RESEARCH   **Open Access**

A smoothing and regularization predictor-corrector method for nonlinear inequalities

Haitao Che[*]

[*]Correspondence:
haitaoche@163.com
School of Mathematics and
Information Science, Weifang
University, Weifang, Shandong
261061, China

Abstract

For a system of nonlinear inequalities, we approximate it by a family of parameterized smooth equations via a new smoothing function. We present a new smoothing and regularization predictor-corrector algorithm. The global and local superlinear convergence of the algorithm is established. In addition, the smoothing parameter μ and the regularization parameter ε in our algorithm are viewed as different independent variables. Preliminary numerical results show the efficiency of the algorithm.

MSC: 90C33; 90C30; 15A06

Keywords: nonlinear inequalities; predictor-corrector algorithm; P_0-function; smoothing function

1 Introduction
Consider the following system of nonlinear inequalities:

$$f(x) \leq 0, \tag{1.1}$$

where $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))^\top$ and $f_i : R^n \to R$ is a continuously differentiable function for $i = 1, 2, \ldots, n$. This problem finds applications in data analysis, set separation problems, computer-aided design problems and image reconstructions [1–3]. Among various solution methods for the inequality problems [4–10], the smoothing-type methods receive much attention [8–10] which first transform the problem as a system of nonsmooth equations and approximate it by a smooth equation and then solve it by the smoothing Newton methods. Since the derivative of the underlying mapping may be seriously ill-conditioned, which may prevent the smoothing methods from converging to a solution of the problem, a perturbed regularization technique is introduced to overcome this drawback [9, 11, 12]. In 2003, Huang *et al.* proposed a predictor-corrector smoothing Newton method for nonlinear complementarity problem with a P_0 function based on the perturbed minimum function [13]. The method was shown to be locally superlinear convergent under the assumptions that all $V \in \partial H(z^*)$ are nonsingular and $f'(x)$ is locally Lipschitz continuous around x^*.

In this paper, motivated by the smoothed penalty function for constrained optimization [14], we construct a new smoothing function for nonlinear inequalities, and thus we

can approximate the nonsmooth system of transformed equations by a system of smooth equations. We develop a regularization smoothing predictor-corrector method for solving the problem by modifying and extending the method in [13]. Besides choosing an arbitrarily starting point, the presented algorithm is simpler than the predictor-corrector noninterior continuation methods developed by Burke and Xu [15].

The rest of this paper is organized as follows. In Section 2, we review some preliminaries to be used in the subsequent analysis and introduce a new smoothing function and its properties. In Section 3, we present a smoothing and regularization predictor-corrector method for solving the nonlinear inequalities and establish the global and local convergence of the proposed algorithm. Preliminary numerical experiments are reported to show the efficiency of the algorithm in Section 5.

To end this section, we introduce some notations used in this paper. The set of $m \times k$ matrices with real entries is denoted by $R^{m \times k}$, R^n_+ (R^n_{++}) denotes the nonnegative (positive) orthant in R^n. The superscript $^\top$ denotes the transpose of a matrix or a vector. Define $N = \{1, 2, \ldots, n\}$, and for any vector $a \in R^n$, we let D_a denote the diagonal matrix whose i-th diagonal element is a_i. $\|u\|$ denotes the 2-norm of a vector $u \in R^n$. For a continuously differentiable function $f : R^n \to R^m$, we denote the Jacobian of f at $x \in R^n$ by $f'(x)$.

2 Smooth reformulation of nonlinear inequalities

In this section, we first review some definitions and basic results, and then introduce a new smoothing function and show its properties.

Definition 2.1 A matrix $M \in R^{n \times n}$ is said to be a P_0-matrix if every principle minor of M is nonnegative.

Definition 2.2 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

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

For a P_0-matrix, the following conclusion holds [16].

Lemma 2.1 *If $M \in R^{n \times n}$ is a P_0-matrix, then every matrix of the form*

$$D_a + D_b M$$

is nonsingular for all positive definite diagonal matrices $D_a, D_b \in R^{n \times n}$.

Definition 2.3 Suppose that $G : R^n \to R^m$ is a locally Lipschitz function. G is said to be semi-smooth at x if G is directionally differentiable at x and

$$\lim_{V \in \partial G(x+th'), h' \to h, t \to 0^+} \big\{Vh'\big\}$$

exists for any $h \in R^n$, where $\partial G(x)$ denotes the generalized derivative in [17].

The concept of semi-smoothness was originally introduced by Mifflin for functions [18]. Qi and Sun extended the definition of a semi-smooth function to vector-valued functions

[19]. Convex functions, smooth functions, piecewise linear functions, convex and concave functions, and sub-smooth functions are examples of semi-smooth functions. A function is semi-smooth at x if and only if all its component functions are. The composition of semi-smooth functions is still a semi-smooth function.

Lemma 2.2 [19] *Suppose that $\varphi : R^n \to R^m$ is a locally Lipschitz function semi-smooth at x. Then*

(a) *for any $V \in \partial\varphi(x + th)$, $h \to 0$,*

$$Vh - \varphi'(x; h) = o(\|h\|),$$

(b) *for any $h \to 0$,*

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

For problem (1.1), based on the function

$$x_+ = \big(\max\{0, x_1\}, \ldots, \max\{0, x_n\}\big)^\top, \quad \text{for } x \in R^n,$$

it can be transformed into the following system of equations [8]:

$$f(x)_+ = 0. \tag{2.1}$$

Since problem (2.1) is a nonsmooth equation, the classical Newton methods cannot be used to solve it. Following the ideas in [14, 20, 21], we adopt the following smoothing function to approximate the nonsmooth equation

$$\phi(\mu, a) = \frac{1}{2}\left(a + \mu\left(\ln 2 + \ln\left(1 + \cosh\frac{a}{\mu}\right)\right)\right), \tag{2.2}$$

where a smoothing parameter $\mu > 0$ and $\cosh x = \frac{e^x + e^{-x}}{2}$.

This new smoothing function has the following properties.

