Comput Optim Appl (2012) 51:1037–1063
DOI 10.1007/s10589-011-9399-x

A proximal point algorithm for the monotone second-order cone complementarity problem

Jia Wu · Jein-Shan Chen

Received: 4 August 2010 / Published online: 1 March 2011

Abstract This paper is devoted to the study of the proximal point algorithm for solving monotone second-order cone complementarity problems. The proximal point algorithm is to generate a sequence by solving subproblems that are regularizations of the original problem. After given an appropriate criterion for approximate solutions of subproblems by adopting a merit function, the proximal point algorithm is verified to have global and superlinear convergence properties. For the purpose of solving the subproblems efficiently, we introduce a generalized Newton method and show that only one Newton step is eventually needed to obtain a desired approximate solution that approximately satisfies the appropriate criterion under mild conditions. Numerical comparisons are also made with the derivative-free descent method used by Pan and Chen (Optimization 59:1173–1197, 2010), which confirm the theoretical results and the effectiveness of the algorithm.

Keywords Complementarity problem · Second-order cone · Proximal point algorithm · Approximation criterion

J.-S. Chen is a member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

J. Wu
School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China
e-mail: jwu_dora@mail.dlut.edu.cn

J.-S. Chen (✉)
Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan
e-mail: jschen@math.ntnu.edu.tw

1 Introduction

The second-order cone complementarity problem (SOCCP) which is a natural extension of nonlinear complementarity problem (NCP), is to find $x \in \Re^n$ satisfying

$$\text{SOCCP}(F): \ \langle F(x), x\rangle = 0, \quad F(x) \in \mathcal{K}, \ x \in \mathcal{K}, \tag{1.1}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product, F is a mapping from $\Re^n$ into $\Re^n$, and $\mathcal{K}$ is the Cartesian product of second-order cones (SOC), in other words,

$$\mathcal{K} = \mathcal{K}^{n_1} \times \cdots \times \mathcal{K}^{n_q}, \tag{1.2}$$

where $q, n_1, \ldots, n_q \geq 1, n_1 + \cdots + n_q = n$, and for each $i \in \{1, \ldots, q\}$

$$\mathcal{K}^{n_i} := \{(x_0, \bar{x}) \in \Re \times \Re^{n_i - 1} | \|\bar{x}\| \leq x_0\},$$

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ denoting the set of nonnegative reals $\Re_+$. If $\mathcal{K} = \Re^n_+$, then (1.1) reduces to the nonlinear complementarity problem. Throughout this paper, corresponding to the Cartesian structure of $\mathcal{K}$, we write $F = (F_1, \ldots, F_q)$ and $x = (x_1, \ldots, x_q)$ with F_i being mappings from $\Re^n$ to $\Re^{n_i}$ and $x_i \in \Re^{n_i}$, for each $i \in \{1, \ldots, q\}$. We also assume that the mapping F is continuously differentiable and monotone.

Until now, a variety of methods for solving SOCCP have been proposed and investigated. They include interior-point methods [1, 2, 13, 18, 28, 31], the smoothing Newton methods [6, 10], the merit function method [5] and the semismooth Newton method [11], where the last three kinds of methods are all based on an SOC complementarity function or a merit function.

The proximal point algorithm (PPA) is known for its theoretically nice convergence properties, which was first proposed by Martinet [16] and further studied by Rockafellar [24]. PPA is a procedure for finding a vector z satisfying $0 \in \mathcal{T}(z)$, where $\mathcal{T}$ is a maximal monotone operator. Therefore, it can be applied to a broad class of problems such as convex programming problems, monotone variational inequality problems, and monotone complementarity problems.

In this paper, motivated by the work of Yamashita and Fukushima [29] for the NCPs, we focus on introducing PPA for solving the SOC complementarity problems. For SOCCP(F), given the current point x^k, PPA obtains the next point x^{k+1} by approximately solving the subproblem

$$\text{SOCCP}(F^k): \langle F^k(x), x\rangle = 0, \quad F^k(x) \in \mathcal{K}, \ x \in \mathcal{K}, \tag{1.3}$$

where $F^k : \Re^n \to \Re^n$ is defined by

$$F^k(x) := F(x) + c_k(x - x^k) \tag{1.4}$$

with $c_k > 0$. It is obvious that F^k is strongly monotone when F is monotone. Then, by [8, Theorem 2.3.3], the subproblem SOCCP(F^k), which is more tractable than the original problem, always has a unique solution. Thus, PPA is well defined. It was pointed out in [15, 24] that with appropriate criteria for approximate solutions of subproblems (1.3), PPA has global and superlinear convergence property under mild conditions. However, those criteria are usually not easy to check. Inspired by [29], we give a practical criterion based on a new merit function for SOCCP proposed by Chen in [3]. Another implementation issue is how to solve subproblems efficiently and obtain an approximate solution such that the approximation criterion for the subproblem is fulfilled. We use a generalized Newton method proposed by De Luca et al. [14] which is used in [29] for the NCP case to solve subproblems. We also give the conditions under which the approximation criterion is eventually approximately fulfilled by a single Newton iteration of the generalized Newton method.

The following notations and terminologies are used throughout the paper. I represents an identity matrix of suitable dimension, $\Re^n$ denotes the space of n-dimensional real column vectors, and $\Re^{n_1} \times \cdots \times \Re^{n_q}$ is identified with $\Re^{n_1+\cdots+n_q}$. Thus, $(x_1, \ldots, x_q) \in \Re^{n_1} \times \cdots \times \Re^{n_q}$ is viewed as a column vector in $\Re^{n_1+\cdots+n_q}$. For any two vectors u and v, the Euclidean inner product is denoted by $\langle u, v\rangle := u^T v$ and for any vector w, the norm $\|w\|$ is induced by the inner product which is called the Euclidean vector norm. For a matrix M, the norm $\|M\|$ is denoted to be the matrix norm induced by the Euclidean vector norm, that is the spectral norm. Given a differentiable mapping $F : \Re^n \to \Re^l$, we denote by $\mathcal{J}F(x)$ the Jacobian of F at x and $\nabla F(x) := \mathcal{J}F(x)^*$, the adjoint of $\mathcal{J}F(x)$. For a symmetric matrix M, we write $M \succ O$ (respectively, $M \succeq O$) if M is positive definite (respectively, positive semidefinite). Given a finite number of square matrices $Q_1, \ldots, Q_q$, we denote the block diagonal matrix with these matrices as block diagonals by $\mathrm{diag}(Q_1, \ldots, Q_q)$ or by $\mathrm{diag}(Q_i, i = 1, \ldots, q)$. If $\mathcal{I}$ and $\mathcal{B}$ are index sets such that $\mathcal{I}, \mathcal{B} \subseteq \{1, 2, \ldots, q\}$, we denote by $P_{\mathcal{I}\mathcal{B}}$ the block matrix consisting of the sub-matrices $P_{ik} \in \Re^{n_i \times n_k}$ of P with $i \in \mathcal{I}$, $k \in \mathcal{B}$, and denote by $x_{\mathcal{B}}$ a vector consisting of sub-vectors $x_i \in \Re^{n_i}$ with $i \in \mathcal{B}$.

The organization of this paper is as follows. In Sect. 2, we recall some notions and background materials. Section 3 is devoted to developing proximal point method to solve the monotone second-order cone complementarity problem with a practical approximation criterion based on a new merit function. In Sect. 4, a generalized Newton method is introduced to solve the subproblems and we prove that the proximal point algorithm in Sect. 3 has approximate genuine superlinear convergence under mild conditions, which is the main result of this paper. In Sect. 5, we report the numerical results for several test problems. Section 6 is to give conclusions.

2 Preliminaries

In this section, we review some background materials that will be used in the sequel. We first recall some mathematical concepts and the Jordan algebra associated with the SOC. Then we talk about the complementarity functions and three merit functions for SOCCP. Finally, we briefly mention the proximal point algorithm.

2.1 Mathematical concepts

Given a set $\Omega \in \Re^n$ locally closed around $\bar{x} \in \Omega$, define the regular normal cone to Ω at $\bar{x}$ by

$$\widehat{\mathcal{N}}_\Omega(\bar{x}) := \left\{ v \in \Re^n \,\middle|\, \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 0 \right\}.$$

The (limiting) normal cone to Ω at $\bar{x}$ is defined by

$$\mathcal{N}_\Omega(\bar{x}) := \limsup_{x \xrightarrow{\Omega} \bar{x}} \widehat{\mathcal{N}}_\Omega(x),$$

where "limsup" is the Painlevé-Kuratowski outer limit of sets (see [25]).

We now recall definitions of monotonicity of a mapping which are needed for the assumptions throughout this paper. We say that a mapping $G : \Re^n \to \Re^n$ is monotone if

$$\langle G(\zeta) - G(\xi), \zeta - \xi \rangle \geq 0, \quad \forall \zeta, \xi \in \Re^n.$$

Moreover, G is strongly monotone if there exists $\rho > 0$ such that

$$\langle G(\zeta) - G(\xi), \zeta - \xi \rangle \geq \rho \|\zeta - \xi\|^2, \quad \forall \zeta, \xi \in \Re^n.$$

It is well known that, when G is continuously differentiable, G is monotone if and only if $\nabla G(\zeta)$ is positive semidefinite for all $\zeta \in \Re^n$ while G is strongly monotone if and only if $\nabla G(\zeta)$ is positive definite for all $\zeta \in \Re^n$. For more details about monotonicity, please refer to [8].

There is another kind of concepts called Cartesian P-properties which have close relationship with monotonicity concept and are introduced by Chen and Qi [4] for a linear transformation. Here we present the definitions of Cartesian P-properties for a matrix $M \in \Re^{n\times n}$ and the nonlinear generalization in the setting of $\mathcal{K}$.

A matrix $M \in \Re^{n\times n}$ is said to have the Cartesian P-property if for any $0 \neq x = (x_1, \ldots, x_q) \in \Re^n$ with $x_i \in \Re^{n_i}$, there exists an index $\nu \in \{1, 2, \ldots, q\}$ such that $\langle x_\nu, (Mx)_\nu \rangle > 0$. And M is said to have the Cartesian P_0-property if the above strict inequality becomes $\langle x_\nu, (Mx)_\nu \rangle \geq 0$ where the chosen index ν satisfies $x_\nu \neq 0$.

