Advances in Mathematical Sciences and Applications
Vol.14, No.2 (2004), pp.513-535

Gakkōtosho
Tokyo, Japan

SEMISMOOTH NEWTON METHODS FOR A CLASS OF UNILATERALLY CONSTRAINED VARIATIONAL PROBLEMS

M. Hintermüller
Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005, USA
(hint@caam.rice.edu)

V. Kovtunenko
Department of Mathematics, University of Graz, A-8010 Graz, Austria
(kovtunen@kfunigraz.ac.at)

K. Kunisch
Department of Mathematics, University of Graz, A-8010 Graz, Austria
(karl.kunisch@uni-graz.at)

Abstract. A class of semismooth Newton methods for quadratic minimization problems subject to non-negativity constraints resulting from discretizing classes of optimization problems in function spaces is considered. For the algorithm, which is equivalent to a primal-dual active set strategy, locally superlinear as well as global convergence results are established. The global convergence assertions rely on matrix properties which characterize classes of discretized differential operators. Further, under an M-matrix property monotonous convergence with respect to the constrained components of the primal iterates is established. A comprehensive report on numerical tests is provided for the scalar valued problem with a boundary obstacle, the vector-valued Signorini problem with an obstacle, and the symmetric crack problem. The numerical results support the theoretical findings.

Communicated by Editors; Received ?????, 2003.
This work is supported by the Austrian Science Fund (FWF) in framework of the SFB project F003 "Optimierung und Kontrolle" and the Lise Meitner project M622/M737 "Variational methods in application to crack problems".
AMS Subject Classification 90C33, 49M29, 74B05.

1. Introduction

A wide class of problems in applications can be expressed as a minimization problem of a convex quadratic objective function subject to non-negativity constraints:

$$\text{minimize} \quad \tfrac{1}{2}u^\top Lu - f^\top u \text{ subject to } u \in \mathbb{R}^N,\ u_B \geq 0, \tag{1}$$

where L is a symmetric, positive definite $N \times N$ matrix, $N \in \mathbb{N}$, $f \in \mathbb{R}^N$, and $B \subset \{1, \ldots, N\}$. By u_B we refer to the components u_i of the vector $u \in \mathbb{R}^N$ with $i \in B$. Let D denote the complement of B in $C := \{1, \ldots, N\}$, *i.e.*, $D = C \setminus B$. It is well-known that the unique solution u^* of (1) is characterized by the existence of a Lagrange multiplier $\lambda^* \in \mathbb{R}^N$ such that

$$Lu^* + \lambda^* = f, \tag{2a}$$

$$u_B^* \geq 0, \quad \lambda_B^* \leq 0, \quad (u_B^*)^\top \lambda_B^* = 0, \tag{2b}$$

$$\lambda_D^* = 0. \tag{2c}$$

From a quadratic programming (QP) point of view the problem class (1) is of a simple structure and many algorithms were proposed for its numerical solution; see the selected references [3, 6, 7, 8, 17, 19, 20, 26, 27, 28]. By now, the research level in this area has reached a high level of sophistication. However, for problems which result from discretizing differential operators, the research level is less complete. First of all some of the classical methods for solving QPs are sorted out by the tremendous size of the problem; e.g. [6, 7]. Frequently, algorithms for solving general QPs do not exploit structural properties of L or B. In fact, if L is related to a discretization of some differential operator by means of finite differences or finite elements then the matrix L will be sparse with a particular block structure. Further, the set B may refer to nodal points at the boundary of the domain of the underlying continuous problem. In this case, thinking of B as some arbitrary subset of the set of indices C will disregard properties of the resulting problem like the existence of Schur complements or invertibility of submatrices of L. Indeed, in some applications the mapping defining u_B^* as a function of u_D^* and f can be related to a trace operator.

In order to substantiate our latter arguments we consider the following problems in elasticity with boundary constraints [10, 11, 14, 12]: Let a solid occupy the domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with the boundary $\partial\Omega$. Consider the quadratic functional of potential energy of the solid under a given load $g \in (L^2(\Gamma_n))^d$ applied to the subset Γ_n of the boundary $\partial\Omega$:

$$J(u) := \tfrac{1}{2} \int_\Omega \sigma(u) : \varepsilon(u) - \int_{\Gamma_n} g \cdot u,$$

where ε and σ are the strain and stress tensors for the displacement vector $u(\mathrm{x}) \in \mathbb{R}^d$, respectively, and are defined by

$$\varepsilon(u) = \tfrac{1}{2}(\nabla u + \nabla u^\top) \quad \text{and} \quad \sigma(u) = c : \varepsilon(u).$$

The tensor $c = (c_{ijkl})$ describes material properties. Above $L^2(\Gamma_n)$ denotes the Hilbert space of square integrable functions over Γ_n. Let $\mathcal{B}$ be a part of the boundary $\partial\Omega$ with $\mathcal{B} \cap \Gamma_n = \emptyset$. We mention here three types of boundary constraints at $\mathcal{B}$:

(S) Signorini condition**:** $u \cdot \nu \geq 0$ on $\mathcal{B}$;
(O) boundary obstacle**:** $u \cdot \nu \geq \psi$ on $\mathcal{B}$;
(NP) non-penetration**:** $[\![u \cdot \nu]\!] \geq 0$ on $\mathcal{B}$,

where the function ψ is an obstacle, ν is the normal vector at $\mathcal{B}$, and $[\![\cdot]\!]$ denotes the jump across $\mathcal{B}$ in the case where $\mathcal{B}$ refers to a crack inside the body; see e.g. Figure 14. Case (O) above can be reduced to (S) by a simple transformation of u . If the crack and the corresponding data are geometrically symmetric with respect to $\mathcal{B}$, then case (NP) can be reduced to (S), too. In such a way, for all mentioned cases we arrive at the constrained minimization problem:

$$\text{(3)} \qquad \text{minimize} \quad J(u) \quad \text{subject to } u \in H \subset (H^1(\Omega))^d,\ u \cdot \nu \geq 0 \text{ on } \mathcal{B} \subset \partial\Omega.$$

Here $H^1(\Omega)$ denotes the Sobolev space of functions in $L^2(\Omega)$ which have distributional derivatives in $L^2(\Omega)$, and H denotes a function space excluding rigid body motions, typically by means of a Dirichlet boundary condition on a part of the boundary $\partial\Omega$. Under standard regularity assumptions the solution u to (3) exists uniquely, and $u_{\mathcal{B}} = u \cdot \nu|_{\mathcal{B}}$ is an element of $(L^2(\mathcal{B}))^d$, or smoother. After appropriate discretization (see Section 4) problem (3) can be written in the form (1). Consequently, in the discrete setting B refers to grid points or nodal points on the boundary $\mathcal{B}$ in case of finite difference respectively finite element discretizations. Further, the matrix L results from discretizing the first integral in the definition of J. Due to the discretization process it typically also includes boundary nodes. For the discretization described in Section 4 L is symmetric and positive definite. Moreover, it is sparse and has a particular block structure corresponding to B and D, with D denoting the grid points or nodal points in $\overline{\Omega} \setminus \mathcal{B}$. The vector f results from the discretization of g on $\mathcal{B}$ and $f = 0$ on D.

In this paper, we devise a semismooth Newton method for computing the solution (u^*, λ^*) of (2) iteratively. It operates on the reformulation of (2b) based on the nonlinear complementarity problem (NCP) function

$$\phi(a, b) = a - \max(\alpha b + a, 0), \quad a, b \in \mathbb{R},$$

where $\alpha > 0$ is arbitrarily fixed. In fact, the following equivalence holds true:

$$\text{(4)} \qquad \phi(a, b) = 0 \quad \Longleftrightarrow \quad a \geq 0,\ b \leq 0,\ ab = 0.$$

For $v, w \in \mathbb{R}^{|B|}$ one defines

$$\Phi(v, w) = (\phi(v_1, w_1), \ldots, \phi(v_{|B|}, w_{|B|}))^\top.$$

If $\Phi(v, w) = 0$, then the complementarity relation (4) is satisfied for the components v_i, w_i of the vectors v, w. This allows us to rewrite (2) as

$$\text{(5a)} \qquad Lu^* + \lambda^* = f,$$

$$\text{(5b)} \qquad \Phi(u^*_B, \lambda^*_B) = 0,$$

$$\text{(5c)} \qquad \lambda^*_D = 0.$$

Due to the nondifferentiability of Φ, Newton techniques for computing the solution of (5) have to work with generalizations of the derivative of a function. The resulting methods are called *generalized Newton methods*; see e.g. [15, 16, 21, 22, 23, 24, 25].

The *semismooth Newton method* which we propose in this paper is related to the algorithm analyzed in [9]. Its local convergence properties rely on the semismoothness property of the max-operator involved in (5b). The concept of semismoothness of a function was originally introduced by Mifflin in [18] and extended to $\mathbb{R}^N$ by Qi and Sun in [24]. In [4] the notion of a slanting function is introduced which is related to semismoothness properties of a function. In fact, according to [4] a mapping $F : D \subset Y \to Z$ is

called *slantly differentiable* in an open subset $U \subset D$ if there exists a family of mappings $G : U \to \mathcal{L}(Y, Z)$ such that

$$\lim_{h \to 0} \frac{1}{\|h\|} \|F(y+h) - F(y) - G(y+h)h\| = 0 \tag{6}$$

for every $y \in U$. The mapping G is called a *slanting function* for F in U. Above, Y, Z denote Banach spaces, and D is an open subset of Y. A significant amount of research work is devoted to generalized or semismooth Newton methods. In addition to the above references we only refer to the recent monograph [13] and the references therein. In general, under an additional boundedness assumption the semismoothness or the slanting property are sufficient for proving the locally superlinear rate of convergence for the respective generalized Newton's method [4, 9, 16, 24].

