The properties of a new smoothing function and a modified smoothing Newton method for the P_0-NCP

RUIJUAN LIU
Nanjing Institute of Technology
Department of Mathematics and Physics
Nanjing, Jiangsu 210028
CHINA
ruijuanliu83@163.com

LI DONG
Xinyang Normal University
College of Mathematics and Information Science
Xinyang, Henan 464000
CHINA
dongli9261128@163.com

JINGYONG TANG
Xinyang Normal University
College of Mathematics and Information Science
Xinyang, Henan 464000
CHINA
jingyongtang@163.com

Abstract: In this paper we introduce a new smoothing function which has many nice properties. Based on this function, a smoothing Newton method is proposed to solve the nonlinear complementarity problem with P_0-function (denoted by P_0-NCP). Our method adopts a variant merit function. Moreover, we use a modified Newton equation to obtain the search direction. Under suitable assumptions, we show that the proposed method is globally and locally quadratically convergent. Some preliminary computational results are reported.

Key–Words: nonlinear complementarity problem, smoothing function, smoothing Newton method, global convergence, quadratic convergence

1 Introduction

We consider the nonlinear complementarity problem with P_0-function (denoted by P_0-NCP) which is to find a vector $(x, y) \in \mathcal{R}^n \times \mathcal{R}^n$ such that

$$(x, y) \geq 0, \;\; y = f(x), \;\; x^T y = 0, \tag{1}$$

where $f : \mathcal{R}^n \to \mathcal{R}^n$ is a continuously differentiable P_0-function. A function $f : \mathcal{R}^n \to \mathcal{R}^n$ is said to be a P_0-function if, for every u and v in $\mathcal{R}^n$ with $u \neq v$, there is an index i_0 such that

$$u_{i_0} \neq v_{i_0}, \; (u_{i_0} - v_{i_0})(f_{i_0}(u) - f_{i_0}(v)) \geq 0.$$

Nonlinear complementarity problems (NCPs) have many important applications in many fields (see, [4] and references therein). Numerous methods have been developed to solve NCPs, in which smoothing Newton methods have attracted a lot of attention partially due to their encouraging convergent properties and numerical results. Recently, Qi, Sun and Zhou [10] presented a class of new smoothing Newton methods for solving NCPs which has global and local quadratical convergence without strict complementarity. Due to its simplicity and weaker assumptions imposed on smoothing functions, the method [10] has been further studied by many authors for NCPs (see, e.g., [1, 2, 5, 6, 12, 13, 14, 15]). Lately, Zhang and Zhang [13] proposed a one-step smoothing Newton method to solve NCPs and proved that the algorithm is globally convergent. The method [13] adopts a merit function which is different with that in [10]. However, the local quadratic convergence of the method [13] is not reported.

Motivated by this direction, in this paper, we first introduce a new smoothing function which has many nice properties. Based on this function, we propose a smoothing Newton method for solving the P_0-NCP. Our method adopts a new merit function which is different with those in [10] and [13]. In addition, the proposed method uses a modified Newton equation to obtain the search direction which contains the usual Newton equation as a special case. Without requiring strict complementarity assumption, the proposed method is shown to be globally and locally quadratically convergent.

This paper is organized as follows. In the next section, we introduce a new smoothing function and study its properties. In Section 3, we reformulate the P_0-NCP concerned as a family of parameterized smooth equations and give its properties. In Section 4, we present a smoothing Newton method for solving the P_0-NCP. The global and local quadratic convergence of the proposed algorithm are investigated in Section 5. Numerical results are reported in Section 6. Some conclusions are given in Section 7.

Throughout this paper, $\mathcal{R}^n_+$ and $\mathcal{R}^n_{++}$ denotes the nonnegative and positive orthant in $\mathcal{R}^n$, respectively. For any vector $w \in \mathcal{R}^n$, we denote w by vec$\{w_i\}$ and the diagonal matrix whose ith diagonal element is w_i by diag$\{w_i\}$.

2 A new smoothing function

Smoothing functions play an important role in designing smoothing Newton methods for the NCP. In this paper, we propose and investigate a new smoothing functions $\phi : \mathcal{R}^3 \to \mathcal{R}$ defined by

$$\phi(\mu, a, b) = a + b - \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}, \quad (2)$$

where

$$A(\mu, a, b) := a \cos^2\mu + b \sin^2\mu,$$
$$B(\mu, a, b) := a \sin^2\mu + b \cos^2\mu.$$

It is worth pointing out that the smoothing function (2) has been discussed by Fang et al. [2] and Tang et al. [11] for second order cone optimization problems. Here, we will study its properties on $\mathcal{R}$.

Lemma 1. *Let ϕ be defined by (2). Then $\phi(\mu, a, b)$ is continuously differentiable at any $(\mu, a, b) \in \mathcal{R}_{++} \times \mathcal{R} \times \mathcal{R}$. Moreover, for any $\mu > 0$*

$$0 \le \phi'_a(\mu, a, b) \le 2, \quad 0 \le \phi'_b(\mu, a, b) \le 2.$$

Proof: It is easy to see that $\phi(\mu, a, b)$ is continuously differentiable at any $(\mu, a, b) \in \mathcal{R}_{++} \times \mathcal{R} \times \mathcal{R}$. From (2), for any $\mu > 0$, a straightforward calculation yields

$$\phi'_\mu = -\frac{-(a-b)^2 \sin 2\mu \cos 2\mu + 2\mu}{\sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}}, \quad (3)$$

$$\phi'_a = 1 - \frac{a - 2(a-b)\sin^2\mu\cos^2\mu}{\sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}}, \quad (4)$$

$$\phi'_b = 1 - \frac{b - 2(b-a)\sin^2\mu\cos^2\mu}{\sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}}. \quad (5)$$

