SIAM J. NUMER. ANAL.
Vol. 40, No. 6, pp. 2292–2310
© 2003 Society for Industrial and Applied Mathematics

A MONOTONIC METHOD FOR THE NUMERICAL SOLUTION OF SOME FREE BOUNDARY VALUE PROBLEMS*

RAPHAÈLE HERBIN†

Abstract. This work presents an efficient monotonic algorithm for the numerical solution of the obstacle problem and the Signorini problems, when they are discretized either by the finite element method or by the finite volume method. The convergence of this algorithm applied to the discrete problem is proven in both cases.

Key words. variational inequalities, iterative algorithm, obstacle problem, Signorini problem, finite element and finite volume methods

AMS subject classifications. 65K10, 49A29

PII. S0036142900380558

1. Introduction.

We are interested here in the numerical solution of some free boundary problems which are discretized by the finite element or the finite volume method. We introduce an efficient monotonic algorithm which applies to both the obstacle problem and the Signorini problem.

The obstacle problem is one of the simplest unilateral problems; it arises when modelling a constrained membrane in the classical linear elasticity theory. Signorini boundary conditions may be encountered in fluid mechanics and heat transfer problems when modelling, for instance, the flow through semipermeable boundaries. They are also encountered in contact problems in elasticity. The Signorini boundary conditions which we deal with here arise from modelling the so-called triple point of an electrochemical reaction (see [23]) and involve a diffusion operator. Both the obstacle and the Signorini problems may be written as variational inequalities.

The obstacle problem appeared in the mathematical literature in the work of Stampacchia [28] (see also [29], [30]), and the first rigorous analysis of a class of Signorini problems was published in 1963 by Fichera [12], [13]. The mathematical analysis including the study of existence, uniqueness, and regularity of the solution for the obstacle problem and Signorini problem may be found in [24], [25], and [7].

The obstacle problem and the Signorini problem are classically discretized by the finite element method formulated in [21], [16]; see also [27], [2], [1], [3] for more recent work (some of them subsequent to the submission of this paper) on elastic contact problems. In the case of diffusion problems, with which we are concerned here, a cell-center finite volume scheme was also recently applied and shown to converge [19].

The approximate problem can be solved by a duality method [16], [15], [22]. In [16], a point overrelaxation method with projection is also studied and found to be cheaper in terms of computational cost than the duality method. Another candidate for the resolution of the approximated Signorini problem is the penalty method (see [21] and references therein); it has the disadvantage of yielding ill-conditioned systems, while our algorithm deals only with submatrices of the whole discretization matrix.

*Received by the editors November 7, 2000; accepted for publication (in revised form) June 3, 2002; published electronically January 14, 2003.
http://www.siam.org/journals/sinum/40-6/38055.html
†Université de Provence, CMI, 39 rue Joliot Curie, 13453 Marseille cedex 13, France (herbin@cmi.univ-mrs.fr).

We present here a particularly simple iterative monotonic algorithm that is inspired by a procedure used for multiphase flow modelling [8]. We show that it may be applied to the finite linear element approximation of the obstacle problem and the finite volume discretization of both the obstacle problem and the Signorini problem. In each case, we prove the monotonicity of the algorithm and its convergence in a finite number of iterations towards the exact solution to the discrete problem.

2. The obstacle problem. We consider here the so-called obstacle problem, which arises for instance in the modelling of contact problems (see [7]):

$$\left\{\begin{array}{l} u \in \mathcal{K} = \{v \in H_0^1(\Omega),\ v \leq \psi \text{ on } \Omega\}, \quad \text{satisfying} \\ \displaystyle\int_\Omega \nabla u(x) \cdot \nabla (v-u)(x) dx \geq \int_\Omega f(x)(v-u)(x) dx \qquad \forall v \in \mathcal{K}, \end{array}\right. \tag{1}$$

where the following holds.

Assumption 2.1.

1. Ω is a bounded open polygonal subset of $\mathbb{R}^d$, with $d = 2$ or 3.
2. $f \in L^2(\Omega)$ and $\psi \in H^1(\Omega) \cap C(\Omega)$ and $\psi \geq 0$ a.e. in the neighborhood of $\partial\Omega$.

Under these assumptions, it is well known that there exists a unique solution to problem (1), thanks to Stampacchia's theorem. Indeed, the set $\mathcal{K}$ is nonempty since $\min(0, \psi)$ belongs to $\mathcal{K}$. Furthermore, it is now classical that the solution to problem (1) belongs to $H^2(\Omega)$ (see [4]). Thanks to this H^2 regularity of the solution, it is easily shown that the variational inequality (1) can be written as a free boundary problem in the following way.

THEOREM 2.1. *Under Assumption* 2.1, *if u is a solution to the free boundary problem*

$$\left\{\begin{array}{l} u \in H^2(\Omega) \cap H_0^1(\Omega), \quad \textit{satisfying} \\ u \leq \psi \quad \textit{a.e. on } \Omega, \\ \Delta u + f \geq 0 \quad \textit{a.e. on } \Omega, \\ (\Delta u + f)(\psi - u) = 0 \quad \textit{a.e. on } \Omega, \end{array}\right. \tag{2}$$

then u is a solution to Problem (1).

Conversely, if $\psi \geq 0$ a.e. on Ω and u is a solution to Problem (1), *then u is a solution to Problem* (2).

We shall study the monotonic algorithm for both the finite element and the finite volume discretization of the above problem. Let us first start with the finite element method.

2.1. Approximation by the finite element method. Let $\mathcal{T}$ denote a "classical" triangulation of Ω (see, e.g., [6]).

DEFINITION 2.2 (triangulation $\mathcal{T}$ of Ω). *Let $\mathcal{T}$ be a finite set of triangles if $d = 2$, or tetrahedra if $d = 3$, such that*

(i) *$T \subset \overline{\Omega}$ $\forall T \in \mathcal{T}$, and $\cup_{T \in \mathcal{T}} = \overline{\Omega}$;*

(ii) *for any $(T_1, T_2) \in \mathcal{T}^2$ with $T_1 \neq T_2$, either the $(d-1)$-dimensional Lebesgue measure of $\overline{T}_1 \cap \overline{T}_2$ is 0, or T_1 and T_2 have only a whole common edge (or face if $d = 3$).*

Let Σ be the set of vertices of triangles (tetrahedra) of $\mathcal{T}$ which belong to Ω (i.e., do not lie on the boundary) and $N = \mathrm{card}(\Sigma)$.

The set $H_0^1(\Omega)$ is classically approximated by

$$V_h = \{v \in H_0^1(\Omega) \cap C^0(\overline{\Omega}),\ v_{|\partial\Omega} = 0,\ v_{|T} \in P_1\}, \tag{3}$$

where $v_{|\partial\Omega}$ is the trace of v on $\partial\Omega$, $v_{|T}$ denotes the restriction of v to T, and P_1 the space of polynomials in x_1 and x_2 of degree less than or equal to one. Assuming that $\Sigma = \{s_i,\ i \in \{1,\dots,N\}\}$, let $(\varphi_i)_{i\in\{1,\dots,N\}}$ be the N basis functions of V_h such that $\varphi_i(s_i) = 1$ and $\varphi_i(s_j) = 0\ \forall i \neq j$; notice that the functions φ_i are linear on each triangle for which s_i is a vertex.

We then consider the following approximate problem:

$$\left\{ \begin{array}{l} \tilde{u} \in K_h = \{v \in V_h,\ v(s) \le \psi(s)\ \forall s \in \Sigma\}, \quad \text{satisfying} \\ \displaystyle\int_\Omega \nabla\tilde{u}(x)\cdot\nabla(v-\tilde{u})(x)dx \ge \int_\Omega f(x)(v-\tilde{u})(x)dx \qquad \forall v \in K_h. \end{array} \right. \tag{4}$$

By Stampacchia's theorem, problem (4) has a unique solution. Indeed, the set K_h is nonempty since the function $\min(\sum_{i=1,N}\psi_i\varphi_i, 0)$ belongs to K_h. Error estimates for the approximate finite element solution of the elliptic variational inequalities can be found in Falk [10], Mosco and Strang [26], Glowinski, Lions, and Trémolières [16], Ciarlet [6], Brezzi, Hager, and Raviart [5], and Falk and Mercier [11]. Error estimates of order 1 in the discretization step are known for the discretization of the obstacle problem using linear elements [10], [5].

Remark 2.1. In the present paper we shall use linear finite elements, and we shall avoid higher order finite elements for three reasons. First, it is well known that the maximum principle does not hold for higher order finite elements. In our underlying application, where the unknown is a concentration, it is absolutely necessary that it hold, since the electrical current, which we need to compute, depends on the logarithm of the concentration. The discrete maximum principle must therefore hold. Second, the precision obtained with the linear elements is, in general, largely sufficient for diffusion problems such as the one we consider. Third, our proof of convergence of the monotonic algorithm makes heavy use of the discrete maximum principle, and it is therefore not clear how the algorithm would behave in a setting where the maximum principle does not hold (in the case of higher order finite elements, or for the elasticity problem, for instance).

The monotonic algorithm is derived on a "strong formulation" of problem (4), which is easily shown to be equivalent to (4) as follows.

PROPOSITION 2.3. *Let $\tilde{u}$ be the unique solution to problem* (4) *and let $U = (u_1,\dots,u_N) \in \mathbb{R}^N$ be defined by $u_i = \tilde{u}(s_i)\ \forall i \in \{1,\dots,N\}$; then $\tilde{u}$ is a solution to* (4) *if and only if U is a solution to the following complementarity problem:*

$$\left\{ \begin{array}{l} u_i \le \psi_i \quad \forall i \in \{1,\dots,N\}, \\ (AU)_i \le F_i \quad \forall i \in \{1,\dots,N\}, \\ ((AU)_i - F_i)(\psi_i - u_i) = 0 \quad \forall i \in \{1,\dots,N\}, \end{array} \right. \tag{5}$$

with $\psi_i = \psi(s_i)$, $F_i = \int_\Omega f(x)\varphi_i(x)dx$, and A being the square matrix of order N whose coefficients satisfy $a_{i,j} = \int_\Omega \nabla\varphi_i(x)\cdot\nabla\varphi_j(x)dx$; therefore

$$(AU)_i = \sum_{j=1}^N u_j \int_\Omega \nabla\varphi_j(x)\cdot\nabla\varphi_i(x)dx.$$

Problem (5) is nonlinear. We shall solve it by an iterative algorithm that is adapted from a similar one used for multiphase flows in porous media [8]. Let us first remark that for $i \in \{1,\dots,N\}$ the last equation in (5) is equivalent to $(AU)_i = F_i$ or $u_i = \psi_i$. Therefore, there exist two disjoint subsets of $\{1,\dots,N\}$ such that $u_i = \psi_i$

and $(AU)_i \leq F_i$ for any i in the first subset, and $(AU)_i = F_i$ and $u_i \leq \psi_i$ for i in the second subset.

