Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 220 (2008) 464–479
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

A regularization semismooth Newton method based on the generalized Fischer–Burmeister function for P_0-NCPs

Jein-Shan Chen[a,*,1], Shaohua Pan[b]

[a]*Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan*
[b]*School of Mathematical Sciences, South China University of Technology, Guangzhou 510641, China*

Received 1 June 2007; received in revised form 29 August 2007

Abstract

We consider a regularization method for nonlinear complementarity problems with F being a P_0-function which replaces the original problem with a sequence of the regularized complementarity problems. In this paper, this sequence of regularized complementarity problems are solved approximately by applying the generalized Newton method for an equivalent augmented system of equations, constructed by the generalized Fischer–Burmeister (FB) NCP-functions ϕ_p with $p > 1$. We test the performance of the regularization semismooth Newton method based on the family of NCP-functions through solving all test problems from MCPLIB. Numerical experiments indicate that the method associated with a smaller p, for example $p \in [1.1, 2]$, usually has better numerical performance, and the generalized FB functions ϕ_p with $p \in [1.1, 2)$ can be used as the substitutions for the FB function ϕ_2.

Keywords: Nonlinear complementarity problem (NCP); Generalized Fischer–Burmeister function; P_0-function; Semismooth Newton method

1. Introduction

The nonlinear complementarity problem (NCP) is to find a point $x \in \mathbb{R}^n$ such that

$$x \geqslant 0, \quad F(x) \geqslant 0, \quad \langle x, F(x)\rangle = 0, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F = (F_1, F_2, \ldots, F_n)^{\mathrm{T}}$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$. We assume that F is continuously differentiable throughout this paper. The NCP has attracted much attention due to its various applications in operations research, economics, and engineering [12,17,24]. There have been many methods proposed for solving the NCP, including merit function approaches [16,21,23,33], nonsmooth Newton methods [11,22,34], smoothing methods [5,18,27,32] and regularization methods [9,19,29,30]. All the aforementioned methods usually exploit so-called NCP-functions defined as below.

Definition 1.1. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an NCP-function (or C-function standing for Complementarity function, see [10]) if it satisfies

$$\phi(a, b) = 0 \quad \Longleftrightarrow \quad a \geqslant 0, \quad b \geqslant 0, \quad ab = 0. \tag{2}$$

* Corresponding author.
E-mail addresses: jschen@math.ntnu.edu.tw (J. Chen), shhpan@scut.edu.cn (S. Pan).
[1] Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office.

0377-0427/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2007.08.020

Over the past two decades, a variety of NCP-functions has been studied; see [15,31] and references therein. Among which, a popular NCP-function intensively studied is the well-known Fischer–Burmeister (FB) NCP-function [13,14] defined as

$$\phi_{\mathrm{FB}}(a, b) = \sqrt{a^2 + b^2} - (a + b). \tag{3}$$

Since ϕ_{FB} satisfies (2), the NCP is equivalent to a system of nonsmooth equations

$$\Phi_{\mathrm{FB}}(x) := \begin{pmatrix} \phi_{\mathrm{FB}}(x_1, F_1(x)) \\ \cdot \\ \cdot \\ \cdot \\ \phi_{\mathrm{FB}}(x_n, F_n(x)) \end{pmatrix} = 0. \tag{4}$$

Then we have the merit function $\Psi_{\mathrm{FB}} : \mathbb{R}^n \to \mathbb{R}_+$ for the NCP, defined by

$$\Psi_{\mathrm{FB}}(x) := \frac{1}{2}\|\Phi_{\mathrm{FB}}(x)\|^2 = \frac{1}{2}\sum_{i=1}^{n} \phi_{\mathrm{FB}}(x_i, F_i(x))^2. \tag{5}$$

Recently, a family of new NCP-functions based on the FB function (3) were studied in [2,6]. In particular, they define $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ by

$$\phi_p(a, b) := \|(a, b)\|_p - (a + b), \tag{6}$$

where p is any fixed real number in the interval $(1, +\infty)$ and $\|(a, b)\|_p$ denotes the p-norm of (a, b), namely, $\|(a, b)\|_p = \sqrt[p]{|a|^p + |b|^p}$. In other words, in the function ϕ_p, we replace the Euclidean norm of (a, b) in the FB function (3) by a more general p-norm with $p \in (1, +\infty)$. Similarly, the NCP is equivalent to the nonsmooth system

$$\Phi_p(x) := \begin{pmatrix} \phi_p(x_1, F_1(x)) \\ \cdot \\ \cdot \\ \cdot \\ \phi_p(x_n, F_n(x)) \end{pmatrix} = 0, \tag{7}$$

which induces a family of merit functions $\Psi_p : \mathbb{R}^n \to \mathbb{R}$ for the NCP as below

$$\Psi_p(x) := \frac{1}{2}\|\Phi_p(x)\|^2 = \frac{1}{2}\sum_{i=1}^{n} \phi_p(x_i, F_i(x))^2. \tag{8}$$

As seen in [6], the merit function Ψ_p for any given $p > 1$ enjoys all favorable properties as the FB merit function Ψ_{FB} holds. Moreover, numerical experiments there indicate that the descent method based on the merit function Ψ_p has better performance when p decreases in $(1, +\infty)$. However, it is still unknown whether such phenomenon occurs in other approaches for the NCP. The main purpose of this paper is to investigate how the generalized FB NCP-functions ϕ_p with $p \in (1, +\infty)$ behave in a regularization semismooth Newton method for solving the NCP.

It is well known that the regularization approach is designed to handle ill-posed problems which substitutes the solution of original problem with the solution of a sequence of well-posed problems whose solutions converging to the solution of the original problem; see [4,3,9,19,30] and references therein. In the context of complementarity problems, if we consider the so-called *Tikhonov regularization*, this scheme consists of solving a sequence of complementarity problems NCP(F_ε):

$$x \geqslant 0, \quad F_\varepsilon(x) \geqslant 0, \quad \langle x, F_\varepsilon(x)\rangle = 0, \tag{9}$$

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

$$F_\varepsilon(x) := F(x) + \varepsilon x. \tag{10}$$

Let $F_{\varepsilon,i}(x)$ denote the ith component of $F_\varepsilon(x)$ and define the map $\Phi_{p,\varepsilon} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\Phi_{p,\varepsilon}(x) := \begin{pmatrix} \phi_p(x_1, F_{\varepsilon,1}(x)) \\ \cdot \\ \cdot \\ \cdot \\ \phi_p(x_n, F_{\varepsilon,n}(x)) \end{pmatrix}. \tag{11}$$

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

$$\Phi_{p,\varepsilon}(x) = 0,$$

which leads to a merit function $\Psi_{p,\varepsilon} : \mathbb{R}^n \to \mathbb{R}_+$ for the NCP(F_ε):

$$\Psi_{p,\varepsilon}(x) := \frac{1}{2}\|\Phi_{p,\varepsilon}(x)\|^2 = \frac{1}{2}\sum_{i=1}^{n} \phi_p(x_i, F_{\varepsilon,i}(x))^2. \tag{12}$$

Therefore, the original NCP is actually equivalent to solving a sequence of nonsmooth systems of equations $\Phi_{p,\varepsilon}(x) = 0$ with ε approaching to 0. From this, we see that the parameter ε plays the same role as the smoothing parameter in smoothing methods for the NCP, except that ε is imposed on the mapping F instead of the NCP-function ϕ_p.

In this paper, the sequence of subproblems $\Phi_{p,\varepsilon}(x) = 0$ with ε tending to 0 will be solved approximately by applying the generalized Newton method for an augmented system of equations equivalent to the NCP. Specifically, we let $z := (\varepsilon, x) \in \mathbb{R}_+ \times \mathbb{R}^n$ by viewing ε as a variable, and define the mapping $H_p : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$H_p(z) := \begin{bmatrix} \varepsilon \\ \phi_p(x_1, F_{\varepsilon,1}(x)) \\ \vdots \\ \phi_p(x_n, F_{\varepsilon,n}(x)) \end{bmatrix}. \tag{13}$$

Notice that if the function $\Phi_{p,\varepsilon}(x)$ defined by (11) is viewed as a function of ε and x, then we may denote it as $\Phi_p(z) := \Phi_p(\varepsilon, x) = \Phi_{p,\varepsilon}(x)$. Hence, (13) is the same as

$$H_p(z) = \begin{bmatrix} \varepsilon \\ \Phi_p(z) \end{bmatrix}.$$

It is easily verified that the NCP is equivalent to the augmented system of equations

$$H_p(z) = H_p(\varepsilon, x) = 0, \tag{14}$$

which naturally induces a merit function $G_p : \mathbb{R}^{n+1} \to \mathbb{R}_+$ given by

$$G_p(z) = \frac{1}{2}\|H_p(z)\|^2 = \frac{1}{2}(\varepsilon^2 + \|\Phi_{p,\varepsilon}(x)\|^2) = \frac{1}{2}\varepsilon^2 + \Psi_p(z). \tag{15}$$

The function H_p is locally Lipschitz continuous since ϕ_p is locally Lipschitz continuous (see [6]). Furthermore, as shown in Section 3, H_p is semismooth. By this, we apply the generalized Newton method developed by [26,28] for (14), and establish a regularized semismooth Newton-type algorithm which in each step solves a regularized problem NCP(F_ε) approximately. Compared with the semismooth Newton method based on (7), the method has a remarkable advantage in handling the P_0-NCPs (see Section 4) since the merit function $\Psi_{p,\varepsilon}(x)$ has bounded level sets for such NCPs. We examine the numerical performance of the algorithm by applying it for all test problems from MCPLIB with three specific NCP-functions $\phi_{1.1}$, ϕ_2 and ϕ_5. Numerical results indicate that the method associated with a smaller p, for example $p \in [1.1, 2]$, usually has better numerical performance, and the generalized FB functions ϕ_p with $p \in [1.1, 2)$ can be used as the substitutions for the FB function ϕ_2.

