Hamburger Beiträge
zur Angewandten Mathematik

Inexact Trust-Region Methods for Nonlinear Complementarity Problems

Christian Kanzow and Martin Zupke

Reihe A
Preprint 127
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

INEXACT TRUST-REGION METHODS FOR NONLINEAR COMPLEMENTARITY PROBLEMS

Christian Kanzow and Martin Zupke

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

September 12, 1997

Abstract: In order to solve the nonlinear complementarity problem, we first reformulate it as a nonsmooth system of equations by using a recently introduced NCP-function. We then apply a trust-region-type method to this system of equations. Our trust-region method allows an inexact solution of the trust-region-subproblem. We show that the algorithm is well-defined for a general nonlinear complementarity problem and that it has some nice global and local convergence properties. Numerical results indicate that the new method is quite promising.

Key Words: Nonlinear complementarity problem, trust-region method, nonsmooth Newton method, global convergence, quadratic convergence.

1 Introduction

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. We consider the problem of finding a solution $x^* \in \mathbb{R}^n$ of the following system of equations and inequalities:

$$x_i \geq 0, \ \ F_i(x) \geq 0, \ \ x_i F_i(x) = 0 \quad \forall i \in I := \{1, \dots, n\}.$$

This is a *nonlinear complementarity problem* and we will denote it by NCP(F) throughout this paper. It has many important applications, and we refer the reader to the recent survey paper [10] by Ferris and Pang for a brief description of several economic and engineering applications.

Due to the importance of the complementarity problem, there exists a huge number of methods for solving problem NCP(F). Here we concentrate on a particular approach which is based on a reformulation of NCP(F) as a nonlinear system of equations. This reformulation is done by means of a so-called *NCP-function* $\varphi : \mathbb{R}^2 \to \mathbb{R}$ which is defined by the following characterization of its zeros:

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

In this paper, we will use

$$\varphi_\lambda(a, b) := \sqrt{(a-b)^2 + \lambda ab} - a - b$$

as an NCP-function, where $\lambda \in (0, 4)$ is a given parameter. This function was introduced by Kanzow and Kleinmichel [16] and covers the well-known Fischer-Burmeister function [11] by taking $\lambda = 2$ as well as the minimum function [23] in the limiting case $\lambda \to 0$.

Since φ_λ is an NCP-function, it is easy to see that a vector $x^* \in \mathbb{R}^n$ solves the nonlinear complementarity problem if and only if it is a solution of the nonsmooth system of equations

$$\Phi_\lambda(x) = 0, \tag{1}$$

where $\Phi_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ is defined by

$$\Phi_\lambda(x) := \begin{pmatrix} \varphi_\lambda(x_1, F_1(x)) \\ \vdots \\ \varphi_\lambda(x_n, F_n(x)) \end{pmatrix}.$$

Alternatively, we may view

$$\min \ \ \Psi_\lambda(x), \quad x \in \mathbb{R}^n,$$

with $\Psi_\lambda : \mathbb{R}^n \to \mathbb{R}$ given by

$$\Psi_\lambda(x) := \frac{1}{2}\Phi_\lambda(x)^T \Phi_\lambda(x)$$

as an unconstrained minimization reformulation of NCP(F). We stress that, despite the nonsmoothness of the equation-operator Φ_λ, it is not difficult to see that its merit function Ψ_λ is continuously differentiable [16, Theorem 3.1].

Many algorithms have been proposed for solving problem NCP(F) by exploiting one of these reformulations, see, e.g., [9, 4, 16]. Most of these algorithms are globalized by using an Armijo-type line search. Another well-known globalization strategy is based on the trust-region idea. Due to the nonsmoothness of the operator Φ_λ and related reformulations of NCP(F), however, it is not possible to use a standard trust-region method from smooth optimization in order to solve the complementarity problem.

So one has to modify these smooth trust-region methods in order to take into account the nonsmoothness of Φ_λ or similar operators. Modifications of this kind were described by Gabriel and Pang [14], Friedlander, Martínez and Santos [13] and Martínez and Santos [20]. On the other hand, our merit function Ψ_λ is smooth, and this additional property was fully exploited in the recent paper [15] by Jiang, Fukushima, Qi and Sun.

The latter paper will therefore be the basis of our current research. In fact, we borrow many ideas from [15]. In contrast to [15], however, our approach is based on the reformulation (1) (Jiang et al. [15] use the Fischer-Burmeister function) and, more important, we allow an inexact solution of the trust-region subproblem. By doing this, we improve the overall efficiency of our algorithm.

Another motivation for this research derives from the fact that trust-region methods are often said to be more reliable than corresponding line-search methods for smooth problems [12]. Hence there seems to be some hope to improve at least the robustness of existing line-search methods by using a trust-region-type globalization.

The paper is organized as follows: In Section 2, we mainly restate some properties of the functions Φ_λ and Ψ_λ from [16]. The inexact trust-region method as well as its global and fast local convergence are considered in Section 3. In Section 4, we give some details about our implementation of the inexact trust-region method and report extensive numerical results. We conclude this paper with some final remarks in Section 5.

Notation: All vector norms used in this paper are Euclidian norms. Matrix norms are assumed to be consistent with the Euclidian vector norm. A mapping $G : \mathbb{R}^n \to \mathbb{R}^m$ is called a C^1 function if G is continuously differentiable, and LC^1 function if G is differentiable with a locally Lipschitzian Jacobian. If G is a C^1 mapping, we denote its Jacobian at a point $x \in \mathbb{R}^n$ by $G'(x)$, whereas $\nabla G(x)$ denotes the transposed Jacobian. In particular, if G is a real-valued mapping, we view its gradient $\nabla G(x)$ as a column vector. If G is only locally Lipschitzian, its Jacobian may not exist on a set of measure zero. In this case, let us denote by D_G the set of differentiable points of G. If $x \notin D_G$, we may view either the B-subdifferential

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

or Clarke's [2] generalized Jacobian

$$\partial G(x) := \mathrm{conv}\partial_B G(x)$$

of G at a point $x \in \mathbb{R}^n$ as a suitable substitute of the Jacobian. In the case $m = n$, we call a solution x^* of the system $G(x) = 0$ BD-regular if all elements in the B-subdifferential $\partial_B G(x^*)$ are nonsingular.

2 Preliminaries

In this section, we recall some of the basic properties of the functions Φ_λ and Ψ_λ and refer to [16] for the corresponding proofs. In addition, we slightly improve the stationary point result from [16] by providing a full characterization for a stationary point of Ψ_λ to be a solution of NCP(F).

We first restate the following overestimation of the generalized Jacobian of Φ_λ at an arbitrary point $x \in \mathbb{R}^n$, see [16, Proposition 2.5].