If we knew two disjoint subsets $\mathcal{J}$ and $\mathcal{I}$ of $\{1, \dots, N\}$ such that

$$\begin{aligned} u_i &\leq \psi_i \quad \forall i \in \mathcal{J}, \\ (AU)_i &\leq F_i \quad \forall i \in \mathcal{I}, \end{aligned}$$

then problem (5) would be solved by the solution of the following linear system:

$$\left\{ \begin{array}{l} u_i = 0 \quad \forall i \in \{1, \dots, N\} \quad \text{s.t. } s_i \in \partial\Omega \cap K_h, \\ (AU)_i = F_i \quad \forall i \in \mathcal{J}, \\ u_i = \psi_i \quad \forall i \in \mathcal{I}. \end{array} \right. \tag{6}$$

The algorithm which we propose here assumes the sets $\mathcal{J}$ and $\mathcal{I}$ to be known at each iteration, solves problem (6), and corrects the sets $\mathcal{J}$ and $\mathcal{I}$ by looking for the nodes where the corresponding constraints are violated. Let us write this algorithm as follows.

MONOTONIC ALGORITHM, OBSTACLE PROBLEM, FINITE ELEMENT DISCRETIZATION.

- Initialization. Let $\mathcal{I}^{(0)}$ and $\mathcal{J}^{(0)}$ be such that

$$\mathcal{I}^{(0)} \subset \{1, \dots, N\} \text{ and } \mathcal{J}^{(0)} = \{1, \dots, N\} \setminus \mathcal{I}^{(0)}. \tag{7}$$

- Step (j), $j \geq 0$. For given sets $\mathcal{I}^{(j)}$ and $\mathcal{J}^{(j)} = \{1, \dots, N\} \setminus \mathcal{I}^{(j)}$, let $U^{(j)} = (u_1^{(j)}, \dots, u_N^{(j)}) \in \mathbb{R}^N$ be the solution to the following set of equations:

$$\left\{ \begin{array}{l} (AU^{(j)})_i = F_i \quad \forall i \in \mathcal{J}^{(j)}, \\ u_i^{(j)} = \psi_i \quad \forall i \in \mathcal{I}^{(j)}, \end{array} \right. \tag{8}$$

where $(AU^{(j)})_i = \sum_{k=1}^N u_k^{(j)} \int_\Omega \nabla\varphi_k(x) \cdot \nabla\varphi_i(x) dx$ and $F_i = \int_\Omega f(x)\varphi_i(x) dx$. Let $\mathcal{I}^{(j+1)}$ and $\mathcal{J}^{(j+1)}$ be defined by

$$\begin{array}{ll} \mathcal{I}^{(j,0)} = \{i \in \mathcal{I}^{(j)}; AU_i^{(j)} \leq F_i\}, & \mathcal{I}^{(j,1)} = \mathcal{I}^{(j)} \setminus \mathcal{I}^{(j,0)}, \\ \mathcal{J}^{(j,0)} = \{i \in \mathcal{J}^{(j)}; u_i^{(j)} \leq \psi_i\}, & \mathcal{J}^{(j,1)} = \mathcal{J}^{(j)} \setminus \mathcal{J}_j^{(0)}, \\ \mathcal{I}^{(j+1)} = \mathcal{I}^{(j,0)} \cup \mathcal{J}^{(j,1)}, & \mathcal{J}^{(j+1)} = \{1, \dots, N\} \setminus \mathcal{I}^{(j+1)}. \end{array} \tag{9}$$

- The algorithm stops if there exists a step n such that $\mathcal{I}^{(n)} = \mathcal{I}^{(n+1)}$.

Let us first remark that this algorithm is well defined.

PROPOSITION 2.4. *Let $\Sigma = \{s_i, i = 1, N\}$ denote the set of nodes of a given triangulation of Ω, let $\mathcal{I}^{(j)} \subset \{1, \dots, N\}$ and $\mathcal{J}^{(j)} = \{1, \dots, N\} \setminus \mathcal{I}^{(j)}$; then problem* (8) *has a unique solution.*

Proof. The proof of this result follows immediately from the Lax–Milgram lemma by noting that under Assumptions 2.1 and with the notations of Definition 2.2 and Proposition 2.4, $U^{(j)} = (u_1^{(j)}, \dots, u_N^{(j)}) \in \mathbb{R}^N$ is a solution to problem (8) if and only if $\tilde{u}^{(j)}(x) = \sum_{i=1}^N u_i^{(j)} \varphi_i(x)$ is a solution to the following variational problem:

$$\left\{ \begin{array}{l} \tilde{u}^{(j)} \in V_h \quad \text{s.t. } u_i^{(j)} = \psi_i \; \forall i \in \mathcal{I}^{(j)}, \\ \displaystyle\int_\Omega \nabla\tilde{u}^{(j)}(x) \cdot \nabla v(x) dx = \int_\Omega f(x) v(x) dx \quad \forall v \in V_h. \end{array} \right. \tag{10}$$ $\square$

Let us now show that the algorithm defined by (7)–(9) is monotonic.

LEMMA 2.5. *Under Assumption* 2.1 *and those of Definition* 2.2, *the sequence* $(U^{(j)})_{j\in\mathbb{N}}$ *constructed by the algorithm* (7)–(9), *where* $U^{(j)}=(u_1^{(j)},\ldots,u_N^{(j)})$, *satisfies*

$$u_i^{(j+1)} \le u_i^{(j)} \quad \forall j\in\mathbb{N},\, \forall i\in\{1,\ldots,N\}. \tag{11}$$

Equivalently, the sequence of functions $(\tilde{u}^{(j)})_{j\in\mathbb{N}}$ *defined by* $\tilde{u}^{(j)}(x)=\sum_{i=1}^N u_i^{(j)}\varphi_i(x)$ *for all* $x\in\Omega$ *satisfies*

$$\tilde{u}^{(j+1)} \le \tilde{u}^{(j)} \quad \forall j\in\mathbb{N}. \tag{12}$$

Proof. Let $j\in\mathbb{N}$ and $w_h=\tilde{u}^{(j)}-\tilde{u}^{(j+1)}$. Then

$$\int_\Omega |\nabla w_h^-(x)|^2dx = -\sum_{i=1}^N w_i^- \int_\Omega \nabla w_h(x)\nabla\varphi_i(x)dx. \tag{13}$$

- If $i\in\mathcal{I}^{(j)}\cap\mathcal{I}^{(j+1)}$, one has $w_i=0$, and therefore $\int_\Omega \nabla w_h(x)\cdot\nabla(w_i^-\varphi_i(x))dx=0$.
- If $i\in\mathcal{J}^{(j)}\cap\mathcal{J}^{(j+1)}$, one has $\int_\Omega\nabla\tilde{u}^{(j)}(x)\cdot\nabla\varphi_i(x)dx=\int_\Omega\nabla\tilde{u}^{(j+1)}(x)\cdot\nabla\varphi_i(x)dx$, and therefore
$$\int_\Omega \nabla w_h(x)\cdot\nabla(w_i^-\varphi_i(x))dx=0.$$
- If $i\in\mathcal{J}^{(j)}\cap\mathcal{I}^{(j+1)}$, one obtains $u_i^{(j)}>\psi_i$ and $u_i^{(j+1)}=\psi_i$, hence $w_i>0$, and therefore
$$\int_\Omega \nabla w_h(x)\cdot\nabla(w_i^-\varphi_i(x))dx=0.$$
- Finally if $i\in\mathcal{I}^{(j)}\cap\mathcal{J}^{(j+1)}$, then $(AU^{(j)})_i>F_i$ and $(AU^{(j+1)})_i=F_i$, so that
$$\int_\Omega \nabla w_h(x)\cdot\nabla(w_i^-\varphi_i(x))dx\ge 0.$$

These inequalities and (13) yield that $\int_\Omega|\nabla w_h^-(x)|^2dx=0$, and since $w_h^-\in H_0^1(\Omega)$, this implies that $w_h\ge 0$, which concludes the proof of the lemma. ∎

We may now turn to the convergence of the algorithm. We first state that if the sets $\mathcal{I}^{(j)}$ and $\mathcal{J}^{(j)}$ are left unchanged from one iteration to the next, then the algorithm has reached the unique solution to problem (5).

PROPOSITION 2.6. *Assume that the sequence of sets* $(\mathcal{I}^{(j)})_{j\in\mathbb{N}}$ *constructed by the algorithm* (7)–(9) *is such that there exists* $n\in\mathbb{N}$ *such that* $\mathcal{I}^{(n)}=\mathcal{I}^{(n+1)}$; *then the solution* $U^{(n)}$ *to* (8) *is the unique solution to problem* (5).

Proof. Under the assumptions of Proposition 2.6, let $\mathcal{I}=\mathcal{I}^{(n)}$, $\mathcal{J}=\mathcal{J}^{(n)}$; let $U^{(n)}=(u_1,\ldots,u_n)$ be the solution to (8) with $j=n$. Since $\mathcal{J}^{(n)}=\mathcal{J}^{(n+1)}$, one has $u_i\le\psi_i$ for any $i\in\mathcal{J}^{(n)}$. Furthermore, $u_i=\psi_i$ for any $i\in\mathcal{I}^{(n)}$, so that $u_i\le\psi_i$ for any $i\in\{1,\ldots,N\}$. In a similar way, one has that $(AU)_i\le F_i$ for any $i\in\{1,\ldots,N\}$, and from (8) one has that $((AU)_i-F_i)(u_i-\psi_i)=0$ for any $i\in\{1,\ldots,N\}$. ∎

Let us now show that the monotonic algorithm terminates in a finite number of iterations.

THEOREM 2.7. *Under Assumption* (2.1), *there exists* $n\in\mathbb{N}$ *such that the sequence* $(U^{(n)})_{n\in\mathbb{N}}$ *constructed by the algorithm* (7)–(9) *is such that* $U^{(n)}$ *is the exact solution to the discrete problem* (4) *for all* $j\ge n$. *Furthermore the integer* n *satisfies*

$$n\le N+1. \tag{14}$$

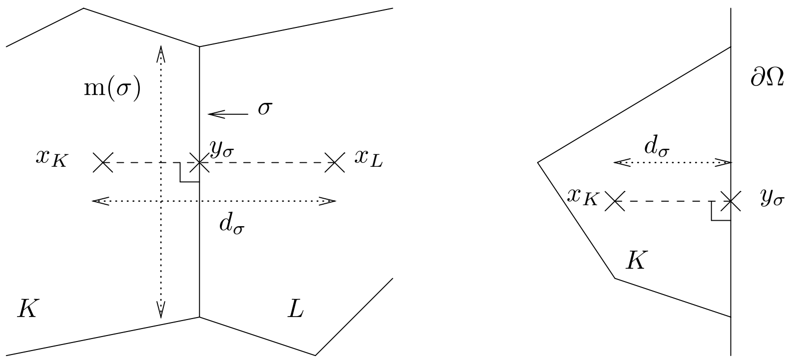


FIG. 1. *Admissible meshes.*

Proof. Let the sets $\mathcal{I}^{(j)}$ and $\mathcal{J}^{(j)}$ be defined by the algorithm (7)–(9) for any step (j); if there exists an integer n such that $\mathcal{I}^{(n)} = \mathcal{I}^{(n+1)}$, then, by Proposition 2.10, $U^{(n)}$ is the exact solution to the discrete problem (4), and the first part of the theorem is proven. It remains to prove that such a step exists and that it satisfies (14).