Throughout this paper, $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the set of nonnegative real numbers and the set of positive real numbers, respectively; $\mathbb{R}^n$ represents the space of n-dimensional real column vectors; and $^{\mathrm{T}}$ is the transpose notation. For any differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x)$ denotes the gradient of f at x. For any differentiable mapping

$F = (F_1, \ldots, F_m)^{\mathrm{T}} : \mathbb{R}^n \to \mathbb{R}^m$, $F'(x)$ means the Jacobian matrix of F at x while $\nabla F(x) = [\nabla F_1(x) \ldots \nabla F_m(x)]$ denotes the transpose Jacobian of F at x. If W is an $n \times n$ matrix with entries W_{jk}, $j, k = 1, 2, \ldots, n$, and $\mathscr{J}$ and $\mathscr{K}$ are index sets such that $\mathscr{J}, \mathscr{K} \subseteq \{1, 2, \ldots, n\}$, we denote by $W_{\mathscr{J}\mathscr{K}}$ the $|\mathscr{J}| \times |\mathscr{K}|$ submatrix of W consisting of entries W_{jk}, $j \in \mathscr{J}, k \in \mathscr{K}$. We denote by $\|x\|_p$ the p-norm of x and by $\|x\|$ the Euclidean norm of x. In addition, unless otherwise stated, we always assume that p is any fixed real number in $(1, +\infty)$ and denote S^* by the solution set of the NCP if it is nonempty.

2. Preliminaries

In this section, we recall some background concepts and materials which will be used in the subsequent analysis. We start with the definition of P-matrix and P_0-matrix.

Definition 2.1. Given a matrix $M \in \mathbb{R}^{n\times n}$, then M is a

(a) P_0-matrix if each of its principal minors is nonnegative;
(b) P-matrix if each of its principal minors is positive.

Clearly, a positive semidefinite matrix is a P_0-matrix, a positive definite matrix is a P-matrix, and every P-matrix is also a P_0-matrix. For more properties about P-matrix and P_0-matrix, please refer to [8]. The two concepts can be extended to nonlinear mappings.

Definition 2.2. Given a mapping $F : \mathbb{R}^n \to \mathbb{R}^n$, then F is a

(a) monotone function if $\langle x - y, F(x) - F(y)\rangle \geqslant 0$ for all $x, y \in \mathbb{R}^n$;
(b) P_0-function if $\max_{\substack{1\leqslant i\leqslant n\\ x_i\neq y_i}} (x_i - y_i)(F_i(x) - F_i(y)) \geqslant 0$ for all $x, y \in \mathbb{R}^n$ and $x \neq y$;
(c) P-function if $\max_{1\leqslant i\leqslant n}(x_i - y_i)(F_i(x) - F_i(y)) > 0$ for all $x, y \in \mathbb{R}^n$ and $x \neq y$;
(d) uniform P-function with modulus $\mu > 0$ if $\max_{1\leqslant i\leqslant n}(x_i - y_i)(F_i(x) - F_i(y)) \geqslant \mu\|x - y\|^2$ for all $x, y \in \mathbb{R}^n$.

From the above definitions, it is obvious that F is a P_0-function if F is monotone, and the Jacobian matrix of every continuously differentiable P_0-function is a P_0-matrix. The following lemma states that the mapping F_ε is a P-function if F is a P_0-function.

Lemma 2.1 (*Facchinei and Kanzow [9, Lemma 3.2]*). *For any $\varepsilon > 0$, let $F_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ be given by* (10). *If F is a P_0-function, then the Jacobian matrices $F'_\varepsilon(x)$ for all $x \in \mathbb{R}^n$ are P-matrices. In particular, the function F_ε is a P-function.*

Next, we review some favorable properties of ϕ_p where the proofs of Property 2.1 can be found in [2, Proposition 3.1] and [25, Lemmas 2.1 and 2.2] whereas the proof of Property 2.2 is given by [25, Lemma 3.1].

Property 2.1. *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be defined as in* (6). *Then, the following results hold.*

(a) *ϕ_p is an NCP-function.*
(b) *ϕ_p is Lipschitz continuous with the Lipschitz constant L given by $L = \sqrt{2} + 2^{(1/p-1/2)}$ when $1 < p < 2$ and $L = 1 + \sqrt{2}$ when $p \geqslant 2$.*
(c) *ϕ_p is strongly semismooth.*
(d) *Given any point $(a, b) \in \mathbb{R}^2$, each element in the generalized gradient $\partial\phi_p(a, b)$ has the representation $(\xi - 1, \zeta - 1)$, where*

$$\xi = \frac{\operatorname{sgn}(a)\cdot|a|^{p-1}}{\|(a,b)\|_p^{p-1}} \quad \text{and} \quad \zeta = \frac{\operatorname{sgn}(b)\cdot|b|^{p-1}}{\|(a,b)\|_p^{p-1}} \quad \text{if } (a,b) \neq (0,0)$$

and otherwise $(\xi, \zeta) \in \mathbb{R}^2$ denotes an arbitrary vector satisfying $|\xi|^{p/(p-1)} + |\zeta|^{p/(p-1)} \leqslant 1$.

(e) *If* $\{(a^k, b^k)\} \subseteq \mathbb{R}^2$ *with* $a^k \to -\infty$, *or* $b^k \to -\infty$, *or* $a^k \to +\infty$ *and* $b^k \to +\infty$, *then we have* $|\phi_p(a^k, b^k)| \to +\infty$ *for* $k \to +\infty$.

Property 2.2. *Let* $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ *be defined as in* (6). *Then, there exists two positive constants* $c_1 > 0$ *and* $c_2 > 0$ *such that* $c_1|\min\{a, b\}| \geqslant |\phi_p(a, b)| \geqslant c_2|\min\{a, b\}|$.

The semismooth property is very important from computational point of view. In particular, it plays a fundamental role in the superlinear convergence analysis of generalized Newton methods [26,28]. If the mapping $G : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous, then G is almost everywhere differentiable by Rademacher's Theorem (see [7]). In this case, the generalized Jacobian $\partial G(x)$ of G at x (in the Clarke sense) can be defined as the convex hull of the generalized Jacobian $\partial_B G(x)$, where

$$\partial_B G(x) := \{V \in \mathbb{R}^{m\times n} | \exists \{x^k\} \subseteq D_G : \{x^k\} \to x \quad \text{and} \quad G'(x^k) \to V\}$$

with D_G denoting the set of differentiable points of G. Assume that $G : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitz continuous. G is called *semismooth* at x if G is directionally differentiable at x and for any $V \in \partial G(x + h)$ and $h \to 0$,

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

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

$$G(x + h) - G(x) - Vh = O(\|h\|^2); \tag{16}$$

G is called a (strongly) semismooth function if it is (strongly) semismooth everywhere.

3. Properties of $H_p(z)$ and $G_p(z)$

In this section, we will study the semismoothness of the mapping H_p and characterize its generalized Jacobian matrix at any point z. In particular, we also give a sufficient condition for the nonsingularity of all generalized Jacobians at a solution of (14). Then, we investigate some favorable properties of the merit function $G_p(z)$ which are crucial to the convergence analysis of the regularized semismooth Newton algorithm described as in the next section.

Proposition 3.1. *The mapping* $H_p : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ *defined as in* (13) *is semismooth. Moreover, it is strongly semismooth if* F' *is locally Lipschitz continuous.*

Proof. Since a function is (strongly) semismooth if and only if its component functions are (strongly) semismooth, to prove that H_p is (strongly) semismooth we only need to prove that $H_{p,i},\ i = 1, 2, \ldots, n+1$ are (strongly) semismooth. Apparently, $H_{p,1}$ is strongly semismooth by formula (16) since $H_{p,1}(z) = \varepsilon$. For $H_{p,i},\ i = 2, 3, \ldots, n + 1$, since ϕ_p is strongly semismooth by Property 2.1 (c) and the composite of two (strongly) semismooth functions is (strongly) semismooth by [14, Theorem 19], we conclude that $H_{p,i},\ i = 2, 3, \ldots, n+1$ are semismooth. If F' is locally Lipschitz continuous, then F_ε is strongly semismooth, and consequently, $H_{p,i},\ i = 2, 3, \ldots, n+1$ are strongly semismooth. □

We next give the estimation of the generalized Jacobian of H_p by Property 2.1 (d).