Lemma 2.3 *For any $(\mu, a) \in R_{++} \times R^n$, it holds that*

(1) *$\phi(\cdot, \cdot)$ is continuously differentiable at any $(\mu, a) \in R_{++} \times R^n$.*
(2) *Let $\phi(0, a) = \lim_{\mu \to 0} \phi(\mu, a)$, then $\phi(0, a) = a_+$.*
(3) *$\frac{\partial \phi(\mu, a)}{\partial a} \geq 0$ at any $(\mu, a) \in R_{++} \times R^n$.*

Proof (1) is straightforward, so we only prove (2) and (3).

For (2), following the ideas in [20], we have

$$|a| \approx \varphi(\mu, a) = \mu\left(\ln 2 + \ln\left(1 + \cosh\frac{a}{\mu}\right)\right).$$

Furthermore, one can obtain the estimate [20]

$$\big||a| - \varphi(\mu, a)\big| \leq \frac{8\mu}{3} e^{-\frac{|a|}{\mu}}.$$

Then $\phi(0,a) = a_+$.

For (3), by a simple calculation, we can show

$$\frac{\partial \varphi(\mu,a)}{\partial a} = \frac{\sinh \frac{a}{\mu}}{1 + \cosh \frac{a}{\mu}} \in (-1,1),$$

then $\frac{\partial \phi(\mu,a)}{\partial a} \geq 0$ at any $(\mu,a) \in R_{++} \times R^n$. We complete the proof. □

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

$$H(z) = H(\mu,\varepsilon,x) = \begin{pmatrix} \mu \\ \varepsilon \\ \Phi(z) \end{pmatrix}, \tag{2.3}$$

where

$$\Phi(z) = \Phi(\mu,\varepsilon x) = \begin{pmatrix} \phi(\mu,f_1(x)) \\ \vdots \\ \phi(\mu,f_n(x)) \end{pmatrix} + \varepsilon x. \tag{2.4}$$

Define a merit function

$$\psi(z) = \left\| H(z) \right\|^2 = \mu^2 + \varepsilon^2 + \left\| \Phi(z) \right\|^2. \tag{2.5}$$

Then the inequalities (1.1) can be reformulated as the following nonlinear equations:

$$H(z) = H(\mu,\varepsilon,x) = 0. \tag{2.6}$$

Theorem 2.1 *Let $H(\mu,\varepsilon,x)$ be defined as (2.3). Then*

(a) *$H(\mu,\varepsilon,x)$ is continuously differentiable at any $z = (\mu,\varepsilon,x) \in R_{++} \times R_{++} \times R^n$ with its Jacobian*

$$H'(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \Phi'_\mu(z) & x & \Phi'_x(z) \end{pmatrix}, \tag{2.7}$$

where

$$\Phi'_\mu(z) = \left\{ \phi'_\mu\big(\mu,f_1(x)\big), \ldots, \phi'_\mu\big(\mu,f_n(x)\big) \right\}^\top,$$

$$\Phi'_x(z) = \phi'_x\big(\mu,f(x)\big) + \varepsilon I = \frac{1}{2} \operatorname{diag}\left\{ \left(1 + \frac{\sinh \frac{f_i(x)}{\mu}}{1 + \cosh \frac{f_i(x)}{\mu}} \right) : i \in N \right\} f'(x) + \varepsilon I.$$

(b) *If f is a P_0-function, then $H'(z)$ is nonsingular at any $R_{++} \times R_{++} \times R^n$.*

Proof (a) is straightforward, so we only prove (b). For (b), we only need to show $\Phi'_x(z)$ is nonsingular. In fact, since f is a P_0-function, then $f'(x)$ is a P_0-matrix for all $x \in R^n$ by

Theorem 3.3 in [22]. We also note that $\mathrm{diag}\{(1 + \frac{\sinh\frac{f_i(x)}{\mu}}{1+\cosh\frac{f_i(x)}{\mu}}) : i \in N\}$ and εI are positive diagonal matrices, we know that $\Phi'_x(z)$ is nonsingular by Lemma 2.1, which implies that $H'(z)$ is also nonsingular. This completes the proof. □

3 Algorithm and convergence

In this section, we first describe our algorithm and then we reveal the global convergence analysis of the algorithm. Now, we are at a position to give the description of our smoothing predictor-corrector algorithm.

Algorithm 3.1

Step 0. Take $\delta \in (0,1)$, $\sigma \in (0,1)$. Let $e^0 = (\mu_0, \varepsilon_0, 0) \in R_{++} \times R_{++} \times R^n$ and $x^0 \in R^n$ is an arbitrary point. Choose $z^0 = (\mu_0, \varepsilon_0, x^0)$ and parameter $\gamma \in (0,1)$ such that $\gamma\|H(z^0)\| \leq 1$, $\gamma\mu_0 + \gamma\varepsilon_0 < 1$. Set $k = 0$.

Step 1. If $\|H(z^k)\| = 0$, then stop. Otherwise, let $\beta_k = \beta(z^k)$ where $\beta(z)$ is defined by $\beta(z) = \gamma\|H(z)\|$.

Step 2. Predictor step. If $\|H(z^k)\| \geq 1$, set $\hat{z}^k := z^k$ and go to Step 3. Otherwise, compute $\Delta\hat{z}^k = (\Delta\hat{\mu}_k, \Delta\hat{\varepsilon}_k, \Delta\hat{x}^k) \in R \times R \times R^n$ by

$$H(z^k) + H'(z^k)\Delta z^k = \beta_k\|H(z^k)\|e^0. \tag{3.1}$$

If

$$\psi(z^k + \Delta\hat{z}^k) \leq \psi^2(z^k), \tag{3.2}$$

then set $\hat{z}^k = z^k + \Delta\hat{z}^k$. Otherwise, set $\hat{z}^k = z^k$.

Step 3. Corrector step. If $\|H(z^k)\| = 0$, then stop. Otherwise, compute $\Delta\tilde{z}^k = (\Delta\tilde{\mu}_k, \Delta\tilde{\varepsilon}_k, \Delta\tilde{x}^k) \in R \times R \times R^n$ by

$$H(\hat{z}^k) + H'(\hat{z}^k)\Delta z^k = \beta(\hat{z}^k)e^0. \tag{3.3}$$

Let l_k be the smallest nonnegative integer l such that

$$\psi(\hat{z}^k + \delta^l\Delta\tilde{z}^k) \leq [1 - \sigma(1 - \gamma(\mu_0 + \varepsilon_0))\delta^l]\psi(\hat{z}^k). \tag{3.4}$$

Set $\lambda_k = \delta^{l_k}$ and $z^{k+1} := \hat{z}^k + \lambda_k\Delta\tilde{z}^k$.

