J Glob Optim (2010) 48:241–261
DOI 10.1007/s10898-009-9489-9

A new smoothing and regularization Newton method for P_0-NCP

Changfeng Ma

Received: 17 August 2008 / Accepted: 2 November 2009 / Published online: 14 November 2009

Abstract The nonlinear complementarity problem (denoted by NCP(F)) can be reformulated as the solution of a nonsmooth system of equations. In this paper, we propose a new smoothing and regularization Newton method for solving nonlinear complementarity problem with P_0-function (P_0-NCP). Without requiring strict complementarity assumption at the P_0-NCP solution, the proposed algorithm is proved to be convergent globally and superlinearly under suitable assumptions. Furthermore, the algorithm has local quadratic convergence under mild conditions. Numerical experiments indicate that the proposed method is quite effective. In addition, in this paper, the regularization parameter ε in our algorithm is viewed as an independent variable, hence, our algorithm seems to be simpler and more easily implemented compared to many previous methods.

Keywords Nonlinear complementarity problem · P_0-Function · Smoothing and regularization Newton method · Global convergence · Superlinear/quadratic convergence · Numerical experiment

1 Introduction

The nonlinear complementarity problem with P_0-function (denoted by P_0-NCP): to find a vector $x \in R^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad \langle x, F(x) \rangle = 0. \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F : R^n \to R^n$ are continuously differentiable P_0-function, that is,

The project supported by Fujian Natural Science Foundation (Grant No. 2009J01002) and National Natural Science Foundation of China (Grant No. 10661005).

C. Ma (✉)
School of Mathematics and Computer Science, Fujian Normal University, 350007 Fuzhou, China
e-mail: macf@fjnu.edu.cn

Definition 1.1

(1) A matrix $M \in R^n$ is said to be a P_0-matrix if all its principal minors are nonnegative.
(2) A function $F : R^n \to R^n$ is said to be a P_0-function if for all $x, y \in R^n$ with $x \neq y$, there exists an index $i_0 \in N$ such that

$$x_{i_0} \neq y_{i_0}, \quad (x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] \geq 0.$$

The nonlinear complementarity problems have various important applications in many fields [1,2]. The NCP has been utilized as a general framework for quadratic programming, linear complementarity, and the other mathematical programming problems. Different concepts have been developed to treat this problem. In the last few years growing attention has been paid to diverse approaches which employ a reformulation of NCP as a system of nonlinear equations or a minimization problem [3–9].

Recently, there have been strong interests in smoothing Newton methods for solving the linear/nonlinear complementarity problems [9–20]. Lastly, Zhang et al. [21] have proposed a one-step smoothing Newton method for solving the nonlinear complementarity problem with P_0-function based on the smoothing symmetric perturbed Fischer function. Their algorithm solves only one linear system of equations and performs only one line search per iteration. Without requiring strict complementarity assumption at the P_0-NCP solution, it has been shown that the algorithm converges globally and superlinearly under mild conditions. Furthermore, the algorithm has local quadratic convergence under suitable conditions. Compared to previous literatures, the algorithm has stronger convergence results under weaker conditions.

It is well known that the regularization method is designed to handle ill-posed problems which substitutes the solution of original problem with the solution of a sequence of well-defined problems whose solutions converging to the solution of the original problem. In the context of this paper, we consider the regularization method of NCP (1.1), which consists of solving a sequence of complementarity problems NCP(F_ε):

$$x \geq 0, \quad F_\varepsilon(x) \geq 0, \quad \langle x, F_\varepsilon(x)\rangle = 0, \tag{1.2}$$

where $\varepsilon > 0$ is a regularization parameter tending to zero and F_ε is given by

$$F_\varepsilon(x) = F(x) + \varepsilon x. \tag{1.3}$$

In this paper, we present a smoothing and regularization Newton method for solving the nonlinear complementarity problem with P_0-function. In our algorithm, the regularization parameter ε is viewed as an independent variable, hence, our algorithm seems to be simpler and more easily implemented compared to many previous literatures. Without requiring strict complementarity assumption at the P_0-NCP solution, the proposed algorithm is proved to be convergent globally and superlinearly under suitable assumptions. Furthermore, the algorithm has local quadratic convergence under mild conditions.

The rest of this paper is organized as follows. In Sect. 2, we state some preliminaries. In Sect. 3, we present a smoothing and regularization Newton method for the P_0-NCP. In Sects. 4 and 5, we establish the global, superlinear/quadratic convergence of the proposed algorithm, respectively. Some numerical results are reported in Sect. 6. Conclusions are given in Sect. 7.

The following notations will be used throughout this paper. All vector are column vectors, the superscript ^{T} denotes transpose, R^n (respectively, R) denotes the space of n-dimensional real column vectors (respectively, real numbers), R^n_+ and R^n_{++} denote the nonnegative and positive orthants of R^n, R_+ (respectively, R_{++}) denotes the nonnegative (respectively, positive) line in R. We define $N := \{1, 2, \ldots, n\}$. For any vector $u \in R^n$, we denote by $\text{diag}\{u_i : i \in N\}$ the diagonal matrix whose ith diagonal element is u_i and $\text{vec}\{u_i : i \in N\}$ the vector u. For simplicity, we use $(u; v)$ for the column vector $(u^T; v^T)^T$. The matrix I represents the identity matrix of arbitrary dimension. The symbol $\|\cdot\|$ stands for the 2-norm. We denote by S the solution set of (1.1). For any $\alpha, \beta \in R_{++}$, $\alpha = O(\beta)$ (respectively, $\alpha = o(\beta)$) means α/β is uniformly bounded (respectively, tends to zero) as $\beta \to 0$.

2 Preliminaries

We consider the following the Fischer–Burmeister NCP-function ϕ_{FB}[22]:

$$\phi_{FB}(a, b) = a + b - \sqrt{a^2 + b^2}. \tag{2.1}$$

The Fischer–Burmeister function has many interesting properties. However, it is not differentiable at $(a, b) = (0, 0)$, which limits its applications in dealing with nonlinear complementarity problems. Many smoothing NCP-functions based on the Fischer–Burmeister function (2.1) have been presented for solving the nonlinear complementarity problems (1.1) (see [12,13,23]).

In this paper, we use the following smoothing function to approximate the Fischer–Burmeister NCP-function ϕ_{FB}:

$$\phi(\mu, a, b) = a + b - \sqrt{a^2 + b^2 + 2\mu^2} \tag{2.2}$$

where $\mu > 0$ is a smoothing parameter. The smoothing NCP-function possesses a few nice properties. The following lemma illustrates some simple properties.

Lemma 2.1 *For any $(\mu, a, b) \in R_+ \times R^2$, we have*

(a) $\phi(0, a, b) = \phi_{FB}(a, b)$;
(b) $\phi(\mu, a, b) = 0$ *if and only if* $a \geq 0,\ b \geq 0$ *and* $ab = \mu^2$.

Lemma 2.1 illustrates that the function $\phi(\mu, a, b)$ defined by (2.2) is indeed a smoothing approximation function of the Fischer-Burmeister function $\phi_{FB}(a, b)$. Moreover, we have the following result.

Lemma 2.2 *For any $\mu_1, \mu_2 \in R_{++}$, we have*

$$|\phi(\mu_1, a, b) - \phi(\mu_2, a, b)| < \sqrt{2}|\mu_1 - \mu_2|. \tag{2.3}$$

By simple calculation, for any $\mu \in R_{++}$, we have

$$\phi'_\mu(\mu, a, b) = -\frac{2\mu}{\sqrt{a^2 + b^2 + 2\mu^2}}, \tag{2.4}$$

$$\phi'_a(\mu, a, b) = 1 - \frac{a}{\sqrt{a^2 + b^2 + 2\mu^2}}, \tag{2.5}$$

$$\phi'_b(\mu, a, b) = 1 - \frac{b}{\sqrt{a^2 + b^2 + 2\mu^2}}. \tag{2.6}$$

It is not difficult to see that ϕ'_μ, ϕ'_a and ϕ'_b are continuous with $\mu > 0$. Then, from (2.5)–(2.7), we have the following results.