Given a mapping $G = (G_1, \ldots, G_q)$ with $G_i : \Re^n \to \Re^{n_i}$, G is said to have the uniform Cartesian P-property if for any $x = (x_1, \ldots, x_q)$, $y = (y_1, \ldots, y_q) \in \Re^n$, there is an index $\nu \in \{1, 2, \ldots, q\}$ and a positive constant $\rho > 0$ such that

$$\langle x_\nu - y_\nu, G_\nu(x) - G_\nu(y) \rangle \geq \rho \|x - y\|^2.$$

In addition, for a single-valued Lipschitz continuous mapping $G : \Re^n \to \Re^m$, the B-subdifferential of G at x denoted by $\partial_B G(x)$, is defined as

$$\partial_B G(x) := \left\{ \lim_{k\to\infty} \mathcal{J} G(x^k) \,\middle|\, x^k \to x, G \text{ is differentiable at } x^k \right\}.$$

The convex hull of $\partial_B G(x)$ is the Clarke's generalized Jacobian of G at x, denoted by $\partial G(x)$, see [7]. We say that G is strongly BD-regular at x if every element of $\partial_B G(x)$ is nonsingular.

There is another important concept named semismoothness which was first introduced in [17] for functionals and was extended in [23] to vector-valued functions. Let $G:\Re^n \to \Re^m$ be a locally Lipschitz continuous mapping. We say that G is semismooth at a point $x \in \Re^n$ if G is directionally differentiable and for any $\Delta x \in \Re^n$ and $V \in \partial G(x+\Delta x)$ with $\Delta x \to 0$,

$$G(x+\Delta x) - G(x) - V(\Delta x) = \mathrm{o}(\|\Delta x\|).$$

Furthermore, G is said to be strongly semismooth at x if G is semismooth at x and for any $\Delta x \in \Re^n$ and $V \in \partial G(x+\Delta x)$ with $\Delta x \to 0$,

$$G(x+\Delta x) - G(x) - V(\Delta x) = \mathrm{O}(\|\Delta x\|^2).$$

2.2 Jordan algebra associated with SOC

It is known that $\mathcal{K}^l$ is a closed convex self-dual cone with nonempty interior given by

$$\mathrm{int}(\mathcal{K}^l) := \{x = (x_0, \bar{x}) \in \Re \times \Re^{l-1} | x_0 > \|\bar{x}\|\}.$$

For any $x = (x_0, \bar{x}) \in \Re^l$ and $y = (y_0, \bar{y}) \in \Re^l$, we define their Jordan product as

$$x \circ y = (x^T y, y_0 \bar{x} + x_0 \bar{y}).$$

We write x^2 to mean $x \circ x$ and write $x + y$ to mean the usual component-wise addition of vectors. Moreover, if $x \in \mathcal{K}^l$, there exists a unique vector in $\mathcal{K}^l$, which we denote by $x^{\frac{1}{2}}$, such that $(x^{\frac{1}{2}})^2 = x^{\frac{1}{2}} \circ x^{\frac{1}{2}} = x$. And we recall that each $x = (x_0, \bar{x}) \in \Re \times \Re^{l-1}$ admits a spectral factorization, associated with $\mathcal{K}^l$, of the form

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

where $\lambda_1(x), \lambda_2(x)$ and u_x^1, u_x^2 are the spectral values and the associated spectral vectors of x, respectively, defined by

$$\lambda_i(x) = x_0 + (-1)^i \|\bar{x}\|, \qquad u_x^i = \frac{1}{2}(1, (-1)^i \omega), \quad i = 1, 2$$

with $\omega = \bar{x}/\|\bar{x}\|$ if $\bar{x} \neq 0$ and otherwise ω being any vector in $\Re^{l-1}$ satisfying $\|\omega\| = 1$.

For each $x = (x_0, \bar{x}) \in \Re^l$, define the matrix L_x by

$$L_x := \begin{bmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{bmatrix}, \tag{2.1}$$

which can be viewed as a linear mapping from $\Re^l$ to $\Re^l$.

Lemma 2.1 *The mapping L_x defined by* (2.1) *has the following properties.*

(a) $L_x y = x \circ y$ *and* $L_{x+y} = L_x + L_y$ *for any* $y \in \Re^l$.
(b) $x \in \mathcal{K}^l$ *if and only if* $L_x \succeq O$. *And* $x \in \mathrm{int}\,\mathcal{K}^l$ *if and only if* $L_x \succ O$.
(c) L_x *is invertible whenever* $x \in \mathrm{int}\,\mathcal{K}^l$.

Proof Please see [5, 10]. □

2.3 Complementarity and merit functions associated with SOC

In this subsection, we discuss three reformulations of SOCCP that will play an important role in the sequel of this paper. We deal with the problem SOCCP($\hat{F}$), where $\hat{F}:\Re^n \to \Re^n$ is a certain mapping that has the same structure with F in Sect. 1, that is, $\hat{F}=(\hat{F}_1,\ldots,\hat{F}_q)$ with $\hat{F}_i:\Re^n \to \Re^{n_i}$.

A mapping $\phi:\Re^l \times \Re^l \to \Re^l$ is called an SOC complementarity function associated with the cone $\mathcal{K}^l$ if

$$\phi(x,y)=0 \quad \Leftrightarrow \quad x\in\mathcal{K}^l, \quad y\in\mathcal{K}^l, \quad \langle x,y\rangle=0. \tag{2.2}$$

A popular choice of ϕ is the vector-valued Fischer-Brumeister (FB) function, defined by

$$\phi_{\mathrm{FB}}(x,y) := (x^2+y^2)^{\frac{1}{2}} - x - y, \quad \forall x,y\in\Re^l. \tag{2.3}$$

The function was shown in [10] to satisfy the equivalence (2.2), and therefore its squared norm

$$\psi_{\mathrm{FB}}(x,y) := \frac{1}{2}\|\phi_{\mathrm{FB}}(x,y)\|^2 \tag{2.4}$$

is a merit function. The functions ϕ_{FB} and ψ_{FB} were studied in the literature [5, 26], in which ϕ_{FB} was shown semismooth in [26] whereas ψ_{FB} was proved smooth everywhere in [5]. Due to these favorable properties, the SOCCP($\hat{F}$) can be reformulated as the following nonsmooth system of equations

$$\Phi_{\mathrm{FB}}(x) := \begin{pmatrix} \phi_{\mathrm{FB}}(x_1,\hat{F}_1(x)) \\ \vdots \\ \phi_{\mathrm{FB}}(x_i,\hat{F}_i(x)) \\ \vdots \\ \phi_{\mathrm{FB}}(x_q,\hat{F}_q(x)) \end{pmatrix} = 0, \tag{2.5}$$

where ϕ_{FB} is defined as in (2.3) with a suitable dimension l. Moreover, its squared norm induces a smooth merit function, given by

$$f_{\mathrm{FB}}(x) := \frac{1}{2}\|\Phi_{\mathrm{FB}}(x)\|^2 = \sum_{i=1}^{q}\psi_{\mathrm{FB}}(x_i,\hat{F}_i(x)). \tag{2.6}$$

Lemma 2.2 *The mappings Φ_{FB} and f_{FB} defined in (2.5) and (2.6) have the following properties.*

(a) *If $\hat{F}$ is continuously differentiable, then Φ_{FB} is semismooth.*
(b) *If $\nabla\hat{F}$ is locally Lipschitz continuous, then Φ_{FB} is strongly semismooth.*
(c) *If $\hat{F}$ is continuously differentiable, then f_{FB} is continuously differentiable everywhere.*

(d) *If* $\hat{F}$ *is continuously differentiable and* $\nabla \hat{F}(x)$ *at any* $x \in \Re^n$ *has the Cartesian* P_0*-property, then every stationary point of* $f_{\rm FB}$ *is a solution to the* SOCCP$(\hat{F})$.
(e) *If* $\hat{F}$ *is strongly monotone and* x^* *is a nondegenerate solution of* SOCCP$(\hat{F})$, *i.e.,* $\hat{F}_i(x^*) + x_i^* \in \operatorname{int}\mathcal{K}^{n_i}$ *for all* $i \in \{1, \ldots, q\}$. *Then* $\Phi_{\rm FB}$ *is strongly BD-regular at* x^*.

Proof Items (a) and (b) come from [26, Corollary 3.3] and the fact that the composite of (strongly) semismooth functions is (strongly) semismooth by [9, Theorem 19]. Item (c) was shown by Chen and Tseng, which is an immediate consequences of [5, Proposition 2]. Item (d) is due to [19, Proposition 5.1].

For item (e), since $\nabla \hat{F}(x^*)$ has Cartesian P-property and is positive definite, which can be obtained from the strongly monotonicity of $\hat{F}$, it follows from [20, Proposition 2.1] that the conditions in [19, Theorem 4.1] are satisfied and hence (e) is proved. □

Since the complementarity function $\Phi_{\rm FB}$ and its induced merit function $f_{\rm FB}$ have many useful properties described as in Lemma 2.2, especially when $\hat{F}$ is strongly monotone, they play a crucial role in solving subproblems by using a generalized Newton method in Sect. 4. On the other hand, in [3], Chen extended a new merit function for the NCP to the SOCCP and studied conditions under which the new merit function provides a global error bound and has property of bounded level sets, which play an important role in convergence analysis. In contrast, the merit function $f_{\rm FB}$ lacks these properties. For this reason, we utilize this new merit function to describe the approximation criterion.

Let $\psi_0 : \Re^l \times \Re^l \to \Re_+$ be defined by

$$\psi_0(x, y) := \frac{1}{2}\|\Pi_{\mathcal{K}^l}(x \circ y)\|^2,$$

where the mapping $\Pi_{\mathcal{K}^l}(\cdot)$ denotes the orthogonal projection onto the set $\mathcal{K}^l$. After taking the fixed parameter as in [3], a new merit function is defined as $\psi(x, y) := \psi_0(x, y) + \psi_{\rm FB}(x, y)$, where $\psi_{\rm FB}$ is given by (2.4). Via the new merit function, it was shown that the SOCCP$(\hat{F})$ is equivalent to the following global minimization:

$$\min_{x \in \Re^n} f(x) \quad \text{where } f(x) := \sum_{i=1}^{q} \psi(x_i, \hat{F}_i(x)). \tag{2.7}$$

Here ψ is defined with a suitable dimension l.

The properties about the function f including the error bound property and the boundedness of level sets which are given in [3] are summarized in the following three lemmas.

Lemma 2.3 *Let* f *be defined as in* (2.7).

