Hamburger Beiträge
zur Angewandten Mathematik

A Penalized Fischer-Burmeister NCP-Function: Theoretical Investigation and Numerical Results

Bintong Chen, Xiaojun Chen and Christian Kanzow

Reihe A
Preprint 126
September 1997

Hamburger Beiträge zur Angewandten Mathematik

Reihe A	Preprints
Reihe B	Berichte
Reihe C	Mathematische Modelle und Simulation
Reihe D	Elektrische Netzwerke und Bauelemente
Reihe E	Scientific Computing
Reihe F	Computational Fluid Dynamics and Data Analysis

A PENALIZED FISCHER-BURMEISTER NCP-FUNCTION: THEORETICAL INVESTIGATION AND NUMERICAL RESULTS[1]

Bintong Chen[2], Xiaojun Chen[3] and Christian Kanzow[4]

[2] Department of Management and Systems
Washington State University
Pullman, WA 99164-4736
USA
e-mail: chenbi@wsu.edu

[3] School of Mathematics
The University of New South Wales
Sydney 2052
Australia
e-mail: X.Chen@unsw.edu.au

[4] University of Hamburg
Institute of Applied Mathematics
Bundesstrasse 55
D-20146 Hamburg
Germany
e-mail: kanzow@math.uni-hamburg.de

September 11, 1997
Revised September 16, 1997

[1]Results of this paper were presented at the International Symposium on Mathematical Programming in Lausanne, Switzerland, August 24-29, 1997.

Abstract: We introduce a new NCP-function that reformulates a nonlinear complementarity problem as a system of semismooth equations $\Phi(x) = 0$. The new NCP-function possesses all the nice properties of the Fischer-Burmeister function for local convergence. In addition, its natural merit function $\Psi(x) = \frac{1}{2}\Phi(x)^T\Phi(x)$ has all the nice features of the Kanzow-Yamashita-Fukushima merit function for global convergence. In particular, the merit function has bounded level sets for a monotone complementarity problem with a strictly feasible point. This property allows the existing semismooth Newton methods to solve this important class of complementarity problems without additional assumptions. We investigate the properties of a semismooth Newton-type method based on the new NCP-function and apply the method to a large class of complementarity problems. The numerical results indicate that the new algorithm is extremely promising.

Key Words: Nonlinear complementarity problems, Newton's method, generalized Jacobians, semismoothness, global convergence, quadratic convergence.

1 Introduction

Let F be a continuously differentiable function from $\mathbb{R}^n$ into itself. The *nonlinear complementarity problem* NCP(F) is to find a vector $x \in \mathbb{R}^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0.$$

Many algorithms developed for NCP(F) or related problems are based on reformulating them as a system of equations using so-called NCP-functions or as an optimization problem using suitable merit functions. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an *NCP-function* if

$$\phi(a, b) = 0 \Longleftrightarrow ab = 0, a \geq 0, b \geq 0.$$

Given an NCP-function, let us define

$$\Phi(x) = \text{vec}\{\phi(x_i, F_i(x))\},$$

where vec$\{z_i\}$ denotes a vector whose ith element is given by z_i. By definition, $x^* \in \mathbb{R}^n$ is a solution of NCP(F) if and only if it solves the following system of equations:

$$\Phi(x) = 0.$$

A function $\Psi : \mathbb{R}^n \to \mathbb{R}$ is called a *merit function* of NCP(F) if the solution set of the minimization problem

$$\min_{x \in \mathbb{R}^n} \Psi(x)$$

coincides with that of NCP(F) provided that the complementarity problem has a nonempty solution set. Clearly, for each NCP-function, there is a natural merit function:

$$\Psi(x) = \frac{1}{2}\Phi(x)^T \Phi(x).$$

Many NCP-functions and merit functions have been discovered in the past two decades [17, 9, 25, 27, 31, 35]. Mentioned below are those closely related to the current paper. The classical NCP-function is the natural residual or "min" function:

$$\phi_{NR}(a, b) = \min\{a, b\}.$$

However, its natural merit function is not differentiable everywhere, which makes it difficult to globalize a corresponding algorithm. The Fischer-Burmeister NCP-function [15]

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

has attracted much attention recently and has been extensively studied. Notice that we reverse the sign of the original Fischer-Burmeister NCP-function in this paper for the convenience of presentation. Many semismooth equation based algorithms were developed based on this function during the last three years. The function has many nice properties. In particular, it is semismooth and its natural merit function is continuously differentiable. These two properties made it possible to design globally and locally fast convergent algorithms for NCP(F) [9, 13, 20]. However, the Fischer-Burmeister NCP-function, as well as some of its variations, has some limitations in dealing with monotone complementarity problems; its natural merit function does not guarantee bounded level sets for this class of problems, an important class which includes, e.g., linearly constrained convex optimization problems as a special case.

Some modifications to the Fischer-Burmeister function have been proposed to overcome the above problem. Luo and Tseng [27] proposed a class of merit functions which include the following function as a special case:

$$\Psi_{LT}(x) := \frac{1}{2}\sum_{i=1}^{n}\left([-\phi_{FB}(x_i, F_i(x))]_+^2\right) + \frac{1}{2}[x^T F(x)]_+^2,$$

where $z_+ := \max\{0, z\}$ for $z \in \mathbb{R}$. Besides many other nice properties, the Luo-Tseng merit function has bounded level sets for a monotone NCP(F) with a strictly feasible point $\hat{x}$ (i.e., $\hat{x} > 0$ and $F(\hat{x}) > 0$). This was shown by Kanzow, Yamashita and Fukushima [25]. These authors further modified the Luo-Tseng merit function. Their modification includes the following merit function as a special case:

$$\Psi_{KYF}(x) := \frac{1}{2}\sum_{i=1}^{n}\left([-\phi_{FB}(x_i, F_i(x))]_+^2 + [x_i F_i(x)]_+^2\right).$$

The Kanzow-Yamashita-Fukushima merit function not only preserves the nice bounded level set property, but also enjoys additional decomposable structure. However, both merit functions lack corresponding NCP-functions, which makes it difficult to design a semismooth equation based algorithm.

In this paper, we propose a new NCP-function that combines the desirable features of the Fischer-Burmeister NCP-function and the Kanzow-Yamashita-Fukushima merit function. This new NCP-function is defined by

$$\phi_\lambda(a, b) := \lambda\phi_{FB}(a, b) + (1 - \lambda)a_+ b_+,$$

where $\lambda \in (0, 1)$ is an arbitrary but fixed parameter, i.e., we define the new NCP-function as a convex combination of the Fischer-Burmeister function ϕ_{FB} and the term a_+b_+; the latter term penalizes violations of the complementarity condition on the positive orthant and is highly important for both theoretical and numerical improvements, see, e.g., the comments given at the end of Section 4. Hence we call the new function the penalized Fischer-Burmeister NCP-function.

Clearly, as λ approaches 1, ϕ_λ approaches the Fischer-Burmeister function ϕ_{FB}. It is also straightforward to verify that ϕ_λ is an NCP-function. Therefore, $x^* \in \mathbb{R}^n$ is a solution of NCP(F) if and only if it solves the following system of equations:

$$\Phi_\lambda(x) := \mathrm{vec}\{\phi_\lambda(x_i, F_i(x))\} = 0.$$

Let

$$\psi_\lambda(a,b) := \frac{1}{2}\phi_\lambda(a,b)^2.$$

Then the natural merit function of Φ_λ is given by:

$$\Psi_\lambda(x) := \frac{1}{2}\Phi_\lambda(x)^T\Phi_\lambda(x) = \sum_{i=1}^{n}\psi_\lambda(x_i, F_i(x)).$$

We show in this paper that Φ_λ and Ψ_λ possess all the nice features of the Fischer-Burmeister NCP-function and the Kanzow-Yamashita-Fukushima merit function, respectively, for the purpose of local and global convergence analysis. In particular, using the new NCP-function, the existing semismooth based algorithms are able to solve larger classes of complementarity problems than using the Fischer-Burmeister function alone. This includes the class of monotone problems with a strictly feasible point.

This paper is organized as follows. In the next section, we study the properties of the NCP-function ϕ_λ and the corresponding operator Φ_λ. In Section 3, we investigate the properties of the natural merit function Ψ_λ. In particular, we give conditions for a stationary point of Ψ_λ to be a solution of NCP(F) as well as conditions for Ψ_λ to have bounded level sets. We also provide a characterization for a monotone complementarity problem to have a strictly feasible point. In Section 4, we apply our new NCP-function to the semismooth equation based algorithm developed in [9]. Extensive numerical tests indicate that the semismooth algorithm based on our new NCP-function is extremely robust and promising. In particular, it is considerably better than, e.g., the same method based on the Fischer-Burmeister NCP-function.

Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm. We use vec$\{x_i\}$ for a vector whose ith element is x_i and diag$\{x_i\}$ for a diagonal matrix with ith diagonal element equal to x_i. We say that a mapping $G : \mathbb{R}^n \to \mathbb{R}^m$ is a C^1 function if G is continuously differentiable, and an LC^1 function if the Jacobian of G is locally Lipschitzian. If the mapping $G : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitzian and if D_G is its set of differentiable points, we use

$$\partial G(x) := \operatorname{conv}\{V \in \mathbb{R}^{m\times n} \,|\, \exists \{x^k\} \in D_G : \{x^k\} \to x \text{ and } G'(x^k) \to V\}$$

to denote Clarke's [8] generalized Jacobian of G at x. If G is a real-valued mapping, i.e., if $m = 1$, the generalized Jacobian will also be called the generalized gradient of G at x. Furthermore, we denote by

$$\partial_C G(x)^T = \partial G_1(x) \times \ldots \times \partial G_m(x)$$

the C-subdifferential of G at x, see Qi [30], where the right-hand side denotes a set of matrices whose ith column can be any element from the generalized gradient $\partial G_i(x)$.

2 Properties of ϕ_λ and Φ_λ

In this section, we first study the semismoothness of ϕ_λ and Φ_λ. We then provide conditions for all elements of the C-subdifferential $\partial_C \Phi_\lambda(x^*)$ of Φ_λ at a solution $x^* \in \mathbb{R}^n$ to be nonsingular as well as a procedure to evaluate an element in $\partial_C \Phi_\lambda(x)$ at an arbitrary point $x \in \mathbb{R}^n$.

Denote $N_\phi = \{(a,b) \mid a \geq 0, b \geq 0, ab = 0\}$.

Proposition 2.1 *The function $\phi_\lambda : \mathbb{R}^2 \to \mathbb{R}$ satisfies the following properties:*

1. $\phi_\lambda(a,b) = 0 \Longleftrightarrow (a,b) \in N_\phi$.
2. $\phi_\lambda(a,b)$ *is continuously differentiable on* $\mathbb{R}^2 \setminus N_\phi$.
3. ϕ_λ *is strongly semismooth on* $\mathbb{R}^2$.
4. *The generalized gradient $\partial\phi_\lambda(a,b)$ of ϕ_λ at a point $(a,b) \in \mathbb{R}^2$ is equal to the set of all (v_a, v_b) such that*

$$(v_a, v_b) = \begin{cases} \lambda(1 - \frac{a}{\|(a,b)\|}, 1 - \frac{b}{\|(a,b)\|}) + (1-\lambda)(b_+ \partial a_+, a_+ \partial b_+) & \text{if } (a,b) \neq (0,0), \\ \lambda(1-\xi, 1-\zeta) & \text{if } (a,b) = (0,0), \end{cases} \tag{1}$$

   where (ξ, ζ) is any vector satisfying $\|(\xi,\zeta)\| \leq 1$ and

$$\partial z_+ = \begin{cases} 1 & \text{if } z > 0, \\ [0,1] & \text{if } z = 0, \\ 0 & \text{if } z < 0. \end{cases}$$

5. *Let $\{a^k\}, \{b^k\} \subset \mathbb{R}$ be any two sequences such that either $a_+^k b_+^k \to \infty$, or $a^k \to -\infty$, or $b^k \to -\infty$. Then $|\phi_\lambda(a^k, b^k)| \to \infty$.*

Proof. Results 1 and 2 are straightforward.

Result 3: The Fischer-Burmeister NCP-function ϕ_{FB} is strongly semismooth by Lemma 20 in [16] (see also [13, 23, 32]). Moreover, it is easy to see that the plus-function $z \mapsto z_+$ is strongly semismooth. Since the composition of two strongly semismooth functions is strongly semismooth by Theorem 19 in [16], it follows that the product of two strongly semismooth functions is also strongly semismooth. Therefore, ϕ_λ is strongly semismooth since the sum of two strongly semismooth functions is obviously strongly semismooth.

Result 4: In the proof of this result, we assume that the reader is familiar with some results on nonsmooth analysis, see Clarke [8].

We first note that ϕ_{FB} is continuously differentiable everywhere except at $(0,0)$. Hence ϕ_{FB} is strictly differentiable except at the origin by the Corollary to Proposition 2.2.1 in [8]. On the other hand, using the characterization of the generalized gradient from Theorem 2.5.1 in [8], it is not difficult to see that the generalized gradient of the function

$$\phi_+(a,b) := a_+ b_+$$

at $(0,0)$ reduces to a singleton, namely

$$\partial\phi_+(0,0) = \{(0,0)^T\}. \tag{2}$$

Together with Proposition 2.2.4 in [8], we therefore obtain the (somewhat surprising) result that ϕ_+ is strictly differentiable at the origin. Since

$$\phi_\lambda(a,b) = \lambda\phi_{FB}(a,b) + (1-\lambda)\phi_+(a,b)$$

by definition and since both ϕ_{FB} and ϕ_+ are obviously locally Lipschitzian functions, we obtain from Corollary 2 of Proposition 2.3.3 in [8] that

$$\partial\phi_\lambda(a,b) = \lambda\partial\phi_{FB}(a,b) + (1-\lambda)\partial\phi_+(a,b).$$

The generalized gradient of the Fischer-Burmeister function is well-known (see, e.g., Lemma 2.2.1 in [23]) to be

$$\partial\phi_{FB}(a,b) = \begin{cases} \left(1 - \frac{a}{\|(a,b)\|}, 1 - \frac{b}{\|(a,b)\|}\right) & \text{if } (a,b) \neq (0,0), \\ (1-\xi, 1-\zeta) & \text{if } (a,b) = (0,0), \end{cases}$$

where $(\xi,\zeta) \in \mathbb{R}^2$ denotes an arbitrary vector with $\|(\xi,\zeta)\| \leq 1$. Since the plus-function $z \mapsto z_+$ is convex, it follows from Proposition 2.3.6 (b) in [8] that it is a regular function in the sense of Definition 2.3.4 in [8]. Taking into account the nonnegativity of this plus-function, Proposition 2.3.13 from [8] therefore gives

$$\partial\phi_+(a,b) = (b_+\partial a_+, a_+\partial b_+).$$

Since $a_+ = b_+ = 0$ for $(a,b) = (0,0)$, Result 4 follows immediately from the previous observations.

Result 5: This statement follows from the fact that ϕ_{FB} is unbounded if either $a^k \to -\infty$ or $b^k \to -\infty$. $\square$

Result 2 implies that the Fischer-Burmeister function is smoother than ϕ_λ. ϕ_{FB} is nondifferentiable only at $(0,0)$, while ϕ_λ is nondifferentiable on the larger set N_ϕ. However, this additional nonsmoothness does not affect the convergence analysis. Indeed, Result 2 is all that is needed for ψ_λ and therefore Ψ_λ to be continuously differentiable, the key property in order to establish global convergence.

Since Φ_λ is (strongly) semismooth if and only if all component functions are (strongly) semismooth and since the composite of (strongly) semismooth functions is (strongly) semismooth ([28, 16]), we obtain the following results as an immediate consequence of Proposition 2.1.

Theorem 2.2 *The equation operator $\Phi_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ satisfies the following properties:*

1. *Φ_λ is semismooth.*

2. *Φ_λ is strongly semismooth if F is LC^1.*

Based on Result 4 of Proposition 2.1, we obtain the following overestimation of the C-subdifferential $\partial_C \Phi_\lambda(x)$.

Proposition 2.3 *For any $x \in \mathbb{R}^n$, we have*

$$\partial_C \Phi_\lambda(x) \subseteq D_a(x) + D_b(x) F'(x),$$

where $D_a(x) = \mathrm{diag}\{a_i(x)\}$ and $D_b(x) = \mathrm{diag}\{b_i(x)\}$ are diagonal matrices with entries

$$(a_i(x), b_i(x)) \in \partial \phi_\lambda(x_i, F_i(x)),$$

where $\partial \phi_\lambda(x_i, F_i(x))$ denotes the set from Proposition 2.1 with (a, b) being replaced by $(x_i, F_i(x))$.

Proof. The proof is similar to Proposition 7 of [13]. By our definition of the C-subdifferential, we have

$$\partial_C \Phi_\lambda(x)^T = \partial \Phi_{\lambda,1}(x) \times \ldots \times \partial \Phi_{\lambda,n}(x),$$

where $\Phi_{\lambda,i}$ is the ith component function of Φ_λ. Based on Result 4 of Proposition 2.1 and Theorem 2.3.9 in [8], we have

$$\partial \Phi_{\lambda,i}(x) \subseteq a_i(x) e_i^T + b_i(x) \nabla F_i(x)^T \tag{3}$$

with $(a_i(x), b_i(x))$ being described by (1) (to this end, note that the set on the right-hand side of (3) is already convex and closed). □

We now provide a procedure to calculate an element of the C-subdifferential $\partial_C \Phi_\lambda(x)$ at an arbitrary point $x \in \mathbb{R}^n$. This procedure will be used in our main algorithm to be described in Section 4.

Algorithm 2.4 *(Procedure to evaluate an element $V \in \partial_C \Phi_\lambda(x)$)*

(S.0) Let $x \in \mathbb{R}^n$ be given, and let V_i denote the ith row of a matrix $V \in \mathbb{R}^{n \times n}$.