Lemma 2.3 *For any $(\mu, a, b) \in R_+ \times R^2$, we have*

$$0 < \phi'_a(\mu, a, b) < 2 \quad \textit{and} \quad 0 < \phi'_b(\mu, a, b) < 2. \tag{2.7}$$

Lemma 2.4 *Let $\mu > 0$ and the function $\phi : R_+ \times R^2$ be defined by (2.2). Let $\{a_k\}$, $\{b_k\}$ be any two sequences such that $a_k,\ b_k \to +\infty$ or $a_k \to -\infty$ or $b_k \to -\infty$. Then for any $(\mu, a, b) \in R_+ \times R^2$, we have $|\phi(\mu, a_k, b_k)| \to +\infty$.*

In this paper, we consider the smoothing regularization method for solving NCP (1.1). Let $F_{\varepsilon,i}$ denote the i-th component of F_ε. Using smoothing NCP-function (2.2), we define the mapping $\Phi : R_+ \times R_+ \times R^n \to R^n$ by

$$\Phi(\varepsilon, \mu, x) = \begin{pmatrix} \phi(\mu, x_1, F_{\varepsilon,1}(x)) \\ \vdots \\ \phi(\mu, x_n, F_{\varepsilon,n}(x)) \end{pmatrix}. \tag{2.8}$$

Then, the regularized problem NCP(F_ε) for any given $\varepsilon \geq 0$ can be reformulated as

$$\Phi(\varepsilon, \mu, x) = 0 \quad \text{and} \quad \mu = 0,$$

which leads to a merit function $\Psi : R^{n+2} \to R_+$ for the NCP(F_ε):

$$\Psi(\varepsilon, \mu, x) = \frac{1}{2}\, \|\Phi(\varepsilon, \mu, x)\|^2 = \frac{1}{2} \sum_{i=1}^{n} \phi(\mu, x_i, F_{\varepsilon,i}(x))^2. \tag{2.9}$$

Let $z := (\varepsilon, \mu, x) \in R_+ \times R_+ \times R^n$ and

$$H(z) := H(\varepsilon, \mu, x) := \begin{pmatrix} \varepsilon \\ \mu \\ \Phi(\varepsilon, \mu, x) \end{pmatrix}. \tag{2.10}$$

Then, it is easily verified that the P_0-NCP (1.1) is equivalent to the following system of equations

$$H(z) = 0, \tag{2.11}$$

which naturally induces a merit function $\Theta : R^{n+2} \to R_+$ given by

$$\Theta(z) = \frac{1}{2}\|H(z)\|^2 = \frac{1}{2}\left(\varepsilon^2 + \mu^2 + \|\Phi(z)\|^2\right) = \frac{1}{2}\left(\varepsilon^2 + \mu^2\right) + \Psi(z). \tag{2.12}$$

By (2.4)–(2.6), it is not difficult to see that $H(\cdot)$ is continuously differentiable at any $z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$ with its Jacobian

$$H'(z) = \begin{pmatrix} 1, & 0, & 0 \\ 0, & 1, & 0 \\ u(z), & v(z), & D_1(z) + D_2(z)[F'(x) + \varepsilon I] \end{pmatrix}, \tag{2.13}$$

where

$$\begin{aligned} u(z) &:= \text{vec}\left\{u_i = \phi'_\varepsilon(\mu, x_i, F_{\varepsilon,i}(x)) : i \in N\right\}, \\ v(z) &:= \text{vec}\left\{v_i = \phi'_\mu(\mu, x_i, F_{\varepsilon,i}(x)) : i \in N\right\}, \\ D_1(z) &:= \text{diag}\left\{a_1(z), a_2(z), \ldots, a_n(z)\right\}, \\ D_2(z) &:= \text{diag}\left\{b_1(z), b_2(z), \ldots, b_n(z)\right\} \end{aligned}$$

with

$$
\begin{aligned}
u_i(z) &= x_i - \frac{x_i(F_i(x)+\varepsilon x_i)}{\sqrt{x_i^2+(F_i(x)+\varepsilon x_i)^2+2\mu^2}} \\
v_i(z) &= -\frac{2\mu}{\sqrt{x_i^2+(F_i(x)+\varepsilon x_i)^2+2\mu^2}} \\
a_i(z) &= 1 - \frac{x_i}{\sqrt{x_i^2+(F_i(x)+\varepsilon x_i)^2+2\mu^2}}, \\
b_i(z) &= 1 - \frac{F_i(x)+\varepsilon x_i}{\sqrt{x_i^2+(F_i(x)+\varepsilon x_i)^2+2\mu^2}}.
\end{aligned}
$$

By Lemma 2.3 we obtain that

$$
0 < a_i(z) < 2 \quad \text{and} \quad 0 < b_i(z) < 2 \tag{2.14}
$$

hold for all $i \in N$.

3 Algorithm

We now give our smoothing and regularization Newton algorithm for solving P_0-NCP (1.1). Let $\gamma \in (0,1)$ and $z := (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$. Define a real-value function $\rho : R_{++} \times R_{++} \times R^n \to R_{++}$ by

$$
\rho(z) := \gamma \|H(z)\| \min\{1, \|H(z)\|\}. \tag{3.1}
$$

Algorithm 3.1 (*A Smoothing Newton Method*)

Step 0 *Choose parameters* $\delta, \sigma \in (0,1)$, $\varepsilon_0 > 0$, $\mu_0 > 0$. *Let* $\bar{u} := (\varepsilon_0, \mu_0, 0) \in R_{++} \times R_{++} \times R^n$ *and* $x^0 \in R^n$ *be an arbitrary initial point. Take* $z^0 = (\varepsilon_0, \mu_0, x^0)$ *and choose parameter* $\gamma \in (0,1)$ *such that* $\gamma\|H(z^0)\| < 1$, $\gamma\varepsilon_0 < 0.5$ *and* $\gamma\mu_0 < 0.5$. *Set* $k := 0$.

Step 1 *Stop if* $\|H(z^k)\| = 0$. *Otherwise, compute* $\rho_k := \rho(z^k)$, *where* $\rho(\cdot)$ *is defined by* (3.1).

Step 2 *Solve the following equation to obtain* $\Delta z^k := (\Delta\mu_k, \Delta x^k)$:

$$
H(z^k) + H'(z^k)\Delta z^k = \rho_k \bar{u}. \tag{3.2}
$$

Step 3 *Let* m_k *be the smallest nonnegative integer* m *such that*

$$
\|H(z^k + \delta^m \Delta z^k)\| \le \left[1 - \sigma(1 - \gamma\mu_0 - \gamma\varepsilon_0)\right] \|H(z^k)\| \tag{3.3}
$$

and let $\lambda_k := \delta^{m_k}$.

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

Assumption 3.1 The solution set $S = \{x \in R^n : x \ge 0,\ F(x) \ge 0,\ x^T F(x) = 0\}$ of NCP (1.1) is nonempty and bounded.

Next, we recall some useful results.

Lemma 3.1 *Let* $H : R_{++} \times R_{++} \times R^n \to R^{n+1}$ *and* $\Phi_\varepsilon : R_{++} \times R_{++} \times R^n \to R^n$ *be defined by* (2.10) *and* (2.8), *respectively. Then*

(a) *Φ_ε is continuously differentiable at any $z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$.*
(b) *H is continuously differentiable at any $z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$ with its Jacobian $H'(z)$ defined by (2.13). If F is a P_0-function, then the matrix $H'(z)$ is nonsingular on $R_{++} \times R_{++} \times R^n$.*

Proof It is not difficult to see that Φ_ε is continuously differentiable at any $z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$. We prove (a).

Next we prove (b). It follows from (2.10) and (a) that H is continuously differentiable on $R_{++} \times R_{++} \times R^n$. And for any $\varepsilon > 0$ and $\mu > 0$, by straightforward calculation we obtain from (2.10) the Jacobian $H'(z)$, which is defined by (2.13). Then we obtain from (2.14) that $D_1(z)$ and $D_2(z)$ are positive diagonal matrices for all $z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n$. By (2.13), in order to show that $H'(z)$ is nonsingular, we need only to prove that the matrix $D_1(z) + D_2(z)[F'(x) + \varepsilon I]$ is. In fact, because F is a P_0-function, $F'(x)$ must be a P_0-matrix for all $x \in R^n$ by Theorem 5.8 in [24], which implies that $F'(x) + \varepsilon I$ is P-matrix. Therefore, the matrix $D_1(z) + D_2(z)[F'(x) + \varepsilon I]$ is nonsingular, which implies that the matrix $H'(z)$ is also nonsingular. Hence, (b) is proved. □

Lemma 3.2 *Assume that F is a P_0-function and $\varepsilon_1, \varepsilon_2, \mu_1, \mu_2$ are given positive numbers satisfying $\varepsilon_1 < \varepsilon_2$, $\mu_1 < \mu_2$. Then, H defined by (2.9) has the property:*

$$\lim_{k \to +\infty} \|H(z^k)\| = +\infty \tag{3.4}$$

for any sequence $\{\varepsilon_k, \mu_k, x^k\}$ such $\varepsilon_k \in [\varepsilon_1, \varepsilon_2]$, $\mu_k \in [\mu_1, \mu_2]$ and $\|x^k\| \to +\infty$.

