NUMERICAL RESULTS FOR A GLOBALIZED ACTIVE-SET NEWTON METHOD FOR MIXED COMPLEMENTARITY PROBLEMS*

A. N. Daryina[1], A. F. Izmailov[2] and M. V. Solodov[3]

IMPA Technical Report A282/2004
April 2, 2004

ABSTRACT

In this report we discuss a globalization scheme for a class of active-set Newton methods for solving the mixed complementarity problem (MCP), which was proposed in [3]. Numerical experiments on some test problems are presented, including results on the MCPLIB collection.

Key words. Mixed complementarity problem, semistability, 2-regularity, weak regularity, error bound, Newton method.
AMS subject classifications. 90C30, 65K05.

---

* Research of the first two authors is supported by the Russian Foundation for Basic Research Grant 01-01-00810 and the Council for the State Support of Leading Scientific Schools Grant 00-15-96141. The third author is supported in part by CNPq Grants 300734/95-6(RN) and 471780/2003-0, and by FAPERJ.

[1] Russian Peoples Friendship University, Miklukho-Maklaya Str. 6, 117198 Moscow, Russia.
Email: anna.daryina@mail.ru

[2] Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Operations Research, Leninskiye Gori, GSP-2, 119992 Moscow, Russia.
Email: izmaf@ccas.ru

[3] Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil.
Email: solodov@impa.br

1 Introduction

The *mixed complementarity problem* (MCP) [6] is the variational inequality on a generalized box, that is

$$\text{find } x \in B \text{ such that } \langle F(x), y - x\rangle \geq 0 \text{ for all } y \in B, \tag{1.1}$$

where $F : \mathbf{R}^n \to \mathbf{R}^n$ and

$$B = \{x \in \mathbf{R}^n \mid l_i \leq x_i \leq u_i,\ i = 1, \ldots, n\},$$

$l_i \in \mathbf{R} \cup \{-\infty\}$, $u_i \in \mathbf{R} \cup \{+\infty\}$, $l_i < u_i$ for all $i = 1, \ldots, n$. Equivalently, it can be stated as

$$\text{find } x \in B \text{ such that } F_i(x) \begin{cases} \geq 0, & \text{if } x_i = l_i, \\ = 0, & \text{if } x_i \in (l_i, u_i), \\ \leq 0, & \text{if } x_i = u_i, \end{cases} \quad i = 1, \ldots, n.$$

As is well known, many important problems can be cast in the format of MCP [8, 6]. As a special case of MCP, we mention the nonlinear complementarity problem (NCP), which corresponds to setting $l_i = 0$, $u_i = +\infty$, $i = 1, \ldots, n$. The systems of nonlinear equations are obtained by choosing $l_i = -\infty$, $u_i = +\infty$, $i = 1, \ldots, n$. Another important example is the primal-dual Karush-Kuhn-Tucker (KKT) optimality system: find $z \in \mathbf{R}^p$ and $\mu \in \mathbf{R}^m$ such that

$$\begin{gathered} g(z) - (G'(z))^T \mu = 0, \\ \mu \geq 0, \quad G(z) \geq 0, \quad \langle \mu, G(z)\rangle = 0, \end{gathered} \tag{1.2}$$

where $g : \mathbf{R}^p \to \mathbf{R}^p$ and $G : \mathbf{R}^p \to \mathbf{R}^m$. The KKT system (1.2) can be written as an MCP if we set $n = p + m$ and

$$F(x) = \begin{pmatrix} g(z) - (G'(z))^T \mu \\ G(z) \end{pmatrix}, \quad x = (z, \mu) \in \mathbf{R}^p \times \mathbf{R}^m,$$

$l_i = -\infty$, $i = 1, \ldots, p$, $l_i = 0$, $i = p + 1, \ldots, n$, $u_i = +\infty$, $i = 1, \ldots, n$. Under well-known assumptions, (1.2) represents the first-order primal-dual necessary conditions characterizing solutions in variational inequality or inequality-constrained optimization problems.

This paper describes a globalization scheme for a local Newton-type method proposed in [3]. Also numerical experiments will be reported on problems from the MCPLIB collection [4], together with results on some additional test examples.

2 The Local Active-Set Method

The algorithm of [3] belongs to the class of *active-set* methods. In the context of MCP, this corresponds to identifying the sets of indices

$$\begin{gathered} A = A(\bar{x}) = \{i = 1, \ldots, n \mid F_i(\bar{x}) = 0\}, \\ N = N(\bar{x}) = \{i = 1, \ldots, n \mid F_i(\bar{x}) \neq 0\}, \\ N_l = N_l(\bar{x}) = \{i \in N \mid \bar{x}_i = l_i\}, \\ N_u = N_u(\bar{x}) = \{i \in N \mid \bar{x}_i = u_i\}, \end{gathered}$$

where $\bar{x}$ is some solution of MCP. If the specified sets can be correctly identified using information available at a point x close enough to the solution $\bar{x}$, then locally MCP can be reduced to a system of nonlinear equations (which is structurally a much simpler problem to solve). In the sequel, we shall also use the following partitioning of the set of active indices:

$$\begin{gathered} A_0 = A_0(\bar{x}) = \{i \in A \mid \bar{x}_i = l_i \text{ or } \bar{x}_i = u_i\}, \\ A_+ = A_+(\bar{x}) = \{i \in A \mid \bar{x}_i \in (l_i, u_i)\}, \\ A_{0l} = A_{0l}(\bar{x}) = \{i \in A_0 \mid \bar{x}_i = l_i\}, \\ A_{0u} = A_{0u}(\bar{x}) = \{i \in A_0 \mid \bar{x}_i = u_i\}. \end{gathered}$$

The analog of the strict complementarity condition in NCP (or KKT) corresponds, in the setting of MCP, to saying that $A_0 = \emptyset$. Under this assumption, locally MCP trivially reduces to a system of nonlinear equations, which simplifies the local structure of MCP significantly. The condition of strict complementarity, however, is restrictive and is not assumed.

