A semismooth Newton method for topology optimization

Samuel Amstutz

A semismooth Newton method for topology optimization

Samuel Amstutz

Laboratoire d'analyse non linéaire et géométrie, Faculté des Sciences, 33 rue Louis Pasteur, 84000 Avignon, France.

Abstract

This paper deals with elliptic optimal control problems for which the control function is constrained to assume values in $\{0,1\}$. Based on an appropriate formulation of the optimality system, a semismooth Newton method is proposed for the solution. Convergence results are proved, and some numerical tests illustrate the efficiency of the method.

Keywords: topology optimization, semismooth Newton method, topological sensitivity, topological derivative
2000 MSC: 49K20, 49M05, 49M15, 49Q10, 49Q12

1. Introduction

Let D be an open and bounded subset of $\mathbb{R}^2$ with Lipschitz boundary. We investigate the numerical solution of the optimal control problem:

$$\underset{(u,y)\in\mathcal{E}\times H_0^1(D)}{\text{Minimize}} \; J(u,y) = \frac{1}{2}\int_D (y-y^\dagger)^2 dx + \nu \int_D u dx \tag{1}$$

subject to the state equation

$$Ey = u. \tag{2}$$

Above the feasible set is defined by

$$\mathcal{E} = \{u_\Omega = \chi_\Omega \text{ a.e. in } D,\ \Omega \subset D\},$$

Email address: samuel.amstutz@univ-avignon.fr (Samuel Amstutz)

where, for every subset Ω of D, χ_Ω is the characteristic function of Ω in D. The data $(y^\dagger, \nu)$ are given in $L^2(D) \times \mathbb{R}$. The operator E stands for the (negative) Dirichlet Laplacian, i.e.,

$$\begin{array}{rcl} E : H_0^1(D) & \to & H^{-1}(D) \\ y & \mapsto & [z \mapsto \int_D \nabla y . \nabla z]. \end{array}$$

We denote by $y_u := E^{-1}u$ the state and by $j(u) := J(u, y_u)$ the reduced cost. Note that, by a straightforward change of variables, this setting also allows to treat controls of the form $u_\Omega = \chi_\Omega u^- + (1 - \chi_\Omega)u^+$, where (u^-, u^+) are given constants.

Due to the binary nature of the control u, the problem in hand falls into the framework of shape optimization, but as the topology (i.e. the number of connected components of the boundary) of the targeted domain Ω^* is a priori unknown, we rather speak of topology optimization. Let us briefly recall the main classes of methods which can be applied in this case. The most commonly used approaches rely on relaxed formulations [1, 8]. They are in principle compatible with all the tools of nonlinear programming, which is a considerable practical advantage, however eventually retrieving a feasible control is not always easy. The so-called classical shape optimization methods [18, 21], based on smooth boundary perturbations, offer a good accuracy as to the parameterization of the domains, but do not allow topological changes. More flexibility is obtained when they are associated with a level-set domain representation [3, 22]. Nevertheless the nucleation of holes still does not occur in this framework. The concept of topological derivative [14, 16, 20] has been precisely introduced to give a quantitative information on the relevance of creating a hole around an arbitrary point of the domain. Therefore, modern topology optimization techniques often couple the topological derivative with the shape derivative and / or the use of level sets [2, 11, 6]. A major drawback of these methods is that they only rely on first order information. To justify this claim, we recall that the topological derivative $g_{u_\Omega}(x)$ for a circular hole is defined by

$$j(u_{\Omega \setminus \bar{B}(x,\rho)}) - j(u_\Omega) = f(\rho) g_{u_\Omega}(x) + o(f(\rho)), \tag{3}$$

where $\bar{B}(x, \rho) = \{x, \|x\| \leq \rho\}$ and f is a nonnegative function that tends to zero when $\rho \to 0$. Higher order topological asymptotic expansions can be derived [9, 10, 13], however the higher order terms are numerically quite difficult to be computed. In addition, as the (first order) topological derivative

is generally not bound to vanish at the optimum, the way to use the higher order information is not obvious. If it vanishes, or gets close to zero as it is the case in least square problems, then the first order methods often prove inefficient, as pointed out in [15]. In [15] a two steps approach is proposed. In the first step the cost function is approximated (or rather weighted) by a Gauss-Newton type method, and in the second step the first order topological derivative of the weighted cost function is computed.

This paper is based on the remark that the topological optimality conditions can be reformulated in the form $F(u) = 0$ with the help of an appropriate function F. Then Newton type methods can be used to solve the equation in an efficient way. These arguments are subsequently developed according to the following outline. The optimality conditions associated with Problem (1)-(2) are specified in Section 2 and the function F is introduced. After some background material is recalled in Section 3, the proposed method is described and analyzed in Sections 4 through 6. Finally, some numerical results are presented in Section 7.

2. Optimality conditions

To begin with, let us specify some definitions and notations.

Definition 1. *We say that a function $u \in \mathcal{E}$ is a local minimizer of* (1)-(2) *if*

$$\exists \eta > 0 \ \forall u' \in \mathcal{E} \ \|u' - u\|_{L^1(D)} \leq \eta \Rightarrow j(u') \geq j(u).$$

We say that it is a global minimizer if $j(u') \geq j(u)$ for all $u' \in \mathcal{E}$.

