Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, pp. 421–441
Edited by M. Fukushima and L. Qi

Numerical Experiments for a Class of Squared Smoothing Newton Methods for Box Constrained Variational Inequality Problems [1]

Guanglu Zhou[†], Defeng Sun[†] and Liqun Qi[†]

Abstract In this paper we present a class of squared smoothing Newton methods for the box constrained variational inequality problem. This class of squared smoothing Newton methods is a regularized version of the class of smoothing Newton methods proposed in [25]. We tested all the test problem collections of GAMSLIB and MCPLIB with all available starting points. Numerical results indicate that these squared smoothing Newton methods are extremely robust and promising.

Key Words variational inequality problem, smoothing approximation, smoothing Newton method, regularization method, convergence.

1 INTRODUCTION

Consider the box constrained variational inequality problem (BVIP for short): Find $y^* \in X = \{y \in \Re^n |\ a \le y \le b\}$, where $a \in \{\Re \cup \{-\infty\}\}^n$, $b \in \{\Re \cup \{\infty\}\}^n$ and $a < b$, such that

$$(y - y^*)^T F(y^*) \ge 0 \quad \text{for all } y \in X, \tag{1.1}$$

where $F : D \to \Re^n$ is a continuously differentiable function on some open set D, which contains X. When $X = \Re^n_+$, BVIP reduces to the nonlinear complementarity problem: Find $y^* \in \Re^n_+$ such that

$$F(y^*) \in \Re^n_+ \quad \text{and} \quad F(y^*)^T y^* = 0. \tag{1.2}$$

[1]This work is supported by the Australian Research Council.
[†]School of Mathematics, The University of New South Wales, Sydney 2052, Australia. E-mail: zhou@maths.unsw.edu.au, sun@maths.unsw.edu.au, L.Qi@unsw.edu.au

Let Π_X be the projection operator on X. It is well known that solving BVIP is equivalent to solving the following Robinson's normal equation

$$E(x) := F(\Pi_X(x)) + x - \Pi_X(x) = 0 \tag{1.3}$$

in the sense that if $x^* \in \Re^n$ is a solution of (1.3) then $y^* := \Pi_X(x^*)$ is a solution of (1.1), and conversely if y^* is a solution of (1.1) then $x^* := y^* - F(y^*)$ is a solution of (1.3) [27]. Let $N := \{1, 2, \cdots, n\}$ and

$$\begin{array}{ll} I_\infty & = \{i \in N \mid a_i = -\infty \text{ and } b_i = +\infty\}, \\ I_{ab} & = \{i \in N \mid a_i > -\infty \text{ and } b_i < +\infty\}, \\ I_a & = \{i \in N \mid a_i > -\infty \text{ and } b_i = +\infty\}, \\ I_b & = \{i \in N \mid a_i = -\infty \text{ and } b_i < +\infty\}. \end{array}$$

Define

$$W(x) := F(\Pi_X(x)) + x - \Pi_X(x) + \alpha T(x), \tag{1.4}$$

where $\alpha \geq 0$ and $T : \Re^n \to \Re^n$ is defined by

$$T_i(x) = \begin{cases} 0 & \text{if } i \in I_\infty \\ 0 & \text{if } i \in I_{ab} \\ [(\Pi_X(x))_i - a_i][F_i(\Pi_X(x))]_+ & \text{if } i \in I_a \\ [(\Pi_X(x))_i - b_i][-F_i(\Pi_X(x))]_+ & \text{if } i \in I_b \end{cases}, \quad i \in N.$$

Properties of $W(x)$ have been studied in [30] in the case that $a_i = 0$ and $b_i = +\infty$ for all $i \in N$.

We can easily prove the following lemma.

Lemma 1.1 $E(x) = 0$ *if and only if* $W(x) = 0$.

By using the Gabriel-Moré smoothing function for $\Pi_X(\cdot)$, we can construct approximations for $W(\cdot)$:

$$G(u, x) := M(u, x) + \alpha S(u, x), \quad (u, x) \in \Re^n \times \Re^n, \tag{1.5}$$

where $M(u, x) := F(p(u, x)) + x - p(u, x)$ and $S : \Re^{2n} \to \Re^n$ is defined by

$$S_i(u, x) = \begin{cases} 0 & \text{if } i \in I_\infty \\ 0 & \text{if } i \in I_{ab} \\ [p_i(u, x) - a_i][p_i(u, F(p(u, x)) + a) - a_i] & \text{if } i \in I_a \\ [p_i(u, x) - b_i][b_i - p_i(u, F(p(u, x)) + b)] & \text{if } i \in I_b \end{cases}, \quad i \in N,$$

where $p(u, x)$ was defined in [25] and will be reviewed in the next section. We note that for any $(u, x) \in \Re^n \times \Re^n$, $p(u, x) \in X$ [25]. So we can assume that F has definition on X only in order that $G(\cdot)$ has definition on $\Re^n \times \Re^n$. This is a very nice feature.

Recently, smoothing Newton methods have attracted a lot of attention in the literature partially due to their superior numerical performance [1], e.g., see [2, 3, 5, 6, 7, 8, 9, 10, 16, 23, 25, 31] and references therein. Among them the first globally and superlinearly (quadratically) convergent smoothing Newton method was proposed by Chen, Qi and Sun in [9]. The result of [9] has been further investigated by Chen and Ye [10]. But they all assumed that F had definition on the whole space $\Re^n$. Qi, Sun and Zhou in [25] avoided this requirement by making use of the mapping $M(\cdot)$ and used one smoothing approximation function instead of using an infinite sequence of those functions.

Regularization methods for solving monotone complementarity problems have been studied by several authors [4, 12, 20, 26, 29]. Facchinei and Kanzow [12] replaced the monotonicity assumption by a P_0-function condition and showed that many properties of regularization methods still hold for this larger class of problems. Sun [29] proposed a regularization smoothing Newton method for solving nonlinear complementarity problem under the assumption that F is a P_0-function and obtained some stronger results for monotone complementarity problems. H.-D. Qi [20] proposed a regularized smoothing Newton method for the nonlinear complementarity problem and the box constrained variational inequality problem by using the developments on regularization methods and smoothing Newton methods. The global convergence of this method was proved under the assumption that F is a P_0-function and the solution set of the problem (1.1) is nonempty and bounded.

In this paper we propose a class of squared smoothing Newton methods for the box constrained variational inequality problem and present the numerical results of this class of methods. This class of squared smoothing Newton methods is a regularized version of the class of smoothing Newton methods proposed in [25]. In the next section we will give some definitions. This class of squared smoothing Newton methods will be proposed in section 3. In section 4 we will report numerical results of these methods. We then make some final remarks in section 5.