Let $\psi : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ be a *complementarity function*, i.e., a function such that

$$\psi(a,\, b) = 0 \quad \Leftrightarrow \quad a \geq 0,\ b \geq 0,\ ab = 0.$$

Assume that ψ satisfies the following additional assumptions:

$$a > 0,\ b < 0 \quad \Rightarrow \quad \psi(a,\, b) < 0,$$

$$a > 0,\ b > 0 \quad \Rightarrow \quad \psi(a,\, b) > 0.$$

Then MCP can be equivalently reformulated as a system of nonlinear equations

$$\Psi(x) = 0, \tag{2.1}$$

where

$$\Psi : \mathbf{R}^n \to \mathbf{R}^n, \quad \Psi_i(x) = \begin{cases} F_i(x), & \text{if } i \in I_F, \\ \psi(x_i - l_i, F_i(x)), & \text{if } i \in I_l, \\ -\psi(u_i - x_i, -F_i(x)), & \text{if } i \in I_u, \\ \psi(x_i - l_i, -\psi(u_i - x_i, -F_i(x))), & \text{if } i \in I_{lu}, \end{cases}$$

$$\begin{gathered} I_F = \{i = 1, \ldots, n \mid -\infty = l_i,\ u_i = +\infty\}, \\ I_l = \{i = 1, \ldots, n \mid -\infty < l_i,\ u_i = +\infty\}, \\ I_u = \{i = 1, \ldots, n \mid -\infty = l_i,\ u_i < +\infty\}, \\ I_{lu} = \{i = 1, \ldots, n \mid -\infty < l_i,\ u_i < +\infty\}. \end{gathered}$$

For some special choices of complementarity functions, this reformulation is well-known (see, e.g., [1, 7]).

Complementarity functions to be used in the sequel are the natural residual $\psi_{NR}(a, b) = \min\{a, b\}$, the Fischer-Burmeister function $\psi_{FB}(a, b) = a + b - \sqrt{a^2 + b^2}$, and $\psi_S(a, b) = 2ab - (\min\{0, a + b\})^2$ (where S stands for "smooth"). The corresponding reformulations of MCP would be denoted by Ψ_{NR}, Ψ_{FB} and Ψ_S, respectively.

Definition 2.1 Let $\Psi : \mathbf{R}^n \to \mathbf{R}^n$ be differentiable in a neighbourhood of $\bar{x} \in \mathbf{R}^n$ and $\Psi' : \mathbf{R}^n \to \mathbf{R}(n,\, n)$ be directionally differentiable at $\bar{x}$. Then Ψ is 2-*regular* at $\bar{x}$ if

$$T = \{0\},$$

where

$$\begin{aligned} T = T(\bar{x}) &= \{\xi \in \ker \Psi'(\bar{x}) \mid (\Psi')'(\bar{x};\, \xi)\xi \in \operatorname{im} \Psi'(\bar{x})\} \\ &= \{\xi \in \ker \Psi'(\bar{x}) \mid P(\Psi')'(\bar{x};\, \xi)\xi = 0\}, \end{aligned} \tag{2.2}$$

with P being the orthogonal projector onto $(\operatorname{im} \Psi'(\bar{x}))^{\perp}$.

The above is a special case of 2-regularity of a nonlinear mapping [10, 9], corresponding to the case when the mapping acts from some space into itself.

Theorem 2.1 *[3] Let $F : \mathbf{R}^n \to \mathbf{R}^n$ be sufficiently smooth near a point $\bar{x} \in \mathbf{R}^n$, which is a solution of MCP.*

The mapping Ψ_S is 2-regular at $\bar{x}$ if, and only if, there exist a neighborhood U of $\bar{x}$ and a constant $M > 0$ such that

$$\|x - \bar{x}\| \le M(\|(E - P)\Psi_S(x)\| + \|P\Psi_S(x)\|^{1/2}) \quad \forall\, x \in U. \tag{2.3}$$

Adjusting M and U, if necessary, the error bound (2.3) can be simplified into the following relation (less accurate, but possibly easier to use):

$$\|x - \bar{x}\| \le M\|\Psi_S(x)\|^{1/2} \quad \forall\, x \in U. \tag{2.4}$$

The assumption of 2-regularity above is very mild. In particular, it is weaker than various alternatives, such as *semistability* [2] of the MCP solution.

Proposition 2.1 *[3] Semistability of a solution $\bar{x}$ of MCP implies 2-regularity of Ψ_S at $\bar{x}$ (equivalently, error bound* (2.3)*), but not vice versa.*

As a consequence, the error bound given above can hold even when MCP reformulations based on Ψ_{NR} and Ψ_{FB} fail to provide a bound.

The following technique for identifying the relevant index sets is based on the ideas of [5], see also [6, Ch. 6.7]. Define the *identification function*

$$\rho : \mathbf{R}_+ \to \mathbf{R}, \quad \rho(t) = \begin{cases} \bar{\rho}, & \text{if } t \ge \bar{t}, \\ -1/\log t, & \text{if } t \in (0,\, \bar{t}), \\ 0, & \text{if } t = 0, \end{cases}$$

where $\bar{t} \in (0,\, 1)$ and $\bar{\rho} > 0$ are fixed numbers (the choice of $\bar{t}$ and $\bar{\rho}$ does not affect theoretical analysis; in our numerical experiments reported in Section 3, we use $\bar{t} = 0.9$ and $\bar{\rho} = -1/\log \bar{t}$, as suggested in [5]). For any $x \in \mathbf{R}^n$, define further the index sets

$$A(x) = \{i = 1, \ldots, n \mid |F_i(x)| \le \rho(\|\Psi_S(x)\|)\}, \tag{2.5}$$

$$N(x) = \{1, \ldots, n\} \setminus A(x), \tag{2.6}$$

$$N_l(x) = \{i \in N(x) \mid x_i - l_i \leq u_i - x_i\}, \quad N_u(x) = N(x) \setminus N_l(x), \tag{2.7}$$

$$A_0(x) = \{i \in A(x) \mid \min\{|x_i - l_i|, |u_i - x_i|\} \leq \rho(\|\Psi_S(x)\|)\}, \quad A_+(x) = A(x) \setminus A_0(x), \tag{2.8}$$

$$A_{0l}(x) = \{i \in A_0(x) \mid x_i - l_i \leq u_i - x_i\}, \quad A_{0u}(x) = A_0(x) \setminus A_{0l}(x). \tag{2.9}$$

Proposition 2.2 *If Ψ_S is 2-regular at a solution $\bar{x}$ of MCP (equivalently, the error bound* (2.3) *holds), then for any $x \in \mathbf{R}^n$ sufficiently close to $\bar{x}$, it holds that*

$$A(x) = A, \quad N(x) = N, \quad N_l(x) = N_l, \quad N_u(x) = N_u, \tag{2.10}$$

$$A_{0l}(x) = A_{0l}, \quad A_{0u}(x) = A_{0u}, \quad A_0(x) = A_0, \quad A_+(x) = A_+. \tag{2.11}$$

Observe that in the implementation of the identification procedure, the following obvious relations can also be taken into account: $I_F \subset A_+$, $I_l \subset (A_{0l} \cup A_+ \cup N_l)$ and $I_u \subset (A_{0u} \cup A_+ \cup N_u)$.

