Hamburger Beiträge
zur Angewandten Mathematik

Strictly Feasible Equation-Based Methods for Mixed Complementarity Problems

Christian Kanzow

Reihe A
Preprint 145
August 1999

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

STRICTLY FEASIBLE EQUATION-BASED METHODS FOR MIXED COMPLEMENTARITY PROBLEMS

Christian Kanzow[1]

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

August 19, 1999

Abstract. We introduce a new algorithm for the solution of the mixed complementarity problem (MCP) which has stronger properties than most existing methods. In fact, typical solution methods for the MCP either generate feasible iterates but have to solve relatively complicated subproblems (like quadratic programs or linear complementarity problems), or they have relatively simple subproblems (like linear systems of equations) but generate not necessarily feasible iterates. The method to be presented here combines the nice features of these two classes of methods: It has to solve only one linear system of equations (of reduced dimension) at each iteration, and it generates feasible (more precisely: strictly feasible) iterates. The new method has some nice global and local convergence properties. Some preliminary numerical results will also be given.

Key Words. Complementarity problems, Newton's method, active set strategy, global convergence, quadratic convergence.

[1]The research of the author was partially supported by the DFG (Deutsche Forschungsgemeinschaft).

1 Introduction

Let $l_i \in \mathbb{R} \cup \{-\infty\}$ and $u_i \in \mathbb{R} \cup \{+\infty\}$ be given lower and upper bounds with $l_i < u_i$ for $i = 1, \dots, n$, define

$$l := (l_1, \dots, l_n)^T \quad \text{and} \quad u := (u_1, \dots, u_n)^T,$$

and suppose that $F : \mathcal{O} \to \mathbb{R}^n$ is a continuously differentiable mapping defined on an open set $\mathcal{O} \subseteq \mathbb{R}^n$ containing the rectangle $[l, u]$. Then the *mixed complementarity problem* (MCP, for short) is to find a vector $x^* \in [l, u]$ such that each component x_i^* satisfies exactly one of the following implications:

$$\begin{array}{rcl} x_i^* = l_i & \Longrightarrow & F_i(x^*) \geq 0, \\ x_i^* \in (l_i, u_i) & \Longrightarrow & F_i(x^*) = 0, \\ x_i^* = u_i & \Longrightarrow & F_i(x^*) \leq 0. \end{array} \tag{1}$$

If $l_i = -\infty$ and $u_i = +\infty$ for all $i = 1, \dots, n$, it is easy to see that the MCP (1) is equivalent to finding a solution of the nonlinear system of equations

$$F(x) = 0,$$

whereas the MCP (1) reduces to the standard nonlinear complementarity problem of finding a feasible vector for the system of equations and inequalities

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

in case $l_i = 0$ and $u_i = +\infty$ for all $i = 1, \dots, n$. For many engineering and economic applications of the mixed and nonlinear complementarity problems, the interested reader is referred to [15].

In this paper, we develop a new algorithm for the solution of the MCP (1) which has a strong global and local convergence theory and which can be applied to an MCP (1) with a general nonlinear (not necessarily monotone) mapping F. Typically, solution methods having these properties fall into one of the following two classes: Either they generate feasible iterates with respect to the box constraints $[l, u]$ but have to solve relatively complicated subproblems (like quadratic programs [25, 3, 20] or linear complementarity problems [9, 31]), or they have relatively simple subproblems (like linear systems of equations [12, 7, 4]) but generate not necessarily feasible iterates.

Having simple subproblems is obviously a desirable feature of any algorithm. On the other hand, there exist many applications of MCPs where the mapping F is not defined outside the feasible region $[l, u]$. Hence it would be extremely nice to have an algorithm that, on the one hand, generates only feasible iterates and, on the other hand, has to solve only simple subproblems. The method to be presented here has these two properties. In fact, as far as we know, there are currently only two related methods available: One was recently proposed by Ulbrich [33], and the other one was presented by Sun [32] on a recent conference.

The method by Ulbrich uses a trust-region globalization and, therefore, has to solve more complicated subproblems in its current version, although it should be possible to adapt his idea within a line search framework with linear systems as subproblems. However, his method will still be different from ours since we use a completely different globalization strategy. Moreover, even within a line search framework, he would have to solve full dimensional linear systems, whereas we have to solve linear systems of reduced dimension only. This is also one of the advantages of our approach when compared to the method proposed by Sun [32].

The algorithm we introduce here is mainly motivated by the recent research [14], but still differs from the work [14] in a couple of respects. In particular, it uses a somewhat different globalization strategy that allows us to stay strictly feasible rather than just feasible. The strict feasibility of the algorithm seems to be highly important for mixed complementarity problems since quite often the mapping F is not even defined on the boundary of the feasible set $[l, u]$. For example, if a component function of F contains a logarithmic term like $\ln(x_i)$, modellers typically introduce a lower bound of $l_i = 0$ for the ith component of the vector x although F is not really defined for $x_i = 0$.

Our algorithm takes into account situations like this and needs F to be well-defined only in the interior (l, u) of the feasible set $[l, u]$ as well as in a neighbourhood of any accumulation point (possibly on the boundary of $[l, u]$) generated by the algorithm. This is not at all a restrictive requirement and will be satisfied for practically any mixed complementarity problem. An alternative would be to place a lower bound of $l_i = \varepsilon, \varepsilon > 0$ a small number, instead of $l_i = 0$ in case of a logarithmic term $\ln(x_i)$, for example. However, this alternative seems to be less elegant and may in some situations even exclude the real solution of the MCP (1).

Our algorithm may be viewed as a Newton-type method applied to a reformulation of the MCP (1) as a constrained nonlinear system of equations. This reformulation is introduced in Section 2, where we also state some preliminary and mostly known results related to this reformulation. The algorithm itself is motivated and stated in detail in Section 3. The global and local convergence properties of the algorithm are investigated in Section 4, whereas numerical results are given in Section 5. We then conclude this paper with some final remarks in Section 6.

Some words about our notation. The n-dimensional real space is denoted by $\mathbb{R}^n$, with $\|x\|$ being the Euclidean norm of a vector $x \in \mathbb{R}^n$. For index sets $I, J \subseteq \{1, \ldots, n\}$ and a matrix $M \in \mathbb{R}^{n \times n}, M = (m_{ij})$, we denote by M_{IJ} the submatrix consisting of the elements m_{ij} for $i \in I, j \in J$. A similar notation is used for subvectors. We further write x_+ or $[x]_+$ for the projection of a vector $x \in \mathbb{R}^n$ on the rectangle $[l, u]$. Finally, if $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function, we view its gradient $\nabla f(x)$ at a point x as a column vector throughout this manuscript.

2 Preliminaries

In this section we first give the precise definition of an operator Φ which is used in order to reformulate the MCP (1) as a nonlinear system of equations. We then state a couple of properties of this operator as well as of its corresponding merit function Ψ which will be useful for our convergence analysis in Section 4. Most of these properties are known and taken from [2] and [13]; hence we can skip most of the proofs. In fact, we include only a proof here if there is no explicit reference available for the corresponding result. Many of the results given in this section are generalizations of standard results for the nonlinear complementarity problem, see, in particular, [12, 7, 18].

Let us introduce the function $\phi : \mathbb{R}^2 \to \mathbb{R}$ by

$$\phi(a,b) := \sqrt{a^2+b^2} - a - b \tag{2}$$

(see [16]) and define $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ componentwise in the following way:

$$\Phi_i(x) := \begin{cases} \phi(x_i - l_i, F_i(x)) & \text{if } i \in I_l, \\ -\phi(u_i - x_i, -F_i(x)) & \text{if } i \in I_u, \\ \phi(x_i - l_i, \phi(u_i - x_i, -F_i(x)) & \text{if } i \in I_{lu}, \\ -F_i(x) & \text{if } i \in I_f; \end{cases}$$

here, the index sets I_l, I_u, I_{lu} and I_f form a partition of the set $\{1, \dots, n\}$ and are defined in the following way:

$$\begin{aligned} I_l &:= \{i \,|\, -\infty < l_i < u_i = +\infty\}, \\ I_u &:= \{i \,|\, -\infty = l_i < u_i < +\infty\}, \\ I_{lu} &:= \{i \,|\, -\infty < l_i < u_i < +\infty\}, \\ I_f &:= \{i \,|\, -\infty = l_i < u_i = +\infty\}, \end{aligned}$$

i.e., I_l denotes the set of indices which have finite lower bounds only, I_u is the set of indices which have finite upper bounds only, I_{lu} contains all indices with both finite lower and finite upper bounds, and I_f is the set of free variables for which there are no finite bounds. Note that, in the definition of Φ, it would also be possible to use some other ϕ-functions (see, e.g., [4, 21, 29]), but for the sake of simplicity and because of some technical details when using different functions, we present our theoretical results only for the example given in (2).