Given a function $f : D \to \mathbb{R}$ and a number $a \in \mathbb{R}$, we define the level set $[f = a] := \{x \in D, f(x) = a\}$. If A is a subset of D we denote by $\mathrm{int}(A)$ its interior. We recall that the operator $E : H_0^1(D) \to H^{-1}(D)$ is continuous and invertible, and that $H^1(D)$ is continuously imbedded in $L^q(D)$ for all $q \in (1, +\infty)$. Consequently, Problem (2) admits a unique solution $y_u \in H_0^1(D)$ for any $u \in L^p(D)$, $p > 1$. For such p we also define the adjoint state z_u as the solution of

$$Ez_u = -(y_u - y^\dagger) \tag{4}$$

and we set

$$g_u = -z_u + \nu. \tag{5}$$

We can show [4] that g_u is the Fréchet derivative of the mapping $u \in L^p(D) \mapsto j(u)$, and that it coïncides with the topological derivative defined by (3) for

$f(\rho) = \mathcal{L}^2(\bar{B}(x,\rho)) = \pi\rho^2$. Here and in the sequel $\mathcal{L}^2$ designates the Lebesgue measure on $\mathbb{R}^2$. The following result is proved in [4].

Proposition 2. *A necessary condition for $u \in \mathcal{E}$ to be a local minimizer of* (1)-(2) *is*

$$\left\{ \begin{array}{l} g_u \geq 0 \ \textit{in} \ \mathrm{int}([u=0]), \\ g_u \leq 0 \ \textit{in} \ \mathrm{int}([u=1]). \end{array} \right. \tag{6}$$

A sufficient condition for $u \in \mathcal{E}$ to be a global minimizer of (1)-(2) *is*

$$\left\{ \begin{array}{l} g_u \geq 0 \ \textit{a.e. in} \ [u=0], \\ g_u \leq 0 \ \textit{a.e. in} \ [u=1]. \end{array} \right. \tag{7}$$

The proposed algorithm relies on an appropriate reformulation of the conditions (6) and (7). To this aim we introduce the functional

$$\begin{array}{rcl} F : L^2(D) & \rightarrow & L^2(D) \\ u & \mapsto & u|g_u| + \min(0, g_u), \end{array} \tag{8}$$

or equivalently

$$F : u \mapsto u \max(0, g_u) + (1-u)\min(0, g_u).$$

In view of the remark 4 below, F is well defined.

Theorem 3. *A necessary condition for $u \in \mathcal{E}$ to be a local minimizer of* (1)-(2) *is*

$$F(u) = 0 \ \textit{in} \ \mathrm{int}([u=0]) \cup \mathrm{int}([u=1]). \tag{9}$$

A sufficient condition for $u \in L^2(D)$ to be a global minimizer of (1)-(2) *is*

$$\left\{ \begin{array}{l} F(u) = 0 \ \textit{a.e. in } D, \\ \mathcal{L}^2([g_u = 0]) = 0. \end{array} \right. \tag{10}$$

Proof. Assume that (6) holds and take $x \in \mathrm{int}([u=0])$. Then $g_u(x) \geq 0$, which entails $F(u)(x) = 0$. Similarly, if $x \in \mathrm{int}([u=1])$, then $g_u(x) \leq 0$, thus $F(u)(x) = 0$. Assume now that $F(u) = 0$ a.e. in D and take $x \in D$ such that $F(u)(x) = 0$. If $g_u(x) > 0$, then $u(x)g_u(x) = 0$, thus $u(x) = 0$. If $g_u(x) < 0$, then $(1-u(x))g_u(x) = 0$, thus $u(x) = 1$. On the one hand we deduce that $u \in \mathcal{E}$ provided that $\mathcal{L}^2([g_u = 0]) = 0$. On the other hand we derive (7) by contradiction. □

The remainder of the paper is devoted to the solution of the equation $F(u) = 0$ a.e. in D (subsequently simply denoted by $F(u) = 0$), which, in non-degenerate cases, is a necessary and sufficient optimality condition for any $u \in L^2(D)$. As F is non-differentiable, a semismooth Newton method will be invoked.

Remark 4. *We recall that, since D is Lipschitz, E^{-1} maps $L^2(D)$ into $H^{3/2}(D)$ (see e.g. [19]). By using for instance the closed graph theorem, it easy to show that this map is continuous for these norms. In addition, $H^{3/2}(D)$ is compactly imbedded in $L^\infty(D)$. It follows that the map $u \in L^2(D) \mapsto (y_u, z_u) \in (L^\infty(D))^2$ is linear and compact.*

3. Preliminary material

In this section a few results related to semismooth Newton methods and collectively compact sets of operators are recalled.

3.1. Semismooth Newton methods

Among the many generalizations of the notion of differentiability and the related extensions of Newton's method, we adopt here the presentation of [17]. The following definition and theorems are extracted from [17], with a few adaptations of the notations. Let X, Y be Banach spaces and U be an open subset of X.

Definition 5. *A function $F : U \to Y$ is called slantly differentiable if there exists a mapping $G : U \to \mathcal{L}(X, Y)$, referred to as slanting function, such that*

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

for all $u \in U$.

Of course, functions which are $\mathcal{C}^1$ in the sense of Fréchet are slantly differentiable. The following theorem provides another particularly useful example.

Theorem 6. *The mapping* $\max(0, .) : L^q(D) \to L^p(D)$ *with* $1 \leq p < q \leq +\infty$ *is slantly differentiable on $L^q(D)$ and $u \mapsto \chi_{[u>0]}$ is a slanting function.*

The following theorem asserts the local convergence of the semismooth Newton method applied to a slantly differentiable function.

