

SHAPE AND TOPOLOGY OPTIMIZATION FOR VARIATIONAL INEQUALITIES WITH POINTWISE BOUNDARY OBSERVATION*

Cornel Marius Murea[†] Dan Tiba[‡]

Dedicated to Prof. Liliana Restuccia on the occasion of her 70th anniversary

DOI 10.56082/annalsarscimag.2025.3.83

Abstract

We study optimal design problems involving variational inequalities with unilateral conditions in the domain and pointwise boundary observation. We use regularizing and penalization techniques in the setting of the Hamiltonian approach to shape/topology optimization problems. Numerical examples are also included.

Keywords: optimal design, pointwise observation, nonsmooth variational inequalities.

MSC: 35J87, 49M41.

1 Introduction

We consider a fixed domain approach, based on level functions for the (unknown) geometry and on the penalization/regularization of the state system. The shape and topology optimization problem is associated to nonlinear

*Accepted for publication on September 15, 2025

[†]`cornel.murea@uha.fr`, University of Haute Alsace, Department of Mathematics, Mulhouse, France

[‡]`dan.tiba@imar.ro`, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania & Academy of Romanian Scientists, Ilfov 3, 050044 Bucharest, Romania

second order elliptic equations with Dirichlet boundary conditions. The optimal design problem can be approximated, via penalization, by an elliptic optimal control problem (see [12]).

Our methodology is underpinned by functional variations [13], [10]. For linear and nonlinear state systems, it is known that both closing and creating new holes is possible. Moreover, we prove differentiability and we apply standard gradient algorithms in the optimization process, that are known to converge locally, [3]. For a detailed discussion of optimal control problems governed by nonlinear and nonsmooth state systems, see [2], [14].

In the setting of functional variations, one assumes that all the admissible domains Ω are contained in a given bounded domain $D \subset \mathbb{R}^2$. We denote by $\mathcal{O}$ their family. We assume that this is generated by a family $\mathcal{F} \subset \mathcal{C}(\overline{D})$ of admissible level functions, as open sublevel sets:

$$\Omega = \Omega_g = \text{int}\{\mathbf{x} \in D : g(\mathbf{x}) \leq 0\}, \quad g \in \mathcal{F}. \quad (1)$$

We recall that functional variations have the form

$$g + \lambda h, \quad \lambda \in \mathbb{R}, \quad g, h \in \mathcal{F}, \quad (2)$$

which enables the work in functional spaces instead of the very specialized geometric variations.

To generate domains, we add to (1) a condition that selects just the component satisfying ($\mathbf{x}_0 \in D$ is fixed)

$$\mathbf{x}_0 \in \partial\Omega_g, \quad \forall g \in \mathcal{F}. \quad (3)$$

Notice that, for any admissible $g \in \mathcal{F}$, we have $g(\mathbf{x}_0) = 0$.

A general shape and topology optimization problem has the form (for $\Omega_g \in \mathcal{O}$ defined as above):

$$\min_{g \in \mathcal{F}} J(g, y_g), \quad (4)$$

$$Ay_g = f \quad \text{in } \Omega_g, \quad (5)$$

$$By_g = 0 \quad \text{on } \partial\Omega_g. \quad (6)$$

In (4) - (6), A is an elliptic partial differential operator, B is a boundary operator, defining the state system and J is some cost functional. The operators may be linear or nonlinear and appropriate assumptions will be imposed as necessity appears.

Assume now that $\mathcal{F} \subset \mathcal{C}^1(\overline{D})$ and

$$\nabla g(\mathbf{x}) \neq 0, \quad \forall g \in \mathcal{F}, \quad \forall \mathbf{x} \in D, \quad \text{with } g(\mathbf{x}) = 0. \quad (7)$$

Assume, as well, that

$$g > 0 \quad \text{on } \partial D, \quad \forall g \in \mathcal{F}, \quad (8)$$

which ensures that $\partial\Omega_g \cap \partial D = \emptyset$. If, moreover, the open cone $\mathcal{F} \subset \mathcal{C}^2(D)$ is given by (7), (8), then the solution of the Hamiltonian system below is periodic and gives a global parameterization of $\partial\Omega_g$

$$z_1'(t) = -\frac{\partial g}{\partial x_2}(z_1(t), z_2(t)), \quad t \in I_g, \quad (9)$$

$$z_2'(t) = \frac{\partial g}{\partial x_1}(z_1(t), z_2(t)), \quad t \in I_g, \quad (10)$$

$$(z_1(0), z_2(0)) = \mathbf{x}_0 \in \partial\Omega_g, \quad (11)$$

by [10], [6]. We may take $I_g = [0, T_g]$, where $T_g > 0$ denotes the principal period in (9)-(11). Moreover, the dependence $g \in \mathcal{F} \rightarrow T_g$ has differentiability properties, due to [10] if the family of shape functions satisfies $\mathcal{F} \subset \mathcal{C}^2(D)$ and (7), (8).

In every $\Omega \in \mathcal{O}$, $\Omega = \Omega_g \subset D$, $g \in \mathcal{F}$, we consider the variational inequality:

$$-\Delta y_g + \beta(y_g) \ni f \quad \text{in } \Omega_g, \quad (12)$$

$$y_g = 0 \quad \text{on } \partial\Omega_g, \quad (13)$$

where $f \in L^2(D)$ is given and $\beta \subset \mathbb{R} \times \mathbb{R}$ is a maximal monotone graph satisfying

$$0 \in \beta(0). \quad (14)$$

The nonlinear boundary value problem (12), (13) includes as an example the obstacle problem with obstacle $\varphi = 0$. In case the obstacle $\varphi : D \rightarrow \mathbb{R}$ is some other function (negative in D , to ensure compatibility with (13) for any $\Omega \in \mathcal{O}$), then $\beta(y_g - \varphi)$ should be written in (12).

It is known [4], that (12), (13) has a unique solution $y_g \in H^2(\Omega_g) \cap H_0^1(\Omega_g)$ with $\beta(y_g) \in L^2(\Omega_g)$, if $\partial\Omega_g$ is in $\mathcal{C}^{1,1}$. Therefore the formulation (4), (12), (13) makes sense. The regularity of $\partial\Omega_g$ may be obtained from the implicit functions theorem since $\mathcal{F} \subset \mathcal{C}^2(D)$ and (7) is satisfied (we get even $\mathcal{C}^2$ regularity for $\partial\Omega_g$).

We penalize and regularize (12) as follows ($\varepsilon > 0$ is “small”):

$$-\Delta y_\varepsilon + \beta_\varepsilon(y_\varepsilon) + \frac{1}{\varepsilon} H_\varepsilon(g) y_\varepsilon = f \quad \text{in } D, \quad (15)$$

$$y_\varepsilon = 0 \quad \text{on } \partial D. \quad (16)$$

Above, β_ε combines the Yosida approximation β^ε with Friedrichs mollifiers:

$$\beta_\varepsilon(y) = \int_{-\infty}^{\infty} \beta^\varepsilon(y - \varepsilon s) \rho(s) ds, \quad (17)$$

with $\rho \in \mathcal{C}_0^\infty(\mathbb{R})$, $\text{supp}(\rho) \subset [-1, 1]$, $\rho(-s) = \rho(s)$, $\rho(s) \geq 0$ and $\int_{-\infty}^{\infty} \rho(s) ds = 1$.

