Journal of Optimization Theory and Applications (2021) 189:679–715
https://doi.org/10.1007/s10957-021-01839-6

A Nonmonotone Smoothing Newton Algorithm for Weighted Complementarity Problem

Jingyong Tang[1] · **Hongchao Zhang**[2]

Received: 17 September 2019 / Accepted: 15 February 2021 / Published online: 20 March 2021

Abstract

The weighted complementarity problem (denoted by WCP) significantly extends the general complementarity problem and can be used for modeling a larger class of problems from science and engineering. In this paper, by introducing a one-parametric class of smoothing functions which includes the weight vector, we propose a smoothing Newton algorithm with nonmonotone line search to solve WCP. We show that any accumulation point of the iterates generated by this algorithm, if exists, is a solution of the considered WCP. Moreover, when the solution set of WCP is nonempty, under assumptions weaker than the Jacobian nonsingularity assumption, we prove that the iteration sequence generated by our algorithm is bounded and converges to one solution of WCP with local superlinear or quadratic convergence rate. Promising numerical results are also reported.

Keywords Smoothing Newton algorithm · Jacobian nonsingularity · Superlinear/quadratic convergence · Weighted complementarity problem · Symmetric cone

Mathematics Subject Classification 90C25 · 90C33 · 65Y20

Communicated by Florian Potra.

This research was partially supported by the National Natural Science Foundation of China (11601466), China Scholarship Council and Nanhu Scholars Program for Young Scholars of XYNU, and by the USA National Science Foundation (1522654, 1819161).

✉ Hongchao Zhang
hozhang@math.lsu.edu

Jingyong Tang
tangjy@xynu.edu.cn

[1] School of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China

[2] Department of Mathematics, Louisiana State University, 303 Lockett Hall, Baton Rouge, LA 70803-4918, USA

1 Introduction

The weighted complementarity problem (WCP) was introduced by Potra [30]. Let $(\mathbb{V}, \langle\cdot,\cdot\rangle, \circ)$ be a Euclidean Jordan algebra [6] and $\mathcal{K} = \{x \circ x : x \in \mathbb{V}\}$ be the symmetric cone formed by the squares of its elements. Given a vector $w \in \mathcal{K}$, WCP is the problem of finding $(x, s, y) \in \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ such that

$$\text{(WCP)} \qquad x \in \mathcal{K},\ s \in \mathcal{K},\ F(x, s, y) = 0,\ x \circ s = w, \tag{1.1}$$

where $F : \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \to \mathbb{V} \times \mathcal{R}^m$ is a continuously differentiable nonlinear map. When $w = 0$, the WCP would reduce to the mixed nonlinear complementarity problem over symmetric cones which, for example, was studied by Yoshise [37]. Moreover, when $x \in \mathcal{K}$ and $s \in \mathcal{K}$, we have $x \circ s = 0$ if and only if $\langle x, s\rangle = 0$ (see, [10, Proposition 6]). Hence, in (1.1) if we have $w = 0$, $m = 0$ and $F(x, s, y) = f(x) - s$ with $f : \mathbb{V} \to \mathbb{V}$ being a continuously differentiable nonlinear map, this WCP would become the well-known symmetric cone complementarity problem (SCCP), which finds $(x, s) \in \mathbb{V} \times \mathbb{V}$ such that

$$\text{(SCCP)} \qquad x \in \mathcal{K},\ \ s \in \mathcal{K},\ \ s = f(x),\ \ \langle x, s\rangle = 0. \tag{1.2}$$

So, WCP significantly extends the scope of general complementarity problems and can model a wider class of problems arising from real applications more conveniently. In addition, the flexibility of WCP model may also lead to more efficient numerical solvers even the problem can be also formulated as a general complementarity problem, e.g., the Fisher market equilibrium problem [30]. Although WCP was proposed in general Euclidean Jordan algebra setting, only the case of linear WCP over the nonnegative orthant (LWCP) was studied in [30], which finds $(x, s, y) \in \mathcal{R}^n \times \mathcal{R}^n \times \mathcal{R}^m$ such that

$$\text{(LWCP)} \qquad x \in \mathcal{R}^n_+,\ s \in \mathcal{R}^n_+,\ Px + Qs + Ry = a,\ xs = w. \tag{1.3}$$

Here, $P \in \mathcal{R}^{(n+m)\times n}$, $Q \in \mathcal{R}^{(n+m)\times n}$, $R \in \mathcal{R}^{(n+m)\times m}$ and $a \in \mathcal{R}^{n+m}$ are given matrices and vector, $w \in \mathcal{R}^n_+$ is a given weight vector, and xs denotes the vector with components $x_i s_i$. In [30], Potra showed that the quadratic programming and weighted centering problem, which generalizes the notion of linear programming and the weighted centering problem proposed by Anstreicher [2], can be formulated as a special LWCP. Two interior-point methods as well as their computational complexities were also studied in [30]. Moreover, sufficient conditions for characterization of solutions of LWCP were given in [31] and a corrector–predictor interior-point method was proposed for its numerical solution. Lately, Asadi et al. [1] introduced a full-Newton step interior-point algorithm for solving the LWCP. When $w = 0$ in (1.1), Yoshise discussed the trajectory of an interior point map in view of homeomorphisms of continuous maps and gave a homogeneous model to solve the problem [37]. However, to the best of our knowledge, also as pointed out in [30], there are no existence results and algorithms for the solution of WCP (1.1) in the general case.

On the other hand, there have been much interests in smoothing Newton-type algorithms for solving optimization problems over symmetric cone, such as the symmetric cone programming (SCP) (e.g., [15,18,20]) and the SCCP (e.g., [13,14,17,19,21,22,24,27,35,36]). And more recently, smoothing Newton algorithm has been proposed for solving LWCP, see [16,34]. The main idea of these classes of smoothing Newton algorithms is to use a smoothing function to reformulate the problem concerned as a system of smooth nonlinear equations $\mathcal{H}(z) = 0$ and then solve it by Newton's method. To obtain local fast convergence, all these smoothing Newton-type algorithms require the following *Jacobian nonsingularity assumption*:

$$\text{All } V \in \partial\mathcal{H}(z^*) \text{ are nonsingular,} \tag{1.4}$$

where z^* is any accumulation point of the iteration sequence generated by smoothing Newton-type algorithms and $\partial\mathcal{H}$ stands for the Clarke's generalized Jacobian [5]. However, this Jacobian nonsingularity assumption may not hold especially when the solution set of WCP is not a singleton. Moreover, even with this Jacobian nonsingularity assumption, many papers do not analyze whether an accumulation point exists or not (e.g., [4,16,18,19,34]). To ensure such an accumulation point exist, smoothing Newton-type algorithms usually require the boundedness of the solution set (e.g., [14,17,21,22,24,27,35]).

In this paper, we aim to design a globally convergent nonmonotone smoothing Newton algorithm to solve WCP (1.1) in the general case and show its local fast convergence without Jacobian nonsingularity assumption. Specifically, we introduce a one-parametric class of smoothing functions which include the weight vector w. Based on these functions, we reformulate WCP in (1.1) as a system of smooth equations $\mathcal{H}(z) = 0$ (see, Sect. 3) and propose a smoothing Newton algorithm combined with nonmonotone line search to solve it. Any accumulation point of the iterates generated by this algorithm is a solution of $\mathcal{H}(z) = 0$. Moreover, under assumptions which are much weaker than the Jacobian nonsingularity assumption, we show that when the solution set of the considered WCP is nonempty,

- the distance between the iteration sequence $\{z^k\}$ and the solution set $\mathcal{Z}^* := \{z|\mathcal{H}(z) = 0\}$ converges to zero locally superlinearly or quadratically;
- furthermore, the iteration sequence $\{z^k\}$ in fact converges to one solution $z^* \in \mathcal{Z}^*$ locally superlinearly or quadratically.

To the best of our knowledge, these convergence results for smoothing Newton-type algorithms have not been studied in the literature, even in the simple case of $\mathcal{K} = \mathcal{R}^n_+$ and the weight vector $w = 0$. Moreover, the iteration points of our algorithm are not required to be interior points of the symmetric cone $\mathcal{K}$ in WCP. Hence, from computational point of view, the new algorithm is much more flexible, easier to use and often finds the solution more efficiently than interior-point methods proposed in [30,31], which are confirmed by our numerical experiments.

The outline of this paper is as follows. In Sect. 2, we briefly recall the Euclidean Jordan algebra. In Sect. 3, we introduce a one-parametric class of smoothing functions including the weight vector w and reformulate WCP as a system of nonlinear smooth equations. We propose and briefly discuss our nonmonotone smoothing Newton algorithm for solving WCP in Sect. 4. In Sect. 5, we analyze the global and local convergence properties of this algorithm. Numerical results are reported in Sect. 6, and some conclusions are given in Sect. 7.

Throughout this paper, we use the following notations. $\mathcal{R}^n$ denotes the set of all n dimensional real vectors, and $\mathcal{R}^n_+$ (respectively, $\mathcal{R}^n_{++}$) denotes the nonnegative (respectively, positive) orthant of $\mathcal{R}^n$. For any $t \in \mathcal{R}$, let $t_+ = \max\{0, t\}$ and $t_- = \min\{0, t\}$. All vectors are column vectors, and for simplicity, the column vector $(u_1^T, \ldots, u_n^T)^T$ is written as $(u_1, \ldots, u_n)$, where u_i is a column vector in $\mathbb{V}$. For a given set $S \subset \mathbb{V}$, $\mathrm{int}S$ and $\mathrm{conv}S$ denote the interior and convex hull of S, respectively, and for any $u \in \mathbb{V}$, $\mathrm{dist}(u, S) = \inf\limits_{v \in S}\{\|u - v\|\}$, where $\|\cdot\|$ is the norm on $\mathbb{V}$ induced by the inner product $\langle \cdot, \cdot \rangle$. For any $u, v \in \mathcal{K}$, we write $u \succeq_{\mathcal{K}} v$ (respectively, $u \succ_{\mathcal{K}} v$) if $u - v \in \mathcal{K}$ (respectively, $u - v \in \mathrm{int}\mathcal{K}$). We use I as the identity operator, i.e., $Ix = x$ for all $x \in \mathbb{V}$. For a differentiable mapping $G : \mathbb{V} \to \mathbb{V}$, we denote $G'(x)$ by the Jacobian of G at $x \in \mathbb{V}$. For any $\alpha, \beta > 0$, $\alpha = \mathcal{O}(\beta)$ (respectively, $\alpha = o(\beta)$) means that $\limsup_{\beta \to 0} \frac{\alpha}{\beta} < \infty$(respectively, $\limsup_{\beta \to 0} \frac{\alpha}{\beta} = 0$).

2 Euclidean Jordan Algebra

In this section, we briefly recall some major properties of Euclidean Jordan algebra which will be used in this paper. Details on the description of Euclidean Jordan algebra can be found in [6].

A *Euclidean Jordan algebra* is a triple $(\mathbb{V}, \langle \cdot, \cdot \rangle, \circ)$, where $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ is a finite-dimensional inner product space over $\mathcal{R}$ and $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying the following conditions:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathbb{V}$;
(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathbb{V}$, where $x^2 := x \circ x$;
(iii) $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$ for all $x, y, z \in \mathbb{V}$.

We assume that there is an element $e \in \mathbb{V}$ (called the *unit element*) such that $x \circ e = e \circ x = x$ for all $x \in \mathbb{V}$.

For a given $x \in \mathbb{V}$, the *Lyapunov transformation* is defined by $\mathcal{L}_x : \mathbb{V} \to \mathbb{V}$ by $\mathcal{L}_x y := x \circ y$, $\forall y \in \mathbb{V}$, which is a symmetric operator in the sense that $\langle \mathcal{L}_x y, z \rangle = \langle y, \mathcal{L}_x z \rangle$ for all $y, z \in \mathbb{V}$. The operator $\mathcal{L}_x$ is positive definite if $\langle u, \mathcal{L}_x u \rangle > 0$ for all $0 \neq u \in \mathbb{V}$. From [37, Proposition 2.1], we have that $\mathrm{int}\mathcal{K} = \{x \in \mathbb{V} :\ \mathcal{L}_x \text{ is positive definite.}\}$.

The *symmetric cone* $\mathcal{K}$ is a self-dual closed convex cone with nonempty interior $\mathrm{int}\mathcal{K}$ and homogeneous, i.e., for any two elements $x, y \in \mathrm{int}\mathcal{K}$, there exists an invertible linear transformation $\ : \mathbb{V} \to \mathbb{V}$ such that $(\mathcal{K}) = \mathcal{K}$ and $(x) = y$. By [6, Theorem III.2.1], the symmetric cone $\mathcal{K}$ coincides with the set of squares $\{x^2 : x \in \mathbb{V}\}$.

For any $x \in \mathbb{V}$, let $m(x) := \min\{k : \{e, x, \ldots, x^k\}$ are linearly dependent$\}$. Since $m(x) \leq \dim\mathbb{V}$ where $\dim\mathbb{V}$ denotes the dimension of $\mathbb{V}$, the rank of $\mathbb{V}$ is well defined by $r := \max\{m(x) : x \in \mathbb{V}\}$. An element $c \in \mathbb{V}$ is said to be *idempotent* if $c^2 = c$. An idempotent is said to be *primitive* if it is nonzero and cannot be written as the sum of two other nonzero idempotents. A *complete system of orthogonal idempotents* is a finite set $\{c_1, \ldots, c_m\}$, where $c_j^2 = c_j,\ c_i \circ c_j = 0,\ \forall i \neq j,\ i, j = 1, \ldots, m$, and $\sum\limits_{i=1}^{m} c_i = e$.

A complete system of orthogonal primitive idempotents is called a *Jordan frame*. We have the following important *spectral decomposition* theorem.

Theorem 1 [6, Theorem III.1.2] *Let* $(\mathbb{V}, \langle\cdot,\cdot\rangle, \circ)$ *be a Euclidean Jordan algebra with rank* r. *Then, for any* $x \in \mathbb{V}$, *there exist a Jordan frame* $\{c_1, \ldots, c_r\}$ *and real numbers* $\lambda_1(x), \ldots, \lambda_r(x)$ *such that* $x = \sum_{i=1}^{r} \lambda_i(x)c_i$. *The numbers* $\lambda_i(x)(i = 1, \ldots, r)$, *which are uniquely determined by* x, *are called the eigenvalues of* x.

Let $x \in \mathbb{V}$ and $\lambda_1(x), \ldots, \lambda_r(x)$ be its eigenvalues. The trace of x is denoted by $\mathrm{Tr}(x) := \sum_{i=1}^{r} \lambda_i(x)$. For any $x, y \in \mathbb{V}$, the inner product of x, y is $\langle x, y\rangle := \mathrm{Tr}(x \circ y)$ and the norm on $\mathbb{V}$ induced by this inner product is $\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{\mathrm{Tr}(x^2)}$.