Let us first remark that for $i \in \{1, \ldots, N\}$ if $u_i^{(0)} \leq \psi_i$, then $u_i^{(1)} \leq \psi_i$ by Lemma 2.5, and if $u_i^{(0)} > \psi_i$, then $u_i^{(1)} = \psi_i$ by (9) in the monotonic algorithm. Hence

$$u_i^{(1)} \leq \psi_i \text{ for any } i \in \{1, \ldots, N\}. \tag{15}$$

Therefore, by an easy induction, one has that $\mathcal{I}^{(j)} = \mathcal{I}_0^{(j)}$ for any $j > 1$, which yields that $\mathcal{I}^{(j)} \subset \mathcal{I}^{(j+1)}$ for any $j > 1$. Since $\mathcal{I}^{(n)} \subset \{1, \ldots, N\}$ is a finite set, this means that there exists an index n such that $\mathcal{I}^{(n)} = \mathcal{I}^{(n+1)}$.

Let us finally show that (14) holds true. Let n be the smallest integer such that $\mathcal{I}^{(n)} = \mathcal{I}^{(n+1)}$. Since $\mathcal{I}^{(j)}$ is strictly included in $\mathcal{I}^{(j+1)}$ for any $j > 1$, one has $N + 1 \geq \operatorname{card}(\mathcal{I}^{(j+1)}) \geq \operatorname{card}(\mathcal{I}^{(j+1)}) + 1$ for any $j < n$, which yields that $n \leq N + 1$. ◻

2.2. Approximation by the finite volume scheme. Let us now define a discretization mesh over Ω, which is assumed (following [9]) to be admissible for finite volumes in the following sense (see Figure 1).

DEFINITION 2.8 (admissible meshes). *Let Ω be an open bounded polygonal domain of $\mathbb{R}^d$. An admissible finite volume mesh of Ω, denoted by $\mathcal{T}$, is given by a family of "control volumes," which are disjoint polygonal convex subsets of Ω, a family of subsets of $\overline{\Omega}$ contained in hyperplanes of $\mathbb{R}^d$, denoted by $\mathcal{E}$ (these are the "sides" of the control volumes), with strictly positive one-dimensional measure, and a family of points of Ω, denoted by $\mathcal{P}$, satisfying the following properties (in fact, we shall denote, somewhat incorrectly, by $\mathcal{T}$ the family of control volumes):*

(i) *The closure of the union of all the control volumes is $\overline{\Omega}$.*
(ii) *For any $K \in \mathcal{T}$, there exists a subset $\mathcal{E}_K$ such that $\partial K = \overline{K} \setminus K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$.*
(iii) *For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the one-dimensional Lebesgue measure of $\overline{K} \cap \overline{L}$ is 0 or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, which will then be denoted by $K|L$.*
(iv) *The family $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$ is such that $x_K \in K$ (for all $K \in \mathcal{T}$) and, if $\sigma = K|L$, it is assumed that $x_K \neq x_L$, and the straight line $\mathcal{D}_{K,L}$ going through x_K and x_L is assumed to be orthogonal to $K|L$.*

In what follows, the following notations are used. Let $\mathrm{size}(\mathcal{T}) = \sup\{\mathrm{diam}(K), K \in \mathcal{T}\}$. *For any* $K \in \mathcal{T}$ *and* $\sigma \in \mathcal{E}$, $\mathrm{m}(K)$ *is the two-dimensional Lebesgue measure of* K, *and* $\mathrm{m}(\sigma)$ *the one-dimensional measure of* σ. *The set of interior (resp., boundary) edges is denoted by* $\mathcal{E}_{\mathrm{int}}$ *(resp.,* $\mathcal{E}_{\mathrm{ext}}$*), that is,* $\mathcal{E}_{\mathrm{int}} = \{\sigma \in \mathcal{E};\ \sigma \not\subset \partial\Omega\}$ *(resp.,* $\mathcal{E}_{\mathrm{ext}} = \{\sigma \in \mathcal{E};\ \sigma \subset \partial\Omega\}$*). The set of neighbors of* K *is denoted by* $\mathcal{N}(K)$, *that is,* $\mathcal{N}(K) = \{L \in \mathcal{T};\ \exists \sigma \in \mathcal{E}_K\ \sigma = K \cap L\}$. *If* $\sigma = K|L$, *we denote by* d_σ *or* $d_{K|L}$ *the Euclidean distance between* x_K *and* x_L *(which is positive). If* $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{\mathrm{ext}}$, *let* d_σ *denote the Euclidean distance between* x_K *and* y_σ. *For any* $\sigma \in \mathcal{E}$, *the transmissivity through* σ *is defined by* $\tau_\sigma = \frac{\mathrm{m}(\sigma)}{d_\sigma}$ *if* $d_\sigma \neq 0$.

Remark 2.2. The condition $x_K \neq x_L$ if $\sigma = K|L$ is in fact quite easy to satisfy: two neighboring control volumes K, L, which do not satisfy it, just have to be collapsed into a new control volume M with $x_M = x_K = x_L$, and the edge $K|L$ removed from the set of edges. The new mesh thus obtained is admissible.

We refer to, e.g., [9] or [14] for examples of admissible meshes. These include rectangular meshes, Delaunay triangulations, and Voronoi meshes.

Let us now define a "discrete" functional space and a discrete H_0^1 norm.

Definition 2.9. *Let* Ω *be an open bounded polygonal domain of* $\mathbb{R}^d$, *and* $\mathcal{T}$ *be an admissible mesh in the sense of Definition* 2.8.

Define $Y(\mathcal{T})$ *as the set of the functions defined a.e. from* Ω *to* $\mathbb{R}$ *which are constant over each control volume of the mesh. We shall denote by* u_K *the value taken by* u *on the control volume* K.

For $u \in Y(\mathcal{T})$, *define the discrete* H_0^1 *norm by*

$$\|u\|_{1,\mathcal{T}}^2 = \sum_{\sigma \in \mathcal{E}} \tau_\sigma (D_\sigma u)^2, \tag{16}$$

with

$$|D_\sigma u| = |u_K - u_L| \text{ if } \sigma \in \mathcal{E}_{\mathrm{int}}, \qquad \sigma = K/L, \tag{17}$$

$$D_\sigma u = -u_K \text{ if } \sigma \subset \partial\Omega. \tag{18}$$

Let $\mathcal{T}$ be an admissible finite volume mesh in the sense of Definition 2.8, let $\psi_K = \psi(x_K)$ and $f_K = \frac{1}{m(K)} \int_K f(x)dx$ for any $K \in \mathcal{T}$. A cell-centered finite volume discretization of problem (1) is written with respect to the discrete unknowns $(u_K)_{K \in \mathcal{T}}$ in the following way (see [19] for a description of how this scheme is obtained):

$$-\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} + m(K) f_K \geq 0 \quad \forall K \in \mathcal{T}, \tag{19}$$

$$\left(-\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} + m(K) f_K\right)(\psi_K - u_K) = 0 \quad \forall K \in \mathcal{T}, \tag{20}$$

$$u_K \leq \psi_K \quad \forall K \in \mathcal{T}, \tag{21}$$

$$F_{K,\sigma} = -\tau_\sigma (u_L - u_K) \quad \forall \sigma \in \mathcal{E}_{\mathrm{int}} \text{ if } \sigma = K/L, \tag{22}$$

$$F_{K,\sigma} = \tau_\sigma u_K \quad \forall \sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K. \tag{23}$$

The proof of the existence and uniqueness of the solution to this scheme was given in [19]. It follows for the following remark: let $(u_K)_{K \in \mathcal{T}} \in \mathbb{R}^{\mathrm{card}(\mathcal{T})}$, and let $u_{\mathcal{T}} \in Y(\mathcal{T})$ be defined by $u_{\mathcal{T}}(x) = u_K$ for $x \in K$ $\forall K \in \mathcal{T}$. Then one may show that $(u_K)_{K\in\mathcal{T}}$ is a solution to problem (19)–(23) if and only if $u_{\mathcal{T}}$ is a solution to the following problem:

$$\left\{\begin{array}{l} u_{\mathcal{T}} \in \mathcal{K}_{\mathcal{T}} = \{v \in Y(\mathcal{T}), \text{ s.t. } v_K \le \psi_K \ \forall K \in \mathcal{T}\}, \\ A(u_{\mathcal{T}}, v - u_{\mathcal{T}}) \ge L(v - u_{\mathcal{T}}) \quad \forall v \in \mathcal{K}_{\mathcal{T}}, \end{array}\right. \tag{24}$$

where for any $u = (u_K)_{K\in\mathcal{T}}$ and $v = (v_K)_{K\in\mathcal{T}} \in Y(\mathcal{T})$,

$$A(u,v) = \sum_{\sigma=K|L\,\in\mathcal{E}_{\text{int}}} \tau_{K|L}(u_K - u_L)(v_K - v_L) + \sum_{\sigma\in\mathcal{E}_{\text{ext}}\cap\mathcal{E}_K} \tau_\sigma u_K v_K, \tag{25}$$

and

$$L(v) = \sum_{K\in\mathcal{T}} m(K) f_K v_K. \tag{26}$$

Our goal here is to construct an algorithm yielding an approximate solution of problem (19)–(23). The iterative process which we described for the finite element discretization is easily adapted to the finite volume framework. Let $K \in \mathcal{T}$; then from (20) one has

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} = m(K) f_K \quad \text{or} \quad u_K = \psi_K.$$

Therefore, from (19) and (21), there exist two disjoint subsets of $\mathcal{T}$ such that on one subset one has

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} = m(K) f_K \text{ and } u_K \le \psi_K \qquad \text{for } K \text{ in the first subset,}$$

and

$$u_K = \psi_K \text{ and } \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} \le m(K) f_K \qquad \text{for } K \text{ in the second subset.}$$

Now assume that we knew two subsets $\mathcal{T}_f$ and $\mathcal{T}_\psi$ of $\mathcal{T}$ such that $\mathcal{T}_f \cup \mathcal{T}_\psi = \mathcal{T}$, $\mathcal{T}_f \cap \mathcal{T}_\psi = \emptyset$, and

$$u_K \le \psi_K \quad \forall K \in \mathcal{T}_f, \tag{27}$$

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} \le m(K) f_K \quad \forall K \in \mathcal{T}_\psi. \tag{28}$$

Then, as in the finite element case, the solution of problem (19)–(23) could be obtained by solving the linear problem

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} = m(K) f_K \quad \forall K \in \mathcal{T}_f, \tag{29}$$

$$u_K = \psi_K \quad \forall K \in \mathcal{T}_\psi, \tag{30}$$

where the numerical fluxes $F_{K,\sigma}$ are defined by (22)–(23). As in the finite element case, we shall solve (29)–(30) at each iteration and iterate on the sets $\mathcal{T}_f$ and $\mathcal{T}_\psi$ by looking at the constraints which are violated after the solution of (29)–(30).

