INEXACT SEMISMOOTH NEWTON METHODS FOR LARGE-SCALE COMPLEMENTARITY PROBLEMS

Christian Kanzow

Preprint 249 January 2003

University of Würzburg
Institute of Applied Mathematics and Statistics
Am Hubland
97074 Würzburg
Germany

e-mail: kanzow@mathematik.uni-wuerzburg.de

January 29, 2003

Abstract. The semismooth Newton method is a nonsmooth Newton-type method applied to a suitable reformulation of the complementarity problem as a nonlinear and nonsmooth system of equations. It is one of the standard methods for solving these kind of problems, and it can be implemented in an inexact way so that all linear systems of equations have to be solved only inexactly. However, from a practical point of view, this inexact Newton method seems to have a significantly worse behaviour than its exact counterpart. The aim of this paper is therefore to show that the inexact Newton method can also be used in a reliable and efficient way at least for some classes of problems. We illustrate this statement by some numerical examples with up to one million variables.

Key Words. Complementarity problems, Newton's method, semismooth functions, obstacle problems, optimal control problems.

1 Introduction

The complementarity problem is to find a solution $x^* \in \mathbb{R}^n$ of the system

$$x_i \geq 0, \;\; F_i(x) \geq 0, \;\; x_i F_i(x) = 0 \quad \forall i = 1, \ldots, n, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ denotes a continuously differentiable function. Many applications from the engineering sciences, economics, game theory etc. lead to problems of this kind, see [15] for a survey.

Most algorithms for the solution of the complementarity problem are based on a suitable reformulation of the complementarity problem either as a system of equations, as an optimization problem or as a fixed-point problem etc. We refer the interested reader to the survey papers [20, 13] for the basic ideas of some algorithms. In fact, many of these reformulations can be obtained for the slightly more general mixed complementarity problem. However, in order to keep our notation as simple as possible, our focus will be on the standard complementarity problem (1).

The semismooth method to be considered in more detail in this paper is based on a reformulation of the complementarity problem (1) as a nonlinear system of equations

$$\Phi(x) = 0, \tag{2}$$

where $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is defined componentwise by

$$\Phi_i(x) := \varphi\big(x_i, F_i(x)\big) \tag{3}$$

for some mapping $\varphi : \mathbb{R}^2 \to \mathbb{R}$ having the property

$$\varphi(a, b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0. \tag{4}$$

Clearly, this property guarantees that a vector $x^* \in \mathbb{R}^n$ is a solution of the complementarity problem (1) if and only if x^* solves the system of equations (2). Applying Newton's method to the system (2) then leads to the class of semismooth methods, see [6, 26, 40, 12, 25, 3, 39] for some references. Depending on the choice of the mapping φ, we obtain different methods with different properties. Some of these methods will be discussed in more detail in Sections 2 and 3.

Most semismooth methods have a very strong theoretical background and seem to be quite reliable and efficient also from a numerical point of view, at least when an exact Newton-type method is applied to the system (2). However, in the large-scale case, we may not be able to find the exact solution of the corresponding linearized equation

$$H_k d = -\Phi(x^k), \tag{5}$$

where $H_k \in \mathbb{R}^{n \times n}$ stands for the Jacobian $\Phi'(x^k)$ or a suitable approximation to it. Fortunately, the theory for the semismooth method can be carried over to the inexact case where we allow iterative solvers to find only approximate solutions of the linear subproblems (5).

Hence the inexact semismooth Newton method seems to be a suitable candidate for the solution of large scale complementarity problems. This is in contrast to the PATH solver, see [9, 14] for more details, which has to work with exact factorizations of a certain matrix. Hence the PATH solver, probably the most widely used solver in the complementarity field, cannot be adapted to the solution of large scale problems, at least not in a direct way.

On the other hand, the inexact semismooth method, being a potential candidate for large-scale problems and having a good theoretical background, seems to have a significantly worse numerical behaviour than its exact counterpart. We will come back to this point in Section 2. The aim of this paper is therefore to show that the inexact semismooth method can nevertheless be applied in a reliable and efficient way at least to certain classes of problems.

The organization of this paper is as follows. In Section 2, we recall the properties of some exact and inexact semismooth Newton methods based on some Fischer-Burmeister-type functions φ. We also discuss its advantages and disadvantages when iterative solvers are used for the approximate solution of the linear subproblems (5). Section 3 reviews some properties of the semismooth method when the minimum function is used for φ. This reformulation leads to completely different properties and is therefore covered in an extra section. Finally, we illustrate the numerical properties of the inexact semismooth methods when applied to some special problems in Section 4. Taking into account the special structure of these problems, we will show that the inexact semismooth solver is able to deal with problems with up to one million variables.

The Euclidean space is denoted by $\mathbb{R}^n$. If $x \in \mathbb{R}^n$ is a vector with components x_i for $i = 1, \ldots, n$ and $\mathcal{I} \subseteq \{1, \ldots, n\}$ denotes a subset, then $x_{\mathcal{I}}$ is the vector in $\mathbb{R}^{|\mathcal{I}|}$ with components $x_i, i \in \mathcal{I}$. Similarly, if $A \in \mathbb{R}^{n \times n}$ is a given matrix with elements a_{ij} and $\mathcal{I}, \mathcal{J} \subseteq \{1, \ldots, n\}$ are two subsets, we write $A_{\mathcal{I}\mathcal{J}}$ for the submatrix comprised by the elements a_{ij} with $i \in \mathcal{I}$ and $j \in \mathcal{J}$. The unique symmetric positive (semi-) definite square root of a symmetric positive (semi-) definite matrix $A \in \mathbb{R}^{n \times n}$ is denoted by $A^{1/2}$. Finally, vector or matrix inequalities like $x \geq 0$ and $A \geq 0$ are to be understood componentwise.

2 Fischer-Burmeister-type Functions

In this section we consider the reformulation (2), (3) of the complementarity problem (1) based on two functions φ from (4), namely the Fischer-Burmeister function

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

introduced in [16] and the penalized Fischer-Burmeister function

$$\varphi_{\mathrm{PFB}}(a, b) := \lambda \varphi_{\mathrm{FB}}(a, b) + (1 - \lambda) \max\{0, a\} \max\{0, b\} \quad \big(\lambda \in (0, 1)\big)$$

from [3]. Consequently, the mapping φ denotes any of these two mappings throughout this section. The properties of the corresponding semismooth methods are discussed in [6] and [3], respectively.

A major advantage of these two reformulations is that the merit function

$$\Psi(x) := \frac{1}{2}\Phi(x)^T\Phi(x) \tag{6}$$

turns out to be continuously differentiable although the equation operator Φ itself is nonsmooth. Hence, if we apply the nonsmooth Newton method from [34, 33] to the equation (2), we can globalize this method in a standard way. One simple example of such a globalized method is given in our following algorithm from [6], where $\partial\Phi(x)$ denotes the generalized Jacobian of Φ at a point x, see [4].