To ease our discussion, we introduce some notation here: If $u \in \Re^n$, $\mathrm{diag}(u)$ is the diagonal matrix whose i-th diagonal element is u_i. For a continuously differentiable function $\Phi : \Re^m \to \Re^m$, we denote the Jacobian of Φ at $x \in \Re^m$ by $\Phi'(x)$, whereas the transposed Jacobian as $\nabla\Phi(x)$. $\|\cdot\|$ denotes the Euclidean norm. If X is a subset in $\Re^n$, we denote by $intX$ the interior of X. If V is an $m \times m$ matrix with entries V_{jk}, $j, k = 1, \ldots, m$, and $\mathcal{J}$ and $\mathcal{K}$ are index sets such that $\mathcal{J}, \mathcal{K} \subseteq \{1, \ldots, m\}$, we denote by $V_{\mathcal{J}\mathcal{K}}$ the $|\mathcal{J}| \times |\mathcal{K}|$ sub-matrix of V consisting of entries V_{jk}, $j \in \mathcal{J}$, $k \in \mathcal{K}$. If $V_{\mathcal{J}\mathcal{J}}$ is nonsingular, we denote by $V/V_{\mathcal{J}\mathcal{J}}$ the Schur-complement of $V_{\mathcal{J}\mathcal{J}}$ in V, i.e., $V/V_{\mathcal{J}\mathcal{J}} := V_{\mathcal{K}\mathcal{K}} - V_{\mathcal{K}\mathcal{J}} V_{\mathcal{J}\mathcal{J}}^{-1} V_{\mathcal{J}\mathcal{K}}$, where $\mathcal{K} = \{1, \ldots, m\} \backslash \mathcal{J}$.

2 SOME PRELIMINARIES

We first restate some definitions.

Definition 2.1 *A matrix $V \in \Re^{n\times n}$ is called a*

(a) P_0-matrix if, for every $x \in \Re^n$ with $x \neq 0$, there is an index $i_0 = i_0(x)$ with

$$x_{i_0} \neq 0 \quad and \quad x_{i_0}[Vx]_{i_0} \geq 0;$$

(b) P-matrix if, for every $x \in \Re^n$ with $x \neq 0$, it holds that

$$\max_i x_i[Vx]_i > 0.$$

Definition 2.2 *A function $F : D \to \Re^n$, $D \subseteq \Re^n$, is called a*

(a) P_0-function on D if, for all $x, y \in D$ with $x \neq y$, there is an index $i_0 = i_0(x, y)$ with

$$x_{i_0} \neq y_{i_0} \quad \text{and} \quad (x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] \geq 0;$$

(b) P-function on D if, for all $x, y \in D$ with $x \neq y$, it holds that

$$\max_i (x_i - y_i)[F_i(x) - F_i(y)] > 0;$$

(c) uniform P-function on D if there is a constant $\mu > 0$ such that

$$\max_i (x_i - y_i)[F_i(x) - F_i(y)] \geq \mu \|x - y\|^2$$

holds for all $x, y \in D$.

Definition 2.3 *A function $F : D \to \Re^n$, $D \subseteq \Re^n$, is called a*

(a) monotone function on D if, for all $x, y \in D$ with $x \neq y$,

$$(x - y)^T[F(x) - F(y)] \geq 0;$$

(b) strictly monotone function on D if, for all $x, y \in D$ with $x \neq y$,

$$(x - y)^T[F(x) - F(y)] > 0.$$

It is known that every uniform P-function is P-function and every P-function is a P_0-function. Moreover, the Jacobian of a continuously differentiable P_0-function (uniform P-function) at a point is a P_0-matrix (P-matrix).

We now restate the definition of $p(u, x)$, $(u, x) \in \Re^n \times \Re^n$, given in [25]. For each $i \in N$, $p_i(u, x) = q(u_i, a_i, b_i, x_i)$ and for any $(\mu, c, d, w) \in \Re \times \{\Re \cup \{-\infty\}\} \times \{\Re \cup \{\infty\}\} \times \Re$ with $c \leq d$, $q(\mu, c, d, w)$ is defined by

$$q(\mu, c, d, w) = \begin{cases} \phi(|\mu|, c, d, w) & \text{if } \mu \neq 0 \\ \Pi_{[c,d]\cap\Re}(w) & \text{if } \mu = 0 \end{cases}, \tag{2.1}$$

and $\phi(\mu, c, d, w)$, $(\mu, w) \in \Re_{++} \times \Re$ is a Gabriel-Moré smoothing approximation function [14]. The definition of $\phi(\cdot)$ is as follows: Let $\rho : \Re \to \Re_+$ be a density function, i.e., $\rho(s) \geq 0$ and $\int_{-\infty}^{\infty} \rho(s)ds = 1$, with a bounded absolute mean, that is

$$\kappa := \int_{-\infty}^{\infty} |s|\rho(s)ds < \infty. \tag{2.2}$$

For any three numbers $c \in \Re \cup \{-\infty\}$, $d \in \Re \cup \{\infty\}$ with $c \leq d$ and $e \in \Re$, the median function $\text{mid}(\cdot)$ is defined by

$$\text{mid}(c, d, e) = \Pi_{[c,d]\cap\Re}(e) = \begin{cases} c & \text{if } e < c \\ e & \text{if } c \leq e \leq d \\ d & \text{if } d < e \end{cases}.$$

Then the Gabriel-Moré smoothing function $\phi(\mu, c, d, w)$ for $\Pi_{[c,d]\cap\Re}(w)$ [14] is defined by

$$\phi(\mu, c, d, w) = \int_{-\infty}^{\infty} \text{mid}(c, d, w - \mu s)\rho(s)ds, \quad (\mu, w) \in \Re_{++} \times \Re. \tag{2.3}$$

If $c = -\infty$ and/or $d = \infty$, the value of ϕ takes the limit of ϕ as $c \to -\infty$ and/or $d \to \infty$, correspondingly. For example, if c is finite and $d = \infty$, then

$$\phi(\mu, c, \infty, w) = \lim_{d' \to \infty} \phi(u, c, d', w), \quad (u, w) \in \Re_{++} \times \Re.$$

For the sake of convenience, let $\phi_{cd} : \Re_{++} \times \Re \to \Re$ be defined by

$$\phi_{cd}(\mu, w) := \phi(\mu, c, d, w), \quad (\mu, w) \in \Re_{++} \times \Re \tag{2.4}$$

and for any given $\mu \in \Re_{++}$, let $\phi_{\mu cd} : \Re \to \Re$ be defined by

$$\phi_{\mu cd}(w) := \phi(\mu, c, d, w), \quad w \in \Re. \tag{2.5}$$

Lemma 2.1 *[14, Lemma 2.3] For any given $\mu > 0$, the mapping $\phi_{\mu cd}(\cdot)$ is continuously differentiable with*

$$\phi'_{\mu cd}(w) = \int_{(w-d)/\mu}^{(w-c)/\mu} \rho(s)ds,$$

where $\phi_{\mu cd}(\cdot)$ is defined by (2.5). In particular, $\phi'_{\mu cd}(w) \in [0, 1]$.

Lemma 2.2 *[25, Lemma 2.2] The mapping $\phi_{cd}(\cdot)$ defined by (2.4) is Lipschitz continuous on $\Re_{++} \times \Re$ with Lipschitz constant $L := 2\max\{1, \kappa\}$.*

Let $q_{cd} : \Re^2 \to \Re$ be defined by

$$q_{cd}(\mu, w) = q(\mu, c, d, w), \quad (\mu, w) \in \Re^2, \tag{2.6}$$

where $q(\mu, c, d, w)$ is defined by (2.1).

Lemma 2.3 *[25, Lemma 2.3] The mapping $q_{cd}(\cdot)$ is globally Lipschitz continuous on $\Re^2$ with the same Lipschitz constant as in Lemma 2.2.*

Some most often used Gabriel-Moré smoothing functions like the neural networks smoothing function, the Chen-Harker-Kanzow-Smale smoothing function and the uniform smoothing function are discussed in [25].