Theorem 7. *Suppose that u^* solves $F(u^*) = 0$ and that F is slantly differentiable in an open set U containing u^*, with slanting function G. If $G(u)$ is nonsingular for all $u \in U$ and $\{\|G(u)^{-1}\|, u \in U\}$ is bounded, then the Newton iteration*

$$u_{n+1} = u_n - G(u_n)^{-1}F(u_n)$$

converges superlinearly to u^*, provided that $\|u_0 - u^*\|$ is sufficiently small.

3.2. Collectively compact sets of operators

Let X be a Banach space and $\mathcal{K}$ be a subset of $\mathcal{L}(X)$, where $\mathcal{L}(X)$ is the set of bounded linear operators from X into itself. We say that $\mathcal{K}$ is collectively compact if the set $\{Kx, x \in X, \|x\| \leq 1, K \in \mathcal{K}\}$ is relatively compact. Obviously, if $\mathcal{K}$ is collectively compact, every $K \in \mathcal{K}$ is compact. We recall a part of Theorem 1.6 of [7].

Theorem 8. *Let $K, (K_n)_{n\in\mathbb{N}} \in \mathcal{L}(X)$. Assume that $K_n \to K$ pointwise, $\{K_n\}$ is collectively compact, and K is compact. Then $(I - K)^{-1}$ exists if and only if for some N and all $n \geq N$ the operators $(I - K_n)^{-1}$ exist and are bounded uniformly.*

The following result can be seen as a particular case of Theorem 4.5 of [12]. For completeness we provide a very simple proof using Theorem 8.

Theorem 9. *Let $\mathcal{K}$ be a collectively compact set of bounded linear operators of X. Assume further that $\mathcal{K}$ is pointwise sequentially compact, i.e., for every sequence (K_n) of $\mathcal{K}$ there exists a subsequence (K_{n_p}) and $K \in \mathcal{K}$ such that $K_{n_p}x \to Kx$ for all $x \in X$. If $I - K$ is invertible for all $K \in \mathcal{K}$, then*

$$\sup_{K\in\mathcal{K}} \|(I - K)^{-1}\| < \infty. \tag{11}$$

Proof. Assume that (11) is not fulfilled. Then we can construct a sequence (K_n) of $\mathcal{K}$ such that

$$\|(I - K_n)^{-1}\| \geq \frac{1}{n} \qquad \forall n \in \mathbb{N}^*. \tag{12}$$

By compactness of $\mathcal{K}$, there exists a subsequence (K_{n_p}) such that $K_{n_p} \to K \in \mathcal{K}$ pointwise. Since $I - K$ is invertible, Theorem 8 implies that $\{(I - K_{n_p})^{-1}\}$ is bounded uniformly, which is in contradiction with (12). □

4. Description and analysis of the method in function spaces

4.1. Direct application of the semismooth Newton method

A natural idea is to apply the semismooth Newton method described in Theorem 7 to the solution of $F(u) = 0$. We shall show that this approach is likely to fail in general.

The direct and adjoint equations (2) and (4) together with the definition of the sensitivity (5) provide

$$g_u = E^{-2}u - E^{-1}y^\dagger + \nu. \tag{13}$$

A straightforward calculation shows that F is locally Lipschitz continuous in $L^2(D)$. But, in order to obtain the slant differentiability of $F : L^2(D) \to L^2(D)$, we would need the slant differentiability of the function $g \in L^\infty(D) \mapsto \max(0, g) \in L^\infty(D)$, which is not fulfilled. However, Theorem 3 could also well be written in arbitrary L^p spaces, thus let us see what happens if we consider $F : L^p(D) \to L^q(D)$ for $1 \leq q < p \leq \infty$. We define

$$G(u) : h \mapsto |g_u|h + \left[u\chi_{[g_u>0]} + (1-u)\chi_{[g_u<0]}\right] E^{-2}h,$$

and we obtain

$$\begin{aligned} F(u+h) - F(u) - G(u+h)h \\ &= u\left[|g_u + E^{-2}h| - |g_u| - (\chi_{[g_u+E^{-2}h>0]} - \chi_{[g_u+E^{-2}h<0]})E^{-2}h\right] \\ &\qquad + h\left[\chi_{[g_u+E^{-2}h>0]} - \chi_{[g_u+E^{-2}h<0]}\right] E^{-2}h \\ &\qquad + \left[\min(0, g_u + E^{-2}h) - \min(0, g_u) - \chi_{[g_u+E^{-2}h<0]}E^{-2}h\right]. \end{aligned}$$

Using successively the Hölder inequality, Theorem 6 and the continuity of $E^{-2} : L^p(D) \to L^\infty(D)$ it comes

$$\begin{aligned} &\|F(u+h) - F(u) - G(u+h)h\|_{L^q(D)} \\ &\leq \|u\|_{L^p(D)} o(\|E^{-2}h\|_{L^\infty(D)}) + 2\|h\|_{L^1(D)}\|E^{-2}h\|_{L^\infty(D)} + o(\|E^{-2}h\|_{L^\infty(D)}) \\ &\leq o(\|h\|_{L^p(D)}). \end{aligned}$$

Therefore G is a slanting function for F. Assume now that $u^* \in L^p(D)$ solves $F(u^*) = 0$. As seen in the proof of Theorem 3, we have $g_{u^*}(x) > 0 \Rightarrow u^*(x) = 0$ and $g_{u^*}(x) < 0 \Rightarrow u^*(x) = 1$, which yields

$$G(u^*)h = |g_{u^*}|h.$$

It appears that $G(u^*)$ is singular, unless we are in finite dimension (we can choose $p = q$) and g_{u^*} does not vanish on D. Consequently the semismooth Newton method applied to the solution of $F(u^*) = 0$ may hardly converge. Numerically, this is rendered by oscillations which occur in the regions where u switches between 0 and 1.