Algorithm 2.1 *(Semismooth Method Based on Fischer-Burmeister-type Functions)*

(S.0) Choose $x^0 \in \mathbb{R}^n, \rho > 0, \beta \in (0,1), \sigma \in (0,1/2), p > 2, \varepsilon \geq 0$, *and set* $k := 0$.

(S.1) If $\|\nabla\Psi(x^k)\| \leq \varepsilon$*: STOP.*

(S.2) Select an element $H_k \in \partial\Phi(x^k)$. *Find a solution* $d^k \in \mathbb{R}^n$ *of the linear system*

$$H_k d = -\Phi(x^k). \tag{7}$$

If this system is not solvable or if the descent condition

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

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

(S.3) Compute $t_k := \max\{\beta^\ell \mid \ell = 0, 1, 2, \ldots\}$ *such that*

$$\Psi(x^k + t_k d^k) \leq \Psi(x^k) + \sigma t_k \nabla\Psi(x^k)^T d^k.$$

(S.4) Set $x^{k+1} := x^k + t_k d^k, k \leftarrow k+1$, *and go to (S.1).*

It can be shown that Algorithm 2.1 is locally superlinearly convergent under a certain assumption (called R-regularity in [6] and taken from [35]). If, in addition, F' is locally Lipschitzian, then this method is locally quadratically convergent. Furthermore, one can show that every accumulation point of a sequence generated by Algorithm 2.1 is at least a stationary point of the merit function Ψ, see [6]. Under further assumptions, satisfied, e.g., for monotone mappings F, such a stationary point is already a solution of the complementarity problem (1). Moreover, the existence of an accumulation point can also be guaranteed. For example, F being strongly monotone implies the existence of an accumulation point when using the Fischer-Burmeister function, and F being monotone implies the same for the penalized Fischer-Burmeister function provided that the complementarity problem (1) has a strictly feasible point, i.e., a point $\hat{x} \in \mathbb{R}^n$ such that $\hat{x} > 0$ and $F(\hat{x}) > 0$. This latter difference shows that the penalized Fischer-Burmeister function has slighlty stronger theoretical properties. Extensive numerical tests also indicate that the

semismooth Newton method based on the penalized Fischer-Burmeister function usually gives better numerical results than the semismooth Newton method based on the Fischer-Burmeister function, see [3, 29].

The theoretical properties of the semismooth Newton method applied to either of these two Fischer-Burmeister-type reformulations of the complementarity problem can be carried over to the inexact semismooth method where we solve the linearized system (5) only inexactly in the sense that

$$\|H_k d + \Phi(x^k)\| \leq \eta_k \|\Phi(x^k)\| \tag{8}$$

for some forcing sequence $\{\eta_k\}$. Like in the classical case for smooth equations (see [8]), we obtain superlinear convergence if $\eta_k \to 0$ and quadratic convergence if $\eta_k = O(\|\Phi(x^k)\|)$ (under the same assumption as for the exact semismooth Newton method), see [27, 11, 12] for more details.

However, the inexact semismooth method applied to either of these two Fischer-Burmeister-type reformulations does not seem to be as successful as its exact counterpart from a numerical point of view. To illustrate this statement, let us apply both the semismooth Newton method from Algorithm 2.1 and the inexact semismooth Newton method to the MCPLIB test problem collection from [10] (note, however, that we use an updated collection here). For the inexact semismooth method, we used the QMR solver as an iterative method for the linear subproblems (5), however, the results are not significantly different for other iterative solvers like CGS, BiCGSTAB, TFQMR or GMRES(m). For more details on these iterative linear system solvers, the reader is referred to [37].

The methods are applied to all problems from the extended MCPLIB with dimension $n \leq 150$ and to the reformulation (2) based on the penalized Fischer-Burmeister function with parameter $\lambda = 0.95$. Our exact semismooth method is a simple MATLAB implementation of Algorithm 2.1 with no further improvements except that the standard Armijo rule from (S.3) is replaced by a nonmonotone Armijo rule, see [19]. The inexact semismooth method uses the same framework but with (7) being replaced by the rule (8) with the choice

$$\eta_k = \min\left\{10^{-2}/k, \|\Phi(x^k)\|\right\}.$$

The other parameters are taken as

$$\rho = 10^{-8}, \beta = 0.5, \sigma = 10^{-4}, p = 2.1,$$

and we terminate the iteration if

$$k > 300 \quad \text{or} \quad \Psi(x^k) \leq 10^{-12}.$$

The following table presents the number of failures of the exact and inexact semismooth Newton methods using different preconditioners for the inexact solver.

Method	Iterative Solver	Preconditioner	# failures
Exact Semismooth	—	—	7
Inexact Semismooth	QMR	none	14
Inexact Semismooth	QMR	Jacobi	15
Inexact Semismooth	QMR	Gauss-Seidel	14
Inexact Semismooth	QMR	Incomplete LU	6

The failures of the exact semismooth method are on test problems `billups, colvdual, duopoly, ne-hard, pgvon106, scarfbsum, simple-ex`. Some of these failures can be avoided by using some further improvements like a preprocessor, a watchdog strategy or a suitable restart as described in [29]. However, this is not the point we are mainly interested in here.

What is more important is the fact that (most of) the inexact semismooth methods produce significantly more errors than the exact semismooth method. Looking at the iteration process, it turns out that the reason for these additional failures lies in the iterative solver (QMR in this case) which is quite often not able to solve the underlying linear system with the required accuracy. This observation is interesting especially since we are only looking at the small problems with $n \leq 150$ and since we allow a maximum number of inner iterations which is equal to the dimension n of the test problem. In exact arithmetic, this implies that the QMR solver should be able to find the exact solution of the corresponding linear system of equations.

Note that these problems occur even if we use suitable standard preconditioners for the QMR method. One exception, however, is the inexact semismooth method with an incomplete LU preconditioner. Here, it turns out that the incomplete LU preconditioner quite often solves the linear system. If we apply the QMR method with the incomplete LU preconditioner to the larger problems from the MCPLIB, however, also this method is not able to solve many of the linear equations (5), resulting in a couple of failures for the inexact semismooth method where the exact semismooth method has no problems at all.

Motivated by these observations, we therefore take a closer look at the structure of the linear equations (5). For both the Fischer-Burmeister and the penalized Fischer-Burmeister function, the matrix $H_k \in \partial\Phi(x^k)$ is of the form

$$H_k = D_a + D_b F'(x^k) \tag{9}$$

with certain diagonal matrices (depending on the iteration index k)

$$D_a = \operatorname{diag}\big(a_1, \ldots, a_n\big), \qquad D_b = \operatorname{diag}\big(b_1, \ldots, b_n\big)$$

which contain the partial derivatives of the mapping φ with respect to the first variable a and the second variable b, respectively, or a suitable approximation to these partial derivatives at those points where φ is not differentiable. The precise values of D_a, D_b may be found in the corresponding references [6, 3]. Here we only note that both for the Fischer-Burmeister function and the penalized Fischer-Burmeister function, we have

$$a_i \geq 0, \;\; b_i \geq 0, \;\; a_i + b_i > 0 \quad \forall i = 1, \ldots, n, \tag{10}$$

i.e., D_a, D_b are positive semidefinite and their sum $D_a + D_b$ is always positive definite.