Proposition 2.1 *For an arbitrary $x \in \mathbb{R}^n$, we have*

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

where $D_a(x) = diag(a_1(x), \ldots, a_n(x)), D_b(x) = diag(b_1(x), \ldots, b_n(x)) \in \mathbb{R}^{n\times n}$ are diagonal matrices whose ith diagonal element is given by

$$a_i(x) = \frac{2(x_i - F_i(x)) + \lambda F_i(x)}{2\sqrt{(x_i - F_i(x))^2 + \lambda x_i F_i(x)}} - 1, \;\; b_i(x) = \frac{-2(x_i - F_i(x)) + \lambda x_i}{2\sqrt{(x_i - F_i(x))^2 + \lambda x_i F_i(x)}} - 1$$

if $(x_i, F_i(x)) \neq (0,0)$, and by

$$a_i(x) = \xi_i - 1, \;\; b_i(x) = \chi_i - 1 \textit{ for any } (\xi_i, \chi_i) \in \mathbb{R}^2 \textit{ such that } \|(\xi_i, \chi_i)\| \leq \sqrt{c_\lambda}$$

if $(x_i, F_i(x)) = (0,0)$, where c_λ is a certain constant (depending on λ) with $c_\lambda \in (0,2)$.

We next want to give a characterization for stationary points of our merit function Ψ_λ to be a solution of the complementarity problem. To this end, we follow the approach by De Luca et al. [4] and define the three index sets

$$\begin{array}{rcll} \mathcal{C}(x) & := & \{i \in I \,|\, x_i \geq 0, F_i(x) \geq 0, x_i F_i(x) = 0\} & (\text{“\underline{c}omplementary indices”}), \\ \mathcal{P}(x) & := & \{i \in I \,|\, x_i > 0, F_i(x) > 0\} & (\text{“\underline{p}ositive indices”}), \\ \mathcal{N}(x) & := & I \setminus (\mathcal{C}(x) \cup \mathcal{P}(x)) & (\text{“\underline{n}egative indices”}), \end{array}$$

where $x \in \mathbb{R}^n$ is any given vector. It is easy to see that the following relationships hold which also motivate the names for these index sets:

$$\begin{array}{lclcl} [D_a(x)\Phi_\lambda(x)]_i > 0 & \Longleftrightarrow & [D_b(x)\Phi_\lambda(x)]_i > 0 & \Longleftrightarrow & i \in \mathcal{P}(x), \\ [D_a(x)\Phi_\lambda(x)]_i = 0 & \Longleftrightarrow & [D_b(x)\Phi_\lambda(x)]_i = 0 & \Longleftrightarrow & i \in \mathcal{C}(x), \\ [D_a(x)\Phi_\lambda(x)]_i < 0 & \Longleftrightarrow & [D_b(x)\Phi_\lambda(x)]_i < 0 & \Longleftrightarrow & i \in \mathcal{N}(x). \end{array} \tag{2}$$

We will sometimes write $\mathcal{C}, \mathcal{P}$ and $\mathcal{N}$ instead of $\mathcal{C}(x), \mathcal{P}(x)$ and $\mathcal{N}(x)$, respectively, if the vector x is clear from the context.

Based on these index sets, we follow De Luca et al. [4] and call a vector $x \in \mathbb{R}^n$ *regular* for NCP(F) if, for any nonzero vector $z \in \mathbb{R}^n$ with

$$z_{\mathcal{C}} = 0, \quad z_{\mathcal{P}} > 0, \quad z_{\mathcal{N}} < 0,$$

there exists a vector $y \in \mathbb{R}^n$ such that

$$y_{\mathcal{P}} \geq 0, \quad y_{\mathcal{N}} \leq 0, \quad y_{\mathcal{P} \cup \mathcal{N}} \neq 0$$

and

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

The regularity of a vector plays a central role in the following characterization result.

Theorem 2.2 *A vector $x \in \mathbb{R}^n$ is a solution of NCP(F) if and only if x is a regular stationary point of Ψ_λ.*

Proof. Based on the relations (2), the proof is essentially the same as the one given in [4, Theorem 14]. □

We mention two simple consequences of this result. These consequences are based on certain classes of matrices for which we refer the reader to the excellent book [3] by Cottle, Pang and Stone.

Corollary 2.3 *Assume that $x^* \in \mathbb{R}^n$ is a stationary point of Ψ_λ.*

(a) If $F'(x^*)$ is a P_0-matrix, then x^* solves NCP(F).

(b) If $F'(x^*)$ is semimonotone and x^* is feasible for NCP(F), then x^* solves NCP(F).

Proof. Part (a) can be verified similarly to [18, Lemma 5.2]. In order to prove part (b), we first note that $\mathcal{N}(x^*) = \emptyset$ due to the feasibility of x^*. Let $z \in \mathbb{R}^n$ be any nonzero vector such that

$$z_{\mathcal{C}} = 0 \quad \text{and} \quad z_{\mathcal{P}} > 0.$$

Since $F'(x^*)$ is semimonotone by assumption, its transpose $\nabla F(x^*)$ is also semimonotone by [3, Corollary 3.9.7]. Hence there exists an index $i_0 \in \mathcal{P}$ such that

$$z_{i_0}[\nabla F(x^*) z]_{i_0} \geq 0.$$

Let $y \in \mathbb{R}^n$ be defined by $y_{i_0} := z_{i_0}$ and $y_i := 0$ for all $i \neq i_0$. Then

$$y^T \nabla F(x^*) z = \sum_{i=1}^n y_i [\nabla F(x^*) z]_i = z_{i_0}[\nabla F(x^*) z]_{i_0} \geq 0.$$

Hence x^* is a regular stationary point of Ψ_λ and therefore a solution of NCP(F) by Theorem 2.2. $\square$

Before restating the next result from [16], let us define the index sets

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

where $x^* \in \mathbb{R}^n$ is a fixed solution of the complementarity problem. Then x^* is called an *R-regular* solution of NCP(F) if the submatrix $F'(x^*)_{\alpha\alpha}$ is nonsingular and the Schur-complement (see [3])

$$F'(x^*)_{\beta\beta} - F'(x^*)_{\beta\alpha} F'(x^*)_{\alpha\alpha}^{-1} F'(x^*)_{\alpha\beta} \in \mathbb{R}^{|\beta|\times|\beta|}$$

is a P-matrix, cf. Robinson [28]. Using the concept of R-regularity, the following result was shown in [16, Theorem 2.7].

Proposition 2.4 *Assume that $x^* \in \mathbb{R}^n$ is an R-regular solution of NCP(F). Then all elements in the generalized Jacobian $\partial\Phi_\lambda(x^*)$ are nonsingular.*

We finally mention a result which follows immediately from the (strong) semismoothness of the equation operator Φ_λ and well-known results for (strongly) semismooth functions, see [16, 27].

Proposition 2.5 *Let $x^* \in \mathbb{R}^n$ be the limit point of a sequence $\{x^k\} \subseteq \mathbb{R}^n$. Then*

$$\|\Phi_\lambda(x^k) - \Phi_\lambda(x^*) - H_k(x^k - x^*)\| = o(\|x^k - x^*\|)$$

for any $H_k \in \partial\Phi_\lambda(x^k)$. If, in addition, F is an LC^1 mapping, then

$$\|\Phi_\lambda(x^k) - \Phi_\lambda(x^*) - H_k(x^k - x^*)\| = O(\|x^k - x^*\|^2).$$

3 Inexact Trust-Region Methods

3.1 Algorithm

In this subsection, we first state our algorithm and then show that it is well-defined for an arbitrary complementarity problem.

