.e. the vector of the form $-\beta\nabla_x\psi_{\mathrm{FB}} - (1-\beta)\nabla_y\psi_{\mathrm{FB}}$ with $\beta \in (0,1)$, as the search direction. Since the computation of the search direction and the step size does not involve the Jacobian of F, our derivative-free algorithm requires less computation work and lower memory in each iteration than the existing methods mentioned above. We show that the algorithm is globally convergent under monotonicity and the uniform Jordan *P*-property of *F*, and particularly that the merit function value sequence $\{\psi_{\mathrm{FB}}(\zeta^k)\}$ generated converges at a linear rate to zero if *F* is strongly monotone. But, unlike the NCP case, the rate of convergence depends on the structure of $\mathcal{K}$ (Remark 5.1 (a)).

The literature on derivative-free methods for solving the NCP is vast; see, for example, [11,15,17,19,24,23]. Nevertheless, to the best of our knowledge, there are no papers to study derivative-free methods for the SOCCP except [3] where a different unconstrained reformulation and a different descent direction were employed, and no rate of convergence result was established. The main difficulty is to extend the growth relation between the FB function and the natural residual function established in [22] to the SOCCP case. In addition, numerical results were not reported for the above derivative-free methods, so the practical performance of these methods cannot be judged. In this article, we obtain the rate of convergence result for the proposed derivative-free descent algorithm by using the favourable properties of the gradients of the function ψ_{FB} (Propositions 3.1 and 3.2), as well as compare the performance of the algorithm with that of the limited BFGS method in [5], which indicates that our method is comparable to the limited BFGS method for some test problems.

Throughout this article, $\mathbb{R}^n$ denotes the space of *n*-dimensional real column vectors, and $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is identified with $\mathbb{R}^{n_1+\cdots+n_m}$. Thus, $(x_1, \ldots, x_m) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ is viewed as a column vector in $\mathbb{R}^{n_1+\cdots+n_m}$. The notation *I* means an identity matrix of suitable dimension, and $\mathrm{int}(\mathcal{K}^n)$ denotes the interior of $\mathcal{K}^n$. For any x, y in $\mathbb{R}^n$, we write $x \succeq_{\mathcal{K}^n} y$ if $x - y \in \mathcal{K}^n$; and write $x \succ_{\mathcal{K}^n} y$ if $x - y \in \mathcal{K}^n$. For a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, $\nabla F(x) \in \mathbb{R}^{n \times m}$ denotes the transposed Jacobian of F at x. For a symmetric matrix A, we write $A \succeq O$ (respectively, $A \succ O$) to mean A is positive semidefinite (respectively, positive definite). In addition, we use $\mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ to denote a diagonal matrix with $\sigma_1, \ldots, \sigma_n$ as the diagonal elements.

2. Preliminaries

This section recalls some background materials that will be used in the subsequent sections. It is known that $\mathcal{K}^n$ is a closed convex self-dual cone with non-empty interior

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

For any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we define their Jordan product [8] by

$$x \circ y := (\langle x, y \rangle,\ y_1 x_2 + x_1 y_2). \tag{8}$$

The Jordan product, unlike scalar or matrix multiplication, is not associative, which is a main source of 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

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

which can be viewed as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ with $L_x y = x \circ y$ for any $y \in \mathbb{R}^n$. It is easy to verify that L_x for $x \in \mathrm{int}(\mathcal{K}^n)$ is invertible with the inverse L_x^{-1} given by

$$L_x^{-1} = \frac{1}{\det(x)} \begin{bmatrix} x_1 & -x_2^T \\ -x_2 & \dfrac{\det(x)}{x_1} I + \dfrac{1}{x_1} x_2 x_2^T \end{bmatrix}, \tag{9}$$

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

We recall from [8,10] 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

$$\lambda_i(x) := x_1 + (-1)^i \|x_2\|, \ u_x^{(i)} := \frac{1}{2}\big(1, \ (-1)^i \bar{x}_2\big), \tag{10}$$

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. The spectral factorization of x and the matrix L_x have various interesting properties; see [10]. We list several ones that will be used later.

Lemma 2.1

(a) *For any* $x \in \mathbb{R}^n, x^2 = (\lambda_1(x))^2 \cdot u_x^{(1)} + (\lambda_2(x))^2 \cdot u_x^{(2)} \in \mathcal{K}^n$.
(b) *For any* $x \in \mathcal{K}^n, x^{1/2} = \sqrt{\lambda_1(x)} \cdot u_x^{(1)} + \sqrt{\lambda_2(x)} \cdot u_x^{(2)} \in \mathcal{K}^n$.
(c) $x \succeq_{\mathcal{K}^n} 0 \Longleftrightarrow \lambda_1(x) \geq 0 \Longleftrightarrow L_x \succeq O$ *and* $x \succ_{\mathcal{K}^n} 0 \Longleftrightarrow \lambda_1(x) > 0 \Longleftrightarrow L_x \succ O$.

The following lemma is a representation of Problem 7 in [13 p. 468] for the real symmetric matrix case. In view of its importance, we here include its proof.

Lemma 2.2 *Let* $B, C \in \mathbb{R}^{n \times n}$ *be symmetric matrices with* $B \succ O$. *Then* $B + C \succ O$ *if and only if every eigenvalue of* CB^{-1} *is greater than* -1.

Proof By Corollary 7.6.5 of [13], there exists a non-singular matrix $D \in \mathbb{R}^{n \times n}$ such that $D^T C D = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ and $D^T B D = \mathrm{I}$. Consequently,

$$\begin{aligned} CB^{-1} &= (D^T)^{-1} \mathrm{diag}(\sigma_1, \ldots, \sigma_n) D^{-1} \big[(D^T)^{-1} D^{-1}\big]^{-1} \\ &= (D^T)^{-1} \mathrm{diag}(\sigma_1, \ldots, \sigma_n) D^T. \end{aligned}$$

This implies that CB^{-1} is similar to the diagonal matrix $\mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, and therefore $\sigma_1, \ldots, \sigma_n$ are the eigenvalues of CB^{-1} including the multiplicities. On the other hand,

$$B + C = (D^T)^{-1} \mathrm{diag}(1 + \sigma_1, \ldots, 1 + \sigma_n) D^{-1},$$

which means that $B + C \succ O$ if and only if $\sigma_i > -1$ for all $i = 1, 2, \ldots, n$. Combining the two sides, we then obtain the desired result. ∎

Next, we review the definitions of the monotonicity and the P-property of a mapping.

Definition 2.1 The mapping $F = (F_1, \ldots, F_m)$ with $F_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ is said to

(a) be monotone if, for every $\zeta, \xi \in \mathbb{R}^n$, $\langle \zeta - \xi, \ F(\zeta) - F(\xi) \rangle \geq 0$;

(b) be strongly monotone if there exists a $\mu > 0$ such that, for every $\zeta, \xi \in \mathbb{R}^n$,

$$\langle \zeta - \xi,\ F(\zeta) - F(\xi)\rangle \geq \mu\|\zeta - \xi\|^2;$$

(c) have the uniform Jordan P-property if there exists a $\rho > 0$ such that, for every $\zeta = (\zeta_1, \ldots, \zeta_m), \xi = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^n$, there exists $\nu \in \{1,2,\ldots,m\}$ such that

$$\lambda_2[(\zeta_\nu - \xi_\nu) \circ (F_\nu(\zeta) - F_\nu(\xi))] \geq \rho\|\zeta - \xi\|^2;$$

(d) have the uniform Cartesian P-property if there exists a $\rho > 0$ such that, for every $\zeta = (\zeta_1, \ldots, \zeta_m), \xi = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^n$, there exist $\nu \in \{1,2,\ldots,m\}$ such that

$$\langle \zeta_\nu - \xi_\nu, F_\nu(\zeta) - F_\nu(\xi)\rangle \geq \rho\|\zeta - \xi\|^2.$$

From Definition 2.1, clearly, the uniform Cartesian P-property implies the uniform Jordan P-property, and if F is strongly monotone with modulus $\mu > 0$, then F has the uniform Jordan P-property and the uniform Cartesian P-property with modulus μ/m. Also, when F is continuously differentiable, F is strongly monotone with modulus $\mu > 0$ if and only if $\nabla F(\zeta)$ is uniformly positive definite with modulus $\mu > 0$, i.e.

$$d^T \nabla F(\zeta) \mathrm{d} \geq \mu\|d\|^2 \quad \text{for all } \zeta, \mathrm{d} \in \mathbb{R}^n.$$

In addition, we see that the uniform Jordan P-property does not imply the monotonicity.

Unless otherwise stated, in the subsequent three sections, we assume $\mathcal{K} = \mathcal{K}^n$, and all analysis can be carried over to the case where $\mathcal{K}$ has the Cartesian structure as in (2).

3. Some properties of ψ_{FB} and Ψ_{FB}

In this section, we present some important properties for the gradient of ψ_{FB} which play a crucial role in analysing the convergence results of the descent algorithm proposed in the next section. In addition, we establish the coerciveness of Ψ_{FB} under two mild conditions. Throughout this section, for any $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R} \times \mathbb{R}^{n-1}$, we write

$$w = (w_1, w_2) := x^2 + y^2 \quad \text{and} \quad z := (z_1, z_2) = (x^2 + y^2)^{1/2}. \tag{11}$$

First, from Propositions 1 and 2 of [5], we know that the function ψ_{FB} is continuously differentiable everywhere and its gradient is given as in the following lemma.

LEMMA 3.1 *The function* ψ_{FB} *in* [4] *is continuously differentiable everywhere. Moreover*, $\nabla_x\psi_{\mathrm{FB}}(0,0) = \nabla_y\psi_{\mathrm{FB}}(0,0) = 0$. *If* $x^2 + y^2 \in \mathrm{int}(\mathcal{K}^n)$, *then*

$$\nabla_x\psi_{\mathrm{FB}}(x,y) = \left(L_x L_z^{-1} - I\right)\phi_{\mathrm{FB}}(x,y),$$
$$\nabla_y\psi_{\mathrm{FB}}(x,y) = \left(L_y L_z^{-1} - I\right)\phi_{\mathrm{FB}}(x,y).$$

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

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

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

For the partial gradients $\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}$ and $\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}$, from [5 Lemma 9] and [4, Theorem 3.1], we readily obtain the following favourable properties whose proofs will be omitted.

PROPOSITION 3.1 *The gradients* $\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}$ *and* $\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}$ *of* $\psi_{\scriptscriptstyle \mathrm{FB}}$ *have the following properties*:

(a) $\langle\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}(x,y),\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)\rangle\geq 0$ *for all* $x,y\in\mathbb{R}^n$, *and furthermore, the equality holds if and only if* $\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)=0$.

(b) *For all* $x,y\in\mathbb{R}^n$, $\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)=\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)=0$ *if and only if* $\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)=0$.