Once the index sets are identified, we have the following relations which are guaranteed to be satisfied at a solution $\bar{x}$ of MCP:

$$F_A(x) = 0, \quad x_{A_{0l} \cup N_l} = l_{A_{0l} \cup N_l}, \quad x_{A_{0u} \cup N_u} = u_{A_{0u} \cup N_u}.$$

For simplicity of notation, suppose that the components of $x \in \mathbf{R}^n$ are ordered in such a way that $x = (x_{A_+}, x_{A_{0l} \cup N_l}, x_{A_{0u} \cup N_u})$. Then MCP locally reduces to the following system of nonlinear equations:

$$F_A(x_{A_+}, l_{A_{0l} \cup N_l}, u_{A_{0u} \cup N_u}) = 0. \tag{2.12}$$

Observe that in the absence of strict complementarity (when $A_0 \neq \emptyset$, i.e., $|A| > |A_+|$), the system is over-determined (the number of equations is larger than the number of unknowns). This opens up a number of options. Of course, one can just solve the system by the Gauss–Newton method (GNM). This possibility will be considered. However, we prefer not to limit ourselves to GNM for the following reason: the Gauss–Newton approach can destroy structure present in F_A (for example, sparsity or the primal-dual structure in the case of KKT).

Our proposal is to consider the following system of nonlinear equations:

$$\Phi_C(x_{A_+}) = 0, \tag{2.13}$$

where

$$\Phi_C : \mathbf{R}^{|A_+|} \to \mathbf{R}^{|A_+|}, \quad \Phi_C(x_{A_+}) = C(x_{A_+}) F_A(x_{A_+}, l_{A_{0l} \cup N_l}, u_{A_{0u} \cup N_u}),$$

with $C : \mathbf{R}^{|A_+|} \to \mathbf{R}(|A_+|, A)$ being a smooth mapping (possibly constant). Clearly, $\bar{x}_{A_+}$ is a solution of (2.13) for any choice of C. The Jacobian of (2.13) at this solution is given by

$$\Phi'_C(\bar{x}_{A_+}) = C(\bar{x}_{A_+}) \frac{\partial F_A}{\partial x_{A_+}}(\bar{x}), \tag{2.14}$$

where we have taken into account that $F_A(\bar{x}) = 0$. Thus $\bar{x}_{A_+}$ can be found by applying Newton-type methods to (2.13) whenever the matrix in (2.14) is nonsingular.

Note that GNM for (2.12) would essentially correspond to choosing in (2.13)

$$C(x_{A_+}) = \left(\frac{\partial F_A}{\partial x_{A_+}}(x_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u})\right)^T, \tag{2.15}$$

and applying to the resulting system an approximate version of the pure Newton method. Indeed, with the notation of (2.15), the Gauss–Newton iteration for (2.12) has the form

$$x^{k+1}_{A_+} = x^k_{A_+} - \left(C(x^k_{A_+})\frac{\partial F_A}{\partial x_{A_+}}(x^k_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u})\right)^{-1} \Phi_C(x^k_{A_+}). \tag{2.16}$$

Observe that the above formula is just an approximation of the standard Newton iteration for (2.13), where the Jacobian $\Phi'_C(x^k_{A_+})$ is replaced by $C(x^k_{A_+})\frac{\partial F_A}{\partial x_{A_+}}(x^k_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u})$. Due to (2.14), this change preserves the superlinear convergence of the pure Newton iteration for (2.13). Note finally that with the choice of (2.15), we have

$$\Phi'_C(\bar{x}_{A_+}) = \left(\frac{\partial F_A}{\partial x_{A_+}}(\bar{x})\right)^T \frac{\partial F_A}{\partial x_{A_+}}(\bar{x}). \tag{2.17}$$

This immediately motivates the following definition.

Definition 2.2 A solution $\bar{x}$ of MCP is referred to as *weakly regular* if

$$\operatorname{rank}\frac{\partial F_A}{\partial x_{A_+}}(\bar{x}) = |A_+|.$$

Weak regularity is implied by semistability, but not vice versa. Moreover, 2-regularity of Ψ_S at $\bar{x}$ and weak regularity, when combined, are still a weaker condition than semistability.

Proposition 2.3 *[3] Let $\bar{x}$ be a solution of MCP. Then semistability of $\bar{x}$ implies weak regularity of $\bar{x}$, but not vice versa.*

Proposition 2.4 *[3] Let $\bar{x}$ be a solution of MCP. Then semistability of $\bar{x}$ implies the combination of 2-regularity of Ψ_S at $\bar{x}$ and weak regularity of $\bar{x}$, but not vice versa.*

We have thus a local algorithm with superlinear convergence under assumptions weaker than semistability of the MCP solution. Specifically, we have the following.

Theorem 2.2 *Let $F : \mathbf{R}^n \to \mathbf{R}^n$ be sufficiently smooth near a point $\bar{x} \in \mathbf{R}^n$, which is a solution of MCP. Suppose that this solution is weakly regular and Ψ_S is 2-regular at $\bar{x}$.*

For any $x^0 \in \mathbf{R}^n$ sufficiently close to $\bar{x}$, if the index sets $A = A(x^0)$, $A_+ = A_+(x^0)$, $A_{0l} = A_{0l}(x^0)$, $A_{0u} = A_{0u}(x^0)$, $N_l = N_l(x^0)$ and $N_u = N_u(x^0)$ are defined according to (2.5)-(2.9)*, then GNM applied to the system* (2.12) *(with $x^0_{A_+}$ as a starting point) is well-defined and superlinearly convergent to $\bar{x}_{A_+}$.*

As already mentioned above, it sometimes can be useful to choose the mapping C differently from the Gauss–Newton option of (2.15). We might want to take $C(\cdot) = C \in \mathbf{R}(|A_+|, |A|)$, a fixed matrix, in order to preserve in the matrix $C\frac{\partial F_A}{\partial x_{A_+}}(x_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u})$ the structure (primal-dual, sparsity, etc.) of the matrix $\frac{\partial F_A}{\partial x_{A_+}}(x_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u})$. This motivates the following considerations.

Proposition 2.5 *Suppose that a solution $\bar{x}$ of MCP is weakly regular.*

Then the set of matrices $C \in \mathbf{R}(|A_+|, |A|)$ such that $\Phi'_C(\bar{x}_{A_+})$ is nonsingular is open and dense in $\mathbf{R}(|A_+|, |A|)$.

Proposition 2.5 justifies choosing C in any desirable way, as the chance that the resulting system would be degenerate is negligible (the set of matrices for which this would happen is of the Lebesgue measure zero). Of course, one should make reasonable choices. For example, it should hold that $\operatorname{rank} C = |A_+|$.