For any $x \in \mathbb{V}$ with the spectral decomposition $x = \sum_{i=1}^{r} \lambda_i(x)c_i$, we define

$$x_+ := \sum_{i=1}^{r} \lambda_i(x)_+ c_i, \quad x_- := \sum_{i=1}^{r} \lambda_i(x)_- c_i, \quad |x| := \sum_{i=1}^{r} |\lambda_i(x)| c_i.$$

Since $x \in \mathcal{K}$ ($x \in \mathrm{int}\mathcal{K}$) if and only if $\lambda_i(x) \geq 0$ ($\lambda_i(x) > 0$) for all $i = 1, \ldots, r$ and $t = t_+ + t_-$ and $|t| = t_+ - t_-$ for any $t \in \mathcal{R}$, we have

$$x_+ \in \mathcal{K}, \quad -x_- \in \mathcal{K}, \quad x = x_+ + x_-, \quad |x| = x_+ - x_-.$$

Moreover, for any $x \in \mathcal{K}$ we define

$$x^2 := \sum_{i=1}^{r} \lambda_i(x)^2 c_i \;\text{ and }\; \sqrt{x} := \sum_{i=1}^{r} \sqrt{\lambda_i(x)} c_i.$$

Then, we have $|x| = \sqrt{x^2}$. More generally, for any real-valued function $f : \mathcal{R} \to \mathcal{R}$, we may define a function associated with the Euclidean Jordan algebra $(\mathbb{V}, \langle\cdot,\cdot\rangle, \circ)$ by

$$f_{\mathbb{V}}(x) := \sum_{i=1}^{r} f(\lambda_i(x))c_i.$$

This function $f_{\mathbb{V}}$ is a *Löwner operator* and inherits many properties from f.

3 Basic Ideas of the Algorithm

3.1 Smoothing Functions for WCP

Since WCP contains the weight vector $w \in \mathcal{K}$, the traditional complementarity functions over the symmetric cone, such as the natural residual function and the

Fischer–Burmeister function (see, [10,33]), cannot be used to carry on the equivalent reformulation. In this subsection, we introduce a one-parametric class of complementarity functions $\phi : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$, including the weight vector $w \in \mathcal{K}$, defined by

$$\phi(x, s) = x + s - \sqrt{x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w}, \quad \forall (x, s) \in \mathbb{V} \times \mathbb{V}, \tag{3.1}$$

where $\tau \in [0, 4)$ is a constant. Note that for any $\tau \in [0, 4)$ and $w \in \mathcal{K}$, we have

$$\begin{aligned} &x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w \\ &\quad = [x + (\tau/2 - 1)s]^2 + \tau(1 - \tau/4)s^2 + (4 - \tau)w \in \mathcal{K}. \end{aligned} \tag{3.2}$$

Hence, the function $\phi(x, s)$ given in (3.1) is well defined. In what follows, we show that the function ϕ is a class of complementarity functions for WCP.

Lemma 1 *For any $u, v \in \mathbb{V}$,* (i) *if $u \succeq_{\mathcal{K}} 0$ and $u^2 \succeq_{\mathcal{K}} v^2$, then $u \succeq_{\mathcal{K}} |v|$ and $u \succeq_{\mathcal{K}} v$;* (ii) *if $u \succeq_{\mathcal{K}} 0$ and $u^2 \succ_{\mathcal{K}} v^2$, then $u \succ_{\mathcal{K}} |v|$ and $u \succ_{\mathcal{K}} v$.*

Proof The results can be directly obtained from [10, Proposition 8]. □

Theorem 2 *Let ϕ be defined by* (3.1) *with $\tau \in [0, 4)$. Then,*

$$\phi(x, s) = 0 \Longleftrightarrow x \in \mathcal{K}, \;\; s \in \mathcal{K}, \;\; x \circ s = w. \tag{3.3}$$

Proof We first suppose that x and s satisfy $\phi(x, s) = 0$. Then, we have

$$x + s = \sqrt{x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w}. \tag{3.4}$$

Squaring the two sides of (3.4) gives

$$x^2 + s^2 + 2x \circ s = x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w,$$

which together with $\tau \in [0, 4)$ yields $x \circ s = w$. By substituting $x \circ s = w$ into (3.4), we have $x + s = \sqrt{x^2 + s^2 + 2w}$. Since $x^2 \succeq_{\mathcal{K}} 0$, $s^2 \succeq_{\mathcal{K}} 0$ and $w \succeq_{\mathcal{K}} 0$, we have $c := x + s = \sqrt{x^2 + s^2 + 2w} \succeq_{\mathcal{K}} 0$, $c^2 \succeq_{\mathcal{K}} x^2$ and $c^2 \succeq_{\mathcal{K}} s^2$. Since $c \succeq_{\mathcal{K}} 0$, it follows from Lemma 1 that $c \succeq_{\mathcal{K}} x$ and $c \succeq_{\mathcal{K}} s$. By noticing that $c = x + s$, we have $x = c - s \succeq_{\mathcal{K}} 0$ and $s = c - x \succeq_{\mathcal{K}} 0$, i.e., $x \in \mathcal{K}$ and $s \in \mathcal{K}$.

Conversely, we suppose that $x \in \mathcal{K}$, $s \in \mathcal{K}$, $x \circ s = w$. Then, $x + s \in \mathcal{K}$ and

$$\sqrt{x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w} = \sqrt{(x + s)^2} = x + s,$$

which implies that $\phi(x, s) = 0$. This completes the proof. □

Note that the complementarity function ϕ in (3.1) is not continuously differentiable everywhere. Hence, to design our algorithm, we introduce a smoothing parameter $\mu \in$

$\mathcal{R}_+$ into ϕ and get a one-parametric class of smoothing functions $\psi : \mathcal{R}_+ \times \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ as follows:

$$\psi(\mu, x, s) = x + s - \sqrt{x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w + 4\mu^t e}, \tag{3.5}$$

where $t \in [1, 2]$ is a constant.

Theorem 3 *Let ψ be defined by (3.5) with $\tau \in [0, 4)$ and $t \in [1, 2]$. Then, the following results hold.*

(i) For any $(x, s) \in \mathbb{V} \times \mathbb{V}$, $\lim\limits_{\mu \to 0} \psi(\mu, x, s) = \psi(0, x, s) = \phi(x, s)$ and

$$\psi(0, x, s) = 0 \Longleftrightarrow x \in \mathcal{K},\ \ s \in \mathcal{K},\ \ x \circ s = w. \tag{3.6}$$

(ii) Let $d = x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w$ with its spectral decomposition given by $d = \sum_{i=1}^r \lambda_i(d)c_i$, where $\{c_1, ..., c_r\}$ is a Jordan frame and the numbers $\lambda_1(d), ..., \lambda_r(d)$ are the eigenvalues of d uniquely determined by d. Let $u, v \in \mathbb{V}$ and $h \in \mathcal{R}$. Then, ψ is continuously differentiable at any $(\mu, x, s) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V}$ with

$$\psi'_\mu(\mu, x, s)h = -\sum_{i=1}^{r} \frac{2t\mu^{t-1}h}{\sqrt{\lambda_i(d) + 4\mu^t}} c_i, \tag{3.7}$$

$$\psi'_x(\mu, x, s)u = u - \mathcal{L}^{-1}_{c(\mu,x,s)}[(x + (\tau/2 - 1)s) \circ u], \tag{3.8}$$

$$\psi'_s(\mu, x, s)v = v - \mathcal{L}^{-1}_{c(\mu,x,s)}[(s + (\tau/2 - 1)x) \circ v], \tag{3.9}$$

where

$$c(\mu, x, s) = \sqrt{x^2 + s^2 + (\tau - 2)x \circ s + (4 - \tau)w + 4\mu^t e} = \sqrt{d + 4\mu^t e}. \tag{3.10}$$

Proof The result (i) obviously holds by Theorem 2. The proof of (3.8) and (3.9) can be obtained by slight modification of the proof of [29, Lemma 4.1]. By (3.2), we have $d \in \mathcal{K}$ and hence $\lambda_i(d) \geq 0$ for all $i = 1, ..., r$. Using this fact, we can prove (3.7) similarly as the proof of [13, Lemma 3.1]. □

3.2 The Reformulation of WCP

Let $z = (\mu, x, s, y) \in \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$. For WCP (1.1), we define the function $\mathcal{H}(z) : \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \to \mathcal{R} \times \mathbb{V} \times \mathcal{R}^m \times \mathbb{V}$ as

$$\mathcal{H}(z) := \begin{pmatrix} \mu \\ F(x, s, y) \\ \psi(\mu, x, s) \end{pmatrix}, \tag{3.11}$$

where ψ is the smoothing function given in (3.5). Then, from (3.6) it holds that

$$\mathcal{H}(z) = 0 \iff \mu = 0 \text{ and } (x, s, y) \text{ is a solution of WCP (1.1).} \tag{3.12}$$

Thus, for solving WCP (1.1), one can apply Newton-type methods to solve the system of nonlinear equations $\mathcal{H}(z) = 0$.

By Theorem 3 (ii), $\mathcal{H}(z)$ is continuously differentiable at any $z \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and its Jacobian is

$$\mathcal{H}'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & F'_x(x, s, y) & F'_s(x, s, y) & F'_y(x, s, y) \\ \psi'_\mu(\mu, x, s) & \psi'_x(\mu, x, s) & \psi'_s(\mu, x, s) & 0 \end{bmatrix}. \tag{3.13}$$

In the rest of the paper, we assume that $F'(x, s, y)$ has the following rank and monotone property.

Assumption 3.1 Suppose that rank $F'_y(x, s, y) = m$ and we have $\langle \Delta x, \Delta s\rangle \geq 0$, for any $(\Delta x, \Delta s, \Delta y) \in \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ with $F'(x, s, y)(\Delta x, \Delta s, \Delta y) = 0$.

For Assumption 3.1, we have the following remarks.

(i) Assumption 3.1 is very standard and has been often used to analyze smoothing Newton-type algorithms for the second-order cone complementarity problem (e.g., [4,9,26]), the LWCP in (1.3) (e.g., [16,34]) and the interior-point methods for solving LWCP [30].

(ii) For SCCP in (1.2), Assumption 3.1 in fact reduces to require that f is monotone, i.e., $\langle f(u) - f(v), u - v\rangle \geq 0$ for all $u, v \in \mathbb{V}$, which has been used in [13,14,17, 19,21,24,27,35,36].

Lemma 2 *Let $a, b \in \mathbb{V}$ with $a \succ_\mathcal{K} 0$, $b \succ_\mathcal{K} 0$ and $a \circ b \succ_\mathcal{K} 0$. Then, there exists a constant $\theta > 0$ such that for all $u, v \in \mathbb{V}$ satisfying $\langle u, v\rangle \geq 0$, we have*

$$\|\mathcal{L}_a u + \mathcal{L}_b v\| \geq \theta(\|u\| + \|v\|). \tag{3.14}$$

Proof Since $b \succ_\mathcal{K} 0$, $\mathcal{L}_b$ is invertible. So, for all $u, v \in \mathbb{V}$ satisfying $\langle u, v\rangle \geq 0$, set $\tilde{u} := \mathcal{L}_b^{-1} u$.

Then, we have

$$\begin{aligned} \|u\| \|\mathcal{L}_b^{-1}\mathcal{L}_a u + v\| &\geq \langle u, \mathcal{L}_b^{-1}\mathcal{L}_a u + v\rangle \\ &\geq \langle u, \mathcal{L}_b^{-1}\mathcal{L}_a u\rangle \\ &= \langle \tilde{u}, \mathcal{L}_a\mathcal{L}_b\tilde{u}\rangle \\ &= \langle \tilde{u}, (\mathcal{L}_a\mathcal{L}_b + \mathcal{L}_b\mathcal{L}_a)\tilde{u}\rangle/2. \end{aligned} \tag{3.15}$$

Since $a \succ_\mathcal{K} 0$, $b \succ_\mathcal{K} 0$ and $a \circ b \succ_\mathcal{K} 0$, it follows from [37, Lemma 2.6 (v)] that $\mathcal{L}_a\mathcal{L}_b + \mathcal{L}_b\mathcal{L}_a$ is positive definite. Let ξ be the minimal eigenvalue of $\mathcal{L}_a\mathcal{L}_b + \mathcal{L}_b\mathcal{L}_a$. Then, $\xi > 0$ and

$$\langle \tilde{u}, (\mathcal{L}_a\mathcal{L}_b + \mathcal{L}_b\mathcal{L}_a)\tilde{u}\rangle \geq \xi\|\tilde{u}\|^2. \tag{3.16}$$

By (3.15) and (3.16), we have

$$\|u\|\|\mathcal{L}_b^{-1}\mathcal{L}_a u + v\| \geq \frac{\xi\|\tilde{u}\|^2}{2} = \frac{\xi\|\mathcal{L}_b^{-1}u\|^2}{2} \geq \frac{\xi}{2\|\mathcal{L}_b\|^2}\|u\|^2,$$

which yields

$$\|\mathcal{L}_b^{-1}\mathcal{L}_a u + v\| \geq \frac{\xi}{2\|\mathcal{L}_b\|^2}\|u\|. \tag{3.17}$$

Notice that (3.17) also holds when $u = 0$. Since

$$\|\mathcal{L}_b^{-1}\mathcal{L}_a u + v\| = \|\mathcal{L}_b^{-1}(\mathcal{L}_a u + \mathcal{L}_b v)\| \leq \|\mathcal{L}_b^{-1}\|\|\mathcal{L}_a u + \mathcal{L}_b v\|,$$

we have from (3.17) that

$$\|\mathcal{L}_a u + \mathcal{L}_b v\| \geq \frac{\xi}{2\|\mathcal{L}_b^{-1}\|\|\mathcal{L}_b\|^2}\|u\|. \tag{3.18}$$

Similarly, we can also show

$$\|\mathcal{L}_a u + \mathcal{L}_b v\| \geq \frac{\xi}{2\|\mathcal{L}_a^{-1}\|\|\mathcal{L}_a\|^2}\|v\|. \tag{3.19}$$

Then, (3.14) follows from (3.18) and (3.19) with

$$\theta = \min\left\{\frac{\xi}{4\|\mathcal{L}_b^{-1}\|\|\mathcal{L}_b\|^2}, \frac{\xi}{4\|\mathcal{L}_a^{-1}\|\|\mathcal{L}_a\|^2}\right\}.$$

This completes the proof. □

Theorem 4 *Let $\mathcal{H}'(z)$ be defined by (3.13). If Assumption 3.1 holds, then $\mathcal{H}'(z)$ is nonsingular at any $z = (\mu, x, s, y) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$.*

Proof For any $(\mu, x, s, y) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$, by the expression of $\mathcal{H}'(z)$, it suffices to prove that the following system

$$F'_x(x, s, y)\Delta x + F'_s(x, s, y)\Delta s + F'_y(x, s, y)\Delta y = 0, \tag{3.20}$$

and

$$\psi'_x(\mu, x, s)\Delta x + \psi'_s(\mu, x, s)\Delta s = 0, \tag{3.21}$$

has only zero solution. By (3.8), (3.9) and (3.21), we have

$$\left[I - \mathcal{L}_{c(\mu,x,s)}^{-1}\mathcal{L}_{x+(\tau/2-1)s}\right]\Delta x + \left[I - \mathcal{L}_{c(\mu,x,s)}^{-1}\mathcal{L}_{s+(\tau/2-1)x}\right]\Delta s = 0,$$

which is equivalent to

$$\mathcal{L}_{c(\mu,x,s)-[x+(\tau/2-1)s]}\Delta x + \mathcal{L}_{c(\mu,x,s)-[s+(\tau/2-1)x]}\Delta s = 0. \tag{3.22}$$