Concerning $H(\cdot)$, it denotes the Heaviside function on $\mathbb{R}$ and its maximal monotone extension in $\mathbb{R} \times \mathbb{R}$ as well (i.e., we take $H(0) = [0, 1]$). The notation H_ε is obtained from $H \subset \mathbb{R} \times \mathbb{R}$ again via Yosida approximation and Friedrichs mollifiers, as in (17). Clearly, (15), (16) has a unique solution $y_\varepsilon \in H^2(D) \cap H_0^1(D)$ if ∂D is supposed to be of class $\mathcal{C}^{1,1}$.

This distributed penalization technique, using H, H_ε , is due to Natori and Kawarada [7], in the setting of free boundary problems, and employed in shape optimization, for instance, in [13], for linear elliptic equations.

We indicate an approximation result, following [11], under weak regularity assumptions (we do not require $g \in \mathcal{F}$, for the moment) on the domain Ω (class $\mathcal{C}$, i.e., continuous boundary or the segment property, see [1], Def. 4.2, p. 66, [15]). We recall here that Ω is of class $\mathcal{C}$, if there exists a locally finite open cover of $\partial\Omega$ and an associated sequence of vectors $\{y_j\}$, not null, such that for any $x \in \partial\Omega$ there is j with $x + ty_j \in \Omega$ for $0 < t < 1$ and this remains valid on the associated neighborhood of x , in the open cover of $\partial\Omega$.

Theorem 1. *If $\Omega = \Omega_g$ is of class $\mathcal{C}$, then, on a subsequence, $y_\varepsilon|_{\Omega_g} \rightarrow y_g$ weakly in $H^1(\Omega_g)$ and $\Delta y_\varepsilon \rightarrow \Delta y_g$, $\beta_\varepsilon(y_\varepsilon) \rightarrow \beta(y_g)$ weakly in $L^2(\Omega_g)$. If $g \in \mathcal{F}$, then $\frac{\partial y_\varepsilon}{\partial \mathbf{n}} \rightarrow \frac{\partial y_g}{\partial \mathbf{n}}$ in $H^{-1/2}(\partial\Omega_g)$ weakly.*

In the next section we discuss a special choice of the cost functional (4), namely in the so called pointwise observation case, that describes realistic, but computationally difficult situations.

2 Pointwise observation

In this section, an important role in the justification of the following computations and arguments is played by the topological stability result from [18], for small functional variational perturbations. Somewhat paradoxical, this stability result allows topological changes in optimal design since the corresponding steps in the descent gradient algorithms are not small, in general.

A natural and maybe the easiest case of pointwise boundary observation (taking into account (9)-(11)) is the following performance index:

$$\min_{g \in \mathcal{F}} \left\{ J_1 = \left| \frac{\partial y_\varepsilon}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right|^2 \right\}, \quad (18)$$

where $\alpha \in \mathbb{R}$ is given, y_ε is the solution of (15), (16) and $\mathbf{x}_0 \in \partial\Omega$ is the initial condition in (9)-(11). Other points may be taken into account as well, see (19). The normal derivative refers to the domain Ω_g and can be expressed as $\frac{\nabla y_\varepsilon \cdot \nabla g}{|\nabla g|}$ in any point of D . Notice that (18) makes sense due to

Remark 1. If $\mathcal{F} \subset \mathcal{C}^2$ and $f \in H^1(D)$, then $y_\varepsilon \in H^3(D)$ since $\beta_\varepsilon(y_\varepsilon)$ and $H_\varepsilon(y_\varepsilon)$ are in $H^2(D)$ due to the regularity of $\beta_\varepsilon, H_\varepsilon$ and we know that $y_\varepsilon \in H^2(D)$. Consequently, $\Delta y_\varepsilon \in H^1(D)$ and we get $y_\varepsilon \in H^3(D)$, by standard regularity results for the Laplace equation and $y_\varepsilon \in \mathcal{C}^1(\bar{D})$ by the Sobolev theorem. That is $\frac{\partial y_\varepsilon}{\partial \mathbf{n}}$ has a pointwise sense.

We have denoted by T_g the main period of the solution of (9)-(11) corresponding to $g \in \mathcal{F}$. Let l be a given natural number and divide $[0, T_g]$ (and implicitly $\partial\Omega_g$) in l parts starting from the initial condition (11) (that we may arbitrarily choose on $\partial\Omega_g$, in this case). The described process is like a discretization procedure. We define in this way a preassigned number of observation points on $\partial\Omega_g$ and another example of pointwise cost functional is

$$\min_{g \in \mathcal{F}} \left\{ J_2 = \sum_{i=0}^{l-1} \left| \frac{\partial y_\varepsilon}{\partial \mathbf{n}} \left(\mathbf{z} \left(\frac{iT_g}{l} \right) \right) - \alpha_i \right|^2 \right\}, \quad (19)$$

where $\alpha_i \in \mathbb{R}$ are given. A variant of (19) is to choose the observation points just on some part of $\partial\Omega_g$, by fixing the indices $l_1 \leq i \leq l_2$, with l_1, l_2 given natural numbers, in the corresponding sum (or by using another appropriate subset of indices).

Of course, one can define pointwise observation functionals, by using points in Ω , not on $\partial\Omega$ (and under the additional condition that $g \in \mathcal{F}$ are negative in such points). But this type of observation is less realistic, in general (it is difficult to place sensors inside a body) and we shall consider just the minimization of tracking type functionals as in (18), (19).

In the sequel, we study the differentiability properties of (18), (for (19) the arguments are quite similar). We recall first the following two results from [9], [16], [10], [11]. We consider the perturbations of the parameters (controls) $g + \lambda h$ and the system below (20)-(23) characterizes the variations of the corresponding solutions of the state system (15), (16), (9)-(11).

Furthermore, $\mathbf{z}_\lambda$ denotes the solution of (9)-(11) associated to $g + \lambda h$, T_λ denotes the main period of $\mathbf{z}_\lambda$ and we shall also write $y_\varepsilon(g), y_\varepsilon(g + \lambda h)$ when we emphasize the dependence on the level functions describing the geometry.