Since the function ϕ has the property that

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

the following result is easy to verify (see [2, Proposition 3.2.7]).

Proposition 2.1 *A vector $x^* \in \mathbb{R}^n$ solves the MCP (1) if and only if x^* is a solution of the nonlinear system of equations $\Phi(x) = 0$.*

Due to the nonsmoothness of ϕ in the origin, the operator Φ is usually nonsmooth as well. However, it turns out to be semismooth or even strongly semismooth, see [23, 28, 27] for the precise definition and some properties of (strongly) semismooth functions.

Proposition 2.2 *The following statements hold:*

(a) If F is continuously differentiable on $\mathcal{O}$, then Φ is semismooth on $\mathcal{O}$.

(b) If F is differentiable with F' being locally Lipschitzian on $\mathcal{O}$, then Φ is strongly semismooth on $\mathcal{O}$.

Proof. The function ϕ is known to be strongly semismooth by [17, Lemma 20]. Furthermore, a continuously differentiable function is also semismooth, and even strongly semismooth if its derivative is locally Lipschitzian [28]. Since the composition of (strongly) semismooth functions is again (strongly) semismooth by [17, Theorem 19], both statements follow immediately. □

As a consequence of Proposition 2.2 and known properties of (strongly) semismooth functions [28, 27], we obtain the following result.

Proposition 2.3 *The following statements hold:*

(a) If F is continuously differentiable on $\mathcal{O}$, then

$$\Phi(x+d) - \Phi(x) - Hd = o(\|d\|)$$

for any $d \to 0$ and any $H \in \partial\Phi(x+d)$, where $x \in \mathcal{O}$ is fixed.

(b) If F is differentiable with F' being locally Lipschitzian on $\mathcal{O}$, then

$$\Phi(x+d) - \Phi(x) - Hd = O(\|d\|^2)$$

for any $d \to 0$ and any $H \in \partial\Phi(x+d)$, where $x \in \mathcal{O}$ is fixed.

We next want to restate a result saying that all elements in the generalized Jacobian $\partial\Phi(x^*)$ of Φ at a solution x^* of the MCP are nonsingular under certain conditions (see [5] for the definition of the generalized Jacobian). To this end, we first introduce a regularity concept which is based on the following notation: Associated with a solution x^* of the MCP (1), let us define the index sets

$$\begin{aligned}
\alpha &:= \{i \,|\, x_i^* \in (l_i, u_i) \text{ and } F_i(x^*) = 0\}, \\
\beta &:= \{i \,|\, x_i^* \in \{l_i, u_i\} \text{ and } F_i(x^*) = 0\}, \\
\gamma &:= \{i \,|\, (x_i^* = l_i \text{ and } F_i(x^*) > 0) \text{ or } (x_i^* = u_i \text{ and } F_i(x^*) < 0)\}.
\end{aligned}$$

Then we can give the following definition.

Definition 2.4 *A solution x^* of the MCP (1) is called* strongly regular *if*

(i) the submatrix $F'(x^*)_{\alpha\alpha}$ is nonsingular, and

(ii) the Schur complement

$$F'(x^*)_{\alpha\cup\beta,\alpha\cup\beta}/F'(x^*)_{\alpha\alpha} := F'(x^*)_{\beta\beta} - F'(x^*)_{\beta\alpha}F'(x^*)_{\alpha\alpha}^{-1}F'(x^*)_{\alpha\beta}$$

is a P-matrix, i.e., has positive principal minors.

Note that the previous definition of strong regularity is not really the original definition given in [30], but can be shown to be equivalent to the original one [10, Theorem 3.4]. For some other characterizations, we refer the reader to [22]. Furthermore, an extensive treatment of P-matrices is given in [6].

Using the strong regularity assumption, it is possible to establish the following nonsingularity result, see [13, Theorem 2.7].

Proposition 2.5 *Let x^* be a strongly regular solution of the MCP (1). Then all elements $H \in \partial\Phi(x^*)$ are nonsingular.*

The previous result can be used in order to establish a local error bound.

Proposition 2.6 *Let x^* be a strongly regular solution of the MCP (1). Then there exists a constant $\kappa > 0$ such that*

$$\|\Phi(x)\| \geq \kappa \|x - x^*\|$$

for all $x \in \mathbb{R}^n$ in a sufficiently small neighbourhood of x^*.

Proof. In view of Proposition 2.5, all elements in the generalized Jacobian $\partial\Phi(x^*)$ are nonsingular. On the other hand, Φ is a semismooth function by Proposition 2.2. Hence the assertion follows from [26, Proposition 3]. □

Next we consider the merit function $\Psi : \mathbb{R}^n \to \mathbb{R}$ associated to the operator Φ and defined by

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

Despite the nonsmoothness of Φ, it turns out that the merit function Ψ is continuously differentiable everywhere. More precisely, we have the following result from [2, Theorem 3.2.8].

Proposition 2.7 *The function Ψ is continuously differentiable at any point $x \in \mathbb{R}^n$ with $\nabla\Psi(x) = H^T\Phi(x)$ for an arbitrary $H \in \partial\Phi(x)$.*

Our algorithm to be presented in the next section tries to find a solution of the constrained nonlinear system

$$\Phi(x) = 0, \quad x \in [l, u],$$

or, equivalently, of the constrained optimization problem

$$\min \Psi(x) \quad \text{subject to} \quad x \in [l, u]. \tag{3}$$

In general, our method will only be able to compute a stationary point of this constrained optimization problem. In our next result, we therefore introduce a condition which says that such a stationary point is already a global minimum and, in fact, a solution of the MCP (1) under relatively mild assumptions. The statement of this result makes use of the index sets

$$\begin{aligned} I_f &:= \{i \,|\, l_i = -\infty \text{ and } u_i = +\infty\}, \\ I_{\bar{f}} &:= \{1, \ldots, n\} \setminus I_f; \end{aligned}$$

note that I_f was already defined in the beginning of this section and contains all the free indices for which there are no finite bounds, whereas its complement $I_{\bar{f}}$ includes all the indices where at least one of the two bounds is finite, i.e., $I_{\bar{f}} = I_l \cup I_u \cup I_{lu}$ in our previous notation.

Using the above two index sets, we can restate the following result from [13, Theorem 3.5].

Proposition 2.8 *Let x^* be a stationary point of the constrained reformulation (3) of the MCP (1) such that*

(a) the submatrix $F'(x^*)_{I_f I_f}$ is nonsingular, and

(b) the Schur complement

$$F'(x^*)/F'(x^*)_{I_f I_f} := F'(x^*)_{I_{\bar{f}} I_{\bar{f}}} - F'(x^*)_{I_{\bar{f}} I_f} F'(x^*)^{-1}_{I_f I_f} F'(x^*)_{I_f I_{\bar{f}}}$$

is a P_0-matrix, i.e., has nonnegative principal minors.

Then x^* is a solution of the MCP (1).

Note that exactly the same conditions guarantee that an unconstrained stationary point of the merit function Ψ is a solution of the MCP (1), see [13, Theorem 3.3]. For a discussion of P_0-matrices in general, the reader is once again referred to [6].

3 Algorithm

In this section, we give a detailed description of an algorithm for the solution of the mixed complementarity problem (1). The main idea is to solve the MCP (1) by using a reformulation of (1) as the constrained nonlinear system of equations

$$\Phi(x) = 0, \quad x \in [l, u].$$

In order to solve this constrained system, we apply a Newton-type method to $\Phi(x) = 0$ and take care of the constraints by an active set strategy. The method is globalized by a projected gradient-type method for the corresponding box constrained optimization problem

$$\min \ \Psi(x) \quad \text{subject to} \quad x \in [l, u].$$

The precise statement is as follows.