Typically, these results on the local rate of convergence are in general terms without exploiting the structure of the problem e.g. resulting from discretizing a variational inequality problem involving partial differential operators. In [9] it was observed that the convergence results can be strengthened when assuming that L comes from discretizing second order elliptic differential operators. In fact, one can prove global convergence with locally superlinear rate for the corresponding semismooth Newton method. Further, under the assumption that L is a nonsingular M-matrix, *i.e.* L is invertible, $L_{ii} > 0$ for all $i \in C$ and $L_{ij} \leq 0$ for $i \neq j$, with $i, j \in C$, for $B = C$ it is shown that the primal iterates converge monotonically. In the present paper we pick up the latter point of view and show that our semismooth Newton method exhibits the aforementioned convergence properties for solving (5) in cases where L is a nonsingular M-matrix and $B \subset C$. Further, monotonicity of the primal iterates will be argued for problems where the Schur complement of L is a sufficiently small perturbation of an M-matrix. These assertions extend the results obtained in [9] and give a theoretical account for the behavior of the method which can be observed in numerical practice.

Motivated by (3), the numerical tests in this paper focus on boundary constrained problems. The test problems cover the scalar-valued problem with boundary obstacle, the vector-valued Signorini problem with an obstacle, and the symmetric crack problem. In the first case, the matrix L is the discrete Laplace operator with homogeneous Dirichlet boundary conditions on a part of the boundary. Hence, L is an M-matrix. For the second problem class mentioned above, L depends on the Lamé parameters and is a discretization of the second order elliptic differential operator occurring in the Lamé equation. Differently from the first case, the Schur complement of L is no longer an M-matrix. It is merely a small perturbation of an M-matrix. A similar observation corresponding to L and its Schur complement is true in the case of the symmetric crack problem.

The rest of the paper is organized as follows. In Section 2 we introduce the algorithm. We first derive it as a semismooth Newton method. Motivated by the nonsmoothness of the involved operator we then interpret the method as a primal-dual active set strategy. We end Section 2 by stating a result on the locally superlinear convergence of our method. Section 3 is devoted to the analysis of global as well as monotone convergence properties of the algorithm. Finally, in Section 4 we report on numerical results attained by our algorithm for the discretizations of a scalar-valued problem with a boundary obstacle, a vector-valued Signorini problem and a symmetric crack problem.

2. The algorithm and its local convergence

In this section we introduce the semismooth Newton method for computing the solution (u^*, λ^*) of (5). In [9] it is shown that it is equivalent to the primal-dual active set method introduced in [1]. In the sequel we use the splitting of vectors and matrices into blocks corresponding to the subsets D and B of C as follows:

$$L = \begin{pmatrix} L_{DD} & L_{DB} \\ L_{BD} & L_{BB} \end{pmatrix}, \quad u = \begin{pmatrix} u_D \\ u_B \end{pmatrix}, \quad f = \begin{pmatrix} f_D \\ f_B \end{pmatrix}.$$

With this notation, the complementarity problem (5) can be written as

$$\begin{pmatrix} L_{DD} & L_{DB} \\ L_{BD} & L_{BB} \end{pmatrix} \begin{pmatrix} u_D \\ u_B \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_B \end{pmatrix} = \begin{pmatrix} f_D \\ f_B \end{pmatrix}, \tag{7a}$$

$$\Phi(u_B, \lambda_B) = 0. \tag{7b}$$

For defining the algorithm we introduce $F : \mathbb{R}^N \times \mathbb{R}^{|B|}$,

$$F(y) = \begin{pmatrix} L_{DD}u_D + L_{DB}u_B - f_D \\ L_{BD}u_D + L_{BB}u_B + \lambda_B - f_B \\ \Phi(u_B, \lambda_B) \end{pmatrix}, \quad y = \begin{pmatrix} u_D \\ u_B \\ \lambda_B \end{pmatrix}.$$

Thus, (7) can be written as

$$F(y) = 0. \tag{8}$$

Let

$$\chi_S = \operatorname{diag}(s_1, \ldots, s_N), \text{ with } s_i = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S. \end{cases}$$

and define

$$A(y) = \{i \in B : u_i + \alpha\lambda_i < 0\}, \quad I(y) = \{i \in B : u_i + \alpha\lambda_i \geq 0\}. \tag{9}$$

Then it can be verified that for every $y \in \mathbb{R}^N \times \mathbb{R}^{|B|}$ the matrix

$$G(y) = \begin{pmatrix} L_{DD} & L_{DB} & 0 \\ L_{BD} & L_{BB} & E_B \\ 0 & \chi_{A(y)} & -\alpha\chi_{I(y)} \end{pmatrix} \tag{10}$$

satisfies

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|} \|F(y+h) - F(y) - G(y+h)h\| = 0. \tag{11}$$

Above, E_B denotes the unit matrix in $\mathbb{R}^{|B| \times |B|}$. According to Definition 1.1 in [9] (compare also (6)), G is a slanting function for F. Thus, G serves as a generalized derivative of the nondifferentiable mapping F.

The semismooth Newton method for computing the solution y^* of $F(y) = 0$ is defined as follows: For some initial guess y^0 sufficiently close to y^* compute

$$y^{k+1} = y^k - G(y^k)^{-1}F(y^k), \quad k = 0, 1, 2, \ldots. \tag{12}$$

From the general local convergence results for Newton iterations like (12), we deduce the following variant. In its formulation, for a $|B| \times |B|$-matrix Q and index sets $R, S \subset B$, we use $Q_{RS} = (Q_{rs})$ with $r \in R$ and $s \in S$.

Theorem 2.1. *The semismooth Newton iteration* (12) *converges superlinearly to* y^* *with* $F(y^*) = 0$ *provided that* y^0 *is sufficiently close to* y^* *and that for all index sets* $I \subset B$ *the inverse matrix of* $L_{BD}L_{DD}^{-1}L_{DB} - L_{BB}$ *exists and that there exists a constant* $\beta > 0$ *independently of* I *such that*

$$\|((L_{BD}L_{DD}^{-1}L_{DB} - L_{BB})_{II})^{-1}\| \leq \beta. \tag{13}$$

Proof. Let y and $g = (g_D, g_B, g_\lambda)^\top \in \mathbb{R}^N \times \mathbb{R}^{|B|}$ be arbitrarily fixed. We show that there exists a unique $z = (z_D, z_B, z_\lambda)^\top \in \mathbb{R}^N \times \mathbb{R}^{|B|}$ such that $G(y)z = g$. Note that the latter equation is equivalent to

$$\chi_{A(y)}z_B + \alpha\chi_{I(y)}Sz_B = g_\lambda + \alpha\chi_{I(y)}(g_B - L_{BD}L_{DD}^{-1}g_D) =: \tilde{g}, \tag{14}$$

with $S := L_{BB} - L_{BD}L_{DD}^{-1}L_{DB}$. Since $I(y) \cap A(y) = \emptyset$, we obtain from (14) that $z_i = \tilde{g}_i$ for all $i \in A(y)$. Hence, $z_{I(y)}$ is uniquely defined by

$$S_{I(y)I(y)}z_{I(y)} = \tfrac{1}{\alpha}\tilde{g}_{I(y)} - S_{I(y)A(y)}\tilde{g}_{A(y)}.$$

The positive definiteness of L and assumption (13) guarantee the existence of a constant $\bar{\beta} > 0$ independently of y, and $I(y), A(y)$ such that

$$\|z\| \leq \bar{\beta}\|g\|.$$

This proves that $G(y)^{-1}$ exists and is uniformly bounded. Now, the standard convergence proof–see e.g. the proof of Theorem 1.1 in [9]–yields the locally superlinear convergence of the Newton iteration (12). □

According to e.g. [5] the matrix

$$S = L_{BB} - L_{BD}L_{DD}^{-1}L_{DB} \tag{15}$$

defined in the proof of Theorem 3.1 is called the *Schur complement* of L_{DD} in L.

In our numerical tests we report on the following implementation of (12). In [9] the subsequent algorithm is referred to as primal-dual active set strategy.

Algorithm 1.

(0) *Choose* $(u^0, \lambda_B^0) \in \mathbb{R}^N \times \mathbb{R}^{|B|}$*; set* $k = 0$.

(1) *Decompose the index set* B *into*

$$A^k = \{i \in B : u_i^k + \alpha\lambda_i^k < 0\}, \tag{16a}$$

$$I^k = \{i \in B : u_i^k + \alpha\lambda_i^k \geq 0\}. \tag{16b}$$

(2) *If* $k \geq 1$ *and* $A^k = A^{k-1}$ *then STOP; else go to step 3.*

(3) *Solve for* $(u^{k+1}, \lambda_B^{k+1}) \in \mathbb{R}^N \times \mathbb{R}^{|B|}$*:*

$$\begin{pmatrix} L_{DD} & L_{DB} \\ L_{BD} & L_{BB} \end{pmatrix} \begin{pmatrix} u_D^{k+1} \\ u_B^{k+1} \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_B^{k+1} \end{pmatrix} = \begin{pmatrix} f_D \\ f_B \end{pmatrix}, \tag{17a}$$

$$u_i^{k+1} = 0 \quad \text{for all } i \in A^k, \qquad \lambda_i^{k+1} = 0 \quad \text{for all } i \in I^k. \tag{17b}$$

(4) *Set* $k = k + 1$ *and go to step 1.*

Note that Algorithm 1 and the Newton process (12) are equivalent. In fact, the first two equations in (12) coincide with (17a). The equations (16) and (17b) realize the Newton step for the nonsmooth, *i.e.* third, equation in (12). In order to see this, recall that the third equation in (12) is given by

$$\chi_{A(y^k)}(u_B^{k+1} - u_B^k) - \alpha\chi_{I(y^k)}(\lambda_B^{k+1} - \lambda_B^k) = -u_B^k + \max(u_B^k + \alpha\lambda_B^k, 0). \tag{18}$$

For $i \in I(y^k) = I^k$ we have $u_i^k + \alpha\lambda_i^k \geq 0$. Thus, (18) yields $\lambda_i^{k+1} = 0$ for $i \in I^k$. For $i \in A(y^k) = A^k$ we have $u_i^k + \alpha\lambda_i^k < 0$, and we obtain from (18) the equality $u_i^{k+1} = 0$. Combining the two cases we recover (17b). As a consequence the system in (17) is well-defined and, under the assumptions of Theorem 2.1, it admits a unique solution.