4.2. Regularization

To circumvent the phenomenon described above we introduce a family of regularized functionals. For $\varepsilon > 0$ we define

$$\begin{array}{rcl} F^\varepsilon : L^2(D) & \to & L^2(D) \\ u & \mapsto & u(\Phi_\varepsilon \circ g_u) + \min(0, g_u), \end{array} \tag{14}$$

with

$$\Phi_\varepsilon(t) = \sqrt{\varepsilon^2 + t^2}. \tag{15}$$

Obviously, Φ_ε maps $L^\infty(D)$ into itself, hence F^ε is well defined in $L^2(D)$. It is also clearly locally Lipschitz continuous. By observing that the first term in the definition of F^ε is $\mathcal{C}^1$ and applying Theorem 6 to the second term, we derive that F^ε is slantly differentiable with slanting function

$$G^\varepsilon(u)h = (\Phi_\varepsilon \circ g_u)h + u(\Phi'_\varepsilon \circ g_u)E^{-2}h + \chi_{[g_u<0]}E^{-2}h,$$

that is,

$$G^\varepsilon(u) = (\Phi_\varepsilon \circ g_u)T^\varepsilon(u), \tag{16}$$

with

$$T^\varepsilon(u) = I + \frac{u(\Phi'_\varepsilon \circ g_u) + \chi_{[g_u<0]}}{\Phi_\varepsilon \circ g_u}E^{-2}. \tag{17}$$

Let us now address the convergence of the semismooth Newton method applied to F^ε. We denote by $\langle ., . \rangle$ the canonical inner product of $L^2(D)$ and by $\|.\|$ the corresponding norm. Given two functions $u, v \in L^2(D)$, we use the notation $u \leq v$ in the pointwise almost everywhere sense. We write $u < v$ if $u \leq v$ and $u \neq v$ (i.e. u and v assume different values on a set on nonzero measure).

Lemma 10. *For all $\varepsilon > 0$, the equation $F^\varepsilon(u^\varepsilon) = 0$ admits at least one solution $u^\varepsilon \in L^2(D)$. In addition, every function $u^\varepsilon \in L^2(D)$ solution of $F^\varepsilon(u^\varepsilon) = 0$ satisfies*

$$0 \leq u^\varepsilon \leq 1. \tag{18}$$

Proof. By definition of F^ε it holds

$$F^\varepsilon(u^\varepsilon) = 0 \;\Leftrightarrow\; u^\varepsilon = -\frac{\min(0, g_{u^\varepsilon})}{\Phi_\varepsilon \circ g_{u^\varepsilon}} =: \Theta^\varepsilon(u^\varepsilon). \tag{19}$$

Equivalently we may write

$$\Theta^\varepsilon(u^\varepsilon) = \begin{cases} 0 & \text{if} \quad g_{u^\varepsilon} \geq 0, \\ \dfrac{|g_{u^\varepsilon}|}{\sqrt{\varepsilon^2 + g_{u^\varepsilon}^2}} & \text{if} \quad g_{u^\varepsilon} < 0, \end{cases}$$

hence $0 \leq \Theta^\varepsilon(u^\varepsilon) \leq 1$. We define the convex set $C = L^2(D, [0,1])$. Clearly, Θ^ε is a continuous mapping from C into itself. Moreover, for all $u \in L^2(D)$, it holds $g_u \in H^1(D) \cap L^\infty(D)$ which implies

$$\nabla\Theta^\varepsilon(u) = -\frac{(\Phi_\varepsilon \circ g_u)\chi_{[g_u<0]}\nabla g_u - \min(0, g_u)(\Phi'_\varepsilon \circ g_u)\nabla g_u}{(\Phi_\varepsilon \circ g_u)^2} \in L^2(D)$$

and subsequently $\Theta^\varepsilon(u) \in H^1(D)$. Therefore $\Theta^\varepsilon(C) \subset \tilde{C} := H^1(D, [0,1])$. Yet, $\tilde{C}$ is a relatively compact subset of $L^2(D)$. By Schauder's fixed point theorem, there exists some $u^\varepsilon \in C$ such that $\Theta^\varepsilon(u^\varepsilon) = u^\varepsilon$.

□

Theorem 11. *Suppose that $u^\varepsilon \in L^2(D)$ solves $F^\varepsilon(u^\varepsilon) = 0$. Then the Newton iteration*

$$u_{n+1} = u_n - G^\varepsilon(u_n)^{-1} F^\varepsilon(u_n)$$

converges superlinearly to u^ε in $L^2(D)$, provided that $\|u_0 - u^\varepsilon\|$ is sufficiently small.

Proof. The assumptions we need to check in order to apply Theorem 7 are the invertibility of the slanting function in a neighborhood of u^ε together with a uniform bound on the norm of the inverse. For all $u \in L^2(D)$ we set

$$w_u = \frac{u(\Phi'_\varepsilon \circ g_u) + \chi_{[g_u<0]}}{\Phi_\varepsilon \circ g_u}, \tag{20}$$

and for all $w \in L^2(D)$ we define the operator

$$K(w) : h \in L^2(D) \mapsto wE^{-2}h \in L^2(D).$$

In this way we have

$$T^\varepsilon(u) = I + K(w_u) \qquad \forall u \in L^2(D).$$

We proceed in four steps.