Proposition 1. *The system in variations corresponding to (15), (16), (9)-(11) is:*

$$\begin{aligned} -\Delta u + \beta'_\varepsilon(y_\varepsilon)u + \frac{1}{\varepsilon}H_\varepsilon(g)u + \frac{1}{\varepsilon}H'_\varepsilon(g)y_\varepsilon h &= 0 \quad \text{in } D, \\ u &= 0 \quad \text{on } \partial D. \end{aligned} \quad (20)$$

$$w'_1 = -\nabla \partial_2 g(\mathbf{z}) \cdot \mathbf{w} - \partial_2 h(\mathbf{z}), \quad \text{in } [0, T_g], \quad (21)$$

$$w'_2 = \nabla \partial_1 g(\mathbf{z}) \cdot \mathbf{w} + \partial_1 h(\mathbf{z}), \quad \text{in } [0, T_g], \quad (22)$$

$$w_1(0) = 0, \quad w_2(0) = 0, \quad (23)$$

where $u = \lim_{\lambda \rightarrow 0} \frac{y_\varepsilon(g + \lambda h) - y_\varepsilon(g)}{\lambda}$, $\mathbf{w} = [w_1, w_2] = \lim_{\lambda \rightarrow 0} \frac{\mathbf{z}_\lambda - \mathbf{z}}{\lambda}$ and the limits exist in $H^2(D)$, respectively $\mathcal{C}^1([0, T_g])^2$.

Proposition 2. *Under the assumptions (7), (8), if T_λ denotes the main period of the solution of (9)-(11), we have:*

$$\lim_{\lambda \rightarrow 0} \frac{T_\lambda - T_g}{\lambda} = -\frac{w_2(T_g)}{z'_2(T_g)}$$

if $z'_2(T_g) \neq 0$.

Remark 2. *If $z'_1(T_g) \neq 0$, the limit is $-\frac{w_1(T_g)}{z'_1(T_g)}$. In general, we denote by $\theta(g, r)$ this limit. Either the last condition in Proposition 2 or this condition here follow by (7). In case both conditions are valid, it is known that the two mentioned results coincide (this is a consequence of Proposition 1 in [17]).*

The directional derivatives of cost functionals can be computed as follows.

We consider the cost (18) and we notice (under the notations of Proposition 1 and by Remark 1 that the computations to obtain (20) can be slightly refined as follows (ε is fixed here):

$$\begin{aligned} &-\Delta(y_\varepsilon(g + \lambda h) - y_\varepsilon(g)) + \beta_\varepsilon(y_\varepsilon(g + \lambda h)) - \beta_\varepsilon(y_\varepsilon(g)) \\ &+ \frac{1}{\varepsilon}H_\varepsilon(g + \lambda h)(y_\varepsilon(g + \lambda h) - y_\varepsilon(g)) \\ &- \frac{1}{\varepsilon}[H_\varepsilon(g + \lambda h) - H_\varepsilon(g)]y_\varepsilon(g) = 0. \end{aligned} \quad (24)$$

Dividing (24) by $\lambda \neq 0$ and using Proposition 1, we get $\Delta \frac{y_\varepsilon(g+\lambda h) - y_\varepsilon(g)}{\lambda}$ bounded in $H^1(D)$ due to the regularity properties of $\beta_\varepsilon, H_\varepsilon$, that is

$$\frac{y_\varepsilon(g + \lambda h) - y_\varepsilon(g)}{\lambda} \rightarrow u$$

strongly in $C^1(D)$, due to the Sobolev theorem. Then, we can compute the derivative of the pointwise cost (18). First, we notice that:

$$\begin{aligned} L &= \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[\left| \frac{\partial y_\varepsilon(g + \lambda h)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right|^2 - \left| \frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right|^2 \right] \\ &= 2 \left(\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}_0). \end{aligned} \quad (25)$$

In (25), we have decomposed the first paranthesis in the product of the sum by the difference of the squared terms and we apply Proposition 1 and the above argument, to pass to the limit. We continue with the other term:

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[\left| \frac{\partial y_\varepsilon(g + \lambda h)}{\partial \mathbf{n}_\lambda}(\mathbf{x}_0) - \alpha \right|^2 - \left| \frac{\partial y_\varepsilon(g + \lambda h)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right|^2 \right]. \quad (26)$$

where $\mathbf{n}_\lambda$ is the normal in $\mathbf{x}_0$ to $\partial\Omega_\lambda$, Ω_λ being the domain associated to $g + \lambda h$ via (1), (3). After the decomposition of the difference of the squares, the important part to be investigated is:

$$\begin{aligned} &\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[\frac{\partial y_\varepsilon(g + \lambda h)}{\partial \mathbf{n}_\lambda}(\mathbf{x}_0) - \frac{\partial y_\varepsilon(g + \lambda h)}{\partial \mathbf{n}}(\mathbf{x}_0) \right] \\ &= \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[\frac{\nabla y_\varepsilon(g + \lambda h) \cdot \nabla(g + \lambda h)}{|\nabla(g + \lambda h)|}(\mathbf{x}_0) - \frac{\nabla y_\varepsilon(g + \lambda h) \cdot \nabla g}{|\nabla g|}(\mathbf{x}_0) \right] \end{aligned} \quad (27)$$

where we apply that $g = 0$ on $\partial\Omega$ and $g + \lambda h = 0$ on $\partial\Omega_\lambda$ and the well known formula for the normal derivative. By some calculations, we get the limit:

$$M = \left[-\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}} \frac{\partial h}{\partial \mathbf{n}} \frac{1}{|\nabla g|} + \frac{\nabla y_\varepsilon(g) \cdot \nabla h}{|\nabla g|} \right](\mathbf{x}_0).$$

For the first term above, we have used that:

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \left[\frac{1}{|\nabla(g + \lambda h)|} - \frac{1}{|\nabla(g)|} \right](\mathbf{x}_0) = -\frac{\partial h}{\partial \mathbf{n}}(\mathbf{x}_0) \frac{1}{|\nabla g(\mathbf{x}_0)|^2}. \quad (28)$$

Combining (25) - (28) with M , we infer:

Proposition 3. Assume that $\mathcal{F} \subset \mathcal{C}^2$, $f \in H^1(D)$ and (7), (8) are valid. Then, the directional derivative of the cost (18) with respect to functional variations $g + \lambda h$, g, h satisfying the above assumptions and $g(\mathbf{x}_0) = 0, h(\mathbf{x}_0) = 0$ is given by:

$$L + 2M \left(\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) = 2 \left(\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) \left[\frac{\partial u}{\partial \mathbf{n}} - \frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}} \frac{\partial h}{\partial \mathbf{n}} \frac{1}{|\nabla g|} + \frac{\nabla y_\varepsilon(g) \cdot \nabla h}{|\nabla g|} \right](\mathbf{x}_0) \quad (29)$$

A similar calculation can be performed for the pointwise cost functional (19), term by term.