In this paper, unless otherwise stated, we always assume that $c \in \Re \cup \{-\infty\}$, $d \in \Re \cup \{\infty\}$ and $c \leq d$. By Lemma 2.2 of [14], for any $(\mu, w) \in \Re_{++} \times \Re$,

$$\phi(\mu, c, d, w) \in [c, d] \cap \Re,$$

and so, for any $(u, x) \in \Re^n \times \Re^n$,

$$p(u, x) \in X. \tag{2.7}$$

Then the mapping $G(\cdot)$ defined in (1.5) is well defined on $\Re^{2n}$ while $F(\cdot)$ is only required to have definition on X, the feasible region.

From Lemma 2.1, for any given $\mu \in \Re_{++}$, $\phi_{\mu cd}(\cdot)$ is continuously differentiable at any $w \in \Re$. Moreover, for several most often used Gabriel-Moré smoothing functions it can be verified that $\phi_{cd}(\cdot)$ is also continuously differentiable at any $(\mu, w) \in \Re_{++} \times \Re$. In this paper, we are interested in smoothing functions with this property, which we make as an assumption.

Assumption 2.1 *The function $\phi_{cd}(\cdot)$ is continuously differentiable at any $(\mu, w) \in \Re_{++} \times \Re$.*

Let $z := (u, x) \in \Re^n \times \Re^n$. For some $\lambda \geq 0$, define $\bar{G} : \Re^{2n} \to \Re^n$ by

$$\bar{G}_i(z) = G_i(z) + \lambda u_i p_i(u, x), \ i \in N,$$

where $G(\cdot)$ is defined in (1.5). Define $H : \Re^{2n} \to \Re^{2n}$ by

$$H(z) := \begin{pmatrix} u \\ \bar{G}(z) \end{pmatrix}. \tag{2.8}$$

Then H is continuously differentiable at any $z = (u, x) \in \Re^n_{++} \times \Re^n$ if Assumption 2.1 is satisfied.

For any $u \in \Re^n_{++}$ and $x \in \Re^n$, define $c(u, x), d(u, x) \in \Re^n$ by

$$c_i(u, x) = \partial p_i(u, x)/\partial x_i, \ d_i(u, x) = \partial p_i(u, x)/\partial u_i, \ i \in N.$$

For any $u \in \Re^n_{++}$ and $x \in \Re^n$, define $D^U(u, x)$, $C^X(u, x)$, $P^N(u, x)$, $P^U(u, x)$, $P^X(u, x) \in \Re^n$ by

$$D_i^U(u, x) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ d_i(u, F(p(u, x)) + a) & \text{if } i \in I_a \\ d_i(u, F(p(u, x)) + b) & \text{if } i \in I_b \end{cases},$$

$$C_i^X(u, x) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ c_i(u, F(p(u, x)) + a) & \text{if } i \in I_a \\ c_i(u, F(p(u, x)) + b) & \text{if } i \in I_b \end{cases},$$

$$P_i^N(u, x) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ p_i(u, x) - a_i & \text{if } i \in I_a \\ b_i - p_i(u, x) & \text{if } i \in I_b \end{cases},$$

$$P_i^U(u, x) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ d_i(u, x)[p_i(u, F(p(u, x)) + a) - a_i] & \text{if } i \in I_a \\ d_i(u, x)[b_i - p_i(u, F(p(u, x)) + b)] & \text{if } i \in I_b \end{cases}$$

and

$$P_i^X(u,x) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ c_i(u,x)[p_i(u,F(p(u,x))+a) - a_i] & \text{if } i \in I_a \\ c_i(u,x)[b_i - p_i(u,F(p(u,x))+b)] & \text{if } i \in I_b \end{cases}, \quad i \in N.$$

Theorem 2.1 *Suppose that Assumption 2.1 holds for a chosen smoothing function* $\phi(\mu,c,d,w)$, $(\mu,w) \in \Re_{++} \times \Re$. *Then*

(i) *The mapping* $H(\cdot)$ *is continuously differentiable at any* $z = (u,x) \in \Re^n_{++} \times \Re^n$ *and*

$$H'(z) = \begin{pmatrix} I & 0 \\ \tilde{G}'_u(z) & \tilde{G}'_x(z) \end{pmatrix}, \tag{2.9}$$

where

$$\begin{aligned} \tilde{G}'_u(z) &= [F'(p(z)) - I + \lambda \mathrm{diag}(u)]\mathrm{diag}(d(u,x)) + \lambda \mathrm{diag}(p(z)) \\ &\quad + \alpha \mathrm{diag}(P^U(u,x)) + \alpha \mathrm{diag}(P^N(u,x))\mathrm{diag}(D^U(u,x)) \\ &\quad + \alpha \mathrm{diag}(P^N(u,x))\mathrm{diag}(C^X(u,x))F'(p(z))\mathrm{diag}(d(u,x)), \end{aligned}$$

$$\begin{aligned} \tilde{G}'_x(z) &= \{[I + \alpha \mathrm{diag}(P^N(u,x))\mathrm{diag}(C^X(u,x))]F'(p(z)) \\ &\quad + \lambda \mathrm{diag}(u)\}\,\mathrm{diag}(c(u,x)) \\ &\quad + I - \mathrm{diag}(c(u,x)) + \alpha \mathrm{diag}(P^X(u,x)) \end{aligned}$$

and for each $i \in N$, $c_i(u,x) \in [0,1]$.

(ii) *If* $\lambda > 0$ *and for some* $z \in \Re^n_{++} \times \Re^n$, $F'(p(z))$ *is a* P_0*-matrix, then* $H'(z)$ *is nonsingular.*

Proof. (i) Since Assumption 2.1 is satisfied for $\phi(\cdot)$, from the definition, we know that $H(\cdot)$ is continuously differentiable at any $z = (u,x) \in \Re^n_{++} \times \Re^n$. By direct computation we have (2.9). From Lemma 2.1 and the definition of $p_i(\cdot)$, $c_i(u,x) \in [0,1]$, $i \in N$.
(ii) Suppose that $\lambda > 0$ and for some $z \in \Re^n_{++} \times \Re^n$, $F'(p(z))$ is a P_0-matrix. From (i) and the definition of $C^X(u,x)$, $P^N(u,x)$ and $P^X(u,x)$, we have $C_i^X(u,x) \in [0,1]$, $P_i^N(u,x) \geq 0$ and $P_i^X(u,x) \geq 0$, for $i \in N$. Then $Q = I + \alpha \mathrm{diag}(P^N(u,x))\mathrm{diag}(C^X(u,x))$ is a positive diagonal matrix. So $QF'(p(z))$ is a P_0-matrix and $QF'(p(z)) + \lambda \mathrm{diag}(u)$ is a P-matrix. From [7, Lemma 2] we have that $\tilde{G}'_x(z)$ is nonsingular. It then follows from (2.9) that $H'(z)$ is also nonsingular. ∎

In order to design high-order convergent Newton methods we need the concept of semismoothness. Semismoothness was originally introduced by Mifflin [19] for functionals. Convex functions, smooth functions, and piecewise linear functions are examples of semismooth functions. The composition of semismooth functions is still a semismooth function [19]. In [24], Qi and J. Sun extended the definition of semismooth functions to $\Phi : \Re^{m_1} \to \Re^{m_2}$. A locally Lipschitz continuous vector valued function $\Phi : \Re^{m_1} \to \Re^{m_2}$ has a generalized Jacobian $\partial\Phi(x)$ as in Clarke [11]. Φ is said to be *semismooth* at $x \in \Re^{m_1}$, if

$$\lim_{\substack{V \in \partial\Phi(x+th') \\ h' \to h,\ t \downarrow 0}} \{Vh'\} \tag{2.10}$$

exists for any $h \in \Re^{m_1}$. It has been proved in [24] that Φ is semismooth at x if and only if all its component functions are. Also, $\Phi'(x;h)$, the directional derivative of Φ at x in the direction h, exists and equals the limit in (2.10) for any $h \in \Re^{m_1}$ if Φ is semismooth at x.