The stopping rule in step 2 of Algorithm 1 is motivated by the following considerations. For $i \in A^{k-1}$ we have $u_i^k = 0$, and for $i \in I^{k-1}$ we obtain $\lambda_i^k = 0$. Hence, if we assume that $A^{k-1} = A^k$, then from (16a) we infer $\lambda_i^k < 0$ for all $i \in A^k$, and $u_i^k \geq 0$ for all $i \in I^k$ by (16b). This, together with (17a), proves that the iterate $(u^k, \lambda_B^k) =: y^k$ upon termination of Algorithm 1 in step 2 satisfies $F(y^k) = 0$. Let us emphasize that the successful termination occurs after a finite number of iterations. Indeed, since there exists only a finite number of choices for A^k with $A^k \neq A^n$ for $n \neq k$ (and analogously for I^k), Theorem 2.1 yields the finite step convergence.

3. Global convergence results

In Theorem 2.1 we investigated the local convergence properties of Algorithm 1. In this section we derive additional global convergence results. The key ingredient is the M-matrix property of the iteration matrix of Algorithm 1 operating on u_B. Let us recall the definition of a nonsingular M-matrix.

Definition 3.1. *A matrix $Q \in \mathbb{R}^{N\times N}$ is a* nonsingular M-matrix *if Q is invertible, $Q_{ii} > 0$ for all $i \in \{1, \ldots, N\}$ and $Q_{ij} \leq 0$ for all $i, j \in \{1, \ldots, N\}$ with $i \neq j$.*

In [2] it is argued that if Q is a nonsingular M-matrix, then $Q^{-1} \geq 0$ in an elementwise sense.

The following convergence result applies in the case where L results from discretizing e.g. the scalar-valued problem with a boundary obstacle or the distributed obstacle problem where the obstacle acts only on a subset of the domain. The second part of the proof of the next theorem is similar to the proof of Theorem 3.2 in [9]. For the sake of completeness and for later reference we provide the entire proof.

Theorem 3.1. *Let $L \in \mathbb{R}^{N\times N}$ be an* M*-matrix. Then the iterates (u^k, λ_B^k) of Algorithm 1 converge to (u^*, λ_B^*) for arbitrary initial data $(u^0, \lambda_B^0) \in \mathbb{R}^N \times \mathbb{R}^{|B|}$. The local rate of convergence is superlinear. Moreover, the following monotonicity and feasibility relations hold true:*

$$u_B^* \geq u_B^{k+1} \geq u_B^k \quad \text{for all } k \geq 1 \quad \text{and} \quad u_B^k \geq 0 \quad \text{for all } k \geq 2. \tag{19}$$

Proof. Note that the subsystem (17a) is equivalent to

$$(L_{BB} - L_{BD}L_{DD}^{-1}L_{DB})u_B + \lambda_B = f_B - L_{BD}L_{DD}^{-1}f_D =: \tilde{f}. \tag{20}$$

Now we make use of the Schur complement S of L defined in (15). Hence, step 3 of Algorithm 1 computes the solution $(u^{k+1}, \lambda_B^{k+1})$ of the system

$$Su_B + \lambda_B = \tilde{f}, \tag{21a}$$

$$u_i = 0 \quad \text{for all } i \in A^k, \quad \lambda_i = 0 \quad \text{for all } i \in I^k. \tag{21b}$$

In [5] it is shown that the Schur complement of a nonsingular M-matrix is a nonsingular M-matrix again. Thus, S is a nonsingular M-matrix. As a consequence we have

$$S_{II}^{-1} \geq 0 \quad \text{and} \quad S_{II}^{-1}S_{IA} \leq 0 \tag{22}$$

for arbitrary index sets $A, I \subset B$. With these sign properties we now argue the monotonicity of $\{u_B^k\}$. First note that for $k \geq 1$ (17b) yields $u_i^k \lambda_i^k = 0$ for all $i \in B$. For $i \in A^k$ we have either $\lambda_i^k = 0$, which implies $u_i^k < 0$, or $\lambda_i^k < 0$, which yields $u_i^k = 0$. As a consequence, we have

$$u_i^k \leq 0 = u_i^{k+1} \text{ for all } i \in A^k. \tag{23}$$

Analogously, we obtain

$$\lambda_i^k \geq 0 = \lambda_i^{k+1} \text{ for all } i \in I^k. \tag{24}$$

The Newton step for (21a) yields

$$S(u_B^{k+1} - u_B^k) + (\lambda_B^{k+1} - \lambda_B^k) = 0. \tag{25}$$

Splitting this equation according to the partition (A^k, I^k) of B results in

$$u_{I^k}^{k+1} - u_{I^k}^k = -S_{I^k I^k}^{-1} S_{I^k A^k}(u_{A^k}^{k+1} - u_{A^k}^k) - S_{I^k I^k}^{-1}(\lambda_{I^k}^{k+1} - \lambda_{I^k}^k) \geq 0$$

Thus, from (22)–(25) we infer $u_B^{k+1} \geq u_B^k$ for all $k \geq 1$.

The feasibility of u_B^k for $k \geq 2$ can be argued as follows: Due to the monotonicity of $\{u_B^k\}_{k\geq 1}$ it is sufficient to show $u_B^2 \geq 0$. For this purpose let $V := \{i \in B : u_i^1 < 0\}$ denote the set of indices for which the constraint is violated. For $i \in V$ we have $\lambda_i^1 = 0$. Hence, $u_i^1 + \alpha\lambda_i^1 < 0$ and consequently $i \in A^1$. Since $u_{A^1}^2 = 0$ and $u_B^2 \geq u_B^1$ it follows that $u_B^2 \geq 0$.

Next we show that $u_B^k \leq u_B^*$ for all $k \geq 2$. To this end, observe that

$$\tilde{f}_{I^k} = \lambda_{I^k}^* + S_{I^k I^k} u_{I^k}^* + S_{I^k A^k} u_{A^k}^* = S_{I^k I^k} u_{I^k}^{k+1}$$

for $k \geq 1$. From this we obtain

$$S_{I^k I^k}(u_{I^k}^{k+1} - u_{I^k}^*) = \lambda_{I^k}^* + S_{I^k A^k} u_{A^k}^*.$$

Since $\lambda^* \leq 0$ and $u_B^* \geq 0$, the M-matrix properties of S imply $u_B^{k+1} \leq u_B^*$.

Next we consider $\{\lambda_B^k\}$. Let (k^-, i), $k^- \geq 1$, denote an index pair with $\lambda_i^{k^-} > 0$. Then $i \in A^{k^- - 1}$. Thus, $i \in I^{k^-}$ and, hence, $\lambda_i^{k^-+1} = 0$. Since $u_B^k \geq 0$ for $k \geq 2$, we have $u_i^{k^-+1} + \alpha\lambda_i^{k^-+1} \geq 0$. Consequently $i \in I^{k^-+1}$ and by induction $i \in I^k$ for all $k \geq k^-$. Hence, there exists an index $\bar{k} \in \mathbb{N}$ such that $\lambda_B^k \leq 0$ for all $k \geq \bar{k}$.

The monotonicity of $\{u_B^k\}_{k\geq 1}$ and $0 \leq u_B^k \leq u_B^*$ for all $k \geq 2$ imply the existence of $\bar{u}_B \geq 0$ with $\lim_k u_B^k = \bar{u}_B$. Further, due to $\lambda_B^k = \tilde{f} - S u_B^k$ there exists $\bar{\lambda}_B \leq 0$ with $\lim_k \lambda_B^k = \bar{\lambda}_B$. Since $u_i^k \lambda_i^k = 0$ for all $i \in B$ and for all $k \geq 1$, we obtain $(\bar{u}_B, \bar{\lambda}_B) = (u_B^*, \lambda_B^*)$.

The locally superlinear convergence follows from Theorem 2.1 and the fact that S is a nonsingular M-matrix. □

Let us interpret the proof of Theorem 3.1. Concerning the properties of $\{\lambda_B^k\}$ observe that there exists an index $k_\lambda \geq 1$ such that $\lambda_B^k \leq 0$ for all $k \geq k_\lambda$. Since $u_B^k \geq 0$ for all $k \geq 2$ and the pair (u_B^k, λ_B^k) satisfies (17a) for all $k \geq 1$, we have that Algorithm 1 stops at iteration $k^* = \max\{k_\lambda, 2\} + 1$ if u_B^0 is infeasible. If the initial choice u_B^0 is feasible, *i.e.* $u_B^0 \geq 0$, then $k^* = \max\{k_\lambda, 1\} + 1$.

In many applications L is positive definite, but not an M-matrix. Important problem classes of this type are given by the vector-valued Signorini problem with a boundary obstacle and the symmetric crack problem. In these cases S (from the proof of Theorem 3.1) is no longer an M-matrix. On the other hand, when studying the matrix entries of S, one often finds that the diagonal elements are positive (like for M-matrices) and only some

off-diagonal elements are positive, thus destroying the M-matrix property of S. Further, typically in each row of S the positive off-diagonal elements are small compared to the absolute value of the other elements of the same row; see Figure 7 in Section 4. Consequently, one may consider $S = M + K$ with $M \in \mathbb{R}^{N\times N}$ an M-matrix and $K \in \mathbb{R}^{N\times N}$ a small perturbation. For scalar-valued continuous problems this situation was considered in [9, Thm. 3.4]. Since the proof of this result does not depend on the scalar-valuedness of the problem, it also applies in the case of a vector-valued problem. Below, for $K \in \mathbb{R}^{N\times N}$ the norm $\|K\|_1$ denotes the subordinate matrix norm when $\mathbb{R}^N$ is endowed with the ℓ_1-norm.