Proposition 3.2. *For any* $z = (\varepsilon, x) \in \mathbb{R}_+ \times \mathbb{R}^n$, *we have*

$$(\partial H_p(z))^{\mathrm{T}} \subseteq \begin{pmatrix} 1 & x^{\mathrm{T}}B(z) \\ 0 & (A(z) - I) + (\nabla F(x) + \varepsilon I)(B(z) - I) \end{pmatrix}, \tag{17}$$

where $A(z)$ *and* $B(z)$ *are possibly multi-valued* $n \times n$ *diagonal matrices with ith diagonal elements* $A_{ii}(z)$ *and* $B_{ii}(z)$ *given by*

$$A_{ii}(z) = \frac{\mathrm{sgn}(x_i) \cdot |x_i|^{p-1}}{\|(x_i, F_{\varepsilon,i}(x))\|_p^{p-1}}, \qquad B_{ii}(z) = \frac{\mathrm{sgn}(F_{\varepsilon,i}(x)) \cdot |F_{\varepsilon,i}(x)|^{p-1}}{\|(x_i, F_{\varepsilon,i}(x))\|_p^{p-1}}$$

if $(x_i, F_{\varepsilon,i}(x)) \neq (0,0)$; *and otherwise given by*

$$A_{ii}(z) = \xi_i, \quad B_{ii}(z) = \zeta_i \quad \text{for any } (\xi_i, \zeta_i) \text{ such that } |\xi_i|^{p/(p-1)} + |\zeta_i|^{p/(p-1)} \leqslant 1.$$

Proof. By the known rules on the evaluation of the generalized Jacobian (see [7, Proposition 2.6.2(e)]), we have

$$\partial H_p(z)^{\mathrm{T}} \subseteq \partial H_{p,1}(z) \times \partial H_{p,2}(z) \times \cdots \times \partial H_{p,n+1}(z),$$

where the right-hand side denotes a set of matrices whose ith column belongs to $\partial H_{p,i}(z)$, and $H_{p,i}$ is the ith component function of H_p. Clearly,

$$\partial H_{p,1}(z) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}^{n+1}.$$

For $j = 2, 3, \ldots, n+1$, letting $i = j-1$ and applying Property 2.1 (d) yield

$$\partial H_{p,j}(z) = \left(\frac{\operatorname{sgn}(x_i) \cdot |x_i|^{p-1}}{\|(x_i, F_{\varepsilon,i}(x))\|_p^{p-1}} - 1 \right) \begin{pmatrix} 0 \\ e_i \end{pmatrix} + \begin{pmatrix} x_i \\ \nabla F_i(x) + \varepsilon e_i \end{pmatrix} \left(\frac{\operatorname{sgn}(F_{\varepsilon,i}(x)) \cdot |F_{\varepsilon,i}(x)|^{p-1}}{\|(x_i, F_{\varepsilon,i}(x))\|_p^{p-1}} - 1 \right)$$

if $(x_i, F_{\varepsilon,i}(x)) \neq (0,0)$; and otherwise

$$\partial H_{p,j}(z) = (\xi_i - 1) \begin{pmatrix} 0 \\ e_i \end{pmatrix} + \begin{pmatrix} x_i \\ \nabla F_i(x) + \varepsilon e_i \end{pmatrix} (\zeta_i - 1)$$

with $|\xi_i|^{p/(p-1)} + |\zeta_i|^{p/(p-1)} \leqslant 1$, where e_i denotes the vector whose ith element is zero and other elements are 1. From these equalities, the conclusion easily follows. □

Now, exploiting the estimation of $\partial H_p(z)$ given by (17), we may present a sufficient condition to guarantee the nonsingularity of all generalized Jacobians of H_p at a solution z^* of (14). This result is important for the superlinear (or quadratic) convergence of the semismooth Newton method (see [11]). Let $z^* = (\varepsilon^*, x^*) \in \mathbb{R}_+ \times \mathbb{R}^n$ be a solution of (14). Clearly, $\varepsilon^* = 0$ and x^* is a solution of the NCP. For the sake of notation, let

$$\begin{aligned} \mathscr{I} &:= \{i \in \{1,2,\ldots,n\} \mid x_i^* > 0,\ F_i(x^*) = 0\}, \\ \mathscr{J} &:= \{i \in \{1,2,\ldots,n\} \mid x_i^* = 0,\ F_i(x^*) = 0\}, \\ \mathscr{K} &:= \{i \in \{1,2,\ldots,n\} \mid x_i^* = 0,\ F_i(x^*) > 0\}. \end{aligned}$$

By rearrangement we assume that $\nabla F(x^*)$ can be written as

$$\nabla F(x^*) = \begin{pmatrix} \nabla F_{\mathscr{I}\mathscr{I}}(x^*) & \nabla F_{\mathscr{I}\mathscr{J}}(x^*) & \nabla F_{\mathscr{I}\mathscr{K}}(x^*) \\ \nabla F_{\mathscr{J}\mathscr{I}}(x^*) & \nabla F_{\mathscr{J}\mathscr{J}}(x^*) & \nabla F_{\mathscr{J}\mathscr{K}}(x^*) \\ \nabla F_{\mathscr{K}\mathscr{I}}(x^*) & \nabla F_{\mathscr{K}\mathscr{J}}(x^*) & \nabla F_{\mathscr{K}\mathscr{K}}(x^*) \end{pmatrix}. \tag{18}$$

The NCP is called R-regular at x^* if $\nabla F_{\mathscr{I}\mathscr{I}}(x^*)$ is nonsingular and its Schur-complement in the matrix $\begin{pmatrix} \nabla F_{\mathscr{J}\mathscr{I}}(x^*) & \nabla F_{\mathscr{J}\mathscr{J}}(x^*) \\ \nabla F_{\mathscr{J}\mathscr{I}}(x^*) & \nabla F_{\mathscr{J}\mathscr{J}}(x^*) \end{pmatrix}$ is a P-matrix.

Proposition 3.3. *Suppose that* $z^* = (\varepsilon^*, x^*) \in \mathbb{R}_+ \times \mathbb{R}^n$ *be a solution of* (14) *and the NCP is R-regular at* x^*, *then all* $V \in \partial H_p(z^*)$ *are nonsingular.*

Proof. From Proposition 3.2, it is easy to see that for any $V \in \partial H_p(z^*)^{\mathrm{T}}$, there exists a vector $u(z^*) \in \mathbb{R}^n$ and a matrix $W(z^*) \in \mathbb{R}^{n \times n}$ such that

$$V = \begin{pmatrix} 1 & u(z^*)^{\mathrm{T}} \\ 0 & W(z^*) \end{pmatrix},$$

where

$$W(z^*) = (A(z^*) - I) + (\nabla F(x^*) + \varepsilon^* I)(B(z^*) - I)$$

with $A(z^*)$ and $B(z^*)$ characterized as in Proposition 3.2. Therefore, proving that V is nonsingular is equivalent to arguing that $W(z^*)$ is nonsingular. Using the expression of $\nabla F(x^*)$ in (18) and noting that $\varepsilon^* = 0$, we can rewrite $W(z^*)$ in the partitioned form

$$W(z^*) = \begin{pmatrix} -\nabla F_{\mathcal{J}\mathcal{J}} & \nabla F_{\mathcal{J}\mathcal{J}}(B_{\mathcal{J}\mathcal{J}} - I_{\mathcal{J}\mathcal{J}}) & 0_{\mathcal{J}\mathcal{K}} \\ -\nabla F_{\mathcal{J}\mathcal{J}} & \nabla F_{\mathcal{J}\mathcal{J}}(B_{\mathcal{J}\mathcal{J}} - I_{\mathcal{J}\mathcal{J}}) + (A_{\mathcal{J}\mathcal{J}} - I_{\mathcal{J}\mathcal{J}}) & 0_{\mathcal{J}\mathcal{K}} \\ -\nabla F_{\mathcal{K}\mathcal{J}} & \nabla F_{\mathcal{K}\mathcal{J}}(B_{\mathcal{J}\mathcal{J}} - I_{\mathcal{J}\mathcal{J}}) & -I_{\mathcal{K}\mathcal{K}} \end{pmatrix},$$

where for convenience we dispense with the notations z^* and x^*. The rest of the proof is identical to that of [11, Proposition 3.2]. □

In what follows, we concentrate on the properties of G_p. First, applying [6, Propositon 3.2 (c)] and Theorem 2.6.6 of [7], we immediately obtain the following conclusion.