(a) *If* $\hat{F}$ *is smooth, then* f *is smooth and* $f^{\frac{1}{2}}$ *is uniformly locally Lipschitz continuous on any compact set.*
(b) $f(\zeta) \geq 0$ *for all* $\zeta \in \Re^n$ *and* $f(\zeta) = 0$ *if and only if* ζ *solves the* SOCCP$(\hat{F})$.

(c) *Suppose that the* SOCCP($\hat{F}$) *has at least one solution, then* ζ *is a global minimization of* f *if and only if* ζ *solves the* SOCCP($\hat{F}$).

Proof From [3, Proposition 3.2], we only need to prove that $f^{\frac{1}{2}}$ is Lipschitz continuous on the set $\{y | f(y) = 0\}$. It follows from [3, Proposition 3.1] that if $f(y) = 0$, then $y_i \circ \hat{F}_i(y) = 0$ for all $i = 1, 2, \ldots, q$. Thus, for any $y \in \{y | f(y) = 0\}$, we have

$$\begin{aligned}
&\left| f(x)^{\frac{1}{2}} - f(y)^{\frac{1}{2}} \right| \\
&\quad = f(x)^{\frac{1}{2}} \\
&\quad \le \frac{1}{\sqrt{2}} \sum_{i=1}^{q} \left(\|\Pi_{\mathcal{K}^{n_i}}(x_i \circ \hat{F}_i(x))\| + \|\phi_{\mathrm{FB}}(x_i, \hat{F}_i(x))\| \right) \\
&\quad = \frac{1}{\sqrt{2}} \sum_{i=1}^{q} \Big(\|\Pi_{\mathcal{K}^{n_i}}(x_i \circ \hat{F}_i(x)) - \Pi_{\mathcal{K}^{n_i}}(y_i \circ \hat{F}_i(y))\| \\
&\qquad + \|\phi_{\mathrm{FB}}(x_i, \hat{F}_i(x)) - \phi_{\mathrm{FB}}(y_i, \hat{F}_i(y))\| \Big).
\end{aligned}$$

Noting that the functions $x_i \circ \hat{F}_i(x)$ and $\phi_{\mathrm{FB}}(x_i, \hat{F}_i(x))$ are Lipschitz continuous provided that $\hat{F}$ is smooth. Then from the Lipschitz continuity of ϕ_{FB} and the nonexpansivity of projective mapping onto a convex set, we obtain that $f^{\frac{1}{2}}$ is Lipschitz continuous at y. □

Lemma 2.4 [3, Proposition 4.1] *Suppose that* $\hat{F}$ *is strongly monotone with the modulus* $\rho > 0$ *and* ζ^* *is the unique solution of* SOCCP($\hat{F}$). *Then there exists a scalar* $\tau > 0$ *such that*

$$\tau \|\zeta - \zeta^*\|^2 \le 3\sqrt{2} f(\zeta)^{\frac{1}{2}}, \quad \forall \zeta \in \Re^n, \tag{2.8}$$

where f *is given by* (2.7) *and* τ *can be chosen as*

$$\tau := \frac{\rho}{\max\{\sqrt{2}, \|\hat{F}(\zeta^*)\|, \|\zeta^*\|\}}.$$

Lemma 2.5 [3, Proposition 4.2] *Suppose that* $\hat{F}$ *is monotone and that* SOCCP($\hat{F}$) *is strictly feasible, i.e., there exists* $\hat{\zeta} \in \Re^n$ *such that* $\hat{F}(\hat{\zeta})$, $\hat{\zeta} \in \operatorname{int} \mathcal{K}$. *Then the level set*

$$\mathcal{L}(r) := \{\zeta \in \Re^n | f(\zeta) \le r\}$$

is bounded for all $r \ge 0$, *where* f *is given by* (2.7).

Another SOC complementarity function which we usually call it the natural residual mapping is defined by

$$\phi_{\mathrm{NR}}(x, y) := x - \Pi_{\mathcal{K}^l}(x - y), \quad \forall x, y \in \Re^l,$$

based on which we define the mapping $\Phi_{\rm NR} : \Re^n \to \Re^n$ as

$$\Phi_{\rm NR}(x) := \begin{pmatrix} \phi_{\rm NR}(x_1, \hat{F}_1(x)) \\ \vdots \\ \phi_{\rm NR}(x_i, \hat{F}_i(x)) \\ \vdots \\ \phi_{\rm NR}(x_q, \hat{F}_q(x)) \end{pmatrix}. \tag{2.9}$$

Then it is straightforward to see that SOCCP($\hat{F}$) is equivalent to the system of equations $\Phi_{\rm NR}(x) = 0$.

Lemma 2.6 *The mapping $\Phi_{\rm NR}$ defined as in (2.9) has the following properties.*

(a) *If $\hat{F}$ is continuously differentiable, then $\Phi_{\rm NR}$ is semismooth.*
(b) *If $\nabla\hat{F}$ is locally Lipschitz continuous, then $\Phi_{\rm NR}$ is strongly semismooth.*
(c) *If $\nabla\hat{F}(x)$ is positive definite, then every $V \in \partial_B \Phi_{\rm NR}(x)$ is nonsingular, i.e., $\Phi_{\rm NR}$ is strongly BD-regular at x.*

Proof Items (a) and (b) are obvious after combining [6, Proposition 4.3] and [9, Theorem 19]. Note that these two items are also proved in [12] in a different approach. The proof of item (c) is similar to that in [27] and [30] for a more general setting, and we omit it. □

From Lemma 2.6(c) we know that the natural residual mapping $\Phi_{\rm NR}$ is strongly BD-regular under weaker conditions than $\Phi_{\rm FB}$. In view of this, we will use $\Phi_{\rm NR}$ to explore the condition of superlinear convergence of PPA in Sect. 3.

2.4 Proximal point algorithm

Let $\mathcal{T} : \Re^n \rightrightarrows \Re^n$ be a set-valued mapping defined by

$$\mathcal{T}(x) := F(x) + \mathcal{N}_{\mathcal{K}}(x). \tag{2.10}$$

Then $\mathcal{T}$ is a maximal monotone mapping and SOCCP(F) defined by (1.1) is equivalent to the problem of finding a point x such that

$$0 \in \mathcal{T}(x).$$

The proximal point algorithm generates, for any starting point x^0, a sequence $\{x^k\}$ by the approximate rule:

$$x^{k+1} \approx P_k(x^k),$$

where $P_k := (I + \frac{1}{c_k}\mathcal{T})^{-1}$ is a single-valued mapping from $\Re^n$ to $\Re^n$, $\{c_k\}$ is some sequence of positive real numbers, and $x^{k+1} \approx P_k(x^k)$ means that x^{k+1} is an approximation to $P_k(x^k)$. Accordingly, for SOCCP(F), $P_k(x^k)$ is given by

$$P_k(x^k) = \left(I + \frac{1}{c_k}(F + \mathcal{N}_{\mathcal{K}})\right)^{-1}(x^k),$$

from which we have

$$P_k(x^k) \in \mathrm{SOL}(\mathrm{SOCCP}(F^k)),$$

where F^k is defined by (1.4) and $\mathrm{SOL}(\mathrm{SOCCP}(F^k))$ is the solution set of $\mathrm{SOCCP}(F^k)$. Therefore, x^{k+1} is given by an approximate solution of $\mathrm{SOCCP}(F^k)$. Two general criteria for the approximate calculation of $P_k(x^k)$ proposed by Rockafellar [24] are as follows:

Criterion 2.1

$$\|x^{k+1} - P_k(x^k)\| \le \varepsilon_k, \qquad \sum_{k=0}^{\infty} \varepsilon_k < \infty.$$

Criterion 2.2

$$\|x^{k+1} - P_k(x^k)\| \le \eta_k \|x^{k+1} - x^k\|, \qquad \sum_{k=0}^{\infty} \eta_k < \infty.$$

Results on the convergence of the proximal point algorithm have already been studied in [15, 24] from which we know that Criterion 2.1 guarantees global convergence while Criterion 2.2, which is rather restrictive, ensures superlinear convergence.

Theorem 2.1 *Let $\{x^k\}$ be any sequence generated by the PPA under Criterion 2.1 with $\{c_k\}$ bounded. Suppose* $\mathrm{SOCCP}(F)$ *has at least one solution. Then $\{x^k\}$ converges to a solution x^* of* $\mathrm{SOCCP}(F)$.

Proof This can be proved by similar arguments as in [24, Theorem 1]. □

Theorem 2.2 *Suppose the solution set $\bar{X}$ of* $\mathrm{SOCCP}(F)$ *is nonempty, and let $\{x^k\}$ be any sequence generated by PPA with Criterions 2.1 and 2.2 and $c_k \to 0$. Let us also assume that*

$$\begin{aligned} &\exists \delta > 0, \quad \exists C > 0, \\ &\text{s.t.} \quad \mathrm{dist}(x, \bar{X}) \le C\|w\| \quad \textit{whenever } x \in \mathcal{T}^{-1}(\omega) \textit{ and } \|\omega\| \le \delta. \end{aligned} \tag{2.11}$$

Then the sequence $\{\mathrm{dist}(x^k, \bar{X})\}$ *converges to* 0 *superlinearly.*

Proof This can be also verified by similar arguments as in [15, Theorem 2.1]. □

3 A proximal point algorithm for solving SOCCP

Based on the previous discussion, in this section we describe PPA for solving $\mathrm{SOCCP}(F)$ as defined in (1.1) where F is smooth and monotone. We first illustrate the related mappings that will be used in the remainder of this paper.