3 Globalization Issues and Numerical Experiments

In this section we report on our numerical experience based on the MCPLIB test problems collection (the newer version of [4]), and on some additional small examples, designed to highlight the case where various standard regularity conditions do not hold, and thus SNM-based methods may have trouble or converge slowly. This is precisely the case where the switch to our local algorithm can be particularly useful.

To perform numerical experiments, we had to implement our method as a final stage of some globally convergent scheme. It seems difficult to suggest a globally convergent scheme directly related to the structure of our local algorithm. In some sense, this is a disadvantage. But on the other hand, our local approach can be combined with *any* globally convergent algorithm satisfying some requirements (see below). In fact, the way we suggest to use our active-set method is precisely for improving the local convergence properties of standard algorithms (say, when they run into difficulties because of the lack of regularity of a solution). Our numerical experiments, reported below, indicate that the resulting strategy fulfills the stated objective.

We now describe the basic *hybrid globalization scheme* and its convergence properties. The scheme employs the linear decrease test for some (computable) merit function $\varphi : \mathbf{R}^n \to \mathbf{R}_+$. Possible choices of a global algorithm and φ are restricted by the following assumptions:

(A1) $\varphi(x) = 0$ if, and only if, $x \in \mathbf{R}^n$ is a solution of MCP.

(A2) φ decreases monotonically along the trajectories of the global algorithm.

(A3) φ decreases superlinearly along the trajectories of our local method near "qualified" solutions, i.e., solutions satisfying the assumptions of the local convergence theorem for our method (Theorem 2.2).

Our choice of a global algorithm and a merit function, as well as the full method, would be stated later. We first give a simplified general discussion.

At each iteration k of the global algorithm, we first identify the relevant index sets. If there is a reason to believe that the identification is correct (e.g., if it did not change when compared to the previous iteration), then we compute the trial point $\tilde{x}^{k+1}_{A_+}$ by the step of GNM applied to (2.12) at $x^k_{A_+}$, and set the remaining components of $\tilde{x}^{k+1} \in \mathbf{R}^n$ equal to the corresponding components of $l_{A_{0l} \cup N_l}$ and $u_{A_{0u} \cup N_u}$. If $\tilde{x}^{k+1}$ is well-defined and

$$\varphi(\tilde{x}^{k+1}) \leq q\varphi(x^k), \quad q \in (0,\, 1), \tag{3.1}$$

then we set $x^{k+1} = \tilde{x}^{k+1}$, and proceed with the next iteration. Otherwise, we compute x^{k+1} by the step of the global algorithm, and proceed with the next iteration.

If the linear decrease test (3.1) is satisfied for a finite number of iterations only, the hybrid globalization scheme behaves essentially as the global algorithm, and the global convergence properties of the latter remain valid. Otherwise, taking into account assumption (A2), we conclude that

$$\varphi(x^{k+1}) \to 0 \text{ as } k \to \infty,$$

and in particular, according to assumption (A1), each accumulation point of the sequence $\{x^k\}$ is a solution of MCP. This characterizes the global convergence properties of the hybrid scheme.

Suppose now that a "qualified" solution $\bar{x}$ of MCP is an accumulation point of $\{x^k\}$. Then, by assumption (A3), the linear decrease test (3.1) is satisfied for all iteration indices k sufficiently large. Furthermore, it is easy to see that the hybrid scheme eventually switches to our local algorithm (with the "starting point" sufficiently close to $\bar{x}$). Hence, by Theorem 2.2, the entire sequence $\{x^k\}$ converges to $\bar{x}$ superlinearly.

In our numerical experiments we adopt the following choices. As a global algorithm, we use the linesearch-SNM for (2.1) with $\Psi = \Psi_{FB}$. More precisely, we essentially follow the implementation suggested in [12] ("General Line Search Algorithm"), with all the parameter values adopted there. Furthermore, we take $\varphi = \varphi_{FB}$, where

$$\varphi_{FB} : \mathbf{R}^n \to \mathbf{R}, \quad \varphi_{FB}(x) = \frac{1}{2}\|\Psi_{FB}(x)\|^2.$$

We emphasize that this is just one example of the various appropriate choices. But this particular algorithm seems reasonable for our purposes. Linesearch SNM is known to be quite efficient, and at the same time, it is easy to implement in its basic form. We note that we do not use any enhancements, such as crashing and nonmonotone linesearch (see, e.g., [13]). The reason is that these are intended to improve global behavior of the algorithm, while we are concerned with local behavior. Thus a simple implementation of the global scheme is sufficient for our purposes, as our principal conclusions refer to the local convergence properties.

Assumptions (A1) and (A2) are evidently satisfied for the adopted choices. In order to guarantee (A3), we need to assume the error bound

$$\|x - \bar{x}\| = O(\|\Psi_{FB}(x)\|). \tag{3.2}$$

Recall that this bound is equivalent to semistability. Under this assumption, for $x^0 \in \mathbf{R}^n$ close enough to a "qualified" solution $\bar{x}$, and the corresponding trajectory $\{x^k\}$ of the local

method, we have

$$\begin{aligned}\varphi_{FB}(x^{k+1}) &= \frac{1}{2}\|\Psi_{FB}(x^{k+1}) - \Psi_{FB}(\bar{x})\|^2 \\ &= O(\|x^{k+1} - \bar{x}\|^2) \\ &= o(\|x^k - \bar{x}\|^2) \\ &= o(\|\Psi_{FB}(x^k)\|^2) \\ &= o(\varphi_{FB}(x^k)),\end{aligned}$$

where the second equality is by the Lipschitz-continuity of Ψ_{FB} near $\bar{x}$, the third is by Theorem 2.2, and the fourth is by (3.2). Thus, assumption (A3) is satisfied.

We proceed with the formal statement of the algorithm. Variable "Alg" below is used to select between the two variants of the algorithm that would be compared to each other. Alg = 1 means that the possibility of switching to our step is blocked, and so the algorithm works as the usual linesearch SNM/FB from [12]. Alg = 2 corresponds to the proposed active-set strategy. Note however that switching to an active-set step is forbidden on the first iteration, and also in the case when the identified index sets differ from the corresponding index sets at the previous iteration. This is done in order to prevent the algorithm from switching to the active-set strategy too early, when the sets are not yet stabilized and are likely to give incorrect identification.

Algorithm 3.1 `Preliminary step.` Set Alg = 1 or 2. Fix q, ε, $\tau \in (0, 1)$, $\delta, \gamma > 0$. Set $k = 0$ and choose $x^0 \in \mathbf{R}^n$.