Proposition 3.4. *For any $\varepsilon \geqslant 0$, the function $\Psi_{p,\varepsilon}$ defined by* (12) *is continuously differentiable everywhere, and consequently, G_p defined as in* (15) *is continuously differentiable everywhere and $\nabla G_p(z) = V^{\mathrm{T}} H_p(z)$ for any $V \in \partial H_p(z)$.*

Proposition 3.5. *Suppose that F is a P_0-function and $\hat{\varepsilon}$, $\tilde{\varepsilon}$ are two given positive numbers such that $\hat{\varepsilon} < \tilde{\varepsilon}$. Then, the merit function G_p defined as in* (15) *has the property*:

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

for any sequence $\{z^k = (\varepsilon^k, x^k)\}$ such that $\varepsilon^k \in [\hat{\varepsilon}, \tilde{\varepsilon}]$ and $\|x^k\| \to +\infty$.

Proof. We prove this by contradiction which is a standard and common technique. Suppose $\lim_{k\to+\infty} G_p(z^k) \neq +\infty$. Then from (15) and (12) it follows that there exists an unbounded sequence $\{x^k\}$ such that $\{\Psi_{p,\varepsilon^k}(x^k)\}$ is bounded. Let

$$J := \{i \in \{1, 2, \ldots, n\} | \{x_i^k\} \text{ is unbounded}\}.$$

Since $\{x^k\}$ is unbounded, we have $J \neq \emptyset$. Without loss of generality, we assume that $\{|x_j^k|\} \to \infty$ for any $j \in J$. Now, we define a bounded sequence by

$$y_i^k := \begin{cases} 0 & \text{if } i \in J, \\ x_i^k & \text{if } i \notin J. \end{cases}$$

From the definition of $\{y^k\}$ and F being a P_0-function, we have

$$\begin{aligned} 0 &\leqslant \max_{\substack{1 \leqslant i \leqslant n \\ x_i^k \neq y_i^k}} (x_i^k - y_i^k)(F_i(x^k) - F_i(y^k)) \\ &= \max_{i \in J} x_i^k \cdot (F_i(x^k) - F_i(y^k)) \\ &= x_{j_0}^k \cdot (F_{j_0}(x^k) - F_{j_0}(y^k)), \end{aligned} \tag{19}$$

where j_0 is one of the indices for which the max is attained. Since $j_0 \in J$, we have that $\{|x_{j_0}^k|\} \to +\infty$ as $k \to +\infty$. If $x_{j_0}^k \to -\infty$ as $k \to +\infty$, using Property 2.1(e) immediately yields that $\phi_p(x_{j_0}^k, F_{\varepsilon^k, j_0}(x^k)) \to +\infty$. If $x_{j_0}^k \to +\infty$ as $k \to +\infty$, noting that $F_{j_0}(y^k)$ is bounded by the continuity of F_{j_0}, we have from (19) that $F_{j_0}(x^k)$ does not tend to $-\infty$, which in turn implies that $\{F_{j_0}(x^k) + \varepsilon^k x_{j_0}^k\} \to +\infty$. From Property 2.1(e) where $\{x_{j_0}^k\} \to +\infty$ and $\{F_{j_0}(x^k) + \varepsilon^k x_{j_0}^k\} \to +\infty$,we also obtain that $\phi_p(x_{j_0}^k, F_{\varepsilon^k, j_0}(x^k)) \to +\infty$. Thus, both cases yield $\phi_p(x_{j_0}^k, F_{\varepsilon^k, j_0}(x^k)) \to +\infty$ which is a contradiction to the boundedness of $\{\Psi_{p,\varepsilon^k}(x^k)\}$. Consequently, we prove that $\lim_{k\to+\infty} G_p(z^k) = +\infty$. □

Remark 3.1. Proposition 3.5 implies that $\Psi_{p,\varepsilon}$ has bounded level sets under the assumption of F being a P_0-function. However, from [6, Proposition 3.5], we know that a stronger condition (i.e., F being a uniform P-function) is needed to guarantee the level sets of Ψ_p to be bounded.

To close this section, we present two results which will be used to analyze the global convergence of the algorithm in the next section. The first result is extracted from Theorem 5.4 of [9], while the second result can be obtained by using Property 2.2 and following the same arguments as in [30, Proposition 2.2].

Proposition 3.6. *Suppose that F is a P_0-function and the solution set S^* of the NCP is nonempty and bounded. Suppose that $\{\varepsilon^k\}$ and $\{x^k\}$ are two infinite sequences such that for each $k \geqslant 0$, $\varepsilon^k > 0$, $\eta^k \geqslant 0$ satisfying $\lim_{k\to+\infty}\varepsilon^k = 0$, $\lim_{k\to+\infty}\eta^k = 0$. For each $k \geqslant 0$, let $x^k \in \mathbb{R}^n$ satisfy $\|\Phi_p(\varepsilon^k, x^k)\| \leqslant \eta^k$. Then $\{x^k\}$ remains bounded and every accumulation point of $\{x^k\}$ is a solution of the NCP.*

Proposition 3.7. *Suppose that F is a monotone function and the solution set S^* of the NCP is nonempty. Suppose that $\{\varepsilon^k\}$ and$\{x^k\}$ are two infinite sequences such that for each $k \geqslant 0$, $\varepsilon^k > 0$, $\eta^k \geqslant 0$, $\eta^k \geqslant C\varepsilon^k$ and $\lim_{k\to+\infty}\varepsilon^k = 0$, where $C > 0$ is a constant. For each $k \geqslant 0$, let $x^k \in \mathbb{R}^n$ satisfy $\|\Phi_p(\varepsilon^k, x^k)\| \leqslant \eta^k$. Suppose that $x^* = \arg\min_{x\in S^*}\|x\|$ and F is Lipschitz continuous. Then $\{x^k\}$ remains bounded and every accumulation point of $\{x^k\}$ is a solution of the NCP.*

4. Regularization semismooth Newton method

From the discussions of last section, we see that $H_p(z)$ and $G_p(z)$ for all $p > 1$ enjoy the same desirable properties. Sun [30] used $H_2(z)$ and $G_2(z)$ to develop a regularization semismooth Newton method for the NCP. In this section, we will develop a regularization semismooth Newton algorithm by any $H_p(z)$ and $G_p(z)$ with $p > 1$. This algorithm is guaranteed to solve P_0-complementarity problems due to Proposition 3.5.

Now we are ready to describe this specific algorithm. We adopt almost the same notations used in [30]. Choose $\bar{\varepsilon} \in (0, +\infty)$ and $\gamma \in (0, 1)$ such that $\gamma\bar{\varepsilon} < 1$. Let $t \in [1/2, 1]$ and $\bar{z} := (\bar{\varepsilon}, 0) \in \mathbb{R}_{++} \times \mathbb{R}^n$. Define $\beta : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$ by

$$\beta(z) := \gamma \min\{1, G_p(z)^t\}. \tag{20}$$

We also denote

$$\Omega := \{z = (\varepsilon, x) \in \mathbb{R}_+ \times \mathbb{R}^n | \varepsilon \geqslant \beta(z)\bar{\varepsilon}\}. \tag{21}$$

Note that $\beta(z) \leqslant \gamma$ for any $z \in \mathbb{R}_+ \times \mathbb{R}^n$ by (20). Hence, $(\bar{\varepsilon}, x) \in \Omega$ for any $x \in \mathbb{R}^n$. In addition, by the definition of $\beta(z)$, it is easily shown the following relation holds.

Proposition 4.1. *Let H_p and β be defined as in* (13) *and* (20), *respectively. Then,*

$$H_p(z) = 0 \iff \beta(z) = 0 \iff H_p(z) = \beta(z)\bar{z}.$$

Algorithm 4.1 (*The Regularization Newton Algorithm*).

(*Step 0*) *Given any $p > 1$ and choose constants $\delta \in (0, 1)$, $t \in [1/2, 1]$ and $\sigma \in (1, 1/2)$. Let $\varepsilon^0 := \bar{\varepsilon}$ and $x^0 \in \mathbb{R}^n$ be an arbitrary point. Set $k := 0$.*

(*Step 1*) *If $H_p(z^k) = 0$, then stop. Otherwise, let*

$$\beta_k := \beta(z^k) = \gamma \min\{1, G_p(z^k)^t\}.$$

(*Step 2*) *Choose $V_k \in \partial H_p(z^k)$ and compute $\Delta z^k = (\Delta\varepsilon^k, \Delta x^k) \in \mathbb{R} \times \mathbb{R}^n$ by*

$$H_p(z^k) + V_k\Delta z^k = \beta_k\bar{z}. \tag{22}$$

(*Step 2*) *Let l_k be the smallest nonnegative integer l such that*

$$G_p(z^k + \delta^l\Delta z^k) \leqslant [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l]G_p(z^k). \tag{23}$$

(*Step 2*) *Set* $z^{k+1} := z^k + \delta^{l_k}\Delta z^k$.
(*Step 4*) *Set* $k := k+1$ *and go to Step* 1.