Lemma 2.4 *[24] Suppose that $\Phi : \Re^{m_1} \to \Re^{m_2}$ is a locally Lipschitzian function and semismooth at x. Then*

(i) *for any $V \in \partial\Phi(x+h)$, $h \to 0$,*

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

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

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

The following lemma is extracted from Theorem 2.3 of [24].

Lemma 2.5 *Suppose that $\Phi : \Re^{m_1} \to \Re^{m_2}$ is a locally Lipschitzian function. Then the following two statements are equivalent:*

(i) *$\Phi(\cdot)$ is semismooth at x.*

(ii) *Φ is directionally differentiable at x, and for any $V \in \partial\Phi(x+h)$, $h \to 0$,*

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

A stronger notion than semismoothness is strong semismoothness. $\Phi(\cdot)$ is said to be *strongly semismooth* at x if Φ is semismooth at x and for any $V \in \partial\Phi(x+h)$, $h \to 0$,

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

(Note that in [22] and [24] different names for strong semismoothness are used.) A function Φ is said to be a (strongly) semismooth function if it is (strongly) semismooth everywhere.

Recall that from Lemma 2.3 the function $q_{cd}(\cdot)$ defined by (2.6) is globally Lipschitz continuous on $\Re^2$. Then, from Lemma 2.5 and the definition of strong semismoothness, we can prove in the above mentioned three usual cases [25] that $q_{cd}(\cdot)$ is strongly semismooth at $x \in \Re^2$, i.e., for any $V \in \partial q_{cd}(x+h)$, $h \to 0$,

$$Vh - q'_{cd}(x;h) = O(\|h\|^2). \tag{2.11}$$

3 A CLASS OF SQUARED SMOOTHING NEWTON METHODS

Choose $\bar{u} \in \Re^n_{++}$ and $\gamma \in (0,1)$ such that $\gamma\|\bar{u}\| < 1$. Let $\bar{z} := (\bar{u}, 0) \in \Re^n \times \Re^n$. Define the merit function $\psi : \Re^{2n} \to \Re_+$ by

$$\psi(z) := \|H(z)\|^2$$

and define $\beta : \Re^{2n} \to \Re_+$ by

$$\beta(z) := \gamma \min\{1, \psi(z)\}.$$

Let

$$\Omega := \{z = (u, x) \in \Re^n \times \Re^n |\ u \geq \beta(z)\bar{u}\}.$$

Then, because for any $z \in \Re^{2n}$, $\beta(z) \leq \gamma < 1$, it follows that for any $x \in \Re^n$,

$$(\bar{u}, x) \in \Omega.$$

Proposition 3.1 *The following relations hold:*

(i) $H(z) = 0 \Longleftrightarrow \beta(z) = 0 \Longleftrightarrow H(z) = \beta(z)\bar{z}$.

(ii) $H(z) = 0 \Longrightarrow u = 0$ *and* $y = \Pi_X(x)$ *is a solution of (1.1).*

(iii) *If* $x = y - F(y)$, *where* y *is a solution of (1.1), then* $H(0, x) = 0$.

The proof of this proposition is similar to that of Proposition 4.1 in [25], so we omit it.

Algorithm 3.1

Step 0. *Choose constants* $\delta \in (0,1)$, $\sigma \in (0, 1/2)$, $\alpha \geq 0$ *and* $\lambda \geq 0$. *Let* $u^0 := \bar{u}$, $x^0 \in \Re^n$ *be an arbitrary point and* $k := 0$.

Step 1. *If* $H(z^k) = 0$ *then stop. Otherwise, let* $\beta_k := \beta(z^k)$.

Step 2. *Compute* $\Delta z^k := (\Delta u^k, \Delta x^k) \in \Re^n \times \Re^n$ *by*

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

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

$$\psi(z^k + \delta^l \Delta z^k) \leq [1 - 2\sigma(1 - \gamma\|\bar{u}\|)\delta^l]\psi(z^k). \tag{3.2}$$

Define $z^{k+1} := z^k + \delta^{l_k} \Delta z^k$.

Step 4. *Replace* k *by* $k+1$ *and go to Step 1.*

Remark. Algorithm 3.1 is actually the smoothing Newton method proposed in [25] for the case that $\alpha = 0$ and $\lambda = 0$. When $\lambda > 0$, Algorithm 3.1 has better properties than the original version of the smoothing Newton method given in [25]. The parameter α is introduced in order to improve the numerical performance.

Lemma 3.1 *Suppose that F is a P_0-function on X, $\lambda > 0$ and that $\bar{u}, \tilde{u} \in \Re^n$ are two positive vectors such that $\bar{u} \geq \tilde{u} > 0$. Then for any sequence $\{z^k = (u^k, x^k)\}$ such that $\tilde{u} \leq u^k \leq \bar{u}$ and $\|x^k\| \to +\infty$ we have*

$$\lim_{k\to\infty} \psi(z^k) = +\infty. \tag{3.3}$$

Proof. For the sake of contradiction, suppose that there exists a sequence $\{z^k = (u^k, x^k) \in \Re^n \times \Re^n\}$ such that $\tilde{u} \leq u^k \leq \bar{u}$, $\|x^k\| \to \infty$ and $\psi(z^k)$ is bounded. It is easy to prove that

$$|\text{mid}(a_i, b_i, x_i^k)| \to \infty \Longrightarrow |x_i^k| \to \infty \text{ and } |x_i^k - \text{mid}(a_i, b_i, x_i^k)| \to 0, \quad i \in N. \tag{3.4}$$

From Lemma 2.3 and the definition of $p(\cdot)$, there exists a constant $L' > 0$ such that

$$|p_i(u^k, x^k) - \text{mid}(a_i, b_i, x_i^k)| \leq L'|u_i^k|, \quad i \in N. \tag{3.5}$$

From (3.4) and (3.5) we have

$$|p_i(u^k, x^k)| \to \infty \Longrightarrow \{|x_i^k - p_i(u^k, x^k)|\} \text{ is bounded.} \tag{3.6}$$

Define the index set J by $J := \{i \in N \mid \{p_i(u^k, x^k)\} \text{ is unbounded}\}$. Then it follows that $J \neq \emptyset$ because otherwise $\|\bar{G}(z^k)\| = \|F(p(z^k)) + x^k - p(z^k) + \alpha S(z^k) + \lambda \text{diag}(u^k) p(z^k)\| \to \infty$. Let $\bar{z}^k = (\bar{u}^k, \bar{x}^k) \in \Re^n \times \Re^n$ be defined by

$$\bar{u}_i^k = \begin{cases} u_i^k & \text{if } i \notin J \\ 0 & \text{if } i \in J \end{cases}$$

and

$$\bar{x}_i^k = \begin{cases} x_i^k & \text{if } i \notin J \\ 0 & \text{if } i \in J \end{cases}, \quad i \in N.$$