Algorithm 3.1 *(Strictly Feasible Newton-type Method)*

(S.0) (Initialization)
Choose $x^0 \in (l, u), \beta \in (0, 1), \sigma \in (0, 1), \tau \in (0, 1), \gamma > 0, \omega \in (0, 1), \rho > 0, p_1 > 1, p_2 > 0, \delta > 0, c > 0$, and set $k := 0$.

(S.1) (Termination Criterion)
If x^k is a stationary point of (3): STOP.

(S.2) (Active Set Strategy)
Define

$$\delta_k := \min\{\delta, c\sqrt{\|\Phi(x^k)\|}\}$$

and

$$\begin{aligned} \mathcal{A}_k &:= \{i \,|\, x_i^k - l_i \le \delta_k \text{ or } u_i - x_i^k \le \delta_k\}, \\ \mathcal{I}_k &:= \{1, \ldots, n\} \setminus \mathcal{A}_k. \end{aligned}$$

(S.3) (Compute Newton-type Search Direction)
Select an element $H^k \in \partial\Phi(x^k)$ and compute a vector $d^k \in \mathbb{R}^n$ in the following way: For $i \in \mathcal{A}_k$, set

$$d_i^k := \begin{cases} l_i - x_i^k & \text{if } x_i^k - l_i \leq \delta_k, \\ u_i - x_i^k & \text{if } u_i - x_i^k \leq \delta_k, \end{cases}$$

then solve the linear system

$$H^k_{\mathcal{I}_k\mathcal{I}_k} d_{\mathcal{I}_k} = -\Phi(x^k)_{\mathcal{I}_k} - H^k_{\mathcal{I}_k\mathcal{A}_k} d^k_{\mathcal{A}_k} \tag{4}$$

in order to get the components d_i^k for $i \in \mathcal{I}_k$. If the linear system (4) is not solvable, then go to (S.6), else go to (S.4).

(S.4) (Acceptance Criterion for Newton-type Search Direction)
Set

$$\tau_k := \max\{\tau, 1 - \|\Phi(x^k)\|\}.$$

If $x^k + \tau_k d^k \in (l, u)$ and

$$\|\Phi(x^k + \tau_k d^k)\| \leq \omega \|\Phi(x^k)\|, \tag{5}$$

then set $x^{k+1} := x^k + \tau_k d^k$ and go to (S.7), else go to (S.5).

(S.5) (Projected Newton-type Search Direction)
Set $\bar{x}_N^k := [x^k + d^k]_+$ and $s_N^k := \bar{x}_N^k - x^k$. If

$$\nabla\Psi(x^k)^T s_N^k \leq -\rho \|s_N^k\|^{p_1} \quad \text{and} \quad \nabla\Psi(x^k)^T s_N^k \leq -\rho \|\Phi(x^k)\|^{p_2}, \tag{6}$$

then compute $t_k := \max\{\tau_k \beta^\ell \mid \ell = 0, 1, 2, \ldots\}$ such that

$$\Psi(x^k + t_k s_N^k) \leq \Psi(x^k) + t_k \sigma \nabla\Psi(x^k)^T s_N^k,$$

set $x^{k+1} := x^k + t_k s_N^k$ and go to (S.7); else go to (S.6)

(S.6) (Projected Gradient-type Search Direction)
Set $\bar{x}_G^k := [x^k - \gamma\nabla\Psi(x^k)]_+$ and $s_G^k := \bar{x}_G^k - x^k$. Compute $t_k := \max\{\tau_k \beta^\ell \mid \ell = 0, 1, 2, \ldots\}$ such that

$$\Psi(x^k + t_k s_G^k) \leq \Psi(x^k) + \sigma t_k \nabla\Psi(x^k)^T s_G^k,$$

set $x^{k+1} := x^k + t_k s_G^k$ and go to (S.7).

(S.7) (Update)
Set $k \leftarrow k + 1$ and go to (S.1).

Throughout this paper, we assume that the constant $\delta > 0$ is chosen sufficiently small such that

$$\delta < \frac{1}{2} \min_{i=1,\ldots,n} |u_i - l_i|.$$

This, in particular, implies that we cannot have $x_i^k - l_i \le \delta_k$ and $u_i - x_i^k \le \delta_k$ for the same index $i \in \mathcal{A}_k$ since this would imply

$$|u_i - l_i| = u_i - l_i = (u_i - x_i^k) + (x_i^k - l_i) \le 2\delta_k \le 2\delta,$$

a contradiction to the choice of δ. Hence, there is no ambiguity in the definition of the components d_i^k for $i \in \mathcal{A}_k$ in Step (S.3) of Algorithm 3.1.

We now give some further comments on each step of Algorithm 3.1 in order to get a better understanding of the method.

The active set strategy in Step (S.2) is motivated by the investigation from [11] for constrained optimization problems. The set $\mathcal{A}_k$ is viewed as a suitable approximation to the correct active set at the current point, whereas its complement $\mathcal{I}_k$ is viewed as a substitute of the inactive indices.

In Step (S.3) of Algorithm 3.1, we try to compute a Newton-type search direction d^k. For the components d_i^k with $i \in \mathcal{A}_k$, we use a simple formula whose aim is, basically, to bring the corresponding components of our iterates closer to the boundary. In order to understand the formula for the computation of the components d_i^k for $i \in \mathcal{I}_k$, note that, after a possible permutation of the rows and columns, we can rewrite the standard (unconstrained) Newton equation $H^k d = -\Phi(x^k)$ as

$$\begin{pmatrix} H^k_{\mathcal{I}_k\mathcal{I}_k} & H^k_{\mathcal{I}_k\mathcal{A}_k} \\ H^k_{\mathcal{A}_k\mathcal{I}_k} & H^k_{\mathcal{A}_k\mathcal{A}_k} \end{pmatrix} \begin{pmatrix} d_{\mathcal{I}_k} \\ d_{\mathcal{A}_k} \end{pmatrix} = - \begin{pmatrix} \Phi(x^k)_{\mathcal{I}_k} \\ \Phi(x^k)_{\mathcal{A}_k} \end{pmatrix}. \tag{7}$$

Now, recalling that we already have an expression for $d^k_{\mathcal{A}_k}$, it is easy to see that the linear system (4) is exactly the first block row in (7).

In Step (S.4), we then test whether we can take an almost full step along the Newton-type search direction d^k computed in Step (S.3); more precisely, if a large step in this direction stays strictly feasible and gives a sufficient decrease for $\|\Phi\|$, we accept this step and turn to the next iteration. We will see in Section 4 that the conditions in Step (S.4) of Algorithm 3.1 will eventually be satisfied around a strongly regular solution of the MCP (1), and this will guarantee the local fast convergence of our method.

If the Newton-type search direction d^k is not acceptable in Step (S.4) either because it brings us out of the strictly feasible region or because it does not provide a sufficient decrease for $\|\Phi\|$, we go to Step (S.5), calculate the projection of a full Newton-type step $x^k + d^k$ onto the feasible set $[l, u]$, and use the line segment between x^k and this projected point as a new search direction. We call this a projected Newton-type search direction and denote it by s_N^k.

In general, it cannot be guaranteed that s_N^k is a good descent direction for our merit function Ψ. Hence, if we do not want to destroy the global convergence properties of Algorithm 3.1, we have to check whether s_N^k is a suitable descent direction for Ψ, and we do this in (6). In case the two relatively mild criteria in (6) are satisfied, we accept s_N^k as our search direction and perform a simple line search by using the standard Armijo rule.

However, if s_N^k is not a good descent direction, we switch to Step (S.6) and perform a similar line search along a standard projected gradient-type search direction, see, e.g., [1]. This procedure will guarantee the global convergence of our method. A detailed convergence analysis follows in the next section.

Remark 3.2 *Given any feasible vector $x^k \in \mathbb{R}^n$ (not necessarily generated by Algorithm 3.1), we can still define corresponding quantities like $\mathcal{I}_k, \mathcal{A}_k, d^k$ etc. in exactly the same way as in Algorithm 3.1. This is precisely the point of view we will take in our convergence analysis below. The reader should keep this in mind.*

4 Convergence Properties

In this section, we investigate the global and local convergence properties of Algorithm 3.1. To this end, we always assume implicitly that Algorithm 3.1 does not terminate after a finite number of iterations with a stationary point of (3).

Our first result states that all iterates generated by Algorithm 3.1 are strictly feasible.