The algorithm that follows determines $\mathcal{T}_f$ and $\mathcal{T}_\psi$ by an iterative method.

Monotonic algorithm, obstacle problem, finite volume discretization.

• Initialization. Let $\mathcal{T}_f^{(0)}$ and $\mathcal{T}_\psi^{(0)}$ be such that

$$\mathcal{T}_f^{(0)} \cap \mathcal{T}_\psi^{(0)} = \emptyset \quad \text{and} \quad \mathcal{T}_f^{(0)} \cup \mathcal{T}_\psi^{(0)} = \mathcal{T} \tag{31}$$

(for example, $\mathcal{T}_f^{(0)} = \mathcal{T}$ and $\mathcal{T}_\psi^{(0)} = \emptyset$).

• Step (j). Assume the sets $\mathcal{T}_f^{(j)}$ and $\mathcal{T}_\psi^{(j)}$ to be known such that $\mathcal{T}_f^{(j)} \cap \mathcal{T}_\psi^{(j)} = \emptyset$ and $\mathcal{T}_f^{(j)} \cup \mathcal{T}_\psi^{(j)} = \mathcal{T}$. Let $(u_K^{(j)})_{K\in\mathcal{T}}$ be the solution to the following set of equations:

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^{(j)} = m(K) f_K \quad \forall K \in \mathcal{T}_f^{(j)}, \tag{32}$$

$$u_K^{(j)} = = \psi_K \quad \forall K \in \mathcal{T}_\psi^{(j)}, \tag{33}$$

$$F_{K,\sigma}^{(j)} = \tau_\sigma (u_K^{(j)} - u_L^{(j)}) \quad \forall \sigma \in \mathcal{E}_{\text{int}} \text{ if } \sigma = K/L, \tag{34}$$

$$F_{K,\sigma}^{(j)} = \tau_\sigma u_K^{(j)} \quad \forall \sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K. \tag{35}$$

Let $\mathcal{T}_f^{(j+1)}$ and $\mathcal{T}_\psi^{(j+1)}$ be defined in the following way:

$$\begin{aligned}
&\mathcal{T}_f^{(j,0)} = \{K \in \mathcal{T}_f^{(j)}, ; u_K^{(j)} \le \psi_K\}, && \mathcal{T}_f^{(j,1)} = \mathcal{T}_f^{(j)} \setminus \mathcal{T}_f^{(j,0)},\\
&\mathcal{T}_\psi^{(j,0)} = \left\{ K \in \mathcal{T}_\psi^{(j)}; \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^{(j)} \le m(K) f_K \right\}, && \mathcal{T}_\psi^{(j,1)} = \mathcal{T}_\psi^{(j)} \setminus \mathcal{T}_\psi^{(j,0)},\\
&\mathcal{T}_f^{(j+1)} = \mathcal{T}_f^{(j,0)} \cup \mathcal{T}_\psi^{(j,1)}, && \mathcal{T}_\psi^{(j+1)} = \mathcal{T} \setminus \mathcal{T}_f^{(j+1)}.
\end{aligned} \tag{36}$$

• The algorithm stops if there exists a step (J) such that $\mathcal{T}_f^{(J)} = \mathcal{T}_f^{(J+1)}$ and $\mathcal{T}_\psi^{(J)} = \mathcal{T}_\psi^{(J+1)}$.

The above algorithm is well defined thanks to the following result.

Proposition 2.10. *Let $\mathcal{T}$ be an admissible finite volume mesh in the sense of Definition* 2.8, *and assume that the sets $\mathcal{T}_f^{(j)}$ and $\mathcal{T}_\psi^{(j)}$ such that $\mathcal{T}_f^{(j)} \cap \mathcal{T}_\psi^{(j)} = \emptyset$ and $\mathcal{T}_f^{(j)} \cup \mathcal{T}_\psi^{(j)} = \mathcal{T}$ are known; then problem* (32)–(35) *admits a unique solution.*

Proof. Under the assumptions of Proposition 2.10, one may find an equivalent "variational" formulation to problem (32)–(35). Let $u_\mathcal{T}^{(j)} \in Y(\mathcal{T})$ be defined by $u_\mathcal{T}^{(j)}(x) = u_K^{(j)}$ for $x \in K$, $\forall K \in \mathcal{T}$; it is easy to prove that $u_\mathcal{T}^{(j)}$ is a solution to problem (32)–(35) if and only if $u_\mathcal{T}^{(j)}$ is a solution to the following problem:

$$\left\{ \begin{array}{l} u_K^{(j)} = \psi_K \quad \forall K \in \mathcal{T}_\psi^{(j)},\\ A(u_\mathcal{T}^{(j)}, v) = L(v) \quad \forall v = (v_K)_{K\in\mathcal{T}} \in Y(\mathcal{T}),\\ \qquad \text{such that } v_K = 0 \quad \forall K \in \mathcal{T}_\psi^{(j)}, \end{array} \right. \tag{37}$$

with A and L defined by (25) and (26). The existence and uniqueness of the solution to (32)–(35) (and (37)) follow from the Lax–Milgram lemma. ∎

The algorithm (31)–(36) is therefore well defined; let us now show its monotonicity. This property is much related to the discrete maximum principle, which holds for finite volume discretizations of the Laplace equation; see, e.g., [18].

LEMMA 2.11 (monotonicity of the scheme). *Under Assumption* 2.1, *let $\mathcal{T}$ be an admissible finite volume mesh in the sense of Definition* 2.8; *the sequences $(u_K^{(j)})_{j\in\mathbb{N},K\in\mathcal{T}}$ which are constructed by the algorithm* (31)–(36) *satisfy*

$$u_K^{(j+1)} \le u_K^{(j)} \quad \text{for } j \in \mathbb{N} \text{ and } K \in \mathcal{T}.$$

Proof. Define $v_{\mathcal{T}} = u_{\mathcal{T}}^{(j)} - u_{\mathcal{T}}^{(j+1)}$, $F_{K,\sigma} = F_{K,\sigma}^{(j)} - F_{K,\sigma}^{(j+1)}$ $\forall K \in \mathcal{T}$, $\forall \sigma \in \mathcal{E}_K$, and $\min(v_{\mathcal{T}}) = \min\{v_K = u_K^{(j)} - u_K^{(j+1)},\ K \in \mathcal{T}\}$; let us show that $\min(v_{\mathcal{T}}) \ge 0$. Let $K_0 \in \mathcal{T}$ such that $\min(v_{\mathcal{T}}) = v_{K_0}$. Then we have the following:

- If $K_0 \in \mathcal{T}_\psi^{(j)} \cap \mathcal{T}_\psi^{(j+1)}$, then $v_{K_0} = 0$ so that $\min(v_{\mathcal{T}}) = 0$.
- Now if $K_0 \in \mathcal{T}_f^{(j)} \cap \mathcal{T}_\psi^{(j+1)}$, one has $u_{K_0}^{(j)} > \psi_{K_0}$ and $u_{K_0}^{(j+1)} > \psi_{K_0}$ so that $\min(v_{\mathcal{T}}) > 0$.
- Assume next that $K_0 \in \mathcal{T}_\psi^{(j)} \cap \mathcal{T}_f^{(j+1)}$; then

  $$\sum_{\sigma\in\mathcal{E}_{K_0}} F_{K_0,\sigma}^{(j)} < m(K_0) f_{K_0} < 0 \quad \text{and} \quad \sum_{\sigma\in\mathcal{E}_{K_0}} F_{K_0,\sigma}^{(j+1)} = m(K_0) f_{K_0}.$$

  Therefore $\sum_{\sigma\in\mathcal{E}_{K_0}} F_{K_0,\sigma} > 0$, and, since $v_{K_0} \le v_K$ $\forall K \in \mathcal{T}$, one has $\sum_{\sigma\in\mathcal{E}_{K_0}} F_{K_0,\sigma} \le 0$, which is impossible.
- Let us finally assume that $K_0 \in \mathcal{T}_f^{(j)} \cap \mathcal{T}_f^{(j+1)}$; in this case one has

  $$\sum_{\sigma\in\mathcal{E}_{K_0}} F_{K_0,\sigma} = 0. \tag{38}$$

  1. If the control volume K_0 lies near the boundary, that is, $\mathcal{E}_{K_0} \cap \mathcal{E}_{\text{ext}} \ne \emptyset$, then (38) becomes

     $$\sum_{\substack{\sigma\in\mathcal{E}_{K_0}\cap\mathcal{E}_{\text{int}} \\ \sigma=K_0|K}} \frac{v_{K_0} - v_{K_\sigma}}{d_\sigma} + v_{K_0}\left(\sum_{\sigma\in\mathcal{E}_{K_0}\cap\mathcal{E}_{\text{ext}}} \frac{1}{d_{K_0,\sigma}}\right) = 0.$$

     Since $\min(v_{\mathcal{T}}) = v_{K_0}$, all the terms in the first sum are nonpositive, and therefore v_{K_0} must be nonnegative, which proves that $\min(v_{\mathcal{T}}) \ge 0$.
  2. Now if the control volume K_0 lies in the interior domain in the sense that $\mathcal{E}_{K_0} \subset \mathcal{E}_{\text{int}}$, then one needs to consider one of the two following subcases:
     (a) There exists a "path" of control volumes, which are all in $\mathcal{T}_f^{(j)} \cap \mathcal{T}_f^{(j+1)}$, leading from K_0 to the boundary; that is, there exists $m \in \mathbb{N}$ and $(K_\ell)_{\ell=0,\dots,m}$ such that $K_\ell \in \mathcal{T}_f^{(j)} \cap \mathcal{T}_f^{(j+1)}$, $\mathcal{E}_{K_\ell} \cap \mathcal{E}_{K_{\ell+1}} \ne \emptyset$ $\forall \ell = 0, \dots, m-1$. In this case, one has $v_{K_0} = v_{K_1} = \cdots = v_{K_m} = \min(v_{\mathcal{T}})$, and since K_m lies near the boundary, $\min(v_{\mathcal{T}}) \ge 0$.
     (b) If there does not exist such a path, then there exists some control volume K which does not belong to $\mathcal{T}_f^{(j)} \cap \mathcal{T}_f^{(j+1)}$ and such that $\min(v_{\mathcal{T}}) = v_K$; this case falls into one of the three cases which were previously analyzed, and for which we proved that $\min(v_{\mathcal{T}}) \ge 0$. ∎

We may now turn to the convergence of the algorithm. As for the finite element discretization, we first state that if the sets $\mathcal{T}_f^{(j)}$ and $\mathcal{T}_\psi^{(j)}$ are left unchanged from one iteration to the next, then the algorithm has reached the unique solution to problem (5).