From (3.2) and (3.10), we have

$$\begin{aligned} c(\mu, x, s) &= \sqrt{[x + (\tau/2 - 1)s]^2 + \tau(1 - \tau/4)s^2 + (4 - \tau)w + 4\mu^t e} \\ &= \sqrt{[s + (\tau/2 - 1)x]^2 + \tau(1 - \tau/4)x^2 + (4 - \tau)w + 4\mu^t e}. \end{aligned}$$

So, it follows from $\tau \in [0, 4)$, $w \in \mathcal{K}$ and $\mu > 0$ that $c(\mu, x, s) \succ_{\mathcal{K}} 0$ and

$$c(\mu, x, s)^2 \succ_{\mathcal{K}} [x + (\tau/2 - 1)s]^2, \ \ c(\mu, x, s)^2 \succ_{\mathcal{K}} [s + (\tau/2 - 1)x]^2,$$

which together with (ii) in Lemma 1 gives

$$c(\mu, x, s) - |x + (\tau/2 - 1)s| \succ_{\mathcal{K}} 0, \ \ c(\mu, x, s) - |s + (\tau/2 - 1)x| \succ_{\mathcal{K}} 0,$$

and hence

$$c(\mu, x, s) - [x + (\tau/2 - 1)s] \succ_{\mathcal{K}} 0, \ \ c(\mu, x, s) - [s + (\tau/2 - 1)x] \succ_{\mathcal{K}} 0. \tag{3.23}$$

In addition, by $\tau \in [0, 4)$, $w \in \mathcal{K}$ and $\mu > 0$, it holds that

$$\begin{aligned} &\left\{c(\mu, x, s) - [x + (\tau/2 - 1)s]\right\} \circ \left\{c(\mu, x, s) - [s + (\tau/2 - 1)x]\right\} \\ &\quad = c(\mu, x, s)^2 - \tau/2c(\mu, x, s) \circ (x + s) + [x + (\tau/2 - 1)s] \circ [s + (\tau/2 - 1)x] \\ &\quad = \tau/4[c(\mu, x, s) - (x + s)]^2 + (2 - \tau/2)^2 w + (4 - \tau)\mu^t e \succ_{\mathcal{K}} 0. \end{aligned} \tag{3.24}$$

Moreover, from Assumption 3.1 and (3.20) we have

$$\langle \triangle x, \triangle s \rangle \geq 0.$$

Hence, by Lemma 2, we can obtain from (3.23) and (3.24) that

$$\|\mathcal{L}_{c(\mu,x,s)-[x+(\tau/2-1)s]} \Delta x + \mathcal{L}_{c(\mu,x,s)-[s+(\tau/2-1)x]} \Delta s\| \geq \xi(\|\Delta x\| + \|\Delta s\|),$$

where $\xi > 0$ is a constant. This together with (3.22) implies $\Delta x = 0$ and $\Delta s = 0$. So, by (3.20), we have $F'_y(x, s, y)\Delta y = 0$, which and the assumption on rank $F'_y(x, s, y) = m$ give $\Delta y = 0$. We complete the proof. □

4 The Algorithm

Let $\mathcal{H}(z)$ be given in (3.11) and define the merit function $\mathcal{M} : \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \to \mathcal{R}_+$ by

$$\mathcal{M}(z) = \|\mathcal{H}(z)\|^2. \tag{4.1}$$

We now describe our algorithm to solve $\mathcal{H}(z) = 0$ by minimizing the merit function $\mathcal{M}(z)$.

Algorithm 4.1 (*A Nonmonotone Smoothing Newton Algorithm (NSNA) for WCP*)

Step 1: Choose $\delta \in (0, 1), \sigma \in (0, 1/2)$ and $z^0 := (\mu_0, x^0, s^0, y^0) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$. Let $\mathcal{C}_0 := \mathcal{M}(z^0)$. Choose $\gamma \in (0, 1)$ such that $\gamma \le \mu_0$. Set $k := 0$.

Step 2: If $\|\mathcal{H}(z^k)\| = 0$, then stop. Else, compute

$$\beta_k := \gamma \min\{1, \mathcal{C}_k\}. \tag{4.2}$$

Step 3: Compute the search direction $\Delta z^k = (\Delta\mu_k, \Delta x^k, \Delta s^k, \Delta y^k) \in \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ by solving the perturbed Newton system:

$$\mathcal{H}'(z^k)\Delta z^k = -\mathcal{H}(z^k) + \beta_k h, \tag{4.3}$$

where $h := (1, 0, 0, 0) \in \mathcal{R} \times \mathbb{V} \times \mathcal{R}^m \times \mathbb{V}$.
Step 4: Let α_k be the maximum of the values $1, \delta, \delta^2, \ldots$ such that

$$\mathcal{M}(z^k + \alpha_k \Delta z^k) \le [1 - 2\sigma(1 - \gamma)\alpha_k]\mathcal{C}_k. \tag{4.4}$$

Step 5: Set $z^{k+1} := z^k + \alpha_k \Delta z^k$. Compute $\mathcal{M}(z^{k+1}) = \|\mathcal{H}(z^{k+1})\|^2$ and set

$$\mathcal{C}_{k+1} := \frac{(\mathcal{C}_k + 1)\mathcal{M}(z^{k+1})}{\mathcal{M}(z^{k+1}) + 1}. \tag{4.5}$$

Set $k := k + 1$. Go to Step 2.

Motivated from the techniques given in [12], the reference value $\mathcal{C}_k$ in the line search of NSNA is updated in an average way, which could not only ensure global convergence but also improve practical performance of NSNA. The following theorem shows NSNA, especially the Newton system (4.3) and the nonmonotone line search (4.4), is well defined.

Theorem 5 *If Assumption 3.1 holds, Algorithm 4.1 is well defined and its generated sequence $\{z_k = (\mu_k, x^k, s^k, y^k)\}$ satisfies $\mu_k > 0$ and $\mathcal{M}(z^k) \le \mathcal{C}_k$ for all $k \ge 0$.*

Proof Suppose that $z^k \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and $\mathcal{M}(z^k) \le \mathcal{C}_k$ for some k. By Theorem 4, $\mathcal{H}'(z^k)$ is nonsingular. So, Step 3 is well defined at the kth iteration. Moreover, from (4.3) we have

$$\mathcal{M}'(z^k)\Delta z^k = 2\mathcal{H}(z^k)^T \mathcal{H}'(z^k)\Delta z^k = -2\mathcal{M}(z^k) + 2\mu_k\beta_k. \tag{4.6}$$

For any $\alpha \in (0, 1]$, we denote

$$R(\alpha) := \mathcal{M}(z^k + \alpha\Delta z^k) - \mathcal{M}(z^k) - \alpha\mathcal{M}'(z^k)\Delta z^k. \tag{4.7}$$

Since $\mathcal{M}(z^k) \le \mathcal{C}_k$, by (3.11) and (4.1) we have $\mu_k \le \sqrt{\mathcal{C}_k}$. From (4.2), it holds that $\beta_k \le \gamma\sqrt{\mathcal{C}_k}$ because $\min\{1, a\} \le \sqrt{a}$ for any $a \ge 0$. Using these results, we can obtain from (4.6) and (4.7) that for any $\alpha \in (0, 1]$

$$\begin{aligned}\mathcal{M}(z^k + \alpha \Delta z^k) &= \mathcal{M}(z^k) + \alpha \mathcal{M}'(z^k)\Delta z^k + R(\alpha) \\ &= (1 - 2\alpha)\mathcal{M}(z^k) + 2\alpha\mu_k\beta_k + R(\alpha) \\ &\le (1 - 2\alpha)\mathcal{C}_k + 2\alpha\gamma\mathcal{C}_k + R(\alpha) \\ &= [1 - 2(1 - \gamma)\alpha]\mathcal{C}_k + R(\alpha). \end{aligned} \tag{4.8}$$

Since $\mathcal{M}$ is continuously differentiable at $z^k \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$, we have $R(\alpha) = o(\alpha)$. This together with (4.8) implies that the line search (4.4) is well defined. So, we can find a step-size $\alpha_k \in (0, 1]$ in Step 4 and get the $(k + 1)$th iteration $z^{k+1} = z^k + \alpha_k \Delta z^k$ in Step 5. Now we prove $z^{k+1} \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and $\mathcal{M}(z^{k+1}) \le \mathcal{C}_{k+1}$. Since $\mu_k > 0$, we have $\mathcal{C}_k \ge \mathcal{M}(z^k) \ge \mu_k^2 > 0$ and hence $\beta_k > 0$. By the first equation of (4.3), we get $\Delta\mu_k = -\mu_k + \beta_k$. Thus,

$$\mu_{k+1} = \mu_k + \alpha_k\Delta\mu_k = (1 - \alpha_k)\mu_k + \alpha_k\beta_k > 0. \tag{4.9}$$

This proves $z^{k+1} \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$. Moreover, by Steps 4 and 5 we have $\mathcal{M}(z^{k+1}) \le \mathcal{C}_k$, which together with (4.5) implies that $\mathcal{C}_{k+1} \ge \mathcal{M}(z^{k+1})$. So, we can conclude that if $z^k \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and $\mathcal{M}(z^k) \le \mathcal{C}_k$, then z^{k+1} can be generated by Algorithm 4.1 with $z^{k+1} \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and $\mathcal{M}(z^{k+1}) \le \mathcal{C}_{k+1}$. This together with $z^0 \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m$ and $\mathcal{M}(z^0) = \mathcal{C}_0$ gives the desired result. □

5 Convergence Analysis

5.1 Global Convergence

Lemma 3 *Suppose that Assumption 3.1 holds. Let* $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ *be the iteration sequence generated by Algorithm 4.1. Then,* $\mathcal{C}_k \ge \mathcal{C}_{k+1}$, $\mu_k \ge \beta_k$ *and* $\mu_k \ge \mu_{k+1}$ *for all* $k \ge 0$.

Proof By Steps 4 and 5, we have $\mathcal{M}(z^{k+1}) \le \mathcal{C}_k$ for all $k \ge 0$. Then, it follows from (4.5) that for all $k \ge 0$

$$\mathcal{C}_{k+1} = \frac{\mathcal{C}_k\mathcal{M}(z^{k+1}) + \mathcal{M}(z^{k+1})}{\mathcal{M}(z^{k+1}) + 1} \le \frac{\mathcal{C}_k\mathcal{M}(z^{k+1}) + \mathcal{C}_k}{\mathcal{M}(z^{k+1}) + 1} = \mathcal{C}_k.$$

Moreover, by Step 1 and (4.2), $\mu_0 \ge \gamma \ge \gamma \min\{1, \mathcal{C}_0\} = \beta_0$. Suppose that $\mu_k \ge \beta_k$ for some k. Then, by (4.9) we have

$$\mu_{k+1} \ge (1 - \alpha_k)\beta_k + \alpha_k\beta_k = \beta_k = \gamma \min\{1, \mathcal{C}_k\} \ge \gamma \min\{1, \mathcal{C}_{k+1}\} = \beta_{k+1}.$$

Thus, $\mu_k \ge \beta_k$ for all $k \ge 0$. Using this result, we can further obtain from (4.9) that

$$\mu_{k+1} \le (1 - \alpha_k)\mu_k + \alpha_k\mu_k = \mu_k,$$

for all $k \geq 0$. □

Based on Lemma 3, we have the following global convergence theorem.

Theorem 6 *Suppose that Assumption 3.1 holds. Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1. If there exists an accumulation point of $\{z^k\}$, we have*

$$\lim_{k\to\infty} \|\mathcal{H}(z^k)\| = 0,$$

and at any accumulation point z^*, we have $\mathcal{H}(z^*) = 0$.

Proof Let z^* be any accumulation point and there exists a subsequence, still denoted as $\{z^k\}$, converging to z^*. From Lemma 3, $\{\mathcal{C}_k\}$ is monotonically decreasing, and hence, it is convergent. So, there exists $\mathcal{C}^* \geq 0$ such that $\lim_{k\to\infty} \mathcal{C}_k = \mathcal{C}^*$. If $\mathcal{C}^* = 0$, then $\lim_{k\to\infty} \|\mathcal{H}(z^k)\| = 0$, since $\mathcal{M}(z^k) = \|\mathcal{H}(z^k)\|^2 \leq \mathcal{C}_k$ for all $k \geq 0$. Then, from the continuity of $\mathcal{H}(z)$ we have $\mathcal{H}(z^*) = 0$. Now we suppose that $\mathcal{C}^* > 0$ and derive a contradiction. By (4.5) we have

$$\lim_{k\to\infty} \mathcal{M}(z^{k+1}) = \lim_{k\to\infty} \left(\frac{\mathcal{C}_{k+1}}{1 + \mathcal{C}_k - \mathcal{C}_{k+1}} \right) = \mathcal{C}^* > 0. \tag{5.1}$$

From Steps 4 and 5, it holds that

$$\mathcal{M}(z^{k+1}) \leq [1 - 2\sigma(1-\gamma)\alpha_k]\mathcal{C}_k, \tag{5.2}$$

which together with (5.1) implies that $\lim_{k\to\infty} \alpha_k = 0$. Let $\hat{\alpha}_k := \alpha_k/\delta$. Then, for all k sufficiently large, $\hat{\alpha}_k$ does not satisfy the search criterion (4.4), i.e.,

$$\mathcal{M}(z^k + \hat{\alpha}_k \Delta z^k) > [1 - 2\sigma(1-\gamma)\hat{\alpha}_k]\mathcal{C}_k \geq \mathcal{M}(z^k) - 2\sigma(1-\gamma)\hat{\alpha}_k\mathcal{C}_k,$$

where the second inequality follows from $\mathcal{M}(z^k) \leq \mathcal{C}_k$ for all $k \geq 0$. Thus,

$$\frac{\mathcal{M}(z^k + \hat{\alpha}_k \Delta z^k) - \mathcal{M}(z^k)}{\hat{\alpha}_k} \geq -2\sigma(1-\gamma)\mathcal{C}_k. \tag{5.3}$$

By Lemma 3 and (4.2), we have

$$\mu^* = \lim_{k\to\infty} \mu_k \geq \lim_{k\to\infty} \beta_k = \beta^* := \gamma \min\{1, \mathcal{C}^*\} > 0.$$

Thus, $\mathcal{M}$ is continuously differentiable at z^*. So, by letting $k \to \infty$ in (5.3) we have

$$2\mathcal{H}(z^*)^T \mathcal{H}'(z^*)\Delta z^* \geq -2\sigma(1-\gamma)\mathcal{C}^*, \tag{5.4}$$

where Δz^* is the solution of $\mathcal{H}'(z^*)\Delta z^* = -\mathcal{H}(z^*) + \beta^* h$. In addition, by (4.6),

$$\begin{aligned}\mathcal{H}(z^*)^T \mathcal{H}'(z^*)\Delta z^* &= -\mathcal{M}(z^*) + \mu^* \beta^* \\ &= -\mathcal{M}(z^*) + \mu^* \gamma \min\{1, \mathcal{C}^*\} \\ &\leq -(1-\gamma)\mathcal{C}^*, \end{aligned} \tag{5.5}$$

where the inequality holds since $\mathcal{M}(z^*) \leq \mathcal{C}^*$, $\mu^* \leq \|\mathcal{H}(z^*)\| \leq \sqrt{\mathcal{C}^*}$ and $\min\{1, \mathcal{C}^*\} \leq \sqrt{\mathcal{C}^*}$. By combining (5.4) and (5.5), we have $\sigma(1-\gamma)\mathcal{C}^* \geq (1-\gamma)\mathcal{C}^*$, which together with $\mathcal{C}^* > 0$ implies that $\sigma(1-\gamma) \geq 1-\gamma$. This contradicts the fact that $\sigma \in (0, 1/2)$ and $\gamma \in (0, 1)$. Thus, $\mathcal{C}^* = 0$ and $\mathcal{H}(z^*) = 0$. □

By Theorem 5 and Lemma 3, we have $\mathcal{M}(z^k) \leq \mathcal{C}_k \leq \mathcal{C}_0 = \mathcal{M}(z^0)$ for all $k \geq 0$. Hence, if the level set $L(z^0) := \{z \in \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \,|\, \mathcal{M}(z) \leq \mathcal{M}(z^0)\}$ is bounded, we will have $\lim\limits_{k\to\infty} \|\mathcal{H}(z^k)\| = 0$.