(c) $\nabla\psi_{\mathrm{FB}}$ *is globally Lipschitz continuous, i.e. there exists a constant* $L>0$ *such that*

$$\|\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)-\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}(\bar{x},\bar{y})\|\leq L\|(x,y)-(\bar{x},\bar{y})\|,$$
$$\|\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}(x,y)-\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}(\bar{x},\bar{y})\|\leq L\|(x,y)-(\bar{x},\bar{y})\|.$$

for all $(x,y),(\bar{x},\bar{y})\in\mathbb{R}^n\times\mathbb{R}^n$, *where* L *is dependent on the dimension* n.

Next we will establish another three important properties for the gradients $\nabla_x\psi_{\scriptscriptstyle \mathrm{FB}}$ and $\nabla_y\psi_{\scriptscriptstyle \mathrm{FB}}$ (Proposition 3.2) which are crucial to analyse the convergent results in Sections 4 and 5. To the end, we need the following technical lemmas. The first one is an extension of [5, Lemma 3], which will be used to give a tighter upper bound for $L_{x+y}L_z^{-1}$.

LEMMA 3.2 *For any* $x=(x_1,x_2), y=(y_1,y_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$ *such that* $w_2\neq 0$, *we have*

$$\left|(x_1+y_1)+(-1)^i(x_2+y_2)^T\bar{w}_2\right|^2\leq\left\|(x_2+y_2)+(-1)^i(x_1+y_1)\bar{w}_2\right\|^2\leq 2\lambda_i(w) \tag{12}$$

for $i=1,2$, *where* $\bar{w}_2=w_2/\|w_2\|$.

Proof The first inequality can be easily obtained by expanding the square on both sides and using the Cauchy–Schwartz inequality. We next show that the second inequality holds when $i=1$, which is equivalent to proving the following inequality:

$$\left\|(x_2+y_2)\|w_2\|-(x_1+y_1)w_2\right\|^2\leq 2\lambda_1(w)\|w_2\|^2. \tag{13}$$

Let L and R denote the left-hand side and the right-hand side of (13), respectively. Then, by plugging in $w_2=2(x_1x_2+y_1y_2)$, it is easy to compute that

$$\begin{aligned}
L=&\|x_2+y_2\|^2\|w_2\|^2+(x_1+y_1)^2\|w_2\|^2\\
&-4\left[x_1^2\|x_2\|^2+x_1y_1x_2^Ty_2+x_1^2x_2^Ty_2+x_1y_1\|x_2\|^2\right]\|w_2\|\\
&-4\left[y_1^2\|y_2\|^2+x_1y_1x_2^Ty_2+y_1^2x_2^Ty_2+x_1y_1\|y_2\|^2\right]\|w_2\|,\\
R=&2(x_1^2+y_1^2)\|w_2\|^2+2(\|x_2\|^2+\|y_2\|^2)\|w_2\|^2\\
&-4\left[2x_1^2\|x_2\|^2+2y_1^2\|y_2\|^2+4x_1y_1x_2^Ty_2\right]\|w_2\|.
\end{aligned}$$

Using the last two equalities, it then follows that

$$\begin{aligned}
R-L &= (x_1-y_1)^2\|w_2\|^2 + \|x_2-y_2\|^2\|w_2\|^2 \\
&\quad - 4\big(x_1^2\|x_2\|^2 + y_1^2\|y_2\|^2 + 2x_1y_1x_2^Ty_2\big)\|w_2\| \\
&\quad + 4\big(x_1^2x_2^Ty_2 + x_1y_1\|y_2\|^2 + y_1^2x_2^Ty_2 + x_1y_1\|x_2\|^2\big)\|w_2\| \\
&= (x_1-y_1)^2\|w_2\|^2 + \|x_2-y_2\|^2\|w_2\|^2 - 2(x_1-y_1)(x_2-y_2)^Tw_2\|w_2\| \\
&= \big\|(x_1-y_1)w_2 - (x_2-y_2)\|w_2\|\big\|^2 \geq 0.
\end{aligned}$$

This implies (13), and consequently the inequality (12) holds for $i=1$. Using similar arguments, we can prove that the inequality (12) holds for $i=2$. ■

Lemma 3.3 *For any* $x=(x_1,x_2), y=(y_1,y_2)\in \mathbb{R}\times\mathbb{R}^{n-1}$ *such that* $x^2+y^2\in \mathrm{int}(\mathcal{K}^n)$,

$$\big\|L_{x+y}L_z^{-1}\big\|_2 \leq 2(\sqrt{n-1}+2\sqrt{2}),$$

where $\|A\|_2$ *denotes the Frobenius norm* (*Euclidean norm*) *of the matrix* $A\in\mathbb{R}^{n\times n}$.

Proof Let λ_1, λ_2 be the spectral values of w. Then, by the definition of z, we have

$$z_1 = \frac{\sqrt{\lambda_2}+\sqrt{\lambda_1}}{2}, \quad z_2 = \frac{\sqrt{\lambda_2}-\sqrt{\lambda_1}}{2}\bar{w}_2 \tag{14}$$

with $\bar{w}_2 = \frac{w_2}{\|w_2\|}$ if $w_2\neq 0$, and otherwise $\bar{w}_2$ being any vector in $\mathbb{R}^{n-1}$ satisfying $\|\bar{w}_2\|=1$.

If $w_2=0$, then $\lambda_1=\lambda_2=w_1=\|x\|^2+\|y\|^2$. From formula (9), it follows that

$$L_{x+y}L_z^{-1} = \frac{1}{\sqrt{w_1}}L_{x+y} = \frac{1}{\sqrt{\|x\|^2+\|y\|^2}}L_{x+y}.$$

Consequently,

$$\big\|L_{x+y}L_z^{-1}\big\|_2^2 = \frac{n(x_1+y_1)^2+2\|x_2+y_2\|^2}{\|x\|^2+\|y\|^2} \leq 2n,$$

which immediately implies the desired result.

If $w_2\neq 0$, then by applying formula (9), it is not difficult to compute that

$$\begin{aligned}
L_{x+y}L_z^{-1} &= \begin{bmatrix} \dfrac{(x_1+y_1)z_1-(x_2+y_2)^Tz_2}{\sqrt{\lambda_1}\sqrt{\lambda_2}} & -\dfrac{(x_1+y_1)z_2^T}{\sqrt{\lambda_1}\sqrt{\lambda_2}} + \dfrac{(x_2+y_2)^T}{z_1} + \dfrac{(x_2+y_2)^Tz_2z_2^T}{z_1\sqrt{\lambda_1}\sqrt{\lambda_2}} \\ \dfrac{(x_2+y_2)z_1-(x_1+y_1)z_2}{\sqrt{\lambda_1}\sqrt{\lambda_2}} & -\dfrac{(x_2+y_2)z_2^T}{\sqrt{\lambda_1}\sqrt{\lambda_2}} + \dfrac{(x_1+y_1)I}{z_1} + \dfrac{(x_1+y_1)z_2z_2^T}{z_1\sqrt{\lambda_1}\sqrt{\lambda_2}} \end{bmatrix} \\
&:= \begin{bmatrix} b_1(x,y) & b_2(x,y)^T \\ c_2(x,y) & B_2(x,y) \end{bmatrix}.
\end{aligned}$$

Substituting the expressions of z_1, z_2 in (14) into the entries of the above matrix, we get

$$
\begin{aligned}
b_1(x,y) &= \frac{(x_1+y_1)+(x_2+y_2)^T\bar{w}_2}{2\sqrt{\lambda_2}} + \frac{(x_1+y_1)-(x_2+y_2)^T\bar{w}_2}{2\sqrt{\lambda_1}},\\
c_2(x,y) &= \frac{(x_2+y_2)+(x_1+y_1)\bar{w}_2}{2\sqrt{\lambda_2}} + \frac{(x_2+y_2)-(x_1+y_1)\bar{w}_2}{2\sqrt{\lambda_1}},\\
b_2(x,y) &= \frac{\lambda_1[(x_1+y_1)+(x_2+y_2)^T\bar{w}_2]\bar{w}_2}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\lambda_2[(x_1+y_1)-(x_2+y_2)^T\bar{w}_2]\bar{w}_2}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})}\\
&\quad + \frac{2(x_2+y_2)-(x_2+y_2)^T\bar{w}_2\bar{w}_2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}},\\
B_2(x,y) &= \frac{\lambda_1[(x_2+y_2)+(x_1+y_1)\bar{w}_2]\bar{w}_2^T}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\lambda_2[(x_2+y_2)-(x_1+y_1)\bar{w}_2]\bar{w}_2^T}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})}\\
&\quad + \frac{(x_1+y_1)}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}(2I-\bar{w}_2\bar{w}_2^T).
\end{aligned}
$$

Now, using Lemma 3.2, we can verify that the following inequalities hold:

$$
\begin{aligned}
\left|\frac{(x_1+y_1)+(x_2+y_2)^T\bar{w}_2}{2\sqrt{\lambda_2}}\right| &\le \left\|\frac{(x_2+y_2)+(x_1+y_1)\bar{w}_2}{2\sqrt{\lambda_2}}\right\| \le \frac{1}{\sqrt{2}},\\
\left|\frac{(x_1+y_1)-(x_2+y_2)^T\bar{w}_2}{2\sqrt{\lambda_1}}\right| &\le \left\|\frac{(x_2+y_2)-(x_1+y_1)\bar{w}_2}{2\sqrt{\lambda_1}}\right\| \le \frac{1}{\sqrt{2}},
\end{aligned}
$$

and

$$
\begin{aligned}
\left\|\frac{\lambda_1[(x_1+y_1)+(x_2+y_2)^T\bar{w}_2]\bar{w}_2}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\lambda_2[(x_1+y_1)-(x_2+y_2)^T\bar{w}_2]\bar{w}_2}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})}\right\| &\le \sqrt{2},\\
\left\|\frac{\lambda_1[(x_2+y_2)+(x_1+y_1)\bar{w}_2]\bar{w}_2^T}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})} - \frac{\lambda_2[(x_2+y_2)-(x_1+y_1)\bar{w}_2]\bar{w}_2^T}{2\sqrt{\lambda_1}\sqrt{\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})}\right\|_2 &\le \sqrt{2}.
\end{aligned}
$$

This together with $|x_1+y_1| \le \sqrt{\lambda_1}+\sqrt{\lambda_2}$ and $\|x_2+y_2\| \le \sqrt{\lambda_1}+\sqrt{\lambda_2}$ implies that

$$
|b_1(x,y)| \le \|c_2(x,y)\| \le \sqrt{2},\ \ \|b_2(x,y)\| \le \sqrt{2}+3,\ \ \|B_2(x,y)\|_2 \le 2\sqrt{n-1}+1+\sqrt{2}.
$$

Consequently, $\|L_{x+y}L_z^{-1}\|_2 \le 2\sqrt{n-1}+4\sqrt{2}$. The proof is thus completed. ■

It should be pointed out that using Lemmas 3–4 of [5] we may also get a upper bound for $\|L_{x+y}L_z^{-1}\|_2$, but such a upper bound is not tighter than the one given here. By using Lemma 3.3, we can further obtain the following result. Its proof is simple, however, as will be shown below, this result is a key to establish Proposition 3.2 (b).