The mappings $\Phi_{\rm NR}$, $\Phi_{\rm FB}$, $f_{\rm FB}$ are defined by (2.9), (2.5) and (2.6), respectively, where the mapping $\hat{F}$ is substituted by F. And the functions f^k, $f_{\rm FB}^k$ and $\Phi_{\rm FB}^k$ are defined by (2.7), (2.6) and (2.5), respectively, where the mapping $\hat{F}$ is replaced by F^k which is given by (1.4), i.e.,

$$\Phi_{\rm NR}(x) := \begin{pmatrix} \phi_{\rm NR}(x_1, F_1(x)) \\ \vdots \\ \phi_{\rm NR}(x_i, F_i(x)) \\ \vdots \\ \phi_{\rm NR}(x_q, F_q(x)) \end{pmatrix}, \qquad \Phi_{\rm FB}^k(x) := \begin{pmatrix} \phi_{\rm FB}(x_1, F_1^k(x)) \\ \vdots \\ \phi_{\rm FB}(x_i, F_i^k(x)) \\ \vdots \\ \phi_{\rm FB}(x_q, F_q^k(x)) \end{pmatrix},$$

$$\Phi_{\rm FB}(x) := \begin{pmatrix} \phi_{\rm FB}(x_1, F_1(x)) \\ \vdots \\ \phi_{\rm FB}(x_i, F_i(x)) \\ \vdots \\ \phi_{\rm FB}(x_q, F_q(x)) \end{pmatrix},$$

$$f^k(x) := \sum_{i=1}^{q} \psi(x_i, F_i^k(x)), \qquad f_{\rm FB}(x) := \frac{1}{2}\|\Phi_{\rm FB}(x)\|^2,$$

$$f_{\rm FB}^k(x) := \frac{1}{2}\|\Phi_{\rm FB}^k(x)\|^2.$$

Now we are in a position to describe the proximal point algorithm for solving Problem (1.1).

Algorithm 3.1

Step 0. Choose parameters $\alpha \in (0,1)$, $c_0 \in (0,1)$ and an initial point $x^0 \in \Re^n$. Set $k := 0$.
Step 1. If x^k satisfies $f_{\rm FB}(x^k) = 0$, then stop.
Step 2. Let $F^k(x) = F(x) + c_k(x - x^k)$. Get an approximation solution x^{k+1} of $\mathrm{SOCCP}(F^k)$ that satisfies the condition

$$f^k(x^{k+1}) \le \frac{c_k^6 \min\{1, \|x^{k+1} - x^k\|^4\}}{18\max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}^2}. \tag{3.1}$$

Step 3. Set $c_{k+1} = \alpha c_k$ and $k := k+1$. Go to Step 1.

Theorem 3.1 *Let $\bar{X}$ be the solution set of* $\mathrm{SOCCP}(F)$. *If $\bar{X} \neq \emptyset$, then the sequence $\{x^k\}$ generated by Algorithm 3.1 converges to a solution x^* of* $\mathrm{SOCCP}(F)$.

Proof From Theorem 2.1, it suffices to prove that such $\{x^k\}$ satisfies Criterion 2.1. Since F^k is strongly monotone with modulus $c_k > 0$ and $P_k(x^k)$ is the unique solution of SOCCP(F^k), it follows from Lemma 2.4 that

$$\|x^{k+1} - P_k(x^k)\|^2 \leq \frac{3\sqrt{2}}{c_k} \max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\} f^k(x^{k+1})^{\frac{1}{2}}, \quad (3.2)$$

which together with (3.1) implies

$$\|x^{k+1} - P_k(x^k)\| \leq c_k. \quad (3.3)$$

□

To obtain superlinear convergence properties, we need to give the following assumption which will be connected to the condition (2.11) in Theorem 2.2.

Assumption 3.1 $\|x - \Pi_{\mathcal{K}}(x - F(x))\|$ provides a local error bound for SOCCP(F), that is, there exist positive constants $\bar{\delta}$ and $\bar{C}$ such that

$$\begin{aligned} &\operatorname{dist}(x, \bar{X}) \leq \bar{C}\|x - \Pi_{\mathcal{K}}(x - F(x))\|, \\ &\quad \text{for all } x \text{ with } \|x - \Pi_{\mathcal{K}}(x - F(x))\| \leq \bar{\delta}, \end{aligned} \quad (3.4)$$

where $\bar{X}$ denotes the solution set of SOCCP(F).

The following lemma can help us to understand Assumption 3.1 as it implies conditions under which Assumption 3.1 holds.

Lemma 3.1 [22, Proposition 3] *If a Lipschitz continuous mapping H is strongly BD-regular at x^*, then there is a neighborhood N of x^* and a positive constant α, such that $\forall x \in N$ and $V \in \partial_B H(x)$, V is nonsingular and $\|V^{-1}\| \leq \alpha$. If, furthermore, H is semismooth at x^* and $H(x^*) = 0$, then there exists a neighborhood N' of x^* and a positive constant β such that $\forall x \in N'$, $\|x - x^*\| \leq \beta\|H(x)\|$.*

Note that when $\nabla F(x)$ is positive definite at one solution x of SOCCP(F), Assumption 3.1 holds by Lemmas 2.6 and 3.1.

Theorem 3.2 *Let $\mathcal{T}$ be defined by (2.10). If $\bar{X} \neq \emptyset$, then Assumption 3.1 implies condition (2.11), that is, there exist $\delta > 0$ and $C > 0$ such that*

$$\operatorname{dist}(x, \bar{X}) \leq C\|\omega\|,$$

whenever $x \in \mathcal{T}^{-1}(\omega)$ and $\|\omega\| \leq \delta$.

Proof For all $x \in \mathcal{T}^{-1}(\omega)$ we have

$$w \in \mathcal{T}(x) = F(x) + \mathcal{N}_{\mathcal{K}}(x).$$

Therefore there exists $v \in \mathcal{N}_{\mathcal{K}}(x)$ such that $w = F(x) + v$. Because $\mathcal{K}$ is a convex set, it is easy to obtain that

$$\Pi_{\mathcal{K}}(x + v) = x. \quad (3.5)$$

Noting that the projective mapping onto a convex set is nonexpansive, we have from (3.5) that

$$\|x - \Pi_{\mathcal{K}}(x - F(x))\| = \|\Pi_{\mathcal{K}}(x + v) - \Pi_{\mathcal{K}}(x - F(x))\| \leq \|v + F(x)\| = \|\omega\|.$$

From Assumption 3.1 and letting $C = \bar{C}$, $\delta = \bar{\delta}$ yield the desired condition (2.11). $\square$

The following theorem gives the superlinear convergence of Algorithm 3.1, whose proof is based on Theorem 3.2 and can be obtained in the same way as Theorem 3.1. We omit the proof here.

Theorem 3.3 *Suppose that Assumption 3.1 holds. Let $\{x^k\}$ be generated by Algorithm 3.1. Then the sequence $\{\mathrm{dist}(x^k, \bar{X})\}$ converges to 0 superlinearly.*

Although we have obtained the global and superlinear convergence properties of Algorithm 3.1 under mild conditions, this does not mean that Algorithm 3.1 is practically efficient, as it says nothing about how to obtain an approximation solution of the strongly monotone second-order cone complementarity problem in Step 2 satisfying (3.1) and what is the cost. We will give the answer in the next section.

4 Generalized Newton method

In this section, we introduce the generalized Newton method proposed by De Luca, Facchinei, and Kanzow [14] for solving the subproblems in Step 2 of Algorithm 3.1. As mentioned earlier, for each fixed k, Problem (1.3) is equivalent to the following nonsmooth equation

$$\Phi_{\mathrm{FB}}^k(x) = 0. \tag{4.1}$$

Now we describe as below the generalized Newton method for solving the nonsmooth system (4.1), which is employed from what was introduced in [29] for solving NCP.

Algorithm 4.1 (Generalized Newton method for SOCCP(F^k))

Step 0. Choose $\beta \in (0, \frac{1}{2})$ and an initial point $x^0 \in \Re^n$. Set $j := 0$.
Step 1. If $\|\Phi_{\mathrm{FB}}^k(x^j)\| = 0$, then stop.
Step 2. Select an element $V^j \in \partial_B \Phi_{\mathrm{FB}}^k(x^j)$. Find the solution d^j of the system

$$V^j d = -\Phi_{\mathrm{FB}}^k(x^j). \tag{4.2}$$

Step 3. Find the smallest nonnegative integer i_j such that

$$f_{\mathrm{FB}}^k(x^j + 2^{-i_j} d^j) \leq (1 - \beta 2^{1-i_j}) f_{\mathrm{FB}}^k(x^j).$$

Step 4. Set $x^{j+1} := x^j + 2^{-i_j} d^j$ and $j := j + 1$. Go to Step 1.

To guarantee the descent sequence of f_{FB}^k must have an accumulation point, Pan and Chen [19] give the following condition under which the coerciveness of f_{FB}^k for SOCCP(F^k) can be established.

Condition 4.1 For any sequence $\{x^j\} \subseteq \Re^n$ satisfying $\|x^j\| \to +\infty$, if there exists an index $i \in \{1, 2, \ldots, q\}$ such that $\{\lambda_1(x_i^j)\}$ and $\{\lambda_1(F_i(x^j))\}$ are bounded below, and $\lambda_2(x_i^j), \lambda_2(F_i(x^j)) \to +\infty$, then

$$\limsup_{j\to\infty} \left\langle \frac{x_i^j}{\|x_i^j\|}, \frac{F_i(x^j)}{\|F_i(x^j)\|} \right\rangle > 0.$$

As F^k is strongly monotone, it then has the uniform Cartesian P-property. From [19], we have the following theorem.

Theorem 4.1 [19, Proposition 5.2] *For* SOCCP(F^k), *if Condition* 4.1 *holds, then the merit function* f_{FB}^k *is coercive.*

To obtain the quadratic convergence of Algorithm 4.1, we need the following two Assumptions which are also essential in the follow-up work.

Assumption 4.1 F is continuously differentiable function with a local Lipschitz Jacobian.

Assumption 4.2 The limit point x^* of the sequence $\{x^k\}$ generated by Algorithm 3.1 is nondegenerate, i.e., $x_i^* + F_i(x^*) \in \operatorname{int} \mathcal{K}^{n_i}$ holds for all $i \in \{1, \ldots, q\}$.

Note that when k is large enough, the unique solution $P_k(x^k)$ of SOCCP(F^k) is nondegenerate, that is, $(P_k(x^k))_i + F_i^k(P_k(x^k)) \in \operatorname{int} \mathcal{K}^{n_i}$ holds for all $i \in \{1, \ldots, q\}$. Because F^k is strongly monotone, we immediately have the following convergence theorem from Lemma 2.2 and [14, Theorem 3.1].

Theorem 4.2 *If the sequence* $\{x^j\}$ *generated by Algorithm* 4.1 *has an accumulation point and Assumptions* 4.1 *and* 4.2 *hold. Then* $\{x^j\}$ *globally converges to the unique solution* $P_k(x^k)$ *and the rate is quadratic.*

Noting that the condition (3.1) in Algorithm 3.1 is equivalent to the following two criteria:

Criterion 4.1

$$f^k(x^{k+1}) \le \frac{c_k^6}{18 \max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}^2}.$$

Criterion 4.2

$$f^k(x^{k+1}) \le \frac{c_k^6 \|x^{k+1} - x^k\|^4}{18 \max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}^2}.$$

It follows from Sect. 3 that Criterion 4.1 guarantees global convergence, while Criterion 4.2, which is rather restrictive, ensures superlinear convergence of PPA.