Since

$$\begin{aligned} & a - 2(a-b)\sin^2\mu\cos^2\mu \\ &= (a\cos^2\mu + b\sin^2\mu)\cos^2\mu + (a\sin^2\mu + b\cos^2\mu)\sin^2\mu \\ &= A(\mu, a, b)\cos^2\mu + B(\mu, a, b)\sin^2\mu, \\ & b - 2(b-a)\sin^2\mu\cos^2\mu \\ &= (a\cos^2\mu + b\sin^2\mu)\sin^2\mu + (a\sin^2\mu + b\cos^2\mu)\cos^2\mu \\ &= A(\mu, a, b)\sin^2\mu + B(\mu, a, b)\cos^2\mu, \end{aligned}$$

we have, for any $\mu > 0$,

$$-1 \le \frac{a - 2(a-b)\sin^2\mu\cos^2\mu}{\sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}} \le 1,$$

$$-1 \le \frac{b - 2(b-a)\sin^2\mu\cos^2\mu}{\sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}} \le 1.$$

Therefore, we obtain the desired result. □

Lemma 2. *Let ϕ be defined by (2). One has*

$$\phi(\mu, a, b) = 0 \iff A(\mu, a, b) \ge 0, \ B(\mu, a, b) \ge 0, \ A(\mu, a, b)B(\mu, a, b) = \mu^2.$$

Proof: First, we assume that $\phi(\mu, a, b) = 0$. Since $A(\mu, a, b) + B(\mu, a, b) = a + b$, it follows from $\phi(\mu, a, b) = 0$ that

$$A(\mu, a, b) + B(\mu, a, b) = \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}. \quad (6)$$

Upon squaring both side of (6), we get

$$A(\mu, a, b)B(\mu, a, b) = \mu^2.$$

It follows that

$$A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2 = [A(\mu, a, b) + B(\mu, a, b)]^2. \quad (7)$$

Moreover, by (6) we have

$$A(\mu, a, b) + B(\mu, a, b) \ge 0.$$

Using this fact, we can conclude from (7) that

$$A(\mu, a, b) = \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2} - B(\mu, a, b) \ge 0,$$
$$B(\mu, a, b) = \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2} - A(\mu, a, b) \ge 0.$$

Now we assume that

$$A(\mu, a, b) \ge 0, \ B(\mu, a, b) \ge 0, \ A(\mu, a, b)B(\mu, a, b) = \mu^2.$$

By $A(\mu, a, b)B(\mu, a, b) = \mu^2$, we get

$$\begin{aligned} & [A(\mu, a, b) + B(\mu, a, b)]^2 \\ &= A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2A(\mu, a, b)B(\mu, a, b) \\ &= A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2. \end{aligned}$$

This, together with $A(\mu, a, b) \ge 0, B(\mu, a, b) \ge 0$, i.e., $A(\mu, a, b) + B(\mu, a, b) \ge 0$, implies that

$$A(\mu, a, b) + B(\mu, a, b) = \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2},$$

that is,

$$a + b = \sqrt{A(\mu, a, b)^2 + B(\mu, a, b)^2 + 2\mu^2}.$$

Hence, we have $\phi(\mu, a, b) = 0$. □

From Lemma 2, we can obtain that

$$\phi(0, a, b) = 0 \iff a \ge 0, \ b \ge 0, \ ab = 0, \quad (8)$$

which together with Lemma 1 shows that ϕ is a class of smoothing functions for the NCP.

Lemma 3. *Let ϕ be defined by (2). Suppose that $\{(\mu_k, a_k, b_k)\} \subset \mathcal{R}_{++} \times \mathcal{R} \times \mathcal{R}$ is a sequence such that $\{\mu_k\}$ is bounded. One has the following results.*
(i) *If either $A(\mu_k, a_k, b_k) \to -\infty$ or $B(\mu_k, a_k, b_k) \to -\infty$ as $k \to \infty$, then $\phi(\mu_k, a_k, b_k) \to -\infty$ as $k \to \infty$;*
(ii) *If $A(\mu_k, a_k, b_k) \to +\infty$ and $B(\mu_k, a_k, b_k) \to +\infty$, then $\phi(\mu_k, a_k, b_k) \to +\infty$ as $k \to \infty$.*

Proof: For simplicity, we define

$$U_k := A(\mu_k, a_k, b_k), \ \ V_k := B(\mu_k, a_k, b_k).$$

Then

$$\phi(\mu_k, a_k, b_k) = U_k + V_k - \sqrt{U_k^2 + V_k^2 + 2\mu_k^2}. \quad (9)$$

It is easy to see that $\phi(\mu_k, a_k, b_k) \to -\infty$, if either $U_k \to -\infty$ or $V_k \to -\infty$ as $k \to \infty$. Now, we suppose that $U_k \to +\infty$ and $V_k \to +\infty$ as $k \to \infty$. In this case, we know that $U_k V_k \geq \mu_k^2$ for sufficiently large k. So, from (9), we obtain that for sufficiently large k

$$\begin{aligned}
&\phi(\mu_k, a_k, b_k)\\
&= \frac{2U_kV_k - 2\mu_k^2}{U_k + V_k + \sqrt{U_k^2 + V_k^2 + 2\mu_k^2}}\\
&= \frac{2\max\{U_k, V_k\}\min\{U_k, V_k\} - 2\mu_k^2}{U_k + V_k + \sqrt{U_k^2 + V_k^2 + 2\mu_k^2}}\\
&\geq \frac{2\max\{U_k, V_k\}\min\{U_k, V_k\} - 2\mu_k^2}{2\max\{U_k, V_k\} + \sqrt{2(\max\{U_k, V_k\})^2 + 2\mu_k^2}}\\
&= \frac{2\min\{U_k, V_k\} - 2\mu_k^2/\max\{U_k, V_k\}}{2 + \sqrt{2 + 2\mu_k^2/(\max\{U_k, V_k\})^2}}.
\end{aligned}$$

Therefore, it follows from the boundedness of $\{\mu_k\}$ that $\phi(\mu_k, a_k, b_k) \to +\infty$ as $k \to \infty$. This completes the proof. □