Denote by $\mu \in L^2(D)$ the last term in (20). In fact, at least $H_0^1(D)$ regularity is valid for μ , under our assumptions, and $u \in H^3(D)$ depends linearly and continuously on μ in these spaces. Then, $\frac{\partial u}{\partial \mathbf{n}}$ and the first term in (29) depend linearly and continuously on $\mu \in H_0^1(D)$. There is a unique $p \in H^{-1}(D)$ such that the linear continuous functional on the left side below, can be expressed as:

$$2 \left(\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}_0) = \langle p, \mu \rangle = \langle p, \frac{1}{\varepsilon} H'_\varepsilon(g) y_\varepsilon h \rangle, \quad (30)$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing in $H^{-1}(D) \times H_0^1(D)$. We define the adjoint equation by rewriting (30) in the form:

$$2 \left(\frac{\partial y_\varepsilon(g)}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}_0) = - \langle p, -\Delta u + \beta'_\varepsilon(y_\varepsilon)u + \frac{1}{\varepsilon} H_\varepsilon(g)u \rangle, \quad (31)$$

where $u \in H^3(D) \cap H_0^1(D)$ has to be considered an arbitrary test function. This approach, called the transposition method, is due to J.L. Lions, see [8], Ch.2.5.4, where a related cost functional is considered. Moreover, one adds the null Dirichlet condition to (31), due to the employed class of test functions and a formal interpretation of equation (31) is:

$$-\Delta p + \beta'_\varepsilon(y_\varepsilon)p + \frac{1}{\varepsilon} H_\varepsilon(g)p = -P(y_\varepsilon)\delta_{\mathbf{x}_0}, \quad (32)$$

in D , where $\delta_{\mathbf{x}_0}$ is the Dirac distribution concentrated in $\mathbf{x}_0$ and $P(y_\varepsilon)$ is generated starting from the linear form in u :

$$2 \int_D \left(\frac{\nabla y_\varepsilon \cdot \nabla g}{|\nabla g|} - \alpha \right) \frac{\nabla u \cdot \nabla g}{|\nabla g|} dx. \quad (33)$$

We have proved:

Proposition 4. *The adjoint equation associated to the cost functional (18) and the state system (15), (16) is given by (31) and it has a unique transposition solution $p \in H^{-1}(D)$.*

Remark 3. *Notice that the arguments used in Proposition 4, Proposition 3 are valid in arbitrary dimension since we don't apply Remark 2 or Proposition 2. The treatment of example (19), where the main period appears explicitly, needs these results and is limited to dimension two.*

In the next section, we shall employ a direct approach based on Proposition 3, in the finite dimensional optimization problem obtained via discretization. Due to the very weak regularity properties of the adjoint state discussed in Proposition 4, the usual method to compute the gradient of the cost via the adjoint system, seems difficult to apply (but we indicate a regularized variant in one example).

3 Discretization and numerical tests

We assume $D \subset \mathbb{R}^2$ to be polygonal domain and $\mathcal{T}_h$ a mesh in D of size $h > 0$, each element T of the mesh is a triangle.

We set the $\mathbb{P}_1$ Lagrangian finite element spaces

$$\begin{aligned} W_h &= \{v_h \in \mathcal{C}(\overline{D}); \forall T \in \mathcal{T}_h, (v_h)|_T \in \mathbb{P}_1(T)\} \\ V_h &= \{v_h \in W_h; v_h = 0 \text{ on } \partial D\}. \end{aligned}$$

Let $\varphi_h \in W_h$ be an approximation of the obstacle φ . We set

$$\mathcal{K}_h = \{v_h \in V_h; v_h \geq \varphi_h \text{ in } D\}. \quad (34)$$

The discretized state equation, associated to (15), (16), in variational formulation, is: find $y_h \in \mathcal{K}_h$ such that

$$\begin{aligned} & \int_D \nabla y_h \cdot \nabla (y_h - v_h) d\mathbf{x} + \frac{1}{\varepsilon} \int_D H_\eta(g_h) y_h (y_h - v_h) d\mathbf{x} \\ & \leq \int_D f(y_h - v_h) d\mathbf{x}, \quad \forall v_h \in \mathcal{K}_h \end{aligned} \quad (35)$$

with $f \in L^2(D)$. Here $H_\eta \in \mathcal{C}^1(\mathbb{R})$, $\eta > 0$, is a regularization of the Heaviside function, given by

$$H_\eta(r) = \begin{cases} 1, & r > \eta \\ \frac{(-2r+3\eta)r^2}{\eta^3}, & r \in [0, \eta] \\ 0, & r < 0. \end{cases} \quad (36)$$

For the following example of maximal monotone graph $\beta \subset \mathbb{R} \times \mathbb{R}$,

$$\beta(r) = \begin{cases} 0, & r > 0 \\] -\infty, 0], & r = 0 \\ \emptyset, & r < 0 \end{cases} \quad (37)$$

we introduce the approximation $\beta_{\eta, \varepsilon_2} \in \mathcal{C}^1(\mathbb{R})$, ($\varepsilon_2 > 0$) given by

$$\beta_{\eta, \varepsilon_2}(r) = \begin{cases} 0, & r > 0 \\ r^2 \left(\frac{-r}{(\eta)^2 \varepsilon_2} - \frac{2}{\eta \varepsilon_2} \right), & r \in [-\eta, 0] \\ \frac{r}{\varepsilon_2}, & r < -\eta. \end{cases} \quad (38)$$

We take $\eta > \varepsilon$ and $\varepsilon_2 > \varepsilon$ to avoid some numerical difficulties, see [11].

Each $g_h \in W_h$ can be written as $g_h = \sum_{i=1}^n G_i \phi_i$ where $(\phi_i)_{1 \leq i \leq n}$ is the basis of the hat functions and $n = \dim(W_h)$. We introduce $\mathcal{J}_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\mathcal{J}_1(G) = J_1(g_h)$$

where $G = (G_i)_{1 \leq i \leq n}$, see (18).

We denote $I = \{1, \dots, n\}$ and let $I_0 \subset I$ be the indices of the vertices A_i of $\mathcal{T}_h$ such that $\|\mathbf{x}_0 - A_i\| < hC$ with $C \geq 2$. As an example, for the triangulation and $\mathbf{x}_0$ of Test 1, we get I_0 of 13 elements for $C = 2$ and 33 elements for $C = 3$.

Another choice for I_0 is the set of indices of the vertices of the triangle $T \in \mathcal{T}_h$ containing $\mathbf{x}_0$. In this case, the parameter C is not needed.

If $i \in I \setminus I_0$, then $\mathbf{x}_0$ doesn't belong to the support of ϕ_i and

$$\forall i \in I \setminus I_0, \phi_i(\mathbf{x}_0) = 0. \quad (39)$$

Let $g_h^0 \in W_h$ be a finite element approximation of the initial guess g^0 in the subsequent algorithm, such that $g_h^0(\mathbf{x}_0) = 0$ and $g_h^0 = \sum_{i=1}^n G_i^0 \phi_i$. For $g_h = \sum_{i \in I_0} G_i^0 \phi_i + \sum_{i \in I \setminus I_0} R_i \phi_i$, $R_i \in \mathbb{R}$, we get

$$\forall R_i \in \mathbb{R}, i \in I \setminus I_0, g_h(\mathbf{x}_0) = \sum_{i \in I_0} G_i^0 \phi_i(\mathbf{x}_0) = g_h^0(\mathbf{x}_0) = 0.$$

In other words, we can choose freely R_i , for all $i \in I \setminus I_0$.