PROPOSITION 2.12. *Assume that the sequence of sets* $(\mathcal{T}_f^{(j)})_{j\in\mathbb{N}}$ *and* $(\mathcal{T}_\psi^{(j)})_{j\in\mathbb{N}}$*, which are constructed by the algorithm* (31)–(36)*, are such that there exists a step* (J) *such that* $\mathcal{T}_f^{(J)} = \mathcal{T}_f^{(J+1)}$ *and* $\mathcal{T}_\psi^{(J)} = \mathcal{T}_\psi^{(J+1)}$*; then the solution* $(u_K^{(J)})_{K\in\mathcal{T}}$ *to* (32)–(35) *is the unique solution to problem* (19)–(23).

Proof. Let $\mathcal{T}_f = \mathcal{T}_f^{(J)}$, $\mathcal{T}_\psi = \mathcal{T}_\psi^{(J)}$, and $u_\mathcal{T} = u_\mathcal{T}^{(J)}$; hence $u_\mathcal{T}$ satisfies the set of equations (32)–(35). Since $\mathcal{T}_\psi^{(J)} = \mathcal{T}_\psi^{(J+1)}$, one has $-\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + m(K)f_K \geq 0$ $\forall K \in \mathcal{T}_\psi$, and, thanks to (32), one has $-\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + m(K)f_K = 0$ $\forall K \in \mathcal{T}_f$; since $\mathcal{T} = \mathcal{T}_\psi \cup \mathcal{T}_f$, $u_\mathcal{T}$ satisfies (19). Similarly, since $\mathcal{T}_f^{(J)} = \mathcal{T}_f^{(J+1)}$, one has $u_K \leq \psi_K$ $\forall K \in \mathcal{T}_f$, and, thanks to (33), $u_\mathcal{T}$ satisfies (21), and finally, since $\mathcal{T}_f \cup \mathcal{T}_\psi = \mathcal{T}$, $u_\mathcal{T}$ satisfies (20). Hence $u_\mathcal{T}$ is the unique solution to problem (19)–(23). ∎

THEOREM 2.13. *Under Assumption* 2.1*, there exists an integer* $J \in \mathbb{N}$ *such that the sequence* $(u_K^{(j)})_{j\in\mathbb{N}}$*, which is constructed by the algorithm* (31)–(36)*, is such that* $(u_K^{(j)}, K \in \mathcal{T})$ *is the exact solution to the discrete problem* (24) *for all* $j \geq J$*. Furthermore the integer* J *satisfies*

$$J \leq \mathrm{card}(\mathcal{T}) + 1, \tag{39}$$

where $\mathrm{card}(\mathcal{T})$ *denotes the number of cells of the mesh.*

Proof. Let the sets $\mathcal{T}_\psi^{(j)}$ and $\mathcal{T}_f^{(j)}$ be defined by the algorithm (31)–(36) for any step (j); if there exists an integer J such that $\mathcal{T}_\psi^{(J)} = \mathcal{T}_\psi^{(J+1)}$, then by Proposition 2.12, $(u_K^{(J)}, K \in \mathcal{T})$ is the exact solution to the discrete problem (24), and the first part of the theorem is proven. There remains to prove that such a step exists and that it satisfies (39).

As in the case of the finite element discretization, let us first remark that for $K \in \mathcal{T}$, if $u_K^{(0)} \leq \psi_K$, then $u_K^{(1)} \leq \psi_i$ by Lemma 2.11, and if $u_K^{(0)} > \psi_K$, then $u_K^{(1)} = \psi_K$ by step (9) of the algorithm. Hence

$$u_K^{(1)} \leq \psi_K \text{ for any } K \in \mathcal{T}; \tag{40}$$

therefore by an easy induction one has that $\mathcal{T}_\psi^{(j)} \subset \mathcal{T}_\psi^{(j+1)}$ for any $j > 1$. Since $\mathcal{T}_\psi^{(j)} \subset \mathcal{T}$, this means that there exists an index J such that $\mathcal{T}_\psi^{(J)} = \mathcal{T}_\psi^{(J+1)}$.

The proof of (39) is identical to the case of the finite element discretization (see the proof of Theorem 2.7). ∎

3. The Signorini problem. Let us now consider the following diffusion problem:

$$-\Delta u(x) = f, \quad x \in \Omega, \tag{41}$$

$$u(x) = 0, \quad x \in \Gamma^1, \tag{42}$$

$$\nabla u(x) \cdot \mathbf{n} = 0, \quad x \in \Gamma^2, \tag{43}$$

with a Signorini condition on a part of the boundary,

$$\left.\begin{array}{rcl} u(x) & \geq & a, \\ \nabla u(x) \cdot \mathbf{n} & \geq & b, \\ (u(x) - a)(\nabla u(x).\mathbf{n} - b) & = & 0, \end{array}\right\} \quad x \in \Gamma_3, \tag{44}$$

where the following holds.

Assumption 3.1.

1. Ω is an open bounded polygonal subset of $\mathbb{R}^d$.
2. The boundary $\partial\Omega$ of Ω is composed of three nonempty, disjoint connected sets Γ^1, Γ^2, and Γ^3 such that $\overline{\Gamma^1} \cup \overline{\Gamma^2} \cup \overline{\Gamma^3} = \overline{\partial\Omega}$.
3. $f \in L^2(\Omega)$, $a \leq 0$, and $b \in \mathbb{R}$.
4. $\mathbf{n}$ is the unit normal vector to $\partial\Omega$ outward to the domain Ω.

Under some regularity assumptions, problem (41)–(44) is equivalent to the following variational problem (see, e.g., [17]):

$$\left\{ \begin{array}{l} u \in \mathcal{K} = \{v \in H^1(\Omega),\ v_{|\partial\Omega} \geq a \text{ a.e.}\}, \quad \text{satisfying} \\ \displaystyle\int_\Omega \nabla u(x) \cdot \nabla(v-u)(x)dx \geq \int_{\partial\Omega} b(\gamma(v) - \gamma(u))(s)ds \quad \forall v \in \mathcal{K}, \end{array} \right. \tag{45}$$

with $v_{\partial\Omega} = \gamma(v)_{\partial\Omega}$, where γ is the trace operator from $H^1(\Omega)$ to $L^2(\partial\Omega)$. By Stampacchia's theorem, problem (45) has a unique solution.

The Signorini problem may be viewed as an obstacle problem in which the obstacle is located on the boundary. However, because the complementarity condition is written on the normal derivative on the boundary, one may not write the monotonic algorithm with piecewise linear finite elements in a straightforward way as in the case of the obstacle problem. Indeed, the normal derivative of the piecewise linear finite element approximate solution is defined on each edge of a triangle neighboring the boundary of the domain, but it is not defined at the nodes of the triangulation lying on the boundary. This problem could be solved by using higher order finite elements, but, as we already mentioned in Remark 2.1, a crucial issue in the underlying electrochemical application is that the maximum principle must hold, and this is not the case with higher order finite element methods. However, there is no such problem when using a finite volume discretization of the Signorini problem; the discrete normal derivative is well defined, and the maximum principle holds (see [19]). Hence the monotonic algorithm may be written quite easily.

We shall use here the same admissible finite volume meshes as for the discretization of the obstacle problem, which were defined in Definition 2.8, with the two following additional assumptions, which are needed because of the Signorini boundary conditions on the boundary:

(v) *For any $\sigma \in \mathcal{E}$ such that $\sigma \subset \partial\Omega$, there exists $i \in \{1, 2, 3\}$ such that $\sigma \subset \Gamma^i$.*

(vi) *For any $\sigma \in \mathcal{E}$ such that $\sigma \subset \partial\Omega$, let K be the control volume such that $\sigma \in \mathcal{E}_K$ and $\mathcal{D}_{K,\sigma}$ be the straight line going through x_K and orthogonal to σ; then $y_\sigma = \mathcal{D}_{K,\sigma} \cap \sigma$.*

Let us then define an appropriate "discrete" functional space.

DEFINITION 3.1 (discrete functional space). *Let Ω be an open bounded polygonal domain of $\mathbb{R}^d$, and $\mathcal{T}$ be an admissible mesh in the sense of Definition* 2.8. *Define $X(\mathcal{T})$ as the set of the functions defined a.e. from $\overline{\Omega}$ to $\mathbb{R}$ which are constant over each control volume of the mesh, and which are constant over each edge in $\mathcal{E}_3 = \mathcal{E}_{\text{ext}}$. We shall denote by u_K the value taken by u on the control volume K, and by u_σ the value taken by u on the edge $\sigma \in \mathcal{E}_{\text{ext}}$, $\sigma \subset \Gamma^3$.*

As in the case of the obstacle problem, a classical finite volume formulation is obtained by integrating the diffusion equation (41) over each control volume $\mathcal{T}$, using Green's formula and approximating the normal fluxes by a consistent difference quotient. Let us denote the discrete unknowns by $(u_K)_{K\in\mathcal{T}}$ for any $K \in \mathcal{T}$ and by $(u_\sigma)_{\sigma\subset\Gamma^3}$ for any $\sigma \in \mathcal{E}_{\text{ext}}$, and the "discrete flux" by $F_{K,\sigma}$, which is expected to approximate the exact flux $-\int_\sigma \nabla u(s).\mathbf{n}ds$; the finite volume scheme can be written

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} = 0 \quad \forall K \in \mathcal{T}, \tag{46}$$

with

$$F_{K,\sigma} = -\tau_\sigma(u_L - u_K) \quad \forall\sigma \in \mathcal{E}_{\text{int}} \text{ if } \sigma = K/L, \tag{47}$$

$$F_{K,\sigma} = \tau_\sigma u_K \quad \forall\sigma \subset \Gamma^1,\ \sigma \in \mathcal{E}_K, \tag{48}$$

$$F_{K,\sigma} = 0 \quad \forall\sigma \subset \Gamma^2,\ \sigma \in \mathcal{E}_K, \tag{49}$$

$$F_{K,\sigma} = -\tau_\sigma(u_\sigma - u_K) \quad \forall\sigma \in \mathcal{E}_3,\ \sigma \in \mathcal{E}_K, \tag{50}$$

with the Signorini boundary condition

$$u_\sigma \geq a \quad \forall\sigma \in \mathcal{E}_3, \tag{51}$$

$$-F_{K,\sigma} \geq \mathrm{m}(\sigma)\, b \quad \forall\sigma \in \mathcal{E}_3, \tag{52}$$

$$(u_\sigma - a)\left(\frac{F_{K,\sigma}}{\mathrm{m}(\sigma)} + b\right) = 0 \quad \forall\sigma \in \mathcal{E}_3, \tag{53}$$

where $\mathcal{E}_3$ denotes the set of edges of the mesh that are included in Γ^3.