Step 4. Set $k := k + 1$ and return to Step 1.

Remark 3.1 If $\|H(z^k)\| \geq 1$, then Algorithm 3.1 solves only one linear system of equations at each iteration. Otherwise, it solves two linear systems of equations at each iteration. However, the coefficient matrices of these two systems are identical when (3.2) is not satisfied. There are the same points as the algorithm in [13], the neighborhood of the path does not appear in the algorithm, thus, it does not need a few additional computations which keep the iteration sequence staying in the given neighborhood.

To prove the convergence of Algorithm 3.1, first, define the set

$$\Omega = \{z = (\mu, \varepsilon, x) \in R_{++} \times R_{++} \times R^n | \mu \geq \beta(z)\mu_0, \varepsilon \geq \beta(z)\varepsilon_0\}.$$

The following lemmas show that Algorithm 3.1 is well defined and generates an infinite sequence with some good features.

Lemma 3.1 *If f is a continuously differentiable P_0-function, then Algorithm* 3.1 *is well defined. In addition,* $\mu_k > 0$, $\varepsilon_k > 0$ *and* $z^k = (\mu_k, \varepsilon_k, x^k) \in \Omega$ *for any* $k \geq 0$.

Proof Since f is a continuously differentiable P_0 function, then it follows from Theorem 2.1 that the matrix $H'(z)$ is nonsingular for $u > 0$, $\varepsilon > 0$. Since $u_0 > 0$, $\varepsilon_0 > 0$ by the choice of an initial point, we may assume, without loss of generality, that $\mu_k > 0$, $\varepsilon_k > 0$, we show that $\hat{\mu}_k > 0$, $\hat{\varepsilon}_k > 0$. If the predictor step is accepted, then by (3.1),

$$\hat{\mu}_k = \beta_k \|H(z^k)\| \mu_0, \tag{3.5}$$

$$\hat{\varepsilon}_k = \beta_k \|H(z^k)\| \varepsilon_0, \tag{3.6}$$

otherwise, we have $z^k = \hat{z}^k$, which means $\mu_k = \hat{\mu}_k$, $\varepsilon_k = \hat{\varepsilon}_k$. Thus, we obtain $\hat{\mu}_k > 0$, $\hat{\varepsilon}_k > 0$. Furthermore, $H'(z^k)$ and $H'(\hat{z}^k)$ are nonsingular which means that (3.1) and (3.3) are well defined.

Given $k \geq 0$, for any $\alpha \in (0, 1]$, let

$$R(\alpha) = H(\hat{z}^k + \alpha \Delta \tilde{z}^k) - H(\hat{z}^k) - \alpha H'(\hat{z}^k) \Delta \tilde{z}^k. \tag{3.7}$$

Noting $H(\cdot)$ is continuously differentiable, we obtain $\|R(\alpha)\| = o(\alpha)$. Then, it follows from (3.3) and (3.7) that

$$
\begin{aligned}
\|H(\hat{z}^k + \alpha \Delta \tilde{z}^k)\| &= \|H(\hat{z}^k) + \alpha H'(\hat{z}^k) \Delta \tilde{z}^k + R(\alpha)\| \\
&\leq (1-\alpha)\|H(\hat{z}^k)\| + \|R(\alpha)\| + \alpha \gamma \|H(\hat{z}^k)\| \|e^0\| \\
&\leq (1-\alpha)\|H(\hat{z}^k)\| + \alpha \gamma (\mu_0 + \varepsilon_0)\|H(\hat{z}^k)\| + o(\alpha) \\
&\leq [1 - (1 - \gamma(\mu_0 + \varepsilon_0))\alpha]\|H(\hat{z}^k)\| + o(\alpha).
\end{aligned}
\tag{3.8}
$$

Therefore, from (3.8), it shows that there exists a positive number $\bar{\alpha} \in (0, 1]$ such that for all $\alpha \in (0, \bar{\alpha}]$ and $\sigma \in (0, 1)$,

$$\|H(\hat{z}^k + \alpha \Delta \tilde{z}^k)\| \leq [1 - \sigma(1 - \gamma(\mu_0 + \varepsilon_0))\alpha]\|H(\hat{z}^k)\| \tag{3.9}$$

holds, which implies

$$\psi(\hat{z}^k + \alpha \Delta \tilde{z}^k) \leq [1 - \sigma(1 - \gamma(\mu_0 + \varepsilon_0))\alpha]\psi(\hat{z}^k).$$

That is, the nonnegative l_k satisfying (3.4) can be found, which demonstrates that (3.4) is well defined.

For $k = 0$, since $\beta(z^0) = \gamma \|H(z^0)\| \leq 1$, we know $z^0 \in \Omega$. Assuming now that $z^i \in \Omega$ is true for $i = 0, 1, \dots, k$, we show that it continues to hold for $k + 1$. If the predictor step is accepted, then it follows from (3.5), (3.6) and (3.2) that

$$
\begin{aligned}
\hat{\mu}_k &= \beta_k \|H(z^k)\| \mu_0 \\
&= \gamma \psi(z^k) \mu_0 \\
&\geq \gamma \left(\psi(z^k + \Delta\hat{z}^k)\right)^{1/2} \mu_0 \\
&= \gamma \|H(z^k + \Delta\hat{z}^k)\| \mu_0 \\
&= \beta(\hat{z}^k) \mu_0
\end{aligned}
\tag{3.10}
$$

and

$$
\begin{aligned}
\hat{\varepsilon}_k &= \beta_k \|H(z^k)\| \varepsilon_0 \\
&= \gamma \psi(z^k) \varepsilon_0 \\
&\geq \gamma \left(\psi(z^k + \Delta\hat{z}^k)\right)^{1/2} \varepsilon_0 \\
&= \gamma \|H(z^k + \Delta\hat{z}^k)\| \varepsilon_0 \\
&= \beta(\hat{z}^k) \varepsilon_0,
\end{aligned}
\tag{3.11}
$$

which implies

$$
\hat{z}^k \in \Omega.
\tag{3.12}
$$

Otherwise, from $z^k = \hat{z}^k$ and the inductive assumption, we obtain that (3.12) also holds. Noting (3.3), we have

$$
\mu_{k+1} = \hat{\mu}_k + \lambda_k \Delta\tilde{\mu}_k = (1 - \lambda_k)\hat{\mu}_k + \lambda_k \beta(\hat{z}_k)\mu_0,
\tag{3.13}
$$

$$
\varepsilon_{k+1} = \hat{\varepsilon}_k + \lambda_k \Delta\tilde{\varepsilon}_k = (1 - \lambda_k)\hat{\varepsilon}_k + \lambda_k \beta(\hat{z}^k)\varepsilon_0.
\tag{3.14}
$$