For $i \in I \setminus I_0$, we set the elliptic problem: find $u_i \in V_h$ such that

$$\begin{aligned} & \int_D \nabla u_i \cdot \nabla v_h \, d\mathbf{x} + \int_D \beta'_{\eta, \varepsilon_2}(y_h - \varphi_h) u_i v_h \, d\mathbf{x} + \frac{1}{\varepsilon} \int_D H_\eta(g_h) u_i v_h \, d\mathbf{x} \\ &= -\frac{1}{\varepsilon} \int_D H'_\eta(g_h) y_h \phi_i v_h \, d\mathbf{x}, \quad \forall v_h \in V_h. \end{aligned} \quad (40)$$

For each $i \in I \setminus I_0$, the matrix of the linear system (40) is the same, just the right-hand side is different. We can solve effectively by factorization of the matrix, see [5], Chap. 4. We point out that $u_i \in V_h$, so the linear system has $\dim(V_h)$ equations.

Corollary 1. *Let g_h in W_h and we assume that $g_h(\mathbf{x}_0) = 0$. Then*

$$\begin{aligned} \partial_i \mathcal{J}_1(G) = \\ 2 \left(\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha \right) \left[\frac{\partial u_i}{\partial \mathbf{n}} - \frac{\partial y_h}{\partial \mathbf{n}} \frac{\partial \phi_i}{\partial \mathbf{n}} \frac{1}{|\nabla g_h|} + \frac{\nabla y_h \cdot \nabla \phi_i}{|\nabla g_h|} \right] (\mathbf{x}_0). \end{aligned} \quad (41)$$

Proof. For given g_h , we set y_h the solution of (35). The partial derivative $\partial_i \mathcal{J}_1(G)$ is in fact $J'_1(g_h)\phi_i$. The last one can be computed using (29), where y_ε , u , h are replaced with y_h , u_i , ϕ_i respectively. Here $\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0)$, $\frac{\partial u_i}{\partial \mathbf{n}}(\mathbf{x}_0)$, $\frac{\partial \phi_i}{\partial \mathbf{n}}(\mathbf{x}_0)$ mean

$$\nabla y_h(\mathbf{x}_0) \cdot \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|}, \quad \nabla u_i(\mathbf{x}_0) \cdot \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|}, \quad \nabla \phi_i(\mathbf{x}_0) \cdot \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|}$$

respectively. $\square$

The optimization problem to be solved is

$$\min_{R_i \in \mathbb{R}} \mathbb{J}_1((R_i)_{i \in I \setminus I_0}) = \mathcal{J}_1((G_i^0)_{i \in I_0}, (R_i)_{i \in I \setminus I_0}) \quad (42)$$

and we have for all $i \in I \setminus I_0$

$$\partial_i \mathbb{J}_1((R_i)_{i \in I \setminus I_0}) = \partial_i \mathcal{J}_1((G_i^0)_{i \in I_0}, (R_i)_{i \in I \setminus I_0}).$$

The steepest descent direction for (42) is given by

$$-\nabla \mathbb{J}_1((R_i)_{i \in I \setminus I_0}) \quad (43)$$

and we can write the **Algorithm**:

Step 1. Let $g_h^0 = \sum_{i=1}^n G_i^0 \phi_i$ be the initial guess, such that $g_h^0(\mathbf{x}_0) = 0$. Set $R_i^0 = G_i^0$ for $i \in I \setminus I_0$ and put $k = 0$.

Step 2. For $i \in I \setminus I_0$, compute u_i using (40), then compute $\nabla \mathbb{J}_1(R^k)$ using (41), where $R^k = (R_i^k)_{i \in I \setminus I_0}$.

Step 3. Find $\lambda_k > 0$ by line search

$$\lambda_k \in \arg \min_{\lambda > 0} \mathbb{J}_1(R^k - \lambda \nabla \mathbb{J}_1(R^k))$$

Step 4. Update

$$R^{k+1} = R^k - \lambda_k \nabla \mathbb{J}_1(R^k)$$

Step 5. We stop the algorithm when $|\mathbb{J}_1(R^{k+1}) - \mathbb{J}_1(R^k)| < tol$. If not, go to **Step 2**.

We notice that the values at the vertices of index $i \in I_0$ of the parameterization function $g_h^k = \sum_{i \in I_0} G_i^0 \phi_i + \sum_{i \in I \setminus I_0} R_i^k \phi_i$ are fixed, in order to satisfy the condition $g_h^k(\mathbf{x}_0) = 0$, for all $k \geq 0$. In fact, g_h^k is fixed in a neighborhood of $\mathbf{x}_0$.

Test 1.

The domain is $D =]0, 1[\times]0, 1[$, the obstacle is $\varphi : D \rightarrow \mathbb{R}$, $\varphi = -0.5$, $f : D \rightarrow \mathbb{R}$, $f = -100$ and $\alpha = 0$. We use a mesh for D of 26870 vertices, 53138 triangles and size $h = 1/150$. The following values have been used: penalization parameter $\varepsilon = 0.0001$, parameters for smoothed Heaviside and β are $\eta = 0.05$, $\varepsilon_2 = 0.01$. The starting domain is a disk of radius $r_0 = 0.25$ and center $(0.5, 0.5)$, while $\mathbf{x}_0 = (0.25, 0.5)$. For the stopping criterion, we use $tol = 10^{-6}$. For the cases a), b), c) we use the descent direction $-\nabla \mathbb{J}_1$ and for the case d) we use the simplified (partial) direction $-y_h p_h$.

Case a). We use $\alpha = 0$ and $C = 2$; I_0 has 13 vertices.

The steepest descent direction algorithm stops after 5 iterations and the objective functions decreases from 36.82 to 2.80×10^{-7} , see Figure 1. The evolution of Ω is presented in Figure 4. The boundary of the final Ω is plotted in Figure 2.

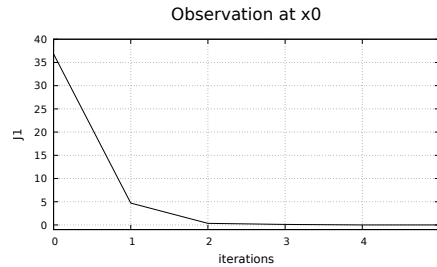


Figure 1: Test 1, case a). The history of the objective function J_1 .

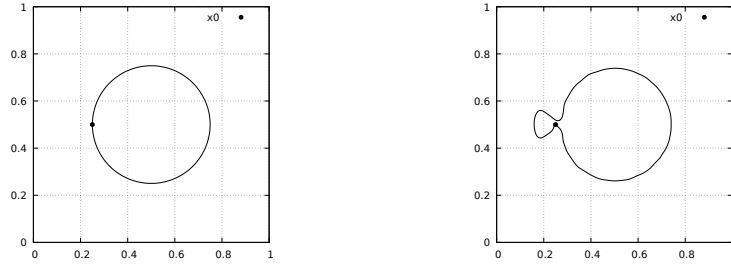


Figure 2: Test 1, case a). The position of $\mathbf{x}_0$ and the boundaries of Ω : initial (left) and final (right).