Proof We prove the lemma by contradiction. In fact, if the lemma is not true, then there exists a sequence $\{z^k = (\varepsilon, \mu_k, x^k)\}$ such that

$$\varepsilon_1 \le \varepsilon_k \le \varepsilon_2, \quad \mu_1 \le \mu_k \le \mu_2, \quad \|\Phi(z^k)\| \le \gamma, \quad \|x^k\| \to \infty, \tag{3.5}$$

where $\gamma > 0$ is certain constant. Since the sequence $\{x^k\}$ is unbounded, then the index set $\mathcal{I} := \{i \in N : \{x_i^k\}$ is unbounded$\}$ is nonempty. Without loss of generality, we can assume that $\{|x_j^k|\} \to +\infty$ for any $j \in \mathcal{I}$. Let the sequence $\{\hat{x}^k\}$ be defined by

$$\hat{x}_i^k = \begin{cases} 0, & \text{if } i \in \mathcal{I}, \\ x_i^k, & \text{if } i \notin \mathcal{I}. \end{cases}$$

Then, $\{\hat{x}^k\}$ is bounded obviously. Noting that F is a P_0-function, by Definition 1.1, we have

$$\begin{aligned} 0 &\le \max_{i \in N} (x_i^k - \hat{x}_i^k)[F_i(x^k) - F_i(\hat{x}^k)] \\ &= \max_{i \in \mathcal{I}} x_i^k [F_i(x^k) - F_i(\hat{x}^k)] \\ &= x_{j_0}^k [F_{j_0}(x^k) - F_{j_0}(\hat{x}^k)], \end{aligned} \tag{3.6}$$

where j_0 is one of the indices for which the max is attained, and j_0 is assumed, without loss of generality, to be independent of k. Since $j_0 \in \mathcal{I}$, we have

$$|x_{j_0}^k| \to +\infty \quad \text{as} \quad k \to +\infty. \tag{3.7}$$

We now consider the following two cases:

(1) If $x_{j_0}^k \to +\infty$ as $k \to +\infty$, noting that $F_{j_0}(\hat{x}^k)$ is bounded by the continuity of F_{j_0}, it follows from (3.6) that $F_{j_0}(\hat{x}^k)$ does not tend to $-\infty$. This implies that $F_{j_0}(x^k) +$

$\varepsilon_k x_{j_0}^k \to +\infty$ as $k \to +\infty$. From Lemma 2.4 where $x_{j_0}^k \to +\infty$ and $F_{j_0}(x^k)+\varepsilon_k x_{j_0}^k \to +\infty$, we obtain that

$$|\phi(\mu_k, x_{j_0}^k, F_{\varepsilon_k, j_0}(x^k))| \to +\infty.$$

(2) If $x_{j_0}^k \to -\infty$, then, by boundedness of $F_{j_0}(\hat{x}^k)$, we obtain from (3.6) that $F_{j_0}(x^k) \le F_{j_0}(\hat{x}^k)$. Noting that $0 < \mu_1 \le \mu_k \le \mu_2$, we also get from Lemma 2.4 that

$$|\phi(\mu_k, x_{j_0}^k, F_{\varepsilon_k, j_0}(x^k))| \to +\infty.$$

In either case we obtain $\|\Phi(z^k)\| \to +\infty$, which contradicts with (3.5). □

Remark 3.1 Lemma 3.2 indicates that, under the assumption of F being a P_0-function, the level set

$$L(\gamma) = \{z = (\varepsilon, \mu, x) \in R_{++} \times R_{+} \times R^n \mid \|H(z)\| \le \gamma\}$$

is bounded.

Lemma 3.3 *Assumption 3.1 holds and F is a P_0-function. Suppose that $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ is a infinite sequence such that for any $k \ge 0$, $\varepsilon_k > 0$, $\mu_k > 0$ and $\gamma_k \ge 0$ satisfying*

$$\lim_{k\to+\infty} \varepsilon_k = 0, \quad \lim_{k\to+\infty} \mu_k = 0, \quad \lim_{k\to+\infty} \gamma_k = 0.$$

Furthermore, assume that x^k satisfies the condition $\|\Phi(\varepsilon_k, \mu_k, x^k)\| \le \gamma_k$ for each $k \ge 0$. Then, the sequence $\{x^k\}$ is bounded and every accumulation point of $\{x^k\}$ is a solution of NCP(F) (1.1).

4 Global convergence

In this section, we consider the global convergence of Algorithm 3.1. First, define the set

$$\Omega := \{z = (\varepsilon, \mu, x) \in R_{++} \times R_{++} \times R^n : \varepsilon \ge \rho(z)\varepsilon_0,\ \mu \ge \rho(z)\mu_0\}, \tag{4.1}$$

where $\rho(\cdot)$ is defined in (3.1) and ε_0, μ_0 are given in Step 0 of Algorithm 3.1. The following theorem show that Algorithm 3.1 is well-defined and generates an infinite sequence with some good feature.

Theorem 4.1 *Let $z^0 = (\varepsilon_0, \mu_0, x^0) \in R_{++} \times R_{++} \times R^n$ be given in Algorithm 3.1. Then Algorithm 3.1 is well-defined and generates an infinite sequence $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ with $\varepsilon_k \in R_{++}$, $\mu_k \in R_{++}$ and $z^k \in \Omega$ for each $k \ge 0$.*

Proof If $\mu_k > 0$, because F is a continuously differentiable P_0-function, it follows from Lemma 3.2 (b) that the matrix $H'(z^k)$ is nonsingular. So, Step 2 of Algorithm 3.1 is well-defined at the k-th iteration. By (3.2) we have

$$\Delta\varepsilon_k = -\varepsilon_k + \rho_k\varepsilon_0 \ \text{ and } \ \Delta\mu_k = -\mu_k + \rho_k\mu_0. \tag{4.2}$$

Therefore, for any $\alpha \in (0, 1]$ we obtain that

$$\varepsilon_k + \alpha\Delta\varepsilon_k = (1-\alpha)\varepsilon_k + \alpha\rho_k\varepsilon_0 > 0$$

and

$$\mu_k + \alpha\Delta\mu_k = (1-\alpha)\mu_k + \alpha\rho_k\mu_0 > 0.$$

Let

$$\omega(\alpha) := H(z^k + \alpha \Delta z^k) - H(z^k) - \alpha H'(z^k)\Delta z^k. \tag{4.3}$$

Note that F is continuously differentiable, which implies from (2.10) and Lemma 3.1(b) that $H(\cdot)$ is continuously differentiable around z^k. Hence, (4.3) implies that

$$\|\omega(\alpha)\| = o(\alpha). \tag{4.4}$$

It follows from the definition of $\rho(\cdot)$ that

$$\rho_k \leq \gamma\|H(z^k)\| \quad \text{and} \quad \rho_k \leq \gamma\|H(z^k)\|^2. \tag{4.5}$$

Hence, for any $\alpha \in (0, 1]$, we obtain from (3.2) and (4.3)–(4.5) that

$$\begin{aligned}\|H(z^k + \alpha\Delta z^k)\| &= \|\omega(\alpha) + (1-\alpha)H(z^k) + \alpha\rho_k\bar{u}\| \\ &\leq (1-\alpha)\|H(z^k)\| + \alpha\gamma(\varepsilon_0 + \mu_0)\|H(z^k)\| + o(\alpha) \\ &= [1 - (1 - \gamma\varepsilon_0 - \gamma\mu_0)\alpha]\|H(z^k)\| + o(\alpha),\end{aligned}$$

which implies that there exists a constant $\bar{\alpha} \in (0, 1]$ such that

$$\|H(z^k + \alpha\Delta z^k)\| \leq [1 - \sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)\alpha]\|H(z^k)\|$$

holds for any $\alpha \in (0, \bar{\alpha}]$. This indicates that Step 3 of Algorithm 3.1 is well-defined at the k-th iteration. Thus, by (4.2) and Steps 3 and 4 of Algorithm 3.1, we have $\lambda_k \in (0, 1]$ and

$$\begin{aligned}\varepsilon_{k+1} &= \varepsilon_k + \lambda_k\Delta\varepsilon_k = (1-\lambda_k)\varepsilon_k + \lambda_k\rho_k\varepsilon_0 > 0, \\ \mu_{k+1} &= \mu_k + \lambda_k\Delta\mu_k = (1-\lambda_k)\mu_k + \lambda_k\rho_k\mu_0 > 0.\end{aligned}$$

Hence, from $\varepsilon_0 > 0, \mu_0 > 0$ and the above statements, we obtain that Algorithm 3.1 is well-defined and generates an infinite sequence $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ with $\varepsilon_k > 0$ and $\mu_k > 0$ for each $k \geq 0$.