1. By definition it holds for all $w, h \in L^2(D)$

$$\langle (I + K(w))h, E^{-2}h\rangle = \langle h, E^{-2}h\rangle + \langle wE^{-2}h, E^{-2}h\rangle.$$

By the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\langle wE^{-2}h, E^{-2}h\rangle &\geq \langle \min(0,w)E^{-2}h, E^{-2}h\rangle \\ &\geq -\|\min(0,w)\|_{L^2(D)}\|E^{-2}h\|^2_{L^4(D)}.\end{aligned}$$

By elliptic regularity (see Remark 4), there exists a constant $c > 0$, depending only on D, such that $\|E^{-2}h\|_{L^4(D)} \leq c\|E^{-1}h\|_{L^2(D)}$. It comes

$$\langle wE^{-2}h, E^{-2}h\rangle \geq -c^2\|\min(0,w)\|_{L^2(D)}\|E^{-1}h\|^2_{L^2(D)},$$

and subsequently by symmetry of E^{-1}

$$\langle (I + K(w))h, E^{-2}h\rangle \geq (1 - c^2\|\min(0,w)\|)\|E^{-1}h\|^2_{L^2(D)}. \tag{21}$$

We set

$$W^- = \left\{ w \in L^2(D), \|\min(0,w)\| \leq \frac{1}{2c^2} \right\}. \tag{22}$$

For all $w \in W^-$, we have $1 - c^2\|\min(0,w)\| \geq 1/2$, thus in view of (21) the operator $I + K(w)$ is injective. In addition, $E^{-2} : L^2(D) \to L^\infty(D)$ is compact, thus $K(w)$ is compact for all $w \in L^2(D)$. Then the Fredholm alternative implies that

$$I + K(w) \text{ is invertible for all } w \in W^-. \tag{23}$$

2. We now examine under which conditions $w_u \in W^-$. Let $u \in L^2(D)$ be arbitrary. Using (18) it comes

$$-|u - u_\varepsilon| \leq u \leq 1 + |u - u^\varepsilon|. \tag{24}$$

We write

$$\begin{aligned}u(\Phi'_\varepsilon \circ g_u) + \chi_{[g_u<0]} &= (u(\Phi'_\varepsilon \circ g_u) + 1)\chi_{[g_u<0]} + u(\Phi'_\varepsilon \circ g_u)\chi_{[g_u\geq 0]} \\ &= \left(1 - u\frac{|g_u|}{\sqrt{\varepsilon^2 + g_u^2}}\right)\chi_{[g_u<0]} + u\frac{|g_u|}{\sqrt{\varepsilon^2 + g_u^2}}\chi_{[g_u\geq 0]}.\end{aligned}$$

Then (24) entails

$$\begin{aligned}
&u(\Phi'_\varepsilon \circ g_u) + \chi_{[g_u<0]} \\
&\geq \left(1 - (1+|u-u^\varepsilon|)\frac{|g_u|}{\sqrt{\varepsilon^2+g_u^2}}\right)\chi_{[g_u<0]} - |u-u^\varepsilon|\frac{|g_u|}{\sqrt{\varepsilon^2+g_u^2}}\chi_{[g_u\geq 0]} \\
&\geq -|u-u^\varepsilon|\chi_{[g_u<0]} - |u-u^\varepsilon|\chi_{[g_u\geq 0]} = -|u-u^\varepsilon|.
\end{aligned}$$

We therefore derive from (20) that

$$w_u \geq -\frac{|u-u^\varepsilon|}{\varepsilon},$$

which implies

$$\|\min(0, w_u)\| \leq \frac{1}{\varepsilon}\|u-u^\varepsilon\|. \tag{25}$$

We denote by U the ball

$$U = \left\{u \in L^2(D), \|u-u^\varepsilon\| < \frac{\varepsilon}{2c^2}\right\}. \tag{26}$$

It stems from (25) and (22) that

$$u \in U \Rightarrow w_u \in W^-. \tag{27}$$

Remembering (23), we conclude that $T^\varepsilon(u)$ is invertible for all $u \in U$.

3. Let $K(w)^*$ be the adjoint operator of $K(w)$, namely

$$K(w)^* : h \in L^2(D) \mapsto E^{-2}(wh) \in L^2(D).$$

Note that the operator E^{-2} involved above is in fact the adjoint of $E^{-2} : L^2(D) \to L^\infty(D)$ which, in particular, defines a compact operator from $L^1(D)$ into $L^2(D)$. The same notation has been kept since it is an extension of the operator E^{-2} defined on $L^2(D)$. As U is bounded, there exists $M > 0$ such that