(S.1) Set $S_1 = \{i |\, x_i = F_i(x) = 0\}$ and $S_2 = \{i |\, x_i > 0, F_i(x) > 0\}$.

(S.2) Set $z \in \mathbb{R}^n$ such that $z_i = 0$ for $i \notin S_1$ and $z_i = 1$ for $i \in S_1$.

(S.3) For $i \in S_1$, set

$$V_i = \lambda \left(1 - \frac{z_i}{\|(z_i, \nabla F_i(x)^T z\|}\right) e_i^T + \lambda \left(1 - \frac{\nabla F_i(x)^T z}{\|(z_i, \nabla F_i(x)^T z)\|}\right) \nabla F_i(x)^T.$$

(S.4) For $i \in S_2$, set

$$\begin{aligned} V_i \;=\; & \left[\lambda \left(1 - \frac{x_i}{\|(x_i, F_i(x))\|}\right) + (1-\lambda) F_i(x)\right] e_i^T \\ & + \left[\lambda \left(1 - \frac{F_i(x)}{\|(x_i, F_i(x))\|}\right) + (1-\lambda) x_i\right] \nabla F_i(x)^T. \end{aligned}$$

(S.5) For $i \notin S_1 \cup S_2$, set

$$V_i = \lambda \left(1 - \frac{x_i}{\|(x_i, F_i(x))\|}\right) e_i^T + \lambda \left(1 - \frac{F_i(x)}{\|(x_i, F_i(x))\|}\right) \nabla F_i(x)^T.$$

Our next result shows that the matrix $V \in \mathbb{R}^{n \times n}$ calculated by Algorithm 2.4 is indeed an element from the C-subdifferential $\partial_C \Phi_\lambda(x)$.

Proposition 2.5 *The element V calculated by Algorithm 2.4 is an element of the C-subdifferential $\partial_C \Phi_\lambda(x)$.*

Proof. We first note that we only calculate an element of the C-subdifferential of Φ_λ (and not an element of its generalized Jacobian or the B-subdifferential). This makes our analysis considerably simpler. In particular, this allows us to treat each column separately.

First assume that we have an index $i \in S_1$. Then the result follows from the fact that the mapping $\phi_+(a, b) = a_+ b_+$ turned out to be strictly differentiable at the origin with $\partial \phi_+(0, 0) = \{(0, 0)^T\}$, see (2). The result then follows from Theorem 27 in [9].

On the other hand, if $i \in S_2$, the expression for V_i is obvious since $\Phi_{\lambda,i}$ is continuously differentiable at x.

It remains to consider the case $i \notin S_1 \cup S_2$. If i is such that $i \notin S_1 \cup S_2$ and $(x_i, F_i(x)) \notin \mathbb{R}^2_+$, the statement follows again from the continuous differentiability of $\Phi_{\lambda,i}$ at the point $x \in \mathbb{R}^n$. Therefore, we only have to deal with those indices $i \notin S_1 \cup S_2$ such that either

$$x_i = 0 \quad \text{and} \quad F_i(x) > 0 \tag{4}$$

or

$$x_i > 0 \quad \text{and} \quad F_i(x) = 0. \tag{5}$$

We first consider situation (4). We choose a sequence $\{y^k\} \subseteq \mathbb{R}^n$ such that

$$y^k := x - \varepsilon_k e_i,$$

where $\varepsilon_k \downarrow 0$ and e_i denotes the ith column of the identity matrix. Then $y_i^k < 0$ and $F_i(y^k) > 0$ for all k sufficiently large, so that $\Phi_{\lambda,i}$ is continuously differentiable at these points y^k with

$$\nabla\Phi_{\lambda,i}(y^k) = \lambda\left(1 - \frac{y_i^k}{\|(y_i^k, F_i(y^k))\|}\right) e_i + \lambda\left(1 - \frac{F_i(y^k)}{\|(y_i^k, F_i(y^k))\|}\right) \nabla F_i(y^k). \qquad (6)$$

Taking the limit $k \to \infty$ gives the desired expression of V_i in Step (S.5) of Algorithm 2.4.

Finally, we consider the case (5). Without loss of generality, we can assume that $\nabla F_i(x) \neq 0$ since otherwise the mapping $x \mapsto \max\{0, F_i(x)\}$ is differentiable so that the expression for V_i in Step (S.5) of Algorithm 2.4 would obviously be correct. Given this situation, we define another sequence $\{y^k\} \subseteq \mathbb{R}^n$ by

$$y^k := x - \varepsilon_k \nabla F_i(x),$$

where, again, $\varepsilon_k \downarrow 0$. Since F is continuously differentiable, we obtain from the mean value theorem that there is a vector ξ^k on the open line segment from x to y^k with

$$F_i(y^k) = F_i(x) - \varepsilon_k \nabla F_i(\xi^k)^T \nabla F_i(x).$$

(Note that the vector ξ^k also depends on the component i but that this dependence is not important in our context.) Since $\xi^k \to x$ for $k \to \infty$, we have $y_i^k > 0$ and $F_i(y^k) < 0$ for all k sufficiently large. Again, this implies that $\Phi_{\lambda,i}$ is continuously differentiable at these points y^k with $\nabla\Phi_{\lambda,i}(y^k)$ being equal to the expression on the right-hand side of (6). The assertion therefore follows by taking the limit $k \to \infty$. □

Since the merit function Ψ_λ is differentiable (to be established in Section 3), a semismooth equation based algorithm can find a descent direction even if the C-subdifferential $\partial_C \Phi_\lambda(x)$ contains singular elements, as long as the current point is not a stationary point of Ψ_λ. Therefore, the nonsingularity of $\partial_C \Phi_\lambda(x)$ is not so important (theoretically) for global convergence. However, to ensure fast local convergence for a semismooth algorithm, we require all elements in the C-subdifferential $\partial_C \Phi_\lambda(x^*)$ to be nonsingular at a solution point x^* of NCP(F). The next result shows that the R-regularity condition [34] is sufficient for this purpose.

To this end, we define the following index sets depending on a given solution x^* of NCP(F):

$$\begin{aligned} \alpha &:= \{i \mid x_i^* > 0, F_i(x^*) = 0\}, \\ \beta &:= \{i \mid x_i^* = 0, F_i(x^*) = 0\}, \\ \gamma &:= \{i \mid x_i^* = 0, F_i(x^*) > 0\}. \end{aligned}$$

Let $M := F'(x^*)$. Then x^* is called an R-regular solution if $M_{\alpha\alpha}$ is nonsingular and the Schur-complement

$$M_{\beta\beta} - M_{\beta\alpha} M_{\alpha\alpha}^{-1} M_{\alpha\beta} \qquad (7)$$

is a P-matrix.

Theorem 2.6 *If x^* is an R-regular solution of NCP(F), then all elements in the C-subdifferential $\partial_C \Phi_\lambda(x^*)$ are nonsingular.*

Proof. Due to Proposition 2.3, any element $V \in \partial_C \Phi_\lambda(x^*)$ can be written as

$$D_a + D_b F'(x^*),$$

for some non-negative diagonal matrices D_a and D_b. Without loss of generality, let

$$D_a = \text{diag}\{D_{a,\alpha}, D_{a,\beta}, D_{a,\gamma}\} \text{ and } D_b = \text{diag}\{D_{b,\alpha}, D_{b,\beta}, D_{b,\gamma}\},$$

where $D_{a,\alpha} := (D_a)_{\alpha\alpha}$ etc. Again by Result 4 of Proposition 2.1 and Proposition 2.3, we have

$$D_{a,\alpha} = 0_\alpha, \;\; D_{a,\gamma} = \lambda I_\gamma + (1-\lambda)\text{diag}\{F_i(x^*)\}_\gamma \text{diag}\{w_i\}_\gamma$$

and, similarly,

$$D_{b,\alpha} = \lambda I_\alpha + (1-\lambda)\text{diag}\{x_i^*\}_\alpha \text{diag}\{w_i\}_\alpha, D_{b,\gamma} = 0_\gamma,$$

where $w_i \in [0,1]$ for all $i \in \alpha \cup \gamma$ and where $\text{diag}\{w_i\}_\gamma$ denotes a diagonal matrix of dimension $|\gamma| \times |\gamma|$ with diagonal elements $w_i, i \in \gamma$. The rest of the proof is identical to that of Theorem 2.7 in [24] or Lemma 5.3 in [9]. □

3 Properties of ψ_λ and Ψ_λ

In this section, we show that Ψ_λ is continuously differentiable. As a result, it is a convenient merit function to globalize a semismooth equation based algorithm. We provide conditions for a stationary point of Ψ_λ to be a solution of NCP(F) as well as conditions for Ψ_λ to have bounded level sets. Based on these results, we also add a discussion on the existence of a strictly feasible point for NCP(F).

Since ϕ_λ is differentiable everywhere except on the set N_ϕ and $\phi_\lambda(a,b) = 0$ for all $(a,b) \in N_\phi$, the following result can be established in essentially the same way as the corresponding results in [9, 24].