Lemma 4.1 *Algorithm 3.1 is well-defined and generates a sequence $\{x^k\}$ belonging to the strictly feasible set (l, u).*

Proof. From the statement of Algorithm 3.1, it follows that it is well-defined if and only if the line searches in Steps (S.5) and (S.6) are finite. This, however, is well-known in case of a projected gradient-type direction (see, e.g., [1, Section 2.3.2]) and follows immediately from (6) in case of a projected Newton-type direction together with the continuous differentiability of Ψ by Proposition 2.7.

The proof that the sequence $\{x^k\}$ is strictly feasible is by induction on k. The starting point x^0 was chosen from (l, u). Suppose that x^k belongs to (l, u) for some $k \geq 0$. If x^{k+1} is computed via Step (S.4), then this new iterate obviously belongs to the interior (l, u). If x^{k+1} is computed in Step (S.5), i.e., $x^{k+1} = x^k + t_k s_N^k$ with $s_N^k = \bar{x}_N^k - x^k$, we obtain

$$x^{k+1} = x^k + t_k(\bar{x}_N^k - x^k) = t_k \bar{x}_N^k + (1 - t_k)x^k.$$

Since x^k is strictly feasible in view of our induction hypothesis, and since $\bar{x}_N^k$ is at least feasible (in view of its very definition as a projection), it follows from $t_k \in (0, 1)$ that x^{k+1} is a strict convex combination of an interior point with a feasible point. Hence x^{k+1} belongs to the interior, too. Finally, the argument is similar if x^{k+1} is computed in Step (S.6). $\square$

We next state a global convergence result for Algorithm 3.1.

Theorem 4.2 *Any accumulation point x^* of a sequence $\{x^k\}$ generated by Algorithm 3.1 is at least a stationary point of the constrained reformulation (3) of the MCP (1). If the conditions of Proposition 2.8 are satisfied at x^*, then the accumulation point x^* is already a solution of the MCP (1).*

Proof. If the descent test (5) is satisfied infinitely many times, then it follows immediately that any accumulation point x^* of a sequence $\{x^k\}$ generated by Algorithm 3.1 satisfies $\Phi(x^*) = 0$, so any accumulation point is a solution of the MCP (1) and therefore, in particular, a stationary point of the constrained reformulation (3) of the MCP (1).

In the rest of this proof we can therefore assume that the descent test (5) is satisfied only a finite number of times; in fact, without loss of generality, we will assume that it is not satisfied at any iteration $k \in \mathbb{N}$. Under these conditions, we show that the sequence of search directions $\{s^k\}$ (where either $s^k = s^k_N$ or $s^k = s^k_G$) is gradient related with respect to $\{x^k\}$. The proof then follows from a general result in, e.g., [1, Proposition 2.2.1].

Let $\{x^k\}_K$ be a subsequence converging to a nonstationary point $\bar{x}$. We have to show that the subsequence $\{s^k\}_K$ is bounded and that $\limsup_{k\to\infty, k\in K} \nabla\Psi(x^k)^T s^k < 0$. From known results [1, Proof of Proposition 2.3.1], it follows that these two conditions are satisfied if $s^k = s^k_G$ denotes the projected gradient-type direction. So, without loss of generality, let us assume that $s^k = s^k_N$ is a projected Newton direction for all $k \in K$. Then we have

$$\nabla\Psi(x^k)^T s^k \leq -\rho\|s^k\|^{p_1} \tag{8}$$

and

$$\nabla\Psi(x^k)^T s^k \leq -\rho\|\Phi(x^k)\|^{p_2} \tag{9}$$

for all $k \in K$. Obviously, (8) and the Cauchy-Schwarz inequality imply that $\{s^k\}_K$ is bounded (recall that $p_1 > 1$). On the other hand, since $\{x^k\}_K \to \bar{x}$, we have

$$\lim_{k\to\infty, k\in K} \|\Phi(x^k)\| = \|\Phi(\bar{x})\|$$

by continuity. Since $\bar{x}$ is nonstationary for (3), it follows that $\bar{x}$ is not a solution of the MCP (1), and (9) therefore implies

$$\limsup_{k\to\infty, k\in K} \nabla\Psi(x^k)^T s^k \leq -\rho\|\Phi(\bar{x})\|^{p_2} < 0.$$

This proves that $\{s^k\}$ is indeed gradient related. $\square$

Given a solution x^* of the MCP (1), let us introduce the active and inactive sets associated with this solution by

$$\begin{aligned} \mathcal{A}_* &:= \{i \,|\, x^*_i = l_i \text{ or } x^*_i = u_i\}, \\ \mathcal{I}_* &:= \{i \,|\, x^*_i \in (l_i, u_i)\}. \end{aligned}$$

We next show that our approximations $\mathcal{A}_k$ and $\mathcal{I}_k$ computed in Step (S.2) of Algorithm 3.1 coincide with the exact sets $\mathcal{A}_*$ and $\mathcal{I}_*$ whenever $x^k \in (l, u)$ is sufficiently close to a strongly regular solution of the MCP (1).

Proposition 4.3 *Let x^* be a strongly regular solution of the mixed complementarity problem MCP (1). Then $\mathcal{A}_k = \mathcal{A}_*$ and $\mathcal{I}_k = \mathcal{I}_*$ for all $x^k \in (l, u)$ sufficiently close to x^*, where $\mathcal{A}_k$ and $\mathcal{I}_k$ denote the index sets computed in Step (S.2) of Algorithm 3.1.*

Proof. The proof is essentially the same as one for a corresponding result in [14], and we include it here only for the sake of completeness.

We first show that $\mathcal{A}_k$ is a subset of $\mathcal{A}_*$ for $x^k \in (l, u)$ close enough to x^* (the proof of this direction is actually independent of the fact that the solution x^* of the MCP (1) is strongly regular). To this end, define

$$\nu := \min\{\min\{x^*_i - l_i, u_i - x^*_i\} \,|\, i \in \mathcal{I}_*\} > 0,$$

i.e., ν is the smallest distance of the inactive components x_i^* to the boundary of the feasible set $[l, u]$. Let $x^k \in (l, u)$ be sufficiently close to x^* such that

$$\|x^k - x^*\| \leq \frac{\nu}{4}$$

and

$$c\sqrt{\|\Phi(x^k)\|} \leq \frac{\nu}{4}.$$

Choose $i \in \mathcal{A}_k$ arbitrarily. Then $x_i^k - l_i \leq \delta_k$ or $u_i - x_i^k \leq \delta_k$. If $x_i^k - l_i \leq \delta_k$, we get from

$$|x_i^k - x_i^*| \leq \|x^k - x^*\| \leq \frac{\nu}{4}$$

and

$$|x_i^k - l_i| \leq \delta_k \leq c\sqrt{\|\Phi(x^k)\|} \leq \frac{\nu}{4}$$

that

$$|x_i^* - l_i| \leq |x_i^* - x_i^k| + |x_i^k - l_i| \leq \frac{\nu}{2},$$

i.e., we have $i \in \mathcal{A}_*$ in view of the very definition of the constant ν.

If, on the other hand, we have $u_i - x_i^k \leq \delta_k$, we can show in a similar way that

$$|x_i^* - u_i| \leq \frac{\nu}{2},$$

so that $i \in \mathcal{A}_*$ follows also in this case. Hence we have $\mathcal{A}_k \subseteq \mathcal{A}_*$ for all $x^k \in (l, u)$ sufficiently close to x^*.

Conversely, we now show that $\mathcal{A}_* \subseteq \mathcal{A}_k$. The verification of this inclusion makes use of the strong regularity assumption for x^*. In fact, since x^* is strongly regular, it follows from Proposition 2.6 that there is a constant $\kappa > 0$ such that

$$|x_i^k - x_i^*| \leq \|\Phi(x^k)\|/\kappa \tag{10}$$

for all $i \in \{1, \dots, n\}$ and all $x^k \in (l, u)$ sufficiently close to x^*. Now let $i \in \mathcal{A}_*$ be any fixed index. Then $x_i^* = l_i$ or $x_i^* = u_i$.