From Proposition 3.2, we know that for any $V \in \partial H_p(z)$ with $z=(\varepsilon, x) \in \mathbb{R}_{++}\times\mathbb{R}^n$, there exists a $W=(u(z)\ W(z)) \in \partial\Phi_p(z)$ with $u(z) \in \mathbb{R}^n$ and $W(z) \in \mathbb{R}^{n\times n}$ such that

$$V = \begin{pmatrix} 1 & 0 \\ u(z) & W(z) \end{pmatrix}. \tag{24}$$

Suppose that F is a P_0-function. Then by Lemma 2.1 $F'_\varepsilon(x)$ is a P-matrix. Hence, for any $x \in \mathbb{R}^n$ and $\varepsilon > 0$, $W(z)$ is nonsingular by the proof of Proposition 2 of [20]. It thus follows that all $V \in \partial H_p(z)$ with $z = (\varepsilon, x) \in \mathbb{R}_{++} \times \mathbb{R}^n$ are nonsingular. Therefore, the Newton step in (22) is well-defined, and moreover, from (22), for any $k \geqslant 0$ and $\varepsilon^k > 0$, there exists a $W_k \in \partial\Phi_p(z^k)$ such that

$$(\nabla\Psi_p(z^k))^{\mathrm{T}}\Delta z^k = \Phi_p(z^k)^{\mathrm{T}}W_k\Delta z^k = -\Phi_p(z^k)^{\mathrm{T}}\Phi_p(z^k) = -2\Psi_p(z^k). \tag{25}$$

Using the equality and Proposition 4.1, we next show that Algorithm 4.1 is well-defined.

Proposition 4.2. *Suppose that* F *is a* P_0*-function and* $z^k = (\varepsilon^k, x^k) \in \mathbb{R}_{++} \times \mathbb{R}^n$ *for* $k \geqslant 0$. *Then* $z^{k+1} \in \mathbb{R}_{++} \times \mathbb{R}^n$ *and Algorithm* 4.1 *is well-defined.*

Proof. Since $\varepsilon^k > 0$, from the definition of $\beta(z)$ it follows that $\beta_k = \beta(z^k) > 0$. From the first component in the relation (22) in Algorithm 4.1, we have

$$\varepsilon^k + \Delta\varepsilon^k = \beta_k\bar{\varepsilon} \implies \triangle\varepsilon^k = -\varepsilon^k + \beta_k\bar{\varepsilon}. \tag{26}$$

Then, for any $\alpha \in [0, 1]$, there has

$$\varepsilon^k + \alpha\Delta\varepsilon^k = (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} > 0. \tag{27}$$

Thus, combining the fact that $\beta(z) \leqslant \gamma G_p(z)^{1/2}$ with (22) and (27) yields that

$$\begin{aligned}(\varepsilon^k + \alpha\Delta\varepsilon^k)^2 &= [(1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon}]^2 \\ &= (1-\alpha)^2(\varepsilon^k)^2 + 2(1-\alpha)\alpha\beta_k\varepsilon^k\bar{\varepsilon} + \alpha^2\beta_k^2\bar{\varepsilon}^2 \\ &\leqslant (1-\alpha)^2(\varepsilon^k)^2 + 2\alpha\beta_k\varepsilon^k\bar{\varepsilon} + \mathrm{O}(\alpha^2) \\ &\leqslant (1-\alpha)^2(\varepsilon^k)^2 + 2\alpha\gamma G_p(z^k)^{1/2}\|H_p(z^k)\|\bar{\varepsilon} + \mathrm{O}(\alpha^2) \\ &= (1-2\alpha)(\varepsilon^k)^2 + 2\sqrt{2}\alpha\gamma\bar{\varepsilon}G_p(z^k) + \mathrm{O}(\alpha^2).\end{aligned} \tag{28}$$

Now, we define

$$\theta(\alpha) := \Psi_p(z^k + \alpha\Delta z^k) - \Psi_p(z^k) - \alpha(\nabla\Psi_p(z^k))^{\mathrm{T}}\Delta z^k.$$

Since Ψ_p is continuously differentiable at any $z^k \in \mathbb{R}_{++} \times \mathbb{R}^n$ by Proposition 3.4, we obtain $\theta(\alpha) = o(\alpha)$. On the other hand, from (22) and (25) it follows that

$$\begin{aligned}\tfrac{1}{2}\|\Phi_p(z^k + \alpha\Delta z^k)\|^2 &= \Psi_p(z^k + \alpha\Delta z^k) \\ &= \Psi_p(z^k) + \alpha(\nabla\Psi_p(z^k))^{\mathrm{T}}\Delta z^k + \theta(\alpha) \\ &= \Psi_p(z^k) - 2\alpha\Psi_p(z^k) + \mathrm{o}(\alpha) \\ &= (1-2\alpha)\Psi_p(z^k) + \mathrm{o}(\alpha)\end{aligned} \tag{29}$$

for any $\alpha \in [0, 1]$. Therefore, using Eqs. (28) and (29), we obtain

$$\begin{aligned} G_p(z^k + \alpha\Delta z^k) &= \tfrac{1}{2}\|H_p(z^k + \alpha\Delta z^k)\|^2 \\ &= \tfrac{1}{2}(\varepsilon^k + \alpha\Delta\varepsilon^k)^2 + \tfrac{1}{2}\|\Phi_p(z^k + \alpha\Delta z^k)\|^2 \\ &\leqslant \tfrac{1}{2}(1-2\alpha)(\varepsilon^k)^2 + \sqrt{2}\alpha\gamma\bar{\varepsilon}G_p(z^k) + (1-2\alpha)\Psi_p(z^k) + \mathrm{o}(\alpha) \\ &\leqslant (1-2\alpha)G_p(z^k) + 2\alpha\gamma\bar{\varepsilon}G_p(z^k) + \mathrm{o}(\alpha) \\ &= [1 - 2(1-\gamma\bar{\varepsilon})\alpha]G_p(z^k) + \mathrm{o}(\alpha) \end{aligned} \tag{30}$$

for any $\alpha \in [0, 1]$. The inequality (30) implies that there exists $\bar{\alpha} \in (0, 1]$ such that

$$G_p(z^k + \alpha\Delta z^k) \leqslant [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]G_p(z^k) \quad \forall \alpha \in [0, \bar{\alpha}],$$

which indicates that Algorithm 4.1 is well-defined. □

Proposition 4.3. *Suppose that F is a P_0-function. For each $k \geqslant 0$, if $\varepsilon^k > 0$ and $z^k \in \Omega$, then for any $\alpha \in [0, 1]$ such that*

$$G_p(z^k + \alpha\Delta z^k) \leqslant [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]G_p(z^k), \tag{31}$$

there holds that $z^k + \alpha\Delta z^k \in \Omega$.

Proof. We prove this proposition by considering the following two cases:

Case (i): $G_p(z^k) > 1$. Then $\beta_k = \gamma$. From $z^k \in \Omega$ and $\beta(z) = \gamma\min\{1, G_p(z)^t\} \leqslant \gamma$ for any $z \in \mathbb{R}_+ \times \mathbb{R}^n$, it follows that for any $\alpha \in [0, 1]$,

$$\begin{aligned} (\varepsilon^k + \alpha\triangle\varepsilon^k) - \beta(z^k + \alpha\Delta z^k)\bar{\varepsilon} &\geqslant (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} - \gamma\bar{\varepsilon} \\ &\geqslant (1-\alpha)\beta_k\bar{\varepsilon} + \alpha\beta_k\bar{\varepsilon} - \gamma\bar{\varepsilon} \\ &= 0. \end{aligned} \tag{32}$$

Case (ii): $G_p(z^k) \leqslant 1$. Then, for any $\alpha \in [0, 1]$ satisfying (31), we have

$$G_p(z^k + \alpha\Delta z^k) \leqslant [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]G_p(z^k) \leqslant 1. \tag{33}$$

Therefore, for any $\alpha \in [0, 1]$ satisfying (31),

$$\beta(z^k + \alpha\Delta z^k) = \gamma G_p(z^k + \alpha\triangle z^k)^t.$$

Using the fact that $z^k \in \Omega$ and the first inequality in (33), we then obtain that for any $\alpha \in [0, 1]$ satisfying (31),

$$\begin{aligned} (\varepsilon^k + \alpha\triangle\varepsilon^k) - \beta(z^k + \alpha\Delta z^k)\bar{\varepsilon} &\geqslant (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} - \gamma G_p(z^k + \alpha\triangle z^k)^t\bar{\varepsilon} \\ &\geqslant (1-\alpha)\beta_k\bar{\varepsilon} + \alpha\beta_k\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t G_p(z^k)^t\bar{\varepsilon} \\ &= \beta_k\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t G_p(z^k)^t\bar{\varepsilon} \\ &= \gamma G_p(z^k)^t\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t G_p(z^k)^t\bar{\varepsilon} \\ &= \gamma\{1 - [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t\}G_p(z^k)^t\bar{\varepsilon} \\ &\geqslant 0. \end{aligned} \tag{34}$$