Next, we give conditions under which a single Newton step of generalized Newton method can generate a point eventually that satisfies the following two criteria for any given $r \in (0, 1)$, i.e., Criterion 4.1 and the following criterion:

Criterion 4.2(r)

$$f^k(x^{k+1}) \le \frac{c_k^6 \|x^{k+1} - x^k\|^{4(1-r)}}{18 \max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}^2}.$$

Thereby the PPA can be practically efficient, which we call that Algorithm 3.1 has approximate genuine superlinear convergence. Firstly, we have the following two lemmas, which indicate the relationship between $\|x^k - P_k(x^k)\|$ and $\mathrm{dist}(x^k, \bar{X})$.

Lemma 4.1 *If* SOCCP(F) *is strictly feasible. Then, for sufficiently large k, there exists a constant $B_1 \ge 2$ such that*

$$2 \le \max\left\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\right\}^2 \le B_1. \tag{4.3}$$

Proof From Lemma 2.5, we obtain that the solution set $\bar{X}$ of SOCCP(F) is bounded, which implies the boundedness of $F(\bar{X})$. Let $m_1 > 0$ be such that

$$\max\left\{\sup_{x\in\bar{X}} \|x\|, \sup_{x\in\bar{X}} \|F(x)\|\right\} \le m_1.$$

Since $c_k \to 0$, it follows from Theorem 3.1 that the two sequences $\{x^k\}$ and $\{P_k(x^k)\}$ have the same limit point $x^* \in \bar{X}$. Then there exists a positive constant m_2 such that

$$\|P_k(x^k) - x^*\| \le m_2,$$

and

$$\|F(P_k(x^k)) - F(x^*)\| \le m_2,$$

when k is large enough. Thus, the following two inequalities

$$\begin{aligned}\|F^k(P_k(x^k))\| &= \|F(P_k(x^k)) + c_k(P_k(x^k) - x^k)\| \\ &\le \|F(P_k(x^k))\| + c_k\|P_k(x^k) - x^k\| \\ &\le \|F(x^*)\| + 2m_2\end{aligned}$$

and

$$\|P_k(x^k)\| \le \|x^*\| + m_2$$

hold for sufficiently large k. Let $B_1 = \max\{2, (m_1 + 2m_2)^2\}$, we complete the proof. □

Lemma 4.2 *If* SOCCP(F) *is strictly feasible, then for sufficiently large* k, *there exists a constant* $B_2 > 0$ *such that*

$$\|x^k - P_k(x^k)\| \le \frac{B_2}{\sqrt{c_k}} \mathrm{dist}(x^k, \bar{X})^{\frac{1}{2}}.$$

Proof Let $\bar{x}^k$ be the nearest point in $\bar{X}$ from x^k. From [8, Theorem 2.3.5] we know that $\bar{X}$ is convex, and hence the mapping $\Pi_{\bar{X}}(\cdot)$ is nonexpansive. Therefore,

$$\|\bar{x}^k - x^*\| = \|\Pi_{\bar{X}}(x^k) - \Pi_{\bar{X}}(x^*)\| \le \|x^k - x^*\|.$$

Since $\{x^k\}$ is bounded, so is $\{\bar{x}^k\}$. Let $\hat{X}$ be a bounded set containing $\{x^k\}$ and $\{\bar{x}^k\}$. From Lemma 2.3, we know that $f^{\frac{1}{2}}$ is uniformly Lipschitz continuous on $\hat{X}$. Then there exists $L_1 > 0$ such that

$$\left(f(x^k)\right)^{\frac{1}{2}} = \left(f(x^k)\right)^{\frac{1}{2}} - \left(f(\bar{x}^k)\right)^{\frac{1}{2}} \le L_1^2 \|x^k - \bar{x}^k\| = L_1^2 \,\mathrm{dist}(x^k, \bar{X}),$$

which implies that

$$\left(f(x^k)\right)^{\frac{1}{4}} \le L_1 \,\mathrm{dist}(x^k, \bar{X})^{\frac{1}{2}}.$$

It follows from Lemma 2.4 that

$$\|x^k - P_k(x^k)\|^2 \le \frac{3\sqrt{2}}{\tau_k} f^k(x^k)^{\frac{1}{2}},$$

where

$$\tau_k = \frac{c_k}{\max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}},$$

which together with Lemma 4.1 yields

$$\frac{\sqrt{2}}{c_k} \le \frac{1}{\tau_k} \le \frac{\sqrt{B_1}}{c_k}.$$

Hence, we have

$$\|x^k - P_k(x^k)\| \le \left(\frac{3\sqrt{2B_1}}{c_k}\right)^{\frac{1}{2}} \left(f^k(x^k)\right)^{\frac{1}{4}}.$$

On the other hand, since $F^k(x^k) = F(x^k)$, we know $f^k(x^k) = f(x^k)$ and hence

$$\|x^k - P_k(x^k)\| \le \left(\frac{3\sqrt{2B_1}}{c_k}\right)^{\frac{1}{2}} \left(f(x^k)\right)^{\frac{1}{4}} \le \left(\frac{3\sqrt{2B_1}}{c_k}\right)^{\frac{1}{2}} L_1 \,\mathrm{dist}(x^k, \bar{X})^{\frac{1}{2}}.$$

Then, letting $B_2 = (3\sqrt{2B_1})^{\frac{1}{2}} L_1$ leads to the desired inequality. □

The next three lemmas give the relationship between $\|x_N^k - P_k(x^k)\|$ and $\|x^k - P_k(x^k)\|$ which is the key to the main result in this section. One attention we should pay to is that we will not be able to obtain the inequality in Lemma 4.3 without twice continuously differentiability. The reason is as explained in [29, Remark 4.1]. To this end, we make the following assumption.

Assumption 4.3 F is twice continuously differentiable.

Lemma 4.3 *Suppose that Assumptions 4.2 and 4.3 hold. Then Φ_{FB}^k is twice continuously differentiable in a neighborhood of x^k for sufficiently large k, and there exists a positive constant B_3 such that*

$$\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)(x^k - P_k(x^k)) - \Phi_{\mathrm{FB}}^k(x^k) + \Phi_{\mathrm{FB}}^k(P_k(x^k))\| \le B_3\|x^k - P_k(x^k)\|^2.$$

Proof It is obvious that when F is twice continuously differentiable and Assumption 4.2 holds, Φ_{FB}^k is twice continuously differentiable near x^k and $P_k(x^k)$ when k is large enough. Then from the second order Taylor expansion and the Lipschitz continuity of $\nabla\Phi_{\mathrm{FB}}^k$ near x^k, there exist positive constants m_3, m_4 such that when k is sufficiently large,

$$\|\Phi_{\mathrm{FB}}^k(x^k) - \Phi_{\mathrm{FB}}^k(P_k(x^k)) - \mathcal{J}\Phi_{\mathrm{FB}}^k(P_k(x^k))(x^k - P_k(x^k))\| \le m_3\|x^k - P_k(x^k)\|^2,$$
$$\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)(x^k - P_k(x^k)) - \mathcal{J}\Phi_{\mathrm{FB}}^k(P_k(x^k))(x^k - P_k(x^k))\| \le m_4\|x^k - P_k(x^k)\|^2.$$

Let $B_3 = m_3 + m_4$, we have for sufficiently large k

$$\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)(x^k - P_k(x^k)) - \Phi_{\mathrm{FB}}^k(x^k) + \Phi_{\mathrm{FB}}^k(P_k(x^k))\| \le B_3\|x^k - P_k(x^k)\|^2. \qquad \square$$

Now let us denote

$$x_N^k := x^k - V_k^{-1}\Phi_{\mathrm{FB}}^k(x^k), \qquad V_k \in \partial_B\Phi_{\mathrm{FB}}^k(x^k). \tag{4.4}$$

Then x_N^k is a point produced by a single Newton iteration of Algorithm 4.1 with the initial point x^k.

Lemma 4.4 *Suppose that Assumption 4.2 holds, then Φ_{FB}^k is differentiable at x^k and the Jacobian $\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)$ is nonsingular for sufficiently large k.*

Proof Let

$$z_i(x) = (x_i^2 + (F_i^k(x))^2)^{\frac{1}{2}}, \tag{4.5}$$

for each $i \in \{1, \ldots, q\}$. From Assumption 4.2 and [19, Lemma 4.2], we have that for every $i \in \{1, \ldots, q\}$,

$$x_i^k + F_i^k(x^k) \in \operatorname{int}\mathcal{K}^{n_i},$$

and

$$(x_i^k)^2 + (F_i^k(x^k))^2 \in \operatorname{int}\mathcal{K}^{n_i},$$

when k is large enough. Thus, Φ_{FB}^k is differentiable at x^k by [19, Proposition 4.2] when k is large enough, and

$$\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^T = \nabla F^k(x^k)(A(x^k) - I) + (B(x^k) - I), \tag{4.6}$$

where $A(x^k) = \mathrm{diag}(A_i(x^k), i = 1, \dots, q)$ and $B(x^k) = \mathrm{diag}(B_i(x^k), i = 1, \dots, q)$ with $A_i(x^k) = L_{F_i^k(x^k)} L_{z_i(x^k)}^{-1}$ and $B_i(x^k) = L_{x_i^k} L_{z_i(x^k)}^{-1}$. For any fixed k, the index sets

$$\mathcal{I}^k := \{i \in \{1, \dots, q\} | F_i^k(x^k) = 0, x_i^k \in \mathrm{int}\,\mathcal{K}^{n_i}\}, \tag{4.7}$$

$$\mathcal{J}^k := \{i \in \{1, \dots, q\} | x_i^k = 0, F_i^k(x^k) \in \mathrm{int}\,\mathcal{K}^{n_i}\}, \tag{4.8}$$

$$\mathcal{B}^k := \{1, \dots, q\} \setminus \{\mathcal{I}^k \cup \mathcal{J}^k\} \tag{4.9}$$

form a partition of $\{1, \dots, q\}$. After rearranging the matrices appropriately, $\nabla F^k(x^k)$ can be rewritten as

$$\nabla F^k(x^k) = \begin{pmatrix} \nabla F^k_{\mathcal{I}\mathcal{I}} & \nabla F^k_{\mathcal{I}\mathcal{B}} & \nabla F^k_{\mathcal{I}\mathcal{J}} \\ \nabla F^k_{\mathcal{B}\mathcal{I}} & \nabla F^k_{\mathcal{B}\mathcal{B}} & \nabla F^k_{\mathcal{B}\mathcal{J}} \\ \nabla F^k_{\mathcal{J}\mathcal{I}} & \nabla F^k_{\mathcal{J}\mathcal{B}} & \nabla F^k_{\mathcal{J}\mathcal{J}} \end{pmatrix}.$$