3 The reformulation of the P_0-NCP

Let $z := (\mu, x, y) \in \mathcal{R}_+ \times \mathcal{R}^n \times \mathcal{R}^n$. We define the function $H : \mathcal{R}^{1+2n} \to \mathcal{R}^{1+2n}$ by

$$H(z) := \begin{pmatrix} \mu \\ \Gamma(z) \end{pmatrix}, \quad (10)$$

where

$$\Gamma(z) := \begin{pmatrix} f(x) - y + \mu x \\ \Phi(\mu, x, y) + \mu y \end{pmatrix} \quad (11)$$

with

$$\Phi(\mu, x, y) := \begin{pmatrix} \phi(\mu, x_1, y_1) \\ \vdots \\ \phi(\mu, x_n, y_n) \end{pmatrix}, \quad (12)$$

and $\phi(\cdot, \cdot, \cdot)$ is defined by (2). Then, it follows from (8) that

$$(x, y) \text{ is the solution of the } P_0\text{-NCP} \iff H(z) = 0.$$

Lemma 4. *Let $H(z)$ be defined by (10). Then the following results hold.*
(i) *$H(z)$ is continuously differentiable at any $z = (\mu, x, y) \in \mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^n$ with its Jacobian*

$$H'(z) = \begin{pmatrix} 1 & 0 & 0 \\ x & f'(x) + \mu I & -I \\ \Phi'_\mu + y & \Phi'_x & \Phi'_y + \mu I \end{pmatrix},$$

where

$$\Phi'_\mu = \text{vec}\left\{-\frac{-(x_i - y_i)^2 \text{sin}2\mu\text{cos}2\mu + 2\mu}{\sqrt{A(\mu, x_i, y_i)^2 + B(\mu, x_i, y_i)^2 + 2\mu^2}}\right\},$$

$$\Phi'_x = \text{diag}\left\{1 - \frac{x_i - 2(x_i - y_i)\text{sin}^2\mu\text{cos}^2\mu}{\sqrt{A(\mu, x_i, y_i)^2 + B(\mu, x_i, y_i)^2 + 2\mu^2}}\right\},$$

$$\Phi'_y = \text{diag}\left\{1 - \frac{y_i - 2(y_i - x_i)\text{sin}^2\mu\text{cos}^2\mu}{\sqrt{A(\mu, x_i, y_i)^2 + B(\mu, x_i, y_i)^2 + 2\mu^2}}\right\}.$$

(ii) *If f is a continuously differentiable P_0-function, then $H'(z)$ is invertible for any $\mu > 0$.*
(iii) *H is strongly semismooth on $\mathcal{R}^{1+2n}$ if f' is Lipschitz continuous on $\mathcal{R}^n$.*

Proof: By Lemma 1, $\Phi'_x(\mu, x, y)$ and $\Phi'_y(\mu, x, y)$ are positive diagonal matrices for any $(\mu, x, y) \in \mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^n$. Using this fact, we can prove the lemma similarly to Lemma 2.4 in [14]. We omit it here. □

Notice that we add the terms μx and μy into $H(z)$. By using such a technique, we can prove that $H(z)$ is coercive, which is a key to prove the global convergence of smoothing Newton methods.

Lemma 5. *Let $H(z)$ be defined by* (10). *Suppose that f is a continuously differentiable P_0-function, then $H(z)$ is coercive, i.e.,*

$$\lim_{k\to\infty} \|H(z^k)\| = +\infty$$

for any sequence $\{z^k = (\mu_k, x^k, y^k)\}$ such that $0 < \mu_k < \pi/2$ and $\lim_{k\to\infty} \|(x^k, y^k)\| = +\infty$.

Proof: Assume that the result of the lemma does not hold. Then, there exists a sequence $\{z^k = (\mu_k, x^k, y^k)\}$ such that

$$0 < \mu_k < \pi/2, \quad \lim_{k\to\infty} \|(x^k, y^k)\| = \infty, \quad \|H(z^k)\| \le \eta,$$

where $\eta > 0$ is certain constant. Since

$$\|H(z^k)\|^2 = \mu_k^2 + \|f(x^k) - y^k + \mu_k x^k\|^2 + \|\Phi(\mu_k, x^k, y^k) + \mu_k y^k\|^2, \quad (13)$$

it follows from the boundedness of $\{\|H(z^k)\|\}$ that $\{\|f(x^k)-y^k+\mu_k x^k\|\}$ and $\{\|\Phi(\mu_k, x^k, y^k)+\mu_k y^k\|\}$ are bounded. Denote

$$\Theta(\mu_k, x^k, y^k) := y^k - f(x^k) - \mu_k x^k,$$

then $\{\|\Theta(\mu_k, x^k, y^k)\|\}$ is bounded and

$$y^k = \Theta(\mu_k, x^k, y^k) + f(x^k) + \mu_k x^k. \quad (14)$$

First, we consider the case where $\lim_{k\to\infty} \|x^k\| = +\infty$. Since f is a continuously differentiable P_0-function, by using Lemma 1 in [3], there exists a subsequence, which we write without loss of generality as $\{x^k\}$, and an index i_0 such that, either $\lim_{k\to\infty} x^k_{i_0} = +\infty$ and $\{f_{i_0}(x^k)\}$ is bounded below, or $\lim_{k\to\infty} x^k_{i_0} = -\infty$ and $\{f_{i_0}(x^k)\}$ is bounded above.