Combining (32) and (34) immediately yields the desired result. □

Proposition 4.4. *Suppose that F is a P_0-function. Then Algorithm* 4.1 *generates an infinite sequence $\{z^k\}$ with $z^k \in \Omega$ for all k and*

$$0 < \varepsilon^{k+1} \leqslant \varepsilon^k \leqslant \bar{\varepsilon} \quad \textit{for all } k. \tag{35}$$

Proof. Since $z^0 = (\bar{\varepsilon}, x^0) \in \Omega$, the first part of the conclusions follows by repeatedly resorting to Propositions 4.2 and 4.3. We next concentrate on the proof of (35). First, $\varepsilon^0 = \bar{\varepsilon} > 0$. From the design of Algorithm 4.1 and the fact that $\beta(z) = \gamma \min\{1, G_p(z)^t\} \leqslant \gamma$ for any $z \in \mathbb{R}_+ \times \mathbb{R}^n$, it then follows that

$$\varepsilon^1 = (1 - \delta^{l_0})\varepsilon^0 + \delta^{l_0}\beta(z^0)\bar{\varepsilon} \leqslant (1 - \delta^{l_0})\bar{\varepsilon} + \delta^{l_0}\gamma\bar{\varepsilon} \leqslant \bar{\varepsilon}.$$

Hence (35) holds for $k = 0$. Suppose that (35) holds for $k = i - 1$. We next prove that (35) holds for $k = i$. From the design of Algorithm 4.1, we have

$$\varepsilon^{i+1} = (1 - \delta^{l_i})\varepsilon^i + \delta^{l_i}\beta(z^i)\bar{\varepsilon}.$$

Noting that $\varepsilon^i \geqslant \beta(z^i)\bar{\varepsilon}$ since $z^i \in \Omega$, we then obtain

$$\varepsilon^{i+1} \leqslant (1 - \delta^{l_i})\varepsilon^i + \delta^{l_i}\varepsilon^i = \varepsilon^i$$

and

$$\varepsilon^{i+1} \geqslant (1 - \delta^{l_i})\beta(z^i)\bar{\varepsilon} + \delta^{l_i}\beta(z^i)\bar{\varepsilon} = \beta(z^i)\bar{\varepsilon} > 0.$$

Therefore, (35) holds for $k = i$. We complete the proof. □

Now, using Propositions 3.5–3.7 and Proposition 4.4 and following the same arguments as in [30], we obtain the following global convergence results of Algorithm 4.1.

Theorem 4.1. *Suppose that F is a P_0-function and the solution set S^* of the NCP is nonempty and bounded. Then the infinite sequence $\{z^k\}$ generated by Algorithm* 4.1 *is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$.*

Theorem 4.2. *Suppose that F is a monotone function and in Algorithm* 4.1 *the parameter $t = \frac{1}{2}$. Then if the iteration sequence $\{z^k\}$ is bounded, then the solution set S^* of the NCP is nonempty. Conversely, if the solution set S^* of the NCP is nonempty and F is Lipschitz continuous, then the infinite sequence $\{z^k\}$ generated by Algorithm* 4.1 *is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$.*

In addition, using Proposition 3.1 and similar proof as for [30, Theorem 5.1], we obtain the following local superlinear (quadratic) convergence results of Algorithm 4.1.

Theorem 4.3. *Suppose that F is a P_0-function and the solution set S^* of the NCP is nonempty and bounded. Suppose that $z^* := (\varepsilon^*, x^*) \in \mathbb{R} \times \mathbb{R}^n$ is an accumulation point of the infinite sequence $\{z^k\}$ generated by Algorithm* 4.1 *and all $V \in \partial H_p(z^*)$ are nonsingular. Then the whole sequence $\{z^k\}$ converges to z^* with*

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

Furthermore, if F' is locally Lipschitz continuous around x^*, then

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

Moreover, from Proposition 3.3, all the conclusions of Theorem 4.3 hold if the assumption that all $V \in \partial H_p(z^*)$ are nonsingular is replaced by that the NCP is R-regular at x^*.

5. Numerical experiments

We implemented Algorithm 4.1 by our codes in MATLAB 6.5 for almost all test problems except the unavailable "pvg105" and "scarfbnum" with the starting points in MCPLIB [1]. All numerical experiments were done at a PC with CPU of 2.8 GHz and RAM of 512 MB. Throughout the experiments, unless otherwise stated, we adopted the following parameters for Algorithm 4.1:

$$\delta = 0.5, \quad t = 1/2, \quad \sigma = 10^{-4}, \quad \gamma = 0.5, \quad \bar{\varepsilon} = 0.1.$$

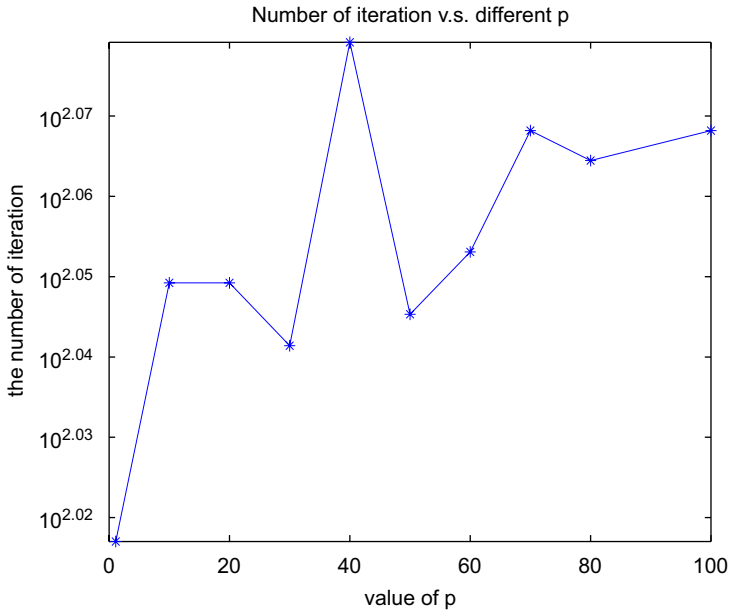

Fig. 1. The number of iterations vs. value of p for Example "bertsekas(3)".

We terminated the iteration if one of the following conditions was satisfied:

(1) $\|H_p(z^k)\| \leqslant \varepsilon_1$ and $\|\min\{x^k, F(x^k)\}\| \leqslant \varepsilon_2$;
(2) the step length $\alpha_k = \delta^{l_k}$ is less than $\alpha_{\min}$.
(3) the number of iteration exceeds $k_{\max}$.

Among others, in our implementation the termination parameters were chosen as follows:

$$\varepsilon_1 = 10^{-10}, \quad \varepsilon_2 = 10^{-6}, \quad \alpha_{\min} = 10^{-25} \quad \text{and} \quad k_{\max} = 1000.$$

During the experiments, we incorporated some strategies to improve the numerical behavior of Algorithm 4.1 to some extent. These strategies are well-accepted and used in basically all suitable implementations of complementarity solvers. The first modification is in the line search step. We replaced the standard (monotone) Armijo-rule by nonmonotone line search described in [35] to seek a suitable steplength, i.e., we computed the smallest nonnegative integer l such that

$$G_p(z^k + \delta^l d^k) \leqslant \mathscr{W}_k - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l G_p(z^k)$$

for all $k \geqslant 0$, where $\mathscr{W}_k$ is given by

$$\mathscr{W}_k = (\eta_{k-1} Q_{k-1} \mathscr{W}_{k-1} + G_p(z^k))/Q_k$$

with

$$Q_k = \eta_{k-1} Q_{k-1} + 1.$$

In our implementation, we used $\mathscr{W}_{-1} = G_p(z^0)$, $Q_{-1} = 1$, $\eta_{-1} = 0.85$ and $\eta_k \equiv 0.85$.

The second modification is necessary since the mapping F is often not defined outside the positive orthant whereas our algorithm assumes that F can be evaluated on the whole space $\mathbb{R}^n$. Hence, in order to avoid possible domain violations, we employed a simple backtracking strategy: Given an iterate $z^k = (\varepsilon^k, x^k) \in \mathbb{R}_{++} \times \mathbb{R}^n$ and a search direction $d^k \in \mathbb{R}^{n+1}$, we first compute the exponent $j_k := \min\{0, 1, 2, \ldots, \}$ such that

$$F(x^k + \beta^{j_k} d^k(2 : n + 1))$$

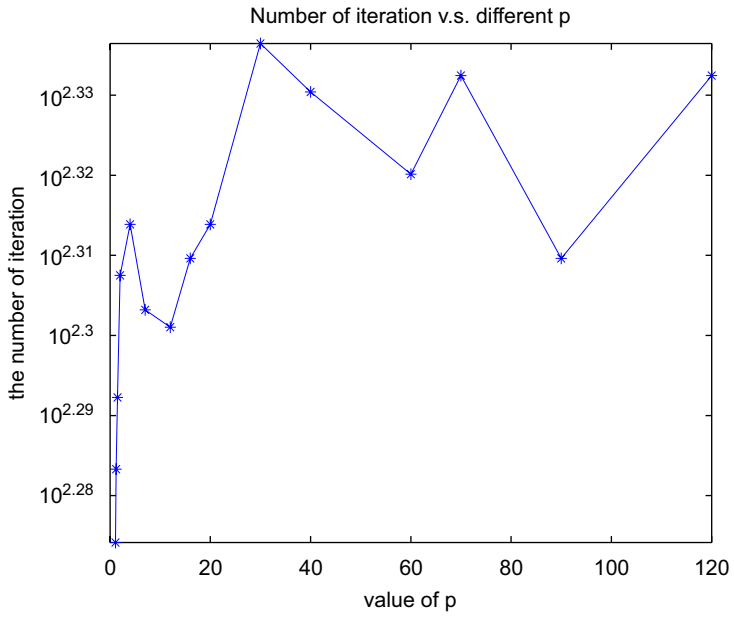

Fig. 2. The number of iterations vs. value of p for Example "freebert(5)".