5.2 Local Superlinear and Quadratic Convergence

In this subsection, we analyze the local convergence properties of NSNA. For $\mathcal{H}(z)$ given by (3.11), let $\mathcal{Z}^*$ be the solution set of $\mathcal{H}(z) = 0$, i.e.,

$$\mathcal{Z}^* := \{z = (0, x, s, y) \in \mathcal{R} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \,|\, \mathcal{H}(z) = 0\}. \tag{5.6}$$

Let $\mathcal{S}^*$ be the solution set of WCP (1.1), i.e.,

$$\mathcal{S}^* := \{(x, s, y) \in \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \,|\, x \in \mathcal{K},\ s \in \mathcal{K},\ F(x, s, y) = 0,\ x \circ s = w\}.$$

Then, by (3.12) we have

$$z = (0, x, s, y) \in \mathcal{Z}^* \Longleftrightarrow (x, s, y) \in \mathcal{S}^*.$$

Thus, $\mathcal{Z}^*$is nonempty if and only if $\mathcal{S}^*$ is nonempty.

Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1. For analyzing the local convergence rate of NSNA, we make the following assumptions.

Assumption 5.1 There exist constants $C > 0$ and $d \in [0, 1/2)$ such that

$$\|\mathcal{H}'(z^k)^{-1}\| \leq \frac{C}{\mu_k^d},$$

for all k sufficiently large.

For Assumption 5.1, we have the following remarks.

(i) To obtain local fast convergence, many smoothing-type algorithms require that $\{\|\mathcal{H}'(z^k)^{-1}\|\}$ is uniformly bounded (e.g., [3,38]) or the nonsingular Jacobian condition

(1.4), while Assumption 5.1 allows $\{\|\mathcal{H}'(z^k)^{-1}\|\}$ to be unbounded when $d \in (0, 1/2)$. In what follows, we show NSNA has local quadratic convergence if Assumption 5.1 holds with $d = 0$, i.e., $\{\|\mathcal{H}'(z^k)^{-1}\|\}$ is uniformly bounded, but it will still has local superlinear convergence when Assumption 5.1 holds with $d \in (0, 1/2)$.
(ii) As an example, Assumption 5.1 holds for the following linear weighted complementarity problem:

$$x \in \mathcal{K}, \quad s \in \mathcal{K}, \quad s = Mx + a, \quad xs = w, \tag{5.7}$$

where $\mathcal{K} = \mathcal{R}^n_+$, $M \in \mathcal{R}^{n\times n}$ is a positive definite matrix, $a \in \mathcal{R}^n$ and $w \in \mathcal{R}^n_+$ is the weight vector. This problem is a special case of LWCP (1.3) with $P = M$, $Q = -I$, $R = 0$. We give the proof in 'Appendix." In fact, with more complicated proof, we can actually show Assumption 5.1 holds for (5.7) with $\mathcal{K}$ being the general second-order cone.

Assumption 5.2 $\lim\limits_{k\to\infty} \|\mathcal{H}(z^k)\| = 0$ and there exists a constant $\eta > 0$ such that

$$\|\mathcal{H}(z^k)\| \geq \eta \mathrm{dist}(z^k, \mathcal{Z}^*), \tag{5.8}$$

for all $\|\mathcal{H}(z^k)\|$ sufficiently small.

Condition (5.8) is a type of *local error bound condition*. Local error bound conditions and its applications were proposed and analyzed in [23,28], which are in general weaker than the nonsingularity assumption on the Jacobian of a nonlinear system of equations at its solution set. Local error bound conditions have been recently used extensively to study local convergence behaviors of Levenberg–Marquardt methods for solving nonlinear system of equations when the Jacobian is singular [7,8]. Assumption 5.2 also assumes the residue $\|\mathcal{H}(z^k)\|$ goes to zero as $k \to \infty$, which by Theorem 6 is ensured if there is one accumulation point of $\{z^k\}$. Under Assumption 5.1, we now give an other condition of ensuring $\|\mathcal{H}(z^k)\|$ goes to zero, which could be satisfied for unbounded solution set.

Theorem 7 *Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1. If Assumption 5.1 holds and for any $0 < \theta < \mu_0$, $\mathcal{M}'(\cdot)$ is Lipschitz continuous on the set*

$$\Theta := \{(\mu, x, s, y) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m |\ \mu \geq \theta,\ \|\mathcal{H}(z)\| \leq 2\|\mathcal{H}(z^0)\|\}, \tag{5.9}$$

we have

$$\lim_{k\to\infty} \|\mathcal{H}(z^k)\| = 0. \tag{5.10}$$

Proof By Lemma 3, $\{\mu_k\}$ and $\{\mathcal{C}_k\}$ are monotonically decreasing. Hence, there exist $\mu^* \geq 0$ and $\mathcal{C}^* \geq 0$ such that $\lim\limits_{k\to\infty} \mu_k = \mu^*$ and $\lim\limits_{k\to\infty} \mathcal{C}_k = \mathcal{C}^*$. Moreover, by Lemma 3 and (4.2), we have $\mu_k \geq \beta_k = \gamma \min\{1, \mathcal{C}_k\}$ for all $k \geq 0$. Then, if $\mu^* = 0$, we have $\lim\limits_{k\to\infty} \mathcal{C}_k = 0$, which implies (5.10). Hence, to complete the proof, we only need to show $\mu^* = 0$.

In the following, assuming $\mu^* > 0$, we derive a contradiction. Since $\mu_k \geq \mu^*$ for all k, it follows from Assumption 5.1 that

$$\|\mathcal{H}'(z^k)^{-1}\| \leq \frac{C}{\mu_k^d} \leq \frac{C}{(\mu^*)^d},$$

for all k sufficiently large. Hence, by $\|\mathcal{H}(z^k)\| \leq \|\mathcal{H}(z^0)\|$ and $\beta_k \leq \gamma$, we have

$$\|\Delta z^k\| \leq \|\mathcal{H}'(z^k)^{-1}\| \|-\mathcal{H}(z^k) + \beta_k h\| \leq \frac{C}{(\mu^*)^d}(\|\mathcal{H}(z^0)\| + \gamma) \tag{5.11}$$

for all k sufficiently large. Now, since $\mu_k^2 \leq \mathcal{M}(z^k) \leq \mathcal{C}_k$ for all k, we have $\mathcal{C}^* = \lim\limits_{k\to\infty} \mathcal{C}_k \geq \lim\limits_{k\to\infty} \mu_k^2 = (\mu^*)^2 > 0$. Then, it follows from (5.1), (5.2) and the same arguments as those in the proof of Theorem 6 that $\lim\limits_{k\to\infty} \hat{\alpha}_k = 0$ and

$$\mathcal{M}(z^k + \hat{\alpha}_k \Delta z^k) > [1 - 2\sigma(1-\gamma)\hat{\alpha}_k]\mathcal{C}_k, \tag{5.12}$$

for all k sufficiently large, where $\hat{\alpha}_k = \alpha_k/\delta$. So, by (5.11), we have

$$\lim_{k\to\infty} \hat{\alpha}_k \|\Delta z^k\| = 0. \tag{5.13}$$

By (5.9), $\mathcal{M}'(\cdot)$ is Lipschitz continuous on

$$\{(\mu, x, s, y) \in \mathcal{R}_{++} \times \mathbb{V} \times \mathbb{V} \times \mathcal{R}^m \mid \mu \geq \mu^*/2,\ \|\mathcal{H}(z)\| \leq 2\|\mathcal{H}(z^0)\|\}$$

with a Lipschitz constant $L > 0$. So, we have from (5.13), $\mu_k \geq \mu^* > 0$ and $\|\mathcal{H}(z^k)\| \leq \|\mathcal{H}(z^0)\|$ that $\mathcal{M}'(\cdot)$ is Lipschitz continuous on the line segment connecting z^k to $z^k + \hat{\alpha}_k \Delta z^k$ for all sufficiently large k. Therefore, we have

$$\left|\mathcal{M}(z^k + \hat{\alpha}_k \Delta z^k) - \mathcal{M}(z^k) - \hat{\alpha}_k \mathcal{M}'(z^k)\Delta z^k\right| \leq \frac{L}{2}(\hat{\alpha}_k \|\Delta z^k\|)^2 \tag{5.14}$$

for all k sufficiently large. Then, we have

$$\begin{aligned}
\mathcal{M}(z^k + \hat{\alpha}_k \Delta z^k) &\leq \mathcal{M}(z^k) + \hat{\alpha}_k \mathcal{M}'(z^k)\Delta z^k + \frac{L}{2}(\hat{\alpha}_k \|\Delta z^k\|)^2 \\
&= (1 - 2\hat{\alpha}_k)\mathcal{M}(z^k) + 2\hat{\alpha}_k \mu_k \beta_k + \frac{L}{2}(\hat{\alpha}_k \|\Delta z^k\|)^2 \\
&\leq (1 - 2\hat{\alpha}_k)\mathcal{C}_k + 2\hat{\alpha}_k \gamma \mathcal{C}_k + \frac{L}{2}(\hat{\alpha}_k \|\Delta z^k\|)^2 \\
&= [1 - 2(1-\gamma)\hat{\alpha}_k]\mathcal{C}_k + \frac{L}{2}(\hat{\alpha}_k \|\Delta z^k\|)^2, \qquad (5.15)
\end{aligned}$$

where the first equality holds by (4.6) and the second inequality uses the fact that $\mathcal{M}(z^k) \le \mathcal{C}_k$, $\mu_k \le \sqrt{\mathcal{C}_k}$ and $\beta_k \le \gamma\sqrt{\mathcal{C}_k}$ for all k. By (5.12) and (5.15), we get

$$2(1-\sigma)(1-\gamma)\mathcal{C}_k \le \frac{L\hat{\alpha}_k\|\Delta z^k\|^2}{2},$$

for all k sufficiently large. Taking $k \to \infty$ in the above equation, it follows from (5.11), (5.13) and $(1-\sigma)(1-\gamma) > 0$ that $\mathcal{C}^* = 0$, which contradicts with $\mathcal{C}^* > 0$. This completes our proof. □

Suppose that the set $\mathcal{Z}^*$ defined by (5.6) is nonempty. Since $\mathcal{Z}^*$ is a closed set, for any $z = (\mu, x, s, y)$, let $\bar{z} \in \mathcal{Z}^*$ be one vector satisfying

$$\|z - \bar{z}\| = \mathrm{dist}(z, \mathcal{Z}^*). \tag{5.16}$$

For local fast convergence, we also need the local strong semismoothness of $\mathcal{H}(\cdot)$.

Assumption 5.3 $\mathcal{H}(\cdot)$ is strongly semismooth with respect to the set $\mathcal{Z}^*$, that is, there exist constants $M > 0$ and $L > 0$ such that

$$\|\mathcal{H}(z)\| = \|\mathcal{H}(z) - \mathcal{H}(\bar{z})\| \le M\|z - \bar{z}\| \tag{5.17}$$

and

$$\|\mathcal{H}(z) - \mathcal{H}'(z)(z - \bar{z})\| \le \frac{L}{2}\|z - \bar{z}\|^2, \tag{5.18}$$

whenever $\|z - \bar{z}\| = \mathrm{dist}(z, \mathcal{Z}^*)$ is sufficiently small.

Local strong semismoothness of $\mathcal{H}(\cdot)$ is indeed a generalization of the standard definition of strongly semismooth of a function at one point. When $\mathcal{Z}^* = \{z^*\}$ is a singleton, Assumption 5.3 will be simply reduced to the assumption that $\mathcal{H}(\cdot)$ is strongly semismooth at z^*. One may refer [32] for the standard definition of strongly semismoothness of a locally Lipschitz continuous function.

We now discuss the local superlinear and quadratic convergence properties of Algorithm 4.1.

Theorem 8 *Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1. Suppose $\mathcal{Z}^*$ is nonempty and Assumptions 3.1, 5.1, 5.2 and 5.3 hold. Then, for all sufficiently large k, we have*

$$z^{k+1} = z^k + \Delta z^k, \tag{5.19}$$

and

$$dist(z^{k+1}, \mathcal{Z}^*) = \mathcal{O}(dist(z^k, \mathcal{Z}^*)^{2-2d}), \tag{5.20}$$

where $d \in [0, 1/2)$.