To this end, we recall that NCP(F) can be written as the unconstrained minimization problem

$$\min \ \Psi_\lambda(x), \quad x \in \mathbb{R}^n.$$

Using a Gauss-Newton-type linearization around a given iterate $x^k \in \mathbb{R}^n$, we obtain the following quadratic approximation of the function $\Psi_\lambda(x^k + \cdot)$:

$$q_k(d) := \frac{1}{2}\|\Phi_\lambda(x^k) + H_k d\|^2,$$

where $H_k \in \partial_B \Phi_\lambda(x^k)$. Since $\nabla \Psi_\lambda(x^k) = H_k^T \Phi_\lambda(x^k)$ for any $H_k \in \partial \Phi_\lambda(x^k)$ by Theorem 3.1 in [16], it follows immediately that q_k can be written as

$$q_k(d) = \Psi_\lambda(x^k) + \nabla \Psi_\lambda(x^k)^T d + \frac{1}{2} d^T H_k^T H_k d;$$

this representation of q_k will be more convenient in our subsequent analysis.

We are now able to state our trust-region-type method.

Algorithm 3.1 *(Inexact Trust-Region Method)*

(S.0) (Initialization)
Choose $\lambda \in (0,4)$, $x^0 \in \mathbb{R}^n$, $\Delta_0 > 0$, $0 < \rho_1 < \rho_2 < 1$, $0 < \sigma_1 < 1 < \sigma_2$, $\Delta_{\min} > 0$, $\varepsilon \geq 0$, *and set* $k := 0$.

(S.1) (Termination Criterion)
If $\|\nabla \Psi_\lambda(x^k)\| \leq \varepsilon$: *STOP.*

(S.2) (Subproblem Solution)
Select an element $H_k \in \partial_B \Phi_\lambda(x^k)$. *Compute a solution* $d^k \in \mathbb{R}^n$ *of the subspace trust-region problem*

$$\min \quad q_k(d) \quad s.t. \quad \|d\| \leq \Delta_k, \quad d \in V_k, \tag{3}$$

where $V_k \subseteq \mathbb{R}^n$ *denotes a suitable subspace which will be specified later.*

(S.3) (Updates)
Compute

$$r_k := \frac{\Psi_\lambda(x^k) - \Psi_\lambda(x^k + d^k)}{\Psi_\lambda(x^k) - q_k(d^k)}.$$

If $r_k \geq \rho_1$, *we call the iteration* k successful *and set* $x^{k+1} := x^k + d^k$; *otherwise we set* $x^{k+1} := x^k$.

(a) If $r_k < \rho_1$, *set* $\Delta_{k+1} := \sigma_1 \Delta_k$.

(b) If $r_k \in [\rho_1, \rho_2)$, *set* $\Delta_{k+1} := \max\{\Delta_{\min}, \Delta_k\}$.

(c) If $r_k \geq \rho_2$, *set* $\Delta_{k+1} := \max\{\Delta_{\min}, \sigma_2 \Delta_k\}$.

Set $k \leftarrow k+1$ *and go to (S.1).*

Basically, Algorithm 3.1 is a standard trust-region method for the solution of nonlinear systems of equations. The only difference lies in the updating rule for the trust-region radius after a successful iteration; Algorithm 3.1 uses a minimal radius $\Delta_{\min} > 0$ as a lower bound for the new radius. This idea was also used by Jiang et al. [15] and some other works on nonsmooth trust-region methods [14, 13, 20]. Since

Φ_λ is not differentiable everywhere, this modified updating rule is necessary in order to ensure global and local superlinear/quadratic convergence of our nonsmooth trust-region method.

Throughout this section, we assume that the termination parameter ε is equal to 0 and that the algorithm does not terminate after a finite number of iterations. We further assume that the subspace V_k from Step (S.2) of Algorithm 3.1 contains at least the gradient direction $\nabla\Psi_\lambda(x^k)$. Then the following result is well-known, see Powell [25, Theorem 4].

Lemma 3.2 *Assume that $d^k \in \mathbb{R}^n$ is a solution of the subproblem (3). Then*

$$\Psi_\lambda(x^k) - q_k(d^k) \geq \frac{1}{2}\|\nabla\Psi_\lambda(x^k)\| \min\left\{\Delta_k, \frac{\|\nabla\Psi_\lambda(x^k)\|}{\|H_k^T H_k\|}\right\}.$$

Note that the denominator on the right-hand side of the displayed formula in Lemma 3.2 is always nonzero; otherwise, we would have $H_k = 0$ which would imply $\nabla\Psi_\lambda(x^k) = H_k^T\Phi_\lambda(x^k) = 0$, so that the algorithm would have stopped in Step (S.1).

As an immediate consequence of Lemma 3.2, we obtain that Algorithm 3.1 can be applied to an arbitrary complementarity problem.

Proposition 3.3 *Algorithm 3.1 is well-defined.*

Proof. We only have to show that the denominator

$$\Psi_\lambda(x^k) - q_k(d^k)$$

in the definition of r_k in Step (S.3) is nonzero for all $k \in \mathbb{N}$. Assume this is not true. Then $\nabla\Psi_\lambda(x^k) = 0$ because of Lemma 3.2. Hence the algorithm would have stopped at this iteration in Step (S.1). $\square$

From Lemma 3.2 and the proof of Proposition 3.3, we actually obtain that

$$\Psi_\lambda(x^k) - q_k(d^k) > 0$$

for all $k \in \mathbb{N}$. This elementary inequality will be used several times in our convergence theory without referring to it explicitly.

3.2 Global Convergence

The aim of this subsection is to present two global convergence results for Algorithm 3.1. We start our analysis with the following lemma.

Lemma 3.4 *Let $\{x^k\}$ be any sequence generated by Algorithm 3.1, and let $\{x^k\}_K$ be a subsequence converging to a point $x^* \in \mathbb{R}^n$. If x^* is not a stationary point of Ψ_λ, then*

$$\liminf_{k\to\infty, k\in K} \Delta_k > 0.$$

Proof. Define the index set

$$\bar{K} := \{k-1 \,|\, k \in K\}.$$

Then the subsequence $\{x^{k+1}\}_{k\in\bar{K}}$ converges to x^*. We have to show that

$$\liminf_{k\to\infty, k\in\bar{K}} \Delta_{k+1} > 0. \tag{4}$$

Suppose (4) does not hold. Subsequencing if necessary, we can therefore assume that

$$\lim_{k\to\infty, k\in\bar{K}} \Delta_{k+1} = 0. \tag{5}$$

In view of the updating rule for the trust-region radius (note that the lower bound $\Delta_{\min} > 0$ plays a central role here), this implies that all iterations $k \in \bar{K}$ with k sufficiently large are not successful. Hence we have

$$r_k < \rho_1 \tag{6}$$

and $x^k = x^{k+1}$ for all $k \in \bar{K}$ large enough. Since $\{x^{k+1}\}_{k\in\bar{K}}$ converges to x^* by assumption, this implies that also the subsequence $\{x^k\}_{k\in\bar{K}}$ converges to x^*. Since $\Delta_{k+1} = \sigma_1\Delta_k$ for nonsuccessful iterations, we obtain

$$\lim_{k\to\infty, k\in\bar{K}} \Delta_k = 0 \tag{7}$$

from (5).