LEMMA 3.4 *For any given* $x, y \in \mathbb{R}^n$ *such that* $x^2+y^2 \in \mathrm{int}(\mathcal{K}^n)$, *let* $A := L_{2z-(x+y)}L_z^{-1}$ *and* $p_A(t) = t^n + a_1(x,y)t^{n-1} + \cdots + a_{n-1}(x,y)t + a_n(x,y)$ *be its characteristic polynomial. Then, there exists a constant* $c_1(n) > 1$ *dependent on n such that*

$$
\|A^{n-1} + a_1(x,y)A^{n-2} + \cdots + a_{n-1}(x,y)A\|_2 \le c_1(n). \tag{15}
$$

Proof For any given $x, y \in \mathbb{R}^n$ such that $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$, since $A = 2I - L_{x+y}L_z^{-1}$, applying Lemma 3.3 yields

$$\|A\|_2 \leq 2(\sqrt{n} + \sqrt{n-1} + 2\sqrt{2}). \tag{16}$$

Let $c_2(n) := 2(\sqrt{n} + \sqrt{n-1} + 2\sqrt{2})$. Then, from the inequality (3.1.11) of [14], we have

$$|\lambda_i(A)| \leq c_2(n), \quad i = 1, 2, \ldots, n,$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A including multiplicities. Since $a_k(x,y)$ is the sum of all $\binom{n}{k}$ k-fold products of distinct items from $\lambda_1(A), \ldots, \lambda_n(A)$, i.e.

$$a_k(x, y) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^{k} \lambda_{i_j}(A), \quad k = 1, 2, \ldots, n,$$

there exists a positive constant $c_3(n)$ only dependent on the dimension n such that

$$|a_k(x, y)| \leq c_3(n), \quad k = 1, 2, \ldots, n. \tag{17}$$

Combining Equations (16) and (17), we immediately obtain (15) with

$$c_1(n) := \max\left\{1, c_2(n)^{n-1} + c_3(n)c_2(n)^{n-2} + \cdots + c_3(n)c_2(n)\right\},$$

and consequently the desired result follows. ■

Now we are in a position to present the three crucial properties of $\nabla_x \psi_{\text{FB}}$ and $\nabla_y \psi_{\text{FB}}$.

Proposition 3.2 *The gradients $\nabla_x \psi_{\text{FB}}$ and $\nabla_y \psi_{\text{FB}}$ of ψ_{FB} have the following properties*:

(a) $\|\nabla_x \psi_{\text{FB}}(x, y) + \nabla_y \psi_{\text{FB}}(x, y)\| \leq 2(\sqrt{n} + \sqrt{n-1} + 2\sqrt{2})\|\phi_{\text{FB}}(x, y)\|$ *for all* $x, y \in \mathbb{R}^n$;

(b) $\left\|\nabla_x \psi_{\text{FB}}(x, y) + \nabla_y \psi_{\text{FB}}(x, y)\right\| \geq \frac{(3-2\sqrt{2})^n}{2^n c_1(n)} \|\phi_{\text{FB}}(x, y)\|$ *for all* $x, y \in \mathbb{R}^n$, *where* $c_1(n)$ *is the constant from Lemma* 3.4.

(c) $\|\nabla_x \psi_{\text{FB}}(x, y) + \nabla_y \psi_{\text{FB}}(x, y)\| = 0$ *if and only if* $x \in \mathcal{K}$, $y \in \mathcal{K}$, $\langle x, y \rangle = 0$.

Proof (a) We prove the result by the following three cases:

Case 1 $(x,y) = (0,0)$. In this case, the result is clear by Lemma 3.1 and $\phi_{\text{FB}}(0,0) = 0$.

Case 2 $x^2 + y^2 \in \text{int}(\mathcal{K}^n)$. Using Lemmas 3.1 and 3.3, it follows that

$$\begin{aligned} \left\|\nabla_x \psi_{\text{FB}}(x, y) + \nabla_y \psi_{\text{FB}}(x, y)\right\| &= \left\|(2I - L_{x+y}L_z^{-1})\phi_{\text{FB}}(x, y)\right\| \\ &\leq \|2I - L_{x+y}L_z^{-1}\|_2 \|\phi_{\text{FB}}(x, y)\| \\ &\leq 2(\sqrt{n} + \sqrt{n-1} + 2\sqrt{2})\|\phi_{\text{FB}}(x, y)\|. \end{aligned} \tag{18}$$

Case 3 $x^2 + y^2 \notin \text{int}(\mathcal{K}^n)$ and $(x,y) \neq (0, 0)$. From Lemma 3.1 we have that

$$\begin{aligned} \left\|\nabla_x \psi_{\text{FB}}(x, y) + \nabla_y \psi_{\text{FB}}(x, y)\right\| &= \left\|\left(\frac{x_1 + y_1}{\sqrt{x_1^2 + y_1^2}} - 2\right)\phi_{\text{FB}}(x, y)\right\| \\ &= \left(2 - \frac{x_1 + y_1}{\sqrt{x_1^2 + y_1^2}}\right)\|\phi_{\text{FB}}(x, y)\| \\ &\leq \|\phi_{\text{FB}}(x, y)\|, \end{aligned} \tag{19}$$

where the second equality is due to $(x_1+y_1)^2 \le 2(x_1^2+y_1^2)$, and the inequality is since $\sqrt{x_1^2+y_1^2} \le x_1+y_1$ by the non-negativity of x_1, y_1.
(b) Similar to part (a), we also proceed the proof by the three cases.

Case 1 $(x,y)=(0,0)$. The result is clear by Lemma 3.1 and $\phi_{\rm FB}(0,0)=0$.

Case 2 $x^2+y^2 \in \mathrm{int}(\mathcal{K}^n)$. In this case, from Lemma 3.1 it follows that

$$\left\|\nabla_x\psi_{\rm FB}(x,y)+\nabla_y\psi_{\rm FB}(x,y)\right\| = \left\|L_{2z-(x+y)}L_z^{-1}\phi_{\rm FB}(x,y)\right\|.$$

Notice that $z \succ_{\mathcal{K}^n} 0$ and $4z^2-(x+y)^2 = 2z^2+(x-y)^2z \succ_{\mathcal{K}^n} 0$. From [10, Proposition 3.4] we have $2z-(x+y) \succ_{\mathcal{K}^n} 0$, which by Lemma 2.1 (c) implies $L_{2z}-(x+y) \succ O$. Consequently,

$$\left\|\nabla_x\psi_{\rm FB}(x,y)+\nabla_y\psi_{\rm FB}(x,y)\right\| = \frac{\|\phi_{\rm FB}(x,y)\|}{\left\|\left(L_{2z-(x+y)}L_z^{-1}\right)^{-1}\right\|_2} = \frac{\|\phi_{\rm FB}(x,y)\|}{\left\|L_zL_{2z-(x+y)}^{-1}\right\|_2}. \tag{20}$$

We next prove that all eigenvalues of $L_zL_{2z-(x+y)}^{-1}$ are bounded. Since $L_{2z}-(x+y)\succ O$ and $L_{2z}-(x+y)+L_z\succ O$, setting $B=L_{2z}-(x+y)$, $C=L_{\rm z}$ and applying Lemma 2.2 then yields that every eigenvalue of CB^{-1} is greater than –1, i.e.

$$\lambda_i\left(L_zL_{2z-(x+y)}^{-1}\right) > -1, \quad i=1,2,\ldots,n. \tag{21}$$

On the other hand, since $z \succ_{\mathcal{K}^n} 0$ and $2z^2-(x+y)^2=(x-y)^2x \succeq_{\mathcal{K}^n} 0$, we have from [10, Proposition 3.4] that $\sqrt{2}z-(x+y)x \succeq_{\mathcal{K}^n} y\sqrt{2}z-|x+y|x \succeq_{\mathcal{K}^n} 0$. Consequently,

$$[2z-(x+y)]-\left(3/2-\sqrt{2}\right)z = (1/2)z+\sqrt{2}z-(x+y) \succ_{\mathcal{K}^n} 0.$$

This in turn implies $L_{2z-(x+y)}-L_{(3/2-\sqrt{2})z} \succ O$. Setting $B=L_{2z-(x+y)}, C=-L_{(3/2-\sqrt{2})z}$ and applying Lemma 2.2 again, we have

$$\lambda_i\left(-L_{(3/2-\sqrt{2})z}L_{2z-(x+y)}^{-1}\right) > -1, \quad i=1,2,\ldots,n,$$

and therefore,

$$\lambda_i\left(L_zL_{2z-(x+y)}^{-1}\right) < \frac{2}{3-2\sqrt{2}}, \quad i=1,2,\ldots,n. \tag{22}$$

Combining (21) and (22) shows that all eigenvalues of $L_zL_{2z-(x+y)}^{-1}$ are bounded and

$$\left|\lambda_i\left(L_zL_{2z-(x+y)}^{-1}\right)\right| < \frac{2}{3-2\sqrt{2}}, \quad i=1,2,\ldots,n. \tag{23}$$

Now let $A=L_{2z-(x+y)}L_z^{-1}$ and $p_{\rm A}(t)$ be the characteristic polynomial of A defined as in Lemma 3.4. Then, using the fact that $p_A(A)=0$, we obtain

$$A^{n-1}+a_1(x,y)A^{n-2}+\cdots+a_{n-1}(x,y)+a_n(x,y)A^{-1}=0,$$

which in turn implies that

$$\begin{aligned}A^{-1} &= -\frac{1}{a_n(x,y)}\left[A^{n-1} + a_1(x,y)A^{n-2} + \cdots + a_{n-1}(x,y)\right]\\ &= -\frac{1}{\lambda_1(A)\cdots\lambda_n(A)}\left[A^{n-1} + a_1(x,y)A^{n-2} + \cdots + a_{n-1}(x,y)\right]\\ &= -\lambda_1(A^{-1})\cdots\lambda_n(A^{-1})\left[A^{n-1} + a_1(x,y)A^{n-2} + \cdots + a_{n-1}(x,y)\right]. \qquad (24)\end{aligned}$$

Note that A^{-1} is precisely $L_z L^{-1}_{2z-(x+y)}$. Hence, from (23) to (24) and Lemma 3.4, we have

$$\|L_z L^{-1}_{2z-(x+y)}\|_2 = \|A^{-1}\|_2 \leq \left(\frac{2}{3-2\sqrt{2}}\right)^n c_1(n).$$

This together with (20) yields the desired result.