This observation may be used in order to show that the matrix H_k from (9) is always nonsingular provided that $F'(x^k)$ is a P_0-matrix, see, for example, [6]. Note that this is true, in particular, if $F'(x^k)$ is positive semidefinite and therefore especially for monotone mappings F. Since a positive semidefinite matrix might be singular, this means that the two diagonal matrices D_a and D_b may be viewed as a kind of regularization, i.e., although the Jacobian $F'(x^k)$ itself might be singular, the matrix H_k from (9) is always nonsingular for certain classes of functions F including the monotone ones.

Furthermore, if $b_i = 0$ for some index $i \in \{1, \ldots, n\}$, we have $a_i > 0$ because of (10). Therefore, it is possible to eliminate the ith row from the linear equation (5) and to obtain a linear system of equations of dimension $n - 1$. Carrying out this elimination process for all i with $b_i = 0$ therefore results in a linear system of equations of the form (5) (with a slightly different right-hand side) whose coefficient matrix has the same structure as in (9) but with the matrix D_b being positive definite. We therefore assume without loss of generality that the matrix $D_b \in \mathbb{R}^{n \times n}$ in (9) is positive definite for the rest of this section.

A major disadvantage of the linear system (5) is, however, that the coefficient matrix H_k from (9) is, in general, nonsymmetric even if the Jacobian $F'(x^k)$ itself is symmetric. For iterative linear system solvers, however, the symmetry of a matrix plays a major role since this allows the application of some reliable methods like CG or MINRES. The situation becomes much more complicated for nonsymmetric systems since then, basically, all iterative solvers based on short-term recurrences have some drawbacks. In fact, unless additional assumptions are satisfied, these methods are no longer guaranteed to find the solution and sometimes even stop with a serious breakdown, see [37].

On the other hand, in the lucky situation where $F'(x^k)$ is symmetric, we can exploit this structure and use our assumption that D_b is positive definite by replacing the linear system

$$H_k d = -\Phi(x^k) \iff \big(D_a + D_b F'(x^k)\big) d = -\Phi(x^k)$$

by the equivalent left-preconditioned system

$$\big(D + F'(x^k)\big) d = -D_b^{-1} \Phi(x^k) \quad \text{with } D := D_b^{-1} D_a, \tag{11}$$

whose coefficient matrix $D + F'(x^k)$ is now again symmetric. In fact, this matrix is symmetric positive definite if $F'(x^k)$ is symmetric positive semidefinite. In general, however, the condition number of this matrix becomes arbitrarily large when x^k converges to a solution x^* of the complementarity problem (1) since one can show that, in this case, some elements of the diagonal matrix D converge to zero whereas other elements are unbounded. This observation does not necessarily exclude the application of iterative methods like CG or MINRES, however, it indicates that one has to expect some difficulties.

A further reformulation of the preconditioned linear system (11) can be obtained if we can factorize the Jacobian $F'(x^k)$ in the form

$$F'(x^k) = LL^T \tag{12}$$

with some matrix $L \in \mathbb{R}^{n\times n}$. Here L is not necessarily the Cholesky factor of $F'(x^k)$. However, the existence of the factorization (12) requires $F'(x^k)$ to be symmetric positive semidefinite and is therefore not always applicable. On the other hand, if we have such a factorization, it is not difficult to see that the linear system (11) is equivalent to the linear least squares problem

$$\min \left\| \begin{pmatrix} L^T \\ D^{1/2} \end{pmatrix} d + \begin{pmatrix} L^{-1} D_b^{-1} \Phi(x) \\ 0 \end{pmatrix} \right\|_2 . \tag{13}$$

In fact, (11) are precisely the normal equations of (13). The standard iterative method for solving this least squares problem is the LSQR method from [30]. Theoretically, LSQR is equivalent to the CG method applied directly to the linear system (11), numerically, however, LSQR applied to the least squares formulation (13) is usually more reliable than the CG method applied to (11).

Such a reformulation will be exploited for the class of optimal control problems in Section 4 where a factorization of the form (12) arises in a very natural way.

3 The Minimum Function

In this section, we investigate a semismooth Newton method for the solution of the complementarity problem (1) when the minimum function

$$\varphi_{\text{MIN}}(a, b) := \min\{a, b\}$$

is used for φ in (3). This minimum function plays a crucial role in many previous papers including [31, 32, 21, 17, 38, 7]. When applied to a linear complementarity problem, where the mapping F is given by $F(x) := Mx + q$ for some constant matrix $M \in \mathbb{R}^{n\times n}$ and a constant vector $q \in \mathbb{R}^n$, it is well-known that the semismooth method applied to the corresponding nonlinear system of equations $\Phi(x) = 0$ has a finite termination property under suitable assumptions, see [21, 17, 38]. Furthermore, local superlinear and quadratic convergence of this method can be obtained under slightly weaker assumptions than for the two Fischer-Burmeister-type approaches discussed in the previous section, see [7].

Hence the minimum function approach has very desirable local convergence properties. A globalization of this approach, however, is more difficult because the corresponding merit function (6) is nondifferentiable in general, although some strategies exist in order to avoid this problem, see, for example [31, 32, 21, 7].

Here our focus will be on the minimum function approach being applied to the linear complementarity problem defined by an M-matrix, i.e., we assume that $F(x) = Mx + q$ and that $M = (m_{ij})$ has the following properties (cf. [24]):

- M is nonsingular with $M^{-1} \geq 0$
- $m_{ij} \leq 0$ for all $i, j = 1, \ldots, n$ with $i \neq j$.

Our notion of an M-matrix coincides with that of a nonsingular M-matrix in [2]. It can also be shown to be equivalent to the class of K-matrices considered in [5].

For a linear complementarity problem defined by an M-matrix, it turns out that a globalization is not necessary and that the method will find the unique solution in a finite number of steps, namely in at most n iterations. This result has recently been proved in the optimal control literature (see [22], where the minimum function approach is called the primal-dual active set method) and does not seem to be well-known in the complementarity field, so that we restate it here within our framework.