- If $\lim_{k\to\infty} x^k_{i_0} = +\infty$ and $\{f_{i_0}(x^k)\}$ is bounded below, then by the boundedness of $\{\|\Theta(\mu_k, x^k, y^k)\|\}$ and $\mu_k > 0$, we obtain from (14) that

$$\lim_{k\to\infty} y^k_{i_0} = +\infty. \quad (15)$$

Furthermore, we get

$$\lim_{k\to\infty} \{x^k_{i_0}\cos^2\mu_k + y^k_{i_0}\sin^2\mu_k\} = +\infty,$$

$$\lim_{k\to\infty} \{x^k_{i_0}\sin^2\mu_k + y^k_{i_0}\cos^2\mu_k\} = +\infty,$$

that is

$$\lim_{k\to\infty} A(\mu_k, x^k_{i_0}, y^k_{i_0}) = +\infty,$$

$$\lim_{k\to\infty} B(\mu_k, x^k_{i_0}, y^k_{i_0}) = +\infty.$$

Then, by Lemma 3 (ii), we have

$$\lim_{k\to\infty} \phi(\mu_k, x^k_{i_0}, y^k_{i_0}) = +\infty,$$

which, together with (15), shows that

$$\lim_{k\to\infty} \{\phi(\mu_k, x^k_{i_0}, y^k_{i_0}) + \mu_k y^k_{i_0}\} = +\infty,$$

i.e., $\{\|\Phi(\mu_k, x^k, y^k)+\mu_k y^k\|\}$ is unbounded. By (13), $\{\|H(z^k)\|\}$ is unbounded, which derives a contradiction.

- If $\lim_{k\to\infty} x^k_{i_0} = -\infty$ and $\{f_{i_0}(x^k)\}$ is bounded above, then by the boundedness of $\{\|\Theta(\mu_k, x^k, y^k)\|\}$ and $\mu_k > 0$, we get from (14) that

$$\lim_{k\to\infty} y^k_{i_0} = -\infty. \quad (16)$$

Furthermore, we have

$$\lim_{k\to\infty} \{x^k_{i_0}\cos^2\mu_k + y^k_{i_0}\sin^2\mu_k\} = -\infty,$$

$$\lim_{k\to\infty} \{x^k_{i_0}\sin^2\mu_k + y^k_{i_0}\cos^2\mu_k\} = -\infty,$$

that is

$$\lim_{k\to\infty} A(\mu_k, x^k_{i_0}, y^k_{i_0}) = -\infty,$$

$$\lim_{k\to\infty} B(\mu_k, x^k_{i_0}, y^k_{i_0}) = -\infty.$$

By Lemma 3 (i), we can also obtain that

$$\lim_{k\to\infty} \phi(\mu_k, x^k_{i_0}, y^k_{i_0}) = -\infty,$$

and thus

$$\lim_{k\to\infty} \{\phi(\mu_k, x^k_{i_0}, y^k_{i_0}) + \mu_k y^k_{i_0}\} = -\infty,$$

which implies that $\{\|\Phi(\mu_k, x^k, y^k) + \mu_k y^k\|\}$ is unbounded. A contradiction is derived.

Second, we consider the case where $\{\|x^k\|\}$ is bounded for all $k \ge 0$. Since $\{\|(x^k, y^k)\|\}$ is unbounded, we obtain that $\{\|y^k\|\}$ is unbounded by the boundedness of $\{\|x^k\|\}$. Hence, there exists a subsequence, which we write as $\{y^k\}$, and an index i_0 such that $\lim_{k\to\infty} |y^k_{i_0}| = +\infty$. Notice that $\sin^2\mu_k > 0$ and $\cos^2\mu_k > 0$ since $0 < \mu_k < \pi/2$. Therefore, when $\lim_{k\to\infty} y^k_{i_0} = +\infty$, we have

$$\lim_{k\to\infty} \{x^k_{i_0}\cos^2\mu_k + y^k_{i_0}\sin^2\mu_k\} = +\infty,$$

$$\lim_{k\to\infty} \{x^k_{i_0}\sin^2\mu_k + y^k_{i_0}\cos^2\mu_k\} = +\infty,$$

and when $\lim_{k\to\infty} y^k_{i_0} = -\infty$, we have

$$\lim_{k\to\infty} \{x^k_{i_0}\cos^2\mu_k + y^k_{i_0}\sin^2\mu_k\} = -\infty,$$

$$\lim_{k\to\infty} \{x^k_{i_0}\sin^2\mu_k + y^k_{i_0}\cos^2\mu_k\} = -\infty.$$

Similarly to the first case, we can obtain the desired contradiction. So, we complete the proof. □

4 A smoothing Newton method

By using the function $\|H(z)\|$ defined by (10), the existing smoothing Newton methods (e.g., [1, 2, 5, 6, 10, 14, 15]) usually defined the merit function as

$$G_1(z) := \|H(z)\|^2 \text{ or } G_2(z) := \|H(z)\|.$$

Zhang and Zhang [13] presented a different merit function as

$$G_3(z) := \mu + \|\Gamma(z)\|^2.$$

Based on this merit function, Zhang and Zhang [13] proposed a smoothing Newton method for solving the NCP and proved that the method is globally convergent. However, the local quadratic convergence of the method [13] is not established.

In this paper, for any $z := (\mu, x, y) \in \mathcal{R}_+ \times \mathcal{R}^{2n}$, we denote the merit function $G(z) : \mathcal{R}_+ \times \mathcal{R}^{2n} \to \mathcal{R}_+$ by

$$G(z) = \mu + \|\Gamma(z)\|. \tag{17}$$

Notice that $H(z) = 0$ if and only if $G(z) = 0$. By using this merit function, we propose a smoothing Newton method for the P_0-NCP and prove that our method is globally and locally quadratically convergent.

Algorithm 6. (A smoothing Newton method)

Step 0. *Choose constants $\delta, \sigma \in (0, 1)$ and $0 < \mu_0 < \frac{\pi}{2}$. Choose constants $\gamma \in (0, 1)$ and $\tau \in [0, 1)$ such that $\gamma < \mu_0$ and $\gamma + \tau < 1$. Choose $(x^0, y^0) \in \mathcal{R}^{2n}$ be an arbitrary initial point and let $z^0 := (\mu_0, x^0, y^0)$. Set $k := 0$.*

Step 1. *If $\|H(z^k)\| = 0$, then stop.*

Step 2. *Compute $\Delta z^k := (\Delta\mu_k, \Delta x^k, \Delta y^k) \in \mathcal{R}^{1+2n}$ by*

$$H(z^k) + H'(z^k)\Delta z^k = \Upsilon_k, \tag{18}$$

where

$$\Upsilon_k := \begin{pmatrix} \beta_k \\ \Lambda_k \end{pmatrix} \tag{19}$$

with

$$\beta_k := \gamma \min\{1, G(z^k)^2\}, \tag{20}$$

$$\Lambda_k := \frac{\tau\|H(z^k)\|}{1 + G(z^k)^2}\Gamma(z^k). \tag{21}$$

Step 3. *Let l_k be the smallest nonnegative integer l such that*

$$G(z^k + \delta^l \Delta z^k) \le [1 - \sigma(1 - t)\delta^l]G(z^k), \tag{22}$$

in which $t := \gamma + \tau$. Let $\alpha_k := \delta^{l_k}$.