In addition, from (3.9) we know that there exists $\lambda_k \in (0, 1)$ such that

$$
\|H(\hat{z}^k + \lambda_k \Delta\tilde{z}^k)\| \leq \left[1 - \sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\lambda_k\right] \|H(\hat{z}^k)\| \leq \|H(\hat{z}^k)\|.
\tag{3.15}
$$

Therefore, it follows from (3.12), (3.13) and (3.15) that

$$
\begin{aligned}
\mu_{k+1} - \beta_{k+1}\mu_0 &= (1 - \lambda_k)\hat{\mu}_k + \lambda_k \beta(\hat{z}^k)\mu_0 - \gamma \|H(z^{k+1})\| \mu_0 \\
&\geq (1 - \lambda_k)\beta(\hat{z}^k)\mu_0 + \lambda_k \beta(\hat{z}^k)\mu_0 - \gamma \|H(z^{k+1})\| \mu_0 \\
&= \beta(\hat{z}^k)\mu_0 - \gamma \|H(z^{k+1})\| \mu_0 \\
&= \gamma \|H(\hat{z}^k)\| \mu_0 - \gamma \|H(z^{k+1})\| \mu_0 \geq 0.
\end{aligned}
\tag{3.16}
$$

Similarly, we can obtain $\varepsilon_{k+1} - \beta_{k+1}\varepsilon_0 \geq 0$. Thus, $z^{k+1} \in \Omega$.

Since $u_0 > 0$, $\varepsilon_0 > 0$, we may assume that $\mu_k > 0$, $\varepsilon_k > 0$ for any given $k \geq 0$. From $\hat{\mu}_k > 0$, $\hat{\varepsilon}_k > 0$, it follows from (3.13) that $\mu_{k+1} > 0$, $\varepsilon_{k+1} > 0$. Hence, $\mu_k > 0$, $\varepsilon_k > 0$ for any $k \geq 0$. □

Lemma 3.2 *Suppose that the infinite sequence $\{z^k = (\mu_k, \varepsilon_k, x^k)\}$ is generated by Algorithm* 3.1, *then* $0 < \mu_{k+1} \leq \mu_k, 0 < \varepsilon_{k+1} \leq \varepsilon_k$ *and the sequence* $\{\|H(z^k)\|\}$ *is monotonically decreasing.*

Proof For any $k \geq 0$, it follows from (3.12), (3.13) and (3.14) that

$$
\begin{aligned}
\mu_{k+1} &= (1 - \lambda_k)\hat{\mu}_k + \lambda_k \beta(\hat{z}^k)\mu_0 \\
&\leq (1 - \lambda_k)\hat{\mu}_k + \lambda_k \hat{\mu}_k \\
&= \hat{\mu}_k,
\end{aligned}
\tag{3.17}
$$

$$
\begin{aligned}
\varepsilon_{k+1} &= (1 - \lambda_k)\hat{\varepsilon}_k + \lambda_k \beta(\hat{z}^k)\varepsilon_0 \\
&\leq (1 - \lambda_k)\hat{\varepsilon}_k + \lambda_k \hat{\varepsilon}_k \\
&= \hat{\varepsilon}_k.
\end{aligned}
\tag{3.18}
$$

If the predictor step (Step 2) is not accepted at the k-th iterate, then (3.17) and (3.18) show the desired result. Otherwise, from (3.5), (3.6), $\|H(z^k)\| < 1$ and $z^k \in \Omega$, one has

$$
\hat{\mu}_k = \beta_k \|H(z^k)\| \mu_0 \leq \beta_k \mu_0 \leq \mu_k,
\tag{3.19}
$$

$$
\hat{\varepsilon}_k = \beta_k \|H(z^k)\| \varepsilon_0 \leq \beta_k \varepsilon_0 \leq \varepsilon_k.
\tag{3.20}
$$

Thus, we obtain that $\mu_{k+1} \leq \mu_k$, $\varepsilon_{k+1} \leq \varepsilon_k$ hold for any $k \geq 0$.

If the predictor step (Step 2) is not accepted at the k-th iterate, then (3.15) implies that

$$
\|H(z^{k+1})\| \leq \|H(\hat{z}^k)\| = \|H(z^k)\|,
$$

and the desired result has been obtained. Otherwise, it follows from (3.2) and $\|H(z^k)\| < 1$ that

$$
\|H(\hat{z}^k)\| \leq \|H(z^k)\|^2 \leq \|H(z^k)\|.
$$

Hence, for any $k \geq 0$, we obtain

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

which means the sequence $\{\|H(z^k)\|\}$ is monotonically decreasing. □

Lemma 3.3 *Assume that f is a P_0-function and μ_1, μ_2, ε_1, ε_2 are given positive numbers satisfying $\mu_1 < \mu_2$, $\varepsilon_1 < \varepsilon_2$. Then, H defined by* (2.3) *has the property*

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

for any sequence $\{(\mu_k, \varepsilon_k, x^k)\}$ such that $\mu_1 \leq \mu_k \leq \mu_2$, $\varepsilon_1 \leq \varepsilon_k \leq \varepsilon_2$ for any k and $\|x^k\| \to +\infty$ as $k \to +\infty$.