Lemma 3.1 *ψ_λ is continuously differentiable on $\mathbb{R}^2$.*

Similarly, we obtain

Theorem 3.2 *The merit function Ψ_λ is continuously differentiable with $\nabla \Psi_\lambda(x) = V^T \Phi_\lambda(x)$ for any $V \in \partial_C \Phi_\lambda(x)$.*

Proof. The proof is basically the same as the one for the Fischer-Burmeister function Φ_{FB}, see [13]. The only difference is that we allow V to be an element from the C-subdifferential (instead of being an element from the generalized Jacobian). It is not difficult to see, however, that the technique of proof used in [13] goes through for the C-subdifferential. □

We need the following key properties of ψ_λ [9, 25, 24] to study the properties of a stationary point of Ψ_λ.

Proposition 3.3 *The following statements hold for ψ_λ:*

1. $\psi_\lambda(a,b)=0 \Longleftrightarrow \nabla\psi_\lambda(a,b)=0 \Longleftrightarrow \frac{\partial\psi_\lambda}{\partial a}(a,b)=0 \Longleftrightarrow \frac{\partial\psi_\lambda}{\partial b}(a,b)=0.$
2. $\frac{\partial\psi_\lambda}{\partial a}(a,b)\frac{\partial\psi_\lambda}{\partial b}(a,b)\geq 0.$

Proof. It is easy to verify that both results hold if $(a,b)=(0,0)$. We therefore assume $(a,b)\neq(0,0)$ in the remaining proof.

Result 1: It suffices to show that

$$\frac{\partial\psi_\lambda}{\partial a}(a,b)=0 \Longrightarrow \psi_\lambda(a,b)=0$$

since $\psi_\lambda(a,b)=0$ implies that (a,b) is a global minimum of ψ_λ, which in turn implies all the equalities. Let $(v_a,v_b)\in\partial\phi_\lambda(a,b)$ as defined in Result 4 of Proposition 2.1. Since $(a,b)\neq(0,0)$, we obtain the following equalities from Lemma 3.1 and [8, Theorem 2.6.6]:

$$\frac{\partial\psi_\lambda}{\partial a}(a,b)=\phi_\lambda(a,b)v_a=\phi_\lambda(a,b)[\lambda(1-\frac{a}{\|(a,b)\|})+(1-\lambda)(b_+\partial a_+)].$$

Hence we have either $\phi_\lambda(a,b)=0$ or $v_a=0$. If the former is true, we proved Result 1. So we assume that the latter is true. Since $b_+\partial a_+\geq 0$, we have

$$1-\frac{a}{\|(a,b)\|}=0.$$

However, this implies $\phi_\lambda(a,b)=0$ and we obtain Result 1.

Result 2: By definition,

$$\frac{\partial\psi_\lambda}{\partial a}(a,b)\frac{\partial\psi_\lambda}{\partial b}(a,b)=\phi_\lambda(a,b)^2 v_a v_b.$$

Since all elements $(v_a,v_b)\in\partial\phi_\lambda(a,b)$ are non-negative, we obtain Result 2. □

Based on the above property and the same proof techniques developed in [9, 25, 24], we have the following condition for a stationary point to be a solution of NCP(F):

Theorem 3.4 *Assume that x^* is a stationary point of Ψ_λ such that the Jacobian $F'(x^*)$ is a P_0-matrix. Then x^* is a solution of NCP(F).*

We next provide conditions for the level sets

$$\mathcal{L}(c) := \{x \in \mathrm{I\!R}^n |\, \Psi_\lambda(x) \leq c\}$$

to be bounded (and therefore compact). We also state some conditions which guarantee the boundedness of the "feasible" level sets

$$\mathcal{L}(c, \mathcal{F}) = \{x \in \mathcal{F} \mid \Psi_\lambda(x) \leq c\},$$

where

$$\mathcal{F} := \{x \in \mathrm{I\!R}^n |\, x \geq 0, F(x) \geq 0\}$$

denotes the feasible set for the complementarity problem NCP(F).

Proposition 3.5 *The following statements hold:*

1. *If*
$$\mathcal{L}_{FB}(c) := \{x \in \mathrm{I\!R}^n \mid \Psi_{FB}(x) \leq c\}$$
is bounded for some $c \geq 0$, then $\mathcal{L}(c')$ is bounded for $c' := \lambda^2 c$.

2. *If*
$$\{x \in \mathcal{F} \mid x^T F(x) \leq c\}$$
is bounded for some $c \geq 0$, then $\mathcal{L}(c', \mathcal{F})$ is bounded for $c' := (1-\lambda)^2 c^2/2n$.

Proof. By definition, we have

$$\begin{aligned}
\Psi_\lambda(x) &= \frac{1}{2}\|\mathrm{vec}\{\lambda\phi_{FB}(x_i, F_i(x)) + (1-\lambda)[x_i]_+[F_i(x)]_+\}\|^2 \\
&\geq \frac{1}{2}\lambda^2\|\Phi_{FB}(x)\|^2 \\
&= \lambda^2\Psi_{FB}(x).
\end{aligned}$$

Result 1 then follows immediately. Since ϕ_{FB} is nonnegative in $\mathcal{F}$, we have

$$\begin{aligned}
\Psi_\lambda(x) &= \frac{1}{2}\|\mathrm{vec}\{\lambda\phi_{FB}(x_i, F_i(x)) + (1-\lambda)x_i F_i(x)\}\|^2 \\
&\geq \frac{1}{2}(1-\lambda)^2\|\mathrm{vec}\{x_i F_i(x)\}\|^2 \\
&\geq \frac{(1-\lambda)^2}{2n}(x^T F(x))^2
\end{aligned}$$

for all $x \in \mathcal{F}$, where the second inequality follows from the well-known fact that $\|z\|_1 \leq \sqrt{n}\|z\|$ for any $z \in \mathrm{I\!R}^n$. Result 2 now follows immediately. $\square$

We next state two simple consequences of Proposition 3.5.

Corollary 3.6 *Let F be a P_0-function and suppose that the solution set of NCP(F) is nonempty and bounded. Then the level sets $\mathcal{L}(c)$ are bounded for every positive c sufficiently small.*

Proof. From Lemma 4.3 in [12], the level set

$$\mathcal{L}_{FB}(c) = \{x \in \mathbb{R}^n \,|\, \Psi_{FB}(x) \le c\},$$

is bounded for every positive c sufficiently small. The assertion then follows from Result 1 of Proposition 3.5. □

Corollary 3.7 *Let F be an affine function. If the solution set of NCP(F) is bounded, then $\mathcal{L}(c, \mathcal{F})$ is bounded for every $c > 0$.*

Proof. By Lemma 2.1 in [7], if the solution set of NCP(F) is bounded, the set

$$\{x \in \mathcal{F} \mid x^T F(x) \le c\}$$

is bounded for every $c > 0$. The result then follows from Proposition 3.5. □

It is easy to see that Result 2 of Proposition 3.5 and Corollary 3.7 fail to hold for Ψ_{FB}; just consider NCP(F) with $F(x) \equiv 1$. Notice that the boundedness of $\mathcal{L}(c)$ ensures the global convergence of a semismooth Newton-type method, while the boundedness of $\mathcal{L}(c, \mathcal{F})$ guarantees an interior point method to find a solution of monotone problems if it exists.

We next provide conditions for $\mathcal{L}(c)$ to be bounded for all $c \ge 0$. It turns out that the following condition on F is sufficient:

Condition 3.8 *For any sequence $\{x^k\}$ such that*

$$\|x^k\| \to \infty, \quad [-x^k]_+ < \infty, \quad [-F(x^k)]_+ < \infty,$$

it holds

$$\max_i [x_i^k]_+ [F_i(x^k)]_+ \to \infty.$$

Theorem 3.9 *If function F satisfies Condition 3.8, then the level sets $\mathcal{L}(c)$ are bounded for all $c \ge 0$.*

Proof. Assume on the contrary there exists an unbounded sequence $\{x^k\} \subseteq \mathcal{L}(c)$ for some $c \ge 0$. Since $\Psi_\lambda(x^k) \le c$ for all $k \in \mathbb{N}$, the sequence $\{\Psi_{FB}(x^k)\}$ is also bounded by the proof of Proposition 3.5. As a result, there is no index i such that $x_i^k \to -\infty$ or $F_i(x^k) \to -\infty$. Since F satisfies Condition 3.8, there is a fixed index j such that $[x_j^k]_+ [F_j(x^k)]_+ \to \infty$ at least on a subsequence. However, this implies

$\Psi_\lambda(x^k)$ is unbounded, a contradiction to the level set assumption. □

To the best of our knowledge, Condition 3.8 on F is the weakest assumption to guarantee bounded level sets for nonlinear complementarity problems. Indeed, both the R_0-function in the sense of Chen and Harker [4] and the monotone function with a strictly feasible point satisfy this condition:

Proposition 3.10 *Condition 3.8 is satisfied if F is*

1. *either a monotone function with a strictly feasible point,*
2. *or an R_0-function.*

Proof. Result 1: Let $\hat{x}$ be any strictly feasible point, i.e., let $\hat{x} > 0$ and $F(\hat{x}) > 0$. Let $\{x^k\}$ be an unbounded sequence as defined in Condition 3.8. By assumption and taking a subsequence if necessary, we can assume that $x_i^k \to \infty$ for all i such that $\{x_i^k\}$ is unbounded. Since F is monotone, we have

$$(x^k)^T F(\hat{x}) + \hat{x}^T F(x^k) \leq (x^k)^T F(x^k) + \hat{x}^T F(\hat{x}).$$

It follows that

$$(x^k)^T F(x^k) \to \infty.$$

Since

$$[-x^k]_+ < \infty \text{ and } [-F(x^k)]_+ < \infty$$

by assumption, there exists an index j such that

$$[x_j^k]_+ [F_j(x^k)]_+ \to \infty$$

and therefore, Condition 3.8 holds.