In Figure 3, we see that the obtained optimal geometry “generates” a saddle point for the optimal y_h and $\mathbf{x}_0$ is close to this saddle point. That’s why the final cost has such a low value. Notice however that the norm of $\nabla y_h(\mathbf{x}_0) \neq 0$, that is the tangential derivative of the optimal state in $\mathbf{x}_0$ is not zero and the Dirichlet boundary condition for y_h is slightly violated due to the numerical errors (but the numerical minimization is correctly performed).

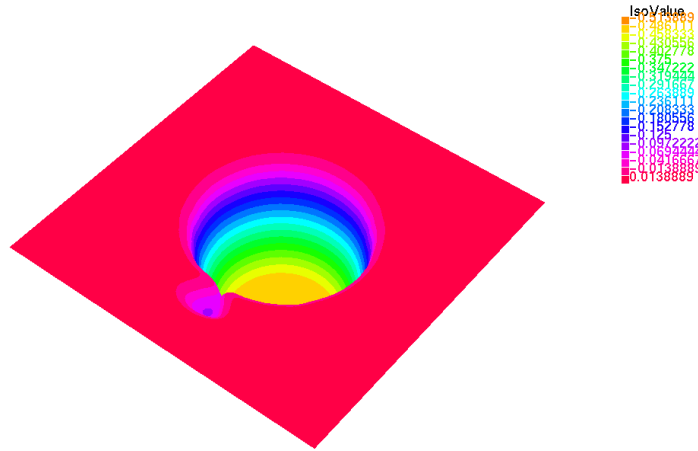


Figure 3: Test 1, case a). The final state solution y_h . For the computed solution, we have $\max(y_h) = 0$ and $\min(y_h) = -0.5$. The visualization software adds at the legend one level greater than max and one level smaller than min.

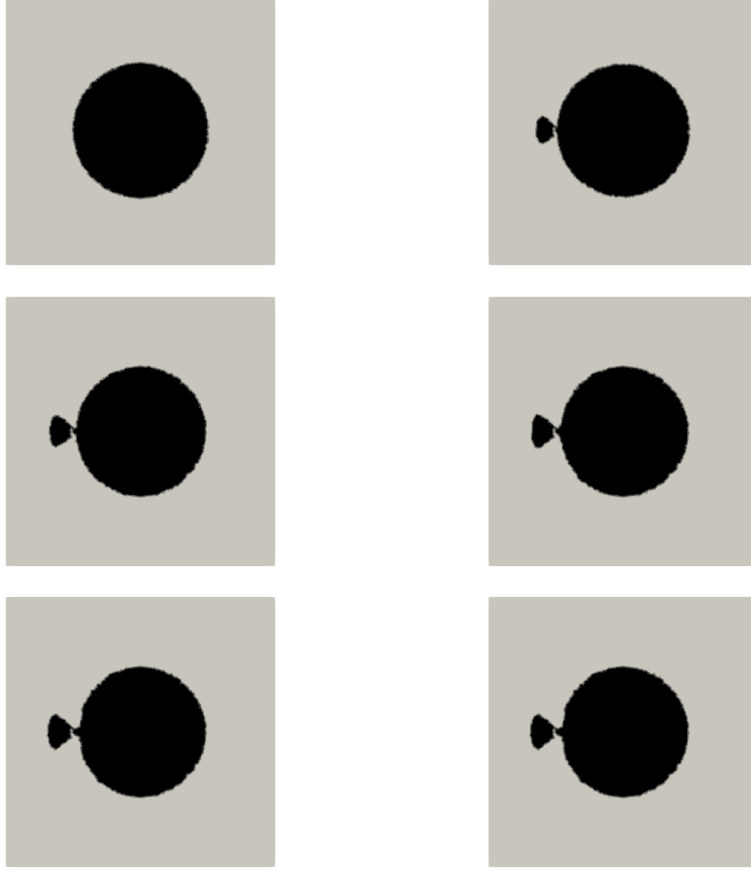


Figure 4: Test 1, case a). Ω at iterations: 0, 1, 2, 3, 4, 5 (from left to right and from top to the bottom).

In Figure 3, we show the final state solution y_h . At the final iteration, we obtained

$$\nabla y_h(\mathbf{x}_0) = (-0.001926, -0.367164), \quad \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|} = (-0.999978, 0.006689)$$

then $\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha = -0.000529 = \frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0)$.

Case b). We use $\alpha = 0$ and I_0 with 3 vertices.

If I_0 is the set of indices of the vertices of the triangle $T \in \mathcal{T}_h$ containing $\mathbf{x}_0$, the algorithm stops after 3 iterations, the values of the objective functions are: 36.82 (initial), 3.40515, 0.000325, 7.05×10^{-7} (final).

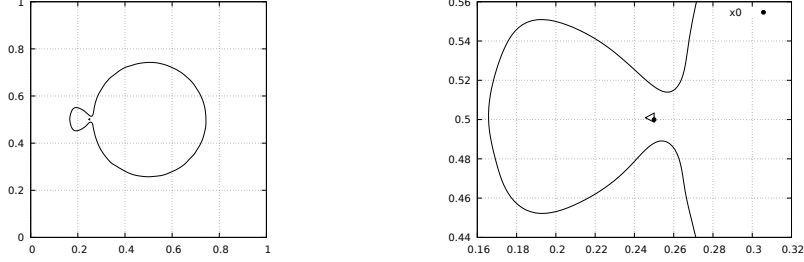


Figure 5: Test 1, case b). The final $\partial\Omega$ (left) and zoom around $\mathbf{x}_0$ (right).

We see in Figure 5 that the final $\partial\Omega$ has two connected boundaries (topological optimization), but one is very small. It is a triangular polygonal line containing $\mathbf{x}_0$. The triangle, having as boundary the polygonal line, represents a hole for the optimal Ω . We point out that this triangle is not an element of the triangulation, see Figure 6, otherwise we get $g_h = 0$ in this triangle, which contradicts (7).

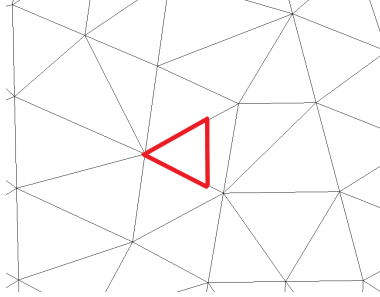


Figure 6: Test 1, case b). The smallest connected boundary (red) of the final $\partial\Omega$ and the mesh (zoom).

At the final iteration, we have

$$\nabla y_h(\mathbf{x}_0) = (-0.004556, -0.806676), \quad \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|} = (-0.999978, 0.006689)$$

then $\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha = -0.000839 = \frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0)$.

Case c). We use $\alpha = 1$ and $C = 2$; I_0 has 13 elements.

The algorithm stops after 3 iterations, the values of the objective functions are: 25.69 (initial), 1.45939, 0.000395, 1.26×10^{-7} (final).