In view of our assumptions, the limit point x^* is not a stationary point of Ψ_λ. Hence there exists a constant $\beta_1 > 0$ such that

$$\|\nabla\Psi_\lambda(x^k)\| \geq \beta_1 \tag{8}$$

for all $k \in \bar{K}$. Due to the upper semicontinuity of the generalized Jacobian (see [2, Proposition 2.6.2 (c)]), there is also a constant $\beta_2 > 0$ with

$$\|H_k^T H_k\| \leq \beta_2 \tag{9}$$

for all $k \in \bar{K}$. Using Lemma 3.2, we obtain from (7), (8) and (9) for all $k \in \bar{K}$ sufficiently large:

$$\begin{aligned}
\Psi_\lambda(x^k) - q_k(d^k) &\geq \frac{1}{2}\|\nabla\Psi_\lambda(x^k)\| \min\left\{\Delta_k, \frac{\|\nabla\Psi_\lambda(x^k)\|}{\|H_k^T H_k\|}\right\} \\
&\geq \frac{1}{2}\beta_1 \min\left\{\Delta_k, \frac{\beta_1}{\beta_2}\right\} \\
&= \frac{1}{2}\beta_1\Delta_k \\
&\geq \frac{1}{2}\beta_1\|d^k\|.
\end{aligned} \tag{10}$$

Since Ψ_λ is continuously differentiable, there exists for each $k \in \mathbb{N}$ a vector $\xi^k = x^k + \theta_k d^k, \theta_k \in (0,1)$, such that

$$\Psi_\lambda(x^k + d^k) = \Psi_\lambda(x^k) + \nabla\Psi_\lambda(\xi^k)^T d^k. \tag{11}$$

Obviously, we have $\{\xi^k\}_{k\in\bar{K}} \to x^*$. It then follows from (8)—(11) and the Cauchy-Schwarz inequality that

$$\begin{aligned}
|r_k - 1| &= \left|\frac{\Psi_\lambda(x^k) - \Psi_\lambda(x^k + d^k)}{\Psi_\lambda(x^k) - q_k(d^k)} - 1\right| \\
&= \left|\frac{q_k(d^k) - \Psi_\lambda(x^k + d^k)}{\Psi_\lambda(x^k) - q_k(d^k)}\right| \\
&= \frac{\left|\Psi_\lambda(x^k) + \nabla\Psi_\lambda(x^k)^T d^k + \frac{1}{2}(d^k)^T H_k^T H_k d^k - \Psi_\lambda(x^k) - \nabla\Psi_\lambda(\xi^k)^T d^k\right|}{\Psi_\lambda(x^k) - q_k(d^k)} \\
&\le \frac{2}{\beta_1\|d^k\|}\left|\nabla\Psi_\lambda(x^k)^T d^k - \nabla\Psi_\lambda(\xi^k)^T d^k + \frac{1}{2}(d^k)^T H_k^T H_k d^k\right| \\
&\le \frac{2}{\beta_1\|d^k\|}\left(\|\nabla\Psi_\lambda(x^k) - \nabla\Psi_\lambda(\xi^k)\|\,\|d^k\| + \frac{1}{2}\|H_k^T H_k\|\,\|d^k\|^2\right) \\
&\le \frac{1}{\beta_1}\left(2\|\nabla\Psi_\lambda(x^k) - \nabla\Psi_\lambda(\xi^k)\| + \beta_2\|d^k\|\right) \\
&\to_{\bar{K}} 0
\end{aligned}$$

for all $k \in \bar{K}$ sufficiently large. Hence the subsequence $\{r_k\}_{k\in\bar{K}}$ converges to 1, a contradiction to (6). □

The following result is an important consequence of Lemma 3.4.

Lemma 3.5 *Let $\{x^k\}$ be any sequence generated by Algorithm 3.1. Then there are infinitely many successful iterations.*

Proof. Assume the number of successful iterations is finite. Then there is an index $k_0 \in \mathbb{N}$ such that $r_k < \rho_1$ and $x^k = x^{k_0}$ for all $k \in \mathbb{N}$ with $k \ge k_0$. Hence $\{\Delta_k\} \to 0$ and $\{x^k\}$ converges to x^{k_0}. Since $\nabla\Psi_\lambda(x^{k_0}) \ne 0$ (otherwise the algorithm would have stopped), this contradicts Lemma 3.4. □

We are now able to prove our first global convergence result.

Theorem 3.6 *Let $\{x^k\}$ be any sequence generated by Algorithm 3.1. Then every accumulation point of $\{x^k\}$ is a stationary point of Ψ_λ.*

Proof. Let $x^* \in \mathbb{R}^n$ denote an accumulation point of $\{x^k\}$, and let $\{x^k\}_K$ be a subsequence converging to x^*. Since $x^{k+1} = x^k$ for all nonsuccessful iterations k and since there are infinitely many successful iterations by Lemma 3.5, we can assume without loss of generality that all iterates x^k with $k \in K$ are successful.

Suppose that $\nabla\Psi_\lambda(x^*) \neq 0$. Using this and the upper semicontinuity of the generalized Jacobian [2, Proposition 2.6.2 (c)], it follows that

$$\|\nabla\Psi_\lambda(x^k)\| \geq \beta_1 \quad \text{and} \quad \|H_k^T H_k\| \leq \beta_2$$

for all $k \in K$, where $\beta_1 > 0$ and $\beta_2 > 0$ are suitable constants. Since the iterations $k \in K$ are successful, we have $r_k \geq \rho_1$ for all $k \in K$. Lemma 3.2 therefore yields

$$\begin{aligned} \Psi_\lambda(x^k) - \Psi_\lambda(x^{k+1}) &\geq \rho_1 \left(\Psi_\lambda(x^k) - q_k(d^k)\right) \\ &\geq \frac{1}{2}\,\rho_1\,\|\nabla\Psi_\lambda(x^k)\| \min\left\{\Delta_k, \frac{\|\nabla\Psi_\lambda(x^k)\|}{\|H_k^T H_k\|}\right\} \\ &\geq \frac{1}{2}\,\rho_1\,\beta_1\,\min\left\{\Delta_k, \frac{\beta_1}{\beta_2}\right\} \end{aligned} \tag{12}$$

for all $k \in K$. Since the entire sequence $\{\Psi_\lambda(x^k)\}$ is obviously decreasing and bounded from below, it is convergent. Using (12), we thus obtain

$$\begin{aligned} \frac{1}{2}\rho_1\beta_1 \sum_{k\in K} \min\left\{\Delta_k, \frac{\beta_1}{\beta_2}\right\} &\leq \sum_{k\in K} \left(\Psi_\lambda(x^k) - \Psi_\lambda(x^{k+1})\right) \\ &\leq \sum_{k=0}^{\infty} \left(\Psi_\lambda(x^k) - \Psi_\lambda(x^{k+1})\right) \\ &< \infty. \end{aligned}$$

This implies $\{\Delta_k\}_{k\in K} \to 0$, a contradiction to Lemma 3.4. □

Although Theorem 3.6 guarantees only convergence to stationary points of Ψ_λ, we stress that Theorem 2.2 and its Corollary 2.3 provide relatively mild conditions for such a stationary point to be a solution of the nonlinear complementarity problem itself. We also note that, for example, F being a uniform P-function ensures the existence of an accumulation point due to Theorem 3.7 in [16].