`Initialization step.` If Alg = 1 or $k = 0$, go to `SNM/FB step`. Otherwise define the index sets $A = A(x^k)$, $N_l = N_l(x^k)$, $N_u = N_u(x^k)$, $A_+ = A_+(x^k)$, $A_{0l} = A_{0l}(x^k)$, $A_{0u} = A_{0u}(x^k)$ according to (2.5)–(2.9). If at least one of these sets does not coincide with its counterpart computed at the previous iteration, go to `SNM/FB step`.

`GNM/active-set step.` Compute $\tilde{x}^{k+1} \in \mathbf{R}^n$ as follows:

$$\tilde{x}^{k+1}_{A_+} = x^k_{A_+} - \left(\left(\frac{\partial F_A}{\partial x_{A_+}}(x^k)\right)^{\mathrm{T}} \frac{\partial F_A}{\partial x_{A_+}}(x^k)\right)^{-1} \frac{\partial F_A}{\partial x_{A_+}}(x^k) F_A(x^k_{A_+}, l_{A_{0l}\cup N_l}, u_{A_{0u}\cup N_u}), \quad (3.3)$$

$$\tilde{x}^{k+1}_{A_{0l}\cup N_l} = l_{A_{0l}\cup N_l}, \quad \tilde{x}^{k+1}_{A_{0u}\cup N_u} = u_{A_{0u}\cup N_u}.$$

If this point is well-defined and

$$\varphi_{FB}(\tilde{x}^{k+1}) \le q\varphi_{FB}(x^k), \quad (3.4)$$

set $x^{k+1} = \tilde{x}^{k+1}$, adjust k by 1, and go to `Initialization step`.

`SNM/FB step.` Compute $\Lambda_k \in \partial_B \Psi_{FB}(x^k)$ and

$$\tilde{x}^{k+1} = x^k - \Lambda_k^{-1} \Psi_{FB}(x^k).$$

If this point is well-defined and (3.4) holds, set $x^{k+1} = \tilde{x}^{k+1}$, adjust k by 1, and go to `Initialization step`.

If $\tilde{x}^{k+1}$ is well-defined but (3.4) does not hold, set $\tilde{d}^k = \tilde{x}^{k+1} - x^k$. If

$$\langle \varphi'_{FB}(x^k),\, \tilde{d}^k \rangle \le -\gamma \|\tilde{d}^k\|^\delta,$$

set $d^k = \tilde{d}^k$ and go to `Linesearch step`.

`Gradient step`. Set $d^k = -\varphi'_{FB}(x^k)$.

`Linesearch step`. Compute the stepsize parameter α_k according to the Armijo rule: $\alpha_k = \tau^s$, where s is the smallest nonnegative integer satisfying

$$\varphi_{FB}(x^k + \tau^s d^k) \le \varphi_{FB}(x^k) + \varepsilon \tau^s \langle \varphi'_{FB}(x^k),\, d^k \rangle.$$

Set $x^{k+1} = x^k + \alpha_k d^k$, adjust k by 1, and go to `Initialization step`.

Note that (3.3) differs slightly from the pure Gauss–Newton iteration given by (2.16) with $C(x_{A_+})$ defined in (2.15). The modification is made in order to reduce the number of evaluations of the Jacobian of F: the described algorithm requires exactly one such evaluation per iteration, whether Alg = 1 or 2. This modification does not affect the rate of convergence. To see this, note that (2.16) can be written in the form

$$x^{k+1}_{A_+} = x^k_{A_+} - (C(x^k_{A_+})(C(x^k_{A_+}))^{\mathrm{T}})^{-1} C(x^k_{A_+}) F_A(x^k_{A_+},\, l_{A_{0l} \cup N_l},\, u_{A_{0u} \cup N_u}).$$

Comparing this iteration with (3.3), observe that

$$\begin{aligned} \left\| \frac{\partial F_A}{\partial x_{A_+}}(x^k) - (C(x^k_{A_+}))^{\mathrm{T}} \right\| &= \left\| \frac{\partial F_A}{\partial x_{A_+}}(x^k) - \frac{\partial F_A}{\partial x_{A_+}}(x^k_{A_+},\, l_{A_{0l} \cup N_l},\, u_{A_{0u} \cup N_u}) \right\| \\ &= O(\|x^k_{A_{0l} \cup N_l} - l_{A_{0l} \cup N_l}\| + \|x^k_{A_{0u} \cup N_u} - u_{A_{0u} \cup N_u}\|) \\ &= O(\|x^k - \bar{x}\|), \end{aligned}$$

provided the index sets are correctly identified for a given solution $\bar{x}$.

To compute $\Lambda_k \in \partial_B \Psi_{FB}(x^k)$ at SNM/FB step, we used the procedure suggested in [1] (with $z_i = 1\ \forall\, i = 1, \ldots, n$, in the notation of [1], and the "computer zero" parameter set to 10^{-10}). The well-known formula $\varphi'_{FB}(x^k) = \Lambda_k^{\mathrm{T}} \Psi_{FB}(x^k)$ is used to compute the gradient of φ_{FB}.

In the numerical experiments reported below, we used the following set of parameters: $q = 0.9$, $\varepsilon = 10^{-4}$, $\tau = 0.5$, $\delta = 2.1$, $\gamma = 10^{-9}$. The stopping criterion is

$$\|\Psi_{FB}(x^k)\| < 10^{-9}. \tag{3.5}$$

The cases when an algorithm did not terminate according to this criterion after 500 iterations are referred to as failures.

The algorithm was implemented in Matlab, making use of the standard option for treating sparse matrices. The Table 1 below reports first the total number of iterations for the two algorithms, and then the number of times the GNM/active-set step has been used. Next to the latter number, in the brackets, stands the number of active-set iterations at the "tail" of the process (right before convergence had been declared). Next, the total number of SNM/FB steps of the globalization scheme is reported for the two algorithms. In the brackets

is the number of complete SNM/FB steps, i.e., those accepted with the unite stepsize in the linesearch step or by the linear decrease test without linesearch. Finally, the number of gradient steps and the number of evaluations of F are reported (recall that for both algorithms, the Jacobian of F is evaluated once per iteration). Failures are marked by "-".

Both variants of the algorithm failed for the following 10 problems: `billups`, `bishop`, `colvdual`, `ehl_k40`, `ehl_k60`, `ehl_k80`, `forcebsm`, `forcedsa`, `lincont`, `simple-ex`. Recall that by failure we mean that (3.5) was not satisfied after 500 iterations. In particular, for `billups` both variants of the algorithm converged to a local minimizer of the merit function, which is not an MCP solution. But for this problem, this is typical for (most) MCP algorithms. We note that the problems mentioned above are considered among the difficult ones in MCP literature. Since we were able to solve essentially all the other problems, this means that our implementation of the global scheme, though simple, is sufficiently robust. The rest of this discussion focuses on the cases for which at least one algorithm did not fail.