Next we prove the second part of conclusion, that is, $z^k \in \Omega$ for all $k \geq 0$. We prove the fact by mathematical induction on k. In fact, it is obvious that $\rho(z^0) \leq \gamma\|H(z^0)\| < 1$, i.e., $\varepsilon_0 \geq \rho_0\varepsilon_0$ and $\mu_0 \geq \rho_0\mu_0$. So, $z^0 \in \Omega$. Suppose that $z^k \in \Omega$, i.e., $\varepsilon_k \geq \rho_k\varepsilon_0$ and $\mu_k \geq \rho_k\mu_0$. Then by (4.2) we have

$$\begin{aligned}\varepsilon_{k+1} - \rho_{k+1}\varepsilon_0 &= (1-\lambda_k)\varepsilon_k + \lambda_k\rho_k\varepsilon_0 - \rho_{k+1}\varepsilon_0 \\ &\geq \varepsilon_0(\rho_k - \rho_{k+1})\end{aligned} \tag{4.6}$$

and

$$\begin{aligned}\mu_{k+1} - \rho_{k+1}\mu_0 &= (1-\lambda_k)\mu_k + \lambda_k\rho_k\mu_0 - \rho_{k+1}\mu_0 \\ &\geq \mu_0(\rho_k - \rho_{k+1}).\end{aligned} \tag{4.7}$$

On one hand, if $\|H(z^k)\| < 1$, then

$$\rho_k = \gamma\|H(z^k)\|^2, \tag{4.8}$$

or else,

$$\rho_k = \gamma\|H(z^k)\|. \tag{4.9}$$

On the other hand, by (3.3) and the definition of $\rho(\cdot)$, we have

$$\|H(z^{k+1})\| \leq \|H(z^k)\| \tag{4.10}$$

and

$$\rho_{k+1} \le \gamma \|H(z^{k+1})\|, \quad \rho_{k+1} \le \gamma \|H(z^{k+1})\|^2. \tag{4.11}$$

Therefore, it follows from (4.6)–(4.7) together with (4.8)–(4.11) that

$$\varepsilon_{k+1} - \rho_{k+1}\varepsilon_0 \ge 0 \text{ and } \mu_{k+1} - \rho_{k+1}\mu_0 \ge 0$$

which proves $z^k \in \Omega$. □

Lemma 4.1 *Let $H(\cdot)$ be defined by (2.10) and $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ be the iteration sequence generated by Algorithm 3.1. Then, the sequence $\{\|H(z^k)\|\}$ is convergent. If it does not converge to zero, then $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ is bounded.*

Proof By (3.3) and Theorem 4.1 we obtain that the sequence $\{\|H(z^k)\|\}$ is monotonically decreasing and $\{z^k\} \subset \Omega$. Therefore, by the definition of ρ_k, it is not difficult to see that both $\{\|H(z^k)\|\}$ and $\{\rho_k\}$ are convergent. Then there exist $h^*, \rho_* \ge 0$ such that

$$\lim_{k\to\infty} \|H(z^k)\| = h^* \quad \text{and} \quad \lim_{k\to\infty} \rho_k = \rho_*. \tag{4.12}$$

If $\{\|H(z^k)\|\}$ does not converge to zero, we have $h^* > 0$ and $\rho_* = \gamma h^* \min\{1, h^*\} > 0$. Using (3.2) and $\{z^k\} \subset \Omega$, we obtain that

$$\varepsilon_{k+1} = \varepsilon_k + \lambda_k \Delta\varepsilon_k = (1-\lambda_k)\varepsilon_k + \lambda_k \rho_k \varepsilon_0 \le \mu_k$$

and

$$\mu_{k+1} = \mu_k + \lambda_k \Delta\mu_k = (1-\lambda_k)\mu_k + \lambda_k \rho_k \mu_0 \le \mu_k,$$

which implies that both $\{\varepsilon_k\}$ and $\{\mu_k\}$ are bounded and

$$0 < \rho_*\varepsilon_0 \le \varepsilon_k \le \varepsilon_0 \quad \text{and} \quad 0 < \rho_*\mu_0 \le \mu_k \le \mu_0, \ \forall k \ge 0. \tag{4.13}$$

If $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ is unbounded, it follows from (4.13) that the sequence $\{x^k\}$ is unbounded. Thus, by Lemma 3.3 we have

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

which violates (4.2). Hence, $\{z^k\}$ must be bounded, which completes the proof of the lemma. □

Now we can prove the global convergence of Algorithm 3.1. we have the following results:

Theorem 4.2 *Assume that the infinite sequence $\{z^k = (\varepsilon_k, \mu_k, x^k)\}$ is generated by Algorithm 3.1. Then*

(a) *$\{\|H(z^k)\|\}$ and $\{\mu_k\}$ converge to zero as $k \to +\infty$, and hence any accumulation point of $\{z^k\}$ is a solution of NCP (1.1);*
(b) *if Assumption 3.1 holds, $\{z^k\}$ is bounded and hence it has at least one accumulation point $z^* = (\varepsilon_*, \mu_*, x^*)$ with $H(z^*) = 0$ and $x^* \in S$.*

Proof By Lemma 4.1 we know that $\{\|H(z^k)\|\}$ converges to h^* as $k \to \infty$. Suppose that $\{\|H(z^k)\|\}$ does not converge to zero. Then, $h^* > 0$ and $\{z^k\}$ is bounded by Lemma 4.1. Assume that $z^* = (\varepsilon_*, \mu_*, x^*)$ is an accumulation point of $z^k = (\varepsilon_k, \mu_k, x^k)$. Without loss of generality, we assume that $\{z^k\}$ converges to z^*. Then, by the continuity of H and the definition of $\rho(\cdot)$, we know that $\{\varepsilon_k\}$, $\{\mu_k\}$ and $\{\rho_k\}$ converge to ε_*, μ_* and ρ_*, respectively and that

$$\begin{cases} h^* = \|H(z^*)\| > 0, & \rho_* = \gamma h^* \min\{1, h^*\} > 0, \\ 0 < \rho_* \varepsilon_0 \le \varepsilon_* \le \varepsilon_0, & 0 < \rho_* \mu_0 \le \mu_* \le \mu_0. \end{cases} \tag{4.14}$$

Therefore, by (3.3), we have

$$\lim_{k\to\infty} \lambda_k = 0. \tag{4.15}$$

On one hand, from Step 3 in Algorithm 3.1, we get

$$\|H(z^k + \delta^{m_k-1}\Delta z^k)\| > [1 - \sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)\delta^{m_k-1}]\|H(z^k)\|,$$

which implies that

$$\frac{\|H(z^k + \delta^{m_k-1}\Delta z^k)\| - \|H(z^k)\|}{\delta^{m_k-1}} > -\sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)\|H(z^k)\|,$$

that is,

$$\begin{aligned} &\frac{\|H(z^k + \delta^{m_k-1}\Delta z^k)\|^2 - \|H(z^k)\|^2}{\delta^{m_k-1}} \\ &\quad > -\sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)\|H(z^k)\| \left(\|H(z^k + \delta^{m_k-1}\Delta z^k)\| + \|H(z^k)\|\right). \end{aligned} \tag{4.16}$$

Let $k \to \infty$ in (4.16), by (4.14), we have

$$2H(z^*)^T H'(z^*)\Delta z^* \ge -2\sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)\|H(z^*)\|^2. \tag{4.17}$$

On the other hand, by (3.2), we have

$$H'(z^*)\Delta z^* = -H(z^*) + \rho_* \bar{u},$$

i.e.,

$$H(z^*)^T H'(z^*)\Delta z^* = -\|H(z^*)\|^2 + \rho_* H(z^*)^T \bar{u}, \tag{4.18}$$

Combining (4.17) and (4.18) we deduce that

$$[1 - \sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)]\|H(z^*)\|^2 \le \rho_* H(z^*)^T \bar{u} \le \rho_* \sqrt{\varepsilon_0^2 + \mu_0^2}\|H(z^*)\|.$$

Since $\|H(z^*)\| > 0$ we have

$$\begin{aligned} &[1 - \sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0)]\|H(z^*)\| \\ &\le \rho_* \sqrt{\varepsilon_0^2 + \mu_0^2} \le \gamma\sqrt{\varepsilon_0^2 + \mu_0^2}\|H(z^*)\| \\ &\le \gamma(\varepsilon_0 + \mu_0)\|H(z^*)\|, \end{aligned} \tag{4.19}$$

where the second inequality follows from $\rho_* \le \gamma\|H(z^*)\|$ and the third inequality follows from $\sqrt{\varepsilon_0^2 + \mu_0^2} \le (\varepsilon_0 + \mu_0)$. Since $\|H(z^*)\| > 0$, then (4.19) implies that $1 - \sigma(1 - \gamma\varepsilon_0 - \gamma\mu_0) \le \gamma(\varepsilon_0 + \mu_0)$, i.e.,

$$(1 - \sigma)(1 - \gamma\varepsilon_0 - \gamma\mu_0) \le 0,$$

which contradicts with the fact $\sigma < 1$, $\gamma\varepsilon_0 < 0.5$ and $\gamma\mu_0 < 0.5$. Hence, we have $h^* = 0$ (i.e., $\|H(z^*)\| = 0$), $\varepsilon_* = 0$ and $\mu_* = 0$. Thus, $H(z^*) = 0$, that is, z^* is a solution of NCP (1.1), which prove (a).