Proof We outline the proof by contradiction. Suppose that the lemma is not true. Then there exists a sequence $\{z^k = (\mu_k, \varepsilon_k, x^k)\}$ such that $\mu_1 \leq \mu_k \leq \mu_2$, $\varepsilon_1 \leq \varepsilon_k \leq \varepsilon_2$, $\psi(z^k) \leq \psi(z^0)$ but $\|x^k\| \to \infty$. Since the sequence $\{x^k\}$ is unbounded, the index set $I = \{i \in N : \{x_i^k\}$ is unbounded$\}$ is nonempty. Without loss of generality, we can assume that $\{|x_i^k|\} \to +\infty$ for all $i \in I$. Then the following sequence $\{\bar{x}^k\}$ is bounded which is defined by

$$\bar{x}^k = \begin{cases} 0, & i \in I, \\ x_i^k, & i \notin I. \end{cases}$$

Since f is a P_0-function, by Definition 2.2, we have

$$\begin{aligned} 0 &\leq \max_{i \in N} \left[(x_i^k - \bar{x}_i^k)(f_i(x^k) - f_i(\bar{x}^k)) \right] \\ &= \max \left\{ 0, \max_{i \in I} \left[x_i^k (f_i(x^k) - f_i(\bar{x}^k)) \right] \right\} \\ &= x_{i_0}^k \left[f_{i_0}(x^k) - f_{i_0}(\bar{x}^k) \right], \end{aligned} \tag{3.21}$$

where i_0 is one of the indices for which the max is attained, and i_0 is assumed, without loss of generality, to be independent of k. Since $i_0 \in I$, one has $\{|x_{i_0}^k|\} \to +\infty$ as $k \to \infty$. We now break up the proof into two cases.

Case 1. If $x_{i_0}^k \to +\infty$ as $k \to \infty$. In this case, since $f_{i_0}(\bar{x}^k)$ is bounded, we deduce from (3.21) that $f_{i_0}(x^k) > f_{i_0}(\bar{x}^k)$.

If $f_{i_0}(\bar{x}^k) < f_{i_0}(x^k) < +\infty$, for $0 < \mu_1 \leq \mu_k \leq \mu_2$, $\varepsilon_1 \leq \varepsilon_k \leq \varepsilon_2$, letting $k \to \infty$ yields

$$\frac{1}{2} \left(f_{i_0}(x^k) + \mu_k \left(\ln 2 + \ln \left(1 + \cosh \frac{f_{i_0}(x^k)}{\mu_k} \right) \right) \right)$$

is bounded and

$$\frac{1}{2} \left(f_{i_0}(x^k) + \mu_k \left(\ln 2 + \ln \left(1 + \cosh \frac{f_{i_0}(x^k)}{\mu_k} \right) \right) \right) + \varepsilon_k x_{i_0}^k \to +\infty.$$

Thus, $\|\Phi(z^k)\| \to +\infty$ as $k \to \infty$.

If $f_{i_0}(x^k) \to +\infty$, for $0 < \mu_1 \leq \mu_k \leq \mu_2$, $\varepsilon_1 \leq \varepsilon_k \leq \varepsilon_2$, we have

$$\frac{1}{2} \left(f_{i_0}(x^k) + \mu_k \left(\ln 2 + \ln \left(1 + \cosh \frac{f_{i_0}(x^k)}{\mu_k} \right) \right) \right) \to +\infty$$

and

$$\frac{1}{2} \left(f_{i_0}(x^k) + \mu_k \left(\ln 2 + \ln \left(1 + \cosh \frac{f_{i_0}(x^k)}{\mu_k} \right) \right) \right) + \varepsilon_k x_{i_0}^k \to +\infty.$$

Thus, $\|\Phi(z^k)\| \to +\infty$ as $k \to \infty$.

Case 2. $x_{i_0}^k \to -\infty$ as $k \to \infty$. In this case, since $f_{i_0}(\bar{x}^k)$ is bounded, we deduce from (3.21) that $f_{i_0}(x^k) < f_{i_0}(\bar{x}^k)$.

If $-\infty < f_{i_0}(x^k) < f_{i_0}(\bar{x}^k)$, for $0 < \mu_1 \leq \mu_k \leq \mu_2$, $\varepsilon_1 \leq \varepsilon_k \leq \varepsilon_2$, we have

$$\frac{1}{2} \left(f_{i_0}(x^k) + \mu_k \left(\ln 2 + \ln \left(1 + \cosh \frac{f_{i_0}(x^k)}{\mu_k} \right) \right) \right)$$

is bounded and

$$\frac{1}{2}\left(f_{i_0}\left(x^k\right) + \mu_k\left(\ln 2 + \ln\left(1 + \cosh\frac{f_{i_0}(x^k)}{\mu_k}\right)\right)\right) + \varepsilon_k x_{i_0}^k \to -\infty.$$

Thus, $\|\Phi(z^k)\| \to +\infty$ as $k \to \infty$.

If $f_{i_0}(x^k) \to -\infty$, for $0 < \mu_1 \le \mu_k \le \mu_2$, $\varepsilon_1 \le \varepsilon_k \le \varepsilon_2$, we have

$$\frac{1}{2}\left(f_{i_0}\left(x^k\right) + \mu_k\left(\ln 2 + \ln\left(1 + \cosh\frac{f_{i_0}(x^k)}{\mu_k}\right)\right)\right)$$

$$= \frac{1}{2}\left(f_{i_0}\left(x^k\right) + \mu_k\left(\ln\left(e^{-\frac{f_{i_0}(x^k)}{\mu_k}}\left(2e^{\frac{f_{i_0}(x^k)}{\mu_k}} + 1 + e^{\frac{2f_{i_0}(x^k)}{\mu_k}}\right)\right)\right)\right)$$

$$= \frac{1}{2}\mu_k\left(\ln\left(2e^{\frac{f_{i_0}(x^k)}{\mu_k}} + 1 + e^{\frac{2f_{i_0}(x^k)}{\mu_k}}\right)\right)$$

is bounded and

$$\frac{1}{2}\left(f_{i_0}\left(x^k\right) + \mu_k\left(\ln 2 + \ln\left(1 + \cosh\frac{f_{i_0}(x^k)}{\mu_k}\right)\right)\right) + \varepsilon_k x_{i_0}^k \to -\infty.$$

Thus, $\|\Phi(z^k)\| \to +\infty$ as $k \to \infty$.