Case 3 $(x,y) \notin \mathrm{int}(\mathcal{K}^n)$ and $(x,y) \neq (0,0)$. Using (19) and $|x_1 + y_1| \leq \sqrt{2}\sqrt{x_1^2 + y_1^2}$,

$$\left\|\nabla_x \psi_{\mathrm{FB}}(x,y) + \nabla_y \psi_{\mathrm{FB}}(x,y)\right\| \geq (2-\sqrt{2})\|\phi_{\mathrm{FB}}(x,y)\|.$$

Noting that $2-\sqrt{2} \geq \frac{(3-2\sqrt{2})^n}{2^n c_1(n)}$ since $c_1(n) > 1$, the desired result follows.

(c) This is direct by using parts (a)–(b) and the equivalence (6). ■

In what follows, we establish the coerciveness of the function Ψ_{FB} under some mild assumptions of F. For this purpose, we assume that $\mathcal{K}$ is given by (2), and corresponding to the Cartesian structure of $\mathcal{K}$, write $\zeta = (\zeta_1, \ldots, \zeta_m)$ and $F = (F_1, \ldots, F_m)$ with $\zeta_i \in \mathbb{R}^{n_i}$ and $F_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$. The following lemma and assumptions will be needed.

Lemma 3.5 [20 Lemma 5.2] *Let ϕ_{FB} be defined by* (5). *For any sequence* $\{(x^k, y^k)\} \subseteq \mathbb{R}^n \times \mathbb{R}^n$, *let $\lambda_1^k \leq \lambda_2^k$ and $\mu_1^k \leq \mu_2^k$ denote the spectral values of x^k* and y^k, *respectively*.

(a) *If $\{\lambda_1^k\} \to -\infty$ or $\{\mu_1^k\} \to -\infty$, then $\{\|\phi_{\mathrm{FB}}(x^k, y^k)\|\} \to \infty$.*

(b) *If $\{\lambda_1^k\}$ and $\{\mu_1^k\}$ are bounded below, but $\{\lambda_2^k\}, \{\mu_2^k\} \to +\infty$ and $\left\{\frac{x^k}{\|x^k\|} \circ \frac{y^k}{\|y^k\|}\right\} \not\to 0$, then $\{\|\phi_{\mathrm{FB}}(x^k, y^k)\|\} \to \infty$.*

Assumption 3.1 *For any sequence $\{\zeta^k\} \subseteq \mathbb{R}^n$ satisfying $\lim_{k\to\infty} \|\zeta^k\| = \infty$, if there exists $\nu \in \{1, \ldots, m\}$ such that the sequences $\{\lambda_1(\zeta_\nu^k)\}, \{\lambda_1(F_\nu(\zeta^k))\}$ are bounded below, but $\{\lambda_2(\zeta_\nu^k)\}, \{\lambda_2(F_\nu(\zeta^k))\} \to \infty$, then there holds that*

$$\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|} \circ \frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|} \not\to 0 \quad as \quad k \to \infty. \tag{25}$$

Assumption 3.2 *There exist $\kappa > 0$ and $r \in (0,1]$ such that the mapping F satisfies*

$$\|F(\zeta)\| \leq \|F(0)\| + \kappa\|\zeta\|^r \quad \textit{for any } \zeta \in \mathbb{R}^n.$$

Proposition 3.3 *Let Ψ_{FB} be given by* (7) *and $F = (F_1, \ldots, F_m)$ with $F_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$. Then, the function Ψ_{FB} is coercive under one of the following conditions:*

(a) *F has the uniform Jordan* P-*property and Assumption* 3.1 *holds*;

(b) *F has the uniform Jordan* P-*property and Assumption* 3.2 *holds.*

Proof The proof is by contradiction. Assume that a sequence $\{\zeta^k\}$ exists such that $\lim_{k\to\infty}\|\zeta^k\|=\infty$ and the sequence $\{\Psi_{\rm FB}(\zeta^k)\}$ is bounded. Corresponding to the structure of $\mathcal{K}$, for each k we write $\zeta^k=(\zeta_1^k,\ldots,\zeta_m^k)$ with $\zeta_i^k\in{\rm I\!R}^{n_i}$. Define the index set

$$J:=\big\{i\in\{1,2,\ldots,m\}\mid\{\zeta_i^k\}\text{ is unbounded}\big\}.$$

Clearly, $J\neq\emptyset$ since $\{\zeta^k\}$ is unbounded. Let $\{\xi^k\}$ be a bounded sequence with $\xi^k=(\xi_1^k,\ldots,\xi_m^k)$ and $\xi_i^k\in{\rm I\!R}^{n_i}$ for $i=1,2,\ldots,m$, where ξ_i^k for each k is defined as follows:

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

(a) From the uniform Jordan P-property of F, there exists $\rho>0$ such that

$$\begin{aligned}\rho\|\zeta^k-\xi^k\|^2&\le\max_{i=1,\ldots,m}\lambda_2\big[(\zeta_i^k-\xi_i^k)\circ(F_i(\zeta^k)-F_i(\xi^k))\big]\\&=\max_{i\in J}\lambda_2\big[\zeta_i^k\circ(F_i(\zeta^k)-F_i(\xi^k))\big]\\&=\lambda_2\big[\zeta_\nu^k\circ(F_\nu(\zeta^k)-F_\nu(\xi^k))\big]\\&\le\|\zeta_\nu^k\circ(F_\nu(\zeta^k)-F_\nu(\xi^k))\|\\&\le\sqrt{2}\|\zeta_\nu^k\|\|F_\nu(\zeta^k)-F_\nu(\xi^k)\|,\end{aligned}\tag{26}$$

where ν is one of the indices for which the maximum is attained and which we have, without loss of generality, assumed to be independent of k, and the last inequality is easily shown by (8). Since $\nu\in J$, we assume without loss of generality that $\{\|\zeta_\nu^k\|\}\to\infty$. Since $\|\zeta^k-\xi^k\|^2\ge\|\zeta_\nu^k-\xi_\nu^k\|^2=\|\zeta_\nu^k\|^2$, dividing the both sides of (26) by $\|\zeta_\nu^k\|$ then yields

$$\rho\|\zeta_\nu^k\|\le\sqrt{2}\|F_\nu(\zeta^k)-F_\nu(\xi^k)\|\le\sqrt{2}\big(\|F_\nu(\zeta^k)\|+\|F_\nu(\xi^k)\|\big).$$

This, together with the boundedness of $\{F_\nu(\xi^k)\}$, implies $\{\|F_\nu(\zeta^k)\|\}\to\infty$. Thus,

$$\{\|\zeta_\nu^k\|\}\to\infty\quad\text{and}\quad\{\|F_\nu(\zeta^k)\|\}\to\infty.\tag{27}$$

Now if $\{\lambda_1(\zeta_\nu^k)\}\to-\infty$ or $\{\lambda_1(F_\nu(\zeta^k))\}\to-\infty$, then using Lemma 3.5 (a) readily yields $\{\|\phi_{\rm FB}(\zeta_\nu^k,F_\nu(\zeta^k))\|\}\to\infty$ and hence $\{\Psi_{\rm FB}(\zeta^k)\}\to\infty$, which gives a contradiction to the boundedness of $\{\Psi_{\rm FB}(\zeta^k)\}$. Otherwise, from (27) we have $\{\lambda_2(\zeta_\nu^k)\}\to\infty$ and $\{\lambda_2(F_\nu(\zeta^k))\}\to\infty$. By the given assumption, condition (25) holds. Then, $\{\zeta^k\}$ satisfies Lemma 3.5 (b), which in turn implies $\{\Psi_{\rm FB}(\zeta^k)\}\to\infty$. This is clearly impossible.

(b) From the above discussions, Equations (26)–(27) still hold for this case. If $\{\lambda_1(\zeta_\nu^k)\}\to-\infty$ or $\{\lambda_2(F_\nu(\zeta^k))\}\to-\infty$, then from part (a) it is impossible. Otherwise, from (27) we have $\{\lambda_2(\zeta_\nu^k)\}\to\infty$ and $\{\lambda_2(F_\nu(\zeta^k))\}\to\infty$. We next show that $\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|}\circ\frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|}\not\to0$ as $k\to\infty$. If not, by the continuity of $\lambda_2(\cdot)$ and Equation (27),

$$\begin{aligned}\lim_{k\to\infty}\frac{\lambda_2\big[\zeta_\nu^k\circ(F_\nu(\zeta^k)-F_\nu(\xi^k))\big]}{\|\zeta_\nu^k\|\|F_\nu(\zeta^k)\|}&\le\lim_{k\to\infty}\lambda_2\left[\frac{\zeta_\nu^k}{\|\zeta_\nu^k\|}\circ\frac{F_\nu(\zeta^k)}{\|F_\nu(\zeta^k)\|}\right]\\&\quad+\lim_{k\to\infty}\lambda_2\left[\frac{-\zeta_\nu^k\circ F_\nu(\xi^k)}{\|\zeta_\nu^k\|\|F_\nu(\zeta^k)\|}\right]\\&=0,\end{aligned}\tag{28}$$

where the inequality is easily shown by (8) and the equality is due to the boundedness of $\{F(\xi^k)\}$. On the other hand, from Assumption 3.2, there exist $\kappa > 0$ and $r \in (0,1]$ such that $\|F_\nu(\zeta^k)\| \leq \|F(\zeta^k)\| \leq \|F(0)\| + \kappa\|\zeta^k\|^r$ for each k, and hence,

$$\lim_{k\to\infty} \frac{\rho\|\zeta^k - \xi^k\|^2}{\|\zeta_\nu^k\|\|F_\nu(\zeta^k)\|} \geq \lim_{k\to\infty} \frac{\rho\|\zeta^k - \xi^k\|^2}{\|\zeta^k\|(\|F(0)\| + \kappa\|\zeta^k\|^r)} \geq \frac{\rho}{\kappa} > 0.$$

This together with the first inequality of (26) yields a contradiction to (28). Thus, we verify that the sequences $\{\zeta_\nu^k\}$ and $\{F_\nu(\zeta^k)\}$ satisfy the conditions of Lemma 3.5 (b). Consequently, we have $\{\Psi_{\mathrm{FB}}(\zeta^k)\} \to \infty$. This is clearly impossible. ■

Since the uniform Cartesian P-property implies the uniform Jordan P-property, the condition of Proposition 3.3 (a) is weaker than that of Proposition 5.2 in [20]. We also see that Assumption 3.2 is weaker than the Lipschitz continuity of F. When $\mathcal{K}$ reduces to the non-negative orthant cone $\mathbb{R}_+^n$ and the Jordan product '$\circ$' becomes the component wise product of the vectors, since Assumption 3.1 automatically holds and the uniform Jordan P-property of F is equivalent to saying that F is a uniform P-function, we readily recover the result of [7, Theorem 4.2] from Proposition 3.3 (a).