Assume first that $x_i^* = l_i$. Then

$$x_i^k - l_i = |x_i^k - x_i^*| \leq \|\Phi(x^k)\|/\kappa$$

because of (10). Since $\|\Phi(x^k)\| \to 0$ for $x^k \to x^*$ and $\delta_k = O(\sqrt{\|\Phi(x^k)\|})$ in view of the definition of δ_k in Step (S.2) of Algorithm 3.1, we have

$$\|\Phi(x^k)\|/\kappa \leq \delta_k$$

for all $x^k \in (l, u)$ sufficiently close to x^*. Therefore, we obtain

$$x_i^k - l_i \leq \delta_k,$$

i.e., $i \in \mathcal{A}_k$.

If, on the other hand, we have $x_i^* = u_i$, we can show in an analogous way that

$$u_i - x_i^k \leq \delta_k.$$

This proves that $\mathcal{A}_* \subseteq \mathcal{A}_k$, so that both sets eventually coincide. This, in turn, implies that we also have $\mathcal{I}_* = \mathcal{I}_k$ for all $x^k \in (l, u)$ close enough to x^*. □

As a simple consequence of Proposition 4.3, we obtain the following technical result which will be used in our subsequent analysis.

Lemma 4.4 *Let x^* be a strongly regular solution of the MCP (1). Then there exists a constant $\kappa_1 > 0$ such that the submatrices $H^k_{\mathcal{I}_k\mathcal{I}_k}$ are nonsingular with*

$$\|(H^k_{\mathcal{I}_k\mathcal{I}_k})^{-1}\| \leq \kappa_1$$

for all $x^k \in (l, u)$ sufficiently close to x^*, where H^k denotes the matrix selected in Step (S.3) of Algorithm 3.1.

Proof. Assume the statement is not true. Then there exists a sequence $\{x^k\} \subseteq (l, u)$ converging to x^* as well as as sequence of matrices $\{H^k\}$ with $H^k \in \partial\Phi(x^k)$ such that $H^k_{\mathcal{I}_k\mathcal{I}_k}$ is either singular for almost all $k \in \mathbb{N}$ or $H^k_{\mathcal{I}_k\mathcal{I}_k}$ is nonsingular on a subsequence with $\|(H^k_{\mathcal{I}_k\mathcal{I}_k})^{-1}\| \to \infty$ on this subsequence.

In view of Proposition 4.3, we have $\mathcal{I}_k = \mathcal{I}_*$ for all $k \in \mathbb{N}$ sufficiently large. Since the generalized Jacobian is upper semicontinuous by [5, Proposition 2.6.2 (c)], it follows that the sequence $\{H^k\}$ is bounded. Subsequencing if necessary, we can therefore assume that $\{H^k\}$ converges to a matrix $H^* \in \mathbb{R}^{n \times n}$. In view of our previous considerations, it follows that the submatrix $H^*_{\mathcal{I}_*\mathcal{I}_*}$ must be singular.

However, since the generalized Jacobian is also a closed mapping by [5, Proposition 2.6.2 (b)], the limiting matrix H^* belongs to $\partial\Phi(x^*)$. Since $\mathcal{I}_* = \alpha$ with the index set α being defined before Definition 2.4, this means that the submatrix $H^*_{\alpha\alpha}$ is singular.

On the other hand, it is not difficult to see that $H_{\alpha\alpha} = -F'(x^*)_{\alpha\alpha}$ for any element $H \in \partial\Phi(x^*)$ (this follows, e.g., from Proposition 2.3 and Lemma 2.4 in [13], but can also be verified by direct calculation). Due to the assumed strong regularity of the solution x^*, the submatrix $F'(x^*)_{\alpha\alpha}$ and therefore also the submatrix $H^*_{\alpha\alpha}$ is nonsingular. This contradiction completes the proof. □

Also the next result is just technical, but will be exploited later.

Lemma 4.5 *Let x^* be a strongly regular solution of the MCP (1). Then*

$$d^k_{\mathcal{A}_k} = x^*_{\mathcal{A}_k} - x^k_{\mathcal{A}_k}$$

for all $x^k \in (l, u)$ sufficiently close to x^*, where $d^k_{\mathcal{A}_k}$ denotes the vector calculated in Step (S.3) of Algorithm 3.1.

Proof. By Proposition 4.3, we have $\mathcal{A}_k = \mathcal{A}_*$ for all $x^k \in (l,u)$ close enough to x^*. Now let $i \in \mathcal{A}_*$ be a fixed index. Then either $x_i^* = l_i$ or $x_i^* = u_i$. Consider the case where $x_i^* = l_i$. Since $\mathcal{A}_k = \mathcal{A}_*$, the definition of d_i^k yields that we have $d_i^k = l_i - x_i^k$ or $d_i^k = u_i - x_i^k$. If $d_i^k = l_i - x_i^k$ for all x^k sufficiently close to x^*, we are done.

Assume this is not true. Then there is a sequence $\{x^k\} \subseteq (l,u)$ converging to x^* such that $d_i^k = u_i - x_i^k$ for all $k \in \mathbb{N}$. The definition of $\mathcal{A}_k$ and δ_k then implies

$$0 \le u_i - x_i^k \le \delta_k \le c\sqrt{\|\Phi(x^k)\|} \to 0$$

and therefore $x_i^k \to u_i$ for $k \to \infty$. On the other hand, we have $x_i^k \to x_i^*$. This implies $x_i^* = u_i$, a contradiction to our choice of the index $i \in \mathcal{A}_*$.

Since the proof is similar for the case where $x_i^* = u_i$, we are done. □

We next give an estimate on the length of the Newton-type search direction d^k computed in Step (S.3) of Algorithm 3.1.

Lemma 4.6 *Let x^* be a strongly regular solution of the MCP (1). Then there is a constant $\kappa_2 > 0$ such that*

$$\|d^k\| \le \kappa_2 \|\Phi(x^k)\|$$

for all $x^k \in (l,u)$ sufficiently close to x^*, where d^k is the vector computed in Step (S.3) of Algorithm 3.1.

Proof. We obviously have

$$\|d^k\| \le \|d^k_{\mathcal{A}_k}\| + \|d^k_{\mathcal{I}_k}\|. \tag{11}$$

Using Lemma 4.5 and Proposition 2.6, it follows that there is a constant $\kappa > 0$ such that

$$\|d^k_{\mathcal{A}_k}\| = \|x^k_{\mathcal{A}_k} - x^*_{\mathcal{A}_k}\| \le \|x^k - x^*\| \le \|\Phi(x^k)\|/\kappa \tag{12}$$

for all $x^k \in (l,u)$ close enough to x^*.

In order to estimate the growth behaviour of the components $d^k_{\mathcal{I}_k}$, we first derive the inequality

$$\begin{array}{rcl} \|d^k_{\mathcal{I}_k}\| & = & \|(H^k_{\mathcal{I}_k\mathcal{I}_k})^{-1}\left(\Phi(x^k)_{\mathcal{I}_k} + H^k_{\mathcal{I}_k\mathcal{A}_k} d^k_{\mathcal{A}_k}\right)\| \\ & \le & \|(H^k_{\mathcal{I}_k\mathcal{I}_k})^{-1}\| \left(\|\Phi(x^k)_{\mathcal{I}_k}\| + \|H^k_{\mathcal{I}_k\mathcal{A}_k}\|\,\|d^k_{\mathcal{A}_k}\|\right) \end{array} \tag{13}$$

from the linear system (4). Since the sequence $\{H^k\}$ with $H^k \in \partial\Phi(x^k)$ is bounded for any bounded sequence $\{x^k\}$, there is a constant $\kappa_3 > 0$ such that

$$\|H^k_{\mathcal{I}_k\mathcal{A}_k}\| \le \kappa_3$$

for all x^k sufficiently close to x^*. Using this, (12), (13), and Lemma 4.4, we obtain

$$\begin{array}{rcl} \|d^k_{\mathcal{I}_k}\| & \le & \kappa_1\left(\|\Phi(x^k)_{\mathcal{I}_k}\| + \kappa_3\|d^k_{\mathcal{A}_k}\|\right) \\ & \le & \kappa_1\left(\|\Phi(x^k)\| + \kappa_3\|\Phi(x^k)\|/\kappa\right) \\ & = & \kappa_4\|\Phi(x^k)\| \end{array} \tag{14}$$

with $\kappa_4 := \kappa_1 + \kappa_3/\kappa$. The statement now follows from (11), (12) and (14) by setting $\kappa_2 := \kappa_4 + 1/\kappa$. $\square$

Using our previous results, we are now in the position to show that the entire sequence $\{x^k\}$ generated by Algorithm 3.1 converges to a solution x^* of the MCP (1) under suitable assumptions.