Step 4. *Set $z^{k+1} := z^k + \alpha_k \Delta z^k$ and $k := k + 1$. Go to Step 1.*

Remark 7. *Similar algorithmic framework was first introduced in [10], and was extensively discussed by many authors for solving the NCP (see, e.g., [1, 5, 6, 12, 13, 14, 15]). Notice that Algorithm 6 uses a modified Newton equation (18) to obtain the direction. If we choose $\tau = 0$, then (18) becomes the usual Newton equation*

$$H(z^k) + H'(z^k)\Delta z^k = \begin{pmatrix} \beta_k \\ 0 \end{pmatrix}$$

which has been extensively used in smoothing Newton methods (see, e.g., [1, 5, 6, 10, 12, 13, 14, 15]).

Lemma 8. *Let Λ_k be defined by (21). If $\mu_k \ge 0$, then*

$$\|\Lambda_k\| \le \tau \min\{1, G(z^k)^2\}. \tag{23}$$

Proof: Since

$$\begin{aligned} \|H(z^k)\|^2 &= \mu_k^2 + \|\Gamma(z^k)\|^2 \\ &\le (\mu_k + \|\Gamma(z^k)\|)^2 \\ &= G(z^k)^2, \end{aligned}$$

we have $\|H(z^k)\| \le G(z^k)$. Notice that $\|\Gamma(z^k)\| \le G(z^k)$ by (17). So, we have

$$\|\Lambda_k\| = \frac{\tau\|H(z^k)\|}{1 + G(z^k)^2}\|\Gamma(z^k)\| \le \frac{\tau G(z^k)^2}{1 + G(z^k)^2},$$

which implies that $\|\Lambda_k\| \le \tau$ and $\|\Lambda_k\| \le \tau G(z^k)^2$ for all $k \ge 0$. This proves the lemma. □

Lemma 9. *Let β_k, Λ_k be defined by (20) and (21), respectively. Then, for all $k \ge 0$, one has*

$$\beta_k + \|\Lambda_k\| \le (\gamma + \tau)G(z^k). \tag{24}$$

Proof: Since $\min\{1, \xi^2\} \le \xi$ for any $\xi \ge 0$, it follows from (20) and (23) that

$$\beta_k = \gamma \min\{1, G(z^k)^2\} \le \gamma G(z^k),$$

$$\|\Lambda_k\| \le \tau \min\{1, G(z^k)^2\} \le \tau G(z^k).$$

Using these facts, we can obtain the result. □

Theorem 10. *Suppose that f is a continuously differentiable P_0-function and that $z^0 = (\mu_0, x^0, y^0) \in \mathcal{R}_{++} \times \mathcal{R}^n \times \mathcal{R}^n$. Then Algorithm 6 is well-defined and generates an infinite sequence $\{z^k := (\mu_k, x^k, y^k)\}$ with $\mu_k > 0$ for all $k \ge 0$.*

Proof: Since $\mu_0 > 0$ by the choice of initial point, we may assume without loss of generality that $\mu_k > 0$ for some k. Since $f(x)$ is a continuously differentiable P_0-function, Lemma 4 implies that the matrix $H'(\tilde{z^k})$ exists and it is invertible. Hence, Step 2 of Algorithm 6 is well-defined at the kth iteration. By (18) we have $\Delta\mu_k = -\mu_k + \beta_k$. Since $\mu_k > 0$, we have $G(z^k) > 0$ and hence $\beta_k > 0$. Then, for any $\alpha \in (0, 1]$ we have

$$\mu_k + \alpha\Delta\mu_k = (1-\alpha)\mu_k + \alpha\beta_k > 0. \tag{25}$$

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

$$F_k(\alpha) := \Gamma(z^k + \alpha\Delta z^k) - \Gamma(z^k) - \alpha\Gamma'(z^k)\Delta z^k,$$

then $F_k(\alpha) = o(\alpha)$ since Γ is continuously differentiable for any $z^k \in \mathcal{R}_{++} \times \mathcal{R}^{2n}$. Hence, we can obtain that for any $\alpha \in (0, 1]$

$$\begin{aligned}
&\|\Gamma(z^k + \alpha\Delta z^k)\| \\
&= \|\Gamma(z^k) + \alpha\Gamma'(z^k)\Delta z^k + F_k(\alpha)\| \\
&= \|(1-\alpha)\Gamma(z^k) + \alpha\Lambda_k + o(\alpha)\| \\
&\le (1-\alpha)\|\Gamma(z^k)\| + \alpha\|\Lambda_k\| + o(\alpha), \qquad (26)
\end{aligned}$$

where the second equality follows from the fact $\Gamma'(z^k)\Delta z^k = -\Gamma(z^k) + \Lambda_k$ by (18). Hence, from (17), also using (24)–(26) we have

$$\begin{aligned}
&G(z^k+\alpha\Delta z^k) \\
&= \mu_k+\alpha\Delta\mu_k+\|\Gamma(z^k+\alpha\Delta z^k)\| \\
&= (1-\alpha)\mu_k+\alpha\beta_k+(1-\alpha)\|\Gamma(z^k)\|+\alpha\|\Lambda_k\|+o(\alpha) \\
&= (1-\alpha)G(z^k)+\alpha(\beta_k+\|\Lambda_k\|)+o(\alpha) \\
&\le (1-\alpha)G(z^k)+\alpha(\gamma+\tau)G(z^k)+o(\alpha) \\
&\le [1-(1-t)\alpha]G(z^k) + o(\alpha).
\end{aligned}$$