For 3 problems, namely `duopoly`, `shubik`, and `ne-hard`, switching to our step at some early stage prevented failure (although for `ne-hard`, the accuracy obtained by the algorithm without switching was of order 10^{-9}, i.e., almost satisfying the stopping test (3.5)). The opposite situation was observed for `games` only, though the obtained accuracy in the latter case was also of order 10^{-9}. We do not have an explanation for the behavior on those 4 problems and regard it as (probably) "accidental". Moreover, the possible negative global effect of switching to the active-set step too early can be avoided by the following simple trick. When an active-set step is accepted, we can store the previous iterate as a back-up, and restart the algorithm from that point if an active-set step is rejected on some subsequent iteration (which indicates that the switch occurred too early).

For 20 test problems the active-set step was never accepted, but trying it never harmed drastically. The number of evaluations of F for the algorithm without the switching option is typically not much less than for the complete variant, especially when the number of iterations is relatively large. We consider this extra work as a price to pay for safeguarding better local convergence properties in situations where SNM/FB runs into difficulties (see below).

The GNM/active-set step was ever accepted "improperly", i.e., far from a solution and with (probably) incorrect identification, for 5 test problems only. As mentioned above, for `duopoly` and `shubik` this actually prevented failure, but for `games` this caused it. For `freebert` this resulted in some extra linesearch steps. This points to the (obvious) fact that switching to the active-set strategy too early should be avoided.

For 19 test problems our step was used "properly" (on the final stage of the process). For `badfree` and `handskoop`, this was evidently rather advantageous. This has to do, of course, with the lack of (BD-)regularity, which affects the SNM/FB algorithm. In the other cases (apparently either BD-regular or not regular even in our sense), the conclusions are overall similar to the situation when our step was never used. In those cases, we do not win or lose much. On the one hand, we typically have to pay the price of a few more evaluations of F (but not always). On the other hand, note that the dimension of the linear system of equations which gives the GNM/active-set step is smaller than the dimension of the system to compute the SNM/FB step. This can be significant for large-scale problems (if the GNM/active-set step is accepted then no further computation is needed at this iteration). Finally, the "proper" use of our step never increased (and sometimes decreased) the overall number of

iterations.

Table 1: *Global algorithm for MCPLIB*

	Iter number		GNM/AS steps (tail)	SNM/FB steps (stepsize = 1)		Gradient steps		Number of evaluations of F	
Alg	1	2	2	1	2	1	2	1	2
badfree	36	8	2(1)			36	6	179	62
bert_oc	8	10	1(0)	8(8)	9(9)			24	56
bertsekas	59	59		59(6)	59(6)			937	1140
billups	–	–	–	–	–	–	–	–	–
bishop	–	–	–	–	–	–	–	–	–
bratu	29	29		29(10)	29(10)			200	258
choi	6	6		6(6)	6(6)			18	32
colvdual	–	–	–	–	–	–	–	–	–
colvnlp	11	11	1(1)	11(5)	10(4)			68	93
cycle	4	4	3(3)	4(3)	1(0)			15	23
degen	6	5	1(1)	6(6)	4(4)			18	30
duopoly	–	59	1(0)	–	56(10)	–	2	–	916
ehl_k40	–	–	–	–	–	–	–	–	–
ehl_k60	–	–	–	–	–	–	–	–	–
ehl_k80	–	–	–	–	–	–	–	–	–
ehl_kost	30	30		30(6)	30(6)			217	289
electric	131	131		131(8)	131(8)			1225	1663
explcp	22	21	1(1)	22(7)	20(5)			132	186
forcebsm	–	–	–	–	–	–	–	–	–
forcedsa	–	–	–	–	–	–	–	–	–
freebert	18	25	1(0)	18(5)	24(5)			124	291
gafni	13	13		13(4)	13(4)			76	114
games	30	–	–	30(7)	–		–	340	–
hanskoop	16	4	1(1)	15(15)	2(2)	1	1	53	26
hydroc06	5	5	4(4)	5(4)	1(0)			18	28
hydroc20	9	9	8(8)	9(8)	1(0)			31	49
jel	10	9	1(1)	10(8)	8(6)			38	57
josephy	4	3	1(1)	4(4)	2(2)			12	16
kojshin	4	3	1(1)	4(4)	2(2)			12	16
lincont	–	–	–	–	–	–	–	–	–
mathinum	8	8	2(2)	8(7)	6(5)			27	47
mathisum	9	9		9(7)	9(7)			33	65
methan08	4	4	3(3)	4(4)	1(1)			12	20
nash	6	6	1(1)	6(5)	5(4)			22	34
ne-hard	–	29	18(4)	–	11(4)	–		–	187
obstacle	8	8		8(6)	8(6)			30	48

Table 1: *Global algorithm for MCPLIB*

	Iter number		GNM/AS steps (tail)	SNM/FB steps (stepsize = 1)		Gradient steps		Number of evaluations of F	
Alg	1	2	2	1	2	1	2	1	2
opt_cont	12	12		12(10)	12(10)			43	71
opt_cont127	13	13		13(7)	13(7)			65	97
opt_cont255	15	15		15(8)	15(8)			75	111
opt_cont31	9	9		9(8)	9(8)			30	48
opt_cont511	20	20		20(6)	20(6)			140	200
pgvon106	28	28		27(2)	27(2)	1	1	206	301
pies	11	11		11(9)	11(9)			41	69
powell	8	8		6(6)	6(6)	2	2	40	68
powell_mcp	3	3	2(2)	3(3)	1(1)			9	15
qp	6	5	1(1)	6(6)	4(4)			18	28
scarfanum	5	5		5(5)	5(5)			15	31
scarfasum	5	5		5(5)	5(5)			15	31
scarfbsum	170	170		170(8)	170(8)			2002	2620
shubik	–	395	1(0)	–	392(6)	–	2	–	7002
simple-ex	–	–	–	–	–	–	–	–	–
simple-red	11	11	2(2)	11(11)	9(9)			33	65
sppe	7	7	2(2)	6(6)	4(4)	1	1	25	42
tinloi	8	8		8(8)	8(8)			24	48
tobin	3	3	1(1)	3(3)	2(2)			9	16
trafelas	58	58		55(7)	55(7)	3	3	471	599

General conclusions are as follows. The option of switching to our step never harms too much, though we certainly have to pay some extra price for computing it at some iterations (at those where the index sets do not change), even if this step is eventually rejected. But this is consistent with the main goal of the presented approach. We emphasize that the goal is not in improving the linesearch SNM/FB (or any other algorithm) *when it works efficiently*. We try not to harm/interfere too much in those cases, while extending the algorithm to the (irregular) cases *when SNM/FB does not work well*.