To this end, let us first derive the method. The matrix H_k has the same structure as in (9) with certain diagonal matrices

$$D_a := D_a^k = \operatorname{diag}\big(a_1^k, \ldots, a_n^k\big), \qquad D_b := D_b^k = \operatorname{diag}\big(a_1^k, \ldots, a_n^k\big),$$

where the diagonal elements are taken as

$$a_i^k := \begin{cases} 1, & \text{if } i \in \mathcal{A}_k, \\ 0, & \text{if } i \in \mathcal{I}_k \end{cases} \quad \text{and} \quad b_i^k := \begin{cases} 0, & \text{if } i \in \mathcal{A}_k, \\ 1, & \text{if } i \in \mathcal{I}_k \end{cases}$$

with $\mathcal{A}_k$ and $\mathcal{I}_k$ being the index sets

$$\mathcal{A}_k := \big\{ i \in \{1, \ldots, n\} \,\big|\, x_i^k < [Mx^k + q]_i \big\}, \tag{14}$$
$$\mathcal{I}_k := \big\{ i \in \{1, \ldots, n\} \,\big|\, x_i^k \geq [Mx^k + q]_i \big\}. \tag{15}$$

Note that these definitions imply that the corresponding matrix H_k is equal to the Jacobian $F'(x^k)$ at any point x^k where F is differentiable. Furthermore, the diagonal elements a_i^k, b_i^k also satisfy property (10) known from the Fischer-Burmeister-type functions.

Using this notation, the linear system (5), or, equivalently,

$$\big(D_a^k + D_b^k M\big) d = -\min\{x^k, Mx^k + q\},$$

may be written as

$$\begin{pmatrix} I_{\mathcal{A}_k \mathcal{A}_k} & 0 \\ M_{\mathcal{I}_k \mathcal{A}_k} & M_{\mathcal{I}_k \mathcal{I}_k} \end{pmatrix} \begin{pmatrix} d_{\mathcal{A}_k} \\ d_{\mathcal{I}_k} \end{pmatrix} = \begin{pmatrix} -x_{\mathcal{A}_k}^k \\ -[Mx^k + q]_{\mathcal{I}_k} \end{pmatrix}.$$

This is equivalent to the decomposed system

$$\begin{aligned} d_{\mathcal{A}_k} &= -x_{\mathcal{A}_k}^k, \\ M_{\mathcal{I}_k \mathcal{A}_k} d_{\mathcal{A}_k} + M_{\mathcal{I}_k \mathcal{I}_k} d_{\mathcal{I}_k} &= -\big[Mx^k + q\big]_{\mathcal{I}_k} \\ &= -M_{\mathcal{I}_k \mathcal{A}_k} x_{\mathcal{A}_k}^k - M_{\mathcal{I}_k \mathcal{I}_k} x_{\mathcal{I}_k}^k - q_{\mathcal{I}_k}. \end{aligned}$$

Since $M_{\mathcal{I}_k \mathcal{A}_k} d_{\mathcal{A}_k}^k = -M_{\mathcal{I}_k \mathcal{A}_k} x_{\mathcal{A}_k}^k$, the second formula simplifies to some extent, and we obtain the following semismooth Newton-type method for the minimum function approach.

Algorithm 3.1 *(Semismooth Method Based on Minimum Function)*

(S.0) Choose $x^0 \in \mathbb{R}^n, \varepsilon \geq 0$, and set $k := 0$.

(S.1) If $\|\min\{x^k, F(x^k)\}\| \leq \varepsilon$: STOP.

(S.2) Define the index sets $\mathcal{A}_k$ and $\mathcal{I}_k$ as in (14) and (15), respectively, and compute $d^k = \left(d^k_{\mathcal{A}_k}, d^k_{\mathcal{I}_k}\right)$ using the formulas

$$\begin{aligned} d^k_{\mathcal{A}_k} &:= -x^k_{\mathcal{A}_k}, \\ d^k_{\mathcal{I}_k} &:= -x^k_{\mathcal{I}_k} - M^{-1}_{\mathcal{I}_k\mathcal{I}_k} q_{\mathcal{I}_k}. \end{aligned}$$

(S.3) Set $x^{k+1} := x^k + d^k, k \leftarrow k+1$, and go to (S.1).

It should be noted that we have to solve only one linear system with coefficient matrix $M_{\mathcal{I}_k\mathcal{I}_k}$ in order to compute the search direction d^k in (S.2) of Algorithm 3.1, i.e., we have to solve only a linear system of reduced dimension. The reduction might be considerable, i.e., at each iteration k, there might be many components i such that $b^k_i = 0$, whereas for the Fischer-Burmeister-type functions such an event is very unlikely. Moreover, the coefficient matrix of this reduced system is symmetric if M itself is symmetric, so the symmetry does not get destroyed as in the previous section. Furthermore, the condition number of the linear system depends on M only and does not depend on the diagonal matrices D^k_a and D^k_b (except that the submatrix $M_{\mathcal{I}_k\mathcal{I}_k}$ depends on these diagonal matrices). In particular, there is no inherent ill-conditioning when x^k approaches the solution of the complementarity problem. Together with the stronger local properties mentioned in the beginning of this section, it therefore follows that the minimum function approach has several advantages over the Fischer-Burmeister-type approaches described in Section 2.

However, Algorithm 3.1 is, in general, only a local algorithm unless relatively strong assumptions are satisfied. If these assumptions do not hold and one has to globalize the semismooth Newton method, the Fischer-Burmeister-type approaches are usually much better than the semismooth method based on the minimum function.

We now state the main convergence result for Algorithm 3.1 which is essentially the result from [22].

Theorem 3.2 *Let $F(x) = Mx + q$ with $q \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n\times n}$ being an M-matrix. Then Algorithm 3.1 is well-defined and finds the unique solution of the corresponding linear complementarity problem in at most n iterations.*

Proof. First note that an M-matrix is a P-matrix, cf. [24, 2, 5]. Hence the linear complementarity problem has a unique solution (see [5]) and all principal submatrices $M_{\mathcal{I}\mathcal{I}}$ are nonsingular, where $\mathcal{I}$ is an arbitrary subset of $\{1, \ldots, n\}$. This implies that Algorithm 3.1 is well-defined.