Since $t = \gamma+\tau < 1$, there exists a constant $\bar{\alpha} \in (0, 1)$ such that

$$G(z^k + \alpha\Delta z^k) \quad\le\quad [1-\sigma(1-t)\alpha]G(z^k)$$

holds for any $\alpha \in (0, \bar{\alpha}]$ and $\sigma \in (0, 1)$. This demonstrates that Step 3 of Algorithm 6 is well-defined at the kth iteration. So, z^{k+1} can be generated by Algorithm 6. Moreover, by Step 3 of Algorithm 6, we have $\alpha_k \in (0, 1]$. This, together with (25), shows that

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

Hence, from $\mu_0 > 0$ and the above statements, we prove that Algorithm 6 is well-defined and generates an infinite sequence $\{z^k\}$ with $\mu_k > 0$. This completes the proof. □

5 Convergence analysis

In this section, we analyze the convergence properties of Algorithm 6. To show its global convergence, we need the following result.

Lemma 11. *The sequence $\{z^k = (\mu_k, x^k, y^k)\}$ generated by Algorithm 6 has the following properties.*
(i) *$\{\beta_k\}$ is monotonically decreasing*;
(ii) *$\mu_k \ge \beta_k$ for all $k \ge 0$*;
(iii) *$\{\mu_k\}$ is monotonically decreasing.*

Proof: From Steps 3 and 4 of Algorithm 6, we know that $\{G(z^k)\}$ is monotonically decreasing. Thus, by the definition of β_k in (20), we have $\beta_k \ge \beta_{k+1}$ for all $k \ge 0$. Notice that $\mu_0 \ge \gamma \ge \beta_0$ by Step 0 of Algorithm 6. Suppose that $\mu_k \ge \beta_k$ for some k, then by (25) we have

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

which, together with (i), implies that $\mu_{k+1} \ge \beta_{k+1}$. So, by mathematical induction, we have $\mu_k \ge \beta_k$ for all $k \ge 0$. Using this result we can further obtain that

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

This completes our proof. □

Theorem 12. *Suppose that f is a continuously differentiable P_0-function and that $\{z^k\}$ is the iteration sequence generated by Algorithm 6. Then*
(i) *$\{G(z^k)\}$ converges to zero as $k \to \infty$, and hence any accumulation point z^* of $\{z^k\}$ is a solution of the P_0-NCP.*
(ii) *If the solution set of the P_0-NCP is nonempty and bounded, then $\{z^k\}$ is bounded and hence it has at least one accumulation point.*

Proof: According to Steps 3 and 4 of Algorithm 6, we know that $\{G(z^k)\}$ is monotonically decreasing and bounded from below by zero. Thus, there exists $G^* \ge 0$ and $\beta^* \ge 0$ such that

$$\lim_{k\to\infty} G(z^k) = G^*, \lim_{k\to\infty} \beta_k = \beta^* := \gamma\min\{1, (G^*)^2\}.$$

If $G^* = 0$, then we obtain the desired result. Suppose that $G^* > 0$, then we have $\beta^* > 0$. By Lemma 11, we obtain that

$$0 < \beta^* \le \beta_k \le \mu_k \le \mu_0 < \frac{\pi}{2}.$$

Also notice that

$$0 \le \|H(z^k)\| \le G(z^k) \le G(z^0).$$

Therefore, from Lemma 5 we know that $\{z^k = (\mu_k, x^k, y^k)\}$ is bounded and hence it has at least one accumulation point $z^* := (\mu^*, x^*, y^*)$. Without loss of generality, we assume that $\{z^k\}$ converges to z^* as $k \to +\infty$. Since $\mu_k \geq \beta_k$, we have $\mu^* \geq \beta^* > 0$. Then, from Lemma 4 we obtain that $H'(z^*)$ exists and is invertible. Hence, there exists a closed neighborhood $N(z^*)$ of z^* such that for any $z \in N(z^*)$ we have $\mu \in \mathcal{R}_{++}$ and $H'(z)$ is invertible. Then, for all sufficiently large k, we have $z^k \in N(z^*)$ and hence $\mu_k \in \mathcal{R}_{++}$ and $H'(z^k)$ is invertible. For all sufficiently large k, let $\Delta z^k \in \mathcal{R}^{1+2n}$ be the unique solution of the system of equations

$$H'(z^k)\Delta z^k = -H(z^k) + \Upsilon_k.$$

Similarly to the proof of Theorem 10, for all sufficiently large k, there exists a nonnegative integer $\bar{l}$ such that $\delta^{\bar{l}} \in (0, 1]$ and

$$G(z^k + \delta^{\bar{l}}\Delta z^k) \leq [1 - \sigma(1 - t)\delta^{\bar{l}}]G(z^k).$$

For all sufficiently large k, since $\alpha_k = \delta^{l_k} \geq \delta^{\bar{l}}$, it follows from Steps 3 and 4 in Algorithm 6 that

$$\begin{aligned} G(z^{k+1}) &\leq [1 - \sigma(1 - t)\delta^{l_k}]G(z^k) \\ &\leq [1 - \sigma(1 - t)\delta^{\bar{l}}]G(z^k). \end{aligned}$$

Taking limits on both sides of the above inequality, we have

$$G^* \leq [1 - \sigma(1 - t)\delta^{\bar{l}}]G^*,$$

which, together with $G^* > 0$, implies that $\sigma(1 - t)\delta^{\bar{l}} \leq 0$, i.e., $t = \gamma + \tau \geq 1$. This contradicts the fact that $\gamma + \tau \leq 1$ in Step 0 of Algorithm 6. Hence, we have $G^* = 0$, that is, $\lim\limits_{k\to\infty} G(z^k) = 0$. This proves the first part of (i). Next, we prove the second part. Assume that z^* is an accumulation point of $\{z^k\}$. Without loss of generality, we assume that $\lim_{k\to\infty} z^k = z^*$. Then, it follows from the continuity of G that $\lim\limits_{k\to\infty} G(z^k) = G(z^*) = 0$. Thus, z^* is a solution of the P_0-NCP.