We also point out that the efficiency of the identification procedure is of crucial importance for the methods presented in this paper. According to our numerical experience, the switching to our local method usually occurs exactly one iteration after the correct identification is obtained. Additional tuning of the identification procedure (i.e., using scaling or different identification functions) might improve the performance significantly. Actually, identification techniques other than the one described above could also be tried. Also, note that we have not been using any heuristic considerations to decide whether or not the active-set step should be computed. Developing such heuristics can certainly save some computational work as well. For example, even if the active sets have not changed from one iteration to the next, we may decide not to compute the active-set step if the residual is relatively large (i.e., we are still far

from a solution). Other important issues are feasible versions of the method, and different globalization schemes which would better fit the structure of the method. This will be the subjects of future research.

To conclude, we illustrate some possible scenarios of the purely *local* behavior of SNM/FB and GNM/AS by applying them to the following small test problems with violated BD-regularity for Ψ_{FB}. We are talking here about the basic SNM/FB and GNM/AS iterations, without any modifications and tricks concerned with globalization. In particular, GNM/AS algorithm is specified precisely by (2.15), (2.16), with the index sets identified at the starting point (thus, this is the algorithm whose local properties are characterized by Theorem 2.2).

The first problem is a slight modification of [3, Example 3.1].

Example 3.1 Let $n = 2$, $l_i = 0$, $u_i = +\infty$, $i = 1, 2$, and let $F(x) = ((x_1-1)^2,\ x_1+x_2+x_2^2-1)$. The point $\bar{x} = (1,\ 0)$ is the solution of this NCP. Semistability (and hence, BD-regularity for Ψ_{NR} and Ψ_{FB}) is violated here, while 2-regularity and weak regularity hold. The starting point is $x^0 = (1.5,\ -0.5)$, with $\|x^0 - \bar{x}\| \approx 7.1\text{e–}01$, $\|\Psi_{FB}(x^0)\| \approx 8.2\text{e–}01$, $\det \Lambda_0 \approx 1.4\text{e+}00$.

SNM/FB converges in 13 steps. At the final step, $\|x^{13} - \bar{x}\| \approx 3.0\text{e–}05$, $\|\Psi_{FB}(x^{13})\| \approx 6.2\text{e–}10$, $\det \Lambda_{13} \approx -8.3\text{e–}05$ (which indicates degeneracy). The rate of convergence is linear, with the ratio approaching 1/2.

The behavior of GNM/AS is reported in Table 2, and it clearly shows fast quadratic convergence.

Table 2: *GNM/AS for Example 3.1*

k	0	1	2	3
$\|x^k - \bar{x}\|$	7.1e–01	1.3e–01	3.7e–03	9.9e–08
$\|\Psi_{FB}(x^k)\|$	8.2e–01	1.6e–02	1.4e–05	9.9e–15
$\frac{\|x^k-\bar{x}\|}{\|x^{k-1}-\bar{x}\|}$		1.8e–01	2.9e–02	2.7e–05

The next four problems, taken from [11, Examples 1–4] (Example 3.2 is slightly modified), are the KKT systems of the form (1.2), with $g(z) = f'(z)$, $z \in \mathbf{R}^p$, where the objective function $f : \mathbf{R}^p \to \mathbf{R}$ will be specified for each example below.

Example 3.2 Let $p = m = 2$, $f(z) = (z_1 + z_2)^2/2 + (z_1 + z_2)^3/3$, $G(z) = (z_1,\ z_2)$, $z \in \mathbf{R}^2$, $\bar{z} = 0$, $\bar{\mu} = 0$. Semistability holds here, but for Ψ_{NR} (and hence, for Ψ_{FB}), BD-regularity is violated. The starting point is $z^0 = (1,\ 2)$, $\mu^0 = (0.01,\ 0.01)$, with $\|x^0 - \bar{x}\| \approx 2.2\text{e+}00$, $\|\Psi_{FB}(x^0)\| \approx 1.7\text{e+}01$, $\det \Lambda_0 \approx 4.3\text{e–}04$.

The behavior of SNM/FB is as follows: $\det \Lambda_1 \approx 7.3\text{e–}10$, $\det \Lambda_2 \approx 0$, but the corresponding linear system is solvable, and the method manages to escape the "bad" region. Specifically, $\det \Lambda_3 \approx 4.3\text{e–}16$, while $\det \Lambda_4 \approx 2.5\text{e–}00$, and the algorithm converges in 7 iterations. At the final step, $\|x^7 - \bar{x}\| \approx 1.0\text{e–}16$, $\|\Psi_{FB}(x^7)\| \approx 1.4\text{e–}16$, $\det \Lambda_7 \approx 2$. The rate of convergence is superlinear.

The behavior of GNM/AS is reported in Table 3, and it also shows the superlinear rate.

Note that while SNM/FB and GNM/AS exhibit similar convergence for this problem, the performance of SNM/FB clearly depends on the specific implementation. In particular,

solution of a given degenerate linear system depends on the linear solver, and can affect the overall convergence. Also, in the cases where BD-regularity is violated, different procedures to compute Λ_k could result (in general, not in this example) in different linear systems and some of them can be ill-conditioned close to the solution, preventing fast convergence of SNM/FB.

Table 3: *GNM/AS for Example 3.2*

k	0	1	2	3	4	5	6	7
$\|x^k - \bar{x}\|$	2.2e+00	9.0e–01	3.2e–01	7.1e–02	5.0e–03	2.8e–05	9.1e–10	9.3e–19
$\|\Psi_{FB}(x^k)\|$	1.7e+01	4.1e+00	9.3e–01	1.6e–01	1.0e–02	5.7e–05	1.8e–09	1.8e–18
$\frac{\|x^k-\bar{x}\|}{\|x^{k-1}-\bar{x}\|}$		4.0e–01	3.5e–01	2.2e–01	7.1e–02	5.6e–03	3.2e–05	1.0e–09

The next example shows that both 2-regularity of Ψ_S and weak regularity are important for fast convergence of GNM/AS.