For simplicity, we omit the notation x^k in the functions and we substitute $\mathcal{B}$ for $\mathcal{B}^k$ here, and also in the sequel of the proof. Thus, the nonsingularity of $\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)$ is equivalent to showing the nonsingularity of the following partitioned form

$$C = \begin{pmatrix} -\nabla F^k_{\mathcal{I}\mathcal{I}} & \nabla F^k_{\mathcal{I}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) & 0_{\mathcal{I}\mathcal{J}} \\ -\nabla F^k_{\mathcal{B}\mathcal{I}} & \nabla F^k_{\mathcal{B}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) + (B_{\mathcal{B}} - I_{\mathcal{B}}) & 0_{\mathcal{B}\mathcal{J}} \\ -\nabla F^k_{\mathcal{J}\mathcal{I}} & \nabla F^k_{\mathcal{J}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) & -I_{\mathcal{J}} \end{pmatrix},$$

where $I_{\mathcal{B}} = \mathrm{diag}(I_i, i \in \mathcal{B})$ with I_i being an $n_i \times n_i$ identity matrix, $A_{\mathcal{B}} = \mathrm{diag}(A_i, i \in \mathcal{B})$ and $B_{\mathcal{B}} = \mathrm{diag}(B_i, i \in \mathcal{B})$. It is not hard to see that C is nonsingular if and only if

$$\hat{C} = \begin{pmatrix} -\nabla F^k_{\mathcal{I}\mathcal{I}} & \nabla F^k_{\mathcal{I}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) \\ -\nabla F^k_{\mathcal{B}\mathcal{I}} & \nabla F^k_{\mathcal{B}\mathcal{B}}(A_{\mathcal{B}} - I_{\mathcal{B}}) + (B_{\mathcal{B}} - I_{\mathcal{B}}) \end{pmatrix}$$

is nonsingular. Suppose that the vector y satisfies the system

$$\hat{C}y = \hat{C}\begin{pmatrix} y_{\mathcal{I}} \\ y_{\mathcal{B}} \end{pmatrix} = 0. \tag{4.10}$$

We only need to argue that y is the zero vector. System (4.10) can be rewritten as the following two equations

$$\nabla F^k_{\mathcal{I}\mathcal{I}} y_{\mathcal{I}} + \nabla F^k_{\mathcal{I}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = 0,$$

$$\nabla F^k_{\mathcal{B}\mathcal{I}} y_{\mathcal{I}} + \nabla F^k_{\mathcal{B}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}}) y_{\mathcal{B}} = -(I_{\mathcal{B}} - B_{\mathcal{B}}) y_{\mathcal{B}}.$$

Since $\nabla F^k(x^k)$ is positive definite, we have that $\nabla F^k_{\mathcal{I}\mathcal{I}}$ is nonsingular. Then we obtain that

$$y_{\mathcal{I}} = -(\nabla F^k_{\mathcal{I}\mathcal{I}})^{-1}\nabla F^k_{\mathcal{I}\mathcal{B}}(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}, \tag{4.11}$$

$$(\nabla F^k_{\mathcal{B}\mathcal{B}} - \nabla F^k_{\mathcal{B}\mathcal{I}}(\nabla F^k_{\mathcal{I}\mathcal{I}})^{-1}\nabla F^k_{\mathcal{I}\mathcal{B}})(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} = -(I_{\mathcal{B}} - B_{\mathcal{B}})y_{\mathcal{B}}. \tag{4.12}$$

Suppose that $y_{\mathcal{B}} \neq 0$, then there exists $i \in \mathcal{B}$ such that $y_i \neq 0$. If $(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} = 0$, then

$$(I_i - A_i)y_i = 0, \qquad (I_i - B_i)y_i = 0.$$

This means that

$$(2I_i - A_i - B_i)y_i = 0.$$

Since for each i, we have

$$\begin{aligned}
2I_i - A_i - B_i &= 2I_i - L_{F_i^k(x^k)}L^{-1}_{z_i(x^k)} - L_{x_i^k}L^{-1}_{z_i(x^k)} \\
&= [2L_{z_i(x^k)} - L_{F_i^k(x^k)} - L_{x_i^k}]L^{-1}_{z_i(x^k)} \\
&= L_{2z_i(x^k) - F_i^k(x^k) - x_i^k}L^{-1}_{z_i(x^k)}
\end{aligned} \tag{4.13}$$

and

$$4z_i^2(x^k) - (F_i^k(x^k) + x_i^k)^2 = 2z_i^2(x^k) + (F_i^k(x^k) - x_i^k)^2 \in \operatorname{int}\mathcal{K}^{n_i} \tag{4.14}$$

for sufficiently large k, using [10, Proposition 3.4] yields $2z_i(x^k) - F_i^k(x^k) - x_i^k \in \operatorname{int}\mathcal{K}^{n_i}$ and we have that

$$L_{2z_i(x^k) - F_i^k(x^k) - x_i^k} \succ O$$

from Lemma 2.1. Therefore, $2I_i - A_i - B_i$ is nonsingular for each $i \in \{1, \ldots, q\}$. This implies that $(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}} \neq 0$. On the other hand, it follows from Lemma 4.1 in [19] that for each $i \in \mathcal{B}$,

$$\langle [(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i, [(I_{\mathcal{B}} - B_{\mathcal{B}})y_{\mathcal{B}}]_i \rangle \geq 0,$$

which together with (4.12) means that

$$\langle [(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i, [(\nabla F^k_{\mathcal{B}\mathcal{B}} - \nabla F^k_{\mathcal{B}\mathcal{I}}(\nabla F^k_{\mathcal{I}\mathcal{I}})^{-1}\nabla F^k_{\mathcal{I}\mathcal{B}})(I_{\mathcal{B}} - A_{\mathcal{B}})y_{\mathcal{B}}]_i \rangle \leq 0. \tag{4.15}$$

Note that F^k is strongly monotone, hence F^k has the uniform Cartesian P-property, which implies that for every $x \in \Re^n$, $\nabla F^k(x)$ has Cartesian P-property. Since $(\nabla F^k_{\mathcal{B}\mathcal{B}} - \nabla F^k_{\mathcal{B}\mathcal{I}}(\nabla F^k_{\mathcal{I}\mathcal{I}})^{-1}\nabla F^k_{\mathcal{I}\mathcal{B}})$ is exactly the Schur-complement of $\nabla F^k_{\mathcal{I}\mathcal{I}}$ in the matrix

$$\begin{pmatrix} \nabla F^k_{\mathcal{I}\mathcal{I}} & \nabla F^k_{\mathcal{I}\mathcal{B}} \\ \nabla F^k_{\mathcal{B}\mathcal{I}} & \nabla F^k_{\mathcal{B}\mathcal{B}} \end{pmatrix},$$

from [20, Proposition 2.1] and the nonsingularity of $\nabla F^k_{\mathcal{II}}$, together with the fact that every principal block matrix of a matrix with Cartesian P-property must have the Cartesian P-property, we obtain that the matrix $(\nabla F^k_{\mathcal{BB}} - \nabla F^k_{\mathcal{BI}}(\nabla F^k_{\mathcal{II}})^{-1}\nabla F^k_{\mathcal{IB}})$ has Cartesian P-property. This leads to a contradiction with (4.15). Thus, we have $y = 0$ and the proof is complete. □

Assumption 4.4 For every sequence $\{x^k\}$ that converges to x^*, we have that either $\mathcal{B}^k = \emptyset$ or

$$(A(x^k)^T - I)\nabla F^k(x^k)^T(B(x^k) - I) \succeq O,$$

when k is large enough, where $\mathcal{B}^k$, $A(x^k)$, $B(x^k)$ are defined in the proof of Lemma 4.4.

Note that when SOCCP(F^k) defined as in (1.3) reduces to NCP, Assumption 4.4 holds automatically because $\mathcal{B}^k = \emptyset$ for sufficiently large k when the limit point x^* is nondegenerate.

Lemma 4.5 *Suppose that Assumptions 4.2–4.4 hold. Then there exists $B_4 > 0$ such that*

$$\|x^k_N - P_k(x^k)\| \le \frac{B_4\|x^k - P_k(x^k)\|^2}{c_k},$$

for sufficiently large k.

Proof From the definition of F^k, we have that $\nabla F^k(x^k)^T$ is positive definite and

$$\langle v, \nabla F^k(x^k)^T v\rangle \ge c_k\|v\|^2,$$

for all $v \in \Re^n$. Let v be an arbitrary vector in $\Re^n$ such that $\|v\| = 1$. Then since

$$\langle v, \nabla F^k(x^k)^T v\rangle \ge c_k,$$

there exists an index $i_0 \in \{1, \dots, q\}$ such that

$$v_{i_0}^T[\nabla F^k(x^k)^T v]_{i_0} \ge \frac{c_k}{q}. \tag{4.16}$$

Since

$$\begin{aligned} v_{i_0}^T[\nabla F^k(x^k)^T v]_{i_0} &\le \|v_{i_0}\|\|[\nabla F^k(x^k)^T v]_{i_0}\| \le \|v_{i_0}\|\|\nabla F^k(x^k)^T v\| \\ &\le \|v_{i_0}\|\|\nabla F^k(x^k)^T\|, \end{aligned}$$

and

$$v_{i_0}^T[\nabla F^k(x^k)^T v]_{i_0} \le \|v_{i_0}\|\|[\nabla F^k(x^k)^T v]_{i_0}\| \le \|[\nabla F^k(x^k)^T v]_{i_0}\|,$$

it follows from (4.16) that

$$\|v_{i_0}\| \ge \frac{c_k}{q\|\nabla F^k(x^k)^T\|}, \qquad \|[\nabla F^k(x^k)^T v]_{i_0}\| \ge \frac{c_k}{q}. \tag{4.17}$$

For any fixed k that is large enough, let the three index sets be defined as (4.7)–(4.9). We consider the following two cases.