Next we prove (b). It follows from (a) that $\|H(z^k)\| \to 0$ as $k \to \infty$. By (2.10) and (2.11), we have

$$\lim_{k\to\infty} \varepsilon_k = 0, \quad \lim_{k\to\infty} \mu_k = 0 \ \text{ and } \ \lim_{k\to\infty} \|\Phi(z^k)\| = 0.$$

Therefore, by the famous mountain pass theorem (Theorem 5.4 in [25]) and by following the similar proof lines of Theorem 3.1 in [26], we get that $\{x^k\}$ is bounded and hence $\{z^k\}$ is. Hence, $\{z^k\}$ has at least one accumulation point $z^* = (\varepsilon_*, \mu_*, x^*)$. By (a), we have $H(z^*) = 0$ and $\varepsilon_* = \mu_* = 0$, $x^* \in S$. □

5 Superlinear/quadratic convergence

In this section, we analyze the rate of convergence for Algorithm 3.1. By Theorem 4.2 (b), we know that Algorithm 3.1 generates a bounded iteration sequence $\{z^k\} \subset \Omega$. Let $z^* = (\mu_*, x^*)$ be an accumulation point of $\{z^k\}$. Then, by Theorem 4.2 we have $\varepsilon_* = \mu_* = 0$ and x^* is a solution of NCP (1.1). To establish the rate of convergence for Algorithm 3.1, we assume that x^* satisfies the nonsingularity condition but may not satisfy the strict complementarity.

In order to analyze the local superlinear/quadratic convergence of Algorithm 3.1, we need the concept of semismoothness for vector value functions. The concept of semismoothness, which was originally introduced by Mifflin [27] for functions and extended by Qi and Sun [28] for vector-valued functions. Convex functions, smooth functions and piecewise linear functions are examples of semismooth function. The composition of semismooth functions is still a semismooth function [27]. Let $\mathcal{F} : R^n \to R^n$ be a locally Lipschitz continuous mapping. Then, from Rademacher's theorem, $\mathcal{F}$ is differentiable almost everywhere and the generalized Jacobian [29] is well-defined such that

$$\partial\mathcal{F}(x) = \mathrm{Co}\left\{ \lim_{x^k \to x, x^k \in D_{\mathcal{F}}} \nabla\mathcal{F}(x^k)^T \right\},$$

where Co denotes a convex hull and $D_{\mathcal{F}}$ denotes a set of points at which $\mathcal{F}$ is differentiable. The function $\mathcal{F}$ is called semismooth at $x \in R^n$, if

$$\lim_{\substack{V \in \partial\mathcal{F}(x+th') \\ h' \to h,\ t \downarrow 0}} \{Vh'\}$$

exists for any $h \in R^n$. The function $\mathcal{F}$ is further said to be strongly semismooth at x if $\mathcal{F}$ is semismooth at x and for any $V \in \partial\mathcal{F}(x+h)$, $h \to 0$,

$$\mathcal{F}(x+h) - \mathcal{F}(x) - Vh = O(\|h\|^2). \tag{5.1}$$

Lemma 5.1 [28] *Suppose that $G : R^n \to R^m$ is a locally Lipschitzian function. Then*

(a) *$G(\cdot)$ has generalized Jacobian $\partial G(x)$ as in Clarke* [29]*. And $G'(x; h)$, the directional derivative of G at x in the direction h, exists for any $h \in R^n$ if G is semismooth at x. Also, $G : R^n \to R^m$ is semismooth at $x \in R^n$ if and only if all it component functions are.*
(b) *$G(\cdot)$ is semismooth at x if and only if for any $V \in \partial G(x+h)$, $h \to 0$,*

$$\|Vh - G'(x; h)\| = o(\|h\|).$$

Also,

$$\|G(x+h) - G(x) - G'(x; h)\| = o(\|h\|).$$

(c) *$G(\cdot)$ is strongly semismooth at x if and only if for any $V \in \partial\Psi(x+h)$, $h \to 0$,*

$$\|Vh - G'(x;h)\| = O(\|h\|^2).$$

Also,

$$\|G(x+h) - G(x) - G'(x;h)\| = O(\|h\|^2).$$

Lemma 5.2 *Let $H : R_{++} \times R_{++} \times R^n \to R^{n+2}$ be defined by (2.10). Then, H is local Lipschitzian and semismooth on $R_{++} \times R_{++} \times R^n$. Furthermore, H is strongly semismooth on $R_{++} \times R_{++} \times R^n$ if $F'(x)$ is Lipschitz continuous on R^n.*

Proof It is not difficult to show that $a+b$, a^2+b^2 and $\sqrt{a^2+b^2}$ are all strongly semismooth for all $(a,b) \in R^2$. By noting that (2.2), the definition of ϕ, and the fact that the composition of strongly semismooth functions is strongly semismooth, we can obtain immediately that $\phi(\cdot,\cdot,\cdot)$ is strongly semismooth at all points $(\mu,a,b) \in R_{++} \times R^2$. Therefore, by Lemma 5.1 (a), we prove the first part of the lemma. If $F'(x)$ is Lipschitz continuous on R^n, then $x_i + F_i(x)$, $x_i^2 + F_i^2(x)$ and $\sqrt{x_i^2 + F_i^2(x)}$ are all strongly semismooth on R^n for all $i \in N$. By Theorem 19 in [12], it easy to see from Lemma 5.1 that the second part of the lemma holds. □

The following is main results of this section.

Theorem 5.1 *Assume that Assumption 3.1 is satisfied and $z^* = (\varepsilon_*, \mu_*, x^*)$ is an accumulation point of the iteration sequence $\{z^k\}$ generated by Algorithm 3.1. If all $V \in \partial H(z^*)$ are nonsingular. Then,*

(a) $\lambda_k \equiv 1$, *for all z^k sufficiently close to z^*;*
(b) *the whole sequence $\{z^k\}$ converges to z^*, that is,*

$$\lim_{k\to\infty} z^k = z^*;$$

(c) *$\{z^k\}$ converges to z^* superlinearly, that is, $\|z^{k+1} - z^k\| = o(\|z^k - z^*\|)$. Moreover, $\mu_{k+1} = o(\mu_k)$;*
(d) *$\{z^k\}$ converges to z^* quadratically if $F'(\cdot)$ is Lipschitz continuous on R^n, that is, $\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2)$. Moreover, $\mu_{k+1} = O(\mu_k^2)$.*

Proof It follows from Theorem 4.2 that $H(z^*) = 0$ and $x^* \in S$. Because all $V \in \partial H(z^*)$ are nonsingular, it follows from Proposition 3.1 in [28] that for all z^k sufficiently close to z^*, we have

$$\|H'(z^k)^{-1}\| \le C, \tag{5.2}$$

where $C > 0$ is some constant. By Lemma 5.2, we know that $H(\cdot)$ is semismooth (strongly semismooth if F' is Lipschitz continuous on R^n, respectively) at z^*. Therefore, for all z^k sufficiently close to z^*, we get

$$\|H(z^k) - H(z^*) - H'(z^k)(z^k - z^*)\| = o(\|z^k - z^*\|) \quad (= O(\|z^k - z^*\|^2)). \tag{5.3}$$

On the other hand, Lemma 5.2 implies that $H(\cdot)$ is locally Lipschitz continuous near z^*. Hence, for all z^k sufficiently close to z^*, we have

$$\|H(z^k)\|^2 = \|H(z^k) - H(z^*)\|^2 = O(\|z^k - z^*\|^2). \tag{5.4}$$

Thus, we obtain from (5.4) and the definition of $\rho(\cdot)$ that

$$\rho_k\mu_0 \leq \gamma\mu_0\|H(z^k)\|^2 = O(\|H(z^k)\|^2) = O(\|z^k - z^*\|^2). \tag{5.5}$$

Then, by (5.2), (5.3) and (5.5), we have

$$\begin{aligned}\|z^k + \Delta z^k - z^*\| &= \|z^k + H'(z^k)^{-1}[-H(z^k) + \rho_k\bar{u}] - z^*\\ &\leq \|H'(z^k)^{-1}\|\left[\|H(z^k) - H(z^*) - H'(z^k)(z^k - z^*)\| + \rho_k\mu_0\right]\\ &= o(\|z^k - z^*\|) \quad (= O(z^k - z^*\|^2)).\end{aligned} \tag{5.6}$$

Similar to the proof of Theorem 3.1 in [30], for all z^k sufficiently close to z^*, we get

$$\|z^k - z^*\| = O(\|H(z^k) - H(z^*)\|). \tag{5.7}$$

Then, because $H(\cdot)$ is semismooth (strongly semismooth if F' is Lipschitz continuous on R^n, respectively) at z^* by Lemma 5.1, H must be local Lipschitz. Therefore, for all z^k sufficiently close to z^*, we obtain that

$$\begin{aligned}\|H(z^k + \Delta z^k)\| &= O(\|z^k + \Delta z^k - z^*\|)\\ &= o(\|z^k - z^*\|) \quad (= O(\|z^k - z^*\|^2))\\ &= o(\|H(z^k) - H(z^*)\|) \quad (= O(\|H(z^k) - H(z^*)\|^2))\\ &= o(\|H(z^k)\|) \quad (= O(\|H(z^k)\|^2)).\end{aligned} \tag{5.8}$$