Example 3.3 Let $p = m = 2$, $f(z) = z_1^2/2 + z_2^3/3$, $G(z) = (z_1 - z_2^2/2,\, z_1 + z_2^2/2)$, $z \in \mathbf{R}^2$, $\bar{z} = 0$, $\bar{\mu} = 0$. Semistability is violated (and hence, BD-regularity for Ψ_{NR} and Ψ_{FB} is violated). For Ψ_S, 2-regularity holds, but weak regularity does not. The starting point is $z^0 = (0.1,\, 0.1)$, $\mu^0 = (0.1,\, 0.1)$, with $\|x^0 - \bar{x}\| \approx$ 2.0e–01, $\|\Psi_{FB}(x^0)\| \approx$ 1.3e–01, $\det \Lambda_0 \approx$ 2.0e–01.

Both SNM/FB and GNM/AS converge in 12 steps, and $\|x^{12} - \bar{x}\| \approx$ 2.4e–05, $\|\Psi_{FB}(x^{12})\| \approx$ 8.4e–10, $\det \Lambda_{12} \approx$ 2.9e–04. The rate of convergence is linear with ratio 1/2.

Example 3.4 Let $p = 2$, $m = 3$, $f(z) = z_1 + (z_1^2 + z_2^2)/2$, $G(z) = (z_1,\, z_2,\, z_1 + z_2)$, $z \in \mathbf{R}^2$, $\bar{z} = 0$, $\bar{\mu} = (1,\, 0,\, 0)$. Semistability holds here, but for Ψ_{NR} (and hence, for Ψ_{FB}), BD-regularity is violated. The starting point is $z^0 = 0$, $\mu^0 = (1,\, 0.01,\, 0.01)$, with $\|x^0 - \bar{x}\| \approx$ 2.2e–01, $\|\Psi_{FB}(x^0)\| \approx$ 1.2e–01, $\det \Lambda_0 = 0$.

Though Λ_0 is singular, the SNM/FB linear system is solvable, and the method manages to escape. Specifically, $\det \Lambda_1 = -2$, and the algorithm converges in 2 iterations: the second step gives the exact solution with $\det \Lambda_2 \approx$ –3.5e–01. Note however that, as already discussed above, this behavior cannot be guaranteed for some other implementation of the linear system solver (and in general, of the procedure to compute Λ_k).

GNM/AS terminates after 1 step at the exact solution. The reason for this is that $A_0 = \{4,\, 5\}$, while x_1^0, x_2^0, and x_3^0 coincide with the corresponding components of $\bar{x}$. The iteration of GNM/AS terminates with the exact solution at the identification phase.

Example 3.5 Let $p = m = 1$, $f(z) = z^4/4$, $G(z) = z$, $z \in \mathbf{R}$, $\bar{z} = 0$, $\bar{\mu} = 0$. Weak regularity holds here, but semistability (and hence, BD-regularity for Ψ_{NR} and Ψ_{FB}) and even 2-regularity for Ψ_S, are violated. The starting point is $z^0 = 1$, $\mu^0 = 0.1$, with $\|x^0 - \bar{x}\| \approx$ 1.0e+00, $\|\Psi_{FB}(x^0)\| \approx$ 9.1e–01, $\det \Lambda_0 \approx$ 2.7e+00.

SNM/FB converges in 18 steps. At the final step, $\|x^{18} - \bar{x}\| \approx$ 6.8e–04, $\|\Psi_{FB}(x^{18})\| \approx$ 3.1e–10, $\det \Lambda_{13} \approx$ 3.1e–06. The rate of convergence is linear with ratio approaching 2/3.

The behavior of GNM/AS is reported in Table 4, and it shows fast quadratic convergence.

Table 4: *GNM/AS for Example 3.5*

k	0	1	2	3	4
$\|x^k - \bar{x}\|$	1.0e+00	6.0e–01	2.2e–01	2.7e–03	9.0e-13
$\|\Psi_{FB}(x^k)\|$	9.1e–01	2.2e–01	1.0e–03	2.0e–08	7.4e-37
$\frac{\|x^k-\bar{x}\|}{\|x^{k-1}-\bar{x}\|}$		6.0e–01	3.6e–01	1.3e–02	3.3e-10

The final example NCP taken from [12, Example 2.1].

Example 3.6 Let $n = 2$, $F(x) = (-x_1 + x_2, -x_2)$, $x \in \mathbf{R}^2$, $\bar{x} = 0$. BD-regularity holds for Ψ_{NR} (and hence, semistability also holds), but not for Ψ_{FB}. The starting point is $x^0 = (2, 4)$, with $\|x^0 - \bar{x}\| \approx 4.5\text{e+}00$, $\|\Psi_{FB}(x^0)\| \approx 5.8\text{e+}00$, $\det \Lambda_0 = 0$.

Here, SNM/FB fails to make a step. At the same time, GNM/AS terminates after 1 step at the exact solution. The reason for this is that $A_0 = \{1, 2\}$. Thus, the iteration of GNM/AS reduces to identifying the index sets.

Note that the problem in Example 3.6 is actually a linear complementarity problem, that is, NCP with affine F. We point out that in the case of affine F, just one step of GNM/AS gives the exact solution, provided the index sets are correctly identified. For example, this behavior is observed also for the problem `badfree` from the MCPLIB collection: once x^k is close to the solution $\bar{x} = (0, 0, 0.5, 0.5, 1)$, GNM/AS produces $x^{k+1} = \bar{x}$. At the same time, for x^k close to $\bar{x}$, a degenerate Λ_k is computed, and SNM/FB fails.

References

[1] S.C. Billups. Algorithms for complementarity problems and generalized equations. Technical Report 95-14, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, August 1995. PhD thesis.

[2] J. F. Bonnans. Local analysis of Newton-type methods for variational inequalities and nonlinear programming. *Applied Mathematics and Optimization*, 29:161–186, 1994.

[3] A.N. Daryina, A.F. Izmailov, and M.V. Solodov. A class of active-set Newton methods for mixed complementarity problems. IMPA preprint A256, October 2003 (revised April 2004). Submitted to *SIAM Journal on Optimization*.

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

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

[6] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer–Verlag, New York, NY, 2003.

[7] M.C. Ferris, C. Kanzow, and T.S. Munson. Feasible descent algorithms for mixed complementarity problems. *Mathematical Programming*, 86:475–497, 1999.

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

[9] A.F. Izmailov and M.V. Solodov. Error bounds for 2-regular mappings with Lipschitzian derivatives and their applications. *Mathematical Programming*, 89:413–435, 2001.

[10] A.F. Izmailov and M.V. Solodov. The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. *Mathematics of Operations Research*, 27:614–635, 2002.

[11] A.F. Izmailov and M.V. Solodov. Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. *Mathematical Programming*, 95:631–650, 2003.

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

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