Table 1
Numerical comparisons among $\phi_{1.1}$, ϕ_2 and ϕ_5 for MCPLIB problems

Problem	$\phi_{1.1}$				ϕ_2				ϕ_5			
	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU
bertsekas (1)	60	171	9.83e-23	0.21	90	219	7.66e-24	0.26	91	231	3.14e-23	0.28
bertsekas (2)	49	157	4.65e-21	0.17	92	214	2.24e-23	0.28	104	257	1.10e-21	0.31
bertsekas (3)	144	344	2.74e-21	0.43	170	356	1.73e-23	0.48	162	386	9.05e-24	0.60
billups	–	–	–	–	234	2443	1.54e-21	0.67	357	8661	2.96e-23	2.23
colvdual (1)	29	67	2.39e-23	0.12	74	88	4.41e-21	0.23	151	219	3.45e-21	0.46
colvdual (2)	120	199	7.10e-24	0.40	279	283	4.90e-21	0.84	556	559	4.59e-21	1.60
colvnlp (1)	25	26	1.51e-21	0.06	67	70	8.61e-23	0.15	62	72	3.27e-22	0.15
colvnlp (2)	21	22	3.37e-21	0.06	38	40	1.65e-22	0.09	36	42	4.98e-21	0.09
cycle	9	11	2.81e-21	0.01	12	15	1.19e-23	0.01	12	14	1.84e-22	0.03
explcp	23	27	1.99e-21	0.07	24	25	6.51e-24	0.07	23	24	8.78e-22	0.04
freebert (1)	71	240	8.60e-23	0.26	82	193	6.79e-23	0.26	86	231	6.98e-22	0.28
freebert (2)	188	402	4.29e-21	0.56	202	365	2.63 e-23	0.54	203	410	3.37e-21	0.57
freebert (3)	78	222	1.70e-23	0.26	93	254	2.56e-22	0.31	96	258	2.57e-21	0.32
freebert (4)	60	171	9.83e-23	0.20	90	219	7.66e-24	0.26	91	231	3.14e-23	0.29
freebert (5)	188	399	8.38e-22	0.56	203	389	5.99e-23	0.56	206	429	9.90e-24	0.59
freebert (6)	69	206	2.41e-21	0.25	98	238	5.70e-22	0.28	101	269	1.82e-22	0.34
hanskoop (1)	207	208	4.92 e-21	0.64	222	229	4.48e-21	0.67	226	238	4.54e-21	0.78
hanskoop (2)	215	216	4.33e-21	0.65	233	235	4.67e-21	0.65	238	241	4.34e-21	0.79
hanskoop (3)	45	78	4.97e-21	0.15	33	35	4.11e-21	0.12	34	36	2.54e-21	0.12

exists and then take $\beta^{j_k} d^k$ as the new search direction d^k, where $d^k(2:n+1)$ denotes the vector composed of the last n components of d^k.

We first took the problems *bertsekas* (3) and *freebert* (5) for examples to test the performance of Algorithm 4.1 on different p. Figs. 1 and 2 depict how the number of iteration varies with the values of p. From the two figures, it seems that the number of iteration will tend to having an increase when p becomes large. In addition, by our numerical experiences, the algorithm will have worse robustness when p tends to 1. In view of these facts, we then compared

Table 2
Numerical comparisons among $\phi_{1.1}$, ϕ_2 and ϕ_5 for MCPLIB problems

Problem	$\phi_{1.1}$				ϕ_2				ϕ_5			
	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU
hanskoop (4)	89	90	4.08e-21	0.40	101	102	3.86e-21	0.31	114	133	4.55e-21	0.35
hanskoop (5)	14	15	1.27e-23	0.03	25	33	1.14e-23	0.07	16	17	5.61e-22	0.04
josephy (1)	11	12	2.96e-22	0.01	21	32	7.99e-22	0.03	11	14	3.59e-22	0.01
josephy (2)	11	14	2.98e-22	0.01	11	15	1.68e-22	0.03	11	17	1.83e-22	0.03
josephy (3)	15	16	4.37e-22	0.03	51	55	4.21e-22	0.04	40	43	4.56e-22	0.06
josephy (4)	9	10	3.12e-22	0.01	9	10	3.05e-21	0.01	9	10	4.99e-22	0.03
josephy (5)	9	10	1.54e-23	0.01	9	10	8.44e-24	0.01	8	9	9.19e-22	0.01
josephy (6)	11	14	1.62e-22	0.01	11	13	1.48e-23	0.01	10	12	7.00e-22	0.01
kojshin(1)	14	15	5.02e-22	0.01	14	17	1.17e-21	0.01	12	15	4.67e-22	0.01
kojshin(2)	12	15	1.30e-22	0.01	12	18	3.49e-21	0.01	10	15	2.88e-21	0.01
kojshin(3)	16	17	6.19e-23	0.02	17	18	3.49e-22	0.02	35	38	1.58e-22	0.04
kojshin(4)	9	10	1.49e-23	0.01	9	10	2.88e-22	0.01	9	10	1.47e-21	0.01
mathinum(1)	9	10	1.21e-21	0.01	9	10	1.83e-23	0.01	9	11	1.08e-21	0.01
mathinum(2)	11	13	1.14e-22	0.01	11	12	9.53e-24	0.01	10	11	1.29e-23	0.01
mathinum(3)	12	13	5.65e-23	0.01	12	15	1.61e-22	0.03	11	13	1.28e-23	0.01
mathinum(4)	12	14	1.69e-23	0.01	12	13	1.09e-22	0.01	11	12	3.66e-22	0.01
mathisum(1)	13	14	6.60e-24	0.02	11	12	5.33e-22	0.01	17	21	3.23e-22	0.03
mathisum(2)	11	12	3.91e-22	0.01	12	13	2.67e-21	0.01	12	13	6.64e-22	0.01
mathisum(3)	10	12	1.82e-21	0.01	10	11	1.77e-23	0.01	9	10	2.23e-21	0.01
mathisum(4)	12	13	3.52e-21	0.01	13	14	2.01e-23	0.02	12	13	8.49e-22	0.01

Table 3
Numerical comparisons among $\phi_{1.1}$, ϕ_2 and ϕ_5 for MCPLIB problems

Problem	$\phi_{1.1}$				ϕ_2				ϕ_5			
	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU	Iter	NF	$G_p(z^f)$	CPU
nash(1)	13	14	9.68e-24	0.03	13	14	1.55e-23	0.03	12	13	4.90e-21	0.03
nash(2)	14	20	8.15e-24	0.03	15	26	2.62e-22	0.04	16	30	3.21e-22	0.04
pgvon106(1)	38	76	3.43e-21	1.06	109	152	3.86e-21	3.17	27	47	4.77e-21	0.73
pgvon106(2)	307	337	1.34e-21	9.06	248	278	2.39e-21	6.78	156	171	8.95e-23	3.98
pgvon106(3)	42	107	4.26e-21	1.25	172	235	5.39e-23	4.46	26	63	1.76e-22	0.76
powell(1)	10	11	3.75e-22	0.03	10	11	7.62e-22	0.03	10	11	4.38e-22	0.03
powell(2)	12	13	4.29e-21	0.04	13	14	1.10e-23	0.06	12	13	2.85e-21	0.05
powell(3)	14	15	3.85e-21	0.04	15	16	9.37e-23	0.04	15	16	4.17e-23	0.04
powell(4)	13	14	2.13e-21	0.04	14	15	3.61e-23	0.04	14	15	1.32e-23	0.04
scarfasum (1)	20	51	2.15e-21	0.09	21	53	1.31e-22	0.12	21	53	4.97e-23	0.09
scarfasum (2)	21	51	3.19e-21	0.11	20	51	3.85e-22	0.07	21	52	4.97e-23	0.07
scarfasum (3)	18	58	1.89e-23	0.14	20	23	1.30e-22	0.06	33	197	1.44e-21	0.35
scarfbsum (1)	17	22	1.09e-22	0.34	34	55	2.82e-23	0.35	13	15	3.21e-21	0.14
scarfbsum (2)	414	971	4.02e-21	3.36	624	2306	2.31e-21	5.58	788	2827	1.54e-21	7.01
sppe (1)	14	15	1.58e-21	0.06	14	15	7.13e-23	0.06	19	20	1.49e-21	0.07
sppe (2)	12	13	1.33e-22	0.04	13	14	3.12e-23	0.04	18	19	4.73e-22	0.06
tobin (1)	13	15	9.90e-23	0.07	18	20	3.46e-22	0.11	14	17	7.01e-24	0.12
tobin (2)	14	15	8.95e-23	0.11	13	15	3.79e-23	0.11	16	18	2.28e-23	0.12

the performance of Algorithm 4.1 on three specific p, i.e., $p = 1.1$, $p = 2$, and $p = 5$, for almost all test problems from MCPLIB [1]. By doing this, we intend to examine from these numerical results whether the conclusion implied by *bertsekas* (3) and *freebert* (5) holds true for other test problems. On the other hand, we wish to examine if the FB NCP-function ϕ_2 is the best.