Then

$$p_i(\bar{z}^k) = \begin{cases} p_i(z^k) & \text{if } i \notin J \\ \text{mid}(a_i, b_i, 0) & \text{if } i \in J \end{cases}, \quad i \in N.$$

Hence $\{\|p(\bar{z}^k)\|\}$ is bounded. Because F is a P_0-function on X, we have

$$\begin{aligned} 0 &\leq \max_{i\in N}[p_i(z^k) - p_i(\bar{z}^k)][F_i(p(z^k)) - F_i(p(\bar{z}^k))] \\ &= \max_{i\in J}[p_i(z^k) - p_i(\bar{z}^k)][F_i(p(z^k)) - F_i(p(\bar{z}^k))] \qquad (3.7) \\ &= [p_i(z^k) - p_i(\bar{z}^k)][F_i(p(z^k)) - F_i(p(\bar{z}^k))], \end{aligned}$$

where $i \in J$ is one of indices for which the maximum is attained, without loss of generality, assumed to be independent of k. Since $i \in J$, we have

$$|p_i(z^k)| \to \infty.$$

From (3.7) and the boundedness of $\{F_i(p(\bar{z}^k))\}$, we have that $F_i(p(z^k))$ does not tend to $-\infty$ if $p_i(z^k) \to +\infty$ and $F_i(p(z^k))$ does not tend to $+\infty$ if $p_i(z^k) \to -\infty$.

We now consider two cases.

Case 1: $p_i(z^k) \to +\infty$.
In this case, we have that $i \in I_\infty \cup I_a$. Since $S_i(z^k) \geq 0$, from (3.6) we have

$$\bar{G}_i(z^k) = F_i(p(z^k)) + x_i^k - p_i(z^k) + \alpha S_i(z^k) + \lambda u_i^k p_i(z^k) \to +\infty.$$

Case 2: $p_i(z^k) \to -\infty$.
In this case, $i \in I_\infty \cup I_b$. Since $S_i(z^k) \leq 0$, from (3.6) we have

$$\bar{G}_i(z^k) = F_i(p(z^k)) + x_i^k - p_i(z^k) + \alpha S_i(z^k) + \lambda u_i^k p_i(z^k) \to -\infty.$$

In either case we get $\psi(z^k) \to +\infty$, which is a contradiction. So we complete our proof. ∎

Remark. Lemma 3.1 is not true if $\tilde{u} = 0$ even if F is strictly monotone. To see this, we may consider the function $F(x) = e^x - 1$, $x \in \Re$. This function was provided by H.-D. Qi [21]. Suppose that $X = \Re$. Then $\psi(z) = u^2 + F(x)^2$ and when $x \to -\infty$ and $u = 0$, $\psi(z) \to 1$. This clearly shows that ψ may have unbounded level sets. However, if the solution set of (1.1) is bounded, we can prove the global convergence of our methods (see Theorem 3.1) under the assumption that F is a P_0-function on X.

Assumption 3.1 (i) *F is a P_0-function on X.*

(ii) *The solution set of the problem (1.1) is nonempty and bounded.*

Theorem 3.1 *Suppose that Assumptions 2.1 and 3.1 are satisfied and $\lambda > 0$. Then the infinite sequence $\{z^k\}$ generated by Algorithm 3.1 is bounded and each accumulation point $\tilde{z}$ of $\{z^k\}$ is a solution of $H(z) = 0$.*

Proof. By combining Lemma 3.1 and the proof of Theorem 4.1 [29] and Theorem 4.6 [20], we can prove this theorem. We omit the details. ∎

Theorem 3.2 *Suppose that Assumptions 2.1 and 3.1 are satisfied, $\lambda > 0$ and z^* is an accumulation point of the infinite sequence $\{z^k\}$ generated by Algorithm 3.1. Suppose that H is semismooth at z^* and that all $V \in \partial H(z^*)$ are nonsingular. Then the whole sequence $\{z^k\}$ converges to z^*,*

$$\|z^{k+1} - z^*\| = o(\|z^k - z^*\|) \tag{3.8}$$

and

$$u_i^{k+1} = o(u_i^k), \; i \in N. \tag{3.9}$$

Furthermore, if H is strongly semismooth at z^*, then

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2) \tag{3.10}$$

and

$$u_i^{k+1} = O\left((u_i^k)^2\right), \ i \in N. \tag{3.11}$$

Proof. See [25, Theorem 7.1] for a similar proof. ∎

Next, we study under what conditions all the matrices $V \in \partial H(z^*)$ are nonsingular at a solution point $z^* = (u^*, x^*) \in \Re^n \times \Re^n$ of $H(z) = 0$. Apparently, $u^* = 0$ and x^* is a solution of (1.3). For convenience of handling notation we denote

$$\mathcal{I} := \{i |\ a_i < x_i^* < b_i \ \&\ F_i(\Pi_X(x^*)) = 0, \ i \in N\},$$

$$\begin{aligned} \mathcal{J} : &= \{i |\ x_i^* = a_i \ \&\ F_i(\Pi_X(x^*)) = 0, \ i \in N\} \\ &\cup \{i |\ x_i^* = b_i \ \&\ F_i(\Pi_X(x^*)) = 0, \ i \in N\} \end{aligned}$$

and

$$\begin{aligned} \mathcal{K} : &= \{i |\ x_i^* < a_i \ \&\ F_i(\Pi_X(x^*)) > 0, \ i \in N\} \\ &\cup \{i |\ x_i^* > b_i \ \&\ F_i(\Pi_X(x^*)) < 0, \ i \in N\}. \end{aligned}$$

Then

$$\mathcal{I} \cup \mathcal{J} \cup \mathcal{K} = N.$$

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

$$\nabla F(\Pi_X(x^*)) = \begin{pmatrix} \nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{I}} & \nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{J}} & \nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{K}} \\ \nabla F(\Pi_X(x^*))_{\mathcal{J}\mathcal{I}} & \nabla F(\Pi_X(x^*))_{\mathcal{J}\mathcal{J}} & \nabla F(\Pi_X(x^*))_{\mathcal{J}\mathcal{K}} \\ \nabla F(\Pi_X(x^*))_{\mathcal{K}\mathcal{I}} & \nabla F(\Pi_X(x^*))_{\mathcal{K}\mathcal{J}} & \nabla F(\Pi_X(x^*))_{\mathcal{K}\mathcal{K}} \end{pmatrix}.$$

BVIP is said to be R-regular at x^* if $\nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{I}}$ is nonsingular and its Schur-complement in the matrix

$$\begin{pmatrix} \nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{I}} & \nabla F(\Pi_X(x^*))_{\mathcal{I}\mathcal{J}} \\ \nabla F(\Pi_X(x^*))_{\mathcal{J}\mathcal{I}} & \nabla F(\Pi_X(x^*))_{\mathcal{J}\mathcal{J}} \end{pmatrix}$$

is a P-matrix, see [28].