In summary, we obtain $\psi(z^k) \to +\infty$ as $k \to \infty$, which contradicts $\psi(z^k) \le \psi(z^0)$, and the proof is completed. $\qquad\square$

Under the assumption of f being a P_0-function, Lemma 3.2 and Lemma 3.3 indicate that the level set $L_\mu(z^0)$ defined by

$$L_\mu\left(z^0\right) = \left\{z \in R_{++} \times R_{++} \times R^n | \psi(z) \le \psi\left(z^0\right)\right\} \tag{3.22}$$

is bounded.

To obtain the global convergence of Algorithm 3.1, we need the following assumption.

Assumption 3.1 The solution $S := \{x \in R^n, f(x) \le 0\}$ of (1.1) is nonempty and bounded.

Note that Assumption 3.1 seems to be the weakest condition used in the previous literature to ensure the bound of iteration sequences (see [23]).

Theorem 3.1 *Assume that the infinite sequence $\{z^k\}$ is generated by Algorithm* 3.1. *Then*
  (a) *The sequences $\{\|H(z^k)\|\}$, $\{\mu_k\}$ and $\{\varepsilon_k\}$ converge to zero as $k \to +\infty$, and hence any accumulation point of $\{x^k\}$ is a solution of* (1.1).
  (b) *If Assumption* 4.1 *is satisfied, then the sequence $\{z^k\}$ is bounded, hence there exists at least one accumulation point $z^* = (\mu_*, \varepsilon_*, x^*)$ with $H(z^*) = 0$ and $x^* \in S$.*

Proof By Lemma 3.2, 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 3.2 and Lemma 3.3. Assume that $z^* = (\mu_*, \varepsilon_*, x^*)$ is an accumulation point of $\{z^k\}$. Without loss of generality, we assume that $\{z^k\}$ converges to z^*. Then, by the continuity of H and the

definition of $\beta(\cdot)$, we know that $\{\mu_k\}$, $\{\varepsilon_k\}$ and $\{\beta_k\}$ converge to μ_*, ε_*, β_* respectively and that $h^* = \|H(z^*)\| > 0$. Therefore, by (3.4), we have

$$\lim_{k \to \infty} \lambda^k = 0.$$

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

$$\psi\left(\hat{z}^k + \delta^{l_k - 1} \Delta \tilde{z}^k\right) \geq \left[1 - \sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\delta^{l_k - 1}\right]\psi\left(\hat{z}^k\right),$$

which implies that

$$\frac{\psi\left(\hat{z}^k + \delta^{l_k - 1} \Delta \tilde{z}^k\right) - \psi\left(\hat{z}^k\right)}{\delta^{l_k - 1}} \geq -\sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\psi\left(\hat{z}^k\right). \tag{3.23}$$

Letting $k \to +\infty$, we have

$$H\left(z^*\right)^T H\left(z^*\right)\Delta z^* \geq -\sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\psi\left(\hat{z}^*\right). \tag{3.24}$$

On the other hand, by (3.3), we have

$$H\left(z^*\right) + H'\left(z^*\right)\Delta z = \beta\left(z^*\right)e^0,$$

i.e.,

$$H\left(z^*\right)^T H\left(z^*\right)\Delta z^* = -\psi\left(\hat{z}^*\right) + \beta\left(z^*\right)H\left(z^*\right)^T e^0. \tag{3.25}$$

Combining (3.24) and (3.25), we deduce that

$$\left[1 - \sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\right]\psi\left(z^*\right) \leq \beta\left(z^*\right)H\left(z^*\right)^T e^0 \leq \beta\left(z^*\right)\sqrt{\mu_0^2 + \varepsilon_0^2}\left\|H\left(z^*\right)\right\|,$$

which means

$$\left[1 - \sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right)\right]\psi\left(z^*\right) \leq \beta\left(z^*\right)\sqrt{\mu_0^2 + \varepsilon_0^2}\left\|H\left(z^*\right)\right\|$$
$$\leq \gamma(\mu_0 + \varepsilon_0)\psi\left(z^*\right). \tag{3.26}$$

Since $\|H(z^*)\| > 0$, then

$$1 - \sigma\left(1 - \gamma(\mu_0 + \varepsilon_0)\right) \leq \gamma(\mu_0 + \varepsilon_0),$$

i.e.,

$$(1 - \sigma)\left(1 - \gamma(\mu_0 + \varepsilon_0)\right) \leq 0.$$

This contradicts the fact that $\sigma \in (0, 1)$ and $\gamma(\mu_0 + \varepsilon_0) < 1$. Hence, we have $h^* = 0$, $\mu_* = 0$, $\varepsilon_* = 0$. Thus, $H(z^*) = 0$, that is, x^* is a solution of (1.1).

Next we prove (b). It follows from (a) that $H(z^k) \to 0$ as $k \to \infty$. By (2.3), one has

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

Therefore, by the famous mountain pass theorem (Theorem 5.4 in [24]) and along the lines of the proof of Theorem 3.1 in [23], we obtain that $\{x^k\}$ is bounded and hence $\{z^k\}$ is. Thus, $\{z^k\}$ has at least one accumulation point $z^* = (\mu_*, \varepsilon_*, x^*)$. By (a), we have $H(z^*) = 0$ and $\mu_* = 0$, $\varepsilon_* = 0$, $x^* \in S$. $\qquad\square$

Next, we show the local superlinear convergence of Algorithm 3.1.

Theorem 3.2 *Suppose that f is a continuously differentiable P_0-function, Assumption 3.1 is satisfied and z^* is an accumulation point of the iteration sequence $\{z^k\}$ generated by Algorithm 3.1. If all $V \in \partial H(z^*)$ are nonsingular and $f'(x)$ is locally Lipschitz continuous around x^*, then the whole sequence $\{z^k\}$ superlinearly converges to z^*, i.e.,*

$$\|z^{k+1} - z^*\| = o(\|z^k - z^*\|)$$

and

$$\mu_{k+1} = o(\mu_k), \qquad \varepsilon_{k+1} = o(\varepsilon_k).$$

Proof First, from Theorem 3.1, we know that z^* is a solution of $H(z) = 0$. Then since all $V \in \partial H(z^*)$ are nonsingular, it follows from [19, 25, 27] that for all z^k sufficiently close to z^*, we have

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

where $C > 0$ is a constant.