Theorem 3.2. *Assume that $S = M + K$ with $M \in \mathbb{R}^{N\times N}$ a nonsingular* M*-matrix and with $K \in \mathbb{R}^{N\times N}$ a perturbation such that $\|K\|_1$ is sufficiently small. Then Algorithm 1 is well-defined, and $\lim_{k\to\infty}(u^k, \lambda_B^k) = (u^*, \lambda_B^*)$ with u^* the unique solution of (1) with corresponding multiplier λ_B^*. Moreover, the local convergence rate is superlinear.*

Proof. See the proof of [9, Thm. 3.4]. □

In our numerical test runs of Algorithm 1 for the vector-valued problems mentioned earlier, we typically observe that the iterates converge monotonically even in cases where $S = M + K$, with M and K as above. The next results give a theoretical justification for this observation in numerical practice. We denote

$$T_{II} = \sum_{l=1}^{\infty} (-M_{II}^{-1} K_{II})^l, \tag{26}$$

which is well-defined for $\|M_{II}^{-1} K_{II}\| < 1$. We further define

$$\begin{aligned} U_{IA} &= T_{II} M_{II}^{-1} M_{IA} + M_{II}^{-1} K_{IA} + T_{II} M_{II}^{-1} K_{IA}, \\ V_{II} &= T_{II} M_{II}^{-1}. \end{aligned}$$

Theorem 3.3. *Assume that $S = M + K$ with $M \in \mathbb{R}^{N\times N}$ a nonsingular* M*-matrix and $K \in \mathbb{R}^{N\times N}$ a perturbation such that $\|K\|_1$ is sufficiently small. For all $k \in \mathbb{N}$ and for all (I^k, A^k) defined by Algorithm 1 suppose that*

$$\left(M_{I^k I^k}^{-1} + V_{I^k I^k}\right) \geq 0 \tag{27a}$$

and there exists an index $k_0 \in \mathbb{N}$ with

$$\text{either } \left(M_{I^{k_0} I^{k_0}}^{-1} M_{I^{k_0} A^{k_0}} + U_{I^{k_0} A^{k_0}}\right) \leq 0 \tag{27b}$$

$$\text{or } u_B^{k_0} \text{ feasible.} \tag{27c}$$

Then

$$u_B^* \geq u_B^{k+1} \geq u_B^k \geq 0 \quad \text{for all } k \geq k_0.$$

Proof. The global and locally superlinear convergence of the iterates (u^k, λ_B^k) to (u^*, λ_B^*) follow from Theorem 3.2.

Let us turn to the monotonicity of the primal iterates. The smallness requirement for $\|K\|_1$ implies the existence of $T_{I^k I^k}$ for all k. Then the inverse of $S_{I^k I^k}$ exists and can be expressed as

$$S_{I^k I^k}^{-1} = (E_{I^k} + T_{I^k I^k}) M_{I^k I^k}^{-1}, \tag{28}$$

where E_{I^k} denotes the unit matrix in $\mathbb{R}^{|I^k|\times|I^k|}$. Utilizing the above identity in equation (25) yields

$$\begin{aligned} du^k_{I^k} &= -S^{-1}_{I^kI^k}S_{I^kA^k}du^k_{A^k} - S^{-1}_{I^kI^k}d\lambda^k_{I^k} \\ &= -(E_{I^k}+T_{I^kI^k})M^{-1}_{I^kI^k}(M_{I^kA^k}+K_{I^kA^k})du^k_{A^k} \\ &\quad -(E_{I^k}+T_{I^kI^k})M^{-1}_{I^kI^k}d\lambda^k_{I^k} \\ &= (M^{-1}_{I^kI^k}M_{I^kA^k}+U_{I^kA^k})(-du^k_{A^k}) + (M^{-1}_{I^kI^k}+V_{I^kI^k})(-d\lambda^k_{I^k}), \end{aligned} \tag{29}$$

where $du^k_{A^k} = u^{k+1}_{A^k} - u^k_{A^k}$, $d\lambda^k_{I^k} = \lambda^{k+1}_{I^k} - \lambda^k_{I^k}$, and analogously $du^k_{I^k}$.

(i) Let us assume that (27a) and (27c) are satisfied. Since $u^{k_0}_B \geq 0$, we have

$$I^{k_0} = \{i \in B : u^{k_0}_i + \alpha\lambda^{k_0}_i \geq 0\} \supseteq I^{k_0-1}, \quad A^{k_0} \subseteq A^{k_0-1}.$$

As a consequence, $u^{k_0+1}_{A^{k_0}} = u^{k_0}_{A^{k_0}}$ which implies $du^{k_0}_{A^{k_0}} = 0$. From (24) we recall that $d\lambda^{k_0}_{I^{k_0}} \leq 0$. Hence (29) and assumption (27a) imply that

$$u^{k_0+1}_{I^{k_0}} = u^{k_0}_{I^{k_0}} + (M^{-1}_{I^{k_0}I^{k_0}} + V_{I^{k_0}I^{k_0}})(-d\lambda^{k_0}_{I^{k_0}}) \geq u^{k_0}_{I^{k_0}} \geq 0.$$

Thus, we obtain $u^{k_0+1}_B \geq u^{k_0}_B \geq 0$, *i.e.* $u^{k_0+1}_B$ is feasible. By induction we have $u^{k+1}_B \geq u^k_B \geq 0$ for all $k \geq k_0$. The property $u^*_B \geq u^k_B$ for all $k \geq k_0$ is an immediate consequence of the convergence and monotonicity of $\{u^k_B\}_{k\geq k_0}$. This proves the assertion in case of (27a) and (27c).

(ii) Now we suppose that (27a) and (27b) hold. Let $V^{k_0} := \{i \in B : u^{k_0}_i < 0\}$. From (17b) we obtain $\lambda^{k_0}_{V^{k_0}} = 0$. Hence, $V^{k_0} \subseteq A^{k_0}$ which yields $u^{k_0+1}_{V^{k_0}} = 0$ again by (17b). Further note that $u^{k_0}_{I^{k_0}} \geq 0$. From (29) we infer

$$\begin{aligned} u^{k_0+1}_{I^{k_0}} &= u^{k_0}_{I^{k_0}} + (M^{-1}_{I^{k_0}I^{k_0}}M_{I^{k_0}A^{k_0}} + U_{I^{k_0}A^{k_0}})(-u^{k_0+1}_{A^{k_0}} + u^{k_0}_{A^{k_0}}) \\ &\quad + (M^{-1}_{I^{k_0}I^{k_0}} + V_{I^{k_0}I^{k_0}})(-d\lambda^{k_0}_{I^{k_0}}) \\ &\geq u^{k_0}_{I^{k_0}} + (M^{-1}_{I^{k_0}I^{k_0}}M_{I^{k_0}A^{k_0}} + U_{I^{k_0}A^{k_0}})u_{A^{k_0}} \\ &\geq u^{k_0}_{I^{k_0}} (\geq 0). \end{aligned}$$

For the last inequality we used $u^{k_0}_{A^{k_0}} \leq 0$ and (27b), and for the next to the last we utilized $d\lambda^{k_0}_{I^{k_0}} \leq 0$ (by (24)) and assumption (27a). Since $u^{k_0+1}_{A^{k_0}} = 0$ we have $u^{k_0+1}_B \geq 0$ and $u^{k_0+1}_B$ is feasible. Now we can argue as in case (i) to prove the assertion.

□

Conditions (27a) and (27b) hold if there exist constants $\epsilon_i \geq 0$, $i = 1,\dots 4$, with $0 \leq \epsilon_1 < \epsilon_2$,$0 \leq \epsilon_4 < \epsilon_3$, and ϵ_1 and ϵ_4 sufficiently small, such that for all $k \in \mathbb{N}$ and for all (I^k, A^k) defined by Algorithm 1

$$|\min(0,(V_{I^kI^k})_{ij})| \leq \epsilon_1, \quad (M^{-1}_{I^kI^k})_{ij} \geq \epsilon_2, \tag{30}$$

$$(M^{-1}_{I^{k_0}I^{k_0}}M_{I^{k_0}A^{k_0}})_{mn} \leq -\epsilon_3, \quad |\max(0,(U_{I^{k_0}A^{k_0}})_{mn})| \leq \epsilon_4, \tag{31}$$

for all $i,j \in I^k$ and $m \in I^{k_0}$, $n \in A^{k_0}$ for some $k_0 \in \mathbb{N}$. In fact, (30) implies that (27a) is satisfied, and from (31) we infer that (27b) is fulfilled. The conditions on $M^{-1}_{I^kI^k}$ and $M^{-1}_{I^kI^k}M_{I^kA^k}$ in (30) and (31) are satisfied for the standard finite difference or finite element discretizations of the differential operators appearing in the vector-valued Signorini problem and the symmetric crack problem; see Section 4. This is also true for many other practically relevant problems involving second order linear elliptic differential operators.

Thus, it is realistic to assume that ϵ_1 (ϵ_4) are small compared to ϵ_2 (ϵ_3). Note that the case $\epsilon_1 = 0$, and $\epsilon_4 = 0$ for all $k \in \mathbb{N}$ corresponds to the monotonicity assertion of Theorem 3.1. Indeed, in this case $K = 0$.

In our numerical practice typically assumptions (27a) and (27c) of Theorem 3.3 as well as assumption (32) for the initialization stated in the following theorem are satisfied. For convenience we recall that step 3 in the kth iteration of Algorithm 1 computes the solution $(u_B^{k+1}, \lambda_B^{k+1})$ of the system

$$\begin{aligned} &Su_B + \lambda_B = \tilde{f}, \\ &u_i = 0 \text{ for all } i \in A^k, \qquad \lambda_i^{k+1} = 0 \text{ for all } i \in I^k. \end{aligned}$$

Theorem 3.4. *Assume that $S = M + K$ with $M \in \mathbb{R}^{N\times N}$ a nonsingular* M*-matrix and with $K \in \mathbb{R}^{N\times N}$ a perturbation such that $\|K\|_1$ is sufficiently small. Suppose that Algorithm 1 is initialized with $\lambda_B^0 = 0$ and $u_B^0 = S^{-1}\tilde{f} \in \mathbb{R}^{|B|}$. If*

$$u_{I^0}^0 + \left(M_{I^0I^0}^{-1} M_{I^0A^0} + U_{I^0A^0}\right) u_{A^0}^0 \geq 0, \tag{32}$$

and (27a) *is satisfied, then the iterates $\{u_B^k\}$ converge monotonically to u_B^* with a locally superlinear rate. Moreover, $u_B^k \geq 0$ for all $k \geq 1$.*

Proof. The global and locally superlinear convergence of the iterates (u^k, λ_B^k) to (u^*, λ_B^*) follow from Theorem 3.2.