To prove finite termination in at most n iterations, we first note that we have

$$x^{k+1}_{\mathcal{A}_k} = x^k_{\mathcal{A}_k} + d^k_{\mathcal{A}_k} = 0 \tag{16}$$

and

$$\begin{aligned}\left[Mx^{k+1}+q\right]_{\mathcal{I}_k} &= M_{\mathcal{I}_k\mathcal{I}_k}x^{k+1}_{\mathcal{I}_k} + M_{\mathcal{I}_k\mathcal{A}_k}\underbrace{x^{k+1}_{\mathcal{A}_k}}_{=0} + q_{\mathcal{I}_k}\\ &= M_{\mathcal{I}_k\mathcal{I}_k}x^{k}_{\mathcal{I}_k} + M_{\mathcal{I}_k\mathcal{I}_k}d^{k}_{\mathcal{I}_k} + q_{\mathcal{I}_k}\\ &= 0.\end{aligned} \tag{17}$$

in view of the updating rules in (S.2) of Algorithm 3.1. In particular, we have

$$x_i^k = 0 \quad \text{or} \quad [Mx^k+q]_i = 0 \tag{18}$$

for each index $i = 1, \ldots, n$ and at each iteration $k \geq 1$. This implies

$$x_i^k \leq 0 \ \ \forall i \in \mathcal{A}_k \quad \text{and} \quad \left[Mx^k+q\right]_i \leq 0 \ \ \forall i \in \mathcal{I}_k \tag{19}$$

at each iteration $k \geq 1$. In fact, if we would have $x_i^k > 0$ for an index $i \in \mathcal{A}_k$, we would obtain $[Mx^k+q]_i > 0$ from the definition of the index set $\mathcal{A}_k$. This, however, contradicts (18) which states that at least one of these two numbers must be zero for each index i. In a similar way, the definition of $\mathcal{I}_k$ implies the second inequality in (19). By an analogous argument, we also get

$$x_i^k \geq 0 \ \ \forall i \in \mathcal{I}_k \quad \text{and} \quad \left[Mx^k+q\right]_i \geq 0 \ \ \forall i \in \mathcal{A}_k \tag{20}$$

at each iteration $k \geq 1$.

We now want to show that the inclusion

$$\mathcal{I}_k \subseteq \mathcal{I}_{k+1} \tag{21}$$

holds for each $k \geq 1$. To this end, we note that M being an M-matrix implies that each submatrix $M_{\mathcal{I}\mathcal{I}}$ is again an M-matrix, see [24]. Hence it follows that

$$\underbrace{M^{-1}_{\mathcal{I}_k\mathcal{I}_k}}_{\geq 0}\underbrace{M_{\mathcal{I}_k\mathcal{A}_k}}_{\leq 0} \leq 0 \tag{22}$$

at each iteration k. Furthermore, (19) and (16) imply

$$x^{k+1}_{\mathcal{A}_k} = 0 \geq x^k_{\mathcal{A}_k}. \tag{23}$$

We therefore obtain

$$\begin{aligned} d^k_{\mathcal{I}_k} &= -x^k_{\mathcal{I}_k} - M^{-1}_{\mathcal{I}_k\mathcal{I}_k}q_{\mathcal{I}_k}\\ &= -M^{-1}_{\mathcal{I}_k\mathcal{I}_k}\left(M_{\mathcal{I}_k\mathcal{I}_k}x^k_{\mathcal{I}_k} + q_{\mathcal{I}_k}\right)\\ &= -\underbrace{M^{-1}_{\mathcal{I}_k\mathcal{I}_k}}_{\geq 0}\Big(\underbrace{\left[Mx^k+q\right]_{\mathcal{I}_k}}_{\leq 0 \text{ by } (19)} - M_{\mathcal{I}_k\mathcal{A}_k}x^k_{\mathcal{A}_k}\Big)\\ &\geq M^{-1}_{\mathcal{I}_k\mathcal{I}_k}M_{\mathcal{I}_k\mathcal{A}_k}x^k_{\mathcal{A}_k}\end{aligned}$$

$$\begin{aligned} &\overset{(16)}{=} \underbrace{-M_{\mathcal{I}_k\mathcal{I}_k}^{-1} M_{\mathcal{I}_k\mathcal{A}_k}}_{\geq 0 \text{ by (22)}} \underbrace{\left(x_{\mathcal{A}_k}^{k+1} - x_{\mathcal{A}_k}^{k}\right)}_{\geq 0 \text{ by (23)}} \\ &\geq \quad 0. \end{aligned}$$

Since $x_{\mathcal{I}_k}^k \geq 0$ for all $k \geq 1$ by (20), it follows that

$$x_{\mathcal{I}_k}^{k+1} = x_{\mathcal{I}_k}^k + d_{\mathcal{I}_k}^k \geq 0 \overset{(17)}{=} \left[Mx^{k+1} + q\right]_{\mathcal{I}_k}.$$

This means that $\mathcal{I}_k \subseteq \mathcal{I}_{k+1}$.

We now prove that the method terminates in at most n iterations. Since $\mathcal{I}_k$ cannot contain more than n elements, it follows from (21) that there is an index $k_0 \in \{1, \ldots, n\}$ such that $\mathcal{I}_{k_0} = \mathcal{I}_{k_0+1}$. Using (20) with $k = k_0 + 1$, we obtain

$$x_{\mathcal{I}_{k_0}}^{k_0+1} = x_{\mathcal{I}_{k_0+1}}^{k_0+1} \geq 0.$$

Similarly, it follows from (20) that

$$\left[Mx^{k_0+1} + q\right]_{\mathcal{A}_{k_0}} = \left[Mx^{k_0+1} + q\right]_{\mathcal{A}_{k_0+1}} \geq 0.$$

Since we also have $x_{\mathcal{A}_{k_0}}^{k_0+1} = 0$ and $\left[Mx^{k_0+1}+q\right]_{\mathcal{I}_{k_0}} = 0$ in view of (16) and (17), respectively, we see that x^{k_0+1} is a solution of the linear complementarity problem. □

Obviously, Theorem 3.2 does not hold when we allow inexact solutions of the linear system in (S.2) of Algorithm 3.1, but it indicates that we can expect almost finite termination if the linear systems are solved very accurately as this is usually done by inexact Newton methods when approaching the solution of the underlying problem.

4 Numerical Results

In this section, we consider two classes of problems, namely the optimal control problem and the obstacle problem. Both classes are infinite-dimensional problems and result into large-scale complementarity problems after a suitable discretization. We describe these problems, give the details of their reformulation as a complementarity problem and present some numerical results obtained with the methods from the previous sections.

4.1 Optimal Control Problems

Let $\Omega \subseteq \mathbb{R}^n$ be a given domain. The class of optimal control problems we consider here is as follows:

$$\begin{array}{ll} \min & J(u) := \frac{1}{2}\|y(u) - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2}\|u - u_d\|_{L^2(\Omega)}^2 \\ \text{s.t.} & u \in \mathcal{F} := \{u \in L^2(\Omega) \,|\, u \leq \psi \text{ a.e. in } \Omega\}, \end{array} \tag{24}$$

where $\alpha > 0$ denotes the regularization parameter, $y = y(u) \in H_0^1(\Omega)$ it the weak solution of

$$-\Delta y = u \quad \text{on } \Omega$$

and

$$y_d, u_d, \psi \in L^2(\Omega)$$

are given functions. Here u denotes the control and y the state variable.

To be more specific, let $\Omega = (0,1) \times (0,1) \subseteq \mathbb{R}^2$ and let A denote the standard five-point finite difference approximation to the negative Laplacian with uniform stepsize $h := 1/(N+1)$ for some $N \in \mathbb{N}$. Then the discretized optimal control problem becomes

$$\min_{u,y} \frac{1}{2}\|y - y_d\|_2^2 + \frac{\alpha}{2}\|u - u_d\|_2^2 \quad \text{s.t.} \quad Ay = u, \;\; \psi - u \geq 0,$$

where the discretized functions u, y etc. are denoted by the same letters as their continuous counterparts.