Proof Assumption 5.2 implies that $\lim_{k\to\infty} \mathrm{dist}(z^k, \mathcal{Z}^*) = \lim_{k\to\infty} \|z^k - \bar{z}^k\| = 0$. Since $\{\mathcal{C}_k\}$ is monotonically decreasing, it is convergent. So, from (4.5) and Assumption 5.2 we have $\lim_{k\to\infty} \mathcal{C}_k = 0$. Thus, by (4.1), (4.2) and (4.5), for all sufficiently large k,

$$\beta_k = \gamma \mathcal{C}_k = \frac{\gamma(\mathcal{C}_{k-1}+1)\mathcal{M}(z^k)}{\mathcal{M}(z^k)+1} \le \gamma(\mathcal{C}_0+1)\|\mathcal{H}(z^k)\|^2. \tag{5.21}$$

Hence, by (5.17), for all sufficiently large k,

$$\beta_k \le \gamma(\mathcal{C}_0+1)\|\mathcal{H}(z^k) - \mathcal{H}(\bar{z}^k)\|^2 \le \gamma(\mathcal{C}_0+1)M^2\|z^k - \bar{z}^k\|^2. \tag{5.22}$$

Then, it follows from (4.3), (5.22), (5.18) and Assumption 5.1 that for all sufficiently large k,

$$\begin{aligned} \|z^k + \Delta z^k - \bar{z}^k\| &= \|z^k + \mathcal{H}'(z^k)^{-1}[-\mathcal{H}(z^k) + \beta_k h] - \bar{z}^k\| \\ &\le \|\mathcal{H}'(z^k)^{-1}\| \Big[\|\mathcal{H}(z^k) - \mathcal{H}'(z^k)(z^k - \bar{z}^k)\| + \beta_k \Big] \\ &\le \frac{\bar{C}}{\mu_k^d}\|z^k - \bar{z}^k\|^2, \end{aligned} \tag{5.23}$$

where $d \in [0, 1/2)$ and $\bar{C} := C(L/2 + \gamma(\mathcal{C}_0+1)M^2)$. By Lemma 3, (5.21), Theorem 5 and Assumption 5.2, for all sufficiently large k,

$$\begin{aligned} \mu_k \ge \beta_k = \gamma\mathcal{C}_k \ge \gamma\mathcal{M}(z^k) &= \gamma\|\mathcal{H}(z^k)\|^2 \\ \ge \gamma\eta^2 \mathrm{dist}(z^k, \mathcal{Z}^*)^2 &= \gamma\eta^2\|z^k - \bar{z}^k\|^2. \end{aligned}$$

So, for all $d \in [0, 1/2)$ and sufficiently large k,

$$\frac{1}{\mu_k^d} \le \frac{1}{\gamma^d\eta^{2d}\|z^k - \bar{z}^k\|^{2d}}. \tag{5.24}$$

Hence, by combining (5.23) and (5.24), we have

$$\|z^k + \Delta z^k - \bar{z}^k\| \le \frac{\bar{C}}{\gamma^d\eta^{2d}}\|z^k - \bar{z}^k\|^{2-2d}. \tag{5.25}$$

So, $\mathrm{dist}(z^k + \Delta z^k, \mathcal{Z}^*)$ is sufficiently small when k is sufficiently large. Hence, it follows (5.17) that for any $d \in [0, 1/2)$ and sufficiently large k,

$$\|\mathcal{H}(z^k + \Delta z^k)\| \le M \mathrm{dist}(z^k + \Delta z^k, \mathcal{Z}^*) \le M\|z^k + \Delta z^k - \bar{z}^k\|. \tag{5.26}$$

Moreover, by Assumption 5.2, for any $d \in [0, 1/2)$ and sufficiently large k,

$$\|\mathcal{H}(z^k)\|^{2-2d} \ge \eta^{2-2d}\mathrm{dist}(z^k, \mathcal{Z}^*)^{2-2d} = \eta^{2-2d}\|z^k - \bar{z}^k\|^{2-2d}. \tag{5.27}$$

Then, we have from (5.25)–(5.27) that for any $d \in [0, 1/2)$ and sufficiently large k,

$$\|\mathcal{H}(z^k + \Delta z^k)\| \leq \tilde{C}\|\mathcal{H}(z^k)\|^{2-2d}, \tag{5.28}$$

where $\tilde{C} := M\bar{C}/(\gamma^d \eta^2)$, which implies

$$\lim_{k\to\infty} \frac{\mathcal{M}(z^k + \Delta z^k)}{\mathcal{M}(z^k)} = \lim_{k\to\infty} \frac{\|\mathcal{H}(z^k + \Delta z^k)\|^2}{\|\mathcal{H}(z^k)\|^2} = 0.$$

Hence, for all sufficiently large k, $\alpha_k = 1$ satisfies

$$\mathcal{M}(z^k + \alpha_k \Delta z^k) \leq [1 - 2\sigma(1-\gamma)\alpha_k]\mathcal{M}(z^k) \leq [1 - 2\sigma(1-\gamma)\alpha_k]\mathcal{C}_k.$$

Therefore, for all sufficiently large k,

$$z^{k+1} = z^k + \Delta z^k.$$

So, by (5.25), for all $d \in [0, 1/2)$ and sufficiently large k, we have

$$\|z^{k+1} - \bar{z}^k\| \leq \frac{\bar{C}}{\gamma^d \eta^{2d}}\|z^k - \bar{z}^k\|^{2-2d},$$

which implies

$$\mathrm{dist}(z^{k+1}, \mathcal{Z}^*) \leq \|z^{k+1} - \bar{z}^k\| = \mathcal{O}(\mathrm{dist}(z^k, \mathcal{Z}^*)^{2-2d}).$$

Thus, the proof is completed. □

Now, under Assumptions 5.1, 5.2 and 5.3, we show that the iteration sequence $\{z^k\}$ is bounded and it converges to some point $z^* \in \mathcal{Z}^*$ locally superlinearly or quadratically.

Theorem 9 *Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1. Suppose the solution set $\mathcal{Z}^*$ is nonempty and Assumptions 3.1, 5.1, 5.2 and 5.3 hold. Then, we have $\{z^k\}$ is bounded and converges to some point $z^* \in \mathcal{Z}^*$. Moreover, for all sufficiently large k, we have*

$$\|z^{k+1} - z^*\| = \mathcal{O}(\|z^k - z^*\|^{2-2d}), \tag{5.29}$$

where $d \in [0, 1/2)$.

Proof By (5.25), for all $d \in [0, 1/2)$ and sufficiently large k,

$$\begin{aligned} \|\Delta z^k\| &\leq \|z^k + \Delta z^k - \bar{z}^k\| + \|z^k - \bar{z}^k\| \\ &= \mathcal{O}(\|z^k - \bar{z}^k\|^{2-2d}) + \|z^k - \bar{z}^k\| \\ &= \mathcal{O}(\|z^k - \bar{z}^k\|) = \mathcal{O}(\mathrm{dist}(z^k, \mathcal{Z}^*)). \end{aligned} \tag{5.30}$$

Moreover, by (5.20), for all sufficiently large k,

$$\frac{\operatorname{dist}(z^{k+1}, \mathcal{Z}^*)}{\operatorname{dist}(z^k, \mathcal{Z}^*)} = \mathcal{O}(\operatorname{dist}(z^k, \mathcal{Z}^*)^{1-2d}), \tag{5.31}$$

where $d \in [0, 1/2)$. Since $\lim_{k\to\infty} \operatorname{dist}(z^k, \mathcal{Z}^*)^{1-2d} = 0$ for $d \in [0, 1/2)$, we have from (5.31) that

$$\sum_{k=1}^{\infty} \operatorname{dist}(z^k, \mathcal{Z}^*) < \infty, \tag{5.32}$$

which together with (5.30) gives

$$\sum_{k=1}^{\infty} \|\Delta z^k\| < \infty. \tag{5.33}$$

Then, it follows from (5.19) that $\{z^k\}$ is a Cauchy sequence. Hence, there exists a z^* such that $\lim_{k\to\infty} z^k = z^*$. And, by Theorem 6, we have $z^* \in \mathcal{Z}^*$.

Now, by (5.19) and (5.31), for all sufficiently large k, we have

$$\begin{aligned}\operatorname{dist}(z^k, \mathcal{Z}^*) &\leq \|z^k - \bar{z}^{k+1}\| = \|z^{k+1} - \bar{z}^{k+1} - \Delta z^k\| \\ &\leq \operatorname{dist}(z^{k+1}, \mathcal{Z}^*) + \|\Delta z^k\| \\ &\leq \frac{1}{2}\operatorname{dist}(z^k, \mathcal{Z}^*) + \|\Delta z^k\|,\end{aligned}$$

which implies

$$\operatorname{dist}(z^k, \mathcal{Z}^*) \leq 2\|\Delta z^k\|. \tag{5.34}$$

Again, by (5.30) and (5.31), for all sufficiently large k, we have

$$\|\Delta z^{k+1}\| \leq \frac{1}{4}\operatorname{dist}(z^k, \mathcal{Z}^*).$$

This together with (5.34) gives

$$\|\Delta z^{k+1}\| \leq \frac{1}{2}\|\Delta z^k\| \tag{5.35}$$

for all sufficiently large k. So, when k is sufficiently large, (5.35) gives

$$\|z^{k+1} - z^*\| = \|\sum_{j=k+1}^{\infty} \Delta z^j\| \leq \sum_{j=k+1}^{\infty} \|\Delta z^j\| \leq 2\|\Delta z^{k+1}\|.$$

This together with (5.20) and (5.30) gives

$$\lim_{k\to\infty} \frac{\|z^{k+1} - z^*\|}{\|z^k - z^*\|^{2-2d}} \leq \lim_{k\to\infty} \frac{2\|\Delta z^{k+1}\|}{\operatorname{dist}(z^k, \mathcal{Z}^*)^{2-2d}} < \infty.$$

Hence, (5.29) holds. □

We can see that if Assumption 5.1 holds for $d = 0$, i.e., $\{\|\mathcal{H}'(z^k)^{-1}\|\}$ is uniformly bounded for all sufficiently large k, then under the conditions of Theorems 8 and 9, NSNA has local quadratic convergence. In addition, from Assumption 5.2, we have $\lim_{k\to\infty} \|\mathcal{H}(z^k)\| = 0$, which implies $\lim_{k\to\infty} \mu_k = 0$. Hence, when $d \in (0, 1/2)$, Assumption 5.1 does allow $\|\mathcal{H}'(z^k)^{-1}\|$ to grow up to infinity but no faster than $\mathcal{O}(1/\mu^d)$. However, in this case, NSNA would still has local superlinear convergence.

In Assumptions 5.2 and 5.3, we assume that $\{\|\mathcal{H}(z^k)\|\}$ converges to zero as $k \to \infty$ and $\mathcal{H}(z)$ is strongly semismooth with respect to the set $\mathcal{Z}^*$. In what follows, we show that under proper conditions these assumptions hold for LWCP (1.3). Here, we consider the smoothing function

$$\psi_c(\mu, a, b) := a + b - \sqrt{a^2 + b^2 + 2c + 4\mu^2},\ \forall\, (\mu, a, b) \in \mathcal{R}^3, \tag{5.36}$$

which corresponds to (3.5) with $\tau = 2$ and $t = 2$, where $c \geq 0$ is some constant. Then, from (3.6) it holds that

$$\psi_c(0, a, b) = 0 \Longleftrightarrow a \geq 0,\ \ b \geq 0,\ \ ab = c.$$

By using ψ_c, we can reformulate LWCP (1.3) as the nonlinear smooth equations

$$\mathcal{H}(z) := \mathcal{H}(\mu, x, s, y) = \begin{pmatrix} \mu \\ Px + Qs + Ry - a \\ \psi_{w_1}(\mu, x_1, s_1) \\ \vdots \\ \psi_{w_n}(\mu, x_n, s_n) \end{pmatrix} = 0 \tag{5.37}$$

and solve it by Algorithm 4.1, where $w = (w_1, ..., w_n)^T$ is the weight vector. Let $\{z^k = (\mu_k, x^k, s^k, y^k)\}$ be the iteration sequence generated by Algorithm 4.1 for solving LWCP (1.3). We first have the following lemma.

Lemma 4 *Let $\xi := (\mu, a, b)^T \in \mathcal{R}^3$ and $\psi_c(\xi)$ be defined by (5.36). Then,*

(i) $\psi_c(\xi)$ is Lipschitz continuous on $\mathcal{R}^3$ with a Lipschitz constant $M = \sqrt{2} + 2$;
(ii) for any $\xi \in \mathcal{R}^3$ any $V \in \partial\psi_c(\xi + h)$ and $h \to 0$,

$$|\psi_c(\xi + h) - \psi_c(\xi) - Vh| \leq L_c(\xi)\|h\|^2,$$

where

$$L_c(\xi) := \begin{cases} \frac{9}{2\sqrt{c}}, & if\ c > 0, \\ 0, & if\ c = 0,\ \xi = 0, \\ \frac{9\sqrt{2}}{\|\xi\|}, & if\ c = 0,\ \xi \neq 0. \end{cases} \tag{5.38}$$

Proof We prove this lemma in "Appendix." □

We show that if Assumption 5.1 holds, then $\{\|\mathcal{H}(z^k)\|\}$ converges to zero as $k \to \infty$. For this purpose, by Theorem 7, we only need to show the following lemma.

Lemma 5 *For* LWCP (1.3)*, let $\mathcal{M}(\cdot)$ be the merit function given in* (4.1) *and $\mathcal{H}(z)$ be defined by* (5.37)*. Then, for any $0 < \theta < \mu_0$, $\mathcal{M}'(\cdot)$ is Lipschitz continuous on the set Θ defined in* (5.9)*.*

Proof The proof is given in "Appendix." □

In terms of the strong smoothness of function $\mathcal{H}(\cdot)$ defined in (5.37), we have the following theorem.

Theorem 10 *The function $\mathcal{H}(\cdot)$ given in* (5.37) *is strongly semismooth with respect to the solution set $\mathcal{Z}^*$, if any one of the following conditions holds:*

(i) The weight vector $w > 0$;
(ii) The solution set $\mathcal{S}^*$ of LWCP *is singleton;*
(iii) The LWCP *is nondegenerate, i.e., for any $(x, s, y) \in \mathcal{S}^*$, we have $x + s > 0$.*

Proof By (i) of Lemma 4, we have $\psi_{w_i}, i = 1, \ldots, n$, is Lipschitz continuous on $\mathcal{Z}^*$ with Lipschitz constant $M = \sqrt{2} + 2$.

Now, let $z = (\mu, x, s, y)$ be any point sufficiently close to $\mathcal{Z}^*$. We have $\|z - \bar{z}\|$ is sufficiently small, where $\bar{z} = (0, \bar{x}, \bar{s}, \bar{y})$ is defined in (5.16). Hence, for $i = 1, \ldots, n$, denoting $u_i = (\mu, x_i, s_i)$, $\bar{u}_i = (0, \bar{x}_i, \bar{s}_i)$, we have

$$\left|\psi_{w_i}(u_i) - \psi_{w_i}(\bar{u}_i) - V(u_i - \bar{u}_i)\right| \leq L_{w_i}(u_i)\|u_i - \bar{u}_i\|^2,$$

where $V \in \partial\psi_{w_i}(u_i)$ and $L_{w_i}(u_i)$ is defined in (5.38). Hence, for case (i), when the weight vector $w > 0$, we have $L_{w_i}(u_i) = 9/(2\sqrt{w_i})$. For case (ii), when the solution set $\mathcal{S}^* = \{x^*, s^*, y^*\}$ is singleton, we have $\bar{u}_i := (0, x_i^*, s_i^*)$. So, if $\bar{u}_i = 0$, we have $L_{w_i}(u_i) = 0$; otherwise, if $\bar{u}_i \neq 0$, we have $L_{w_i}(u_i) = 9\sqrt{2}/\|u_i\| \leq 18\sqrt{2}/\|\bar{u}_i\|$ when u_i is sufficiently close to $\bar{u}_i$ such that $\|u_i\| \geq \|\bar{u}_i\|/2 > 0$. For case (iii), since $\mathcal{S}^*$ is a closed set, we have $\sqrt{x_i^2 + s_i^2} > \theta > 0$ for any $(x, s, y) \in \mathcal{S}^*$ and some $\theta > 0$. So, we have $L_{w_i}(u_i) = 9\sqrt{2}/\|u_i\| \leq 18\sqrt{2}/\theta$ when u_i is sufficiently close to $\bar{u}_i$ such that $\|u_i\| \geq \|\bar{u}_i\|/2 > \theta/2$. By the above discussions, based on our definition of strongly semismooth with respect to $\mathcal{Z}^*$ in Assumption 5.3, we can see the vector function $\mathcal{H}(\cdot)$ given in (5.37) is strongly semismooth with respect to the solution set $\mathcal{Z}^*$. □

6 Numerical Results

In this section, we report some numerical results of Algorithm 4.1. All experiments are carried on a PC with CPU of Inter(R) Core(TM)i7-7700 CPU @ 3.60 GHz and RAM of 8.00GB. The program codes are written in MATLAB and run in MATLAB R2018a environment. The parameters used in Algorithm 4.1 are chosen as $\mu_0 = 10^{-4}$, $\sigma = 0.2$, $\delta = 0.5$, $\gamma = 10^{-5}$. Moreover, unless particularly specified, we use $\|\mathcal{H}(z^k)\| \leq 10^{-6}$ as the stopping criterion.