Proposition 3.2 *Suppose that $z^* = (u^*, x^*) \in \Re^n \times \Re^n$ is a solution of $H(z) = 0$. If BVIP is R-regular at x^*, then all $V \in \partial H(z^*)$ are nonsingular.*

Proof. It is easy to see that for any $V \in \partial H(z^*)$ there exists a matrix $W = (W_u, W_x) \in \partial \tilde{G}(z^*)$ with $W_u, W_x \in \Re^{n \times n}$ such that

$$V = \begin{pmatrix} I & 0 \\ W_u & W_x \end{pmatrix}.$$

Hence, proving that V is nonsingular is equivalent to proving that W_x is nonsingular. For any $U = (U_u, U_x) \in \partial p(z^*)$, by the definition of p, we have

$$U \in \partial p_1(z^*) \times \partial p_2(z^*) \times \cdots \times \partial p_n(z^*) = \partial p(z^*).$$

Then for each $i \in N$, the ith row of U, $U_i \in \partial p_i(z^*)$. Apparently, from the definition of p and Lemma 2.1,

$$U_x = \text{diag}\{(u_x)_i,\ i \in N\},$$

where $(u_x)_i$ is defined by

$$\begin{cases} (u_x)_i = 1 & \text{if } i \in \mathcal{I} \\ (u_x)_i \in [0,1] & \text{if } i \in \mathcal{J} \\ (u_x)_i = 0 & \text{if } i \in \mathcal{K} \end{cases}.$$

Define $P^N(z^*) \in \Re^n$ by

$$P_i^N(z^*) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ p_i(z^*) - a_i & \text{if } i \in I_a \\ b_i - p_i(z^*) & \text{if } i \in I_b \end{cases}, \quad i \in N.$$

For any $W = (W_u, W_x) \in \partial \bar{G}(z^*)$ with $W_u, W_x \in \Re^{n \times n}$ there exist $U = (U_u, U_x) \in \partial p(z^*)$ and $C^X(z^*) \in \Re^n$ such that

$$\begin{aligned} W_x \;=\; & F'(p(z^*))U_x + I - U_x \\ & +\alpha \text{diag}(P^N(z^*))\text{diag}(C^X(z^*))F'(p(z^*))U_x \\ & +\alpha \text{diag}(P^X(z^*)), \end{aligned}$$

where

$$P_i^X(z^*) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ (u_x)_i[(\Pi_X(F(p(z^*)) + a)_i - a_i] & \text{if } i \in I_a \\ (u_x)_i[b_i - (\Pi_X(F(p(z^*)) + b)_i] & \text{if } i \in I_b \end{cases}, \quad i \in N$$

and

$$C_i^X(z^*) = \begin{cases} 0 & \text{if } i \in I_\infty \cup I_{ab} \\ A_i & \text{if } i \in I_a \\ B_i & \text{if } i \in I_b \end{cases}, \quad i \in N,$$

where

$$A_i = \begin{cases} 1 & \text{if } F_i(\Pi_X(x^*)) > 0 \\ 0 & \text{if } F_i(\Pi_X(x^*)) < 0 \\ \epsilon_1 \in [0,1] & \text{if } F_i(\Pi_X(x^*)) = 0 \end{cases}$$

and

$$B_i = \begin{cases} 0 & \text{if } F_i(\Pi_X(x^*)) > 0 \\ 1 & \text{if } F_i(\Pi_X(x^*)) < 0 \\ \epsilon_2 \in [0,1] & \text{if } F_i(\Pi_X(x^*)) = 0 \end{cases}.$$

Let $D = \alpha \text{diag}(P^N(z^*))\text{diag}(C^X(z^*))$. We have that D is a nonnegative diagonal matrix. By inspecting the structure of P^X we have that $P_i^X = 0$, for all $i \in N$. Then we have

$$W_x = (I + D)F'(p(z^*))U_x + I - U_x.$$

Let Q $= W_x^T(I+D)^{-1}$. Then

$$Q = U_x \nabla F(p(z^*)) + (I - U_x)(I+D)^{-1}.$$

Thus, for each $i \in \mathcal{J}$, there exists $\lambda_i \in [0,1]$ such that

$$Q_i = \begin{cases} \nabla F(p(z^*))_i & \text{if } i \in \mathcal{I} \\ \lambda_i \nabla F(p(z^*))_i + (1-\lambda_i)(1+D_{ii})^{-1} e_i & \text{if } i \in \mathcal{J} \\ (1+D_{ii})^{-1} e_i & \text{if } i \in \mathcal{K} \end{cases},$$

where e_i is the ith unit row vector of $\Re^n$ and $\nabla F(p(z^*))_i$ is the ith row vector of $\nabla F(p(z^*))$, $i \in N$. Then, by [13, Proposition 3.2] we can prove that Q, and so W_x, is nonsingular under the assumption of R-regularity (note that $p(z^*) = \Pi_X(x^*)$). Hence, any $V \in \partial H(z^*)$ is nonsingular. So, we complete our proof. ■

4 NUMERICAL RESULTS

Algorithm 3.1 was implemented in MATLAB and was run on a SUN Sparc Server 3002 for all test problems with all available starting points from the test problem collections GAMSLIB and MCPLIB [1] (note that there are three starting points in `ehl_kost` with the same data and so we only list the results for `ehl_kost` with the first starting point in Table 1.2). Throughout the computational experiments, unless otherwise stated, we chose the Chen-Harker-Kanzow-Smale smoothing function and used the following parameters:

$$\delta = 0.5, \sigma = 10^{-4}, \bar{u} = 0.2e, \gamma = \min\{10^{-5}, 0.2/\|\bar{u}\|\} \text{ and } \lambda = 0.05,$$

where e is the vector of all ones.

To improve the numerical behaviour of Algorithm 3.1, we replaced the standard (monotone) Armijo-rule by a nonmonotone line search as described in Grippo, Lampariello and Lucidi [15], i.e., we computed the smallest nonnegative integer l such that

$$z^k + \delta^l \Delta z^k \in \Omega \tag{4.1}$$

and

$$\psi(z^k + \delta^l \Delta z^k) \le \mathcal{W}_k - 2\sigma(1 - \gamma\|\bar{u}\|)\delta^l \psi(z^k), \tag{4.2}$$

where $\mathcal{W}_k$ is given by

$$\mathcal{W}_k = \max_{j=k-m_k,\ldots,k} \psi(z^j)$$

and where, for given nonnegative integers m and s, we set

$$m_k = 0$$

if $k \le s$ at the kth iteration, whereas we set

$$m_k := \min\{m_{k-1} + 1, m\}$$

at all other iterations. In our implementation, we use

$$m = 8 \quad \text{and} \quad s = 2.$$

We terminated our iteration when one of the following conditions was satisfied

$$k > 3000, R(x^k) := \|p(z^k) - \Pi_X[p(z^k) - F(p(z^k))]\|_\infty \leq 10^{-6} \quad \text{or} \quad ls > 80,$$

where ls was the number of line search at each step.

Using this algorithmic environment, we made some preliminary test runs using different values of the parameter α. In view of these preliminary experiments, it seems that α should be large if the iteration point is far away from a solution and α should be reduced if the iteration point is getting closer to a solution of the problem. This motivated us to use a dynamic choice of α for our test runs. More precisely, we updated α using the following rules:

(a) Set $\alpha = 10^4$ at the beginning of each iteration.

(b) If $R(x^k) < 10$, then set $\alpha = 100$.

(c) If $R(x^k) < 10^{-2}$ or $k \geq 80$, then set $\alpha = 10^{-3}$.

(d) If $R(x^k) < 10^{-3}$, then set $\alpha = 10^{-6}$.

The numerical results which we obtained are summarized in Tables 1.1-1.3. In these tables, **Dim** denotes the number of the variables in the problem, **Start. point** denotes the starting point, **Iter** denotes the number of iterations, which is also equal to the number of Jacobian evaluations for the function F, **NF** denotes the number of function evaluations for the function F and $R(x^k)$ denotes the value of $R(x)$ at the final iteration.