Using $u = Ay$ in order to remove the control variable, we obtain

$$\min_{y} \frac{1}{2}\|y - y_d\|_2^2 + \frac{\alpha}{2}\|Ay - u_d\|_2^2 \quad \text{s.t.} \quad \psi - Ay \geq 0.$$

Setting $v := \psi - Ay$ then gives

$$\min_{v} \frac{1}{2}\|A^{-1}(\psi - v) - y_d\|_2^2 + \frac{\alpha}{2}\|\psi - v - u_d\|_2^2 \quad \text{s.t.} \quad v \geq 0.$$

Defining $v_d := y_d - A^{-1}\psi$ and $\psi_d := u_d - \psi$, we finally obtain the problem

$$\min_{v} f(v) := \frac{1}{2}\|A^{-1}v + v_d\|_2^2 + \frac{\alpha}{2}\|v + \psi_d\|_2^2 \quad \text{s.t.} \quad v \geq 0$$

which is obviously equivalent to the linear complementarity problem

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

with

$$F(v) := \nabla f(v) := \underbrace{\big(A^{-1}A^{-1} + \alpha I\big)}_{=:M} v + \underbrace{A^{-1}v_d + \alpha\psi_d}_{=:q}.$$

Of course, we do not want to calculate the matrix A^{-1} explicitly. To avoid this, let us take a closer look at the corresponding linear equation (5) resulting from any of the semismooth approaches described in the previous sections.

To this end, let us remove the iteration index k for a while, and let D_a, D_b be the diagonal matrices resulting from the function φ. Without loss of generality, we assume that D_b is positive definite since otherwise we can consider a reduced linear system of similar structure. Then

$$Hd = -\Phi(x) \qquad \Longleftrightarrow \qquad \big(D_a + D_b M\big)d = -\Phi(x)$$

$$
\begin{aligned}
&\iff && \big(D_a + D_b\big(A^{-1}A^{-1} + \alpha I\big)\big)d = -\Phi(x) \\
&\iff && \big(D_b^{-1}D_a + A^{-1}A^{-1} + \alpha I\big)d = -D_b^{-1}\Phi(x) \\
&\overset{D:=D_b^{-1}D_a+\alpha I}{\iff} && \big(A^{-1}A^{-1} + D\big)d = -D_b^{-1}\Phi(x) \\
&\iff && \big(I + DAA\big)\tilde{d} = -D_b^{-1}\Phi(x), \quad d := AA\tilde{d} \\
&\iff && \big(D^{-1} + AA\big)\tilde{d} = -D^{-1}D_b^{-1}\Phi(x), \quad d := AA\tilde{d}.
\end{aligned}
$$

Now we may use the observation from the end of Section 2 in order to see that the last system is equivalent to the linear least squares problem

$$
\min \left\| \begin{pmatrix} A \\ D^{-1/2} \end{pmatrix} \tilde{d} + \begin{pmatrix} A^{-1}D^{-1}D_b^{-1}\Phi(x) \\ 0 \end{pmatrix} \right\|_2, \quad d := AA\tilde{d}
$$

At this stage, we can apply the LSQR method from [30] in order to solve this least squares problem inexactly.

The only other problem is that we have to compute, at each outer iteration, a few vectors of the form $A^{-1}b$ for some right-hand sides b. This means that we have to solve a linear system of equations with the coefficient matrix A. Fortunately, since A corresponds to the five-point finite difference approximation of the negative Laplacian, these systems can be solved, e.g., by a fast sine transform, in only $O(N^2 \log_2 N) \approx O(n \log_2 n)$ arithmetic operations. Altogether, it follows that one outer iteration of a semismooth method applied to the discretized optimal control problem is quite cheap.

Taking into account these considerations, let us apply our method to an example from [1] with the following data:

$$
y_d(x_1, x_2) := \frac{1}{6}\sin(2\pi x_1)\sin(2\pi x_2)\exp(2x_1), \quad u_d \equiv 0, \quad \psi \equiv 0, \quad \alpha := 10^{-2}. \tag{25}
$$

We use different discretizations $N \in \mathbb{N}$. Note that the dimension of the corresponding complementarity problem is $n = N^2$. Table 1 contains the number of iterations needed by our different methods using either the minimum function φ_{MIN} or the Fischer-Burmeister function φ_{FB} or the penalized Fischer-Burmeister function φ_{PFB}. For the minimum function, we also include the results if no globalization is used.

φ	line search	$N = 30$	$N = 60$	$N = 90$	$N = 120$	$N = 150$
φ_{MIN}	yes	5	8	15	28	47
φ_{FB}	yes	8	8	11	15	20
φ_{PFB}	yes	8	9	11	15	20
φ_{MIN}	no	4	5	8	9	12

Table 1: Numerical results for the optimal control example from (25)

Table 2 contains the corresponding numerical results for our second example, also taken

from [1] and given by the data

$$y_d(x_1,x_2) := \begin{cases} 200x_1x_2(x_1-\frac{1}{2})^2(1-x_2), & \text{if } 0 < x_1 \le \frac{1}{2}, \\ 200x_2(x_1-1)(x_1-\frac{1}{2})^2(1-x_2), & \text{if } \frac{1}{2} < x_1 \le 1, \end{cases} \quad u_d \equiv 0,\ \psi \equiv 1,\ \alpha := 10^{-2}. \tag{26}$$

Both tables indicate that our inexact semismooth methods may be applied very successfully to optimal control problems of the form (24). The minimum function approach seems to work better for smaller problems, whereas the number of iterations seems to stay smaller for the two Fischer-Burmeister-type approaches when the dimension is getting larger. In any case, however, the minimum function approach works best if no line search globalization is used. In fact, the number of iterations increases very slowly in this case.

φ	line search	$N = 30$	$N = 60$	$N = 90$	$N = 120$	$N = 150$
φ_{MIN}	yes	5	9	12	19	31
φ_{FB}	yes	8	9	12	15	20
φ_{PFB}	yes	7	9	12	15	21
φ_{MIN}	no	4	6	7	8	11

Table 2: Numerical results for the optimal control example from (26)

The resulting optimal controls and optimal states for the two examples from (25) and (26) are shown in Figures 1 and 2, respectively. These figures clearly show the areas where the upper bound ψ on the control u is active.