Let $V = \{i \in B : u_i^0 < 0\}$. Since $\lambda_B^0 = 0$ and $\alpha > 0$, we have

$$A^0 = \{i \in B : u_i^0 + \alpha\lambda_i^0 < 0\} = V \text{ and } I^0 = \{i \in B : u_i^0 \geq 0\}.$$

The update strategy of Algorithm 1 yields $u_i^1 = 0$ for all $i \in A^0$ and $\lambda_i^1 = 0$ for all $i \in I^0$. Thus, $du_i^0 = u_i^1 - u_i^0 > 0$ for all $i \in A^0$ and $d\lambda_i^0 = \lambda_i^1 - \lambda_i^0 = 0$ for all $i \in I^0$. From (29) and assumption (32) we infer

$$u_{I^0}^1 = u_{I^0}^0 + \left(M_{I^0I^0}^{-1} M_{I^0A^0} + U_{I^0A^0}\right) u_{A^0}^0 \geq 0.$$

As a consequence $u_B^1 \geq 0$, *i.e.* u_B^1 is feasible. Since, in addition, $\lambda_i^1 = 0$ for all $i \in I^0$, we have $I^0 \subseteq I^1$. This implies $A^1 \subseteq A^0$ and further $u_{A^1}^2 - u_{A^1}^1 = du_{A^1}^1 = 0$. From (24) we infer that $d\lambda_{I^1}^1 \leq 0$. Hence, assumption (27a) and (29) yield $u_{I^1}^2 - u_{I^1}^1 = du_{I^1}^1 \geq 0$. Therefore we have $u_B^2 \geq u_B^1 \geq 0$. By induction we get $u_B^{k+1} \geq u_B^k \geq 0$ for all $k \geq 1$ which proves the assertion. □

Observe that condition (32) is automatically satisfied in the case where S is an M-matrix. In general, (32) excludes the situation where $U_{ij} > 0$ and $-u_j^0 > 0$, for $i \in I^0$ and $j \in A^0$, are large.

4. Numerical results

In this section we report on numerical results obtain from an implementation of Algorithm 1. We also discuss the convergence results of Sections 2 and 3 on the numerical level. The test problems are a scalar-valued problem with a boundary obstacle, a vector-valued Signorini problem, and a symmetric crack problem. For the respective problem we give a brief description of the continuous formulation, an adequate discretization, and we relate the properties of the discrete operator L to the assumptions of our convergence theorems.

4.1. **The scalar-valued problem with a boundary obstacle.** Let $\Omega \subset \mathbb{R}^2$ denote a bounded domain with a Lipschitz boundary $\partial\Omega$. We assume that $\partial\Omega$ consists of disjoint and nonempty components Γ_d, Γ_n and Γ_c. Similar to (3) we introduce the objective functional

$$J(u) := \tfrac{1}{2}\int_\Omega \nabla u^\top \nabla u \, d\mathrm{x} - \int_{\Gamma_n} gu \, ds.$$

We define $H_1 := \{u \in H^1(\Omega) : u = 0 \text{ a.e. on } \Gamma_d\}$ and consider the minimization problem

$$\begin{aligned} &\text{minimize} \quad J(u) \quad \text{over } u \in H_1 \\ &\text{subject to} \quad u \geq \psi \text{ a.e. on } \Gamma_c, \end{aligned} \tag{33}$$

with the obstacle $\psi \in C^{0,1}(\overline{\Gamma_c})$ satisfying $\psi \leq 0$ a.e. on $\overline{\Gamma}_c \cap \overline{\Gamma}_d$. It is well-known that (33) admits a unique solution in H_1.

For our numerical calculations we take the following data: $\Omega = (0,1)^2$, $\Gamma_d = \{x = 1, 0 \leq y \leq 1\}$, $\Gamma_c = \{0 < x < 1, y = 0\}$, $\Gamma_n = \{x = 0, 0 \leq y \leq 1\} \cup \{0 < x < 1, y = 1\}$. We assume that $g(x,y) = -0.001$ on $\{x = 0, 0 \leq y \leq 1\}$ and $g(x,y) = 0$ on $\{0 < x < 1, y = 1\}$. The obstacle is prescribed on Γ_c and is defined as $\psi(x,y) = 0.004(\sin(\pi x) - 1)$. For convenience, in Figure 1 we illustrate the geometrical situation.

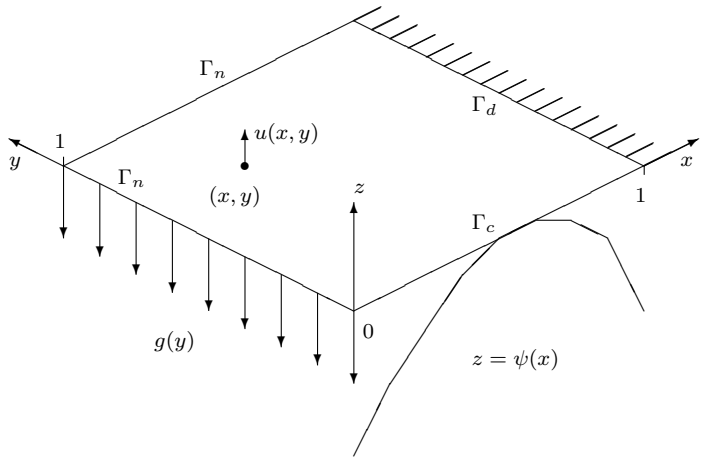

FIGURE 1. Example 1: The scalar-valued problem with the boundary obstacle.

For the discretization of H_1 we use linear finite elements which reflect the homogeneous Dirichlet boundary condition on Γ_d. The resulting matrix L is positive definite, and it is a nonsingular M-matrix. It corresponds to the discretization of the Laplace operator with homogeneous Dirichlet boundary conditions on Γ_d. As a consequence, the results of Theorem 3.1 apply. Let N denote the number of nodes in $\bar{\Omega}\backslash\Gamma_d$. We split $C = \{1, \ldots, N\}$ into B, which corresponds to nodes on Γ_c, and D, which contains the indices of nodes in $\bar{\Omega} \setminus (\Gamma_d \cup \Gamma_c)$. Let ψ_B denote the restriction of ψ onto the nodal points on Γ_c yielding $\psi_B \in \mathbb{R}^{|B|}$. The simple transformation $\tilde{u}_B = u_B - \psi_B$, $\tilde{f}_D = f_D - L_{DB}\psi_B$, and $\tilde{f}_B = f_B - L_{BB}\psi_B$ allows us to recast the discrete version of (33) as a problem of the type (1). For ease of reference we rename the quantities $\tilde{u}$ and $\tilde{f}$ by u and f again.

We initialize Algorithm 1 by the solution to the unconstrained problem, *i.e.*, $u^0 = L^{-1}f$. Further we set $\lambda_B^0 = 0$. The parameter α is fixed to $\alpha = 0.001(\max u^0 / \max \lambda^0)$. For the mesh-size $h = 0.025$, with $N = \frac{1}{h}(1 + \frac{1}{h}) = 1640$, Algorithm 1 terminates at iteration 5

with the solution to the problem. In Figure 2 and Figure 3 we display the iterates u_B^k and λ_B^k, respectively. The primal iterates in Figure 2 clearly exhibit the monotonous behavior as expected from Theorem 3.1. Further, combining both figures one finds $u_i^k \lambda_i^k = 0$ for all $i \in B$. Figure 4 shows the indicator functions for I^k for all iterations k. Again,

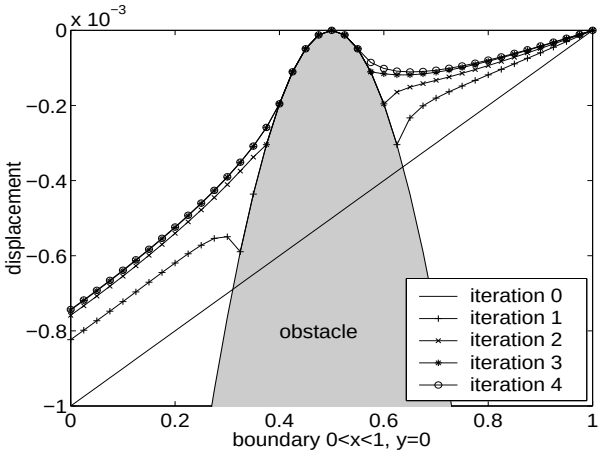

FIGURE 2. Iterations u_B^k of the displacements.

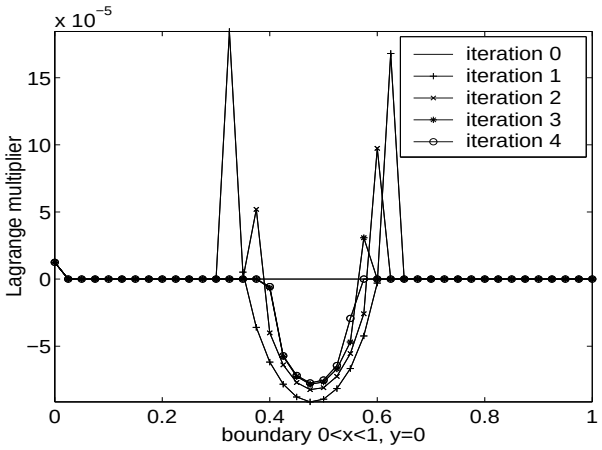

FIGURE 3. Iterations λ_B^k of the multipliers.

the monotonicity becomes evident. In fact, one observes $I^0 \subseteq I^1 \subseteq \ldots \subseteq I^4$. Finally, in Figure 5 we depict the quotients

$$q^k := \frac{\|u^{k+1} - u^*\|^2_{H^1(\Omega)}}{\|u^k - u^*\|^2_{H^1(\Omega)}} = \frac{\int_\Omega |\nabla(u^{k+1} - u^*)|^2 \, d\mathrm{x}}{\int_\Omega |\nabla(u^k - u^*)|^2 \, d\mathrm{x}}.$$

Here u^* denotes the numerical solution of the discrete problem. We present the results for several mesh-sizes h. The corresponding number of iterations until the successful termination of Algorithm 1 can be found in Table 1. For fixed mesh size h the results in Figure 5 suggest a superlinear convergence behavior of $\{u_B^k\}$. Further, it appears that q^k depends only moderately on the mesh-size h of the discretization. This observation is also supported by the results report on in Table 1

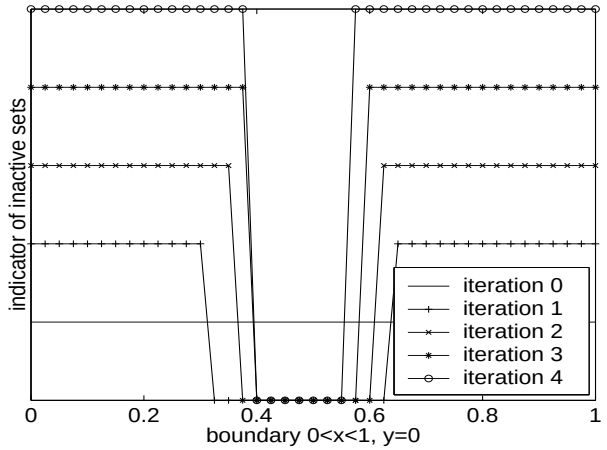

FIGURE 4. History of the indicator functions for I^k.