Result 2: Clearly, $x_i^k \to \infty$ and $F_i(x^k) \to \infty$ implies

$$[x_i^k]_+ [F_i(x^k)]_+ \to \infty.$$

The result then follows immediately in view of the definition of an R_0-function [4]. □

In addition, Condition 3.8 is also weaker than the conditions used in Theorem 4.2 of [25] and Theorem 4.1 of [27]. This is because of the following implications:

$$\begin{aligned}
\liminf_{k\to\infty} \frac{(x^k)^T F(x^k)}{\|x^k\|} > 0 \implies & \liminf_{k\to\infty} \frac{\max_i\{x_i^k F_i(x^k)\}}{\|x^k\|} > 0 \\
\implies & \max_i \{x_i^k F_i(x^k)\} \to \infty \\
\implies & \max_i \{[x_i^k]_+ [F_i(x^k)]_+\} \to \infty,
\end{aligned}$$

where the last implication holds only for the sequence $\{x^k\}$ defined in Condition 3.8.

The next result shows that $\sqrt{\Psi_\lambda}$ provides a global error bound for a complementarity problem with a uniform P-function.

Theorem 3.11 *If F is a uniform P-function, then there exists a constant $\tau > 0$ such that*

$$\|x - x^*\|^2 \le \tau \Psi_\lambda(x)$$

for all $x \in \mathbb{R}^n$, where x^* is the unique solution of NCP(F) .

Proof. By Corollary 2 of [2], $(\|\min\{x, F(x)\}\| + [x]_+^T [F(x)]_+)$ is a global error bound for NCP(F) with a uniform P-function F. Therefore, for any $x \in \mathbb{R}^n$, we have

$$\begin{aligned}
& \|x - x^*\|^2 \\
\le\ & \tau_1 (\|\min\{x, F(x)\}\| + [x]_+^T [F(x)]_+)^2 \\
\le\ & \tau_2 (\|\min\{x, F(x)\}\|_1 + [x]_+^T [F(x)]_+)^2 \\
=\ & \tau_2 \|\min\{x, F(x)\} + \mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|_1^2 \\
\le\ & \tau_3 \|\min\{x, F(x)\} + \mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|^2 \\
=\ & \tau_3 (\|\min\{x, F(x)\}\|^2 + \|\mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|^2) \\
& +2 \sum_{i=1}^n |\min\{x_i, F_i(x)\}| [x_i]_+ [F_i(x)]_+ \\
\le\ & \tau_3 (c^2 \|\Phi_{FB}(x)\|^2 + \|\mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|^2) + 2c \sum_{i=1}^n |\phi_{FB}(x_i, F_i(x)| [x_i]_+ [F_i(x)]_+ \\
=\ & \tau_3 (\|c\Phi_{FB}(x) + \mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|^2) \\
\le\ & \tau (\frac{1}{2} \|\lambda \Phi_{FB}(x) + (1-\lambda) \mathrm{vec}\{[x_i]_+ [F_i(x)]_+\}\|^2) \\
=\ & \tau \Psi_\lambda(x),
\end{aligned}$$

where the second and third inequalities follow from the equivalence of norms, the fourth inequality follows from Lemma 3.1 of Tseng [36] with $c > 0$ being a constant. $\tau_1, \tau_2, \tau_3, \tau$ are appropriate positive constants that depend on the dimension n, and τ also depends on λ. □

Notice that the above result does not require F to be Lipschitz continuous, a condition often needed for similar results based on merit functions derived from the Fischer-Burmeister function or related NCP-functions, see [24, 35]. On the other hand, however, the class of merit functions considered in [25] has similar global error bound properties.

In the remaining part of this section, we will discuss some extensions to Result 1 of Proposition 3.10. More precisely, we show that the assumption on the existence of a strictly feasible point is also necessary for $\mathcal{L}(c)$ to be bounded for monotone complementarity problems.

Proposition 3.12 *If F is a P_0-function and NCP(F) has a nonempty and bounded solution set, then there is a strictly feasible point for NCP(F).*

Proof. For a fixed parameter $\mu > 0$, define

$$\phi_{FB}(a, b; \mu) := a + b - \sqrt{a^2 + b^2 + 2\mu}$$

and

$$\Phi_{FB}(x; \mu) := \mathrm{vec}\{\phi_{FB}(x_i, F_i(x); \mu)\}.$$

Note that, for $\mu = 0$, $\phi_{FB}(a, b; \mu)$ and $\Phi_{FB}(x; \mu)$ reduce to the functions $\phi_{FB}(a, b)$ and $\Phi_{FB}(x)$, respectively. These perturbed functions were introduced in Kanzow [21] and have the following property:

$$\phi_{FB}(a, b; \mu) = 0 \Longleftrightarrow a > 0, \;\; b > 0, \;\; ab = \mu.$$

Therefore,

$$\Phi_{FB}(x; \mu) = 0 \Longleftrightarrow x > 0, \;\; F(x) > 0, \;\; x_i F_i(x) = \mu \;\; \forall i, \tag{8}$$

see [21]. By Lemma 4.3 of [12], the level set

$$\mathcal{L}_{FB}(c) := \{x \in \mathbb{R}^n \,|\, \|\Phi_{FB}(x)\| \leq c\}$$

is (nonempty and) bounded for $c > 0$ sufficiently small. We want to show that this implies the compactness of the level set

$$\mathcal{L}_{FB}(c/2; \mu) := \{x \in \mathbb{R}^n \,|\, \|\Phi_{FB}(x; \mu)\| \leq c/2\}$$

for $\mu > 0$ sufficiently small (to be specified later). To this end, we first recall that the following relationship holds for all $\mu > 0$ (see [22]):

$$\|\Phi_{FB}(x) - \Phi_{FB}(x; \mu)\| \leq \kappa\sqrt{\mu},$$

where $\kappa := \sqrt{2n}$. Let

$$\mu := (c/(2\kappa))^2 > 0.$$

Then, for any $x \in \mathcal{L}_{FB}(c/2; \mu)$, we have

$$\begin{aligned} \|\Phi_{FB}(x)\| &\leq \|\Phi_{FB}(x) - \Phi_{FB}(x; \mu)\| + \|\Phi_{FB}(x; \mu)\| \\ &\leq \kappa\sqrt{\mu} + c/2 \\ &= c. \end{aligned}$$

Hence, $x \in \mathcal{L}_{FB}(c)$ so that $\mathcal{L}(c/2; \mu)$ is also compact. Moreover, for any solution x^* of NCP(F), we have

$$\begin{aligned} \|\Phi_{FB}(x^*; \mu)\| &\leq \|\Phi_{FB}(x^*) - \Phi_{FB}(x^*; \mu)\| + \|\Phi_{FB}(x^*)\| \\ &\leq \kappa\sqrt{\mu} \\ &= c/2. \end{aligned}$$

This implies that the level set $\mathcal{L}_{FB}(c/2;\mu)$ is nonempty, since the solution set of NCP(F) , nonempty by assumption, belongs to the level set.

The continuous mapping $\|\Phi_{FB}(\cdot;\mu)\|$ therefore attains a global minimum x_μ which is obviously also a global minimum of the corresponding merit function

$$\Psi_{FB}(x;\mu) := \frac{1}{2}\Phi_{FB}(x;\mu)^T\Phi_{FB}(x;\mu).$$

Hence, x_μ is a stationary point of this merit function. Since F is a P_0-function by assumption, it follows from [21] that x_μ is actually a solution of the nonlinear system of equations

$$\Phi_{FB}(x;\mu) = 0.$$

In particular, x_μ satisfies the conditions

$$x_\mu > 0 \quad \text{and} \quad F(x_\mu) > 0,$$

see (8). Hence, x_μ is a strictly feasible point for NCP(F). □

As an interesting consequence of the previous proposition, we have the following result.