Computational results are summarized in Tables 1–3. In these tables, the first column lists the name of the problems and the number of the starting point in MCPLIB, *NF* indicates the number of function evaluations of G_p for solving

each problem, *Iter* denotes the number of iteration, $G_p(z^f)$ represents the function value of G_p at the final iterate z^f, and CPU records the CPU time in second for solving each problem.

The results listed in Tables 1–3 indicate that the regularization semismooth Newton algorithms based on the generalized FB functions $\phi_{1.1}$, ϕ_2 and ϕ_5 work well and are able to solve almost all complementarity problems from MCPLIB. More specifically, two problems (*scarfanum* and *billups*) fail for $\phi_{1.1}$ and one problem (*scarfanum*) fails for ϕ_2 and ϕ_5. Among others, when solving *billups* by the algorithm associated with ϕ_2, we used $\bar{\varepsilon} = 0.5$, and when solving (*scarfasum* (1)) and (*scarfasum* (2)) by the algorithm associated with $\phi_{1.1}$, and solving (*scarfasum* (3)) by the algorithm with ϕ_5, we used $\bar{\varepsilon} = 0.3$. From Tables 1–3, it is not hard to see that the semismooth Newton algorithm associated with $p = 1.1$ requires less iteration and function evaluation than the ones with $p = 2$ and $p = 5$ for almost all test problems, whereas the algorithm associated with $p = 2$ requires less iteration and function evaluation than the one with $p = 5$ for most of test problems. This implies that the regularization semismooth Newton algorithm associated with a smaller p, for example $p \in [1.1, 2]$, has better numerical behavior, and the generalized FB NCP-functions ϕ_p with $p \in [1.1, 2)$ can be used as the substitutions for the FB NCP function ϕ_2. Notice that the value of p can not be too small since the semismooth Newton algorithm will have worse robustness for those p.

6. Conclusions

In this paper, we have considered a regularization semismooth Newton method based on the generalized FB NCP-functions ϕ_p with $p > 1$ for the P_0 nonlinear complementarity problems. The global convergence and local superlinear (quadratic) convergence results are established by easy extensions of existing arguments in the regularization method [30]. Our main concern is on the numerical side. The numerical results show that the algorithm associated with a smaller p usually has better numerical behavior in terms of the number of iteration and function evaluations. Of course, the value of p cannot be too small since the algorithm will have worse robustness when p approaches to 1. In addition, the numerical results with the algorithm based on the three specific NCP-functions $\phi_{1.1}$, ϕ_2 and ϕ_5 indicate that the algorithm associated with $\phi_{1.1}$ requires less iteration and function evaluation for almost all test problems, especially for those difficult problems such as *bertsekas*, *colvdual*, *pgvon106*, and the generalized FB NCP-functions ϕ_p with $p \in [1.1, 2)$ can be used as the substitutions of the FB NCP-function ϕ_2.

Acknowledgments

The authors would like to thank two anonymous referees for their helpful comments and suggestions on the revision of this paper.

References

[1] S.C. Billups, S.P. Dirkse, M.C. Soares, A comparison of algorithms for large scale mixed complementarity problems, Comput. Optim. Appl. 7 (1997) 3–25.
[2] J.-S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, J. Global Optim. 36 (2006) 565–580.
[3] J.T. Chen, K.H. Chen, Analytical study and numerical experiments for Laplace equation with overspecified boundary conditions, Appl. Math. Modelling 22 (1998) 703–725.
[4] L.Y. Chen, J.T. Chen, H.K. Hong, C.L. Chen, Application of Cesaro mean and the *L*-curve for the deconvolution problem, Soil Dynamics Earthquake Eng. 14 (1995) 361–373.
[5] C. Chen, O.L. Magasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, Comput. Optim. Appl. 5 (1996) 97–138.
[6] J.-S. Chen, S. Pan, A family of NCP functions and a descent method for the nonlinear complementarity problem, Comput. Optim. Appl. 2006, to appear.
[7] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[8] R.W. Cottle, J.-S. Pang, R.-E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
[9] F. Facchinei, C. Kanzow, Beyond monotonicity in regularization methods for the nonlinear complementarity problems, SIAM J. Optim. 37 (1999) 1150–1161.
[10] F. Facchinei, F. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I and II, Springer, New York, 2003.
[11] F. Facchinei, J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, SIAM J. Optim. 7 (1997) 225–247.
[12] M.C. Ferris, O.L. Mangasarian, J.-S. Pang, editors, Complementarity: Applications, Algorithms and Extensions, Kluwer Academic Publishers, Dordrecht, 2001.

[13] A. Fischer, A special Newton-type optimization methods, Optimization 24 (1992) 269–284.
[14] A. Fischer, Solution of the monotone complementarity problem with locally Lipschitzian functions, Math. Programming 76 (1997) 513–532.
[15] M. Fukushima, Merit functions for variational inequality and complementarity problems, in: G. Di Pillo, F. Giannessi (Eds.), Nonlinear Optimization and Applications, Plenum Press, New York, 1996, pp. 155–170.
[16] C. Geiger, C. Kanzow, On the resolution of monotone complementarity problems, Comput. Optim. Appl. 5 (1996) 155–173.
[17] P.T. Harker, J.-S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: A survey of theory, algorithms and applications, Math. Programming 48 (1990) 161–220.
[18] K. Hotta, A. Yoshise, Global convergence of a noninterior-point algorithm using Chen-Harker-Kanzow functions for nonlinear complementarity problems, Math. Programming 86 (1999) 105–133.
[19] Z.-H. Huang, Locating a maximally complementary solutions of the monotone NCP by using non-interior-point smoothing algorithms, Math. Methods Oper. Res. 61 (2005) 41–55.
[20] H. Jiang, L. Qi, A new nonsmooth equation approach to nonlinear complementarity problems, SIAM J. Control Optim. 35 (1997) 178–193.
[21] C. Kanzow, Nonlinear complementarity as unconstrained optimization, J. Optim. Theory Appl. 88 (1996) 139–155.
[22] O.L. Mangasarian, Equivalence of the complementarity problem to a system of nonlinear equations, SIAM J. Appl. Math. 31 (1976) 89–92.
[23] O.L. Mangasarian, M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, Math. Programming 62 (1993) 277–297.
[24] J.-S. Pang, Complementarity problems, in: R. Horst, P. Pardalos (Eds.), Handbook of Global Optimization, Kluwer Academic Publishers, Boston, MA, 1994, pp. 271–338.
[25] S. Pan, J.-S. Chen, A family of penalized NCP functions based on the generalized Fischer–Burmeister function: theoretical investigation and numerical results, Manuscript, 2007.
[26] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993) 227–244.
[27] H.-D. Qi, L.-Z. Liao, A smoothing Newton method for general nonlinear complementarity problems, Comput. Optim. Appl. 17 (2000) 231–253.
[28] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Programming 58 (1993) 353–367.
[29] P. Subramanian, A note on least two norm solutions of monotone complementarity problems, Appl. Math. Lett. 1 (1988) 395–397.
[30] D. Sun, A regularization Newton method for solving nonlinear complementarity problems, Appl. Math. Optim. 40 (1999) 315–339.
[31] D. Sun, L.-Q. Qi, On NCP-functions, Comput. Optim. Appl. 13 (1999) 201–220.
[32] P. Tseng, Analysis of a noninterior continuation method based on Chen-Magasarian smoothing function for complementarity problems, in: M. Fukushima, L. Qi (Eds.), Reformulation: Nonsmooth, Piecewise-Smooth, Semismooth, and Smoothing Methods, Kluwer Academic Publishers, Boston, 1998, pp. 381–404.
[33] N. Yamashita, M. Fukushima, On stationary points of the implicit Lagrangian for the nonlinear complementarity problems, J. Optim. Theory Appl. 84 (1995) 653–663.
[34] N. Yamashita, M. Fukushima, Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems, Math. Programming 76 (1997) 469–491.
[35] H.-C. Zhang, W.W. Hager, A nonmontone line search technique and its application to unconstrained optimization, SIAM J. Optim. 14 (2004) 1043–1056.