Note that $\|H(z^k)\| \to 0$ as $k \to \infty$ by Theorem 4.2, hence, (5.8) implies that when z^k sufficiently close to z^*, $\lambda_k = 1$ can satisfy (3.3), which proves (a). Thus, for all z^k sufficiently close to z^* we have

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

which, together with (5.6), proves (b) and

$$\|z^{k+1} - z^*\| = o(\|z^k - z^*\|) \quad (\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2), \text{ respectively}).$$

Next, from (a), (b) and (5.5), we obtain for all sufficiently large k that

$$\mu_{k+1} = \mu_k + \Delta z^k = \rho_k\mu_0 = \gamma\mu_0\|H(z^k)\|^2,$$

which, together with (5.7), implies that

$$\frac{\mu_{k+1}}{\mu_k} = \frac{\|H(z^k)\|^2}{\|H(z^{k-1})\|^2} = \frac{o(\|H(z^{k-1})\|^2)}{\|H(z^{k-1})\|^2} \quad \left(\frac{O(\|H(z^{k-1})\|^4)}{\|H(z^{k-1})\|^2}, \text{ respectively}\right).$$

Therefore, for all z^k sufficiently close to z^* we obtain that

$$\mu_{k+1} = o(\mu_k) \quad (\mu_{k+1} = O(\mu_k^2), \quad \text{respectively}),$$

which completes whole proof. □

6 Numerical experiments

In this section, we present some numerical experiments with Algorithm 3.1 described in the previous section. All the program codes were written in MATLAB and run in MATLAB 7.6 environment. All numerical experiments were done at a PC with Celeron(R) D CPU of

3.06 GHz and RAM of 1,024 MB. Throughout the computational experiments, the parameters in Algorithm 3.1 were set as

$$\varepsilon_0 = 0.05, \quad \mu_0 = 0.05, \quad \delta = 0.9, \quad \sigma = 0.25, \quad \gamma = 0.1.$$

The stopping criterion for the algorithm was $\|H(z^k)\| \le 10^{-6}$.

In the tables of test results, DIM denotes the dimension of the problem (the dimension of the variable x), SP denotes the starting point of x^0, IN denotes the number of iteration, FV denotes the final value of $\Theta(z^k) = \|H(z^k)\|^2/2$ when the algorithm terminates, CPU records the CPU time in second for solving each problem, and RESIDUAL denotes the final residual of $\|x^k - x^*\|$ when the algorithm stops, where x^k is the final value of x and x^* is an accurate solution of the NCP. In the following, we give a detailed description of the test problems.

Example 6.1 Murty Problem. This test problem is the fifth example of Kanzow [31] in Sect. 5 with n variables. The solution is $x^* = (0, \ldots, 0, 1)^T$. The matrix in this example is a P-matrix.

$$M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \quad q = (-1, \ldots, -1)^T.$$

This example was also tested by Kanzow [32], Xu [33], and Burke and Xu [34]. We test this problem by using $x^0 = (1, 1, \ldots, 1)$ as a starting point. The test results are listed in Table 1.

Example 6.2 Fathi Problem. This test problem is from Fathi [35], which was also tested by Xu [33], Kanzow [32]. The solution is $x^* = (1, 0, \ldots, 0)^T$. The matrix M of this example is positive definite.

Table 1 Numerical results for Example 6.1

SP	DIM	IN	FV	RESIDUAL
$(1, 1, \ldots, 1)$	32	4	7.6786×10^{-18}	3.9187×10^{-9}
	64	4	1.6753×10^{-17}	5.7890×10^{-9}
	128	4	2.0518×10^{-16}	2.0248×10^{-8}
	256	5	4.7428×10^{-19}	9.7327×10^{-10}
	512	6	3.9645×10^{-21}	8.9189×10^{-11}
	1024	7	1.0697×10^{-22}	1.4643×10^{-11}
	2048	8	2.8486×10^{-24}	2.3877×10^{-12}
	4096	8	6.8463×10^{-15}	1.1535×10^{-7}

$$
\begin{aligned}
&[M]_{ii} = 4(i-1)+1, \quad i = 1,\ldots,n;\\
&[M]_{ij} = [M]_{ii}+1, \quad i = 1,\ldots,n-1, \quad j = i+1,\ldots,n;\\
&[M]_{ij} = [M]_{jj}+1, \quad j = 1,\ldots,n-1, \quad i = j+1,\ldots,n;\\
&q = (-1,-1,\ldots,-1)^T.
\end{aligned}
$$

We test this problem by using $x^0 = (0, 0, \ldots, 0)$ as a starting point. The test results are listed in Table 2.

Example 6.3 This test problem is from Ahn [36]. The matrix M of this example is:

$$
M = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ 0 & 1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}, \quad q = \begin{pmatrix} -1 \\ -1 \\ -1 \\ \vdots \\ -1 \\ -1 \end{pmatrix}.
$$

We test this problem by using $x^0 = (0, 0, \ldots, 0)$ as a starting point. The test results are listed in Table 3.

Table 2 Numerical results for Example 6.2

SP	DIM	IN	FV	RESIDUAL
$(0, 0, \ldots, 0)$	64	8	7.5798×10^{-21}	8.3088×10^{-10}
	128	8	1.1736×10^{-18}	1.5451×10^{-8}
	256	9	2.3448×10^{-24}	4.4385×10^{-11}
	512	9	1.1830×10^{-13}	1.2477×10^{-6}
	1024	10	5.7002×10^{-16}	9.9874×10^{-7}
	2048	11	1.8393×10^{-18}	5.9860×10^{-8}
	4096	12	5.5722×10^{-22}	9.5803×10^{-10}

Table 3 Numerical results for Example 6.3

SP	DIM	IN	FV	CPU
$(0, 0, \ldots, 0)$	64	5	1.4066×10^{-13}	0.0286
	128	5	5.9168×10^{-15}	0.0749
	256	5	9.1990×10^{-15}	0.2357
	512	5	1.9198×10^{-14}	0.8844
	1024	6	4.1281×10^{-25}	4.9571
	2048	6	2.7852×10^{-20}	26.646
	4096	6	7.4777×10^{-14}	163.59

Table 4 Numerical results for Example 6.4

SP	IN	FV
$(0, 0, 0, 0, 0, 0, 0)$	7	4.8396×10^{-24}
$(1, 1, 1, 1, 1, 1, 1)$	7	2.8586×10^{-24}
$(-1, -1, -1, -1, -1, -1, -1)$	6	8.5192×10^{-16}
$(10, 10, 10, 10, 10, 10, 10)$	7	2.1789×10^{-16}
$(-10, -10, -10, -10, -10, -10, -10)$	10	8.8400×10^{-25}
$(10, 20, 30, 40, 50, 60, 70)$	11	1.9157×10^{-14}
$(-70, -60, -50, -40, -30, -20, -10)$	10	9.1518×10^{-15}
$(10^3, 10^3, 10^3, 10^3, 10^3, 10^3, 10^3)$	13	9.8667×10^{-15}
$(10^5, 10^5, 10^5, 10^5, 10^5, 10^5, 10^5)$	18	9.8753×10^{-16}

Example 6.4 This test problem is the LCP reformulation for 76th problem in the Hock–Schittkowski collection. Let

$$F(x) = \begin{pmatrix} 2x_1 - x_3 + x_5 + 3x_6 - 1 \\ x_2 + 2x_5 + x_6 - x_7 - 3 \\ -x_1 + 2x_3 + x_4 + x_5 + 2x_6 - 4x_7 + 1 \\ x_3 + x_4 + x_5 - x_6 - 1 \\ -x_1 - 2x_2 - x_3 - x_4 + 5 \\ -3x_1 - x_2 - 2x_3 + x_4 + 4 \\ x_2 + 4x_3 - 1.5 \end{pmatrix}.$$

This example was also tested by Pieraccini et al. [37]. The problem has the nondegenerate solution $x^* \simeq (0.2727272727, 2.0909090909,\ 0, 0.5454545454, 0.4545454545,\ 0,\ 0)$. The test results are listed in Table 4 using different points.

Example 6.5 Kojima–Shindo Problem. This test problem is the third example of Jiang and Qi [38] with four variables, which was also tested by Xu [33], Kanzow [4], Pang and Gabriel [39], Huang et al. [40] and Pieraccini et al. [37]. Let

$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

This example has one nondegenerate solution $(1, 0, 3, 0)$ and one degenerate solution $(\sqrt{6}/2, 0, 0, 1/2)$. The results are summarized in Table 5 using different starting points. The asterisk (*) denotes that the limit point generated by Algorithm 3.1 is the degenerate solution, otherwise it is the nondegenerate solution.