A second global convergence result is given by the following theorem. Note that this result also applies to possibly unbounded sequences $\{x^k\}$.

Theorem 3.7 *Let $\{x^k\}$ be any sequence generated by Algorithm 3.1, and suppose that the sequence $\{H_k\}$ is bounded. Then*

$$\liminf_{k\to\infty} \|\nabla\Psi_\lambda(x^k)\| = 0.$$

Proof. The proof is similar to the one of Theorem 5.2 in [15], so we omit it here. □

3.3 Local Convergence

This subsection considers the local rate of convergence of Algorithm 3.1. In order to prove our main result, we need the following two lemmas. The first one is due to Moré and Sorensen [22], whereas the second one was shown by Facchinei and Soares [9] (see also [17]); however, it is restated here under a slightly weaker condition, cf. [7].

Lemma 3.8 *Assume that $x^* \in \mathbb{R}^n$ is an isolated accumulation point of a sequence $\{x^k\}$ (not necessarily generated by Algorithm 3.1), and suppose that $\{\|x^{k+1} - x^k\|\}_K \to 0$ for any subsequence $\{x^k\}_K$ converging to x^*. Then the entire sequence $\{x^k\}$ converges to x^*.*

Lemma 3.9 *Assume that $G : \mathbb{R}^n \to \mathbb{R}^n$ is semismooth and suppose that $x^* \in \mathbb{R}^n$ is a BD-regular solution of $G(x) = 0$. Let $\{x^k\}$ and $\{d^k\}$ be any two sequences (not necessarily generated by Algorithm 3.1) such that*

$$\{x^k\} \to x^* \quad \text{and} \quad \|x^k + d^k - x^*\| = o(\|x^k - x^*\|).$$

Then

$$\|G(x^k + d^k)\| = o(\|G(x^k)\|).$$

These two lemmas enable us to prove the following main local convergence result. In this result, we assume that the subspace V_k contains at least the gradient direction $\nabla\Psi_\lambda(x^k)$ and, if existing, the generalized Newton direction $d_N^k := -H_k^{-1}\Phi_\lambda(x^k)$.

Theorem 3.10 *Let $\{x^k\}$ be any sequence generated by Algorithm 3.1. If $\{x^k\}$ has an accumulation point x^* which is a BD-regular solution of $\Phi_\lambda(x) = 0$, then the following statements hold:*

- *(a) The entire sequence $\{x^k\}$ converges to x^*.*
- *(b) The generalized Newton direction $d_N^k := -H_k^{-1}\Phi_\lambda(x^k)$ exists and is the unique solution of subproblem (3) for all $k \in \mathbb{N}$ sufficiently large.*
- *(c) Eventually all iterations are successful, and we have $x^{k+1} = x^k + d_N^k$ for all $k \in \mathbb{N}$ large enough.*
- *(d) The sequence $\{x^k\}$ converges Q-superlinearly to x^*; if F is an LC^1 mapping, the rate of convergence is Q-quadratic.*

Proof. Let x^* be an accumulation point of the sequence $\{x^k\}$ which is a BD-regular solution of the system $\Phi_\lambda(x) = 0$.

(a) We prove statement (a) by verifying the assumptions of Lemma 3.8. To this end, we first note that the BD-regularity implies that x^* is an isolated solution of

NCP(F) by [26, Proposition 2.5]. Since $\{\Psi_\lambda(x^k)\}$ is monotonically decreasing, it is easy to see that x^* is therefore also an isolated accumulation point of the sequence $\{x^k\}$.

Now let $\{x^k\}_K$ be any subsequence converging to x^*. For the solution $d^k \in \mathbb{R}^n$ of the subspace trust-region subproblem (3), we have

$$\Psi_\lambda(x^k) + \nabla\Psi_\lambda(x^k)^T d^k + \frac{1}{2}(d^k)^T H_k^T H_k d^k = q_k(d^k) \le q_k(0) = \Psi_\lambda(x^k)$$

and therefore

$$\frac{1}{2}(d^k)^T H_k^T H_k d^k \le -\nabla\Psi_\lambda(x^k)^T d^k \le \|\nabla\Psi_\lambda(x^k)\|\,\|d^k\|. \tag{13}$$

The assumed BD-regularity of x^* further implies that there is a constant $\alpha > 0$ such that

$$\alpha\|d^k\|^2 \le \frac{1}{2}(d^k)^T H_k^T H_k d^k$$

for all $k \in K$ sufficiently large. In view of (13), we thus obtain

$$\|d^k\| \le \frac{1}{\alpha}\|\nabla\Psi_\lambda(x^k)\| \tag{14}$$

for all $k \in K$ large enough.

Due to our global convergence result from Theorem 3.6, we have $\{\nabla\Psi_\lambda(x^k)\}_K \to \nabla\Psi_\lambda(x^*) = 0$. Hence $\{d^k\}_K \to 0$ by (14) so that

$$\{\|x^{k+1} - x^k\|\}_K \to 0$$

since $\|x^{k+1} - x^k\| \le \|d^k\|$. Part (a) therefore follows from Lemma 3.8.

(b), (c): We verify statements (b) and (c) simultaneously. The proof is devided into three steps, where we denote by $\bar{K}$ the set of indices $k \in \mathbb{N}$ such that the $(k-1)$st iteration is successful. In step (i), we show that the generalized Newton direction $d_N^k := -H_k^{-1}\Phi_\lambda(x^k)$ exists and is the unique solution of the subproblem (3) for all $k \in \bar{K}$ sufficiently large. Part (ii) shows that, if the $(k-1)$st iteration is successful, then also the kth iteration is successful. Statements (b) and (c) are then shown by using an induction argument in Step (iii).