Corollary 3.13 *If F is a monotone function, then the following two statements are equivalent:*

1. *NCP(F) has a nonempty and bounded solution set.*
2. *NCP(F) has a strictly feasible point.*

Proof. It is well-known (see, e.g., [26]) that if NCP(F) has a strictly feasible point, then the solution set of NCP(F) is nonempty and bounded. The other direction follows immediately from Proposition 3.12. □

This corollary clearly shows that the assumption on the existence of a strictly feasible point cannot be removed from Statement 1 of Proposition 3.10 since the level set with level $c = 0$ is equal to the solution set of NCP(F).

However, in case F is a P_0-function, the existence of a strictly feasible point does not imply that the solution set of NCP(F) is bounded, as shown by the following example taken from [7]. Consider NCP(F) with $F(x) = Mx + q$, where

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \quad \text{and} \quad q = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

It is not difficult to verify that M is a P_0-matrix and that NCP(F) has the strictly feasible point $e = (1,1,1)^T$ with $F(e) = e$. However, the solution set of NCP(F) contains the unbounded line $(x_1,0,0)^T$ for all $x_1 \geq 0$.

We finally stress that the proof of Proposition 3.12 implies that, if the complementarity problem with a P_0-function has a nonempty and bounded solution set, then the system

$$x_i > 0, F_i(x) > 0, x_iF_i(x) = \mu \ \forall i$$

has a solution for each $\mu > 0$ sufficiently small, i.e., the central path used in interior-point methods exists for all $\mu > 0$ small enough. This observation might be useful for interior-point methods to solve the class of P_0-function complementarity problems. We formally state this result in the following corollary.

Corollary 3.14 *Suppose that F is a P_0-function such that NCP(F) has a nonempty and bounded solution set. Then the central path exists for all $\mu > 0$ sufficiently small, i.e., the system*

$$x_i > 0, \ \ F_i(x) > 0, \ \ x_iF_i(x) = \mu \quad \forall i$$

has a (necessarily unique) solution for all $\mu > 0$ sufficiently small.

Remark 3.15 *Ravindran and Gowda generalized Lemma 4.3 in [12] to continuous P_0 function in the recent version of [33, Remark 6]. By using their results, we can reduce the assumption that F is continuously differentiable to that F is continuous in Proposition 3.12, Corollary 3.13 and Corollary 3.14. Proposition 3.12 for a continuous P_0-function was independently derived by Ravindran and Gowda in [33].*

4 Semismooth Newton Method and Numerical Results

Based on the discussion of the previous two sections, the NCP-function ϕ_λ as well as the merit function Ψ_λ possess all the nice features of the Fischer-Burmeister function and the corresponding merit function. Therefore, by replacing ϕ_{FB} and Ψ_{FB} by ϕ_λ and Ψ_λ, respectively, in any semismooth based algorithm designed for the former functions, we can preserve and in some cases improve the convergence properties of the algorithm. In this section, we choose the modified damped Newton method proposed by De Luca, Facchinei and Kanzow [9] to test the new NCP-function. This algorithm, based on the new function, is guaranteed to solve not only complementarity problems with P_0- and R_0-functions, but also monotone problems with a strictly feasible point.

Algorithm 4.1 *(Semismooth Newton Method)*

(S.0) (Initialization)
Let $\beta \in (0,1)$, $\sigma \in (0, \frac{1}{2})$, $p > 2$, $\rho > 0$ and $\epsilon \geq 0$. Choose any $x^0 \in \mathbb{R}^n$. Set $k := 0$.

(S.1) (Termination Check)
If $\|\nabla \Psi_\lambda(x^k)\| \leq \epsilon$: STOP.

(S.2) (Search Direction Calculation)
Choose $V_k \in \partial_C \Phi_\lambda(x^k)$ *and let* $d^k \in \mathbb{R}^n$ *be a solution of the following linear system of equations:*

$$V_k d = -\Phi_\lambda(x^k).$$

If we cannot find a solution d^k *or if the descent test*

$$\nabla \Psi_\lambda(x^k)^T d^k \leq -\rho \|d^k\|^p$$

is not satisfied, set $d^k = -\nabla \Psi_\lambda(x^k)$.

(S.3) (Line Search)
Let ℓ_k *be the smallest nonnegative integer* ℓ *such that*

$$\Psi_\lambda(x^k + \beta^\ell d^k) \leq \Psi_\lambda(x^k) + \sigma \beta^\ell \nabla \Psi_\lambda(x^k)^T d^k.$$

Set $x^{k+1} = x^k + \beta^{\ell_k} d^k$, $k \leftarrow k+1$, *and go to (S.1).*

The convergence properties of this method are summarized in the following Theorem.

Theorem 4.2 *The following results hold for Algorithm 4.1:*

1. *The algorithm is well defined for an arbitrary NCP(F).*
2. *Any accumulation point is a stationary point of* Ψ_λ. *In particular, if* F *satisfies Condition 3.8, such an accumulation point exists.*
3. *Let* x^* *be any accumulation point such that* $F'(x^*)$ *is a* P_0*-matrix. Then* x^* *is a solution of NCP(F) .*
4. *If, in addition,* x^* *is an R-regular solution of NCP(F), then the whole sequence generated by Algorithm 4.1 converges to* x^*, *and the rate of convergence is Q-superlinear (Q-quadratic if* F *is an* LC^1 *function).*

Proof. Follows from Theorems 3.9, 3.4, 2.6 and the same arguments used for Theorem 11 of [9]. Instead of using Qi's [29] theory on the B-subdifferential, however, we have to apply the local convergence results for the C-subdifferential, also due to Qi [30]. □

We implemented Algorithm 4.1 in MATLAB using the following parameters:

$$\lambda = 0.95, \;\; \beta = 0.5, \;\; \sigma = 10^{-4}, \;\; \rho = 10^{-10}, \;\; p = 2.1.$$

We applied Algorithm 2.4 to compute an element in $\partial_C \Phi_\lambda(x^k)$. In addition, we incorporated some strategies to improve the numerical behaviour 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 a nonmonotone line search as introduced by Grippo, Lampariello and Lucidi [18], i.e., we compute $t_k = \max\{\beta^\ell \,|\, \ell = 0, 1, 2, \ldots\}$ such that

$$\Psi_\lambda(x^k + t_k d^k) \le \mathcal{R}_k + \sigma t_k \nabla\Psi_\lambda(x^k)^T d^k, \tag{9}$$

where the reference value $\mathcal{R}_k$ is given by

$$\mathcal{R}_k := \max_{j=k-m_k,\ldots,k} \Psi_\lambda(x^j)$$

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

$$m_k = 0$$

if either $k \le s$ or $d^k = -\nabla\Psi_\lambda(x^k)$ at the kth iterate (note that (9) reduces to the standard Armijo rule for this choice of m_k), whereas we set

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

at all other iterations. In our implementation, we use

$$m = 8 \quad \text{and} \quad s = 1.$$

This nonmonotone line search rule usually improves the efficiency of the algorithm, i.e., the number of iterations which are needed to solve a test example is often less than for the "monotone" Algorithm 4.1. However, sometimes the nonmonotone line search may also help to escape from local-nonglobal minima. As a result, it often improves the robustness of the overall algorithm.

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 use a simple backtracking strategy: Given an iterate $x^k \in \mathbb{R}^n$ and a search direction $d^k \in \mathbb{R}^n$, we first compute the exponent

$$j_k := \min\{0, 1, 2, \ldots\}$$

such that

$$F(x^k + \beta^{j_k} d^k)$$

exists and then take $\beta^{j_k} d^k$ as the new (shorter) search direction d^k. Despite its simplicity, this backtracking strategy works quite effectively.

In Tables 1 and 2, we report our results for all complementarity problems from the MCPLIB and GAMSLIB libraries, respectively (see [10, 14]), using all the different starting points which are available within the MATLAB framework. In particular, the columns in these tables have the following meanings:

problem:	name of test example,
n :	dimension of test example,
SP:	number of starting point (see the M-file `cpstart.m`),
k:	number of iterations,
F-eval.:	number of function evaluations,
Newton:	number of Newton steps,
gradient:	number of gradient steps,
$\Psi_\lambda(x^f)$:	function value of Ψ_λ at the final iterate x^f.

Note that the number of Jacobian evaluations of F is always equal to the number of iterations plus one more evaluation at the starting point. We also note that the major computational effort of Algorithm 4.1 is on solving the linear systems, and the number of the systems solved is obviously equal to the number of iterations.

We terminate the iteration if

$$\Psi_\lambda(x^k) \le \varepsilon \quad \text{or} \quad k > k_{\max} \quad \text{or} \quad t_k < t_{\min},$$

with

$$\varepsilon = 10^{-12}, \;\; k_{\max} = 500 \quad \text{and} \quad t_{\min} = 10^{-25}.$$

Note that there is just one starting point available for all examples from the GAMSLIB collection. We therefore excluded the problems `cammge`, `co2mge`, `finmge`, `sammge`, `shovmge`, and `threemge` from Table 2 since the starting points for these examples are either very close or sometimes even equal to the solution.