Proposition 4.7 *Let x^* be a strongly regular solution of the MCP (1), and assume that x^* is an accumulation point of a sequence $\{x^k\}$ generated by Algorithm 3.1. Then the entire sequence $\{x^k\}$ converges to x^*.*

Proof. Let $\{x^k\}_{k\in K}$ be a subsequence converging to the strongly regular solution x^* of the MCP (1). In order to prove that the entire sequence $\{x^k\}$ converges to x^*, it is enough to show that $\{\|x^{k+1} - x^k\|\}_{k\in K} \to 0$, see, e.g., [24, Lemma 4.10]. To this end, we consider three cases depending on how x^{k+1} gets calculated in Algorithm 3.1.

Case 1: For all $k \in K, x^{k+1}$ is computed in Step (S.4) of Algorithm 3.1.
Then we have

$$\|x^{k+1} - x^k\| = \tau_k\|d^k\| \le \|d^k\| \le \|d^k_{\mathcal{I}_k}\| + \|d^k_{\mathcal{A}_k}\|. \tag{15}$$

Using Lemma 4.5, we have

$$d^k_{\mathcal{A}_k} = x^*_{\mathcal{A}_k} - x^k_{\mathcal{A}_k}$$

for all $k \in K$ sufficiently large. It therefore follows that

$$\|d^k_{\mathcal{A}_k}\| \to 0 \tag{16}$$

for $k \in K, k \to \infty$.

Next consider the subvector $d^k_{\mathcal{I}_k}$. Since the generalized Jacobian is upper semicontinuous, the sequence $\{H^k\}_{k\in K}$ is bounded. Hence, there exists a constant $\kappa_5 > 0$ such that

$$\|H^k_{\mathcal{I}_k\mathcal{A}_k}\| \le \kappa_5$$

for all $k \in K$ sufficiently large. We therefore obtain from Lemma 4.4 that

$$\begin{aligned} \|d^k_{\mathcal{I}_k}\| &\le \|(H^k)^{-1}_{\mathcal{I}_k\mathcal{I}_k}\| \left(\|\Phi(x^k)_{\mathcal{I}_k}\| + \|H^k_{\mathcal{I}_k\mathcal{A}_k}\| \, \|d^k_{\mathcal{A}_k}\|\right) \\ &\le \kappa_1 \left(\|\Phi(x^k)_{\mathcal{I}_k}\| + \kappa_5\|d^k_{\mathcal{A}_k}\|\right), \end{aligned} \tag{17}$$

cf. (13). From (16), we already know that $d^k_{\mathcal{A}_k} \to 0$ for $k \in K, k \to \infty$. Furthermore, since $\Phi(x^*) = 0$ at the solution x^* of MCP, we also have $\|\Phi(x^k)_{\mathcal{I}_k}\| \to 0$ for $k \in K, k \to \infty$. Hence (17) yields

$$\|d^k_{\mathcal{I}_k}\| \to 0 \tag{18}$$

for $k \in K, k \to \infty$. Now, using (16) and (18), we immediately obtain from (15) that

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

for $k \in K, k \to \infty$.

Case 2: For all $k \in K, x^{k+1}$ is computed in Step (S.5) of Algorithm 3.1.
Then we have $x^{k+1} = x^k + t_k s_N^k$, and the search direction s_N^k satisfies, in particular, the descent condition

$$\nabla\Psi(x^k)^T s_N^k \leq -\rho \|s_N^k\|^{p_1} \tag{19}$$

for all $k \in K$. Since the subsequence $\{x^k\}_{k\in K}$ converges to x^* and $\{\Psi(x^k)\}$ is monotonically decreasing, it follows that the entire sequence $\{\Psi(x^k)\}$ converges to $\Psi(x^*)$. Therefore, we have $\{\Psi(x^{k+1}) - \Psi(x^k)\} \to 0$. Hence, the Armijo line search rule in Step (S.5) of Algorithm 3.1 implies

$$t_k \nabla\Psi(x^k)^T s_N^k \to 0$$

for $k \in K, k \to \infty$. In view of (19), it is easy to see that this implies

$$t_k \|s_N^k\| \to 0.$$

Consequently, we obtain

$$\|x^{k+1} - x^k\| = t_k \|s_N^k\| \to 0$$

for $k \in K, k \to \infty$.

Case 3: For all $k \in K$, x^{k+1} is computed in Step (S.6) of Algorithm 3.1.
Then we have $x^{k+1} = x^k + t_k s_G^k$, and the definition of s_G^k yields

$$\begin{array}{rcl} \|x^{k+1} - x^k\| & = & t_k \|s_G^k\| \\ & = & t_k \|\bar{x}_G^k - x^k\| \\ & = & t_k \|[x^k - \gamma\nabla\Psi(x^k)]_+ - x^k\| \\ & \leq & \|[x^k - \gamma\nabla\Psi(x^k)]_+ - [x^k]_+\| \\ & \leq & \gamma \|\nabla\Psi(x^k)\| \\ & \to & 0 \end{array}$$

for $k \in K, k \to \infty$; here we used that fact that $x^k = [x^k]_+$ (since $x^k \in (l, u)$ by Lemma 4.1), that the projection operator is nonexpansive, that $t_k \in (0, 1)$ and that $\nabla\Psi(x^*) = 0$ (since Ψ is continuously differentiable around x^* and x^* is a global minimum and therefore an unconstrained stationary point of Ψ).

Finally, if the subsequence $\{x^k\}_{k\in K}$ contains finitely or infinitely many iterates k with x^{k+1} being calculated either in Step (S.4) or in Step (S.5) or in Step (S.6) of Algorithm 3.1, then we can easily combine the above three cases and still obtain

$$\{\|x^{k+1} - x^k\|\}_{k\in K} \to 0.$$

This completes the proof. □

Our next result basically says that the Newton-type search direction d^k computed in Step (S.3) of Algorithm 3.1 is a "superlinearly/quadratically convergent search direction".

Lemma 4.8 *Let x^* be a strongly regular solution of the MCP (1), and let $\{x^k\}$ be any sequence (not necessarily generated by Algorithm 3.1) converging to x^*. Then*

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|),$$

where d^k denotes the Newton-type search direction calculated in Step (S.3) of Algorithm 3.1. Furthermore, if F' is locally Lipschitzian around x^*, then

$$\|x^k + d^k - x^*\| = O(\|x^k - x^*\|^2).$$

Proof. The proof is essentially the same as one given for a corresponding result in [14], and is presented here for the sake of completeness.

Since $x^k \to x^*$ by assumption, it follows from Lemma 4.4 that the search directions d^k as computed in Step (S.3) of Algorithm 3.1 actually exist for all k sufficiently large. Moreover, Proposition 4.3 guarantees that we have $\mathcal{A}_k = \mathcal{A}_*$ and $\mathcal{I}_k = \mathcal{I}_*$ for all k large enough. In the rest of this proof, we therefore assume implicitly that k is sufficiently large so that the above two statements hold.