In [19], we prove the following existence result.

PROPOSITION 3.2. *Let $\mathcal{T}$ be an admissible mesh of Ω; problem* (46)–(53) *admits a unique solution $(u_K)_{K\in\mathcal{T}}, (u_\sigma)_{\sigma\in\mathcal{E}_3}$.*

We may therefore define the approximate solution $u_\mathcal{T}$ from a.e. in $\Omega \cup \Gamma^3$ to $\mathbb{R}$ by

$$u_\mathcal{T}(x) = u_K \text{ for } x \in K \text{ and } K \in \mathcal{T}, \qquad u_\mathcal{T}(x) = u_\sigma \text{ for } x \in \sigma \text{ and } \sigma \in \mathcal{E}_3. \tag{54}$$

Remark 3.1. Under regularity assumptions on the exact solution, we give in [19] an estimate of order 1 with respect to the mesh size for the "discrete" H^1 norm and L^2 norm of the error on the solution. If the exact solution is no longer assumed to be regular, the convergence of the discrete solution towards the exact solution may still be proven; see [19].

The monotonic algorithm is again based on the obvious remark that, for a given $\sigma \in \mathcal{E}_3$, (53) is equivalent to $u_\sigma = a$ or $-F_{K,\sigma} = \mathrm{m}(\sigma)\, b$. Hence there exist two disjoint subsets of $\mathcal{E}_3$ such that on one subset one has

$$u_\sigma = a \quad \text{and} \quad -F_{K,\sigma} \geq \mathrm{m}(\sigma)\, b,$$

and on the other one,

$$-F_{K,\sigma} = \mathrm{m}(\sigma)\, b \quad \text{and} \quad u_\sigma \geq a.$$

Now if the subsets $\mathcal{E}_a$ and $\mathcal{E}_b$ of $\mathcal{E}_3$ such that $\mathcal{E}_a \cup \mathcal{E}_b = \mathcal{E}_3$, $\mathcal{E}_a \cap \mathcal{E}_b = \emptyset$ and such that

$$-F_{K,\sigma} \geq \mathrm{m}(\sigma)\, b \quad \forall\sigma \in \mathcal{E}_a, \tag{55}$$

$$u_\sigma \geq a \quad \forall\sigma \in \mathcal{E}_b \tag{56}$$

were known, then the solution to problem (46)–(53) could be obtained by solving the linear problem (46)–(50), together with

$$u_\sigma = a \quad \forall\sigma \in \mathcal{E}_a, \tag{57}$$

$$-F_{K,\sigma} = \mathrm{m}(\sigma)\, b \quad \forall\sigma \in \mathcal{E}_b. \tag{58}$$

The algorithm which follows determines the sets $\mathcal{E}_a$ and $\mathcal{E}_b$ by an iterative method.

MONOTONIC ALGORITHM, SIGNORINI PROBLEM, FINITE VOLUME DISCRETIZATION.

• Initialization. Let $\mathcal{E}_a^{(0)}$ and $\mathcal{E}_b^{(0)} \subset \mathcal{E}_3$ be such that

$$\mathcal{E}_a^{(0)} \cap \mathcal{E}_b^{(0)} = \emptyset \quad \text{and} \quad \mathcal{E}_a^{(0)} \cup \mathcal{E}_b^{(0)} = \mathcal{E}_3. \tag{59}$$

• Step (j). Assume that the sets $\mathcal{E}_a^{(j)}$ and $\mathcal{E}_b^{(j)}$ are known such that $\mathcal{E}_a^{(j)} \cap \mathcal{E}_b^{(j)} = \emptyset$ and $\mathcal{E}_a^{(j)} \cup \mathcal{E}_b^{(j)} = \mathcal{E}_3$.

Let $u_\mathcal{T}^{(j)} \in X(\mathcal{T})$ be defined by $u_\mathcal{T}^{(j)}(x) = u_K^{(j)}$ for $x \in K$, $\forall K \in \mathcal{T}$, and by $u_\mathcal{T}^{(j)}(x) = u_\sigma^{(j)}$ for $x \in \sigma$, $\forall \sigma \in \mathcal{E}_3$, and let $u_\mathcal{T}^{(j)}$ be the solution to the set of equations (46)–(50) and

$$u_\sigma^{(j)} = a \quad \forall \sigma \in \mathcal{E}_a^{(j)}, \tag{60}$$

$$F_{K,\sigma}^{(j)} = -\mathrm{m}(\sigma)\, b \quad \forall \sigma \in \mathcal{E}_b^{(j)}. \tag{61}$$

Let $\mathcal{E}_a^{(j+1)}$ and $\mathcal{E}_b^{(j+1)}$ be defined in the following way:

$$\begin{array}{ll} \mathcal{E}_a^{(j,0)} = \{\sigma \in \mathcal{E}_a^{(j)}; -F_{K,\sigma}^{(j)} \geq \mathrm{m}(\sigma)\, b\}, & \mathcal{E}_a^{(j,1)} = \mathcal{E}_a^{(j)} \setminus \mathcal{E}_a^{(j,0)}, \\ \mathcal{E}_b^{(j,0)} = \{\sigma \in \mathcal{E}_b^{(j)}; u_\sigma^{(j)} \geq a\}, & \mathcal{E}_b^{(j,1)} = \mathcal{E}_b^{(j)} \setminus \mathcal{E}_b^{(j,0)}, \\ \mathcal{E}_a^{(j+1)} = \mathcal{E}_a^{(j,0)} \cup \mathcal{E}_b^{(j,1)}, & \mathcal{E}_b^{(j+1)} = \mathcal{E}_3 \setminus \mathcal{E}_b^{(j+1)}. \end{array} \tag{62}$$

• The algorithm stops if there exists a step (J) such that $\mathcal{E}_a^{(J)} = \mathcal{E}_a^{(J+1)}$ and $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$.

The above algorithm is well defined, thanks to the following result.

PROPOSITION 3.3. *Under Assumption* 3.1, *let $\mathcal{T}$ be an admissible finite volume mesh in the sense of Definition* 2.8; *then problem* (46)–(50), (60)–(61) *has a unique solution $u_\mathcal{T}^{(j)}$.*

Proof. Under Assumption 3.1, let $\mathcal{T}$ be an admissible finite volume mesh in the sense of Definition 2.8, and $u_\mathcal{T}^{(j)} \in X(\mathcal{T})$ be defined by $u_\mathcal{T}^{(j)}(x) = u_K^{(j)}$ for $x \in K$, $\forall K \in \mathcal{T}$, and by $u_\mathcal{T}^{(j)}(x) = u_\sigma^{(j)}$ for $x \in \sigma$, $\forall \sigma \in \mathcal{E}_3$, and let the sets $\mathcal{E}_a^{(j)}$ and $\mathcal{E}_b^{(j)}$ be such that $\mathcal{E}_a^{(j)} \cap \mathcal{E}_b^{(j)} = \emptyset$ and $\mathcal{E}_a^{(j)} \cup \mathcal{E}_b^{(j)} = \mathcal{E}_3$. It is easily seen that $u_\mathcal{T}^{(j)}$ is solution to problem (46)–(50), (60)–(61) if and only if $u_\mathcal{T}^{(j)}$ is a solution to the following problem:

$$\left\{ \begin{array}{l} u_\mathcal{T}^{(j)} \in \mathcal{K}_\mathcal{T}^{(j)} = \{v \in X(\mathcal{T}) \text{ s.t. } v_\sigma = a\, \forall \sigma \subset \mathcal{E}_a^{(j)}\} \text{ such that} \\ \mathcal{A}(u_\mathcal{T}^{(j)}, v) = \mathcal{L}^{(j)}(v) \quad \forall v \in X(\mathcal{T}) \text{ s.t. } v_\sigma = 0\, \forall \sigma \in \mathcal{E}_a^{(j)}, \end{array} \right. \tag{63}$$

with

$$\begin{aligned} \mathcal{A}(u,v) = & \sum_{\sigma = K|L \in \mathcal{E}_{\mathrm{int}}} \tau_{K|L}(u_K - u_L)(v_K - v_L) + \sum_{\sigma \in \mathcal{E}_K,\, \sigma \subset \Gamma^1} \tau_\sigma u_K v_K \\ & + \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_3} \tau_\sigma (u_\sigma - u_K)(v_\sigma - v_K) \qquad \forall u,\ v \in X(\mathcal{T}), \end{aligned} \tag{64}$$

$$\mathcal{L}^{(j)}(v) = \sum_{\sigma \in \mathcal{E}_b^{(j)}} b v_\sigma \mathrm{m}(\sigma) \quad \forall v \in X(\mathcal{T}). \tag{65}$$

Then the existence and uniqueness of the solution to (46)–(50) follows by the Lax–Milgram lemma. $\square$

Let us now turn to the monotonicity property of the algorithm.

LEMMA 3.4 (monotonicity). *Under Assumption* 3.1, *let* $\mathcal{T}$ *be an admissible finite volume mesh in the sense of Definition* 2.8; *the sequences* $(u_K^{(j)})_{j\in\mathbb{N}}$, $K \in \mathcal{T}$, *and* $(u_\sigma^{(j)})_{j\in\mathbb{N}}$, $\sigma \in \mathcal{E}_3$, *which are constructed by the algorithm* (59)–(62), *satisfy*

$$
\begin{aligned}
&u_K^{(j)} \le u_K^{(j+1)} \quad \forall j \in \mathbb{N} \text{ and for } K \in \mathcal{T},\\
&u_\sigma^{(j)} \le u_\sigma^{(j+1)} \quad \forall j \in \mathbb{N} \text{ and for } \sigma \in \mathcal{E}_3.
\end{aligned}
\tag{66}
$$

Proof. Let v be defined by $v = u_{\mathcal{T}}^{(j+1)} - u_{\mathcal{T}}^{(j)}$, and let $\min(v_{\mathcal{T}})$ be defined by

$$
\min(v_{\mathcal{T}}) = \min\left\{\min_{K\in\mathcal{T}} v_K, \min_{\sigma\in\mathcal{E}_3} v_\sigma\right\}.
$$

We note that v satisfies the set of equations (46)–(50).