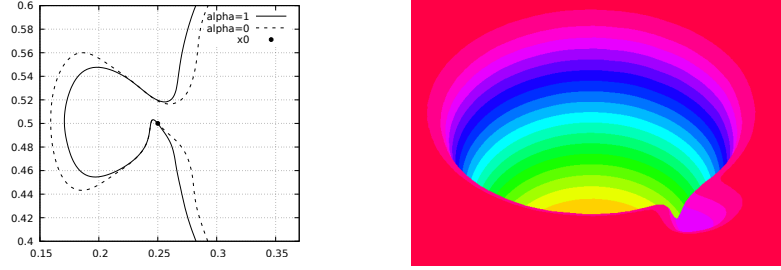


Figure 7: Test 1, case c). The final $\partial\Omega$ (left, zoom) and the final state y_h (right, zoom).

At the final iteration, we obtained

$$\nabla y_h(\mathbf{x}_0) = (-1.00285, -0.475403), \quad \frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|} = (-0.999978, 0.006689)$$

then $\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0) - \alpha = -0.000356$. Some images for the final solution are plotted in Figure 7. Although in Figure 7, we also see a saddle point for y_h , $\mathbf{x}_0$ is sufficiently far from it such that $\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}_0) - 1 = -0.000356$.

We see that $\frac{\nabla g_h(\mathbf{x}_0)}{|\nabla g_h(\mathbf{x}_0)|}$ is the same for the three cases a), b), c). It is correct, because g_h is not modified during the algorithm in the vertices $i \in I_0$. We have that $\mathbf{x}_0$ belongs to $\cup_{i \in I_0} \text{supp}(\phi_i)$. The set I_0 in the case b) is included in the set I_0 of cases a) and c). The triangulation, $\mathbf{x}_0$ and the initial g_h are the same for the three cases.

The holdall domain D is defined by the user. For rectangular shape, we can build a mesh composed only of right triangles or rectangular elements. Here, we use an unstructured triangulation (irregular grid) which can be employed also for non rectangular D .

Case d). We use $\alpha = 0$, $C = 2$ and $-y_h p_h$ as descent direction.

We can increase the regularity in the adjoint equation (32) by using $\zeta_{\varepsilon_1}(\mathbf{x} - \mathbf{x}_0)$, a mollifier approximation of the Dirac functional $\delta_{\mathbf{x}_0}$. We write $\zeta_{\varepsilon_1}(\mathbf{x}) = \frac{1}{\varepsilon_1^2} \zeta\left(\frac{\mathbf{x}}{\varepsilon_1}\right)$, where $\zeta \in \mathcal{C}^\infty(\mathbb{R}^2)$ has support in the unitary disk centered at $(0, 0)$ and $\int_{\mathbb{R}^2} \zeta(\mathbf{x}) d\mathbf{x} = 1$, $\varepsilon_1 > 0$.

An approximate version of the adjoint state equation (32) is: find $p \in H_0^1(D)$ such that

$$\begin{aligned} & \int_D \nabla p \cdot \nabla v d\mathbf{x} + \int_D \beta'_\varepsilon(y_\varepsilon) p v d\mathbf{x} + \frac{1}{\varepsilon} \int_D H_\varepsilon(g) p v d\mathbf{x} \\ &= - \int_D 2 \left(\frac{\partial y_\varepsilon}{\partial \mathbf{n}}(\mathbf{x}) - \alpha \right) \frac{\partial v}{\partial \mathbf{n}}(\mathbf{x}) \zeta_{\varepsilon_1}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x}, \quad \forall v \in H_0^1(D). \end{aligned} \quad (44)$$

Taking $v = u$ in (44) and using (20), we get that

$$\frac{1}{\varepsilon} \int_D H'_\varepsilon(g) h y_\varepsilon p \, d\mathbf{x} = \int_D 2 \left(\frac{\partial y_\varepsilon}{\partial \mathbf{n}}(\mathbf{x}) - \alpha \right) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) \zeta_{\varepsilon_1}(\mathbf{x} - \mathbf{x}_0) \, d\mathbf{x}$$

and the right-hand side is an approximation of L defined by (25). Since $H'_\varepsilon \geq 0$, the left-hand side is negative, for $h = -y_\varepsilon p$ and can be used in the gradient-type algorithms, inspired by [10], although it has a “partial” character.

All the data are the same as in the case a), but here we use the discretization (45) and in the Algorithm, the simplified descent direction $-y_h p_h$ instead of $-\nabla \mathbb{J}_1$. That is, we use just a part of the gradient computed in the previous section (but the descent properties are preserved according to Fig. 8). The objective functions decreases from 36.82 to 1.15×10^{-7} , see Figure 8. The Algorithm stops after 4 iterations.

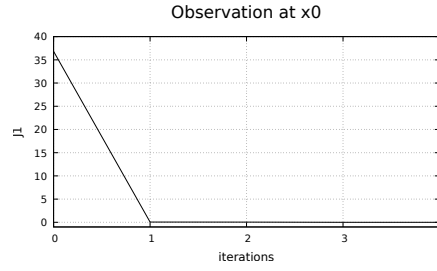


Figure 8: Test 1, case d). The history of the objective function.

The finite element approximation of the adjoint state is: find $p_h \in V_h$ such that

$$\begin{aligned} & \int_D \nabla p_h \cdot \nabla v_h \, d\mathbf{x} + \int_D \beta'_{\eta, \varepsilon_2}(y_h - \varphi_h) p_h v_h \, d\mathbf{x} + \frac{1}{\varepsilon} \int_D H_\eta(g_h) p_h v_h \, d\mathbf{x} \\ &= - \int_D 2 \left(\frac{\partial y_h}{\partial \mathbf{n}}(\mathbf{x}) - \alpha \right) \frac{\partial v_h}{\partial \mathbf{n}}(\mathbf{x}) \zeta_{\varepsilon_1}(\mathbf{x} - \mathbf{x}_0) \, d\mathbf{x}, \quad \forall v_h \in V_h. \end{aligned} \quad (45)$$

For the mollifier approximation of the Dirac function, we use $\varepsilon_1 = 0.05$.

We notice in Figure 9 that the final $\partial\Omega$ has three connected components of the boundary (topological optimization) and the smallest one is containing $\mathbf{x}_0$. That is, the final Ω has two holes.

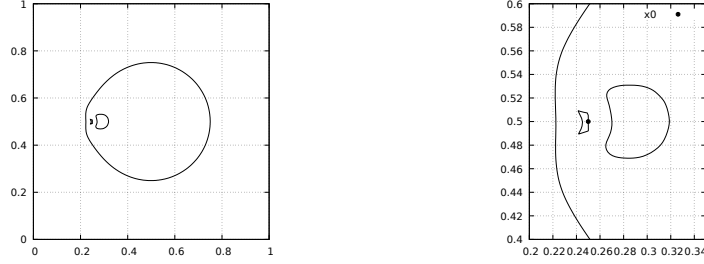


Figure 9: Test 1, case d). The final boundary of Ω (left), the position of $\mathbf{x}_0$ and the final boundary (right, zoom).