$$\|w_u\| \leq M \qquad \forall u \in U. \tag{28}$$

We define the sets

$$W^+ = \{w \in L^2(D), \|w\| \leq M\}, \tag{29}$$

$$W = W^- \cap W^+, \tag{30}$$
$$\mathcal{K} = \{K(w)^*, w \in W\}.$$

We shall check the assumptions of Theorem 9 for $\mathcal{K}$. Thanks to the compactness of the operator $E^{-2} : L^1(D) \to L^2(D)$, we immediately see that $\mathcal{K}$ is collectively compact. Let now (w_n) be a sequence of W. Since W is bounded, convex and closed in $L^2(D)$, there exists a subsequence still denoted by (w_n) such that $w_n \rightharpoonup w \in W$ weakly in $L^2(D)$. It follows that $E^{-2}(w_n h) \rightharpoonup E^{-2}(wh)$ weakly in $L^2(D)$ for all $h \in L^2(D)$. By compactness of $E^{-2} : L^1(D) \to L^2(D)$, the convergence holds actually in norm (for a subsequence), hence $K(w_n)^* h \to K(w)^* h$. Therefore Theorem 9 provides $\sup_{K\in\mathcal{K}} \|(I + K)^{-1}\| < \infty$, that is, $\sup_{w\in W} \|(I + K(w)^*)^{-1}\| < \infty$. Passing to the adjoint yields $\sup_{w\in W} \|(I + K(w))^{-1}\| < \infty$. Yet, due to (27), (28), (29) and (30) it holds $u \in U \Rightarrow w_u \in W$. We arrive at

$$\sup_{u\in U} \|T^\varepsilon(u)^{-1}\| = \sup_{u\in U} \|(I + K(w_u))^{-1}\| < \infty. \tag{31}$$

4. Let us now come back to the operator $G^\varepsilon(u)$. Using (16) and the invertibility of $T^\varepsilon(u)$ we straightforwardly derive that $G^\varepsilon(u)$ is invertible for all $u \in U$ with

$$G^\varepsilon(u)^{-1} : h \mapsto T^\varepsilon(u)^{-1}\left(\frac{h}{\Phi_\varepsilon \circ g_u}\right).$$

Hence

$$\|G^\varepsilon(u)^{-1}\| \leq \frac{1}{\varepsilon}\|T^\varepsilon(u)^{-1}\|, \tag{32}$$

which thanks to (31) yields

$$\sup_{u\in U} \|G^\varepsilon(u)^{-1}\| < \infty.$$

□

Remark 12. *Throughout the proof we can see that the role of the regularization parameter ε is twofold: on the one hand it is directly involved in the definition of U by* (26)*, on the other hand it acts on the conditioning of the system through* (32)*.*

4.3. Convergence of the regularized solutions

We now examine the behavior of u_ε when ε goes to zero.

Lemma 13. *If $F^\varepsilon(u^\varepsilon) = 0$ then*

$$-\varepsilon \leq F(u^\varepsilon) \leq 0.$$

Proof. We have

$$|g_{u^\varepsilon}| \leq \Phi_\varepsilon \circ g_{u^\varepsilon} = \sqrt{\varepsilon^2 + g_{u^\varepsilon}^2} \leq \varepsilon + |g_{u^\varepsilon}|,$$

which, as $u^\varepsilon \geq 0$, yields

$$F^\varepsilon(u^\varepsilon) - \varepsilon u^\varepsilon \leq F(u^\varepsilon) \leq F^\varepsilon(u^\varepsilon).$$

Using that $u^\varepsilon \leq 1$ and $F^\varepsilon(u^\varepsilon) = 0$ completes the proof. □

Theorem 14. *Let $(\varepsilon_k)_{k\in\mathbb{N}}$ be a sequence of positive numbers such that $\varepsilon_k \to 0$. For each k consider a solution $u^{\varepsilon_k} \in L^2(D)$ of $F^{\varepsilon_k}(u^{\varepsilon_k}) = 0$ and set $y^{\varepsilon_k} = E^{-1}u^{\varepsilon_k}$.*

1. *There exists a subsequence (ε_{k_p}) and $u^* \in L^2(D, [0,1])$ such that*

$$u^{\varepsilon_{k_p}} \overset{L^2(D)}{\rightharpoonup} u^*, \qquad y^{\varepsilon_{k_p}} \overset{L^2(D)}{\to} y^* := E^{-1}u^*,$$

$$u^{\varepsilon_{k_p}} \to \chi_{[g_{u^*}<0]} \text{ a.e. in } [g_{u^*} \neq 0].$$

   If $\mathcal{L}^2([g_{u^*} = 0]) = 0$, then

$$u^{\varepsilon_{k_p}} \to u^* = \chi_{[g_{u^*}<0]} \text{ a.e. in } D \qquad \text{and} \qquad u^{\varepsilon_{k_p}} \overset{L^2(D)}{\to} u^*. \tag{33}$$

   In particular, $F(u^*) = 0$ and (u^*, y^*) is a solution of (1)-(2).
2. *Every accumulation point u^* of the sequence (u^{ε_k}) in $L^2(D)$ satisfies $F(u^*) = 0$.*

Proof. The weak convergence of $(u^{\varepsilon_{k_p}})$ stems from (18) and the weak closedness of $L^2(D, [0,1])$. The strong convergence of $(y^{\varepsilon_{k_p}})$ stems from the compactness of the operator $E^{-1} : L^2(D) \to L^2(D)$. Analogously $g_{u^{\varepsilon_{k_p}}}$ strongly converges to g_{u^*} in $L^\infty(D)$. The pointwise a.e. convergence in $[g_{u^*} \neq 0]$ can be straightforwardly deduced from (19). Lebesgue's dominated convergence theorem ensures the convergence in $L^2(D)$ provided that $\mathcal{L}^2([g_{u^*} = 0]) = 0$. The second assertion of the theorem is a direct consequence of Lemma 13 together with the continuity of $F : L^2(D) \to L^2(D)$. □