4. A descent method and global convergence

In this section, we propose a derivative-free descent algorithm based on the minimization reformulation (7). The algorithm will make use of the vector of the following form:

$$d(\zeta,\beta) := -\beta\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)) - (1-\beta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta)) \tag{29}$$

as the search direction, where $\beta \in [0,1)$ is a parameter. Note that $d(\zeta,\beta)$ for any $\beta \in [0,1)$ may not be a descent direction of Ψ_{FB} at ζ. But, the following lemma states that, when F is monotone, there always exists $\bar{\beta}(\zeta) \in (0,1]$ such that $d(\zeta, \beta)$ for any $\beta \in [0, \bar{\beta}(\zeta))$ is a descent direction. The idea for constructing such a direction is borrowed from [23].

Lemma 4.1 *Suppose that F is monotone. If ζ is not a solution of the SOCCP, then there exists $\bar{\beta}(\zeta) \in (0,1]$ such that $\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta,\beta) < 0$ for all $\beta \in [0, \bar{\beta}(\zeta))$.*

Proof Since F is continuously differentiable, the function $\Psi_{\mathrm{FB}}(\zeta)$ is also continuously differentiable by Lemma 3.1. Using the chain rule, the gradient of Ψ_{FB} at ζ is

$$\nabla\Psi_{\mathrm{FB}}(\zeta) = \nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)) + \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta)). \tag{30}$$

This together with the definition of $d(\zeta, \beta)$ yields that

$$\begin{aligned}\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta,\beta) = &-\beta\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2 - \beta\big\langle\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle \\ &- (1-\beta)\big\langle\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle \\ &- (1-\beta)\big\langle\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle.\end{aligned} \tag{31}$$

Let

$$q(\zeta) := -\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2 - \big\langle\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle$$

and

$$p(\zeta) := -\langle \nabla_x \psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla_y \psi_{\mathrm{FB}}(\zeta, F(\zeta))\rangle - \langle \nabla_y \psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y \psi_{\mathrm{FB}}(\zeta, F(\zeta))\rangle.$$

Then, (31) can be rewritten as

$$\nabla \Psi_{\mathrm{FB}}(\zeta)^T d(\zeta, \beta) = (1-\beta)p(\zeta) + \beta q(\zeta).$$

Note that the first term of $p(\zeta)$ is negative by Proposition 3.1 (a) since ζ is not a solution of the SOCCP, whereas the second term is non-positive since F is monotone. Therefore, we have $p(\zeta) < 0$. Let $\bar{\beta}(\zeta)$ be defined as follows:

$$\bar{\beta}(\zeta) := \begin{cases} \dfrac{-p(\zeta)}{q(\zeta)-p(\zeta)} & \text{if } q(\zeta) > p(\zeta) \text{ and } \dfrac{-p(\zeta)}{q(\zeta)-p(\zeta)} \leq 1; \\ 1 & \text{otherwise.} \end{cases}$$

We see that for all $\beta \in [0, \bar{\beta}(\zeta))$, the search direction $d(\zeta, \beta)$ defined by (29) satisfies the descent condition $\nabla \Psi_{\mathrm{FB}}(\zeta)^T d(\zeta,\beta) < 0$. The proof is thus completed. ■

Lemma 4.1 motivates us to propose the following descent algorithm with $d(\zeta, \beta)$.

Algorithm 4.1

Step 0. Choose $\zeta^0 \in \mathbb{R}^n$, $\epsilon \geq 0$, $\sigma \in (0, 1/2)$ and $\gamma, \beta \in (0,1)$ with $\gamma > \beta$. Set $k := 0$.

Step 1. If $\Psi_{\mathrm{FB}}(\zeta^k) \leq \epsilon$, then stop and ζ^k is an approximate solution of the SOCCP.

Step 2. Let l_k be the smallest non-negative integer l satisfying

$$\begin{aligned} &\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k) \\ &\quad \leq -\sigma\gamma^{2l} \|\nabla_x \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2, \end{aligned} \tag{32}$$

where $d(\zeta, \beta)$ is defined as in (29), and set

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

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

Algorithm 4.1 is similar to the one proposed in [23] for the NCP with a regularized FB merit function. Since there is no need to compute the gradient of Ψ_{FB} and the Jacobian of $F(\cdot)$, Algorithm 4.1 is suitable for large-scale problems, as well as applications where the Jacobians of $F(\cdot)$ are not available or are costly to compute. In addition, the stepsize and the search direction are adjusted during the backtracking search of Armijo-type, which may be regarded as a kind of curvilinear search.

In what follows, we analyse the global convergence of Algorithm 4.1. Without loss of generality, we assume that $\epsilon = 0$. We first show that under the monotonicity of F every accumulation point of the sequence $\{\zeta^k\}$ is a solution of the SOCCP.

THEOREM 4.1 *Suppose that F is monotone. Then, Algorithm* 4.1 *is well-defined for any initial point ζ^0. Furthermore, if ζ^* is an accumulation point of the sequence $\{\zeta^k\}$ generated by Algorithm* 4.1, *then ζ^* is a solution of the SOCCP.*

Proof The proofs are similar to those of [23, Theorem 4.1]. We first show that, whenever ζ^k is not a solution, there exists a non-negative integer l_k in Step 3 of

Algorithm 4.1 such that (32) holds. Suppose not, then for any positive integer l, we have

$$\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k) > -\sigma\gamma^{2l}\|\nabla_x \Psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y \Psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2.$$

Dividing the above inequality by γ^l and passing to the limit $l \to \infty$, we get

$$\lim_{l\to\infty} \frac{\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k)}{\gamma^l} \geq 0. \tag{33}$$

On the other hand, using the mean-value theorem, it follows that

$$\begin{aligned}
&\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, 0)) \\
&\quad = \gamma^l \nabla\Psi_{\mathrm{FB}}\big(\zeta^k + \gamma^l d(\zeta^k, 0) + t\gamma^l\,(d(\zeta^k, \beta^l) - d(\zeta^k, 0))\big)^T \big(d(\zeta^k, \beta^l) - d(\zeta^k, 0)\big) \\
&\quad = \gamma^l \beta^l \nabla\Psi_{\mathrm{FB}}\big(\zeta^k + \gamma^l d(\zeta^k, 0) + t\gamma^l \beta^l h(\zeta^k)\big)^T h(\zeta^k),
\end{aligned}$$

where t is a constant such that $t \in (0, 1)$ and $h(\zeta^k) := \nabla_y \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) - \nabla_x \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))$. From this and the continuity of $\nabla\Psi_{\mathrm{FB}}$, we immediately obtain

$$\lim_{l\to\infty} \frac{\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, 0))}{\gamma^l} = 0.$$

Consequently,

$$\begin{aligned}
&\lim_{l\to\infty} \frac{\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k)}{\gamma^l} \\
&= \lim_{l\to\infty} \frac{\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, 0))}{\gamma^l} \\
&\quad + \lim_{l\to\infty} \frac{\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, 0)) - \Psi_{\mathrm{FB}}(\zeta^k)}{\gamma^l} \\
&= \nabla\Psi_{\mathrm{FB}}(\zeta^k)^T d(\zeta^k, 0).
\end{aligned} \tag{34}$$

Combining (34) with (33) then yields $\nabla\Psi_{\mathrm{FB}}\ (\zeta^k)^T d(\zeta^k, 0) \geq 0$. This gives a contradiction, since, by Lemma 4.1, $d(\zeta^k, 0)$ must be a descent direction of Ψ_{FB} at ζ^k if ζ^k is not a solution of the SOCCP. Thus, Algorithm 4.1 is well defined.

Next, we prove that any accumulation point ζ^* of $\{\zeta^k\}$ is a solution of the SOCCP. Let $\{\zeta^k\}_{k\in K\}}$ be a subsequence converging to ζ^*. From the definition of $d(\zeta, \beta)$, we see that $d(\cdot,\cdot)$ is continuous, which implies that $d^k(\beta^{l_k}) = d(\zeta^k, \beta^{l_k}) \to d^*$ as $k (\in K) \to \infty$. Since $\Psi_{\mathrm{FB}}\ (\zeta^k)$ decreases at each iteration, the right-hand side of (32) tends to 0. We next proceed the discussions by two cases: $\{l_k\}_{k\in K}$ is bounded and $\{l_k\}_{k\in K}$ is unbounded.

Case 1 $\{l_k\}_{k\in K}$ is bounded. In this case, $\{\gamma^{l_k}\}_{k\in K}$ does not approach 0. Consequently,

$$\|\nabla_x \psi_{\mathrm{FB}}(\zeta^*, F(\zeta^*)) + \nabla_y \psi_{\mathrm{FB}}(\zeta^*, F(\zeta^*))\|^2 = 0.$$

From Proposition 3.2 (c), it then follows that ζ^* is a solution of the SOCCP.

Case 2 $\{l_k\}_{k\in K}$ is unbounded. Without loss of generality, assume that $\{l_k\}_{k\in K}\to\infty$. Now we have $\{\gamma^{l_k}\}_{k\in K}\to 0$. In addition, from Step 3 of Algorithm 4.1, it follows that

$$\Psi_{\text{FB}}(\zeta^k+\gamma^{l_k-1}d^k(\beta^{l_k-1}))-\Psi_{\text{FB}}(\zeta^k)> -\sigma\gamma^{2(l_k-1)}\|\nabla_x\psi_{\text{FB}}(\zeta^k,F(\zeta^k)) + \nabla_y\psi_{\text{FB}}(\zeta^k,F(\zeta^k))\|^2$$

for all $k\in K$. Dividing the above inequality by γ^{l_k-1} and passing to the limit $k\ (\in K)\to\infty$ then yields

$$\lim_{k\to\infty}\frac{\Psi_{\text{FB}}(\zeta^k+\gamma^{l_k-1}d^k(\beta^{l_k-1}))-\Psi_{\text{FB}}(\zeta^k)}{\gamma^{l_k-1}}\geq 0. \tag{35}$$

On the other hand, by the mean-value theorem, there exists a $\widehat{\gamma}^k\in[0,\gamma^{l_k-1}]$ such that

$$\frac{\Psi_{\text{FB}}(\zeta^k+\gamma^{l_k-1}d^k(\beta^{l_k-1}))-\Psi_{\text{FB}}(\zeta^k)}{\gamma^{l_k-1}}=\nabla\Psi_{\text{FB}}(\zeta^k+\widehat{\gamma}^k d^k(\beta^{l_k-1}))^T d^k(\beta^{l_k-1}).$$

Since $\{\gamma^{l_k-1}\}\to 0$ as $k(\in K)\to\infty$, we have $\widehat{\gamma}^k\to 0$, which in turn implies

$$\lim_{k\to\infty}\widehat{\gamma}^k d^k(\beta^{l_k-1})=0.$$

Combining the last two equations then yields that

$$\lim_{k\to\infty}\frac{\Psi_{\text{FB}}(\zeta^k+\gamma^{l_k-1}d^k(\beta^{l_k-1}))-\Psi_{\text{FB}}(\zeta^k)}{\gamma^{l_k-1}}=\nabla\Psi_{\text{FB}}(\zeta^*)^T d^*. \tag{36}$$

Now, from (35) and (36), we get $\nabla\Psi_{\text{FB}}(\zeta^*)^T d^*\geq 0$. Noting that $d^*=-\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*))$ and $\nabla\Psi_{\text{FB}}(\zeta^*)=\nabla_x\Psi_{\text{FB}}(\zeta^*,F(\zeta^*))+\nabla F(\zeta^*)\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*))$, we then have that

$$\begin{aligned}0\leq\nabla\Psi_{\text{FB}}(\zeta^*)^T d^* &= -\langle\nabla_x\psi_{\text{FB}}(\zeta^*,F(\zeta^*)),\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*))\rangle\\ &\quad-\langle\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*)),\nabla F(\zeta^*)\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*))\rangle\\ &\leq 0,\end{aligned}$$

where the last inequality is by Proposition 3.1 (a) and the monotonicity of F. Thus,

$$\langle\nabla_x\psi_{\text{FB}}(\zeta^*,F(\zeta^*)),\nabla_y\psi_{\text{FB}}(\zeta^*,F(\zeta^*))\rangle=0.$$

From Proposition 3.1 (a), it then follows that ζ^* is a solution of the SOCCP. ■

Theorem 4.1 together with Proposition 3.3 leads to the following result.