- Assume first that $\min(v_{\mathcal{T}}) = v_{K_0}$, with $K_0 \in \mathcal{T}$ such that $\partial K_0 \cap \Gamma^1 = \emptyset$ or $\partial K_0 \cap \Gamma^1$ is a point. Since $v_{K_0} \le v_K$ $\forall K \in \mathcal{T}$ and $v_{K_0} \le v_\sigma$ $\forall \sigma \in \mathcal{E}_3$, one has $\min(v_{\mathcal{T}}) = v_K$ $\forall K \in \mathcal{T}$, and $\min(v_{\mathcal{T}}) = v_\sigma$ $\forall \sigma \in \mathcal{E}_3$. Therefore, the minimum is reached on a control volume neighboring Γ^1, or on an edge included in Γ^3.
- Assume next that $\min(v_{\mathcal{T}}) = v_{K_0}$, with $K_0 \in \mathcal{T}$ such that there exists $\sigma \subset \partial K_0 \cap \Gamma^1$; from (46)–(50), we deduce that $\min(v_{\mathcal{T}}) \ge 0$.
- Now assuming $\sigma \in \mathcal{E}_b^{(j)} \cap \mathcal{E}_b^{(j+1)}$ and $\min(v_{\mathcal{T}}) = v_\sigma$, we obtain $-\tau_\sigma(v_\sigma - v_K) = 0$ with K such that $\partial K \cap \sigma = \sigma$; then $\min(v_{\mathcal{T}}) = v_K\, \forall K \in \mathcal{T}$ and $\min(v_{\mathcal{T}}) = v_\sigma$ $\forall \sigma \in \mathcal{E}_3$.
- Next if $\sigma \in \mathcal{E}_b^{(j)} \cap \mathcal{E}_a^{(j+1)}$ and $\min(v_{\mathcal{T}}) = v_\sigma$, one has $u_\sigma^{(j)} < a$ and $u_\sigma^{(j+1)} = a$; hence $\min(v_{\mathcal{T}}) > 0$.
- Finally if $\sigma \in \mathcal{E}_a^{(j)} \cap \mathcal{E}_b^{(j+1)}$ and $\min(v_{\mathcal{T}}) = v_\sigma$, one has $-\tau_\sigma(u_\sigma^{(j)} - u_K^{(j)}) < \mathrm{m}(\sigma) b$ and $-\tau_\sigma(u_\sigma^{(j+1)} - u_K^{(j+1)}) = \mathrm{m}(\sigma) b$; therefore $-\tau_\sigma(v_\sigma - v_K) > 0$, which is in contradiction with $\min(v_{\mathcal{T}}) = v_\sigma$. ∎

We now turn to the convergence of the algorithm.

PROPOSITION 3.5. *Assume that there exists a step* (J) *such that* $\mathcal{E}_a^{(J)} = \mathcal{E}_a^{(J+1)}$ *and* $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$ *and let* $u_{\mathcal{T}}^{(J)}$ *be the solution to* (46)–(50), (60)–(61); *then* $u_{\mathcal{T}}^{(J)}$ *is the unique solution to problem* (46)–(53).

Proof. Let $\mathcal{E}_a = \mathcal{E}_a^{(J)}$, $\mathcal{E}_b = \mathcal{E}_b^{(J)}$, and $u_{\mathcal{T}} = u_{\mathcal{T}}^{(J)}$; hence $u_{\mathcal{T}}$ satisfies the set of equations (46)–(50) and

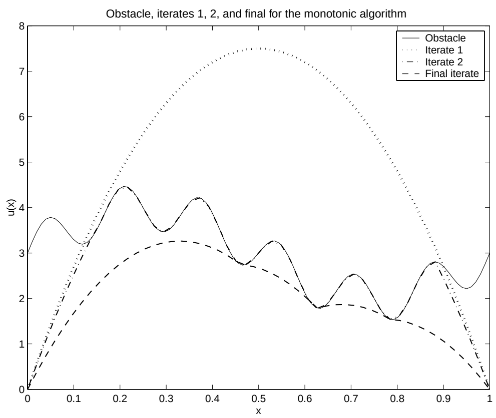
$$
\begin{aligned}
u_\sigma &= a \quad \forall \sigma \in \mathcal{E}_a,\\
F_{K,\sigma} &= -\mathrm{m}(\sigma)\, b \quad \forall \sigma \in \mathcal{E}_b.
\end{aligned}
$$

Since $\mathcal{E}_a^{(J)} = \mathcal{E}_a^{(J+1)}$, one has $F_{K,\sigma} \ge -\mathrm{m}(\sigma)\, b$ $\forall \sigma \in \mathcal{E}_a$, and since $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$, one has $u_\sigma \ge a$ $\forall \sigma \in \mathcal{E}_b$. Therefore, since $\mathcal{E}_a \cup \mathcal{E}_b = \mathcal{E}_3$, $u_{\mathcal{T}}$ satisfies the set of equations (46)–(53). ∎

THEOREM 3.6. *Under Assumption* 3.1, *there exists an integer* $J \in \mathbb{N}$ *such that the sequences* $(u_K^{(j)})_{j\in\mathbb{N}}$ $(u_\sigma^{(j)})_{j\in\mathbb{N}}$, *which are constructed by the algorithm* (31)–(36), *are such that* $(u_K^{(j)}, K \in \mathcal{T})$, $(u_\sigma^{(j)}, K \in \mathcal{E}_3)$ *is the exact solution to the discrete problem* (24) *for all* $j \ge J$. *Furthermore the integer* J *satisfies*

$$
J \le \mathrm{card}(\mathcal{E}_3) + 1,
\tag{67}
$$

where $\mathrm{card}(\mathcal{E}_3)$ *denotes the number of edges of the mesh which are on the Signorini boundary* Γ_3.



N	ITER
20	6
50	7
100	8
500	31
1000	130

FIG. 2. *One-dimensional obstacle problem: the obstacle and some iterates (left), the number of discretization points N and number of iterations to convergence ITER (right).*

Proof. Let the sets $\mathcal{E}_a^{(j)}$ and $\mathcal{E}_b^{(j)}$ be defined by the algorithm (59)–(62) for any step (j); if there exists an integer J such that $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$, then by Proposition 2.12, $(u_K^{(J)}, K \in \mathcal{T})$ is the exact solution to the discrete problem (24), and the first part of the theorem is proven. There remains to prove that such a step exists and that it satisfies (67).

Similarly to the remark on the obstacle problem, let us first note that for $\sigma \in \mathcal{E}_3$ if $u_\sigma^{(0)} \geq a$, then $u_\sigma^{(1)} \geq a$ by Lemma 2.11, and if $u_\sigma^{(0)} < a$, then $u_\sigma^{(1)} = a$ by step (62) of the algorithm. Hence

$$u_\sigma^{(1)} \geq a \text{ for any } \sigma \in \mathcal{E}_3; \tag{68}$$

therefore by an easy induction one has that $\mathcal{E}_a^{(j)} \subset \mathcal{E}_a^{(j+1)}$ for any $j > 1$. Since $\mathcal{E}_a^{(j)} \subset \mathcal{E}_3$, this implies that there exists an index J such that $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$.

Let us now prove that (67) holds. Let J be the smallest integer such that $\mathcal{E}_b^{(J)} = \mathcal{E}_b^{(J+1)}$. Since $\mathcal{E}_b^{(J)} \subset \mathcal{E}_3$ $\forall j \in \mathbb{N}$, one has $\mathrm{card}(\mathcal{E}_3) + 1 \geq \mathrm{card}(\mathcal{E}_b^{(j+1)}) \geq \mathrm{card}(\mathcal{E}_b^{(j)}) + 1$ for any $j > 1$, which yields that $J \leq \mathrm{card}(\mathcal{E}_3) + 1$. ◻

3.1. Numerical tests. In order to test the efficiency of this new algorithm, some numerical experiments were performed. We first tested the algorithm on a one-dimensional obstacle problem, discretized by either the finite volume or the finite element method (in the one-dimensional case the two schemes differ by only the right-hand side and the boundary conditions). The results proved excellent. We show in Figure 2 a few iterations for $\psi(x) = 3 + \frac{1}{2}\sin(12\varphi_i x) + \sin(2\varphi_i x)$ and for a right-hand side equal to 1. We also give the number of iterations (ITER) required to convergence versus the number N of discretization points. Recall that we have a theoretical bound ITER $\leq$ N: the results show that this bound is far from optimal. Note also that we have taken the solution of the unconstrained problem (i.e., the solution of $-u'' = f$) as an initial guess; of course, one could decrease the number of iterations by taking a better-chosen initial guess.

A less academic study was performed by Herbin and Marchand [20] for a finite volume discretization of an electrochemical problem involving a Signorini boundary