5. Finite dimensional version

Proposition 2 relies on the concept of topological sensitivity, i.e., on the possibility to perturb the control variable in an infinitesimal region. Hence it has no equivalent in finite dimension. In this section we rather discuss the solution of a finite dimensional approximation of the optimality conditions given by Proposition 2, or equivalently Theorem 3. Therefore we are concerned with the solution of $F(u) = 0$ with

$$\begin{array}{rcl} F : \mathbb{R}^N & \to & \mathbb{R}^N \\ u & \mapsto & u.|g_u| + \min(0, g_u). \end{array} \tag{34}$$

Above, g_u is given by (13), E is a symmetric and invertible real matrix of order N, $y^\dagger \in \mathbb{R}^N$, and $\nu \in \mathbb{R}$. The dot stands for the coordinatewise product. We introduce the regularized functional

$$F^\varepsilon : u \mapsto u.\Phi_\varepsilon(g_u) + \min(0, g_u),$$

with Φ_ε given by (15) and applied coordinatewise. Similarly to the function space setting, F^ε is slantly differentiable on $\mathbb{R}^N$, and we obtain the slanting matrix

$$G^\varepsilon(u) = \mathrm{diag}(\Phi_\varepsilon(g_u)) + \mathrm{diag}(u.\Phi'_\varepsilon(g_u) + \chi_{[g_u<0]})E^{-2}. \tag{35}$$

By inspection of the proofs, it appears that Lemma 10, Theorem 11, Lemma 13 and Theorem 14 can be readily translated in the finite dimensional setting, which is left to the reader. As a corollary of Theorem 14, we deduce that the equation $F(u^*) = 0$ admits at least one solution, which is obtained as an accumulation point of the sequence (u^{ε_k}).

6. Computational issues

The preceding discussions suggest an algorithm made of two nested loops: an inner loop to solve $F^{\varepsilon_k}(u^{\varepsilon_k}) = 0$ for $\varepsilon_k > 0$ fixed, and an outer loop to let ε_k go to zero.

Outer loop for solving $F(u^*) = 0$ starting from $u_{ini} \in L^2(D)$

- Generate a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ of positive numbers such that $\lim_{k\to\infty} \varepsilon_k = 0$. Set $u^{\varepsilon_0} = u_{ini}$.

- Iterate for $k = 1, 2, ...$
  Solve $F^{\varepsilon_k}(u^{\varepsilon_k}) = 0$ by the semismooth Newton method initialized at $u^{\varepsilon_{k-1}}$ (see below).

- Set $u^* = \lim_{k \to +\infty} u^{\varepsilon_k}$.

Inner loop for solving $F^\varepsilon(u^\varepsilon) = 0$ starting from $u_0 \in L^2(D)$

- Iterate for $n = 0, 1, ...$
  1. Compute g_{u_n}, $F^\varepsilon(u_n)$ and $G^\varepsilon(u_n)$ by (13), (14) and (16).
  2. Solve
  $$G^\varepsilon(u_n) d_n = -F^\varepsilon(u_n). \tag{36}$$
  3. Set $u_{n+1} = u_n + d_n$.

- Set $u^\varepsilon = \lim_{n \to +\infty} u_n$.

Above, the algorithm is described in the function space setting, but the discrete version is completely similar. Let us now comment on the solution of (36). At iteration n, the slanting matrix $G^\varepsilon(u_n)$ is of the form $G^\varepsilon(u_n) = A_n + B_n E^{-2}$, where A_n and B_n are diagonal matrices according to (35). Hence the computation of $G^\varepsilon(u_n)$ requires the complete inversion of E. This task has to be done only once, but it may still be costly in case of fine discretization. In addition, the obtained matrix is full, which may rise a problem of memory. In fact, one can bypass these difficulties by replacing (36) by

$$\begin{pmatrix} A_n & 0 & B_n \\ -I & E & 0 \\ 0 & -I & E \end{pmatrix} \begin{pmatrix} d_n \\ d'_n \\ d''_n \end{pmatrix} = \begin{pmatrix} -F^\varepsilon(u_n) \\ 0 \\ 0 \end{pmatrix}. \tag{37}$$

This corresponds to the Newton iteration applied to the solution of the optimality system

$$\begin{cases} u\Phi_\varepsilon(g) + \min(0, g) = 0, \\ Ey = u, \\ E(g - \nu) = y - y^\dagger. \end{cases}$$

Although (37) is of higher dimension than (36), it is sparse and does not require any preliminary computation. One can reduce its size by reversing completely (13). Note that, in the continuous setting, this is only possible if $y^\dagger \in H_0^1(D)$, and, at the discrete level, numerical instabilities can be observed

when $y^\dagger{}_{|\partial D} \neq 0$. In this case, steps (2) and (3) of the inner loop can be replaced by solving successively

$$\begin{array}{l} (A_n E^2 + B_n) e_n = -F^\varepsilon(u_n), \\ g_{n+1} = g_n + e_n, \\ u_{n+1} = E^2(g_{n+1} - \nu) + E y^\dagger. \end{array} \tag{38}$$

In the numerical experiments which are presented in Section 7, E is chosen as the finite difference approximation of the Dirichlet Laplacian with the five points stencil. The regularization parameter is updated by

$$\varepsilon_k = \min(\varepsilon_{k-1}, a\|g_{u^{\varepsilon_{k-1}}}\| r^{k-1})$$

with $a > 0$, $0 < r < 1$ and $\varepsilon_0 = a\|g_{u^{\varepsilon_0}}\|$. Obviously, the sequence (ε_k) is nonincreasing and, due to the boundedness of $\{\|g_u\|, 0 \leq u \leq 1\}$, it goes to zero as k goes to infinity. The following parameters are used: $a = 10^{-3}$, $r = 0.5$, $k = 1, ..., 10$. The algorithm is initialized by $u_{ini} \equiv 0.5$. The stopping criterion of the inner loop is chosen as $\|u_{n+1} - u_n\|/\|u_n\| \leq b$, with $b = 10^{-3}$.