(i) From the BD-regularity of the solution x^* and statement (a), it follows that there is a constant $c > 0$ such that, for all $k \in \mathbb{N}$ sufficiently large, the matrix H_k is nonsingular with

$$\|H_k^{-1}\| \le c. \tag{15}$$

In particular, the generalized Newton direction $d_N^k := -H_k^{-1}\Phi_\lambda(x^k)$ exists for $k \in \mathbb{N}$ large enough. Since $\{x^k\} \to x^*$ by part (a), the continuity of Φ_λ and (15) imply that

$$\|d_N^k\| = \|H_k^{-1}\Phi_\lambda(x^k)\| \le c\|\Phi_\lambda(x^k)\| \to 0$$

so that
$$\|d_N^k\| \le \Delta_{\min}$$
for all $k \in \mathbb{N}$ sufficiently large. The definition of the set $\bar{K}$ together with the updating rule for the trust-region radius therefore gives
$$\|d_N^k\| \le \Delta_{\min} \le \Delta_k$$
for all $k \in \bar{K}$ large enough. Since $d_N^k \in V_k$ by assumption, this implies that the direction d_N^k is at least (strictly) feasible for the subproblem (3). Since the matrix $H_k^T H_k$ is positive definite due to the BD-regularity of the solution x^*, it follows for all $k \in \bar{K}$ suffiently large that d_N^k is indeed the unique solution of the subspace trust-region subproblem (3).

(ii) Assume that $k \in \bar{K}$ so that the $(k-1)$st iteration is successful. We want to show that the kth iteration is also successful, i.e., we want to show that $r_k \ge \rho_1$.

Since x^* is a solution of the nonlinear system of equations $\Phi_\lambda(x) = 0$, we obtain from (15), $\{x^k\}_{\bar{K}} \to x^*$, Proposition 2.5 and part (i) of this proof:
$$\begin{aligned}
\|x^k + d^k - x^*\| &= \|x^k + d_N^k - x^*\| \\
&= \left\| -(H_k^{-1}) \left(\Phi_\lambda(x^k) - \Phi_\lambda(x^*) - H_k(x^k - x^*) \right) \right\| \\
&\le c\|\Phi_\lambda(x^k) - \Phi_\lambda(x^*) - H_k(x^k - x^*)\| \\
&= o(\|x^k - x^*\|).
\end{aligned}$$
Since Φ_λ is semismooth by [16, Theorem 2.3], Lemma 3.9 therefore implies
$$\|\Phi_\lambda(x^k + d^k)\| = o(\|\Phi_\lambda(x^k)\|).$$
Hence
$$\Psi_\lambda(x^k + d^k) = o(\Psi_\lambda(x^k)).$$
In particular, we have
$$\Psi_\lambda(x^k + d^k) \le (1 - \rho_1)\Psi_\lambda(x^k)$$
for $k \in \bar{K}$ sufficiently large since $\rho_1 < 1$. From $q_k(d^k) = q_k(d_N^k) = 0$ we therefore get
$$\begin{aligned}
r_k &= \frac{\Psi_\lambda(x^k) - \Psi_\lambda(x^k + d^k)}{\Psi_\lambda(x^k) - q_k(d^k)} \\
&= 1 - \frac{\Psi_\lambda(x^k + d^k)}{\Psi_\lambda(x^k)} \\
&\ge \rho_1
\end{aligned}$$
for all $k \in \bar{K}$ large enough, i.e., the kth iteration is also successful.

(iii) It follows from parts (i), (ii) and an induction argument that, for all $k \in \mathbb{N}$ sufficiently large, the iteration k is successful and the trust-region subproblem (3) is solved by the generalized Newton direction d_N^k.

(d) In view of part (c), we have $x^{k+1} = x^k + d_N^k$ for all $k \in \mathbb{N}$ sufficiently large. Hence our inexact trust region method 3.1 reduces to the standard nonsmooth Newton method [27, 26] for the solution of a semismooth system of equations. The Q-superlinear/Q-quadratic rate of convergence follows therefore directly from known results by Qi [26]. □

We finally note that the assumed BD-regularity condition in Theorem 3.10 is satisfied, in particular, if x^* is an R-regular solution of NCP(F), see Proposition 2.4.

4 Numerical Results

4.1 Implementation Issues

This subsection provides a short description of our implementation of Algorithm 3.1. In particular, we give some details on the solution of the subspace trust-region subproblem (3).

To this end, let us define the two search directions

$$d_G^k := -\nabla \Psi_\lambda(x^k)$$

and

$$d_N^k := -H_k^{-1} \Phi_\lambda(x^k),$$

where $H_k \in \partial_B \Phi_\lambda(x^k)$. Note that H_k can be computed in essentially the same way as described in [4] for the Fischer-Burmeister function, see [31] for further details.

Assuming that the generalized Newton direction d_N^k exists, we define the subspace $V_k \subseteq \mathbb{R}^n$ from Step (S.2) of Algorithm 3.1 as

$$V_k := \text{span}\{d_G^k, d_N^k\}.$$

If d_N^k cannot be computed, we just set

$$V_k := \text{span}\{d_G^k\}.$$

Due to the theoretical results from the previous section, this choice of V_k guarantees the global and local fast convergence of our inexact trust region method.

Using this set V_k, it is well-known (see, e.g., [29, 21, 19]) that our subspace trust-region subproblem

$$\min \quad q_k(d) := \Psi_\lambda(x^k) + \nabla \Psi_\lambda(x^k)^T d + \frac{1}{2} d^T H_k^T H_k d \quad \text{s.t.} \quad \|d\| \leq \Delta_k, \;\; d \in V_k \qquad (16)$$

can be reformulated as a standard trust-region subproblem of dimension $r_k := \dim V_k \in \{1, 2\}$. More precisely, if $r_k = 2$ (the case $r_k = 1$ is obvious) and $\{v^{k,1}, v^{k,2}\}$

denotes an orthonormal basis of V_k, a vector $d^k \in \mathbb{R}^n$ solves problem (16) if and only if

$$d^k = \alpha_1^k v^{k,1} + \alpha_2^k v^{k,2},$$

where $\alpha^k := (\alpha_1^k, \alpha_2^k)^T$ is a solution of the 2-dimensional problem

$$\min \quad \tilde{q}_k(\alpha) := f^k + (g^k)^T \alpha + \frac{1}{2}\alpha^T M_k \alpha \quad \text{s.t.} \quad \|\alpha\| \le \Delta_k; \tag{17}$$

here, we used the notation

$$\begin{array}{rcl} f^k & := & \Psi_\lambda(x^k) \in \mathbb{R}, \\ B_k & := & (v^{k,1}, v^{k,2}) \in \mathbb{R}^{n \times 2}, \\ g^k & := & B_k^T \nabla \Psi_\lambda(x^k) \in \mathbb{R}^2, \\ M_k & := & B_k^T H_k^T H_k B_k \in \mathbb{R}^{2 \times 2}. \end{array}$$

It is therefore sufficient to solve the small-dimensional problem (17) in order to get a vector d^k as required in Step (S.2) of Algorithm 3.1.