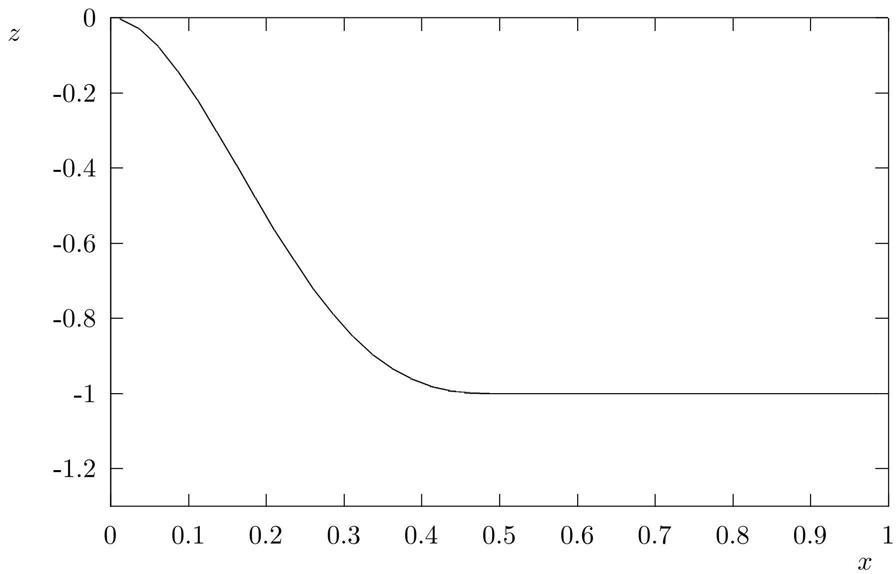


FIG. 3. $z = u(x, 0)\ \forall x \in [0, x_m]$ *(on* Γ^3*: Signorini boundary).*

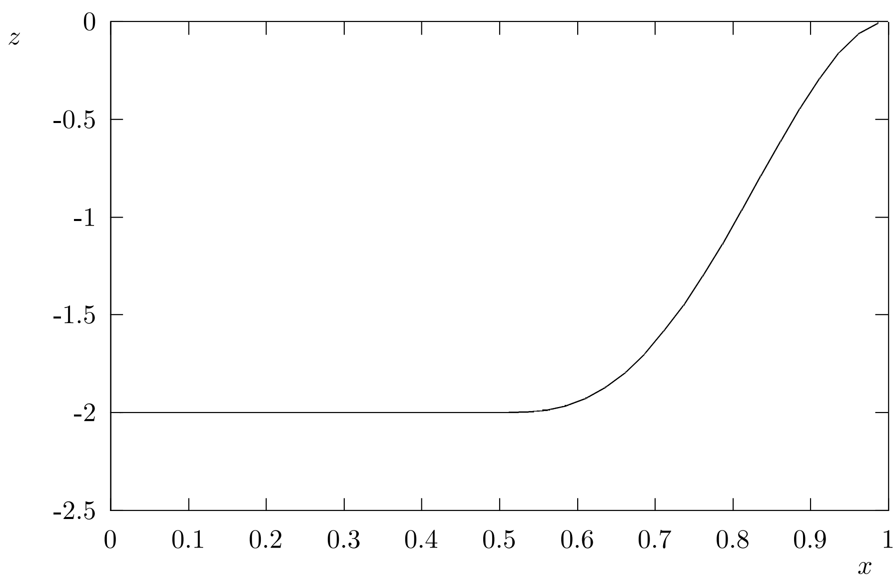


FIG. 4. $z = \nabla u \cdot \mathbf{n}(x, 0)\ \forall x \in [0, x_m]$ *(on* Γ^3*: Signorini boundary).*

condition, which was introduced in [23], and which involved a two-dimensional problem. These results illustrate the performance of the monotonic algorithm quite well.

The domain Ω is taken to be the rectangle ($\Omega =]0, x_m[\ \times\ ]0, y_m[$). We set the data such that the exact solution $u \in C^2(\overline{\Omega})$ is known. We show in Figures 3 and 4 the plots of the traces of u and of $\nabla u \cdot \mathbf{n}$ on the Signorini boundary Γ^3.

A rectangular mesh is used on the domain Ω. We vary the discretization step and the initial guess $\mathcal{E}_a^{(0)}$ and give the number of iterates required to converge to the (exact) solution of the discrete problem. Table 1 gives some results when taking the number of cells between 100 and 2500 with a uniform step.

TABLE 1
Number of iterations needed for the monotonicity algorithm.

$\mathcal{E}_3$ / grid size	10×10	20×20	30×30	40×40	50×50
$\mathcal{E}_{\text{ext}}$	4	6	7	7	8
$\emptyset$	4	6	7	7	7
$\{\sigma \subset \{x_m\} \times [0, x_m/2]\}$	4	7	7	7	8
$\{\sigma \subset \{x_m\} \times [x_m/2, x_m]\}$	2	3	2	2	2

These results show that the algorithm is quite efficient. The number of iterations does not vary much with respect to the grid size, and it is considerably less than the number of cells on $\mathcal{E}_3$, which was the upper bound given by Theorem 3.6. Of course, if one has a hint of how the Signorini boundary should be, then a good initial guess lowers the number of iterations, as may be seen from the last line of the table.

Let us also recall that at each iteration there is only one solve of a linear subproblem to be performed. Hence there are no ill-conditioned systems involved, such as those in penalty methods. Finally, let us point out that this algorithm may also be successfully implemented for other free boundary problems: we also tested it on the dam problem and it performs well. It is also used in multiphase problems [8], although in this last case no theoretical convergence result is known (in fact, existence and uniqueness of the solution are an open problem in this last case).

4. Conclusion. The monotonic algorithm which we have introduced for both the obstacle problem and the Signorini problems has been shown to be convergent for the linear finite element and finite volume discretizations in the case of the obstacle problem and the finite volume discretization in the case of the Signorini problem. Furthermore, a bound of the number of iterations is known. An important advantage of this algorithm is that, at each iteration, it necessitates only a linear solve involving a submatrix of the diffusion operator, and therefore no ill-conditioned system must be solved as would be the case with a penalty method. The actual implementation of the algorithm is very easy, and the computational cost is low, since the number of iterations is much lower than the theoretical bound, that is, the number of cells for which the constraint holds, even when the initial guess is not well chosen.

The limitation of the method is linked to the fact that it is proven to converge thanks to the discrete maximum principle, which holds for adequate discretizations of diffusion problems such as the ones we considered here. Note that the discretization is originally chosen such that the maximum principle holds, not because of the monotonic algorithm, but because the discrete maximum principle reflects a physical constraint which the approximate solution needs to satisfy (in the case of a chemical diffusion, the concentration should stay between 0 and 1). Hence it is natural in this type of problem to use the monotonic algorithm.

However, the efficiency (and proof of convergence) of the monotonic algorithm when the maximum principle does not hold (elasticity, higher order finite elements) is still an open question.

REFERENCES

[1] Z. BELHACHMI AND F. BEN BELGACEM, *Eléments finis d'ordre deux pour l'inéquation variationelle de Signorini*, C. R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 727–732.

[2] F. BEN BELGACEM, *Numerical simulation of some variational inequalities arisen from unilateral contact problems by the finite element methods*, SIAM J. Numer. Anal., 37 (2000), pp. 1198–1216.

[3] F. BEN BELGACEM, *Méthodes d'éléments finis pour les inéquations variationelles de contact unilatéral*, C. R. Acad. Sci. Paris Sér. I Math., 328 (1999), pp. 811–816.

[4] H. BREZIS AND G. STAMPACCHIA, *Sur la régularité de la solution d'inéquations elliptiques*, Bull.

Soc. Math. France, 96 (1968), pp. 153–180.

[5] F. Brezzi, W. W. Hager, and P. A. Raviart, *Error estimates for the finite element solution of variational inequalities Part* I—*Primal theory*, Numer. Math., 28 (1977), pp. 431–443.

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1998.

[7] G. Duvaut and J. L. Lions, *Inequalities in Mechanics and Physics*, Springer–Verlag, Berlin, 1976.

[8] R. Eymard and T. Gallouët, *Traitement de changement de phase dans la modelisation de gisements pétroliers*, Journées Numériques de Besançon, J.M. Crolet and P. Lesaint, eds., Université de Besançon, Besançon, France, 1991.

[9] R. Eymard, T. Gallouët, and R. Herbin, *Finite volume methods*, in Handb. Numer. Anal. 7, P. G. Ciarlet and J. L. Lions, eds., North–Holland, Amsterdam, 2000, pp. 713–1020.

[10] R. Falk, *Error estimates for the approximation of a class of variational inequalities*, Math. Comp., 28 (1974), pp. 963–997.

[11] R. Falk and B. Mercier, *Error estimates for elastoplastic problems*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér., 11 (1977), pp. 135–144.

[12] G. Fichera, *Problemi elastotatici con vincoli unilateral: Il problema di Signorini con ambigue condizioni al contorno*, Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Natur. Sez. Ia., 7 1963–1964, pp. 91–140.

[13] G. Fichera, *Boundary value problems of elasticity with unilateral constraints*, in Handbuch der Physik, Vol. VI a/2, Springer–Verlag, Berlin, 1972, pp. 391–424.

[14] T. Gallouët, R. Herbin, and M. H. Vignal, *Error estimates on the approximate finite volume solution of convection diffusion equations with general boundary conditions*, SIAM J. Numer. Anal., 37 (2000), pp. 1935–1972.

[15] R. Glowinski, *Lectures on Numerical Methods for Non-Linear Variational Problems*, notes by M.G. Vijayasundaram and M. Adimurthi, Tata Institute of Fundamental Research Lectures on Mathematics and Physics 65, Tata Institute of Fundamental Research, Bombay, India, Springer-Verlag, Berlin, New York, 1980.

[16] R. Glowinski, J. L. Lions, and R. Trémolières, *Analyse numérique des inequations variationnelles*, Dunod, Paris, 1976.

[17] S. Gerbi, R. Herbin, and E. Marchand, *Existence of a solution to a coupled elliptic system with a Signorini condition*, Adv. Differential Equations, 4 (1999), pp. 225–250.

[18] R. Herbin, *An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh*, Numer. Methods Partial Differential Equations, 11 (1995), pp. 165–173.

[19] R. Herbin and E. Marchand, *Finite volume approximation of a class of variational inequalities*, IMA J. Numer. Anal., 21 (2001), pp. 553–585.

[20] R. Herbin and E. Marchand, *Numerical approximation of a nonlinear problem with a Signorini boundary condition*, in Iterative Methods in Scientific Computations, J. Wang, M. B. Allen, B. Chen, and T. Matthew, eds., IMACS Series in Computational and Applied Mathematics 4, IMACS, New Brunswick, NJ, 1998, pp. 283–288.

[21] N. Kikuchi and J. T. Oden, *Contact Problems in Elasticity, A Study of Variational Inequaltities and Finite Element Methods*, SIAM Stud. Appl. Math. 8, SIAM, Philadelphia, 1988.

[22] N. Kikuchi and Y. J. Song, *Contact problems involving forces and movements for incompressible linearly elastic materials*, Internat. J. Engrg. Sci., 18 (1980), pp. 357–377.

[23] M. Kleitz, L. Dessemond, R. Jimenez, F. Petitbon, R. Herbin, and E. Marchand, *Micromodelling of the cathode and experimental approaches*, in Proceedings of the Second European Solid Oxide Fuel Cell Forum, Oslo, Norway, 1996, B. Thortensen, ed., Dr. Ulf Bossel, Oberrohrdorf, Switzerland, pp. 317–324.

[24] H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Comm. Pure Appl. Math., 22 (1969), pp. 153–188.

[25] J. L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math., 20 (1969), pp. 493–519.

[26] U. Mosco and G. Strang, *One sided approximation and variational inequalities*, Bull. Amer. Math. Soc., 80 (1974), pp. 308–312.

[27] L. Slemane, A. Bendali, and P. Laborde, *Mixed formulations for a class of variational inequalities*, C. R. Acad. Sci. Paris Sér. I Math., 334 (2002), pp. 87–92.

[28] G. Stampacchia, *Formes bilinéaire coercivitives sur les ensembles convexes*, C. R. Acad. Sci. Paris, 258 (1964), pp. 4413–4416.

[29] G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.

[30] G. Stampacchia, *Equations Elliptiques du Second Ordre à coefficients discontinus*, les Presses de l'université de Montréal, Montréal, 1966.