Now, we prove (ii). Since $\{\mu_k\}$ is monotonically decreasing by Lemma 11, we have

$$0 \leq \mu_k \leq \mu_0 < \pi/2$$

holds for all $k \geq 0$. We only need to prove that $\{(x^k, y^k)\}$ is bounded. By the result (i), it is easy to see that $\lim\limits_{k\to\infty} \|H(z^k)\| = 0$, which gives

$$\lim_{k\to\infty} \mu_k = 0, \quad \lim_{k\to\infty} \|\Gamma(\mu_k, x^k, y^k)\| = 0.$$

From this fact, also using the famous mountain pass theorem (see, Theorem 9.2.7 in [7]), we can prove that $\{(x^k, y^k)\}$ is bounded similarly as Theorem 3.1 in [5]. Therefore, $\{z^k\}$ is bounded and hence it has at least one accumulation point. □

Now, we discuss the convergence rate of Algorithm 6. For this purpose, we need the following result.

Lemma 13. *Let Υ_k be defined by (19). Then, for all sufficiently large k,*

$$\|\Upsilon_k\| = O(\|H(z^k)\|^2).$$

Proof: From the result (i) in Theorem 12, we have $\lim\limits_{k\to\infty} G(z^k) = 0$. So, according to (20) and (23), we have for all sufficiently large k,

$$\beta_k = \gamma G(z^k)^2, \quad \|\Lambda_k\| \leq \tau G(z^k)^2. \tag{27}$$

Moreover, by the definition of H, we can obtain that $\mu_k \leq \|H(z^k)\|$ and $\|\Gamma(z^k)\| \leq \|H(z^k)\|$. Hence, we get

$$\begin{aligned} G(z^k)^2 &= (\mu_k + \|\Gamma(z^k)\|)^2 \\ &= \mu_k^2 + 2\mu_k\|\Gamma(z^k)\| + \|\Gamma(z^k)\|^2 \\ &\leq 4\|H(z^k)\|^2. \end{aligned}$$

It follows from (19) and (27) that for all sufficiently large k,

$$\|\Upsilon_k\| \leq \beta_k + \|\Lambda_k\| \leq (\gamma + \tau)G(z^k)^2 = O(\|H(z^k)\|^2).$$

This proves the lemma. □

Theorem 14. *Suppose that f is a continuously differentiable P_0-function and that z^* is an accumulation point of the iteration sequence $\{z^k\}$ generated by Algorithm 6. If all $V \in \partial H(z^*)$ are nonsingular and f' is locally Lipschitz continuous around x^*, then $\{z^k\}$ converges to z^* quadratically with $\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2)$.*

Proof: First, from Theorem 12, we have $H(z^*) = 0$. Since all $V \in \partial H(z^*)$ are nonsingular, it follows from Proposition 3.1 in [9] that

$$\|H'(z^k)^{-1}\| = O(1) \tag{28}$$

holds for all z^k sufficiently close to z^*. Second, under the assumption that $f'(x)$ is locally Lipschitz continuous around x^*, it follows from Lemma 4 that H is strongly semismooth at z^*. Hence, for all z^k sufficiently close to z^*,

$$\begin{aligned} &\|H(z^k) - H(z^*) - H'(z^k)(z^k - z^*)\| \\ &= O(\|z^k - z^*\|^2). \end{aligned} \tag{29}$$

Third, since H is strongly semismooth at z^*, H is locally Lipschitz continuous near z^*, So,

$$\|H(z^k)\| = ||H(z^k) - H(z^*)\| = O(\|z^k - z^*\|) \quad (30)$$

holds for all z^k sufficiently close to z^*. This, together with Lemma 13, gives that for all z^k sufficiently close to z^*

$$\|\Upsilon_k\| = O(\|z^k - z^*\|^2). \quad (31)$$

Therefore, using (28), (29) and (31), we get from (18) that for all z^k sufficiently close to z^*,

$$\begin{aligned}
&\|z_k + \Delta z^k - z^*\| \\
&= \|z_k + H'(z^k)^{-1}[-H(z^k) + \Upsilon_k] - z^*\| \\
&\leq \|H'(z^k)^{-1}\| \Big[\|H(z^k) - H(z^*) \\
&\qquad -H'(z^k)(z^k - z^*)\| + \|\Upsilon_k\|\Big] \\
&= O(\|z^k - z^*\|^2). \quad (32)
\end{aligned}$$

By a similar way as the proof of Theorem 3.1 in [8], we have

$$\|z^k - z^*\| = O(\|H(z^k)\|) \quad (33)$$

for all z^k sufficiently close to z^*. Hence, it follows from (30), (32) and (33) that for all z^k sufficiently close to z^*,

$$\begin{aligned}
\|H(z^k + \Delta z^k)\| &= O(\|z^k + \Delta z^k - z^*\|) \\
&= O(\|z^k - z^*\|^2) \\
&= O(\|H(z^k)\|^2). \quad (34)
\end{aligned}$$

Notice that $\sqrt{a^2 + b^2} \leq a + b \leq \sqrt{2(a^2 + b^2)}$ holds for any $a, b > 0$. Using this result, we can obtain from (10) and (17) that

$$\|H(z^k)\| \leq G(z^k) \leq \sqrt{2}\|H(z^k)\|. \quad (35)$$

So, it follows from (34) and (35) that

$$G(z^k + \Delta z^k) = O(G(z^k)^2). \quad (36)$$

From Theorem 12, we know that $G(z^k) \to 0$ as $k \to \infty$. Thus, (22) and (36) imply that $\alpha_k = 1$ for all z^k sufficiently close to z^*. This, together with (32) , indicates that for all z^k sufficiently close to z^*,

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2).$$

This completes the proof. □

6 Numerical results

In this section, we implement Algorithm 6 for solving some NCPs. All experiments are done using a PC with CPU of 2.5 GHz and RAM of 512 MB, and all codes are written in MATLAB.