Then, since $H(z)$ is semi-smooth at z^*, $H(z)$ is locally Lipschitz continuous near z^*, for all z^k sufficiently close to z^*,

$$\psi(z^k) = \|H(z^k)\|^2 = \|H(z^k) - H(z^*)\|^2 = O(\|z^k - z^*\|^2). \tag{3.27}$$

For all z^k sufficiently close to z^*, we have

$$\begin{aligned}
&\|z^k + \Delta z^k - z^*\| \\
&= \|z^k + H'(z^k)^{-1}[-H(z^k) + \beta_k \|H(z^k)\| e^0] - z^*\| \\
&= \|H'(z^k)^{-1}\| \|H'(z^k)(z^k - z^*) - H(z^k) + \beta_k \|H(z^k)\| e^0\| \\
&\le C[\|H'(z^k)(z^k - z^*) - H(z^k)\| + \beta_k \|H(z^k)\| \|e^0\|] \\
&= C[\|H(z^k) - H(z^*) - H'(z^k)(z^k - z^*)\| + \beta_k \|H(z^k)\| \|e^0\|] \\
&= o(\|z^k - z^*\|). \tag{3.28}
\end{aligned}$$

Thus, for z^k sufficiently close to z^*, we obtain

$$
\begin{aligned}
\psi\left(z^k + \Delta z^k\right) &= \left\|H\left(z^k + \Delta z^k\right)\right\|^2 \\
&= O\left(\left\|z^k - \Delta z^k - z^*\right\|^2\right) \\
&= o\left(\left\|z^k - z^*\right\|^2\right) \\
&= o\left(\left\|H\left(z^k\right) - H\left(z^*\right)\right\|^2\right) \\
&= o\left(\psi\left(z^k\right)\right).
\end{aligned}
\tag{3.29}
$$

Hence, for z^k sufficiently close to z^*, we have $z^{k+1} = z^k + \Delta z^k$. By (3.28), we prove that

$$
\left\|z^{k+1} - z^*\right\| = o\left(\left\|z^k - z^*\right\|\right)
$$

holds.

Next, when k is sufficiently large, then $z^{k+1} = z^k + \Delta z^k$, so

$$
\mu_{k+1} = \mu_k + \Delta\mu_k = \beta_k \left\|H\left(z^k\right)\right\| \mu_0,
$$

and

$$
\varepsilon_{k+1} = \varepsilon_k + \Delta\varepsilon_k = \beta_k \left\|H\left(z^k\right)\right\| \varepsilon_0.
$$

Hence, for all k sufficiently large,

$$
\mu_{k+1} = \gamma\mu_0 \left\|H\left(z^k\right)\right\|^2, \qquad \varepsilon_{k+1} = \gamma\varepsilon_0 \left\|H\left(z^k\right)\right\|^2,
$$

which, together with (3.29), yields

$$
\lim_{k\to+\infty} \frac{\mu_{k+1}}{\mu_k} = \lim_{k\to+\infty} \frac{\|H(z^k)\|^2}{\|H(z^{k-1})\|^2} = 0,
$$

and

$$
\lim_{k\to+\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \lim_{k\to+\infty} \frac{\|H(z^k)\|^2}{\|H(z^{k-1})\|^2} = 0.
$$

This means that $\mu_{k+1} = o(\mu_k)$, $\varepsilon_{k+1} = o(\varepsilon_k)$ and the desired result follows. $\qquad\square$

4 Numerical experiments

In this section, we test our algorithm for solving the systems of inequalities. In our implementation, we adopt the strategy in [8], the function H defined by (2.3) is replaced by

$$
H(z) = H(\mu, \varepsilon, x) = \begin{pmatrix} \mu \\ \varepsilon \\ \phi(\mu, f(x)) + c\varepsilon x \end{pmatrix},
\tag{4.1}
$$

where c is a constant. It is easy to see that such a change does not destroy any theoretical results obtained in Section 3.

Table 1 Numerical results of Example 4.1

st	Our proposed algorithm						Smoothing algorithm [8]		
	c	ε_0	ic	ip	sol	cpu	iter	sol	cpu
$(0,0)^\top$	10	0.5	1	2	$(0.0424, -1.0101)^\top$	0.14	23	$(0.0592, -0.9961)^\top$	1.1
$(1,-1)^\top$	10	0.4	1	2	$(0.5090, -0.8655)^\top$	0.18	26	$(0.3440, -0.9392)^\top$	1.3
$(1,-1)^\top$	10	0.8	1	2	$(0.6132, -0.7744)^\top$	0.18	28	$(0.2568, -0.9667)^\top$	1.4

Table 2 Numerical results of Example 4.2

st	Our proposed algorithm						Smoothing algorithm [8]		
	c	ε_0	ic	ip	sol	cpu	iter	sol	cpu
$(0,0)^\top$	0.5	1	2	1	$(0.0442, 0.3356)^\top$	0.12	fail	fail	fail
$(0,0)^\top$	0.5	0.5	2	1	$(-0.0105, 0.9541)^\top$	0.15	21	$(-0.0000, 1.2045)^\top$	0.87
$(1,1)^\top$	0.5	0.1	2	1	$(-0.0019, 1.5663)^\top$	0.16	21	$(0.0004, 1.5704)^\top$	0.87
$(1,1)^\top$	0.5	1	2	1	$(-0.0206, 0.8605)^\top$	0.17	22	$(0.0006, 1.5698)^\top$	0.90

Table 3 Numerical results of Example 4.3

st	Our proposed algorithm						Smoothing algorithm [8]		
	c	ε_0	ic	ip	sol	cpu	iter	sol	cpu
$(0,1)^\top$	10	0.5	1	2	$(0.1237, 0.5902)^\top$	0.17	fail	fail	fail
$(1,1)^\top$	10	1	1	2	$(0.0590, 0.5236)^\top$	0.16	23	$(0.5023, 0.5153)^\top$	1.26
$(1,1)^\top$	1	1	1	2	$(0.4533, 0.4973)^\top$	0.20	23	$(0.5274, 0.5080)^\top$	1.22

In our numerical experiments, the parameters used in the algorithm are chosen as follows: $\sigma = 0.06$, $\delta = 0.3$, $\mu_0 = 1$, $\gamma = 0.01 \min\{1, 1/\|H(z^0)\|\}$. The algorithm terminates when $\|\psi(z^k)\| \leq 10^{-3}$. In the tables of test results, *st* denotes the starting point of x^0, *ic* denotes the corrector iteration numbers in Step 3 followed directly from Step 1, *ip* denotes the predictor iteration numbers, *iter* denotes the iteration numbers of smoothing method (in [8]), *cpu* denotes the CPU time for solving the underlying problems in seconds, and *sol* denotes a solution of the test problem. In the following, we reveal a detailed description of the tested problems.