First consider an index $i \in \mathcal{A}_*$. Then $x_i^* = l_i$ or $x_i^* = u_i$. If $x_i^* = l_i$, then $d_i^k = l_i - x_i^k$ by Lemma 4.5, and it follows that

$$|x_i^k + d_i^k - x_i^*| = |x_i^k + l_i - x_i^k - l_i| = 0 = o(\|x^k - x^*\|). \tag{20}$$

Similarly, if $x_i^* = u_i$, then $d_i^k = u_i - x_i^k$, and we obtain

$$|x_i^k + d_i^k - x_i^*| = |x_i^k + u_i - x_i^k - u_i| = 0 = o(\|x^k - x^*\|). \tag{21}$$

Next consider indices $i \in \mathcal{I}_*$. Using the notation $\mathcal{N} := \{1, \ldots, n\}$, we obtain from the linear system (4), Lemma 4.5 and the fact that $\mathcal{I}_k = \mathcal{I}_*$:

$$\begin{aligned} H^k_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} + d^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) &= H^k_{\mathcal{I}_*\mathcal{I}_*} d^k_{\mathcal{I}_*} + H^k_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -\Phi_{\mathcal{I}_*}(x^k) - H^k_{\mathcal{I}_*\mathcal{A}_*} d^k_{\mathcal{A}_*} + H^k_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -\Phi_{\mathcal{I}_*}(x^k) + \Phi_{\mathcal{I}_*}(x^*) + H^k_{\mathcal{I}_*\mathcal{A}_*}(x^k_{\mathcal{A}_*} - x^*_{\mathcal{A}_*}) \\ &\quad + H^k_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -\Phi_{\mathcal{I}_*}(x^k) + \Phi_{\mathcal{I}_*}(x^*) + H^k_{\mathcal{I}_*\mathcal{N}}(x^k - x^*), \end{aligned} \tag{22}$$

since the solution x^* satisfies $\Phi(x^*) = 0$. In view of Proposition 2.3 (a), we have

$$|\Phi_i(x^k) - \Phi_i(x^*) - H^k_{i\cdot}(x^k - x^*)| = o(\|x^k - x^*\|)$$

for all $i \in \{1, \ldots, n\}$, where $H^k_{i\cdot}$ denotes the i-th row of the matrix H^k. Hence we obtain

$$\|\Phi_{\mathcal{I}_*}(x^k) - \Phi_{\mathcal{I}_*}(x^*) - H^k_{\mathcal{I}_*\mathcal{N}}(x^k - x^*)\| = o(\|x^k - x^*\|),$$

so it follows from (22) and Lemma 4.4 that

$$\|x^k_{\mathcal{I}_*} + d^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}\| = o(\|x^k - x^*\|). \tag{23}$$

Using (20), (21) and (23), we have

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

In a similar way, one can exploit Proposition 2.3 (b) and show that

$$\|x^k + d^k - x^*\| = O(\|x^k - x^*\|^2)$$

holds if F' is locally Lipschitzian around x^*. □

We now show that we maintain the superlinear/quadratic convergence properties of our Newton-type search direction d^k if we combine it with the corresponding stepsize τ_k computed in Step (S.4) of Algorithm 3.1.

Lemma 4.9 *Let x^* be a strongly regular solution of the MCP (1), and let $\{x^k\}$ be any sequence (not necessarily generated by Algorithm 3.1) converging to x^*. Then*

$$\|x^k + \tau_k d^k - x^*\| = o(\|x^k - x^*\|),$$

where d^k denotes the Newton-type search direction calculated in Step (S.3) of Algorithm 3.1 and $\tau_k > 0$ is the stepsize computed in Step (S.4) of Algorithm 3.1. Furthermore, if F' is locally Lipschitzian around x^*, then

$$\|x^k + \tau_k d^k - x^*\| = O(\|x^k - x^*\|^2).$$

Proof. Since $\{x^k\} \to x^*$ by assumption, we have $\|\Phi(x^k)\| \to \|\Phi(x^*)\| = 0$ and therefore eventually

$$\tau_k = 1 - \|\Phi(x^k)\| \to 1.$$

Hence Lemma 4.8 implies

$$\begin{aligned} \|x^k + \tau_k d^k - x^*\| &\leq \|x^k + d^k - x^*\| + (1 - \tau_k)\|d^k\| \\ &= o(\|x^k - x^*\|) + o(\|d^k\|). \end{aligned}$$

However, we have $\|d^k\| = O(\|\Phi(x^k)\|)$ by Lemma 4.6. Moreover, since Φ is locally Lipschitz continuous around the solution x^*, there is a constant $L > 0$ such that

$$\|\Phi(x^k)\| = \|\Phi(x^k) - \Phi(x^*)\| \leq L\|x^k - x^*\|.$$

This implies that

$$\|x^k + \tau_k d^k - x^*\| = o(\|x^k - x^*\|) + o(\|\Phi(x^k)\|) = o(\|x^k - x^*\|).$$

The proof of the second part is similar by using the second part of Lemma 4.8 and exploiting the precise definition of τ_k (note that, in order to verify the first statement, it was enough to have $\tau_k \to 1$, whereas the proof of the second statement requires τ_k to converge to 1 with a certain speed). □

Summarizing the previous results in an appropriate way, we are now able to state our main local convergence result for Algorithm 3.1.

Theorem 4.10 *Let x^* be a strongly regular solution of the MCP (1), and assume that x^* is an accumulation point of a sequence $\{x^k\}$ generated by Algorithm 3.1. Then the following statements hold:*

(a) The entire sequence $\{x^k\}$ converges to x^*.

(b) Eventually, the algorithm takes only the Newton-type directions from Step (S.3) of Algorithm 3.1.

(c) The rate of convergence is Q-superlinear.

(d) If F' is locally Lipschitzian around x^*, the rate of convergence is actually Q-quadratic.

Proof. Statement (a) is a direct consequence of Proposition 4.7.

In order to prove Statement (b), we first recall that the Newton-type search direction d^k from Step (S.3) eventually exists by Lemma 4.4, and that $x^k + \tau_k d^k$ stays strictly feasible as a relatively simple consequence of Proposition 4.3. To see that $x^k + \tau_k d^k$ satisfies the descent test from (5), we first note that Φ is locally Lipschitz around x^*. Hence we have

$$\|\Phi(x^k + \tau_k d^k) - \Phi(x^*)\| \leq L_\Phi \|x^k + \tau_k d^k - x^*\| \tag{24}$$

for some constant $L_\Phi > 0$ since $x^k \to x^*$ by Part (a) and $\tau_k d^k \to 0$ by, e.g., Lemma 4.6. Moreover, all elements $H \in \partial\Phi(x^*)$ are nonsingular by Proposition 2.5. Therefore, it follows from the Inverse Function Theorem for locally Lipschitzian mappings (see [5, Theorem 7.1.1]) that Φ is invertable around $\Phi(x^*)$, and that the inverse function Φ^{-1} is also locally Lipschitzian. Consequently, we have

$$\|\Phi^{-1}(\Phi(x^k)) - \Phi^{-1}(\Phi(x^*))\| \leq L_{\Phi^{-1}} \|\Phi(x^k) - \Phi(x^*)\| \tag{25}$$

for some constant $L_{\Phi^{-1}}$ and all $k \in \mathbb{N}$ sufficiently large. Now, using (24), (25), Lemma 4.9 and the fact that $\Phi(x^*) = 0$, we immediately obtain

$$\begin{aligned}
\|\Phi(x^k + \tau_k d^k)\| &= \|\Phi(x^k + \tau_k d^k) - \Phi(x^*)\| \\
&= O(\|x^k + \tau_k d^k - x^*\|) \\
&= o(\|x^k - x^*\|) \\
&= o(\|\Phi^{-1}(\Phi(x^k)) - \Phi^{-1}(\Phi(x^*))\|) \\
&= o(\|\Phi(x^k) - \Phi(x^*)\|) \\
&= o(\|\Phi(x^k)\|).
\end{aligned}$$

Therefore, the descent test (5) in Step (S.4) of Algorithm 3.1 will eventually be satisfied, and we have $x^{k+1} = x^k + \tau_k d^k$ for all $k \in \mathbb{N}$ sufficiently large with τ_k being computed in Step (S.4) of Algorithm 3.1. Hence Statement (b) holds.

Statements (c) and (d) now follow immediately from Lemma 4.9 and the previous considerations. □

5 Numerical Results

We implemented Algortihm 3.1 in MATLAB for nonlinear complementarity problems using the parameter setting

$$\beta = 0.5,\ \sigma = \delta = 10^{-4},\ \tau = \omega = 0.995,\ \rho = 10^{-12},\ p_1 = 2.1,\ p_2 = c = \gamma = 1$$

and the termination criteria

$$\Psi(x^k) \leq 10^{-12} \quad \text{or} \quad k > 500.$$

The actual implementation differs slightly from the description in Algorithm 3.1. In fact, we allow a nonmonotone line search in Step (S.5) instead of the standard (monotone) Armijo rule (see [19] for the original reference), and we try to accept our Newton-type search direction in Step (S.4) even if the acceptance criterion (5) is not satisfied, i.e., we allow the stepsize τ_k used in (S.4) to become a bit smaller.