We test some NCPs with $f(x) = P(x) + Mx + q$, where $P(x)$ and $Mx + q$ are the nonlinear and linear parts of $f(x)$, respectively. We form the matrix M as $M = A^T A + B$, where

$$A = 40\textbf{rand}(\text{n,n}) - 20\textbf{ones}(\text{n,n}),$$

$$B = 40\textbf{rand}(\text{n,n}) - 20\textbf{ones}(\text{n,n}),$$

and the vector q as

$$q = 10\textbf{rand}(\text{n,1}) - 5\textbf{ones}(\text{n,1}).$$

The components of $P(x)$, the nonlinear part of $f(x)$, are $P_j(x) = p_j \cdot \arctan(x_j)$, where p_j is a random variable in $(0, 4)$. For the initial point, we choose

$$x^0 = \textbf{rand}(\text{n,1}) \quad \text{and} \quad y^0 = \textbf{ones}(\text{n,1}).$$

Throughout the experiments, the parameters used in Algorithm 6 are $\mu_0 = 10^{-3}, \gamma = 5 \times 10^{-4}, \tau = 10^{-3}, \sigma = 0.2, \delta = 0.8$. We use $\|H(z^k)\| \leq 10^{-6}$ as the stopping criterion.

In the experiments, we generate 3 problem instances for each size of n. The test results are listed in Table 1, in which **IT** denotes the iteration numbers; **CPU** denotes the CPU time in seconds; **HK** denotes the value of $\|H(z^k)\|$ when the algorithm terminates.

Table 1 Test results of Algorithm 6

n	IT	CPU	HK
50	19	0.09	2.9638×10^{-8}
	33	0.08	2.4386×10^{-7}
	29	0.06	7.9090×10^{-8}
100	46	0.54	4.6097×10^{-9}
	39	0.42	1.6051×10^{-10}
	39	0.41	3.1825×10^{-10}
150	66	2.13	6.3729×10^{-11}
	44	1,38	1.6024×10^{-9}
	41	1.27	4.6621×10^{-7}
200	69	4.89	6.8591×10^{-7}
	62	4.33	1.0608×10^{-9}
	40	2.78	6.1262×10^{-8}
250	82	11.02	7.2019×10^{-7}
	69	9.27	1.3076×10^{-9}
	89	11.91	5.2743×10^{-9}
300	101	23.61	1.8499×10^{-9}
	74	17.27	2.8132×10^{-7}
	65	15.13	7.6130×10^{-7}
400	117	61.94	1.4907×10^{-7}
	84	44.42	9.6036×10^{-8}
	95	50.24	7.7582×10^{-7}

7 Conclusions

Based on a new smoothing function, we propose a modified smoothing Newton algorithm for solving the P_0-NCP. The proposed algorithm can start from an arbitrary point and it solves only one linear system of equations and performs only one line search per iteration. In addition, it adopts a new merit function. We show that the proposed algorithm is globally and locally quadratically convergent under suitable assumptions. Numerical results demonstrate that our algorithm is promising.

Acknowledgements: This paper was partly supported by National Natural Science Foundation of China (11371306), Basic and Frontier Technology Research Project of Henan Province (142300410437) and Key Scientific Research Projects of Institutions of Higher Education of Henan Province (15A110039).

References:

[1] L. Fang, A new one-step smoothing Newton method for nonlinear complementarity problem with P_0-function, *Appl. Math. Comput*. 216, 2010, pp. 1087–1095.

[2] L. Fang, G.P. He and Y.H. Hu, A new smoothing Newton-type method for second-order cone programming problems, *Appl. Math. Comput*. 215, 2009, pp. 1020–1029.

[3] M.S. Gowda and M.A. Tawhid, Existence and limiting behavior of trajectories associated with P_0-equations, *Comput. Optim. Appl*. 12, 1999, pp. 229–251.

[4] P.T. Harker and J.S. Pang, Finite-dimensional variational inequality and non-linear complementarity problems: A survey of theory, algorithms and applications, *Math. Program*. 48, 1990, pp. 161–220.

[5] Z.H. Huang, J. Han, D. Xu and L. Zhang, The non-linear continuation methods for solving the P_0- functions non-linear complementarity problem, *Sci. China Ser. A: Math*. 44(2), 2001, pp. 1107–1114.

[6] C.F. Ma, A new smoothing and regularization Newton method for P_0-NCP, *J. Global Optim*. 48, 2010, pp. 241–261.

[7] R.S. Palais and C.-L. Terng. *Critical point theory and submanifold geometry*, Lecture Notes in Mathematics, vol. 1353. Springer, Berlin, 1988.

[8] L. Qi, Convergence analysis of some algorithm for nonsmooth equations, *Math. Oper. Res*. 18, 1993, pp. 227–244.

[9] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Math. Program*. 58, 1993, pp. 353–367.

[10] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Math. Program*. 87, 2000, pp. 1–35.

[11] J.Y. Tang, L. Dong, J.C. Zhou and L. Fang. A new non-interior continuation method for solving the second-order cone complementarity problem, *Appl. Math. Comput.* 236, 2014, pp. 287–299.

[12] J.Y. Tang, L. Dong, J.C. Zhou and L. Fang, A smoothing Newton method for nonlinear complementarity problems, *Comput. Appl. Math.* 32, 2013, pp. 107–118.

[13] J. Zhang and K.C. Zhang, A variant smoothing Newton method for P_0-NCP based on a new smoothing function, *J. Comput. Appl. Math.* 225, 2009, pp. 1–8.

[14] L. Zhang, J. Han and Z. Huang. Superlinear/quadratic one-step smoothing Newton method for P_0-NCP, *Acta Math. Sinica*, 26(2), 2005, pp. 117–128.

[15] L.P. Zhang, S.Y. Wu and T.R. Gao, Improved smoothing Newton methods for P_0 nonlinear complementarity problems, *Appl. Math. Comput.* 215, 2009, pp. 324–332.