THEOREM 4.2 *If F is monotone, has the uniform Jordan* P*-property and satisfies Assumption* 3.1 (*or Assumption* 3.2, *then the sequence* $\{\zeta^k\}$ *generated by Algorithm* 4.1 *at least has one accumulation point and any accumulation point is a solution of the SOCCP.*

Since strong monotonicity implies monotonicity and uniform Jordan P-property, Theorem 4.2 also holds if F is strongly monotone and satisfies Assumptions 3.1 or 3.2.

5. Linear convergence rate

In this section, we show that the merit function value sequence $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ generated by Algorithm 4.1 converges linearly to the solution of the SOCCP if F is strongly monotone by using similar analysis techniques to those of [23]. We still assume that $\epsilon = 0$ in this section. The following technical lemma will be needed.

LEMMA 5.1 *Let S be any given bounded set. Suppose that F is strongly monotone with modulus $\mu < 0$. Then, for any $\zeta \in S$, there exists an integer $\bar{l} > 0$ such that for all $l \geq \bar{l}$,*

$$\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta, \beta^l) \leq -\frac{\beta^l}{2}\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)) + \nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2.$$

Proof Since ∇ F is continuous, there exists a constant $\vartheta > 0$ such that

$$\|\nabla F(\zeta)\| \leq \vartheta \quad \text{for all } \zeta \in S. \tag{37}$$

Now, from Equation (31) and Proposition 3.1 (a), it follows that

$$\begin{aligned}
&\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta, \beta^l)\\
&\leq -\beta^l\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2 - \beta^l\big\langle\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle\\
&\quad - (1-\beta^l)\big\langle\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta)), \nabla F(\zeta)\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\big\rangle\\
&\leq -\beta^l\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2 - (1-\beta^l)\mu\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2\\
&\quad + \beta^l\vartheta\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\\
&= -\frac{1}{2}\beta^l\big(\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\| + \|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\big)^2\\
&\quad -\frac{1}{2}\beta^l\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2 - \frac{2(1-\beta^l)\mu - \beta^l}{2}\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|^2\\
&\quad + \beta^l(\vartheta+1)\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\| \quad \text{for all } \zeta \in S,
\end{aligned} \tag{38}$$

where the second inequality is by (37) and the strong monotonicity of F. If

$$\beta^l \leq \frac{\mu}{2\mu+1},$$

then the inequality (38) can be rewritten as

$$\begin{aligned}
&\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta, \beta^l)\\
&\leq -\frac{1}{2}\beta^l\big(\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\| + \|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\big)^2\\
&\quad - \left(\sqrt{\frac{\beta^l}{2}}\Big\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\Big\| - \sqrt{\frac{2(1-\beta^l)\mu - \beta^l}{2}}\Big\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\Big\|\right)^2\\
&\quad + \Big(\beta^l(\vartheta+1) - \sqrt{2\mu\beta^l - (2\mu+1)\beta^{2l}}\Big)\|\nabla_x\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|\|\nabla_y\psi_{\mathrm{FB}}(\zeta, F(\zeta))\|.
\end{aligned} \tag{39}$$

Suppose that $\beta^l(\vartheta+1) - \sqrt{2\mu\beta^l - (2\mu+1)\beta^{2l}} \leq 0$, i.e.

$$\beta^l \leq \frac{2\mu}{(\vartheta+1)^2 + (2\mu+1)}.$$

Then, from Equation (39) and the Cauchy–Schwartz inequality, we obtain that

$$\begin{aligned}\nabla\Psi_{\mathrm{FB}}(\zeta)^T d(\zeta,\beta^l) &\leq -\tfrac{1}{2}\beta^l\big(\|\nabla_x\psi_{\mathrm{FB}}(\zeta,F(\zeta))\| + \|\nabla_y\psi_{\mathrm{FB}}(\zeta,F(\zeta))\|\big)^2 \\ &\leq -\tfrac{1}{2}\beta^l\big\|\nabla_x\psi_{\mathrm{FB}}(\zeta,F(\zeta)) + \nabla_y\psi_{\mathrm{FB}}(\zeta,F(\zeta))\big\|^2. \end{aligned} \tag{40}$$

Summing up the above discussions, whenever

$$\beta^l \leq \min\left\{\frac{2\mu}{2\mu+1}, \frac{2\mu}{(\vartheta+1)^2+(2\mu+1)}\right\} = \frac{2\mu}{(\vartheta+1)^2+(2\mu+1)},$$

or

$$l \geq \bar{l} := \left\lceil \log_\beta \frac{2\mu}{(\vartheta+1)^2+(2\mu+1)} \right\rceil, \tag{41}$$

we have that (40) holds. Thus, the proof is completed. ■

Let ζ^0 be any starting point of Algorithm 4.1. By Proposition 3.3, if F is strongly monotone and Assumption 3.1 (or Assumption 3.2) holds, then the level set

$$\mathcal{L}(\Psi_{\mathrm{FB}}, \Psi_{\mathrm{FB}}(\zeta^0)) := \left\{\zeta \in \mathrm{IR}^n \mid \Psi_{\mathrm{FB}}(\zeta) \leq \Psi_{\mathrm{FB}}(\zeta^0)\right\}$$

is bounded. By the continuity of ∇F and $\nabla\psi_{\mathrm{FB}}$, it further follows that the quantity

$$d_{\max} := \sup\left\{\|d(\zeta,\beta)\| \mid \zeta \in \mathcal{L}(\Psi_{\mathrm{FB}}, \Psi_{\mathrm{FB}}(\zeta^0))\right\}$$

is finite for any $\beta \in [0,1]$. Consequently, the following set

$$\mathcal{B}(\zeta^0) := \mathcal{L}(\Psi_{\mathrm{FB}}, \Psi_{\mathrm{FB}}(\zeta^0)) + \left\{\mathrm{d} \in \mathrm{IR}^n \mid \|d\| \leq d_{\max}\right\}$$

is also bounded. We are now ready to state and prove the linear convergence result.

THEOREM 5.1 *Suppose that F is strongly monotone with modulus $\mu > 0$ and satisfies Assumption* 3.1 (*or Assumption* 3.2). *Let $\zeta^0 \in \mathrm{IR}^n$ be the starting point of Algorithm* 4.1. *If ∇F is Lipschitz continuous on the set $\mathcal{B}(\zeta^0)$, then the sequence $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ converges Q-linearly to zero.*

Proof Since $\nabla F(\cdot)$ is Lipschitz continuous on $\mathcal{B}(\zeta^0)$ and F is continuous, by Proposition 3.1 (c) it is easily shown that $\nabla\Psi_{\mathrm{FB}}(\cdot)$ is Lipschitz continuous on this bounded set. In particular, there exists a positive constant $L_1(n)$ dependent on n such that

$$\|\nabla\Psi_{\mathrm{FB}}(\zeta) - \nabla\Psi_{\mathrm{FB}}(\xi)\| \leq L_1(n)\|\zeta-\xi\| \quad \forall \zeta,\xi \in \mathcal{B}(\zeta^0). \tag{42}$$

From the construction of Algorithm 4.1, the sequence $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ is non-increasing, and hence $\{\zeta^k\} \subseteq \mathcal{L}(\Psi_{\mathrm{FB}}, \Psi_{\mathrm{FB}}(\zeta^0))$. This implies that $\zeta^k, \zeta^k + td(\zeta^k, \beta^l) \in \mathcal{B}(\zeta^0)$ for any $t \in [0,1]$. From the mean-value theorem, it follows that

$$\begin{aligned}
&\Psi_{\mathrm{FB}}(\zeta^k + td(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k)\\
&= \int_0^t \nabla\Psi_{\mathrm{FB}}(\zeta^k + sd(\zeta^k, \beta^l))^T d(\zeta^k, \beta^l)\mathrm{d}s\\
&= t\nabla\Psi_{\mathrm{FB}}(\zeta^k)^T d(\zeta^k, \beta^l)\\
&\quad + \int_0^t \left(\nabla\Psi_{\mathrm{FB}}(\zeta^k + sd(\zeta^k, \beta^l)) - \nabla\Psi_{\mathrm{FB}}(\zeta^k)\right)^T d(\zeta^k, \beta^l)\mathrm{d}s\\
&\le t\nabla\Psi_{\mathrm{FB}}(\zeta^k)^T d(\zeta^k, \beta^l) + L_1(n)\int_0^t s\|d(\zeta^k, \beta^l)\|^2\mathrm{d}s\\
&= t\left(\nabla\Psi_{\mathrm{FB}}(\zeta^k)^T d(\zeta^k, \beta^l) + \frac{1}{2}L_1(n)t\|d(\zeta^k, \beta^l)\|^2\right),
\end{aligned} \tag{43}$$

where the inequality is by the Cauchy–Schwartz inequality and (42). Note that

$$\begin{aligned}
\|d(\zeta^k, \beta^l)\|^2 &= \|\beta^l\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + (1-\beta^l)\nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2\\
&= \beta^{2l}\|\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2 + (1-\beta^l)^2\|\nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2\\
&\quad + 2\beta^l(1-\beta^l)\langle\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)), \nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\rangle\\
&\le \|\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2 + \|\nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\|^2\\
&\quad 2\left\langle\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)), \nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\right\rangle\\
&= \left\|\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\right\|^2,
\end{aligned} \tag{44}$$

where the inequality is due to $\beta^l \in (0,1]$ and Proposition 3.1 (a). Let $\bar{l}$ be defined as in (41). Then, from equations (43) to (44) and Lemma 5.1, we have for all $l \ge \bar{l}$,

$$\begin{aligned}
&\Psi_{\mathrm{FB}}(\zeta^k + \gamma^l d(\zeta^k, \beta^l)) - \Psi_{\mathrm{FB}}(\zeta^k)\\
&\le \gamma^l\left(-\frac{\beta^l}{2} + \frac{L_1(n)\gamma^l}{2}\right)\left\|\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\right\|^2\\
&= -\frac{\gamma^l}{2}\left(\beta^l - L_1(n)\gamma^l\right)\left\|\nabla_x\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y\psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k))\right\|^2.
\end{aligned} \tag{45}$$