The results reported in Tables 1.1-1.3 show that the squared smoothing Newton methods are extremely promising and robust. The algorithm was able to solve almost all the problems. There are just three problems with superscript triple-asterisk which our algorithm was not able to solve because the steplength was getting too small. They are `gemmge`, `vonthmcp` and `hydroc20`. However we can solve these problems if we change some of the parameters. The results reported for these three problems were obtained by letting $\alpha = 10^{-10}$ and $\lambda = 10^{-6}$ while keeping other parameters unchanged.

Lastly, it is deserved to point out that domain violation phenomenon does not occur during our computation because $p(u, x) \in intX$ for all $(u, x) \in \Re^n_{++} \times \Re^n$. This is a very nice feature of our methods.

5 CONCLUSIONS

In this paper we present a regularized version of a class of squared smoothing Newton methods, originally proposed in [25], for the box constrained variational

Table 1.1 Numerical results for the problems from GAMSLIB

Problem	Dim	Iter	NF	$R(x^k)$
cafemge	101	29	30	8.0×10^{-7}
cammcp	242	8	9	9.1×10^{-8}
cammge	128	15	16	1.2×10^{-7}
cirimge	9	3	4	4.5×10^{-10}
co2mge	208	36	140	4.1×10^{-8}
dmcmge	170	13	21	3.4×10^{-13}
ers82mcp	232	6	8	1.9×10^{-9}
etamge	114	16	84	4.8×10^{-9}
finmge	153	10	11	7.8×10^{-10}
gemmcp	262	1	2	1.9×10^{-7}
gemmge***	178	15	18	6.6×10^{-7}
hansmcp	43	9	18	6.2×10^{-8}
hansmge	43	72	75	2.4×10^{-10}
harkmcp	32	14	18	6.5×10^{-13}
harmge	11	11	48	8.7×10^{-8}
kehomge	9	17	35	4.0×10^{-8}
kormcp	78	5	6	4.0×10^{-10}
mr5mcp	350	9	11	2.6×10^{-9}
nsmge	212	13	14	7.5×10^{-10}
oligomcp	6	7	11	1.1×10^{-7}
sammge	23	4	5	6.1×10^{-8}
scarfmcp	18	10	14	4.9×10^{-9}
scarfmge	18	14	19	1.5×10^{-10}
shovmge	51	88	90	2.3×10^{-9}
threemge	9	9	10	1.6×10^{-10}
transmcp	11	5	24	4.6×10^{-10}
two3mcp	6	6	7	1.3×10^{-8}
unstmge	5	8	9	1.3×10^{-8}
vonthmcp***	125	2609	19066	5.9×10^{-9}
vonthmge	80	94	516	1.8×10^{-8}
wallmcp	6	4	5	1.1×10^{-8}

inequality problem. As can be seen from the numerical results, these methods are fairly robust and promising. The global convergence of these methods were proved under the assumption that F is a P_0-function on X and the solution set of the problem (1.1) is nonempty and bounded. This assumption may be the weakest one known in the literature.

In Algorithm 3.1 we always assume that the iteration matrix $H'(z)$ is nonsingular. This is guaranteed by assuming that F is a P_0-function on X. In this paper, We have not discussed how to handle the case that $H'(z)$ is singular. By introducing a gradient direction if necessary, Kanzow and Pieper [18] described a strategy for handling the singularity issue of the iteration matrices for the smoothing Newton method proposed in [9]. Whether or not the idea introduced in [18] is applicable to our method is an interesting question.

Acknowledgements

The authors would like to thank Robert S. Womersley for his help in preparing the computation and two referees for their helpful comments.

References

[1] S.C. Billups, S.P. Dirkse and M.C. Ferris, "A comparison of algorithms for large-scale mixed complementarity problems", *Computational Optimization and Applications*, *7* (1997), 3–25.

[2] B. Chen and X. Chen, "A global linear and local quadratic continuation smoothing method for variational inequalities with box constraints", Preprint, Department of Management and Systems, Washington State University, Pullman, March 1997.

[3] B. Chen and X. Chen, "A global and local superlinear continuation-smoothing method for $P_0 + R_0$ and monotone NCP", *SIAM J. Optimization*, to appear.

[4] B. Chen, X. Chen and C. Kanzow, "A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results", AMR 97/28, Applied Mathematics Report, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, September 1997.

[5] B. Chen and P.T. Harker, "A non-interior-point continuation method for linear complementarity problems", *SIAM Journal on Matrix Analysis and Applications*, *14* (1993), 1168–1190.

[6] B. Chen and P.T. Harker, "Smooth approximations to nonlinear complementarity problems", *SIAM Journal on Optimization*, *7* (1997), 403–420.

[7] B. Chen, P.T. Harker and M. C. Pınar, "Continuation methods for nonlinear complementarity problems via normal maps", Preprint, Department of Management and Systems, Washington State University, Pullman, November 1995.

[8] C. Chen and O.L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems", *Computational Optimization and Applications*, *5* (1996), 97–138.

[9] X. Chen, L. Qi, and D. Sun, "Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities", *Mathematics of Computation*, *67* (1998), 519-540.

[10] X. Chen and Y. Ye, "On homotopy-smoothing methods for variational inequalities", *SIAM J. Control Optim.*, to appear.

[11] F.H. Clarke, *Optimization and Nonsmooth Analysis,* Wiley, New York, 1983.

[12] F. Facchinei and C. Kanzow, "Beyond monotonicity in regularization methods for nonlinear complementarity problems", DIS Working Paper 11-97, Università di Roman "La Sapienza", Roma, Italy, May 1997.

[13] F. Facchinei and J. Soares, "A new merit function for nonlinear complementarity problems and a related algorithm", *SIAM Journal on Optimization, 7* (1997), 225–247.

[14] S.A. Gabriel and J.J. Moré, "Smoothing of mixed complementarity problems", in *Complementarity and Variational Problems: State of the Art,* M.C. Ferris and J.S. Pang, eds., SIAM, Philadelphia, Pennsylvania, pp. 105–116, 1997.

[15] L. Grippo, F. Lampariello and S. Lucidi, "A nonmonotone line search technique for Newton's method", *SIAM Journal on Numerical Analysis, 23* (1986), pp. 707-716.

[16] K. Hotta and A. Yoshise "Global convergence of a class of non-interior-point algorithms using Chen-Harker-Kanzow functions for nonlinear complementarity problems", Discussion Paper Series No. 708, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305, Japan, December 1996.

[17] H. Jiang, "Smoothed Fischer-Burmeister equation methods for the nonlinear complementarity problem", Preprint, Department of Mathematics, the University of Melbourne, Victoria 3052, Australia, June 1997.

[18] C. Kanzow and H. Pieper, "Jacobian smoothing methods for general complenmentarity problems", Technical Report 97-08, Computer Sciences Department, University of Wisconsin, Madison, WI, October 1997.

[19] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization", *SIAM Journal Control and Optimization, 15* (1977), 957–972.

[20] H.-D. Qi, "A regularized smoothing Newton method for box constrained variational inequality problems with P_0 functions", Preprint, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, July 1997.

[21] H.-D. Qi, Private communication, August 1997.

[22] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", *Mathematics of Operations Research, 18* (1993), 227–244.

[23] L. Qi and D. Sun, "Improving the convergence of non-interior point algorithms for nonlinear complementarity problems", *Mathematics of Computation,* to appear.