Case 1. $\mathcal{B}^k = \emptyset$. This means that either $i_0 \in \mathcal{I}^k$ or $i_0 \in \mathcal{J}^k$. For every $i \in \{1, \dots, q\}$, we have

$$\begin{aligned}\|[\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)v]_i\| = \|(&L_{z_i(x^k)}^{-1} L_{F_i^k(x^k)} - I_i)[\nabla F^k(x^k)^T v]_i \\ &+ (L_{z_i(x^k)}^{-1} L_{x_i^k} - I_i)v_i\|,\end{aligned}$$

where $z_i(x^k)$ is defined in (4.5). If $i_0 \in \mathcal{I}^k$, then

$$\|[\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)v]_{i_0}\| = \|[\nabla F^k(x^k)^T v]_{i_0}\| \geq \frac{c_k}{q}.$$

And if $i_0 \in \mathcal{J}^k$, then

$$\|[\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)v]_{i_0}\| = \|v_{i_0}\| \geq \frac{c_k}{q\|\nabla F^k(x^k)^T\|}.$$

Since the spectral norm is self-adjoint, we have in this case, that

$$\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)v\| \geq \frac{c_k}{q\max\{1, \|\nabla F^k(x^k)\|\}},$$

which implies

$$\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^{-1}\| = \frac{1}{\inf_{\|v\|=1}\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)v\|} \leq \frac{q\max\{1, \|\nabla F^k(x^k)\|\}}{c_k}.$$

Case 2. $\mathcal{B}^k \neq \emptyset$.

$$\begin{aligned}
\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^T v\|^2 &= v^T \mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^T v \\
&= v^T[(A(x^k)^T - I)\nabla F^k(x^k)^T + (B(x^k)^T - I)] \\
&\quad \times [\nabla F^k(x^k)(A(x^k) - I) + (B(x^k) - I)]v \\
&= v^T[(A(x^k)^T - I)\nabla F^k(x^k)^T \nabla F^k(x^k)(A(x^k) - I)]v \\
&\quad + v^T[(B(x^k)^T - I)(B(x^k) - I)]v \\
&\quad + 2v^T[(A(x^k)^T - I)\nabla F^k(x^k)^T(B(x^k) - I)]v \\
&\geq v^T[(A(x^k)^T - I)\nabla F^k(x^k)^T \nabla F^k(x^k)(A(x^k) - I)]v \\
&\quad + v^T[(B(x^k)^T - I)(B(x^k) - I)]v \\
&\geq \frac{[v^T(A(x^k)^T - I)\nabla F^k(x^k)^T(A(x^k) - I)v]^2}{\|(A(x^k) - I)v\|^2} \\
&\quad + \|(B(x^k) - I)v\|^2 \\
&\geq c_k^2\|(A(x^k) - I)v\|^2 + \|(B(x^k) - I)v\|^2
\end{aligned}$$

$$
\begin{aligned}
&\geq c_k^2(\|(A(x^k)-I)v\|^2+\|(B(x^k)-I)v\|^2)\\
&\geq \frac{c_k^2}{2}(\|(A(x^k)-I)v\|+\|(B(x^k)-I)v\|)^2\\
&\geq \frac{c_k^2}{2}\|[2I-A(x^k)-B(x^k)]v\|^2,
\end{aligned}
\tag{4.18}
$$

where the first inequality is from Assumption 4.4.

Suppose that x^* is the limit point of the sequence $\{x^k\}$, and hence by Criterion 4.1, it is the limit point of the sequence $\{P_k(x^k)\}$. Next we prove that $2I-A(x^*)-B(x^*)$ is nonsingular. Since $2I-A(x^*)-B(x^*)=\mathrm{diag}(2I_i-A_i(x^*)-B_i(x^*), i\in\{1,\ldots,q\})$ and for each $i\in\{1,\ldots,q\}$, $2I_i-A_i(x^*)-B_i(x^*)$ is nonsingular for the same reason as that in the proof of Lemma 4.4, then we obtain the nonsingularity of $2I-A(x^*)-B(x^*)$, which, together with the Von Neumann Lemma, implies that $2I-A(x^k)-B(x^k)$ is nonsingular and

$$
\|[2I-A(x^k)-B(x^k)]^{-1}\|\leq\sqrt{2}\|[2I-A(x^*)-B(x^*)]^{-1}\|
\tag{4.19}
$$

for sufficiently large k. Combining (4.18) and (4.19) yields

$$
\begin{aligned}
\|[\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^T]^{-1}\| &= \frac{1}{\inf_{\|v\|=1}\|\mathcal{J}\Phi^k(x^k)^Tv\|}\\
&\leq \frac{\sqrt{2}}{c_k\inf_{\|v\|=1}\|[2I-A(x^k)-B(x^k)]v\|}\\
&= \frac{\sqrt{2}\|[2I-A(x^k)-B(x^k)]^{-1}\|}{c_k}.
\end{aligned}
$$

From all the above discussion, we obtain that

$$
\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^{-1}\|\leq\frac{\max\{q,q\|\nabla F^k(x^k)\|,\sqrt{2}\|[2I-A(x^k)-B(x^k)]^{-1}\|\}}{c_k}.
$$

It follows from Assumption 4.3 and (4.19) that there exists a positive number m_5 such that

$$
\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^{-1}\|\leq\frac{m_5}{c_k}
$$

when k is large enough. Now, from Lemma 4.3, we have

$$
\begin{aligned}
\|x_N^k-P_k(x^k)\| &= \|x^k-P_k(x^k)-\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^{-1}(\Phi_{\mathrm{FB}}^k(x^k)-\Phi_{\mathrm{FB}}^k(P_k(x^k)))\|\\
&\leq \|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)^{-1}\|\\
&\quad\times\|\mathcal{J}\Phi_{\mathrm{FB}}^k(x^k)(x^k-P_k(x^k))-\Phi_{\mathrm{FB}}^k(x^k)+\Phi_{\mathrm{FB}}^k(P_k(x^k))\|\\
&\leq \frac{m_5B_3\|x^k-P_k(x^k)\|^2}{c_k}.
\end{aligned}
$$

Then, letting $B_4=m_5B_3$ gives desired result. □

Now we are in a position to give the main result of this section which shows that only a single Newton step of generalized Newton method can generate the point satisfying condition (3.1) in Algorithm 3.1.

Theorem 4.3 *Suppose that Assumptions* 3.1, 4.2–4.4 *hold and* SOCCP(F) *is strictly feasible. Let* x_N^k *be given by* (4.4). *Then, for any* $r \in (0,1)$, *when* k *is large enough, one has that* x_N^k *satisfies Criterion* 4.1 *and Criterion* 4.2(r), *that is,*

$$f^k(x_N^k) \le \frac{c_k^6 \min\{1, \|x_N^k - x^k\|^{4(1-r)}\}}{18 \max\{\sqrt{2}, \|F^k(P_k(x^k))\|, \|P_k(x^k)\|\}^2}.$$

Proof From Lemma 4.1, it is sufficient to argue that

$$[f^k(x_N^k)]^{\frac{1}{4}} \le \frac{\sqrt{c_k^3}\|x_N^k - x^k\|^{(1-r)}}{(18B_1)^{\frac{1}{4}}},$$

when k is large enough. Let $\tau > 0$ be arbitrary. Since $\{\mathrm{dist}(x^k, \bar{X})\}$ converges to 0 superlinearly by Theorem 3.3, we have that

$$\mathrm{dist}(x^k, \bar{X})^{2r} \le \tau c_k^6$$

for sufficiently large k. It follows from Lemmas 4.2 and 4.5 that

$$\begin{aligned}
\|x_N^k - P_k(x^k)\| &\le \frac{B_4\|x^k - P_k(x^k)\|^2}{c_k} \\
&\le \frac{B_2 B_4 \,\mathrm{dist}(x^k, \bar{X})\|x^k - P_k(x^k)\|}{\sqrt{c_k^3}} \\
&\le \tau B_2 B_4 c_k^2 \|x^k - P_k(x^k)\|.
\end{aligned}$$

Moreover,

$$\|x^k - P_k(x^k)\| \le \|x_N^k - P_k(x^k)\| + \|x^k - x_N^k\|,$$

which says

$$(1 - \tau B_2 B_4 c_k^2)\|x^k - P_k(x^k)\| \le \|x^k - x_N^k\|.$$

Also, when τ is chosen sufficiently small, we have

$$\begin{aligned}
\|x_N^k - P_k(x^k)\| &\le \frac{B_4\|x^k - P_k(x^k)\|^2}{c_k} \\
&\le \frac{B_2 B_4 \,\mathrm{dist}(x^k, \bar{X})^{2r}\|x^k - P_k(x^k)\|^{2(1-r)}}{\sqrt{c_k^3}} \\
&\le \tau B_2 B_4 c_k^4 \|x^k - P_k(x^k)\|^{2(1-r)}
\end{aligned}$$

$$\leq \frac{\tau B_2 B_4 c_k^4}{(1-\tau B_2 B_4 c_k^2)^{2(1-r)}} \|x^k - x_N^k\|^{2(1-r)}$$
$$\leq c_k^4 \|x^k - x_N^k\|^{2(1-r)}.$$

On the other hand, the function $[f^k(x)]^{\frac{1}{2}}$ is uniformly locally Lipschitz continuous from Lemma 2.3. Then there exists $L_2 > 0$ such that

$$[f^k(x_N^k)]^{\frac{1}{2}} \leq L_2^2 \|x_N^k - P_k(x^k)\|.$$

Hence

$$[f^k(x_N^k)]^{\frac{1}{4}} \leq L_2 \|x_N^k - P_k(x^k)\|^{\frac{1}{2}} \leq L_2 c_k^2 \|x^k - x_N^k\|^{(1-r)} \leq \frac{\sqrt{c_k^3}\|x_N^k - x^k\|^{(1-r)}}{(18B_1)^{\frac{1}{4}}}.$$

Then, the proof is complete. □

We point it out that this theorem together with Theorem 3.3 implies that the proximal point algorithm in Sect. 3 has approximate genuine superlinear convergence.

5 Numerical experiments

In this section, we report numerical results of Algorithm 3.1 for solving SOCCP(F) defined by (1.1) and compare the performance with that of the derivative-free descent method used by [21]. To construct SOCs of various types, we chose n_i and q such that $n_1 = n_2 = \cdots = n_q$. Our numerical experiments are carried out in Matlab (version 7.8) running on a PC Intel core 2 Q8200 of 2.33 GHz CPU and 2.00 GB Memory.