This implies that (32) is satisfied whenever $-\frac{1}{2}\gamma^l(\beta^l - L_1(n)\gamma^l) \le -\sigma(\gamma^l)^2$ or

$$\left(\frac{\gamma}{\beta}\right)^l \le \frac{1}{2\sigma + L_1(n)}.$$

Consequently, the condition in (32) is satisfied for all $l \ge \hat{l}$, where $\hat{l}$ is defined as

$$\hat{l} := \max\left\{\bar{l}, \left\lceil \log_{\frac{\beta}{\gamma}}(2\sigma + L_1(n)) \right\rceil\right\}. \tag{46}$$

Observing that $\hat{l}$ does not depend on k, we have $l_k \leq \hat{l}$ for all k since l_k is the smallest non-negative integer l satisfying (32). From (32) and Proposition 3.2 (b), it follows that

$$\begin{aligned}\Psi_{\mathrm{FB}}(\zeta^{k+1}) - \Psi_{\mathrm{FB}}(\zeta^k) &\leq -\sigma\gamma^{2l_k} \left\| \nabla_x \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) \right\|^2 \\ &\leq -\sigma\gamma^{2\hat{l}} \left\| \nabla_x \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) + \nabla_y \psi_{\mathrm{FB}}(\zeta^k, F(\zeta^k)) \right\|^2 \\ &\leq -\sigma\gamma^{2\hat{l}} \frac{(3-2\sqrt{2})^{2n}}{(2^n c_1(n))^2} 2\Psi_{\mathrm{FB}}(\zeta^k),\end{aligned}$$

where $c_1(n) > 1$ is the constant from Lemma 3.4. From this, we immediately obtain

$$0 \leq \Psi_{\mathrm{FB}}(\zeta^{k+1}) \leq \left[1 - \sigma\gamma^{2\hat{l}} \frac{2(3-2\sqrt{2})^{2n}}{(2^n c_1(n))^2}\right] \Psi_{\mathrm{FB}}(\zeta^k). \tag{47}$$

Notice that $\sigma\gamma^{2\hat{l}} \in (0, 1)$ and $\sqrt{2}(3-2\sqrt{2})^n < 2^n c_1(n)$, and consequently

$$0 < 1 - \sigma\gamma^{2\hat{l}} \frac{2(3-2\sqrt{2})^{2n}}{(2^n c_1(n))^2} < 1.$$

Thus, (47) shows that the sequence $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ converges Q-linearly to zero. ■

Remark 5.1

(a) From (47), we observe that the convergence rate of Algorithm 4.1 is related to the dimension n when $\mathcal{K} = \mathcal{K}^n$. By this, when $\mathcal{K}$ has the Cartesian structure as (2), it is not hard to verify that the convergence rate of Algorithm 4.1 depends on the value of

$$\min_{1 \leq i \leq m} \left\{ \left(\frac{3-2\sqrt{2}}{2}\right)^{2n_i} \left(\frac{1}{c_1(n_i)}\right)^2 \right\},$$

where $c_1(n_i)$ is determined as in Lemma 3.4 and only related to n_i. This means that the convergence rate of Algorithm 4.1 depends on the structure of $\mathcal{K}$.

(b) From (46) and (47), we know that for the same test problems, Algorithm 4.1 will have better rate of convergence results if γ is larger and the ratio of β/γ is smaller.

6. Numerical experiments

In this section, we apply Algorithm 4.1 for solving the SOCCP and compare its numerical performance with that of the limited Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method used by Chen and Tseng [5]. Since the corresponding test instances cannot be found in the literature, we consider the case where $F(\zeta) = M\zeta + b$ with the matrix $M \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$ generated randomly.

In the experiments, the matrix M was generated by the following procedure: choose the positive semidefinite matrices $M_i \in \mathbb{R}^{n_i \times n_i}$ for $i = 1, 2, \ldots, m$, and then let M be the block diagonal matrix with $M_1, \ldots, M_m$ as block diagonals, i.e.

$M = \text{diag}(M_1, \ldots, M_m)$. The positive semidefinite matrix M_i for $i = 1, 2, \ldots, m$ was set to be $M_i = N_i N_i^T$, where $N_i \in \mathbb{R}^{n_i \times n_i}$ was a square matrix whose non-zero elements were chosen randomly from a normal distribution with mean -1 and variance 4. We can verify that the matrix M generated by such way is positive semidefinite, and furthermore, it cannot be positive definite by controlling the non-zero density of N_i such that every block matrix M_i has at least zero eigenvalues. During the tests, the non-zero density of N_i for $i = 1, 2, \ldots, m$ was chosen as 1%. This means that the corresponding F is monotone but does not necessarily have the uniform Jordan P-property. The vector b was obtained by letting $b = -Mw$, where $w = (w_1, \ldots, w_m) \in \mathcal{K}$ with $w_i \in \mathcal{K}^{n_i}$ generated in the following way: let the elements of w_i be chosen randomly from a normal distribution with mean -1 and variance 4, and then set $w_{i1} = \|w_{i2}\|$ where w_{i1} is the first element of w_i and w_{i2} is a vector composed of the rest $n_i - 1$ components of w_i. In this way, the affine SOCCP was guaranteed to have a solution $\zeta^* = w$. To construct SOCs of various types, we chose n_i and m such that $n_1 = n_2 = \cdots = n_m$. All experiments were done with a PC of 2.8 GHz CPU and 512 MB memory. The computer codes were all written in Matlab 6.5.

We first used Algorithm 4.1 with $\beta = 0.9$, $\gamma = 0.8$ and $\beta = 0.9$, $\gamma = 0.1$, respectively, to solve a test problem generated as above with $n = 1000$ and $m = 100$. The parameters ϵ and σ in Algorithm 4.1 were chosen as: $\epsilon = 10^{-8}$ and $\sigma = 10^{-4}$. The starting point ζ^0 is set to be $(\bar{\zeta}^{n_i}, \ldots, \bar{\zeta}^{n_m})$, where $\bar{\zeta}^{n_i} = (10, \omega_i / \|\omega_i\|)$ for $i = 1, 2, \ldots, m$ with $\omega_i \in \mathbb{R}^{n_i - 1}$ generated randomly by Matlab's rand.m. The Figures 1 and 2 below plot the corresponding convergence of $\{\Psi_{\text{FB}}(\zeta^k)\}$ *versus* the iteration number. From the two figures, when $\beta = 0.9$ and $\gamma = 0.8$, $\Psi_{\text{FB}}(\zeta^k)$ has a faster decrease once its value is less than 10^{-2}. This implies that Algorithm 4.1 with a

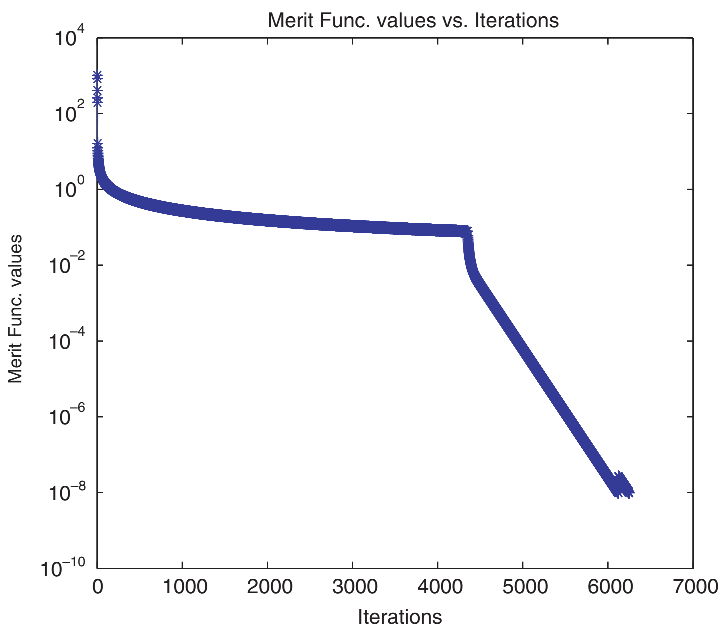

Figure 1. Convergence behaviour of $\{\Psi_{\text{FB}}(\zeta^k)\}$ with $\beta = 0.9$ and $\gamma = 0.8$.

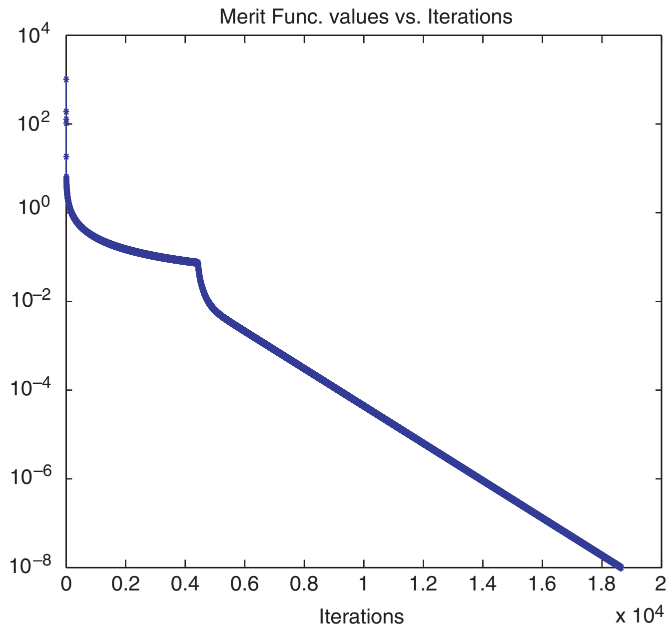

Figure 2. Convergence behavior of $\{\Psi_{\mathrm{FB}}(\zeta^k)\}$ with $\beta = 0.9$ and $\gamma = 0.1$.

larger γ and a smaller ratio β/γ indeed has a better rate of convergence, which coincides with the analysis in Remark 5.1 (b).