6.1 WCP Over the Second-Order Cone

Optimization problems over the second-order cone have received considerable attention in recent years for its wide applications in many fields such as engineering, optimal control and design, machine learning, robust optimization and combinatorial optimization (e.g., [4,26]). Now, we consider the problem of finding a pair $(x, s, y) \in \mathcal{R}^n \times \mathcal{R}^n \times \mathcal{R}^m$ such that

$$x \in \mathcal{K}^n,\ s \in \mathcal{K}^n,\ F(x, s, y) = 0,\ x \circ s = w, \tag{6.1}$$

with

$$F(x, s, y) = \begin{pmatrix} \nabla f(x) - s + A^T y \\ Ax - b \end{pmatrix}, \tag{6.2}$$

where $\mathcal{K}^n$ is the n dimensional second-order cone defined by

$$\mathcal{K}^n := \{(x_1, x_{2:n}^T)^T \in \mathcal{R} \times \mathcal{R}^{n-1} \,|\, x_1 \geq \|x_{2:n}\|\},$$

"$\circ$" denotes the Jordan product associated with $\mathcal{K}^n$ (e.g., [4]), $A \in \mathcal{R}^{m\times n}$, $b \in \mathcal{R}^m$ and $f : \mathcal{R}^n \to \mathcal{R}$ is a twice continuously differentiable function. Notice that if $w = 0$, then the system (6.1)–(6.2) is the KKT conditions for the nonlinear second-order cone programming: $\{\min\ f(x),\ \text{ s.t. } Ax = b,\ x \in \mathcal{K}^n\}$.

We set up WCP (6.1)–(6.2) using the data that $w = (w_1, w_{2:n}^T)^T \in \mathcal{K}^n$ with $w_{2:n} = \textbf{rand}(n-1, 1)$ and $w_1 = \|w_{2:n}\| + \textbf{rand}(1, 1)$, $A = \textbf{randn}(m, n)$, $b = Au$ with $u \in \mathcal{K}^n$ being generated by the same way as w and $f(x) = f_i(x) (i = 1, 2, 3)$, respectively, where $f_i(x)$ is defined by

(i) Quadratic Function:

$$f_1(x) = \frac{1}{2} x^T Q x + c^T x,$$

where $Q = nBB^T/\|BB^T\|$ with $B = \textbf{rand}(n, n)$ and $c = \textbf{rand}(n, 1)$;

(ii) Extended Powell Function [11]:

$$f_2(x) = \sum_{i=1}^{n/4} [(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 \\ +10(x_{4i-3} - x_{4i})^4];$$

(iii) Oren Function [11]:

$$f_3(x) = \left[\sum_{i=1}^{n} i x_i^2\right]^2.$$

As one example of observing local convergence behavior of NSNA, we first generate one test problem for the quadratic function (i) with $n = 100$, $m = 50$ and solve this

Table 1 Value of $\|\mathcal{H}(z^k)\|$ at the kth iteration

	$\tau = 0, t = 1.5$	$\tau = 2,\ t = 2$
$k = 1$	12.8285	8.7837
$k = 2$	1.2962	1.6764
$k = 3$	0.1674	0.2244
$k = 4$	0.0061	0.0142
$k = 5$	8.5043e−06	1.0043e−04
$k = 6$	1.6948e−11	5.1236e−09
$k = 7$	4.9846e−14	6.3705e−14

problem by using $x^0 = s^0 = (1, 0, ..., 0)^T$ and $y^0 = (1, ..., 1)^T$ as the starting point. Table 1 gives the value of $\|\mathcal{H}(z^k)\|$ at the kth iteration, in which τ and t are parameter values used in the smoothing function ψ. We can clearly see the local fast, at least superlinear, convergence of NSNA.

Next, for each $f_i(x)(i = 1, 2, 3)$, we generate 100 instances with different sizes and test these problems by using the starting point: (1) $x^0 = s^0 = (1, 0, ..., 0)^T$, $y^0 = (1, ..., 1)^T$; (2) $x^0 = \mathbf{rand}(n, 1), s^0 = \mathbf{rand}(n, 1), y^0 = \mathbf{rand}(m, 1)$. In the experiments, we use the smoothing function ψ with $\tau = 4\mathbf{rand}(1, 1)$ and $t = 2$. Table 2 shows the numerical results of 100 trials for each case, where f denotes the test functions defined by (i)–(iii), SP denotes the starting point, n and m denote the problem size, **AIT** and **ACPU** denote the average number of iterations and the average CPU time in seconds, respectively, and **AHK** denotes the average value of $\|\mathcal{H}(z^k)\|$. From Table 2, we can see that NSNA is quite efficient and robust by using different starting points and algorithm parameters for solving WCP (6.1)–(6.2) over the second-order cone.

6.2 The LWCP

In this subsection, we consider to solve the quadratic programming and weighted centering problem [30] (denoted by QPWCP):

$$\begin{aligned} \min \quad & \varphi(x) := \frac{1}{2}x^T Mx + c^T x - \sum_{i=1}^{n} w_i \log x_i \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0, \end{aligned}$$

whose dual is

$$\begin{aligned} \max \quad & \phi(u, s, y) := -\frac{1}{2}u^T Mu + b^T y + \sum_{i=1}^{n} w_i \log s_i + \sum_{i=1}^{n} w_i (1 - \log w_i) \\ \text{s.t.} \quad & s = Mu - A^T y + c, \quad s \geq 0, \end{aligned}$$

where M is an $n \times n$ symmetric positive semidefinite matrix, $A \in \mathcal{R}^{m\times n}$ is a full row rank matrix with $m < n$ and $c \in \mathcal{R}^n, w \in \mathcal{R}^n_+, b \in \mathcal{R}^m$. For $M = 0$, the

Table 2 Numerical results of the WCP given in (6.1)–(6.2)

$f(x)$	SP	n	m	AIT	ACPU	AHK
Quadratic	(1)	1000	500	6.33	3.20	4.9948e−08
		1500	750	6.32	8.02	4.0829e−08
		2000	1000	6.33	16.02	4.0747e−08
	(2)	1000	500	6.51	3.13	3.7571e−08
		1500	750	6.63	8.69	4.0645e−08
		2000	1000	6.65	17.76	6.8175e−08
Extended Powell	(1)	100	100	38.26	0.19	3.7031e−08
		100	50	12.52	0.05	9.0767e−08
		100	20	9.97	0.04	6.8521e−08
	(2)	100	100	14.75	0.06	8.3200e−08
		100	50	14.01	0.04	8.2427e−08
		100	50	10.72	0.03	4.4850e−08
Oren	(1)	30	30	473.54	0.60	8.4200e−09
		30	20	254.56	0.31	5.2597e−09
		20	20	192.45	0.15	3.9084e−08
	(2)	30	30	7.19	0.01	1.0016e−07
		30	20	7.18	0.01	7.9289e−08
		20	20	7.02	0.01	9.7322e−08

QPWCP reduces to the notion of a linear programming and weighted centering problem introduced by Anstreicher [1]. By [30, Theorem 2.1], the optimality conditions of the QPWCP are equivalent to LWCP in (1.3) with

$$P = \begin{pmatrix} A \\ M \end{pmatrix}, \quad Q = \begin{pmatrix} 0 \\ -I \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ -A^T \end{pmatrix}, \quad a = \begin{pmatrix} b \\ -f \end{pmatrix}. \tag{6.3}$$

We apply our algorithm to solve this LWCP with problem data generated by the following way. We choose $A = \mathbf{randn}(m, n)$ with the rank of A being m and set $M = UU^T/\|UU^T\|$ with $U = \mathbf{rand}(n, n)$. Then, we choose $\hat{x} = \mathbf{rand}(n, 1)$, $f = \mathbf{rand}(n, 1)$ and set $b := A\hat{x}$, $\hat{s} := M\hat{x} + f$ and $w := \hat{x}\hat{s}$. For each problem with different sizes, we generate 100 instances and test them by using the starting points $x^0 = s^0 = (1, 0, ..., 0)^T$ and $y^0 = (0, ..., 0)^T$. In these experiments, we use the smoothing function ψ with different τ and t. Numerical results are listed in Table 3. From Table 3, it appears that the iteration numbers and the CPU time vary slightly for different t and the best numerical results occur in the case of $\tau = 0$. However, it is yet unknown how to find a better τ in general.

Now, we would like compare our algorithm for solving LWCP with the interior point methods proposed by Potra [30]. In this experiment, we choose $A = [I\ -B] \in \mathcal{R}^{m\times n}$ where $I = \mathbf{eye}(m)$ and $B = \mathbf{rand}(m, n - m)$, and choose $M = \mathbf{diag}(\mathbf{rand}(n, 1))$. Then, we choose $\hat{x} = \mathbf{rand}(n, 1)$, $f = \mathbf{rand}(n, 1)$ and set $b := A\hat{x}$, $\hat{s} := M\hat{x} + f$ and $w := \hat{x}\hat{s}$. To generate a strictly feasible starting point, we first generate a vector

Table 3 Numerical results of the LWCP given in (1.3) with (6.3)

τ	n	m	$t=1$		$t=1.5$		$t=2$	
			AIT	ACPU	AIT	ACPU	AIT	ACPU
0.0	1000	500	5.00	1.17	5.00	1.31	5.00	1.28
	1500	1000	5.51	4.19	5.92	5.03	5.52	4.92
	2000	1000	5.00	6.60	6.00	8.12	5.00	7.02
	2000	1500	5.97	9.55	6.10	10.13	5.97	9.98
2.0	1000	500	6.02	1.46	6.16	1.89	6.02	1.55
	1500	1000	6.83	5.21	6.98	6.12	6.81	5.62
	2000	1000	6.09	8.13	7.00	10.23	6.05	9.01
	2000	1500	7.02	12.15	7.03	13.52	7.02	12.93
3.5	1000	500	8.32	2.03	8.33	2.56	8.32	2.12
	1500	1000	8.64	6.60	8.64	6.78	8.64	6.81
	2000	1000	8.65	11.45	8.67	11.67	8.67	11.98
	2000	1500	8.87	14.51	8.87	14.55	8.87	14.76

$\tilde{x} = [x_I; x_B]$ with $x_B = \mathbf{rand}(n-m, 1)$, $x_I = Bx_B$. Then, we set $x^0 = \tilde{x}+\hat{x}$, $y^0 = 0$ and $s^0 = Mx^0 + f$. Since $A\tilde{x} = 0$, we have that (x^0, s^0, y^0) is a strictly feasible point for the LWCP. It is worth pointing out that this class of LWCP has been also tested by Zhang [16]. Let $gap = \|x^k s^k - w\|_\infty$, $res = \|Px^k + Qs^k + Ry^k - a\|_\infty$, $fea = \max\{\|x^k_-\|_\infty, \|s^k_-\|_\infty\}$ and $iter$ be the iteration number. In the experiments, we use $\max(gap,\ res,\ fea) < 10^{-9}$ and $iter > 20$ as the stopping criterion for NSNA and the interior point methods in [30]. For each problem with different sizes, we test 3 instances. For NSNA, we use the smoothing function ψ with $\tau = 0$ and $t = 1$. Numerical results are listed in Table 4, in which **A1** and **A2** denote NSNA using a general starting point $x^0 = s^0 = (1, 0, ..., 0)^T$, $y^0 = (0, ..., 0)^T$ and the strictly feasible starting point generated by the above way, respectively. **LS-IPM** and **PC-IPM** denote the largest step interior-point method and the predictor–corrector interior point method in [30] using the above strictly feasible starting point, respectively. **IT** and **CPU** denote iteration number and the CPU time in seconds, respectively, and $*$ stands for that the iteration number is greater than 20.

From Table 4, we can see that NSNA is very robust and effective compared with interior point methods for solving LWCP. This is probably due to the nonmonotone line search and non-requirement of keeping feasibility of the iterates. Moreover, we can clearly see from Table 4 that by starting with a non-interior point could not only simplify the application of the algorithm but also often significantly improve its performance.

7 Conclusion and Final Discussion

In this paper, we have introduced a one-parametric class of smoothing functions which include the weight vector w. These functions can be used to reformulate WCP in the

Table 4 Comparison of algorithms for the LWCP given in (1.3) with (6.3)

n	m	A1		A2		LS-IPM		PC-IPM	
		IT	CPU	IT	CPU	IT	CPU	IT	CPU
1000	800	7	2.16	9	2.81	14	3.96	14	7.92
		7	2.12	9	2.72	*	*	14	7.82
		6	1.82	10	3.42	15	4.31	14	7.83
1500	1000	7	5.24	9	6.52	*	*	15	20.17
		7	5.17	10	7.29	14	9.67	15	20.08
		7	5.12	10	7.32	16	10.93	15	20.20
2000	1800	7	12.85	10	18.34	15	25.34	14	46.07
		7	12.65	10	18.23	17	29.25	14	46.04
		8	14.57	11	20.31	14	24.36	14	47.37

general case as a nonlinear smooth equation. By the equivalent reformulation, we have proposed a nonmonotone smoothing Newton algorithm (NSNA) to solve WCP. We have showed that any accumulation point of the iteration sequence is a solution of WCP. Moreover, when the solution set of WCP is nonempty, under proper assumptions which are much weaker than Jacobian nonsingularity assumption, we have proved that the iteration sequence is bounded and it converges to one solution of WCP locally superlinearly or quadratically. Hence, compared with existing smoothing Newton-type algorithms, our algorithm has stronger convergence properties under weaker assumptions. To the best of our knowledge, our algorithm is the first effective algorithm to solve the general WCP. Due to the global nonmonotone line search strategy, non-interior requirement of the iterates and the fast local convergence, NSNA is very robust and efficient for solving WCP in our numerical experiments.

Finally, we would like to finish the paper with some brief discussion. It is well known that interior point methods (IPMs) have become attractive due to their well established complexity analysis. For instance, different IPMs have been investigated in [1,30,31] for solving LWCP with nice global complexity results. Whether these complexity results can be extended for general WCP is still not clear. Global complexities for smoothing Newton methods (SNMs) to solve WCP do not exist either. These complexity results have tight connections with the globalization strategies and the level of accuracy for solving the Newton system in the methods. To study the global complexity of NSNA as well as other SNMs and IPMs combined with nonmonotone line search will be an interesting topic for future research. In addition, some important optimization problems recently arising from data minting or machine learning could be also formulated as WCP, for example, the L1-SQSSVM problem proposed in [25]. However, such problems often involve large data. So, how to apply SNMs, which usually involves solving linear systems at each iteration, such as NSNA to solve these WCPs by exploiting the problem and data structure also deserves further investigation.