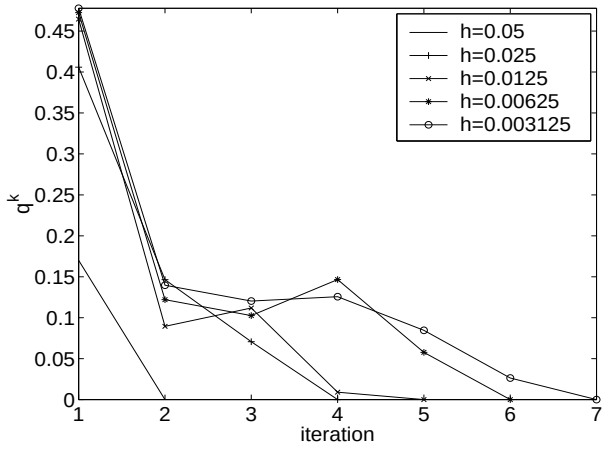

FIGURE 5. Behavior of q^k for various mesh-sizes.

h	0.05	0.025	0.0125	0.00625	0.003125
# it	2	4	5	6	7

TABLE 1. Number of iterations (# it) for different mesh-sizes h.

4.2. **The vector-valued Signorini problem with an obstacle.** With Ω, $\partial\Omega$, Γ_d, Γ_n and Γ_c as in Section 4.1 we consider the energy functional

$$J(u) := \tfrac{1}{2}b(u,u) - \langle g, u\rangle_{\Gamma_n},$$

where b denotes the bilinear form

$$\begin{aligned} b(u,v) = \int_\Omega [&(\kappa+1)u_{1,x}v_{1,x} + u_{1,y}v_{1,y} + u_{1,y}v_{2,x} + (\kappa-1)u_{1,x}v_{2,y} \\ &+ (\kappa-1)u_{2,y}v_{1,x} + u_{2,x}v_{1,y} + u_{2,x}v_{2,x} + (\kappa+1)u_{2,y}v_{2,y}]\,d\mathrm{x} \end{aligned}$$

and the pairing $\langle \cdot, \cdot \rangle_{\Gamma_n}$ is given by

$$\langle g, u \rangle_{\Gamma_c} = \int_{\Gamma_c} g^\top u \, ds.$$

Above u, v and g denote the vectors $u = (u_1, u_2)^\top$, $v = (v_1, v_2)^\top$ and $g = (g_1, g_2)^\top \in (L^2(\Gamma_n))^2$. Further subscripts x and y denote differentiation with respect to the respective variables. Note that the bilinear form b corresponds to the Lamé system

$$\begin{aligned} -\Delta u_1 - \kappa(u_{1,x} + u_{2,y})_x &= 0, \\ -\Delta u_2 - \kappa(u_{1,x} + u_{2,y})_y &= 0. \end{aligned} \tag{34}$$

Here and above $\kappa = (\mu + \lambda)/\mu$ depends on the Lamé coefficients $\mu > 0$ and λ with $\mu + \lambda > 0$. See e.g. [12] for more details on equations of the Lamé type.

We define $H_2 := \{u \in (H^1(\Omega))^2 : u_1 = u_2 = 0 \text{ a.e. on } \Gamma_d\}$ and consider the minimization problem

$$\begin{aligned} &\text{minimize} \quad J(u) \quad \text{over } u \in H_2 \\ &\text{subject to} \quad u_2 \geq \psi \text{ a.e. on } \Gamma_c, \end{aligned} \tag{35}$$

with the obstacle $\psi \in C^{0,1}(\overline{\Gamma_c})$ satisfying $\psi \leq 0$ a.e. on $\overline{\Gamma}_c \cap \overline{\Gamma}_d$. The bilinear form b is elliptic on H_2, and it is well-known that (33) admits a unique solution in H_2. In Figure 6 we give a graphical account for the problem under consideration.

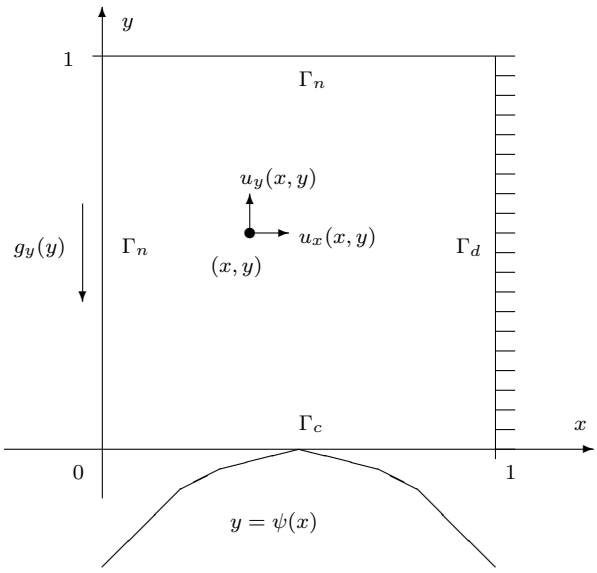

FIGURE 6. Example 2: The vector-valued Signorini problem with an obstacle.

Note that u_1 and u_2 describe horizontal displacements in the plane $z = 0$.

We discretize H_2 by linear finite elements which satisfy the homogeneous Dirichlet boundary conditions satisfied on Γ_d. Let $\hat{N}$ denote the total number of nodes. We suppose that the vector of unknowns is ordered as

$$u = ((u_1)_1, (u_2)_1, (u_1)_2, (u_2)_2, (u_1)_3, (u_2)_3, \ldots, (u_1)_{\hat{N}}, (u_2)_{\hat{N}})^\top \in \mathbb{R}^{2\hat{N}}.$$

The stiffness matrix $L \in \mathbb{R}^{2\hat{N} \times 2\hat{N}}$ is obtained from the finite element representation of the bilinear form b. Due to the ellipticity property of b, the matrix L is positive definite.

But, in contrast to the scalar-valued case in Section 4.1, it is *not* an M-matrix. We denote by B the set of indices corresponding to $(u_2)_i$ with nodal points located on Γ_c. Further $C = \{1, \ldots, 2\hat{N}\}$ and $D = C \setminus B$. Let ψ_B denote the restriction of ψ to the nodal points on Γ_c, and let $f \in \mathbb{R}^{2\hat{N}}$ with $f^\top u$ representing the discretization of $\langle g, u\rangle_{\Gamma_c}$. Setting $N = 2\hat{N}$ and performing the simple transformation of $u_B \geq \psi_B$ into $u_B \geq 0$ as in Section 4.1, we arrive at a problem of type (1) which is equivalent to an appropriate discretization of (35).

Let us categorize the problem under consideration with respect to our convergence results in Section 3. As noted above, L fails to be an M-matrix. Also the Schur complement S, which corresponds to the iteration matrix with respect to u_B, lacks the M-matrix property. However, the assumption of Theorem 3.2 is likely to be satisfied. In fact, $S_{ii} > 0$ for all $i \in B$. Let $K_{ij} = \max(0, S_{ij})$ with $i, j \in B$ and $i \neq j$, and define $K_{ii} = 0$ for all $i \in B$. Then one obtains $S = M + K$ with M a nonsingular M-matrix. In Figure 7 we display the ℓ_1-norm of S and the ℓ_1-norm of K multiplied by 100, respectively. We plot these norms in dependence on κ and the mesh-size h. The problem data are specified below. First note that the norms depend only slightly on h for sufficiently small h. With respect

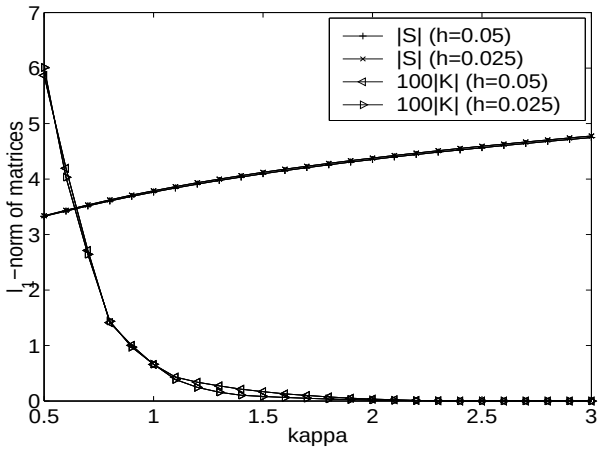

FIGURE 7. ℓ_1-norms of S (upper graphs) and K (lower graphs).

to κ one finds that the larger κ becomes the smaller the ℓ_1-norm of K is. This latter fact is interesting since for $\kappa = 0$ the Lamé system admits no solution. From Figure 7 we can see that $\|K\|_1$ becomes large when $\kappa > 0$ is small. On the other hand, we typically have $\|K\|_1/\|S\|_1 \leq 0.01$ for $\kappa \geq 1$. This clearly indicates that K can be considered as a small perturbation of M, and the assumptions of Theorem 3.2 are likely to be satisfied. In fact, the results reported on below show that Algorithm 1 is globally convergent with a fast local rate.

For our test of Algorithm 1 discussed below, we recall that Ω, Γ_d, Γ_n, and Γ_c are like in Section 4.1. Unless otherwise stated, we use $\kappa = 1$ and $\alpha = 0.001(\max u^0 / \max \lambda^0)$. Further we have

$$g = (g_1, g_2)^\top = \begin{cases} (0, -0.001)^\top & \text{on } \{x = 0, 0 \leq y \leq 1\} \\ (0, 0)^\top & \text{on } \{0 < x < 1, y = 1\} \end{cases}$$

and $\psi(x, y) = 0.004(\sin(\pi x) - 1)$ for $(x, y) \in \Gamma_c$. The algorithm is initialized by $u^0 = L^{-1} f$ and $\lambda_B = 0$.

In Figures 8–10 we show the iterates $(u_2^k)_B$, the multipliers λ_B^k, and the indicator functions of I^k for all iterations k.