We consider the case where $F(x) = Mx + b$ with the matrix $M \in \Re^{n\times n}$ and $b \in \Re^n$ generated randomly, whose generating procedure was described as in [21]. In our numerical experiments, the stopping criterions for both Algorithms 3.1 and 4.1 for solving subproblems are Tol. $= 10^{-8}$. In Algorithm 3.1, we set the parameters as $\alpha = 0.5$, $c_0 = 0.5$ and the initial point is chosen as $x^0 = (x_1^0, \ldots, x_q^0)$, where $x_i^0 = (10, \frac{\omega_i}{\|\omega_i\|})$ for $i = 1, 2, \ldots, q$ with $\omega_i \in \Re^{n_i - 1}$ being generated randomly by Matlab's *rand.m*. In Algorithm 4.1, we set the parameter $\beta = 10^{-4}$ and the initial point for Newton's method is selected as the current iteration point in the main algorithm, i.e., Algorithm 3.1. In additional, the main task of Algorithm 4.1 for solving the Subproblem, at each iterate, is solve the linear system (4.2). In numerical implementation, we apply the preconditioner conjugate gradient square method for solving system (4.2).

We first used Algorithm 3.1 to solve a test problem with $n = 1000$ and $q = 100$. The Fig. 1 below plot the corresponding convergence of $\{f_{\rm FB}(x^k)\}$ versus the iteration number of PPA and Table 1 reported its corresponding iteration performance, where k denotes the kth iteration ($k = 0$ stands for the initial iteration) of Algorithm 3.1, $f_{\rm FB}(x^k)$ indicates the current value of the merit function at the kth iteration, Gap reports the value of $(x^k)^T F(x^k)$ at the kth iteration and Num means the number of

Fig. 1 Convergence behavior of $\{f_{\mathrm{FB}}(x^k)\}$

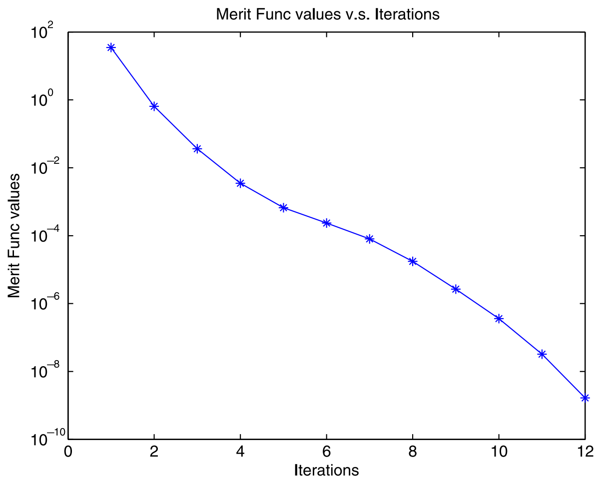

Table 1 Iteration performance of affine monotone SOCCP

k	$f_{\mathrm{FB}}(x^k)$	Gap	Num	k	$f_{\mathrm{FB}}(x^k)$	Gap	Num
0	3.24e+4	1.21e+3	–	1	3.49e+1	1.18e+1	5
2	6.40e−1	9.69e−1	3	3	3.62e−2	8.51e−1	2
4	3.51e−3	3.11e−1	2	5	6.65e−4	8.19e−2	2
6	2.36e−4	3.19e−2	1	7	8.01e−5	1.96e−2	1
8	1.74e−5	5.67e−3	1	9	2.65e−6	2.51e−3	1
10	3.61e−7	2.11e−3	1	11	3.24e−8	5.48e−4	1
12	1.65e−9	4.80e−5	1				

Newton steps needed in Algorithm 4.1 at the kth iteration. From Table 1, we see that only a single Newton step of generalized Newton method can generate the point with the desired accuracy, which coincides with the analysis in Sect. 4.

To further test how the performance of Algorithm 3.1 varies with the structure of $\mathcal{K}$ and the total dimension, we used Algorithm 3.1 to solve several test problems with different n and q. Also, we compared the numerical performance of the group of test problems when $n = 1000$ and $q = 100$ with that of the derivative-free descent method used by [21]. The numerical results were reported in Tables 2 and 3, where $f_{\mathrm{FB}}(x^*)$, Gap, NF, Time, stand for, respectively, the merit function value at the final iteration, the value of $|(x^*)^T F(x^*)|$ at the final iteration, the number of function evaluations of f_{FB}, the total CPU time in second. From Tables 2 and 3, we see that when n is fixed, Algorithm 3.1 requires less function evaluations and CUP time for those problems with larger q. Moreover, Algorithm 3.1 is superior to the derivative-free descent method in terms of the number of function evaluations and CPU time for the test problems with $n = 1000$ and $q = 100$.

Table 2 Numerical results for affine monotone SOCCPs

n	q	$f_{\mathrm{FB}}(x^*)$	Gap	NF	Time
1000	100	4.89e−10	8.64e−5	65	21.1
1000	20	5.29e−9	2.69e−4	97	46.5
1000	10	1.59e−9	5.52e−4	124	89.5
2000	100	8.98e−9	2.68e−4	73	166.2
2000	50	6.67e−9	2.77e−4	101	338.4
2000	20	6.41e−9	8.37e−4	160	809.1
3000	100	1.24e−9	3.53e−4	87	702.8

Table 3 Numerical comparisons for affine monotone SOCCPs with 100 SOCs

Problem	Algorithm 4.1				Derivative-free method			
	$f_{\mathrm{FB}}(x^*)$	Gap	NF	Time	$f_{\mathrm{FB}}(x^*)$	Gap	NF	Time
1	4.68e−9	1.73e−4	60	8.5	9.99e−9	2.49e−4	19972	109.4
2	1.53e−9	4.08e−4	59	8.7	9.99e−9	1.16e−3	29626	156.6
3	4.80e−10	1.54e−4	69	21.1	9.99e−9	5.82e−4	62084	323.8
4	1.76e−9	3.96e−5	67	18.4	9.99e−9	1.62e−4	70949	361.9
5	3.34e−9	2.20e−4	68	18.4	9.99e−9	3.29e−4	79244	420.4

6 Conclusions

The proximal point algorithm has nice theoretical convergence results under appropriate criteria for approximate solutions of subproblems. However, it is usually not easy to check those criteria. In this paper, we introduce PPA for solving monotone SOCCP and construct a practical approximation criterion. Moreover, we adopt the generalized Newton method to solve subproblems and show that only one Newton step is eventually needed to obtain an approximation solution of the subproblem that approximately satisfies the criterion. Our work, though is motivated by that of Yamashita and Fukushima for solving NCP, is not a direct extension from NCP to SOCCP as many results that are easy to achieve for NCP are no longer hold for SOCCP. For example, it is easy to derive a global error bound for NCP from Fischer-Burmeister function, but this is not true for SOCCP. Besides, the nonsingularity of complementarity function and the boundedness of its inverse are much more difficult to be verified in SOCCP case than in NCP case.

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

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**, 3–51 (2003)
2. Andersen, E.D., Roos, C., Terlaky, T.: On implementing a primal-dual interior-point method for conic quadratic optimization. Math. Program. **95**, 249–277 (2003)

3. Chen, J.-S.: A new merit function and its related properties for the second-order cone complementarity problem. Pac. J. Optim. **2**, 167–179 (2006)
4. Chen, X., Qi, H.: Cartesian P-property and its applications to the semidefinite linear complementarity problem. Math. Program. **106**, 177–201 (2006)
5. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problems. Math. Program. **104**, 293–327 (2005)
6. Chen, X.-D., Sun, D., Sun, J.: Complementarity functions and numerical experiments for second-order cone complementarity problems. Comput. Optim. Appl. **25**, 39–56 (2003)
7. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
8. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementary Problems. Springer, New York (2003)
9. Fischer, A.: Solution of monotone complementarity problems with locally Lipschitzian functions. Math. Program. **76**, 513–532 (1997)
10. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
11. Kanzow, C., Ferenczi, I., Fukushima, M.: On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity. SIAM J. Optim. **20**, 297–320 (2009)
12. Hayashi, S., Yamashita, N., Fukushima, M.: A combined smoothing and regularization method for monotone second-order cone complementarity problems. SIAM J. Optim. **15**(2), 593–615 (2005)
13. Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H.: Application of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
14. De Luca, T., Facchinei, F., Kanzow, C.: A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407–439 (1996)
15. Luque, F.J.: Asymptotic convergence analysis of the proximal point algorithm. SIAM J. Control Optim. **22**, 277–293 (1984)
16. Martinet, B.: Perturbation des méthodes d'opimisation. RAIRO Anal. Numér. **12**, 153–171 (1978)
17. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 957–972 (1997)
18. Monteiro, R.D.C., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone programs based on the MZ-family of directions. Math. Program. **88**, 61–83 (2000)
19. Pan, S., Chen, J.-S.: A damped Gauss-Newton method for the second-order cone complementarity problem. Appl. Math. Optim. **59**, 293–318 (2009)
20. Pan, S., Chen, J.-S.: A regularization method for the second-order cone complementarity problem with the Cartesian P_0-property. Nonlinear Anal, Theory Methods Appl. **70**, 1475–1491 (2009)
21. Pan, S., Chen, J.-S.: A linearly convergent derivative-free descent method for the second-order cone complementarity problem. Optimization **59**, 1173–1197 (2010)
22. Pang, J.-S., Qi, L.: Nonsmooth equations: motivation and algorithms. SIAM J. Optim. **3**, 443–465 (1993)
23. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
24. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
25. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
26. Sun, D., Sun, J.: Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions. Math. Program. **103**, 575–581 (2005)
27. Sun, J., Sun, D., Qi, L.-Q.: A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. SIAM J. Optim. **14**, 783–806 (2004)
28. Tsuchiya, T.: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Softw. **11**, 141–182 (1999)
29. Yamashita, N., Fukushima, M.: The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem. SIAM J. Optim. **11**(2), 364–379 (2000)
30. Yan, T., Fukushima, M.: Smoothing method for mathematical programs with symmetric cone complementarity constraints. Technical report 2009-017, Department of Applied Mathematics and Physics, Kyoto University (2009)
31. Yoshise, A.: Interior point trajectories and a homogenous model for nonlinear complementarity problems over symmetric cones. SIAM J. Optim. **17**, 1129–1153 (2006)