In the following, we reveal a detailed description of the tested problems. For Example 4.1, 4.2 and 4.3, we compare the results obtained by our method with which obtained by smoothing method [8]. The results are summarized in Table 1, Table 2 and Table 3.

Example 4.1 [9, 10] Consider (1.1), where $f = (f_1, f_2)^\top$ with $x \in R^2$ and

$$f_1(x) = x_1^2 + x_2^2 - 1, \qquad f_2(x) = -x_1^2 - x_2^2 + (0.999)^2.$$

Example 4.2 [9, 10] Consider (1.1), where $f = (f_1, f_2)^\top$ with $x \in R^2$ and

$$f_1(x) = \sin(x_1), \qquad f_2(x) = -\cos(x_2).$$

Example 4.3 [26] Consider (1.1), where $f = (f_1, f_2)^\top$ with $x \in R^2$ and

$$f_1(x) = x_1 - 0.7 \sin(x_1) - 0.2 \cos(x_2), \qquad f_2(x) = x_2 - 0.7 \cos(x_1) + 0.2 \sin(x_2).$$

5 Conclusion

In this paper, we present a new smoothing and regularization predictor-corrector algorithm to solve the nonlinear inequalities, the global and local convergence are obtained. Furthermore, the smoothing parameter μ and the regularization parameter ε in our algorithm are viewed as independent variables. Preliminary numerical results show the efficiency of the algorithm.

Authors' contributions
The author carried out the proof. The author conceived of the study and participated in its design and coordination. The author read and approved the final manuscript.

References
 1. Zakian, V, Nail, UA: Design of dynamical and control systems by the method of inequalities. Proc. Inst. Electr. Eng. **120**, 1421-1472 (1973)
 2. Dennis, JE, Schnabel, RB: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs (1983)
 3. Neumaier, A: Interval Methods for System of Equations. Cambridge University Press, Cambridge (1990)
 4. Daniel, JW: Newtons method for nonlinear inequalities. Numer. Math. **21**, 381-387 (1973)
 5. Mayne, DQ, Polak, E, Heunis, AJ: Solving nonlinear inequalities in a finite number of iterations. J. Optim. Theory Appl. **33**, 207-221 (1981)
 6. Sahba, M: On the solution of nonlinear inequalities in a finite number of iterations. Numer. Math. **46**, 229-236 (1985)
 7. Yin, HX, Huang, ZH, Qi, L: The convergence of a Levenberg-Marquardt method for nonlinear inequalities. Numer. Funct. Anal. Optim. **29**, 687-716 (2008)
 8. Huang, ZH, Zhang, Y, Wu, W: A smoothing-type algorithm for solving system of inequalities. J. Comput. Appl. Math. **220**, 355-363 (2008)
 9. Zhu, JG, Liu, HW, Li, XL: A regularized smoothing-type algorithm for solving a system of inequalities with a P_0-function. J. Comput. Appl. Math. **233**, 2611-2619 (2010)
10. He, C, Ma, CF: A smoothing self-adaptive Levenberg-Marquardt algorithm for solving system of nonlinear inequalities. Appl. Math. Comput. **216**, 3056-3063 (2010)
11. Huang, ZH, Qi, L, Sun, D: Sub-quadratic convergence of a smoothing Newton algorithm for the P_0 and monotone LCP. Math. Program. **99**, 423-441 (2004)
12. Zhao, N, Huang, ZH: A nonmonotone smoothing Newton algorithm for solving box constrained variational inequalities with a P_0 function. J. Ind. Manag. Optim. **7**(2), 467-482 (2011)
13. Huang, ZH, Han, J, Chen, Z: Predictor-corrector smoothing Newton method, based on a new smoothing function, for solving the nonlinear complementarity with a P_0 function. J. Optim. Theory Appl. **117**, 39-68 (2003)
14. Herty, M, Klar, A, Singh, AK, Spellucci, P: Smoothed penalty algorithms for optimization of nonlinear models. Comput. Optim. Appl. **37**, 157-176 (2007)
15. Burke, J, Xu, S: A noninterior predictor-corrector path-following algorithm for the monotone linear complementarity problem. Math. Program. **87**, 113-130 (2000)
16. Luca, TD, Facchinei, F, Kanzow, C: A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407-439 (1996)
17. Clarke, FH: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
18. Mifflin, R: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**(6), 957-972 (1977)
19. Qi, L, Sun, J: A nonsmooth version of Newton's method. Math. Program. **58**(3), 353-367 (1993)
20. Lee, YJ, Mangasarin, OL: SSVM: A smooth support vector machine for classification. Comput. Optim. Appl. **20**, 5-22 (2001)
21. Che, H, Li, M: A smoothing and regularization Broyden-like method for nonlinear inequalities. J. Appl. Math. Comput. (2012). doi:10.1007/s12190-012-0588-2
22. Chen, B, Harker, PT: A non-interior-point continuation method for linear complementarity problems. SIAM J. Matrix Anal. Appl. **14**, 1168-1190 (1993)
23. Huang, ZH, Han, J, Xu, D, Zhang, L: The non-linear continuation methods for solving the P0-functions non-linear complementarity problem. Sci. China **44**(2), 1107-1114 (2001)
24. Facchinei, F, Kanzow, C: Beyond monotonicity in regularization methods for nonlinear complementarity problems. SIAM J. Control Optim. **37**(2), 1150-1161 (1999)
25. Qi, L: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227-244 (1993)

26. Zhang, Y, Huang, ZH: A nonmonotone smoothing-type algorithm for solving a system of equalities and inequalities. J. Comput. Appl. Math. **233**, 2312-2321 (2010)
27. Qi, L, Sun, DF, Zhou, GL: A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities. Math. Program., Ser. A **87**, 1-35 (2000)