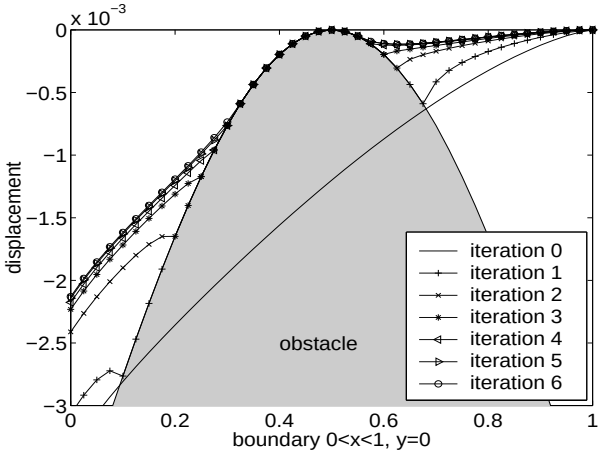

FIGURE 8. Iterations $(u_2^k)_B$ of the displacements.

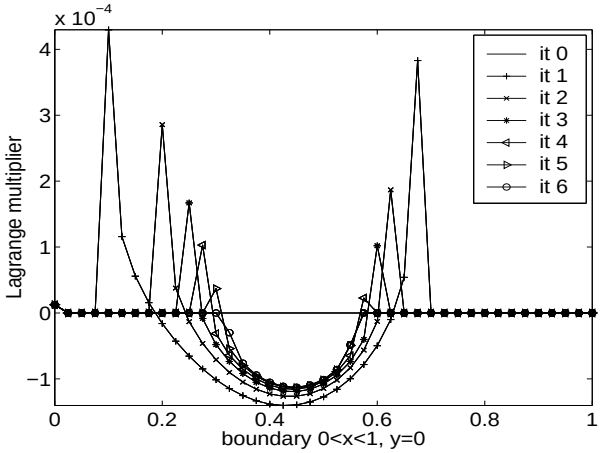

FIGURE 9. Iterations λ_B^k of the multipliers.

The conclusions concerning the monotonicity of $\{u_B^k\}$ and of the indicator functions of I^k are as in Section 4.1. Further, we checked the conditions of Theorem 3.4. In fact, (27a) and (28) are satisfied; see Figure 8. As a consequence we observe monotone convergence of the primal iterates $(u_2^k)_B$. In Figure 11 we depict the quotients

$$q^k = \frac{\|u^{k+1} - u^*\|^2_{(H^1(\Omega))^2}}{\|u^k - u^*\|^2_{(H^1(\Omega))^2}}.$$

for various mesh-sizes h. The results presented in Figure 11 clearly indicate a fast local convergence property of Algorithm 1 for solving the vector-valued Signorini problem.

Next we investigate the dependence of the number of iterations until successful termination for various mesh-sizes h and coefficients $\kappa > 0$. The corresponding results can be found in Table 2 and Table 3. The figures in Table 2 indicate a moderate dependence

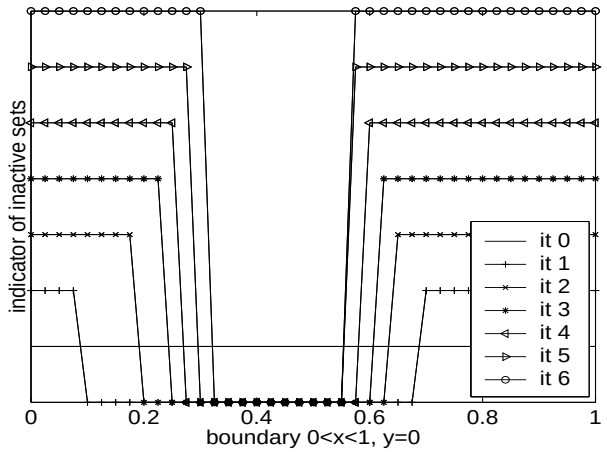

FIGURE 10. History of the indicator functions for I^k.

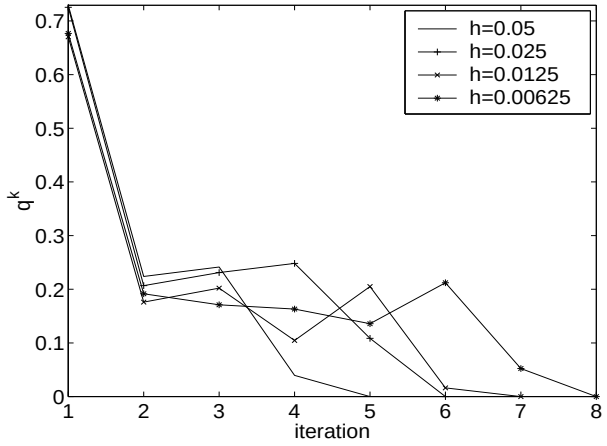

FIGURE 11. Behavior of q^k for various mesh-sizes.

h	0.05	0.025	0.0125	0.00625
# it	5	6	7	8

TABLE 2. Number of iterations (# it) for different mesh-sizes h.

κ	0.5	1.0	1.5	2	2.5	3
# it	6	6	6	5	6	6

TABLE 3. Number of iterations (# it) for different coefficients $\kappa > 0$.

of the number of iterations #it on the mesh-size h of the discretization. From Table 3 we see that #it is essentially independent of $\kappa > 0$. We also tested alternative initializations and found that Algorithm 1 converged for all of our initializations.

Finally, we test whether our observations depend on the structure of the active respectively inactive sets. In Figure 12 we display the iterates $(u_2^k)_B$ for the obstacle $\psi(x, y) = 0.004(\sin(9\pi x) - 1)$ with $(x, y) \in \Gamma_c$. As can be seen from Figure 13, which

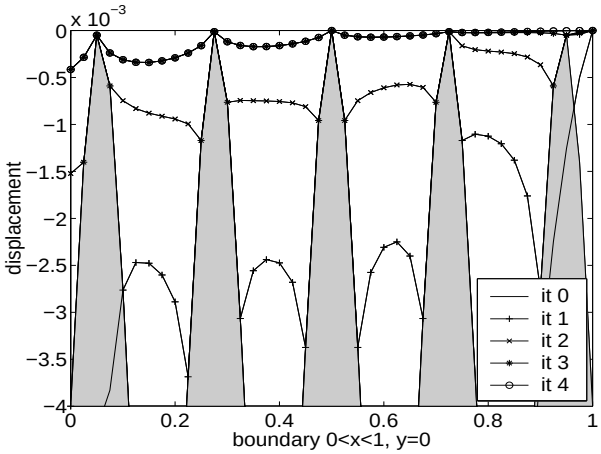

FIGURE 12. Iterations $(u_2^k)_B$ of the displacements.

shows the indicator functions of I^k for all iterations k, the inactive set consists of several disjoint components. Like in the previous case, the iterates depicted in Figure 12 con-

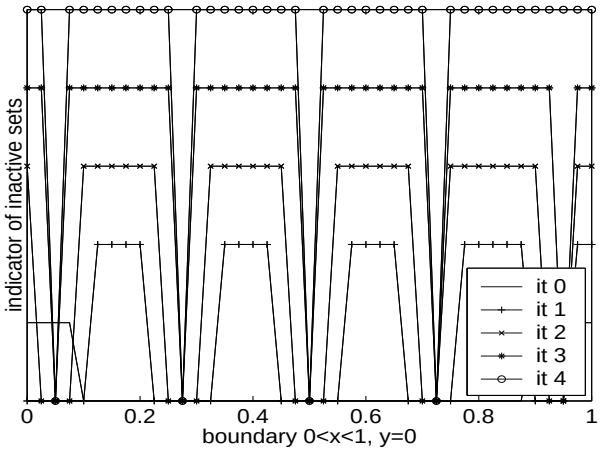

FIGURE 13. History of the indicator functions for I^k.

verge monotonically. The latter example shows that Algorithm 1 copes successfully with complex active and inactive set structures.

4.3. **The symmetric crack problem.** Symmetric crack problems are commonly considered in fracture mechanics when investigating the mode-1 model of a crack. We consider the following version. Let $O \in \mathbb{R}^2$ be a bounded domain which is symmetric with respect to the x-axis, *i.e.*,

$$\overline{O} = \overline{O}^+ \cup \overline{O}^-, \quad O^\pm = O \cap \{\pm y > 0\}, \quad \overline{O}^+ \cap \overline{O}^- = \Gamma_s.$$

Here $\Gamma_s \subset \{y = 0\}$ is an interface between O^+ and O^-. We assume that the crack Γ_c occupies a part of Γ_s. Under symmetric boundary conditions imposed on ∂O one can

consider the following crack problem which is stated in $\Omega := O^+$ only:

$$\begin{aligned} &\text{minimize} \quad J(u) \quad \text{over } u \in H_2 \\ &\text{subject to} \quad u_2 \geq 0 \text{ a.e. on } \Gamma_c, \quad u_2 = 0 \text{ a.e. on } \Gamma_s \setminus \Gamma_c, \end{aligned} \tag{36}$$

where the objective functional is defined by

$$J(u) := \tfrac{1}{2} b(u,u) - \langle g, u \rangle_{\Gamma_n},$$

see Section 4.2 for the precise definitions of J and H_2. Note that $u_2 \geq 0$ a.e. on Γ_c is called non-penetration condition, since it is related to non-penetration of opposite crack faces. Problem (36) is well-defined and admits a unique solution in H_2.

We apply a finite element discretization similar to the one in the previous section. The resulting problem is of the type (1) with the additional condition $(u_2)_i = 0$ for indices i which belong to nodal points on $\Gamma_s \setminus \Gamma_c$. Since J and H_2 are like in Section 4.2, the matrix L is positive definite, but it is not an M-matrix. A similar investigation to the one carried out in the previous section yields that the Schur complement S can be expressed as $S = M + K$ with M a nonsingular M-matrix and K a sufficiently small perturbation. Consequently, Theorems 3.2 and 3.3 are likely to be applicable.