In our MATLAB implementation of Algorithm 3.1, we modified the termination criterion in Step (S.1) and stop the iteration if

$$\Psi_2(x^k) < \varepsilon \quad \text{or} \quad \Delta_k < \Delta_{\text{stop}} \quad \text{or} \quad k > k_{\max},$$

where

$$\varepsilon := 10^{-12}, \quad \Delta_{\text{stop}} := 10^{-15}, \quad k_{\max} := 300,$$

i.e., we terminate the iteration if either an approximate solution was found or if the trust-region radius Δ_k becomes too small or if we need too many iterations. The other parameters from Algorithm 3.1 were chosen as follows:

$$\Delta_0 := 10, \;\; \rho_1 = 10^{-4}, \;\; \rho_2 = 0.75, \;\; \sigma_1 = 0.5, \;\; \sigma_2 = 2, \;\; \Delta_{\min} = 10^{-2}.$$

Apart from the stopping criterion, the implementation of our inexact trust-region method differs from its description in Subsection 3.1 in the following points:

(a) We used a nonmonotone trust-region strategy as described by Toint [30]. We do not give a precise description of this strategy here, instead we only note that the basic idea of a nonmonotone trust-region method is to accept more iterations as being successful. Compared to the standard (monotone) trust-region method, this nonmonotone variant needs less many iterations, and the improvement is sometimes considerable, see [31] for a numerical comparison.

(b) We incorporated a simple heuristic backtracking strategy in order to avoid possible domain violations, see [8] for a more detailed description.

(c) Similar to [16], we used a dynamic choice of the parameter λ. The main idea is to start the iteration with $\lambda = 2$ (which seems to give the best global behaviour) and to reduce $\lambda \to 0$ if we get close to a solution (this seems to improve the local behaviour to some extent). We refer to [31] for further details.

4.2 Computational Results

We tested the previously described implementation of Algorithm 3.1 on a SUN SPARC 20 station. We used all complementarity problems and all available starting points from the MCPLIB test problem collection by Dirkse and Ferris [6]. Our results are summarized in Table 1, where we present the following data:

problem:	name of test example in MCPLIB
n:	dimension of test example
SP:	number of starting point in the M-file `cpstart.m`
k:	number of iterations
k_{suc}:	number of successful iterations
N:	number of Newton steps
F-ev.:	number of function evaluations of F
F'-ev.:	number of Jacobian evaluations of F
$\Psi_2(x^f)$:	value of $\Psi_2(x)$ at the final iterate $x = x^f$
$\|\nabla\Psi_2(x^f)\|$:	value of $\|\nabla\Psi_2(x)\|$ at the final iterate $x = x^f$
B:	number of iterations using a backtracking step.

Table 1: Numerical results for Algorithm 3.1

problem	n	SP	k	k_{suc}	N	F-ev.	F'-ev.	$\Psi_2(x^f)$	$\|\nabla\Psi_2(x^f)\|$	B
bertsekas	15	1	34	20	7	35	21	5.7e-25	9.6e-11	0
bertsekas	15	2	32	20	6	33	21	1.3e-16	9.2e-7	0
bertsekas	15	3	20	18	12	21	19	1.1e-28	9.1e-13	0
billups	1	1	—	—	—	—	—	—	—	—
colvdual	20	1	—	—	—	—	—	—	—	—
colvdual	20	2	105	76	40	106	77	4.8e-14	7.9e-5	0
colvnlp	15	1	30	19	6	31	20	8.5e-18	4.3e-7	0
colvnlp	15	2	14	11	4	15	12	2.8e-18	2.4e-7	0
cycle	1	1	7	5	6	8	6	6.2e-22	3.5e-11	0
explcp	16	1	25	19	21	26	20	2.4e-18	2.2e-9	0
hanskoop	14	1	10	10	9	12	11	3.1e-20	2.9e-9	1
hanskoop	14	2	13	13	12	14	14	1.3e-21	6.1e-10	0
hanskoop	14	3	15	11	10	16	12	3.1e-20	2.8e-9	0
hanskoop	14	4	11	8	8	13	9	7.2e-13	8.2e-6	1
hanskoop	14	5	11	11	10	13	12	7.4e-23	9.1e-11	1
josephy	4	1	14	8	11	15	9	7.6e-27	1.1e-12	0
josephy	4	2	10	6	7	11	7	6.5e-17	1.0e-7	0
josephy	4	3	54	36	48	55	37	1.6e-15	5.1e-7	0
josephy	4	4	5	5	5	6	6	8.2e-17	1.2e-7	0
josephy	4	5	4	4	4	5	5	1.5e-16	1.6e-7	0
josephy	4	6	14	7	8	15	8	6.2e-23	1.0e-10	0

Table 1: Numerical results for Algorithm 3.1 (continued)

problem	n	SP	k	k_{suc}	N	F-ev.	F'-ev.	$\Psi_2(x^f)$	$\|\nabla\Psi_2(x^f)\|$	B
kojshin	4	1	16	9	11	17	10	1.7e-19	7.2e-9	0
kojshin	4	2	11	7	7	12	8	1.1e-15	5.8e-7	0
kojshin	4	3	16	16	12	17	17	1.5e-20	1.6e-9	0
kojshin	4	4	4	4	4	5	5	2.1e-13	6.8e-6	0
kojshin	4	5	11	5	10	12	6	7.7e-27	1.1e-12	0
kojshin	4	6	13	8	10	14	9	2.0e-24	2.5e-11	0
mathinum	3	1	4	4	4	5	5	5.1e-16	6.4e-8	0
mathinum	3	2	5	5	5	6	6	2.2e-14	5.8e-7	0
mathinum	3	3	12	7	6	13	8	2.9e-13	1.5e-6	0
mathinum	3	4	7	7	7	8	8	6.3e-19	3.1e-9	0
mathisum	4	1	9	5	7	10	6	6.9e-13	2.8e-6	0
mathisum	4	2	6	6	6	7	7	1.6e-13	1.3e-6	0
mathisum	4	3	13	8	12	14	9	1.3e-23	8.8e-12	0
mathisum	4	4	6	6	6	7	7	2.6e-23	1.7e-11	0
nash	10	1	8	8	8	9	9	6.7e-24	2.7e-10	0
nash	10	2	21	11	10	22	12	5.9e-19	8.2e-8	0
pgvon105	105	1	53	45	47	103	46	5.1e-13	4.3e-2	49
pgvon106	106	1	30	30	27	58	31	8.3e-13	2.3e+05	27
powell	16	1	10	7	9	11	8	6.0e-20	2.5e-9	0
powell	16	2	9	9	8	10	10	3.5e-15	7.3e-7	0
powell	16	3	13	13	12	14	14	1.1e-22	1.2e-10	0
powell	16	4	10	10	9	11	11	8.1e-15	1.0e-6	0
scarfanum	13	1	11	8	8	12	9	1.6e-19	9.2e-9	0
scarfanum	13	2	19	12	7	20	13	2.6e-18	3.7e-8	0
scarfanum	13	3	13	9	11	15	10	3.3e-20	4.1e-9	1
scarfasum	14	1	12	5	11	13	6	1.5e-13	3.2e-5	0
scarfasum	14	2	15	9	8	17	10	7.5e-19	7.2e-8	1
scarfasum	14	3	21	14	6	22	15	4.9e-18	1.8e-7	0
scarfbnum	39	1	29	25	21	30	26	7.2e-28	3.9e-12	0
scarfbnum	39	2	28	26	25	29	27	1.5e-21	4.8e-9	0
scarfbsum	40	1	27	20	16	28	21	2.7e-27	1.4e-11	0
scarfbsum	40	2	30	22	21	31	23	5.7e-27	1.6e-11	0
sppe	27	1	9	9	6	10	10	3.4e-16	3.8e-8	0
sppe	27	2	8	8	6	9	9	4.7e-29	1.7e-14	0
tobin	42	1	10	10	7	11	11	3.7e-13	1.3e-6	0
tobin	42	2	11	11	8	12	12	1.0e-13	6.1e-7	0

The results in Table 1 are quite promising. The number of iterations seems comparable to the full-dimensional trust-region method by Jiang et al. [15] (note that Jiang et al. [15] only count the number of successful iterations in their paper). Also the robustness of the two algorithms seems to be very similar, whereas our method is more efficient for obvious reasons.