The results in Tables 1 and 2 indicate that the new algorithm works extremely well. In fact, the algorithm was able to solve all test examples. Well, we should say that the termination criterion

$$\Psi_\lambda(x^k) \le 10^{-12}$$

was not fully reached for example `pgvon106`, but the final value of Ψ_λ is very close to this stopping criterion, so we view this problem as being solved (note also that we received a small function value of approximately 10^{-10} after 45 steps).

For all other test examples, $\Psi_\lambda(x^f)$ is less than $\epsilon = 10^{-12}$. In particular, this means that we were able to solve problems `billups`, `vonthmge`, `bertsekas`, `colvdual`, `hanskoop`, `pgvon105`, `scarfbnum`, `scarfbsum`, `dmcmge`, `vonthmcp` and `vonthmge` which, in general, are the difficult test problems (we note, however, that problem `billups` was solved just by chance). In particular, we stress that, as known to the authors, there is currently no other algorithm which can solve problem `vonthmcp` (up to the desired accuracy and by using the default values for the parameters of the corresponding algorithm).

Besides these advantages, we also stress that we did not observe any irregular behaviour of our algorithm on problems `josephy` and `kojshin`, which often occured to many other semismooth methods based on the Fischer-Burmeister function.

We also tried a couple of further starting points for the difficult test problems (which are not part of the file `cpstart.m` available via anonymous ftp from Madison)

Table 1: Results for MCPLIB test problems

problem	n	SP	k	F-eval.	Newton	gradient	$\Psi_\lambda(x^f)$
bertsekas	15	1	28	219	28	0	1.8e-15
bertsekas	15	2	19	72	19	0	1.1e-14
bertsekas	15	3	27	34	27	0	2.7e-16
billups	1	1	441	5260	441	0	4.2e-13
colvdual	20	1	14	16	14	0	5.0e-15
colvdual	20	2	37	49	37	0	2.6e-19
colvnlp	15	1	14	16	14	0	4.4e-16
colvnlp	15	2	15	16	15	0	6.5e-16
cycle	1	1	7	8	7	0	8.8e-21
explcp	16	1	19	20	19	0	5.2e-14
hanskoop	14	1	10	12	9	1	8.2e-16
hanskoop	14	2	13	14	12	1	1.3e-14
hanskoop	14	3	10	12	9	1	1.9e-20
hanskoop	14	4	12	18	12	0	2.7e-19
hanskoop	14	5	11	15	10	1	3.1e-14
josephy	4	1	7	8	7	0	2.7e-20
josephy	4	2	6	8	6	0	2.7e-25
josephy	4	3	15	16	15	0	1.1e-23
josephy	4	4	5	6	5	0	4.6e-14
josephy	4	5	4	5	4	0	9.0e-14
josephy	4	6	6	8	6	0	1.6e-14
kojshin	4	1	13	15	13	0	5.1e-20
kojshin	4	2	7	10	7	0	2.3e-22
kojshin	4	3	15	16	15	0	8.4e-21
kojshin	4	4	4	5	4	0	6.0e-13
kojshin	4	5	4	5	4	0	4.0e-13
kojshin	4	6	5	7	5	0	5.0e-15
mathinum	3	1	5	7	5	0	1.3e-23
mathinum	3	2	5	6	5	0	3.4e-15
mathinum	3	3	10	15	10	0	4.7e-13
mathinum	3	4	6	7	6	0	2.8e-15
mathisum	4	1	5	7	5	0	1.9e-18
mathisum	4	2	6	7	6	0	9.9e-15
mathisum	4	3	8	9	8	0	7.4e-23
mathisum	4	4	6	7	6	0	6.3e-22
nash	10	1	8	9	8	0	8.7e-20
nash	10	2	10	11	10	0	2.5e-18

Table 1 (continued): Results for MCPLIB test problems

problem	n	SP	k	F-eval.	Newton	gradient	$\Psi_\lambda(x^f)$
pgvon105	105	1	43	88	43	0	2.0e-15
pgvon106	106	1	500	3864	179	121	5.2e-11
powell	16	1	10	12	10	0	7.9e-19
powell	16	2	9	10	8	1	1.9e-18
powell	16	3	15	18	14	1	1.2e-19
powell	16	4	11	12	10	1	6.4e-17
scarfanum	13	1	11	13	11	0	1.5e-16
scarfanum	13	2	12	17	12	0	1.5e-16
scarfanum	13	3	10	12	10	0	6.6e-17
scarfasum	14	1	7	9	7	0	1.6e-16
scarfasum	14	2	11	14	11	0	1.6e-16
scarfasum	14	3	11	13	11	0	1.6e-16
scarfbnum	39	1	21	23	21	0	4.9e-16
scarfbnum	39	2	18	21	17	1	4.7e-13
scarfbsum	40	1	20	29	20	0	5.9e-20
scarfbsum	40	2	26	28	25	1	2.1e-14
sppe	27	1	10	32	8	2	2.1e-18
sppe	27	2	6	7	6	0	9.7e-22
tobin	42	1	8	11	8	0	1.4e-19
tobin	42	2	10	13	10	0	5.9e-16

and, again, the new algorithm behaves favourably compared to many other existing methods. So the theoretical advantages of the new NCP-function and, especially, of its corresponding merit function can also be seen numerically. On the other hand, we admit that we are actually a bit surprised about these very good numerical behaviour since taking $\lambda = 0.95$ means that the new NCP-function is quite close to the Fischer-Burmeister NCP-function, so that we had not expected such a drastic difference in the numerical behaviour. The nice bounded level set property of Φ_λ can be one reason for the good numerical results.

In order to give another reasoning why the new function works better than the Fischer-Burmeister NCP-function, consider a simple problem with $n = 1$ and $F(x) = 1$. Clearly, $x^* = 0$ is the unique solution of the corresponding complementarity problem. Now suppose that we have an iterate x^k with, say, $x^k \approx 10^{12}$. Then, especially due to rounding errors and cancellation, the Fischer-Burmeister merit function Ψ_{FB} has a very small function value at the iterate x^k. This indicates that x^k is quite close to the solution of NCP(F). However, in reality, x^k is far away from the solution $x^* = 0$, so the information provided by the Fischer-Burmeister function is totally wrong.

Table 2: Results for GAMSLIB test problems

problem	n	SP	k	F-eval.	Newton	gradient	$\Psi_\lambda(x^f)$
cafemge	101	1	9	11	9	0	5.7e-19
dmcmge	170	1	22	51	22	0	7.4e-16
etamge	114	1	16	25	16	0	9.4e-14
hansmcp	43	1	19	46	17	2	1.2e-16
hansmge	43	1	16	19	16	0	1.0e-20
harkmcp	32	1	9	10	8	1	3.7e-14
kehomge	9	1	10	11	10	0	1.9e-17
mr5mcp	350	1	9	11	9	0	3.8e-14
nsmge	212	1	18	21	18	0	6.2e-18
oligomcp	6	1	6	7	6	0	4.9e-17
scarfmcp	18	1	11	12	11	0	1.5e-16
scarfmge	18	1	11	12	11	0	6.6e-17
transmcp	11	1	13	38	11	2	9.0e-17
two3mcp	6	1	7	8	7	0	4.1e-21
unstmge	5	1	9	11	9	0	7.7e-22
vonthmcp	125	1	58	335	48	10	1.5e-14
vonthmge	80	1	37	324	33	4	9.9e-22

On the other hand, the function value of our new merit function Ψ_λ is very large at x^k since we put more penalization onto the positive orthant, and this gives us a much better information. Since cancellation does not occur outside the positive orthant, there is no need to modify the Fischer-Burmeister function outside this region, and this is exactly what our new NCP-function does.

5 Final Remarks and Extensions

In this paper, we proposed a new NCP-function Φ_λ by adding an extra term a_+b_+ to the Fischer-Burmeister function Φ_{FB}. The new NCP-function Φ_λ possesses all the nice properties of Φ_{FB} for local convergence and its natural merit function Ψ_λ has all the nice features of the Kanzow-Yamashita-Fukushima merit function for bounded level sets and global convergence. We are not aware of any other NCP-function which has the same strong theoretical properties as our function. Numerical results indicate that the semismooth Newton-type method based on the new NCP-function is extremely promising. In fact, we believe that our new NCP-function is close to be an "optimal" NCP-function.

We finally discuss some possible extensions of the new NCP-function. We first note that there is no need to define ϕ_λ as a convex combination of the Fischer-

Burmeister function ϕ_{FB} and the term a_+b_+. Any positive linear combination would also work. By taking a convex combination, however, we only need to control one parameter instead of two, which is simpler from a computational point of view.

We further observe that the extra term a_+b_+ can also be added to any other NCP-function in order to improve the global bounded level set properties of that NCP-function. For example, by adding the extra term to Qi's piecewise rational NCP-function [31], we obtain a new piecewise rational NCP-function, which has all the nice properties of Φ_λ.