Example 6.6 Modified Mathiesen Problem. This test problem is the fifth example of Jiang and Qi [38] with four variables, which was also tested by Huang et al. [40], Kanzow [4], and Pieraccini et al. [37]. Let

$$F(x) = \begin{pmatrix} -x_2 + x_3 + x_4, \\ x_1 - (4.5x_3 + 2.7x_4)/(x_2 + 1), \\ 5 - x_1 - (0.5x_3 + 0.3x_4)/(x_3 + 1), \\ 3 - x_1 \end{pmatrix}.$$

Table 5 Numerical results for Example 6.5

SP	IN	FV	RESIDUAL
(0, 0, 0, 0)	7*	5.1460×10^{-20}	3.6552×10^{-10}
(1, 1, 1, 1)	6	1.9325×10^{-17}	8.8782×10^{-9}
(0, 1, 0, 1)	6*	1.4356×10^{-18}	5.6248×10^{-10}
(1, 1, 0, 0)	7	4.0822×10^{-18}	9.973061×10^{-10}
(1, 0, 1, 0)	11*	3.8025×10^{-26}	3.5685×10^{-14}
(0, 0, 1, 1)	16	6.3633×10^{-23}	1.5152×10^{-11}
$(-1, -1, -1-, 1)$	10*	7.8863×10^{-18}	2.0615×10^{-9}
(10, 10, 10, 10)	7	2.7783×10^{-22}	2.5518×10^{-11}
$(-10, -10, -10, -10)$	12	9.8099×10^{-18}	6.3209×10^{-9}
$(10^3, 10^3, 10^3, 10^3)$	7*	9.7271×10^{-17}	1.4535×10^{-8}
$(10^6, 10^6, 10^6, 10^6)$	11	1.7713×10^{-13}	2.7334×10^{-7}

Note: The asterisk (*) denotes the degenerate solution, otherwise it is the nondegenerate solution

Table 6 Numerical results for Example 6.6

SP	IN	FV	x^k
(1, 1, 1, 1)	4	1.0547×10^{-16}	(0.0779, 0.0000, 0.0000, 0.0000)
(3, 3, 3, 3)	6	1.8260×10^{-24}	(3.0000, 0.0000, 0.0000, 0.0000)
(1, 2, 3, 4)	5	7.5814×10^{-15}	(0.7883, 0.0000, 0.0000, 0.0000)
$(2, 1, 0, -1)$	4	4.4198×10^{-17}	(1.5355, 0.0000, 0.0000, 0.0000)
(10, 10, 10, 10)	6	1.5495×10^{-16}	(3.0000, 0.0000, 0.0000, 0.0000)
$(-30, -30, -30, -30)$	7	3.7697×10^{-17}	(3.0000, 0.0000, 0.0000, 0.0000)
$(10^2, 10^2, 10^2, 10^2)$	10	1.6217×10^{-22}	(0.0000, 0.0000, 0.0000, 0.0000)
$(-10^3, -10^3, -10^3, -10^3)$	9	4.5414×10^{-22}	(0.4205, 0.0000, 0.0000, 0.0000)
$(10^4, 10^4, 10^4, 10^4)$	12	8.1585×10^{-20}	(0.3125, 0.0000, 0.0000, 0.0000)
$(10^6, 10^6, 10^6, 10^6)$	21	4.2162×10^{-20}	(0.0453, 0.0000, 0.0000, 0.0000)

This example has infinitely many solutions $(\lambda, 0, 0, 0)$, where $\lambda \in [0, 3]$. For $\lambda = 0, 3$, the solutions are degenerate, and for $\lambda \in (0, 3)$ nondegenerate. The test results for Example 6.6. are listed in Table 6 using different starting points.

Example 6.7 Nash equilibrium model. This test problem is from MCPLIB with ten variables. The test function $F(x) = (F_1(x), \ldots, F_{10}(x))^T$ is defined by

$$F_i(x) = c_i + (L_i x_i)^{1/\beta_i} - \left[\frac{5000}{\sum_{k=1}^{10} x_k}\right]^{1/\gamma} + \frac{x_i}{\gamma \sum_{k=1}^{10} x_k}\left[\frac{5000}{\sum_{k=1}^{10} x_k}\right]^{1/\gamma}, \quad 1 \le i \le 10,$$

where $\gamma = 1.2$, $L_i = 10(1 \le i \le 10)$, and $c = (5.0, 3.0, 8.0, 5.0, 1.0, 3.0, 7.0, 4.0, 6.0, 3.0)^T$, $\beta = (1.2, 1.0, 0.9, 0.6, 1.5, 1.0, 0.7, 1.1, 0.95, 0.75)^T$. In the example, we take the parameter $\delta = 0.7$, and starting point x^0: (1) e; (2)$10e$; (3) $(1.0, 1.2, 1.4, 1.6, 1.8, 2.1, 2.3, 2.5, 2.7, 2.9)^T$; (4) $(7, 4, 3, 1, 8, 4, 1, 6, 3, 2)^T$; (5) $(5, 4, 3, 2, 1, 6, 7, 8, 9, 10)^T$, and other

Table 7 Numerical results for Example 6.7

SP	IN	FV	CPU
(1)	26	2.8970×10^{-13}	0.0275
(2)	27	3.7600×10^{-13}	0.0771
(3)	24	4.6561×10^{-13}	0.0259
(4)	24	1.6004×10^{-13}	0.0258
(5)	24	2.9653×10^{-13}	0.0290

Table 8 Numerical results for Example 6.8

SP	IN	FV	CPU
(1)	7	1.7046×10^{-21}	0.0145
(2)	11	6.3498×10^{-17}	0.0899
(3)	7	7.4251×10^{-17}	0.0149
(4)	8	2.2719×10^{-17}	0.0153
(5)	8	3.8764×10^{-17}	0.0155
(6)	8	1.0439×10^{-15}	0.0154
(7)	10	8.2963×10^{-15}	0.0180

parameters are same as previous example. The test results for Example 6.7 are listed in Table 7 using these starting points (1)–(5).

Example 6.8 Hanshoop problem. This test problem is from MCPLIB. The test function is:

$$F(x, y, u) = \begin{bmatrix} -\nabla v(x) \\ 0 \\ w \end{bmatrix} + \begin{bmatrix} 0 & A^T - \alpha B^t & C^T \\ B - A & 0 & 0 \\ -C & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ u \end{bmatrix},$$

where $v(x) = (x_1 + 2.5x_2)^p (2.5x_3 + x_4)^p (2x_5 + 3x_6)^p, \alpha = 0.7, p = 0.2, w = (0.8, 0.8)^T$, and

$$A = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 2 & 2 & 1 & 1 & 1 & 0.5 & 1.5 & 0.5 \end{bmatrix},$$

$$B = \begin{bmatrix} 1.5 & 1.5 & 1.5 & 1.5 & 1.5 & 1.5 & 4 & 3 & 1.5 & 1.5 \\ 2.7 & 2.7 & 1.8 & 1.8 & 0.9 & 0.9 & 0.9 & 0.4 & 2 & 1.5 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.5 & 1.5 & 1.5 & 0.5 & 0.5 & 1.5 & 1.5 & 0.5 & 0.5 & 1.5 \end{bmatrix}.$$

In the example, we take the same parameters as ones of Example 6.7. We take the starting point $y^0 = (0, 0)^T$, $u^0 = (0, 0)^T$ and x^0 : (1) $0.3e$; (2) $(0, 0.3, 0.3, 0, 0, 0.1, 0.3, 0, 0, 0.3)^T$; (3) $(0.3, 0, 0.3, 0, 0.3, 0, 0.3, 0, 0.3, 0)^T$; (4) $0.5e$; (5) e; (6) $4e$; (7) $8e$. The test results for Example 6.8 are listed in Table 8 using these starting points (1)–(7).

Example 6.9 Harker and Pang Problem. This example was tested by Kanzow in [41].The matrix M is computed as follows: let $A, B \in R^{n\times n}$, and $q, \eta \in R^n$ be randomly generated such that $a_{ij}, b_{ij}(-5, 5), q_i \in (-500, 500)$, and $\eta_i \in (0.0, 0.3)$ and that B is skew-symmetric. Define $M = A^T A + B + \text{diag}(\eta)$. Then, M is a P-matrix. For several values

Table 9 Numerical results for Example 6.9

DIM. (n)	25	50	100	200	400	600	800	1000	2000
Max.-Iter.	13	17	18	20	21	23	24	24	27
Avg.-Iter.	7.4	9.2	9.5	10.1	10.3	10.7	10.9	11.2	13.1
Min.-Iter.	2	2	2	2	2	2	2	2	2

Table 10 Numerical results for Example 6.10

DIM. (n)	25	50	100	200	400	600	800	1000	2000
Max.-Iter.	14	15	21	20	24	24	25	26	30
Avg.-Iter.	11.8	12.8	17.3	18.7	20.3	22.5	22.7	24.1	26.4
Min.-Iter.	9	10	15	17	19	21	21	23	22

Table 11 Numerical results for Example 6.11

DIM. (n)	25	50	100	200	400	600	800	1000	2000
Max.-Iter.	9	10	10	11	12	12	14	15	15
Avg.-Iter.	7.4	8.2	8.9	9.8	10.9	11.1	12.1	12.4	13.9
Min.-Iter.	6	7	8	8	10	10	11	11	13

of n, 10 examples have been generated in this way. The maximum, average, and minimum numbers of iterations needed by the algorithms are summarized in Table 9. In all test runs, $x^0 = (0, \dots, 0)^T$ has been chosen as starting vector.