Appendix: The proof of the example given in remark (ii) for Assumption 5.1

Since M is positive definite, all of its diagonal elements are greater than zero. Without loss of generality, here we assume that $M - I_n$ is positive definite where I_n represents the $n \times n$ identity matrix. Consider the smoothing function

$$\psi_c(\mu, a, b) = a + b - \sqrt{a^2 + b^2 + (\tau - 2)ab + (4 - \tau)c + 4\mu^t},\ \forall\,(\mu, a, b) \in \mathcal{R}_+ \times \mathcal{R}^2,$$

where $t \in [1, 2]$ and $\tau \in [0, 4)$. By (3.6), we have

$$\psi_c(0, a, b) = 0 \Longleftrightarrow a \geq 0,\ b \geq 0,\ ab = c.$$

Then we can reformulate the problem (5.7) as the following nonlinear smooth equations:

$$\mathcal{H}(z) := \mathcal{H}(\mu, x, s) = \begin{pmatrix} \mu \\ s - Mx - a \\ \psi_{w_1}(\mu, x_1, s_1) \\ \vdots \\ \psi_{w_n}(\mu, x_n, s_n) \end{pmatrix} = 0$$

and apply Algorithm 4.1 to solve it. Let

$$g(\mu, a, b) := \sqrt{a^2 + b^2 + (\tau - 2)ab + (4 - \tau)c + 4\mu^t},\ \forall\,(\mu, a, b) \in \mathcal{R}_+ \times \mathcal{R}^2.$$

Then, $\mathcal{H}(z)$ is continuously differentiable at any $z \in \mathcal{R}_{++} \times \mathcal{R}^{2n}$ with its Jacobian

$$\mathcal{H}'(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -M & I \\ d_\mu & \mathbf{diag}(d_x) & \mathbf{diag}(d_s) \end{bmatrix},$$

where

$$\begin{aligned}
d_\mu &:= \left(-\frac{2t}{\mu^{1-t} g(\mu, x_1, s_1)}, \ldots, -\frac{2t}{\mu^{1-t} g(\mu, x_n, s_n)} \right)^T, \\
d_x &:= \left(1 - \frac{x_1 + (\tau/2 - 1)s_1}{g(\mu, x_1, s_1)}, \ldots, 1 - \frac{x_n + (\tau/2 - 1)s_n}{g(\mu, x_n, s_n)} \right)^T, \\
d_s &:= \left(1 - \frac{s_1 + (\tau/2 - 1)x_1}{g(\mu, x_1, s_1)}, \ldots, 1 - \frac{s_n + (\tau/2 - 1)x_n}{g(\mu, x_n, s_n)} \right)^T.
\end{aligned}$$

Since M is positive definite, $\mathcal{H}'(z)$ is nonsingular at any $z \in \mathcal{R}_{++} \times \mathcal{R}^{2n}$ and we can find its inverse

$$\mathcal{H}'(z)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & z_{33} \end{bmatrix},$$

where

$$\begin{aligned}
z_{21} &= -[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1} d_\mu, \\
z_{22} &= -[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1} \mathbf{diag}(d_s), \\
z_{23} &= [\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1}, \\
z_{31} &= -M[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1} d_\mu, \\
z_{32} &= I - M[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1} \mathbf{diag}(d_s), \\
z_{33} &= M[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1}.
\end{aligned}$$

In what follows, we divide the analysis into three parts.

Part 1. We show that $\mathbf{diag}(d_x)$ and $\mathbf{diag}(d_s)$ are positive semidefinite and bounded when $\mu \to 0^+$. This result holds by noticing that $g(\mu, a, b)$ can be written as

$$\begin{aligned}
g(\mu, a, b) &= \sqrt{[a + (\tau/2 - 1)b]^2 + \tau(1 - \tau/4)b^2 + (4 - \tau)c + 4\mu^t} \\
&= \sqrt{[b + (\tau/2 - 1)a]^2 + \tau(1 - \tau/4)a^2 + (4 - \tau)c + 4\mu^t},
\end{aligned}$$

and hence for any $(\mu, a, b) \in \mathcal{R}_+ \times \mathcal{R}^2$,

$$0 \le 1 - \frac{a + (\tau/2 - 1)b}{g(\mu, a, b)} \le 2, \ 0 \le 1 - \frac{b + (\tau/2 - 1)a}{g(\mu, a, b)} \le 2.$$

Part 2. We show that $\|[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1}\| \le \frac{\sqrt{n}\,\mathrm{cond}(M - I_n)}{2 - \sqrt{\tau}}$ when $\mu \to 0^+$, where $\mathrm{cond}(M - I_n)$ is the *conditional number* of $M - I_n$. First, we have

$$\mathbf{diag}(d_x) + \mathbf{diag}(d_s) = \mathbf{diag}(d_{xs}),$$

where

$$d_{xs} := \left(2 - \frac{\tau/2(x_1 + s_1)}{g(\mu, x_1, s_1)}, ..., 2 - \frac{\tau/2(x_n + s_n)}{g(\mu, x_n, s_n)}\right)^T.$$

Since $\left|\frac{\tau/2(a+b)}{g(\mu,a,b)}\right| \le \left|\frac{\tau/2(a+b)}{\sqrt{a^2+b^2+(\tau-2)ab}}\right| \le \sqrt{\tau}$ holds for any $\tau \in [0, 4)$ and $(\mu, a, b) \in \mathcal{R}_+ \times \mathcal{R}^2$, we have

$$0 < 2 - \sqrt{\tau} < 2 - \frac{\tau/2(x_i + s_i)}{g(\mu, x_i, s_i)} < 2 + \sqrt{\tau}, \ \ \forall\, i = 1, ..., n.$$

Hence, $\mathbf{diag}(d_{xs})$ is positive definite and so is $\mathbf{diag}(d_{xs})^{-1}$. Since $M - I_n$ is positive definite, it is invertible and $(M - I_n)^{-1}$ is also positive definite. So, we have

$$\begin{aligned}\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M &= \mathbf{diag}(d_{xs}) - \mathbf{diag}(d_s) + \mathbf{diag}(d_s)M \\ &= \mathbf{diag}(d_{xs}) + \mathbf{diag}(d_s)(M - I_n) \\ &= \mathbf{diag}(d_{xs})[(M - I_n)^{-1} + \mathbf{diag}(d_{xs})^{-1}\mathbf{diag}(d_s)](M - I_n).\end{aligned}$$

When $\mu \to 0^+$, since $\mathbf{diag}(d_s)$ is positive semidefinite by Part 1, $\mathbf{diag}(d_{xs})^{-1}\mathbf{diag}(d_s)$ is positive semidefinite. So, $(M - I_n)^{-1} + \mathbf{diag}(d_{xs})^{-1}\mathbf{diag}(d_s)$ is positive definite. Hence, when $\mu \to 0^+$, $\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M$ is nonsingular and

$$\begin{aligned}[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1} = (M - I_n)^{-1}[(M - I_n)^{-1} \\ + \mathbf{diag}(d_{xs})^{-1}\mathbf{diag}(d_s)]^{-1}\mathbf{diag}(d_{xs})^{-1},\end{aligned}$$

which implies that

$$\|[\mathbf{diag}(d_x) + \mathbf{diag}(d_s)M]^{-1}\| \le \frac{\sqrt{n}\mathrm{cond}(M - I_n)}{2 - \sqrt{\tau}},$$

where $\mathrm{cond}(M - I_n) := \|M - I_n\|\|(M - I_n)^{-1}\|$.

Part 3. We show that when $\mu \to 0^+$, if $w > 0$, then $\|d_\mu\| \le \varrho$ where $\varrho > 0$ is a constant, and if $w \ge 0$, then $\|d_\mu\| \le \frac{\sqrt{n}t}{\mu^{1-\frac{t}{2}}}$. This result holds since for any $(\mu, a, b) \in \mathcal{R}_+ \times \mathcal{R}^2$,

$$\begin{aligned}\frac{2t}{\mu^{1-t}g(\mu, a, b)} &= \frac{2t}{\mu^{1-t}\sqrt{[a + (\tau/2 - 1)b]^2 + \tau(1 - \tau/4)b^2 + (4 - \tau)c + 4\mu^t}} \\ &\le \begin{cases} \frac{2t\mu^{t-1}}{\sqrt{(4-\tau)c}}, & \text{if } c > 0, \\ \frac{t}{\mu^{1-\frac{t}{2}}}, & \text{if } c = 0. \end{cases}\end{aligned}$$

Therefore, when $\mu \to 0^+$, from Parts 1,2 and 3, we have that z_{22}, z_{23}, z_{32} and z_{33} are bounded, and z_{21}, z_{31} are bounded or

$$\|z_{21}\| \le \frac{n\mathrm{cond}(M - I_n)t}{(2 - \sqrt{\tau})\mu^{1-\frac{t}{2}}} \quad \text{and} \quad \|z_{31}\| \le \frac{n\|M\|\mathrm{cond}(M - I_n)t}{(2 - \sqrt{\tau})\mu^{1-\frac{t}{2}}}.$$

Hence, for any $t \in [1, 2]$, we can conclude that for any $z = (\mu, x, s) \in \mathcal{R}_{++} \times \mathcal{R}^{2n}$, when $\mu \to 0^+$, $\|\mathcal{H}'(z)^{-1}\|$ is bounded or there exists a constant $C_t > 0$ such that

$$\|\mathcal{H}'(z)^{-1}\| \le \frac{C_t}{\mu^{1-\frac{t}{2}}}.$$

Since $1 - \frac{t}{2} \in [0, 1/2]$, there exist constants $C > 0$ and $d \in [0, 1/2)$ such that Assumption 5.1 holds. This completes the proof. □

The proof of Lemma 4 For any $\xi = (\mu, a, b)^T, \xi' = (\mu', a', b')^T \in \mathcal{R}^3$, we have

$$\begin{aligned}
|\psi_c(\xi) - \psi_c(\xi')| &= |\psi_c(\mu, a, b) - \psi_c(\mu', a', b')| \\
&= |a + b - \|(a, b, \sqrt{2c}, 2\mu)^T\| - (a' + b') + \|(a', b', \sqrt{2c}, 2\mu')^T\|| \\
&\le |a - a'| + |b - b'| + |\|(a, b, \sqrt{2c}, 2\mu)^T\| - \|(a', b', \sqrt{2c}, 2\mu')^T\|| \\
&\le |a - a'| + |b - b'| + |\|(a, b, \sqrt{2c}, 2\mu)^T - (a', b', \sqrt{2c}, 2\mu')^T\|| \\
&= |a - a'| + |b - b'| + \|(a - a', b - b', 2(\mu - \mu'))^T\| \\
&\le \sqrt{2[(a - a')^2 + (b - b')^2]} + \sqrt{(a - a')^2 + (b - b')^2 + 4(\mu - \mu')^2} \\
&\le (\sqrt{2} + 2)\|\xi - \xi'\|.
\end{aligned}$$

This proves (i). Now we prove (ii) by considering the following three cases.

Case 1. If $c > 0$, then ψ_c is differentiable at any $\xi \in \mathcal{R}^3$ with

$$\begin{aligned}
\psi_c'(\xi) = \Bigg(&- \frac{4\mu}{\sqrt{a^2 + b^2 + 2c + 4\mu^2}}, 1 - \frac{a}{\sqrt{a^2 + b^2 + 2c + 4\mu^2}}, \\
&1 - \frac{b}{\sqrt{a^2 + b^2 + 2c + 4\mu^2}} \Bigg)^T
\end{aligned}$$

and

$$\psi_c''(\xi) = \frac{1}{(\sqrt{a^2 + b^2 + 2c + 4\mu^2})^3} \times M$$

where

$$M := \begin{bmatrix} -4(a^2 + b^2 + 2c) & 4a\mu & 4b\mu \\ 4a\mu & -(b^2 + 2c + 4\mu^2) & ab \\ 4b\mu & ab & -(a^2 + 2c + 4\mu^2) \end{bmatrix},$$

which yields

$$\begin{aligned}
\|\psi_c''(\xi)\| &\le \frac{\|M\|}{(\sqrt{a^2 + b^2 + 2c + 4\mu^2})^3} \\
&\le \frac{\sqrt{18(a^2 + b^2 + 2c + 4\mu^2)^2}}{(\sqrt{a^2 + b^2 + 2c + 4\mu^2})^3} \\
&= \frac{3\sqrt{2}}{\sqrt{a^2 + b^2 + 2c + 4\mu^2}} \\
&\le \frac{3}{\sqrt{c}}.
\end{aligned} \tag{7.1}$$

By (7.1), for any $u, v \in \mathcal{R}^3$ we have

$$\|\psi_c'(u) - \psi_c'(v)\| \leq \frac{3}{\sqrt{c}}\|u - v\|. \tag{7.2}$$

For any $\xi \in \mathcal{R}^3$ and $h \in \mathcal{R}^3$, ψ_c is differentiable at $\xi + h$ and hence $\partial\psi_c(\xi + h) = \{\psi_c'(\xi + h)^T\}$. So, from (7.2) we have for any $V \in \partial\psi_c(\xi + h)$ and $h \to 0$,

$$\begin{aligned}
|\psi_c(\xi + h) - \psi_c(\xi) - Vh| &= |\psi_c(\xi + h) - \psi_c(\xi) - \psi_c'(\xi + h)^T h| \\
&\leq |\psi_c(\xi + h) - \psi_c(\xi) - \psi_c'(\xi)^T h| + |[\psi_c'(\xi + h) - \psi_c'(\xi)]^T h| \\
&\leq |\psi_c(\xi + h) - \psi_c(\xi) - \psi_c'(\xi)^T h| + \frac{3}{\sqrt{c}}\|h\|^2 \\
&= \left| \int_0^1 [\psi_c'(\xi + th) - \psi_c'(\xi)]^T h dt \right| + \frac{3}{\sqrt{c}}\|h\|^2 \\
&\leq \frac{3}{\sqrt{c}}\|h\|^2 \int_0^1 t dt + \frac{3}{\sqrt{c}}\|h\|^2 \\
&= \frac{9}{2\sqrt{c}}\|h\|^2.
\end{aligned}$$

Case 2. If $c = 0$ and $\xi = 0$, then for any nonzero direction vector $h = (\tilde{\mu}, \tilde{a}, \tilde{b})^T \in \mathcal{R}^3$, ψ_c is smooth at the point $0 + h = h$. Thus, $V \in \partial\psi_c(0 + h) = \{\psi_c'(h)^T\}$ is uniquely given by

$$V = \left(-\frac{4\tilde{\mu}}{\sqrt{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2}}, 1 - \frac{\tilde{a}}{\sqrt{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2}}, 1 - \frac{\tilde{b}}{\sqrt{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2}} \right).$$

Then, for any $V \in \partial\psi_c(0 + h)$ and $h \to 0$,

$$\begin{aligned}
|\psi_c(0 + h) - \psi_c(0) - Vh| &= \left| -\sqrt{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2} + \frac{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2}{\sqrt{\tilde{a}^2 + \tilde{b}^2 + 4\tilde{\mu}^2}} \right| \\
&= 0.
\end{aligned}$$