We may also use the new NCP-function to construct a smoothing method. For example, if we replace the Fischer-Burmeister function by its smooth counterpart from [21] (cf. the proof of Proposition 3.12) and if we replace the plus-function by any smooth function taken from the class considered by Chen and Mangasarian [5], we obtain a smooth version of ϕ_λ. Due to our bounded level set result for monotone problems, this new function should simplify the local convergence analysis of some recent smoothing methods considerably. In particular, it is probably no longer necessary to dynamically adjust neighbourhoods for monotone problems, as done in the smoothing method by Chen and Chen [3].

Finally, it is also possible to extend the idea to box constrained variational inequality problems (sometimes also called mixed complementarity problems, see, e.g., [10, 11]). Let

$$X = \{x \in \mathbb{R}^n \mid l \leq x \leq u\}$$

where $l \in \{\mathbb{R} \cup \{-\infty\}\}^n$, $u \in \{\mathbb{R} \cup \{-\infty\}\}^n$ and $l < u$ are given vectors. The box constrained variational inequality problem, denoted by BVIP(F), is to find an $x \in X$ such that for all $y \in X$,

$$(y - x)^T F(x) \geq 0.$$

Obviously, if $l = 0$ and $u = \infty$, BVIP(F) reduces to NCP(F). Central to a semismooth Newton-type algorithm for BVIP(F) is a so-called BVIP-function, which is a generalization of the concept of an NCP-function. Let $\alpha \in \mathbb{R} \cup \{-\infty\}$, $\beta \in \mathbb{R} \cup \{+\infty\}$ and $\alpha \leq \beta$. Denote

$$B_0 := \{(a, b) : b \geq 0\} \cup \{(a, 0) : \alpha \leq a \leq \beta\} \cup \{(\beta, b) : b \leq 0\}.$$

$\phi_{(\alpha,\beta)} : \mathbb{R}^2 \to \mathbb{R}$ is called a BVIP-function [31] if $\phi_{(\alpha,\beta)}(a, b) = 0$ if and only if $(a, b) \in B_0$. Billups [1] and Sun and Womersley [35] generalized the Fischer-Burmeister function to corresponding BVIP-functions, and Qi [31] constructed a BVIP-function based on his piecewise rational NCP-function. The following observation allows us to generalize the extra term a_+b_+ to the case of BVIP. It is shown in [6] that $(a, b) \in B_0$ if and only if

$$q(a, b) := \min(a - \alpha, \beta - b) \geq 0$$

$$p(a, b) := \mathrm{sign}(\frac{\alpha + \beta}{2} - a)b \geq 0$$

and

$$q(a,b)p(a,b) = 0.$$

As a result, $[q(a,b)]_+[p(a,b)]_+$ corresponds to the extra term a_+b_+ for complementarity problems. Indeed, if $\alpha = 0$ and $\beta = \infty$, $[q(a,b)]_+[p(a,b)]_+$ reduces to a_+b_+. Therefore, by adding this extra term to the Billups BVIP-function or other BVIP-functions properly, we may improve their global bounded level set properties.

Acknowledgement: The authors are grateful to Prof. J. Stoer for pointing out that the new NCP function ϕ_λ gives penalties in the positive orthant, and to Prof. M.S. Gowda for his helpful comments on an earlier version of this paper, which provide Remark 4.1.

References

[1] S.C. Billups: *Algorithms for Complementarity Problems and Generalized Equations.* Ph.D. Thesis, Computer Sciences Department, University of Wisconsin, Madison, WI, 1995.

[2] B. Chen: *Error bounds for R_0-type and monotone nonlinear complementarity problems.* Technical Report, Department of Management and Systems, Washington State University, Pullman, WA, 1997.

[3] B. Chen and X. Chen: *A global and local superlinear continuation-smoothing method for $P_0 + R_0$ and monotone NCP.* Technical Report, Department of Mangagement and Systems, Washington State University, Pullman, WA, 1997.

[4] B. Chen and P.T. Harker: *Smooth approximations to nonlinear complementarity problems.* SIAM Journal on Optimization 7, 1997, pp. 403–420.

[5] C. Chen and O.L. Mangasarian: *A class of smoothing functions for nonlinear and mixed complementarity problems.* Computational Optimization and Applications 2, 1996, pp. 97–138.

[6] X. Chen and Y. Ye: *On homotopy-smoothing methods for variational inequalities.* Technical Report AMR 96/39, The University of New South Wales, Sydney, Australia, 1996.

[7] X. Chen and Y. Ye: *On homogeneous methods for the P_0 matrix linear complementarity problem.* The University of New South Wales, Sydney, Australia, 1997.

[8] F.H. Clarke: *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, NY, 1983 (reprinted by SIAM, Philadelphia, PA, 1990).

[9] T. De Luca, F. Facchinei and C. Kanzow: *A semismooth equation approach to the solution of nonlinear complementarity problems.* Mathematical Programming 75, 1996, pp. 407–439.

[10] S.P. Dirkse and M.C. Ferris: *MCPLIB: A collection of nonlinear mixed complementarity problems.* Optimization Methods and Software 5, 1995, pp. 319–345.

[11] S.P. Dirkse and M.C. Ferris: *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems.* Optimization Methods and Software 5, 1995, pp. 123–156.

[12] F. Facchinei: *Structural and stability properties of P_0 nonlinear complementarity problems.* Technical Report, Università di Roma "La Sapienza", Dipartimento di Informatica e Sistemistica, Rome, Italy, 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, pp. 225–247.

[14] M.C. Ferris and T.F. Rutherford: *Accessing realistic mixed complementarity problems within MATLAB.* In: G. Di Pillo and F. Giannessi (eds.): *Nonlinear Optimization and Applications.* Plenum Press, New York, 1996, pp. 141–153.

[15] A. Fischer: *A special Newton-type optimization method.* Optimization 24, 1992, pp. 269–284.

[16] A. Fischer: *Solution of monotone complementarity problems with locally Lipschitzian functions.* Mathematical Programming 76, 1997, pp. 513–532.

[17] M. Fukushima: *Merit functions for variational inequality and complementarity problems.* In: G. Di Pillo and F. Giannessi (eds.): *Nonlinear Optimization and Applications.* Plenum Press, New York, NY, 1996, pp. 155–170.

[18] 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.

[19] P.T. Harker and J.-S. Pang: *Finite dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications.* Mathematical Programming 48, 1990, pp. 161–220.

[20] H. Jiang and D. Ralph: *Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares.* Technical Report, Department of Mathematics, The University of Melbourne, Melbourne, Australia, 1997.

[21] C. Kanzow: *Some noninterior continuation methods for linear complementarity problems.* SIAM Journal on Matrix Analysis and Applications 17, 1996, pp. 851–868.

[22] C. Kanzow: *A new approach to continuation methods for complementarity problems with uniform P-functions.* Operations Research Letters 20, 1997, pp. 85–92.

[23] C. Kanzow: *Semismooth Newton-type Methods for the Solution of Nonlinear Complementarity Problems.* Habilitation Thesis, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, 1997.

[24] C. Kanzow and H. Kleinmichel: *A new class of semismooth Newton-type methods for nonlinear complementarity problems.* Preprint 118, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, 1997.

[25] C. Kanzow, N. Yamashita and M. Fukushima: *New NCP-functions and their properties.* Journal of Optimization Theory and Applications 94, 1997, pp. 115–135.

[26] M. Kojima, S. Mizuno and T. Noma: *Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems.* Mathematics of Operations Research 15, 1990, pp. 662–675.

[27] Z.-Q. Luo and P. Tseng: *A new class of merit functions for the nonlinear complementarity problem.* In: M.C. Ferris and J.-S. Pang (eds.): *Complementarity and Variational Problems: State of the Art.* SIAM, Philadelphia, PA, 1997, pp. 204–225.

[28] R. Mifflin: *Semismooth and semiconvex functions in constrained optimization.* SIAM Journal on Control and Optimization 15, 1977, pp. 959–972.

[29] L. Qi: *Convergence analysis of some algorithms for solving nonsmooth equations.* Mathematics of Operations Research 18, 1993, pp. 227–244.

[30] L. Qi: *C-differentiability, C-differential operators and generalized Newton methods.* Technical Report, The University of New South Wales, Sydney, Australia, 1996.

[31] L. Qi: *Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems.* Technical Report AMR 97/14, The University of New South Wales, Sydney, Australia, 1997.

[32] L. Qi and H. Jiang: *Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton methods and quasi-Newton methods for solving these equations.* Mathematics of Operations Research 22, 1997, pp. 301–325.

[33] G. Ravindran and M.S. Gowda: *Regularization of P_0-functions in box variational inequality problems.* Technical Report, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD, September 1997.

[34] S.M. Robinson: *Strongly regular generalized equations.* Mathematics of Operations Research 5, 1980, pp. 43–62.

[35] D. Sun and R.S. Womersley: *A new unconstrained differentiable merit function for box constrained variational inequality problems and a damped Gauss-Newton method.* Technical Report AMR 96/37, The University of New South Wales, Sydney, Australia, 1996.

[36] P. Tseng: *Growth behaviour of a class of merit functions for the nonlinear complementarity problem.* Journal of Optimization Theory and Applications 89, 1996, pp. 17–37.