Example 6.10 Harker and Pang Problem. This example was also tested by Kanzow in [41]. In this example, M is computed in the same way as in Example 6.9, and $q \in R^n$ is randomly generated with entries $q_i \in (-500, 0)$. Table 10 contains our numerical results, which we have obtained using the starting vector $x^0 = (0, \dots, 0)^T$.

Example 6.11 Pang Problem. This example was tested by Chen and Harker in [42]. In this example, $M = (m_{ij})$ is computed as follows:

$$m_{ij} = \begin{cases} 6, & \text{if } i = j, \\ -4, & \text{if } j = i - 1 \geq 1 \text{ or } j = i + 1 \leq n, \\ 1, & \text{if } j = i - 2 \geq 1 \text{ or } j = i + 2 \leq n \\ 0, & \text{else} \end{cases}$$

For each $1 \leq i \leq n$, by 0.5 probability, we generate $x_i^* = 0$ and Pseudo random number $x_i^* \in (0, 100)$. If $x_i^* > 0$, let $y_i^* = 0$; else if $x_i^* = 0$, let $y_i^* \in (0, 100)$ be generated randomly. Then we take $q = y^* - Mx^*$. Table 11 contains our numerical results, which we have obtained using the starting vector $x^0 = (0, \dots, 0)^T$.

7 Conclusions

Based on the ideas developed in smoothing Newton methods, we propose a new smoothing and regularization Newton method for solving nonlinear complementarity problem with P_0-function (P_0-NCP). Without requiring strict complementarity assumption at the P_0-NCP solution, the proposed algorithm is proved to be convergent globally and superlinearly under suitable assumptions. Furthermore, the algorithm has local quadratic convergence under mild conditions. Numerical experiments indicate that the proposed method is quite effective. In addition, in this paper, the regularization parameter ε in our algorithm is viewed as an independent variable, hence, our algorithm seems to be simpler and more easily implemented compared to many previous methods.

Acknowledgments The author expresses his heartfelt thanks to two anonymous referees for many constructive proposals that improved the paper greatly.

References

1. Harker, P.T., Pang, J.-S.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Program. **48**(1), 161–220 (1990)
2. Ferris, M.C., Pang, J.-S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**(3), 669–713 (1997)
3. Geiger, C., Kanzow, C.: On the solution of monotone complementarity problems. Comput. Optim. Appl. **5**, 155–173 (1996)
4. Kanzow, C.: Some equation-based methods for the nonlinear complementarity problem. Optim. Methods Softw. **3**, 327–340 (1994)
5. Ma, C.-F., Nie, P.-Y., Liang, G.-P.: A new smoothing equations approach to the nonlinear complementarity problems. J. Comput. Math. **21**, 747–758 (2003)
6. Pang, J.S., Gabriel, S.A.: NE/SQP: a robust algorithm for nonlinear complementarity problem. Math. Program. **60**, 295–337 (1993)
7. Pang, J.S.: A B-differentiable equations based, globally and locally quadratically convergent algorithm for nonlinear programming, complementarity, and variational inequality problems. Math. Program. **51**, 101–131 (1991)
8. Nie, P.-Y.: A null space approach for solving nonlinear complementarity problems. Acta Mathematicae Applicatae Sinica, English Series **22**(1), 9–20 (2006)
9. Fischer, A.: Solution of monotone complementarity problems with locally Lipschitz functions. Math. Program. **76**(2), 513–532 (1997)
10. Chen, B., Xiu, N.: A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen–Mangasarian smoothing functions. SIAM J. Optim. **9**(2), 605–623 (1999)
11. Chen, X., Qi, L., Sun, D.: Global and superlinear convergence of the smoothing Newton method and its application to general box-constrained variational inequalities. Math. Comput. **67**(1), 519–540 (1988)
12. Chen, B., Harker, P.T.: Smoothing approximations to nonlinear complementarity problems. SIAM J. Optim. **7**(1), 403–420 (1997)
13. Qi, H.: A regularized smoothing Newton method for box constrained variational inequality problems with P0-functions. SIAM J. Optim. **10**(1), 315–330 (2000)
14. Qi, L., Sun, D.: Improving the convergence of non-interior point algorithm for nonlinear complementarity problems. Math. Comput. **69**(1), 283–304 (2000)
15. Tseng, P.: Error bounds and superlinear convergence analysis of some Newton-type methods in optimization. In: Di Pillo, G., Giannessi, F. (eds.) Nonlinear Optimization and Related Topics, pp. 445–462. Kluwer, Boston (2000)
16. Derkse, S., Ferris, M.: MCPLIB: a collection of nonlinear mixed complementarity problems. Optim. Methods Softw. **5**, 319–345 (1995)
17. Derkse, S., Ferris, M.: The PATH solver: a nonmonotone stabilization schems for mixed complementarity problems. Optim. Methods Softw. **5**, 123–156 (1995)

18. Kojima, M., Megiddo, N., Nomo, T., Yoshise, A.: A Unified Approach to Interior-point Algorithms for Linear Complementarity Problems. Lecture Notes in Computer Science, vol. 538. Springer, New York (1991)
19. Pardalos, P.M., Rosen, J.B.: Global optimization approach to the linear complementarity problem. SIAM J. Sci. Stat. Comput. **9**(2), 341–353 (1988)
20. Pardalos, P.M., Ye, Y., Han, C.-G., Kalinski, J.: Solution of P-matrix linear complementarity problems using a potential reduction algorithm. SIAM J. Matrix Anal. Appl. **14**(4), 1048–1060 (1993)
21. Zhang, L., Han, J., Huang, Z.: Superlinear/quadratic one-step smoothing Newton method for P_0-NCP. Acta Mathematica Sinica **26**(2), 117–128 (2005)
22. Fischer, A.: A special Newton-type optimization method. Optimization **24**, 269–284 (1992)
23. Qi, L., Sun, D., Zhou, G.: A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequality problems. Math. Program. **87**(1), 1–35 (2000)
24. Moré, J.J., Rheinboldt, W.C.: On P- and S-functions and related classes of n-dimensional nonlinear mappings. Linear Algebra Appl. **6**(1), 45–68 (1973)
25. Facchinei, F., Kanzow, C.: Beyond monotonicity in regularization methods for nonlinear complementarity problems. SIAM J. Control Optim. **37**(2), 1150–1161 (1999)
26. Huang, Z., Han, J., Xu, D., Zhang, L.: The noninterior continuation methods for solving the P_0-function nonlinear complementarity problem. Sci. China **44**(2), 1107–1114 (2001)
27. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**(1), 957–972 (1977)
28. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**(2), 353–367 (1993)
29. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
30. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**(1), 227–244 (1993)
31. Kanzow, C.: Global convergence properties of some iterative methods for linear complementarity problems. Optimization **6**, 326–334 (1996)
32. Kanzow, C.: Some noninterior continuation methods for linear complementarity problems. SIAM J. Matrix Anal. Appl. **17**, 851–868 (1996)
33. Xu, S.: The global linear convergence of an infeasible noninterior path-following algorithm for complementarity problems with uniform P-funcitons. Math. Program. **87**, 501–517 (2000)
34. Burke, J., Xu, S.: The global linear convergence of a non-interior path-following algorithm for linear complementarity problems. Math. Oper. Res. **23**, 719–734 (1998)
35. Fathi, Y.: Computational complexity of LCPs associated with positive definite matrices. Math. Program. **17**, 335–344 (1979)
36. Ahn, B.H.: Iterative methods for linear complementarity problem with upperbounds and lowerbounds. Math. Prog. **26**, 265–315 (1983)
37. Pieraccini, S., Gasparo, M.G., Pasquali, A.: Global Newton-type methods and semismooth reformulations for NCP. Appl. Numer. Math. **44**, 367–384 (2003)
38. Jiang, H., Qi, L.: A new nonsmooth equations approach to nonlinear complementarity problems. SIAM J. Control Optim. **35**, 178–193 (1997)
39. Pang, J.S., Gabriel, S.A.: NE/SQP: a robust algorithm for the nonlinear complementarity problem. Math. Program. **60**, 295–337 (1993)
40. Huang, Z.H., Han, J., Chen, Z.: Predictor-corrector smoothing Newton method, based on a new smoothing function, for solving the nonlinear complementarity problem with a P_0 function. J. Optim. Theory Appl. **117**, 39–68 (2003)
41. Kanzow, C.: Some noninterior continuation methods for linear complementarity problems. SIAM J. Matrix Anal. Appl. **17**(4), 851–868 (1996)
42. Chen, B., Harker, P.: A non-interior-point continuation method for linear complementarity problems. SIAM J. Matrix Anal. Appl. **14**(4), 1168–1190 (1993)