To test how the performance of Algorithm 4.1 varies with the structure of $\mathcal{K}$, we used Algorithm 4.1 to solve two groups of test problems generated as above with $n = 1000$ and $m = 100$ and $m = 20$, respectively, and compared its numerical performance with that of the L-BFGS method used by Chen and Tseng [5]. The parameters in Algorithm 4.1 were set as:

$$\epsilon = 10^{-8}, \quad \beta = 0.5, \quad \gamma = 0.4, \quad \delta = 10^{-4}.$$

We started Algorithm 4.1 from the initial point $\zeta^0 = (\bar{\zeta}^{n_i}, \ldots, \bar{\zeta}^{n_m})$, where $\bar{\zeta}^{n_i} = (10, \frac{\omega_i}{\|\omega_i\|})$ for $i = 1, 2, \ldots, m$ with $\omega_i \in \mathbb{R}^{n_i - 1}$ generated randomly by Matlab's rand.m. The two methods were terminated whenever $\Psi_{\mathrm{FB}}(\zeta^k) \leq \epsilon$ or the number of iteration is over 10^5.

The numerical results were summarized in Tables 1 and 2, in which $\Psi_{\mathrm{FB}}(\zeta^*)$ denotes the merit function value at the final iteration, NF indicates the number of function evaluations of ψ_{FB}, Iter reports the number of iteration required in order to satisfy the termination condition, Gap means the value of $|\zeta^T F(\zeta)|$ at the final iteration and Time denotes the CPU time in second for solving each problem. From Tables 1 and 2, we see that Algorithm 4.1 and the L-BFGS method in [5] require fewer function evaluations and iterations for those problems with $m = 100$, and moreover the L-BFGS method has better numerical performance, but for those problems with $m = 20$, Algorithm 4.1 is superior to the L-BFGS method in terms of the number of iterations and the CPU time. We also find that the matrix M in those problems with $m = 20$ has more non-zero entries and fewer zero eigenvalues.

Table 1. Numerical results for the affine monotone SOCCPs with 100 SOCs.

	Algorithm 4.1					L-BFGS Method				
Problem	$\Psi_{FB}(\zeta^*)$	NF	Iter	Gap	Time	$\Psi_{FB}(\zeta^*)$	NF	Iter	Gap	Time
1	9.98e–9	5744	5031	3.79e–4	123.9	9.27e–9	256	246	9.80e–4	8.23
2	9.88e–9	6897	5788	4.76e–4	145.8	9.12e–9	232	197	1.59e–4	7.00
3	9.96e–9	9068	6687	2.19e–4	184.4	9.50e–9	397	385	2.41e–4	12.71
4	9.86e–9	6359	5310	3.09e–4	137.9	9.95e–9	636	615	1.64e–4	19.89
5	9.57e–9	6637	5812	3.97e–4	145.8	9.97e–9	452	380	3.81e–4	13.40
6	9.99e–9	19195	11143	1.64e–4	359.1	9.54e–9	745	730	2.62e–4	25.02
7	9.99e–9	23359	9283	1.72e–4	405.2	9.54e–9	1192	1157	8.97e–5	37.64
8	9.77e–9	4443	4033	3.04e–4	97.93	9.94e–9	205	194	9.94e–5	6.59
9	9.82e–9	5809	5162	5.07e–4	126.9	9.61e–9	191	159	1.44e–4	5.50
10	9.99e–9	11448	7517	5.32e–5	226.1	7.89e–9	1069	999	1.47e–4	32.98

Table 2. Numerical results for the affine monotone $\mathrm{SOCCP_s}$ with $m = 20$ SOCs.

	Algorithm 4.1					L-BFGS Method				
Problem	$\Psi_{FB}(\zeta^*)$	NF	Iter	Gap	Time	$\Psi_{FB}(\zeta^*)$	NF	Iter	Gap	Time
1	9.99e–9	85240	25457	3.19e–4	840.7	9.99e–9	48339	48291	3.22e–4	990.5
2	9.99e–9	216544	50741	2.23e–4	2073.4	1.40e–6	100160	100000	2.78e–3	2088.1
3	9.99e–9	67285	20743	4.99e–4	667.4	9.99e–9	97133	97044	4.35e–4	2038.8
4	9.99e–9	91968	21835	3.49e–4	884.9	9.99e–9	43620	43538	2.07e–4	858.6
5	9.99e–9	196049	53584	2.51e–4	1913.7	1.26e–6	100071	100000	1.63e–3	2067.1
6	9.99e–9	205716	59262	3.21e–4	2024.1	7.61e–7	100089	100000	2.73e–3	2039.6
7	9.99e–9	127062	33457	3.42e–4	1254.7	7.07e–7	100072	100000	3.53e–3	2085.8
8	9.99e–9	123650	36698	4.10e–4	1238.5	1.16e–6	100051	100000	4.09e–3	2031.8
9	9.99e–9	59916	18344	1.19e–4	578.7	2.37e–8	100088	100000	3.91e–5	1973.5
10	9.99e–9	233878	64759	3.13e–4	2347.3	1.53e–6	100082	100000	3.99e–3	2071.3

7. Conclusions

We have extended the derivative-free descent method in [23] to solve the SOCCP, based on the FB unconstrained minimization reformulation (7) and the descent direction given by a convex combination of negative partial gradients $-\nabla_x \psi_{FB}$ and $-\nabla_y \psi_{FB}$. We showed that for the strongly monotone case, the sequence $\{\psi_{FB}(\zeta^k)\}$ generated converges globally to zero at a linear rate, and analysed that the rate of convergence depends on the structure of $\mathcal{K}$. Numerical comparisons indicated that the derivative-free descent algorithm is comparable to that of the limited BFGS method used by Chen and Tseng [5] for some affine SOCCPs.

We want to point out that the arguments and the convergence results in Sections 4 and 5 are parallel to those for the NCP discussed in [23], but our analysis relies on the important properties of the gradient of Ψ_{FB}, instead of the growth relation between the FB merit function and the natural residual merit function. We did not establish the result that Ψ_{FB} provides a global error bound for the SOCCP. This is possible to achieve for some other merit functions, but is always hard to obtain for Ψ_{FB}. Nonetheless, from Theorem 5.1, if the function Ψ_{FB} is able to provide a global error bound for the solution of the SOCCP under the strong monotonicity of F, then we can expect that the sequence $\{\zeta^k\}$ itself will converge R-linearly to the unique solution of the SOCCP. Such an issue deserves further investigation in the future.

In addition, in view of Proposition 3.1 (a)–(b), it is not difficult to obtain the global convergence of the proposed derivative-free method for the case where F is non-smooth by using the similar arguments as in [9] for the NCP. However, unlike the NCP case, if $\psi_{\rm FB}$ is replaced by the penalized FB merit function $\psi_\lambda(x,y) = \frac{1}{2}\|\lambda\phi_{\rm FB}(x,y) - (1-\lambda)(x)_+ \circ (y)_+\|^2$, where $\lambda \in (0,1)$ is a fixed parameter and $(x)_+$ means the Euclidean projection onto $\mathcal{K}$, we cannot expect Algorithm 4.1 and the corresponding convergence results since ψ_λ is not continuously differentiable everywhere.

Acknowledgements

The authors would like to thank one of anonymous referees for his helpful suggestions on the revision of this article. The author's work is partially supported by National Science Council of Taiwan.

References

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Math. Program. 95 (2003), pp. 3–51.
[2] J.-S. Chen, *Alternative proofs for some results of vector-valued functions associated with second-order cone*, J. Nonlinear Convex Anal. 6 (2005), pp. 297–325.
[3] J.-S. Chen and S.-H. Pan, *A derivative-free descent method for the second-order cone complementarity problem*, J. Comput. Appl. Mathematics 64 (2008), pp. 495–519.
[4] J.-S. Chen, D. Sun, and J. Sun, *The SC^1 property of the squared norm of the SOC Fischer–Burmeister function*, Oper. Res. Lett. 36 (2008), pp. 385–392.
[5] J.-S. Chen and P. Tseng, *An unconstrained smooth minimization reformulation of the second-order cone complementarity problem*, Math. Program. 104 (2005), pp. 293–327.
[6] X.-D. Chen, D. Sun, and J. Sun, *Complementarity functions and numerical experiments for second-order cone complementarity problems*, Comput. Optimiz. Appl. 25 (2003), pp. 39–56.
[7] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problem and a related algorithm*, SIAM J. Optimiz. 7 (1997), pp. 225–247.
[8] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, *Oxford Mathematical Monographs*, Oxford University Press, New York, 1994.
[9] A. Fischer, V. Jeyakumar, and D.T. Luc, *Solution point characterizations and convergence analysis of a descent algorithm for nonsmooth continuous complementarity problems*, J. Optimiz. Theory Appl. 110 (2001), pp. 493–513.
[10] M. Fukushima, Z.-Q. Luo, and P. Tseng, *Smoothing functions for second-order cone complementarity problems*, SIAM J. Optimiz. 12 (2002), pp. 436–460.
[11] C. Geiger and C. Kanzow, *On the resolution of monotone complementarity problems*, Comput. Optimiz. Appl. 5 (1996), pp. 155–173.
[12] S. Hayashi, N. Yamashita, and M. Fukushima, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM J. Optimiz. 15 (2005), pp. 593–615.
[13] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1986.

[14] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[15] H. Jiang, *Unconstrained minimization approaches to nonlinear complementarity problems*, J. Global Optimiz. 9 (1996), pp. 169–181.
[16] C. Kanzow, I. Ferenczi, and M. Fukushima, *Semismooth Newton methods for linear and nonlinear second-order cone programs*, Tech. Rep., Department of Applied Mathematics and Physics, Kyoto University, 2006.
[17] C. Kanzow, Y. Yamashita, and M. Fukushima, *New NCP functions and their properties*, J. Optimiz. Theory Appl. 97 (1997), pp. 115–135.
[18] M.S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Application of second-order cone programming*, Linear Algebra Appl. 284 (1998), pp. 193–228.
[19] O.L. Mangasarian and M.V. Solodov, *A linearly convergent derivative-free descent method for strongly monotone complementarity problems*, Comput. Optimiz. Appl. 64 (1999), pp. 5–16.
[20] S.-H. Pan, and J.-S. Chen, *A damped Guass–Newton method for the second-order cone complementarity problem*, Appl. Math. Optimiz. 59 (2009), pp. 293–318.
[21] D. Sun and J. Sun, *Strong semismoothness of Fischer–Burmeister SDC and SOC function*, Math. Program. 103 (2005), pp. 575–581.
[22] P. Tseng, *Growth behaviour of a class of merit functions for the nonlinear complementarity problem*, J. Optimiz. Theory Appl. 89 (1996), pp. 17–37.
[23] 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. 463–487.
[24] N. Yamashita and M. Fukushima, *On stationary points of the implicit Lagrangian for nonlinear complementarity problems*, J.Optimiz. Theory Appl. 84 (1995), pp. 653–663.
[25] A. Yoshise, *Interior point trajectories and a homogenous model for nonlinear complementarity problems over symmetric cones*, SIAM J. Optimiz. 17 (2006), pp. 1129–1153.