Test 2.

We can use similar computations, when we have pointwise observations at a finite number of fixed points $\mathbf{x}_0^j$, $1 \leq j \leq \ell$. We assume that $g(\mathbf{x}_0^j) = 0$, for all j . The objective function is the sum of the expression (18) computed at each point $\mathbf{x}_0^j$. The set I_0 is, in this case, the indices of the vertices A_i of $\mathcal{T}_h$ such that there exists $1 \leq j \leq \ell$, $\|\mathbf{x}_0^j - A_i\| < hC$ with $C \geq 2$. The descent direction is the sum of the terms in (43) computed at each point $\mathbf{x}_0^j$.

We have $\ell = 3$, the observation points are:

$$\begin{aligned} \mathbf{x}_0^1 &= (0.5 + r_0 \cos(\pi), 0.5 + r_0 \sin(\pi)), \\ \mathbf{x}_0^2 &= \left(0.5 + r_0 \cos\left(\pi - \frac{\pi}{6}\right), 0.5 + r_0 \sin\left(\pi - \frac{\pi}{6}\right)\right), \\ \mathbf{x}_0^3 &= \left(0.5 + r_0 \cos\left(\pi + \frac{\pi}{6}\right), 0.5 + r_0 \sin\left(\pi + \frac{\pi}{6}\right)\right). \end{aligned}$$

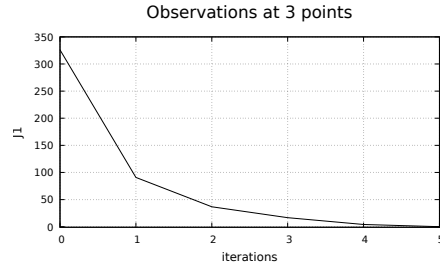


Figure 10: Test 2. The history of the objective function.

For $C = 3$ we get I_0 of 96 elements. The others parameters are the same as in Test 1 and $\alpha = 0$. The algorithm stops after 5 iterations, the objective functions decreases from 326.12 to 1.64×10^{-5} , see Figure 10.

In Figure 11 we plot the initial and final $\partial\Omega$ and in Figure 12 we show the final state solution y_h . Final Ω seems to be not connected, while the support of y_h is connected. This phenomena appears since in the state variational inequality we use in the penalization term $H_\eta(g_h)$ which is an approximation of the characteristic function of the final Ω , that is the approximations yield important perturbations on the result without affecting its essence.

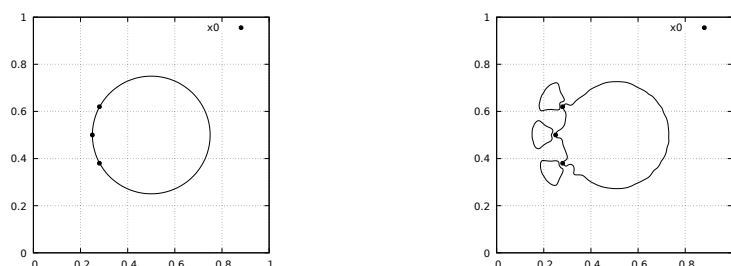


Figure 11: Test 2. The boundaries of Ω : initial (left) and final (right).

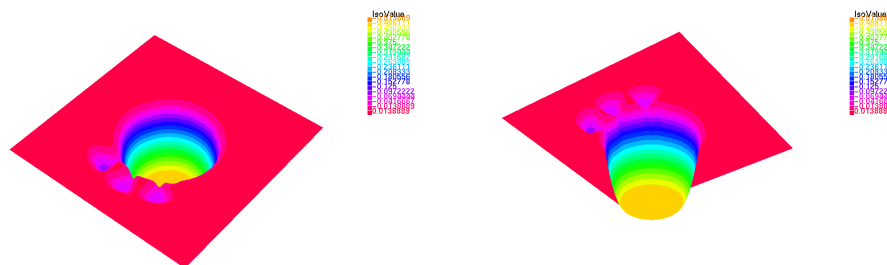


Figure 12: Test 2. The final state solution y_h , top view (left) and bottom view (right).

Acknowledgements. Paper written with financial support of ECO-MATH France-Roumanie.

References

- [1] R. Adams, *Sobolev Spaces*, Academic Press, 1975.
- [2] V. Barbu, *Optimal Control of Variational Inequalities*, Pitman Advanced Publishing Program, London, 1984.

- [3] D.P. Bertsekas, *Nonlinear Programming*, 3rd edition, Athena Scientific, Nashua, MASS, 2016.
- [4] H. Brezis, Problèmes unilatéraux, *J. Math. Pures Appl.* 51 (1972), 1–168.
- [5] Ph. G. Ciarlet, *Introduction to Numerical Linear Algebra and Optimization*, Cambridge University Press, New York, 1989.
- [6] A. Halanay, C.M. Murea and D. Tiba, Some properties of the period for certain ordinary differential systems and applications to topology optimization of variational inequalities. In: Migorski, S., Sofonea, M. (eds.) Nonsmooth problems with applications in mechanics, *Banach Center Publications*, 127 (2024), 129–146.
- [7] H. Kowarada and M. Natori, An application of the integrated penalty method to free boundary problems of Laplace equation, *Numer. Funct. Anal. Optim.* 3 (1981), 1–17.
- [8] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.
- [9] C.M. Murea and D. Tiba, Topological optimization via cost penalization, *Topol. Methods Nonlinear Anal.* 54 (2019), 1023–1050.
- [10] C.M. Murea and D. Tiba, Periodic Hamiltonian systems in shape optimization problems with Neumann boundary conditions, *J. Differ. Equations* 321 (2022), 1–39.
- [11] C.M. Murea and D. Tiba, Topological optimization with nonlinear state equation, *SeMA* 82 (2025), 219–240.
- [12] P. Neittaanmäki, J. Sprekels and D. Tiba, *Optimization of Elliptic Systems. Theory and Applications*, Springer, New York, 2006.
- [13] P. Neittaanmäki, A. Pennanen and D. Tiba, Fixed domain approaches in shape optimization problems with Dirichlet boundary conditions, *Inverse Probl.* 25 (2009), 1–18.
- [14] D. Tiba, *Optimal Control of Nonsmooth Distributed Parameter Systems*, Springer, Berlin, 1990.
- [15] D. Tiba, Domains of class C: properties and applications, *Ann. Univ. Buchar. Math. Ser.* 4(LXII) (2013), 89–102.

- [16] D. Tiba, The implicit function theorem and implicit parametrizations, *Ann. Acad. Rom. Sci. Ser. Math. Appl.* 5 (2013), 193–208.
- [17] D. Tiba, Iterated Hamiltonian type systems and applications, *J. Differ. Equations* 264 (2018), 5465–5479.
- [18] D. Tiba, Optimality conditions and Lagrange multipliers for shape and topology optimization problems, *Atti della Accademia Peloritana dei Pericolanti Classe di Scienze Fisiche, Matematiche e Naturali* 101 (2023), A9.