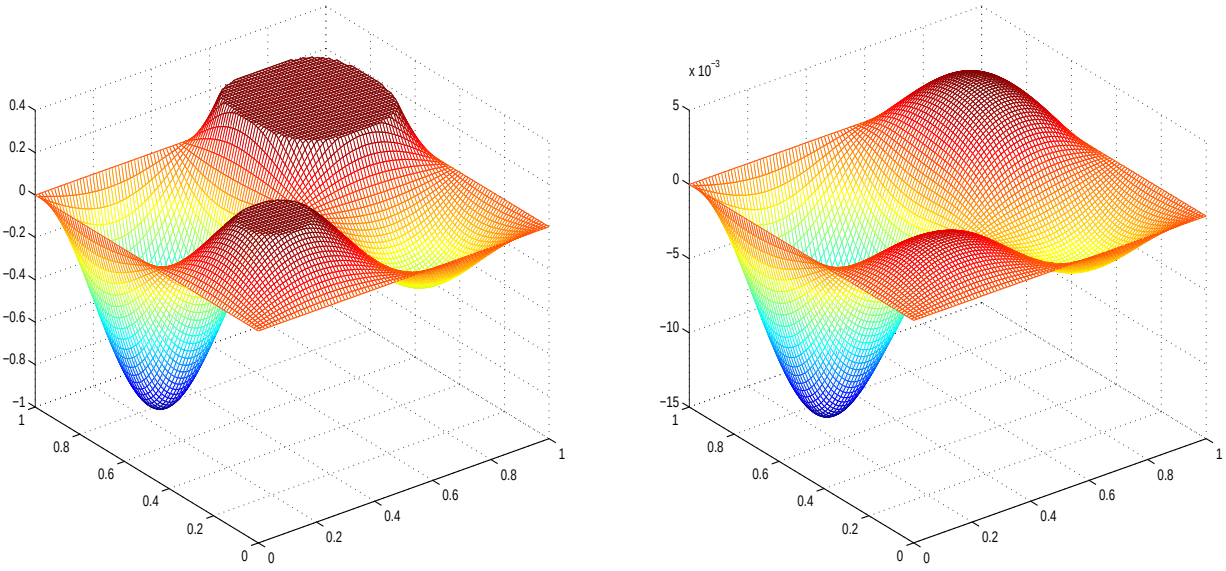

Figure 1: Optimal control (left) and optimal state (right) for the example from (25)

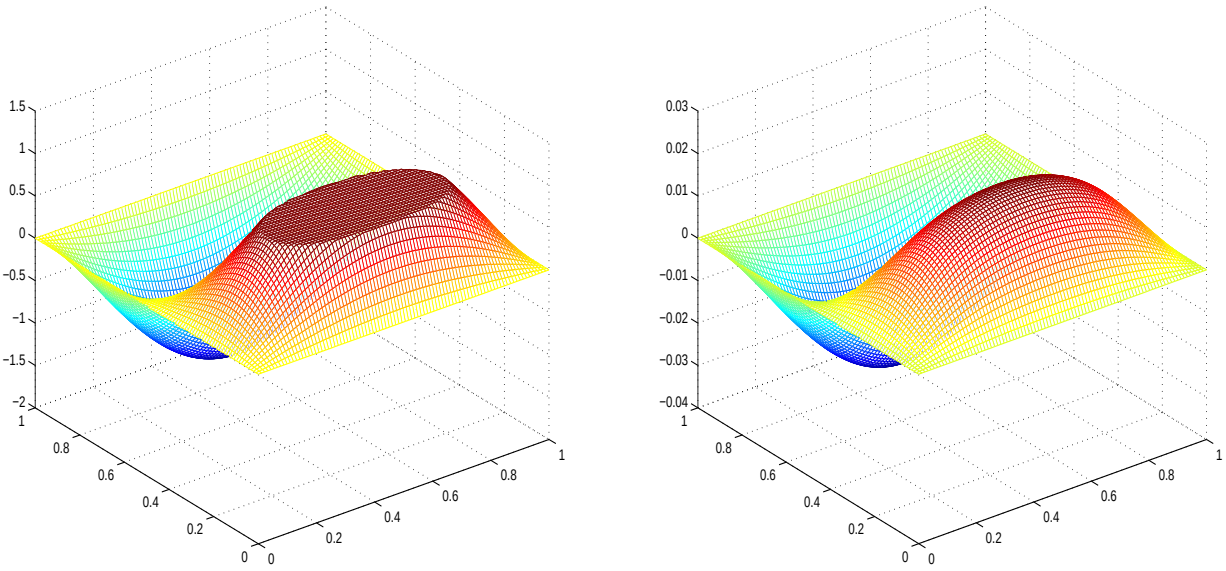

Figure 2: Optimal control (left) and optimal state (right) for the example from (26)

4.2 Obstacle Problems

Let $\Omega \subseteq \mathbb{R}^n$ be a given domain. The obstacle problem (see, e.g., [36]) consists in finding the equilibrium position of an elastic membrane subject to an external force f and an obstacle ψ. Hence the infinite-dimensional problem is to minimize the total energy

$$E(u) := \frac{1}{2}\int_\Omega \|\nabla u\|^2 dx - \int_\Omega f u dx$$

such that $u \in K$, where K denotes the cone

$$K := \big\{ v \in H_0^1(\Omega) \,\big|\, v \geq \psi \text{ a.e. in } \Omega \big\}.$$

The optimality conditions for this optimization problem lead to the variational inequality of finding an element $u \in K$ such that

$$\int_\Omega \nabla u \cdot \nabla(v-u) dx \geq \int_\Omega f(v-u) dx \quad \forall v \in K. \tag{27}$$

Under a weak regularity condition, this is equivalent to the complementarity formulation

$$\begin{array}{rcll} -\Delta u & \geq & f & \text{on } \Omega \\ u & \geq & \psi & \text{on } \Omega \\ \big(-\Delta u - f\big)(u-\psi) & = & 0 & \text{on } \Omega \\ u & \equiv & 0 & \text{on } \partial\Omega. \end{array} \tag{28}$$

In fact, the variational inequality (27) is the weak formulation of this complementarity problem.

N	300	400	500	600	700	800	900	1000
# outer it.	6	6	6	6	6	6	6	6
# inner it. (av.)	470,5	627,3	784,3	942,0	1.133,7	1.297,0	1.415,0	1.624,0

Table 3: Numerical results for the obstacle Bratu problem with data from (29)

Now let us take $n = 2$ and let $\Omega = (0,1) \times (0,1)$ be the standard rectangle. As before, the matrix A denotes the five-point finite difference approximation to the negative Laplace operator on a uniform grid with stepsize $h := 1/(N+1), N \in \mathbb{N}$. Then the discretized form of (28) becomes

$$\begin{aligned} & Au \geq f, u \geq \psi, (Au - f)(u - \psi) = 0 \\ \overset{v:=u-\psi}{\Longleftrightarrow} \quad & Av + A\psi \geq f, v \geq 0, \big(Av + A\psi\big)^T v = 0 \\ \Longleftrightarrow \quad & v \geq 0, F(v) \geq 0, v^T F(v) = 0 \\ & \text{with } F(v) := A(v + \psi) - f. \end{aligned}$$

The obstacle problem we are interested in is the obstacle Bratu problem from [28, 23]. Here f depends on v and is given by

$$f = f(v) := -\lambda e^{-\psi - v}$$

for some parameter $\lambda \geq 0$. In [28, 23], this problem has been investigated as a function of λ. Here we are only interested in the particular instance given by the data

$$\psi \equiv -4, \quad \lambda := 1 \tag{29}$$

which are also used as a complementarity test problem in, for example, [18]. The mapping F is nonlinear and not monotone. Hence the Jacobian $F'(x)$ is not positive semidefinite in general, but is is always symmetric. Since the CG methods works quite well on symmetric matrices, we solve the corresponding linear systems (5) using conjugate gradients. An alternative would be the MINRES method.