The test problems whose names are given in Table 5.1 are taken from the MCPLIB collection, see [8]. However, we had to exclude a few examples (namely problems `powell` [second, third and fourth starting point], and `scarfasum` [all three starting points]) because domain violations occured during the iteration (although all iterates were strictly feasible!). The results of our test runs are given in Table 5.1 whose columns have the following meanings:

problem:	name of the test example
n:	dimension of the test example
SP:	number of starting point
k:	number of iterations until termination
F-ev.	number of function evaluations until termination
$\Psi(x^f)$:	value of $\Psi(x)$ at final iterate $x = x^f$
$\|\nabla\Psi(x^f)\|$	value of $\|\nabla\Psi(x)\|$ at final iterate $x = x^f$

Note that we had to change the starting point for some examples by a small perturbation in order to begin with a strictly feasible point, as required by our algorithm.

We believe that the numerical results are quite reasonable, especially since the parameters have not been tuned very much. For some problems, we have a relatively high number of function evaluations; however, this has mainly to do with our simple implementation of the projected gradient-type method.

6 Concluding Remarks

We presented a new method for the solution of mixed complementarity problems which, besides having a strong local and global convergence theory, generates only strictly feasible iterates and has to solve only one linear system of equations of reduced dimension at each iteration. Theoretically, it seems possible to generalize our idea to solve systems of equations with more complicated constraints (as they occur, e.g., in the context of variational inequalities) than just box constraints. However, from a practical point of view, this does not seem

Table 5.1: Numerical results for Algorithm 3.1.

problem	n	SP	k	F-ev.	$\Psi(x^f)$	$\|\nabla\Psi(x^f)\|$
bertsekas	15	1	30	212	3.7e-15	3.4e-6
bertsekas	15	2	29	166	3.7e-15	3.4e-6
bertsekas	15	3	47	227	5.6e-17	5.3e-7
billups	1	1	—	—	—	—
colvdual	20	1	39	179	9.7e-14	2.7e-6
colvdual	20	2	29	135	9.7e-14	2.7e-6
colvnlp	15	1	49	258	9.4e-14	3.6e-6
colvnlp	15	2	18	46	9.4e-14	3.6e-6
cycle	1	1	6	7	8.6e-22	3.3e-11
explcp	16	1	6	7	3.2e-13	8.0e-7
hanskoop	14	1	48	160	1.2e-21	4.1e-10
hanskoop	14	2	41	121	1.1e-13	3.1e-6
hanskoop	14	3	25	132	8.9e-15	3.5e-7
hanskoop	14	4	39	195	3.7e-13	7.4e-6
hanskoop	14	5	460	3,614	8.0e-16	1.1e-7
josephy	4	1	5	8	2.9e-16	9.7e-8
josephy	4	2	6	7	4.7e-22	2.2e-10
josephy	4	3	15	16	1.3e-13	1.1e-6
josephy	4	4	5	6	6.5e-18	1.9e-8
josephy	4	5	6	7	7.2e-15	2.5e-7
josephy	4	6	7	10	7.6e-21	2.9e-10
kojshin	4	1	7	10	1.9e-18	4.9e-9
kojshin	4	2	6	7	1.5e-14	2.2e-6
kojshin	4	3	17	18	5.9e-15	2.1e-6
kojshin	4	4	6	8	2.1e-19	5.9e-9
kojshin	4	5	6	7	1.7e-17	2.0e-8
kojshin	4	6	6	12	5.9e-14	5.5e-6
mathinum	3	1	7	9	6.4e-14	7.8e-7
mathinum	3	2	5	6	5.1e-18	7.0e-9
mathinum	3	3	11	14	2.1e-20	4.4e-10
mathinum	3	4	5	6	9.0e-14	9.3e-7
mathisum	4	1	6	7	5.2e-16	6.1e-8
mathisum	4	2	6	7	2.0e-17	1.2e-8
mathisum	4	3	11	13	2.6e-18	4.3e-9
mathisum	4	4	5	6	1.7e-13	1.1e-6
nash	10	1	8	9	1.0e-16	7.4e-7
nash	10	2	11	12	1.1e-13	2.8e-5
pgvon105		1	—	—	—	—
pgvon106		1	—	—	—	—

Table 5.1 (continued): Numerical results for Algorithm 3.1

problem	n	SP	k	F-ev.	$\Psi(x^f)$	$\|\nabla\Psi(x^f)\|$
powell	16	1	95	691	6.8e-13	1.0e-5
scarfanum	13	1	35	288	9.9e-17	1.0e-7
scarfanum	13	2	35	270	9.9e-17	1.0e-7
scarfanum	13	3	37	289	9.9e-17	1.0e-7
scarfbnum	39	1	42	121	5.7e-15	4.5e-6
scarfbnum	39	2	31	144	2.4e-15	2.0e-5
scarfbsum	40	1	29	167	2.7e-21	8.1e-9
scarfbsum	40	2	—	—	—	—
sppe	27	1	66	365	8.6e-13	9.4e-7
sppe	27	2	11	25	1.7e-14	5.0e-7
tobin	42	1	9	12	2.3e-14	9.1e-7
tobin	42	2	9	16	6.4e-16	8.2e-8

to be very useful due to the fact that it can be quite expensive to calculate projections onto these more general constraint sets, as required by our algorithm. So we are currently looking for a somewhat different strategy in order to avoid the calculation of any projections.

References

[1] D.P. Bertsekas: *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995.

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

[3] S.C. Billups and M.C. Ferris: *QPCOMP: A quadratic programming based solver for mixed complementarity problems.* Mathematical Programming 76, 1997, pp. 533–562.

[4] B. Chen, X. Chen and C. Kanzow: *A penalized Fischer-Burmeister NCP-function.* Preprint 126, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, September 1997 (revised May 1998 and March 1999).

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

[6] R.W. Cottle, J.-S. Pang and R.E. Stone: *The Linear Complementarity Problem.* Academic Press, San Diego, CA, 1992.

[7] 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.

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

[9] 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.

[10] 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.

[11] F. Facchinei, A. Fischer and C. Kanzow: *On the accurate identification of active constraints.* SIAM Journal on Optimization 9, 1999, pp. 14–32.

[12] 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.

[13] M.C. Ferris, C. Kanzow and T.S. Munson: *Feasible descent algorithms for mixed complementarity problems.* Mathematical Programming, to appear.

[14] M.C. Ferris, C. Kanzow and T.S. Munson: *An active set-type method for constrained nonlinear systems.* Technical Report, forthcoming.

[15] M.C. Ferris and J.-S. Pang: *Engineering and economic applications of complementarity problems.* SIAM Review 39, 1997, pp. 669–713.

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

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

[18] A. Fischer: *A new constrained optimization reformulation for complementarity problems.* Journal of Optimization Theory and Applications 97, 1998, pp. 105–117.

[19] 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.

[20] C. Kanzow: *An inexact QP-based method for nonlinear complementarity problems.* Numerische Mathematik 80, 1998, pp. 557–577.

[21] C. Kanzow and H. Kleinmichel: *A new class of semismooth Newton-type methods for nonlinear complementarity problems.* Computational Optimization and Applications 11, 1998, pp. 227–251.

[22] J. Liu: *Strong stability in variational inequalities.* SIAM Journal on Control and Optimization 33, 1995, pp. 725–749.

[23] R. Mifflin: *Semismooth and semiconvex functions in constrained optimization.* SIAM Journal on Control and Optimization 15, 1977, pp. 957–972.

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

[25] J.-S. Pang and S.A. Gabriel: *NE/SQP: A robust algorithm for the nonlinear complementarity problem.* Mathematical Programming 60, 1993, pp. 295–337.

[26] J.-S. Pang and L. Qi: *Nonsmooth equations: Motivation and algorithms.* SIAM Journal on Optimization 3, 1993, pp. 443–465.

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

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

[29] L. Qi and D. Sun: *On NCP-functions.* Computational Optimization and Applications 13, 1999, pp. 201–220.

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

[31] T.F. Rutherford: *MILES: A mixed inequality and nonlinear equation solver.* Working Paper, Department of Economics, University of Colorado, Boulder, 1993.

[32] D. Sun: *Projected Newton-type methods for constrained nonsmooth equations.* Talk presented at the "Workshop on Nonsmooth and Smoothing Methods" in Hongkong, China, December 1998.

[33] M. Ulbrich: *Non-monotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems.* Technical Report TUM-M9906, Technical University of Munich, Munich, Germany, April 1999.