Case 3. If $c = 0$ and $\xi \neq 0$, then ψ_c is differentiable at ξ and from (7.1) we have

$$\begin{aligned}
\|\psi_c''(\xi)\| &\leq \frac{3\sqrt{2}}{\sqrt{a^2 + b^2 + 4\mu^2}} \\
&\leq \frac{3\sqrt{2}}{\sqrt{a^2 + b^2 + \mu^2}} = \frac{3\sqrt{2}}{\|\xi\|}.
\end{aligned} \tag{7.3}$$

Next we show that $\psi_c'(\xi)$ is locally Lipschitz continuous at ξ with the constant $\frac{6\sqrt{2}}{\|\xi\|}$. In fact, let $u, v \in N(\xi, \frac{\|\xi\|}{2})$, we have

$$\psi_c'(u) = \psi_c'(v) + \int_0^1 \psi_c''(v + t(u - v))(u - v)dt. \tag{7.4}$$

By (7.3), we have

$$\|\psi_c''(v + t(u - v))\| \leq \frac{3\sqrt{2}}{\|v + t(u - v)\|}. \tag{7.5}$$

Since $u, v \in N(\xi, \frac{\|\xi\|}{2})$, we have

$$\|v + t(u - v)\| = \|tu + (1 - t)v\| \in N(\xi, \frac{\|\xi\|}{2}).$$

It follows that

$$\|v + t(u - v)\| \geq \|\xi\| - \|v + t(u - v) - \xi\| \geq \|\xi\| - \frac{\|\xi\|}{2} = \frac{\|\xi\|}{2}.$$

This together with (7.5) yields

$$\|\psi_c''(v + t(u - v))\| \leq \frac{6\sqrt{2}}{\|\xi\|}. \tag{7.6}$$

Thus, from (7.4) and (7.6), we have for any $u, v \in N(\xi, \frac{\|\xi\|}{2})$,

$$\|\psi_c'(u) - \psi_c'(v)\| \leq \frac{6\sqrt{2}}{\|\xi\|}\|u - v\|. \tag{7.7}$$

Now we show that $\psi_c(\xi)$ is strongly semismooth at $\xi \neq 0$. Since $\xi \neq 0$, when $h \to 0$, $\xi + h \neq 0$ and hence $\partial\psi_c(\xi + h) = \{\psi_c'(\xi + h)^T\}$. So, from (7.7), similarly as the proof of Case 1, we have for any $V \in \partial\psi_c(\xi + h)$ and $h \to 0$,

$$|\psi_c(\xi + h) - \psi_c(\xi) - Vh| \leq \frac{6\sqrt{2}}{\|\xi\|}\|h\|^2 \int_0^1 tdt + \frac{6\sqrt{2}}{\|\xi\|}\|h\|^2 = \frac{9\sqrt{2}}{\|\xi\|}\|h\|^2.$$

Thus, the proof is completed. □

The proof of Lemma 5 Let $\mathcal{H}(z)$ be defined by (5.37). Then, $\mathcal{H}(z)$ is continuously differentiable at any $z = (\mu, x, s, y) \in \mathcal{R}_{++} \times \mathcal{R}^{2n+m}$ and its Jacobian is

$$\mathcal{H}'(z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & P & Q & R \\ D_\mu & \mathbf{diag}(D_x) & \mathbf{diag}(D_s) & 0 \end{bmatrix},$$

where

$$D_\mu := \left(-\frac{4\mu}{\sqrt{x_1^2+s_1^2+2w_1+4\mu^2}}, ..., -\frac{4\mu}{\sqrt{x_n^2+s_n^2+2w_n+4\mu^2}}\right)^T,$$

$$D_x := \left(1-\frac{x_1}{\sqrt{x_1^2+s_1^2+2w_1+4\mu^2}}, ..., 1-\frac{x_n}{\sqrt{x_n^2+s_n^2+2w_n+4\mu^2}}\right)^T,$$

$$D_s := \left(1-\frac{s_1}{\sqrt{x_1^2+s_1^2+2w_1+4\mu^2}}, ..., 1-\frac{s_n}{\sqrt{x_n^2+s_n^2+2w_n+4\mu^2}}\right)^T.$$

We now divide the proof by the following three parts.

Part (i) We show that $\mathcal{H}(z)$ is Lipschitz continuous on $\mathcal{R}^{1+2n+m}$. In fact, since $(Px + Qs + Ry - a)' = [P, Q, R]$, we have that $Px + Qs + Ry - a$ is Lipschitz continuous on $\mathcal{R}^{2n+m}$. Moreover, from (i) of Lemma 4, ψ_c is Lipschitz continuous on $\mathcal{R}^3$. Hence, $\mathcal{H}(z)$ is Lipschitz continuous on $\mathcal{R}^{1+2n+m}$.

Part (ii) For any $\theta > 0$, we show that $\mathcal{H}'(z)$ is bounded and Lipschitz continuous on the set

$$\Omega := \{z = (\mu, x, s, y) \in \mathcal{R}_{++} \times \mathcal{R}^{2n+m} | \ \mu \geq \theta\}.$$

In fact, for any $i = 1, ..., n$, since

$$0 \leq \frac{4\mu}{\sqrt{x_i^2+s_i^2+2w_i+4\mu^2}} \leq 2,$$

$$0 \leq 1-\frac{x_i}{\sqrt{x_i^2+s_i^2+2w_i+4\mu^2}} \leq 2,$$

$$0 \leq 1-\frac{s_i}{\sqrt{x_i^2+s_i^2+2w_i+4\mu^2}} \leq 2,$$

D_μ, D_x and D_s are bounded and hence $\mathcal{H}'(z)$ is bounded. Let

$$\Gamma := \{\upsilon = (\mu, a, b) \in \mathcal{R}_{++} \times \mathcal{R}^2 | \ \mu \geq \theta\}.$$

For any $\upsilon \in \Gamma$, define

$$f_c(\upsilon) := 1 - \frac{a}{\sqrt{a^2+b^2+2c+4\mu^2}}.$$

Since $\mu \geq \theta > 0$, $f_c(\upsilon)$ is continuously differentiable and

$$f_c'(\upsilon) = \frac{1}{(\sqrt{a^2+b^2+2c+4\mu^2})^3}\Big[4a\mu, -(b^2+2c+4\mu^2), \ ab\Big],$$

which yields

$$\|f_c'(\upsilon)\| = \frac{\sqrt{16a^2\mu^2 + (b^2+2c+4\mu^2)^2 + a^2b^2}}{(\sqrt{a^2+b^2+2c+4\mu^2})^3}.$$

By noticing that

$$\begin{aligned} 16a^2\mu^2 &\le 2(a^2+4\mu^2)^2 \le 2(a^2+b^2+2c+4\mu^2)^2, \\ (b^2+2c+4\mu^2)^2 &\le (a^2+b^2+2c+4\mu^2)^2, \\ a^2b^2 \le (a^2+b^2)^2 &\le (a^2+b^2+2c+4\mu^2)^2, \end{aligned}$$

we have

$$\|f_c'(\upsilon)\| \le \frac{2}{\sqrt{a^2+b^2+2c+4\mu^2}} \le \frac{1}{\mu} \le \frac{1}{\theta}.$$

Thus, for any $\tilde{\upsilon}, \upsilon \in \Gamma$, we have

$$\|f_c(\tilde{\upsilon}) - f_c(\upsilon)\|| \le \frac{1}{\theta}\|\tilde{\upsilon} - \upsilon\|.$$

This implies that f_c is Lipschitz continuous on Γ and hence D_x is Lipschitz continuous on Ω. By a similar way, we can show that D_s and D_μ are also Lipschitz continuous on Ω and so is $\mathcal{H}'(z)$.

Part (iii) We show that $\mathcal{M}'(z)$ is Lipschitz continuous on the set Θ defined by (5.9). In fact, $\mathcal{M}(z)$ is continuous differentiable at any $z \in \mathcal{R}_{++} \times \mathcal{R}^{2n+m}$ and

$$\mathcal{M}'(z) = 2\mathcal{H}(z)^T\mathcal{H}'(z).$$

So, for any $\tilde{z}, z \in \Theta$, by Parts (i) and (ii) and $\|\mathcal{H}(\tilde{z})\| \le 2\|\mathcal{H}(z^0)\|$, there exists a constant $M > 0$ such that

$$\begin{aligned} \|\mathcal{M}'(\tilde{z}) - \mathcal{M}'(z)\| &= 2\|\mathcal{H}(\tilde{z})^T\mathcal{H}'(\tilde{z}) - \mathcal{H}(z)^T\mathcal{H}'(z)\| \\ &= 2\|\mathcal{H}(\tilde{z})^T[\mathcal{H}'(\tilde{z}) - \mathcal{H}'(z)] - [\mathcal{H}(z) - \mathcal{H}(\tilde{z})]^T\mathcal{H}'(z)\| \\ &\le 2[\|\mathcal{H}(\tilde{z})\|\|\mathcal{H}'(\tilde{z}) - \mathcal{H}'(z)\| + \|\mathcal{H}(z) - \mathcal{H}(\tilde{z})\|\|\mathcal{H}'(z)\|] \\ &= M\|z - \tilde{z}\|. \end{aligned}$$

This completes the proof. □

References

1. Asadi, S., Darvay, Z., Lesaja, G., Mahdavi-Amiri, N., Potra, F.A.: A full-Newton step interior-point method for monotone weighted linear complementarity problems. J. Optim. Theory Appl. (2020). https://doi.org/10.1007/s10957-020-01728-4

2. Anstreicher, K.M.: Interior-point algorithms for a generalization of linear programming and weighted centering. Optim. Methods Softw. **27**, 605–612 (2012)
3. Burke, J., Xu, S.: A non-interior predictor–corrector path following algorithm for the monotone linear complementarity problem. Math. Program. **87**, 113–130 (2000)
4. Chen, J.-S., Pan, S.H.: A survey on SOC complementarity functions and solution methods for SOCPs and SOCCPs. Pac. J. Optim. **8**, 33–74 (2012)
5. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
6. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs, Oxford University Press, New York (1994)
7. Fan, J.Y.: The modified Levenberg–Marquardt method for nonlinear equations with cubic convergence. Math. Comput. **81**, 447–466 (2012)
8. Fan, J.Y.: Accelerating the modified Levenberg–Marquardt method for nonlinear equations. Math. Comput. **83**, 1173–1187 (2014)
9. Fukushima, M., Luo, Z.Q., Tseng, P.: Smoothing functions for second-order cone complementarity problems. SIAM J. Optim. **12**, 436–460 (2002)
10. Gowda, M.S., Sznajder, R., Tao, J.: Some P-properties for linear transformations on Euclidean Jordan algebras. Linear Algebra Appl. **393**, 203–232 (2004)
11. Grippo, L., Lampariello, F., Lucidi, S.: A truncated Newton method with nonmonotone line search for unconstrained optimization. J. Optim. Theory Appl. **60**, 401–419 (1989)
12. Hager, W.W., Zhang, H.: A nonmonotone line search technique and its application to unconstrained optimization. SIAM J. Optim. **14**, 1043–1056 (2004)
13. Huang, Z.H., Hu, S.L., Han, J.Y.: Convergence of a smoothing algorithm for symmetric cone complementarity problems with a nonmonotone line search. Sci. China Ser. A: Math. **52**, 833–848 (2009)
14. Huang, Z.H., Ni, T.: Smoothing algorithms for complementarity problems over symmetric cones. Comput. Optim. Appl. **45**, 557–579 (2010)
15. Huang, Z.H., Liu, X.H.: Extension of smoothing Newton algorithms for solve linear programming over symmetric cones. J. Syst. Sci. Complex. **24**, 195–206 (2011)
16. Jian, Z.: A smoothing Newton algorithm for weighted linear complementarity problem. Optim. Lett. **10**, 499–509 (2016)
17. Kong, L., Sun, J., Xiu, N.: A regularized smoothing Newton method for symmetric cone complementarity problems. SIAM J. Optim. **19**, 1028–1047 (2008)
18. Kong, L.: Quadratic convergence of a smoothing Newton method for symmetric cone programming without strict complementarity. Positivity **16**, 297–319 (2012)
19. Li, Y.M., Wei, D.Y.: A generalized smoothing Newton method for the symmetric cone complementarity problem. Appl. Math. Comput. **264**, 335–345 (2015)
20. Liu, X.H., Huang, Z.H.: A smoothing Newton algorithm based on a one-parametric class of smoothing functions for linear programming over symmetric cones. Math. Math. Oper. Res. **70**, 385–404 (2009)
21. Liu, L.X., Liu, S.Y., Wu, Y.: A smoothing Newton method for symmetric cone complementarity problem. J. Appl. Math. Comput. **47**, 175–191 (2015)
22. Lu, N., Huang, Z.H.: A smoothing Newton algorithm for a class of non-monotonic symmetric cone linear complementarity problems. J. Optim. Theory Appl. **161**, 446–464 (2014)
23. Luo, Z.Q.: New error bounds and their applications to convergence analysis of iterative algorithms. Math. Program. Ser. B **88**, 341–355 (2000)
24. Ma, C.F.: A regularized smoothing Newton method for solving the symmetric cone complementarity problem. Math. Comput. Model. **54**, 2515–2527 (2011)
25. Mousavi, S., Gao, Z., Han, L., Lim, A.: Quadratic Surface Support Vector Machine with L1 Norm Regularization. arXiv preprint arXiv:1908.08616 (2019)
26. Narushima, Y., Sagara, N., Ogasawara, H.: A smoothing Newton method with Fischer–Burmeister function for second-order cone complementarity problems. J. Optim. Theory Appl. **149**, 79–101 (2011)
27. Ni, T., Gu, W.Z.: Smoothing Newton algorithm for symmetric cone complementarity problems based on a one-parametric class of smoothing functions. J. Appl. Math. Comput. **35**, 73–92 (2011)
28. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**, 299–332 (1997)
29. Pan, S., Chen, J.-S.: A one-parametric class of merit functions for the symmetric cone complementarity problem. J. Math. Anal. Appl. **355**, 195–215 (2009)
30. Potra, F.A.: Weighted complementarity problems—a new paradigm for computing equilibria. SIAM J. Optim. **22**, 1634–1654 (2012)

31. Potra, F.A.: Sufficient weighted complementarity problems. Comput. Optim. Appl. **64**(2), 467–488 (2016)
32. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)
33. Sun, D., Sun, J.: Löwner's operator and spectral functions in Euclidean Jordan algebras. Math. Oper. Res. **33**(2), 421–445 (2008)
34. Tang, J.Y.: A variant nonmonotone smoothing algorithm with improved numerical results for large-scale LWCPs. Comput. Appl. Math. **37**, 3927–3936 (2018)
35. Tang, J.Y., Dong, L., Fang, L., Zhou, J.C.: The convergence of a modified smoothing-type algorithm for the symmetric cone complementarity problem. J. Appl. Math. Comput. **43**, 307–328 (2013)
36. Tang, J.Y., Zhou, J.C., Fang, L.: Strong convergence properties of a modified nonmonotone smoothing algorithm for the SCCP. Optim. Lett. **12**, 411–424 (2018)
37. Yoshise, A.: Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. SIAM J. Optim. **17**, 1129–1153 (2006)
38. Zhang, J.L., Chen, J.: A smoothing Levenberg–Marquardt type method for LCP. J. Comput. Math. **22**, 735–752 (2004)