[24] L. Qi and J. Sun, "A nonsmooth version of Newton's method", *Mathematical Programming, 58* (1993), 353–367.

[25] L. Qi, D. Sun and G. Zhou, "A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational

inequalities", AMR 97/13, Applied Mathematics Report, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, June 1997.

[26] G. Ravindran and M.S. Gowda, "Regularization of P_0 functions in box variational inequality problem", Preprint, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, Maryland 2150, USA, September 1997.

[27] S. M. Robinson, "Normal maps induced by linear transformation", *Mathematics of Operations Research, 17* (1992), 691–714.

[28] S.M. Robinson, "Generalized equations", in *Mathematical Programming: The State of the Art*, A. Bachem, M. Grötschel and B. Korte, eds., Springer-Verlag, Berlin, pp. 346–347, 1983.

[29] D. Sun, "A regularization Newton method for solving nonlinear complementarity problems", *Applied Mathematics and Optimization*, to appear.

[30] D. Sun and L. Qi, "On NCP-functions", AMR 97/32, Applied Mathematics Report, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, October 1997.

[31] S. Xu, "The global linear convergence of an infeasible non-interior path-following algorithm for complementarity problems with uniform P-functions", Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, December 1996.

Table 1.2 Numerical results for the problems from MCPLIB

Problem	Dim	Start. point	Iter	NF	$R(x^k)$
bertsekas	15	(1)	8	9	4.1×10^{-8}
bertsekas	15	(2)	10	11	5.5×10^{-9}
bertsekas	15	(3)	60	87	3.8×10^{-9}
billups	1	(1)	9	19	7.3×10^{-10}
bert_oc	5000	(1)	11	30	8.4×10^{-7}
bratu	5625	(1)	43	156	3.4×10^{-10}
choi	13	(1)	5	6	1.2×10^{-10}
colvdual	20	(1)	7	8	1.8×10^{-14}
colvdual	20	(2)	8	10	5.5×10^{-9}
colvnlp	15	(1)	7	8	1.1×10^{-14}
colvnlp	15	(2)	8	10	1.8×10^{-9}
cycle	1	(1)	7	12	0
ehl_k40	41	(1)	11	13	9.2×10^{-8}
ehl_k60	61	(1)	13	15	1.5×10^{-8}
ehl_k80	81	(1)	13	15	4.6×10^{-13}
ehl_kost	101	(1)	14	17	1.3×10^{-12}
explcp	16	(1)	6	8	1.0×10^{-7}
freebert	15	(1)	9	10	2.3×10^{-8}
freebert	15	(2)	25	70	4.7×10^{-7}
freebert	15	(3)	9	10	2.4×10^{-8}
freebert	15	(4)	9	10	5.5×10^{-8}
freebert	15	(5)	88	987	1.9×10^{-7}
freebert	15	(6)	9	10	5.6×10^{-8}
gafni	5	(1)	5	10	1.4×10^{-9}
gafni	5	(2)	5	8	1.7×10^{-10}
gafni	5	(3)	6	15	1.5×10^{-11}
hanskoop	14	(1)	22	35	1.5×10^{-7}
hanskoop	14	(2)	20	31	1.5×10^{-12}
hanskoop	14	(3)	21	37	1.0×10^{-7}
hanskoop	14	(4)	17	29	5.1×10^{-11}
hanskoop	14	(5)	9	11	2.6×10^{-8}
hydroc06	29	(1)	7	10	2.5×10^{-7}
hydroc20***	99	(1)	8	10	3.1×10^{-7}
jel	6	(1)	6	7	1.3×10^{-8}
josephy	4	(1)	10	30	6.1×10^{-10}
josephy	4	(2)	6	7	1.6×10^{-7}
josephy	4	(3)	17	18	1.9×10^{-12}
josephy	4	(4)	5	6	1.2×10^{-8}
josephy	4	(5)	4	6	9.9×10^{-7}
josephy	4	(6)	6	7	9.2×10^{-10}
kojshin	4	(1)	6	10	2.6×10^{-9}
kojshin	4	(2)	7	8	5.5×10^{-8}
kojshin	4	(3)	18	19	7.9×10^{-14}
kojshin	4	(4)	5	7	1.3×10^{-7}
kojshin	4	(5)	6	9	3.5×10^{-7}

Table 1.3 (continued) Numerical results for the problems from MCPLIB

Problem	Dim	Start. point	Iter	NF	$R(x^k)$
kojshin	4	(6)	6	9	2.5×10^{-7}
mathinum	3	(1)	4	5	3.7×10^{-9}
mathinum	3	(2)	5	6	4.1×10^{-12}
mathinum	3	(3)	9	13	2.6×10^{-8}
mathinum	3	(4)	5	6	1.3×10^{-9}
mathisum	4	(1)	4	5	1.6×10^{-9}
mathisum	4	(2)	5	6	2.3×10^{-9}
mathisum	4	(3)	10	11	5.0×10^{-10}
mathisum	4	(4)	4	5	7.6×10^{-9}
methan08	31	(1)	5	6	6.6×10^{-13}
nash	10	(1)	6	7	4.0×10^{-9}
nash	10	(2)	9	10	1.4×10^{-7}
obstacle	2500	(1)	7	8	5.5×10^{-13}
obstacle	2500	(2)	9	13	5.8×10^{-13}
opt_cont31	1024	(1)	11	16	4.1×10^{-10}
opt_cont127	4096	(1)	12	24	1.6×10^{-8}
opt_cont255	8193	(1)	14	33	3.1×10^{-8}
opt_cont511	16384	(1)	16	54	5.6×10^{-8}
pgvon105	105	(1)	71	200	6.3×10^{-7}
pgvon105	105	(2)	13	34	1.0×10^{-7}
pgvon105	105	(3)	13	34	1.0×10^{-7}
pgvon106	106	(1)	40	175	9.9×10^{-7}
pies	42	(1)	36	340	4.5×10^{-13}
powell	16	(1)	87	304	1.0×10^{-7}
powell	16	(2)	20	22	5.0×10^{-7}
powell	16	(3)	25	35	8.0×10^{-7}
powell	16	(4)	18	20	4.2×10^{-7}
powell_mcp	8	(1)	6	7	6.5×10^{-12}
powell_mcp	8	(2)	7	8	1.8×10^{-12}
powell_mcp	8	(3)	8	9	4.7×10^{-8}
powell_mcp	8	(4)	7	8	1.2×10^{-7}
scarfanum	13	(1)	9	18	5.4×10^{-10}
scarfanum	13	(2)	9	16	3.0×10^{-10}
scarfanum	13	(3)	8	9	6.3×10^{-9}
scarfasum	14	(1)	15	20	2.8×10^{-9}
scarfasum	14	(2)	20	43	1.6×10^{-10}
scarfasum	14	(3)	15	28	2.1×10^{-10}
scarfbnum	39	(1)	18	76	5.2×10^{-12}
scarfbnum	39	(2)	26	154	4.3×10^{-14}
scarfbsum	40	(1)	27	259	2.0×10^{-12}
scarfbsum	40	(2)	31	283	5.1×10^{-7}
sppe	27	(1)	60	386	1.9×10^{-7}
sppe	27	(2)	10	11	7.3×10^{-13}
tobin	42	(1)	18	86	2.5×10^{-13}
tobin	42	(2)	11	29	2.7×10^{-10}