On the other hand, if we compare our algorithm with the corresponding line search method from [16], we usually need a few more iterations. However, the number of function evaluations is often considerably lower than for the method in [16]. Moreover, our trust-region method seems to be more reliable than the line search method from [16].

In fact, our method has only two failures: One for the example by Billups [1], which is impossible to solve for almost all state-of-the-art solvers since the starting point is very close to a local-nonglobal minimum of basically any reasonable merit function, and one for problem `colvdual` (first starting point) which is also known to be a very difficult test problem.

5 Final Remarks

In this paper, we investigated the theoretical and numerical properties of a nonsmooth Newton-type method for the solution of nonlinear complementarity problems. In contrast to many earlier works [9, 4, 16], we used a trust-region-type globalization instead of a line search globalization. Our nonsmooth trust-region method is very similar to the one of Jiang et al. [15]. However, in order to reduce the computational overhead, we did not solve the trust-region subproblem exactly like in [15]. Instead, we followed an idea from Shultz et al. [29] and solved the trust-region problem only on a subspace. Alternatively, we could have used the dogleg step by Powell [24] or the double dogleg step by Dennis and Mei [5]; however, the subspace trust-region method is more general.

From a theoretical point of view, our inexact trust-region method has the same convergence properties as the (exact) trust-region method by Jiang et al. [15]. The convergence properties are also similar to those of the corresponding line search method by Kanzow and Kleinmichel [16]. From a numerical point of view, the new method should be considerably more efficient than the algorithm from [15]. Moreover, it seems to be more reliable and almost as efficient as the method from [16].

We therefore believe that a sophisticated implementation of a suitable trust-region method could actually outperform corresponding line search methods. At least we think that future studies in the numerical solution of complementarity problems should put more attention on trust-region-type methods than this was done before.

References

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

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

[3] R.W. Cottle, J.-S. Pang and R.E. Stone: *The Linear Complementarity Problem.* Academic Press, Boston, 1992.

[4] 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.

[5] J.E. Dennis and H.H.W. Mei: *Two new unconstrained optimization algorithms which use function and gradient values.* Journal of Optimization Theory and Applications 28, 1979, pp. 453–482.

[6] S.P. Dirkse and M.C. Ferris: *MCPLIB: A collection of nonlinear mixed complementarity problems.* Optimization Methods and Software 5, 1995, pp. 123–156.

[7] F. Facchinei, A. Fischer and C. Kanzow: *Inexact Newton methods for semismooth equations with applications to variational inequality problems.* In: G. Di Pillo and F. Giannessi (eds.): *Nonlinear Optimization and Applications.* Plenum Press, New York, NY, 1996, pp. 125–139.

[8] F. Facchinei, A. Fischer and C. Kanzow: *A semismooth Newton method for variational inequalities: The case of box constraints.* In: M.C. Ferris and J.-S. Pang (eds.): *Complementarity and Variational Problems. State of the Art.* SIAM, Philadelphia, PA, 1997, pp. 76–90.

[9] 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.

[10] M.C. Ferris and J.-S. Pang: *Engineering and economic applications of complementarity problems.* SIAM Review, to appear.

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

[12] R. Fletcher: *Practical Methods of Optimization.* John Wiley & Sons, Chichester, England, 1987.

[13] A. Friedlander, J.M. Martínez and A. Santos: *A new algorithm for bound constrained minimization.* Journal of Applied Mathematics and Optimization 30, 1994, pp. 235–266.

[14] S.A. Gabriel and J.-S. Pang: *A trust region method for constrained nonsmooth equations.* In: W.W. Hager, D.W. Hearn and P.M. Pardalos (eds.): *Large Scale Optimization. State of the Art.* Kluwer Academic Publishers, Dordrecht, 1994, pp. 155–181.

[15] H. Jiang, M. Fukushima, L. Qi and D. Sun: *A trust region method for solving generalized complementarity problems.* SIAM Journal on Optimization, to appear.

[16] 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, January 1997.

[17] C. Kanzow and H.-D. Qi: *A QP-free constrained Newton-type method for variational inequality problems.* Preprint 121, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, March 1997.

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

[19] O. Knoth: *Marquardt-ähnliche Verfahren zur Minimierung nichtlinearer Funktionen.* Ph.D. Thesis, Martin-Luther-University, Halle-Wittenberg, Germany, 1983 (in German).

[20] J.M. Martínez and A. Santos: *A trust region strategy for minimization on arbitrary domains.* Mathematical Programming 68, 1995, pp. 267–302.

[21] J.J. Moré: *Recent developments in algorithms and software for trust region methods.* In: A. Bachem, M. Grötschel and B. Korte (eds.): *Mathematical Programming. The State of the Art.* Springer-Verlag, Berlin, 1983, pp. 258–287.

[22] J.J. Moré and D.C. Sorensen: *Computing a trust-region step.* SIAM Journal on Scientific and Statistical Computing 4, 1983, pp. 553–572.

[23] J.-S. Pang: *Newton's method for B-differentiable equations.* Mathematics of Operations Research 15, 1990, pp. 311–341.

[24] M.J.D. Powell: *A new algorithm for unconstrained optimization.* In: J.B. Rosen, O.L. Mangasarian and K. Ritter (eds.): *Nonlinear Programming.* Academic Press, New York, NY, 1970, pp. 31–65.

[25] M.J.D. Powell: *Convergence properties of a class of minimization algorithms.* In: O.L. Mangasarian, R.R. Meyer and S.M. Robinson (eds.): *Nonlinear Programming 2.* Academic Press, New York, NY, 1975, pp. 1–27.

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

[27] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.

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

[29] G.A. Shultz, R.B. Schnabel and R.H. Byrd: *A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties.* SIAM Journal on Numerical Analysis 22, 1985, pp. 47–67.

[30] Ph.L. Toint: *Non-monotone trust-region algorithms for nonlinear optimization subject to convex constraints.* Mathematical Programming 77, 1997, pp. 69–94.

[31] M. Zupke: *Trust-Region-Verfahren zur Lösung nichtlinearer Komplementaritätsprobleme.* Diploma Thesis, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, July 1997 (in German).