We present our numerical results for this example using the penalized Fischer-Burmeister function which seems to give the best results. The minimum function approach does not seem to be a good idea in this case since the mapping F does not have any of the desirable properties usually needed for the minimum function to work well.

The numerical results for this example are given in Table 3 for different dimensions N. There we also present the average number of CG iterations needed to solve the corresponding linear equations (5). Note that this number is very small, although the dimension of the linear systems are pretty large. In particular, for $N = 1000$ we solve a nonlinear and nonmonotone example with one million variables.

5 Final Remarks

In this paper, we have shown that the inexact semismooth Newton method for a nonsmooth reformulation of the complementarity problem may be applied quite successfully to certain large-scale problems. However, the solution of these large-scale problems within the inexact Newton framework depends very much on the availability of a good iterative linear system solver. For symmetric positive definite systems, the CG method is a good choice. For symmetric and possibly indefinite systems, also the CG method or the MINRES method is a suitable candidate. The numerical results from Section 2 indicate, however, that the nonsymmetric iterative linear system solvers do not seem to work quite well in our context. On the other hand, the behaviour of these nonsymmetric solvers can be improved dramatically by using good preconditioners. On the other hand, none of the standard preconditioners seems to do a good job for complementarity problems. Further research is therefore necessary in order to find either a good general purpose preconditioner for complementarity problems or at least good preconditioners for special classes of complementarity problems.

References

[1] M. Bergounioux, M. Haddou, M. Hintermüller and K. Kunisch: *A comparison of a Moreau-Yosida based active set strategy and interior point methods for constrained optimal control problems.* SIAM Journal on Optimization 11, 2001, pp. 495–521.

[2] A. Berman and R.J. Plemmons: *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York, NY, 1979.

[3] B. Chen, X. Chen and C. Kanzow: *A penalized Fischer-Burmeister NCP-function.* Mathematical Programming 88, 2000, pp. 211–216.

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

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

[6] 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.

[7] T. De Luca, F. Facchinei and C. Kanzow: *A theoretical and numerical comparison of some semismooth algorithms for complementarity problems.* Computational Optimization and Applications 16, 2000, pp. 173–205.

[8] R.S. Dembo, S.C. Eisenstat and T. Steihaug: *Inexact Newton methods.* SIAM Journal on Numerical Analysis 19, 1982, pp. 400–408.

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

[11] 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.

[12] F. Facchinei and C. Kanzow: *A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems.* Mathematical Programming 76, 1997, pp. 493–512.

[13] M.C. Ferris and C. Kanzow: *Complementarity and related problems.* In: P.M. Pardalos and M.G.C. Resende (eds.): *Handbook of Applied Optimization.* Oxford University Press, New York, NY, 2002, pp. 514–530.

[14] M.C. Ferris and T.S. Munson: *Interfaces to PATH 3.0: Design, implementation and usage.* Computational Optimization and Applications 12, 1999, pp. 207–227.

[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 and C. Kanzow: *On finite termination of an iterative method for linear complementarity problems.* Mathematical Programming 74, 1996, pp. 279–292.

[18] S.A. Gabriel and J.-S. Pang: *An inexact NE/SQP method for solving the nonlinear complementarity problem.* Computational Optimization and Applications 1, 1992, pp. 67–91.

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

[21] P.T. Harker and J.-S. Pang: *A damped-Newton method for the linear complementarity problem.* In: E.L. Allgower and K. Georg: (eds.): *Computational Solution of Nonlinear Systems of Equations.* Lectures in Applied Mathematics 26, AMS, Providence, RI, 1990.

[22] M. Hintermüller, K. Ito and K. Kunisch: *The primal-dual active set strategy as a semi-smooth Newton method.* SIAM Journal on Optimization, to appear.

[23] R.H.W. Hoppe and H.D. Mittelmann: *A multi-grid continuation strategy for parameter-dependent variational inequalities.* Journal of Computational and Applied Mathematics 26, 1989, pp. 35–46.

[24] R.A. Horn and Ch.R. Johnson: *Topics in Matrix Analysis.* Cambridge University Press, New York, NY, 1991.

[25] H. Jiang, M. Fukushima, L. Qi and D. Sun: *A trust region method for solving generalized complementarity problems.* SIAM Journal on Optimization 8, 1998, pp. 140–157.

[26] H. Jiang and L. Qi: *A new nonsmooth equations approach to nonlinear complementarity problems.* SIAM Journal on Control and Optimization 35, 1997, pp. 178–193.

[27] J.M. Martínez and L. Qi: *Inexact Newton methods for solving nonsmooth equations.* Journal of Computational and Applied Mathematics 60, 1995, pp. 127–145.

[28] E. Miersemann and H.D. Mittelmann: *Continuation for parameterized nonlinear variational inequalities.* Journal of Computational and Applied Mathematics 26, 1989, pp. 23–34.

[29] T.S. Munson, F. Facchinei, M.C. Ferris, A. Fischer and C. Kanzow: *The semismooth algorithm for large scale complementarity problems.* INFORMS Journal on Computing 13, 2001, pp. 294–311.

[30] C.C. Paige and M.A. Saunders: *LSQR: An algorithm for sparse linear equations and sparse least squares.* ACM Transactions on Mathematical Software 8, 1982, pp. 43–71.

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

[32] J.-S. Pang: *A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems.* Mathematical Programming 51, 1991, pp. 101–131.

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

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

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

[36] J.-F. Rodrigues: *Obstacle problems in mathematical physics.* North-Holland, Amsterdam, 1987.

[37] Y. Saad: *Iterative Methods for Sparse Linear Systems.* PWS Publishing Company, Boston, MA, 1996.

[38] D. Sun, J. Han and Y.B. Zhao: *On the finite termination of the damped-Newton algorithm for the linear complementarity problem.* Acta Mathematica Numerica Applicatae 21, 1998, pp. 148–154.

[39] M. Ulbrich: *Nonmonotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems.* SIAM Journal on Optimization 11, 2001, pp. 889–917.

[40] N. Yamashita and M. Fukushima: *Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems.* Mathematical Programming 76, 1997, pp. 469–491.