Below we report on the results obtain from Algorithm 1 applied to the following example:

$$\begin{aligned} &\Omega = \{0 < x < 1, 0 < y < 0.5\}, \quad \Gamma_d = \{x = 1, 0 \leq y \leq 0.5\}, \\ &\Gamma_s = \{0 < x < 1, y = 0\}, \quad \Gamma_n = \Gamma_n^1 \cup \Gamma_n^2, \\ &\Gamma_n^1 = \{x = 0, 0 \leq y \leq 0.5\}, \quad \Gamma_n^2 = \{0 < x < 1, y = 0.5\}. \end{aligned}$$

We consider a so-called multi-crack Γ_c consisting of three pieces:

$$\begin{aligned} &\Gamma_c = \Gamma_c^1 \cup \Gamma_c^2 \cup \Gamma_c^3, \quad \Gamma_c^1 = \{0 < x < 0.1, y = 0\}, \\ &\Gamma_c^2 = \{0.2 < x < 0.8, y = 0\}, \quad \Gamma_c^3 = \{0.9 < x < 1, y = 0\}. \end{aligned}$$

The geometrical situation is depicted in Figure 14.

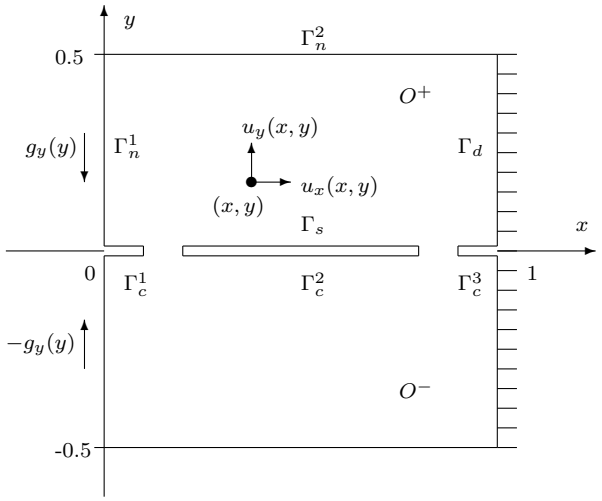

FIGURE 14. Example 3: The symmetric problem with the multi-crack.

We choose $\kappa = 1$ and $\alpha = 0.001(\max u^0 / \max \lambda^0)$. Further we have

$$g = (g_1, g_2)^\top = \begin{cases} (0, -0.001)^\top & \text{on } \Gamma_n^1, \\ (0, 0)^\top & \text{on } \Gamma_n^2. \end{cases}$$

In Figure 15–17 we show, for $h = 0.025$, $(u_2^k)_B$, λ_B^k and the indicator functions of the inactive sets I^k for all iterations k. The algorithm stops after 4 iterations at the optimal

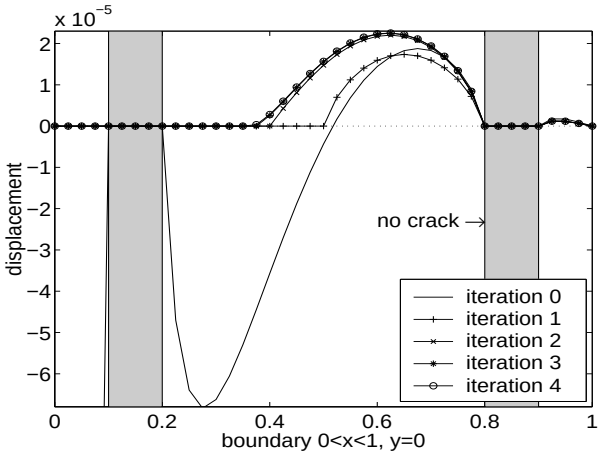

FIGURE 15. Iterations $(u_2^k)_B$ of the displacements.

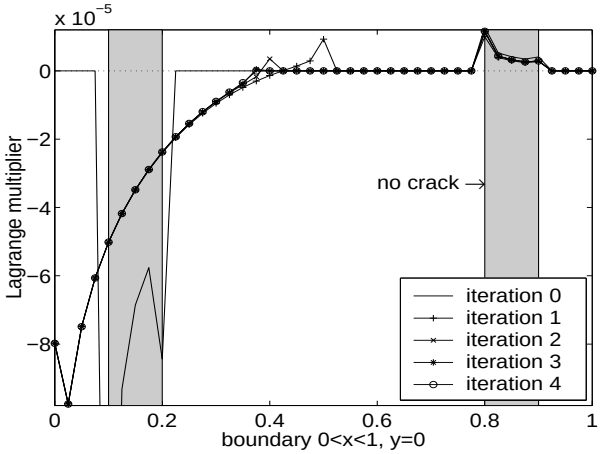

FIGURE 16. Iterations λ_B^k of the multipliers.

solution. From Figure 16 it can be seen that the Lagrange multipliers takes on negative as well as positive values in nodes on $\Gamma_s \setminus \Gamma_c$. This is due to the fact that $(u_2)_i = 0$ in these nodes is kept as an explicit constraint. Since it is of equality type, the corresponding multipliers admit no sign condition. Figure 17 indicates that all nodes are active on Γ_c^1. The nodes on Γ_c^3 are all inactive, and on Γ_c^2 there are both active nodes and inactive ones. This behavior can also be inferred from Figure 15. In Table 4 we report on the #it for various mesh-sizes h. Similar to the problem considered in the previous section, we conclude that #it depends only moderately on the mesh-size of discretization.

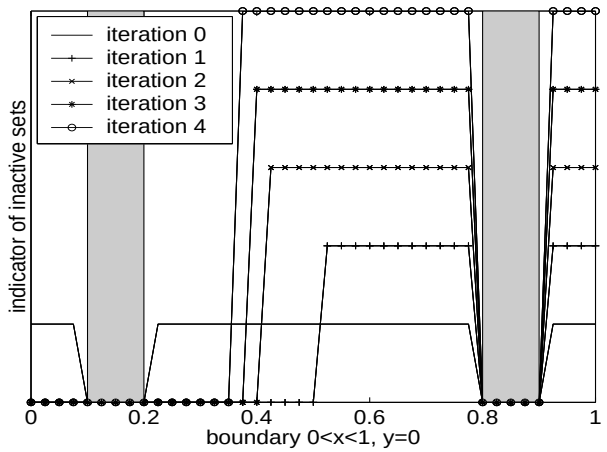

FIGURE 17. History of the indicator functions for I^k.

h	0.05	0.025	0.0125	0.00625
# it	3	4	5	7

TABLE 4. Number of iterations (# it) for different mesh-sizes h.

Finally, let us point out that this example illustrates that there is no maximum principle for the Lamé problem in general. In fact, the negative load g_2 applied on Γ_n^1 yields a positive displacement u_2 on Γ_c. This behavior would not occur in case of the existence of a maximum principle.

REFERENCES

[1] M. Bergounioux, K. Ito, K. Kunisch, Primal-dual strategy for constrained optimal control, SIAM J. Control Optim. 37 (1999), pp. 1176–1194.

[2] A. Berman, R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* Computer Science and Scientific Computing Series, Academic Press, New York, 1979.

[3] D. P. Bertsekas, Projected Newton methods for optimization problems with simple constraints, SIAM J. Control Optim. 20 (1982), pp. 221–246.

[4] X. Chen, Z. Nashed, L. Qi, Smoothing methods and semismooth methods for nondifferentiable operator equations, SIAM J. Numer. Anal. 38 (2000), pp. 1200–1216.

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

[6] P. Gill, W. Murray, M. H. Wright, *Practical optimization*, Academic Press, New York, 1999 (reprint).

[7] D. Goldfarb, A. Idnani, A numerically stable dual method for solving strictly convex quadratic programs, Math. Prog. 21 (1983), pp. 1–33.

[8] W. W. Hager, The dual active set algorithm, Advances in optimization and parallel computing, 1992, pp. 137–142.

[9] M. Hintermüller, K. Ito, K. Kunisch, The primal-dual active set strategy as a semi-smooth Newton method, SIAM J. Optim. 13 (2003), pp. 865–888.

[10] M. Hintermüller, V. Kovtunenko, K. Kunisch, The primal-dual active set method for a crack problem with non-penetration, IMA J. Appl. Math., accepted for publication.

[11] A.M. Khludnev, V.A. Kovtunenko, *Analysis of Cracks in Solids*, WIT-Press, Southampton, Boston, 2000.

[12] N. Kikuchi, T. Oden, *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*, SIAM, Philadelphia, 1988.
[13] D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization*, Kluwer Publishers, Dordrecht, 2002.
[14] V. Kovtunenko, Numerical simulation of the nonlinear crack problem with non-penetration, Math. Meth. Appl. Sci., to appear.
[15] B. Kummer, Newton's method for nondifferentiable functions, in *Advances in mathematical optimization*, Akademie-Verlag, Berlin, 1988, pp. 114–125.
[16] B. Kummer, Generalized Newton and NCP-methods: convergence, regularity, actions, Discuss. Math. Differ. Incl. Control Optim. 20 (2000), pp. 209–244.
[17] C. Lin, J. J. Moré, Newton's method for large bound-constrained optimization problems, preprint, Argonne National Laboratory, Mathematics and Computer Science Division, 1999.
[18] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, SIAM J. Control Optim. 15 (1977), pp. 959–972.
[19] J. J. Moré, G. Toraldo, Algorithms for bound constrained quadratic programming problems, Num. Math. 55 (1989), pp. 377–400.
[20] J. J. Moré, G. Toraldo, On the solution of large quadratic programming problems with bound constraints, SIAM J. Optim. (1) 1 (1991), pp. 93–113.
[21] J.-S. Pang, Newton's method for B-differentiable equations, Math. Oper. Res. 15 (1990), pp. 311–341.
[22] J.-S. Pang, L. Qi, Nonsmooth equations: Motivation and algorithms, SIAM J. Optim. 3 (1993), pp. 443–465.
[23] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Math. Oper. Res. 18 (1993), pp. 227–244.
[24] L. Qi, J. Sun, A nonsmooth version of Newton's method, Math. Prog. 58 (1993), pp. 353–367.
[25] S. M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994), pp. 291–305.
[26] R. J. Vanderbei, *Linear Programming: Foundations and Extensions*, Kluwer, Dordrecht, 1997.
[27] S. J. Wright, *Primal-Dual Interior Point Methods*, SIAM, Philadelphia, 1997.
[28] Y. Ye, *Interior Point Algorithms: Theory and Analysis*, Wiley, New York, 1997.