7. Numerical tests

7.1. First example

The computational domain is the square $D = [0, 1] \times [0, 1]$ discretized by a 100×100 regular grid. We consider the data

$$y^\dagger(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2), \qquad \nu = 10^{-2}.$$

As $y^\dagger{}_{|\partial D} = 0$, we use the reduced version of the algorithm given by (38). The functions u^{ε_k} obtained at the end of iterations 1 and 10 (of the outer loop) are depicted in Figure 1. The convergence history of $\|F(u^{\varepsilon_k})\|^2$ is shown in Figure 2. Some indications of the CPU time used are reported in Table 1. They have been obtained on a standard desktop computer with 2.4 GHz processor.

7.2. Second example

The computational domain and the grid are unchanged, but we now choose

$$y^\dagger = E^{-1}\chi_B, \qquad \nu = 0,$$

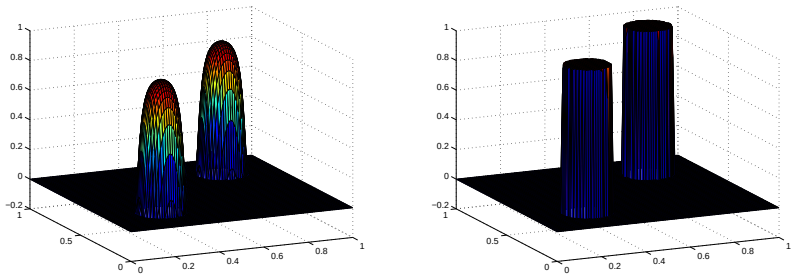

Figure 1: Example 1, control u^ε at iterations 1 (left) and 10 (right).

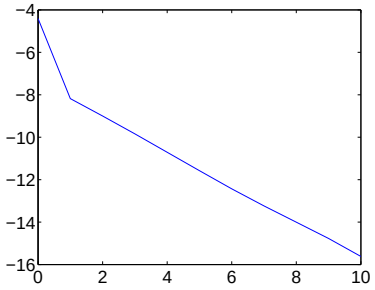

Figure 2: Example 1, convergence history of $\log_{10} \|F(u^{\varepsilon_k})\|^2$.

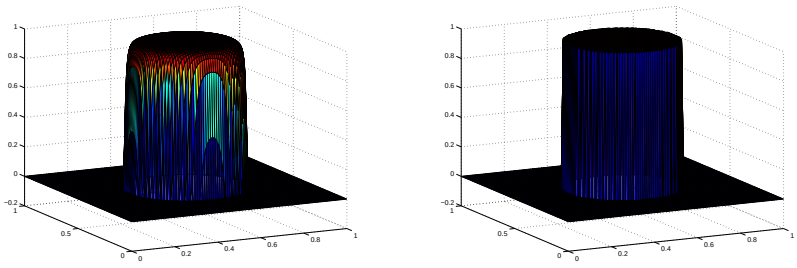

Figure 3: Example 2, control u^ε at iterations 3 (left) and 10 (right).

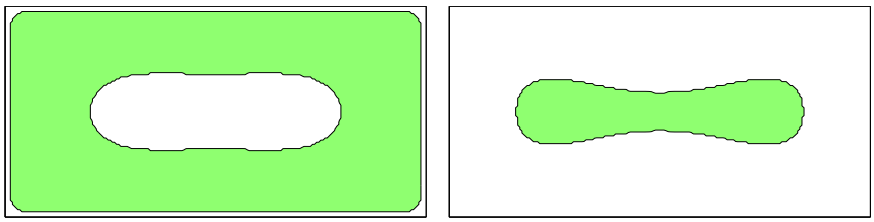

Figure 4: Example 3, control u^ε at iteration 10 for case (a) (left) and case (b) (right).

where B is the ball of center $(0.5, 0.5)$ and radius 0.25. It is a difficult problem, because the solution is $(u^*, y^*) = (\chi_B, y^\dagger)$, thus $g_{u^*} \equiv 0$. Nevertheless the algorithm still behaves properly, as shown in Figure 3. At iteration 10 we obtain the relative errors $\|y - y^\dagger\|/\|y^\dagger\| \approx 4.5 \times 10^{-7}$ and $\|u - \chi_B\|/\|\chi_B\| \approx 1.5 \times 10^{-4}$.

7.3. Third example

The computational domain is now the rectangle $D = [0, 2] \times [0, 1]$ discretized by a 200×100 regular grid. We choose

$$y^\dagger \equiv 0.05, \qquad (a)\ \nu = 10^{-4}, \qquad (b)\ \nu = 2 \times 10^{-3}.$$

Since ${y^\dagger}_{|\partial D} \neq 0$ we do not use the reduced version, but the update (37). The obtained results are displayed in Figure 4. For readability we have represented the level set $[u^\varepsilon > 0.5]$ (in dark), but in fact, like in the other examples, this region corresponds to the set of points where $u^\varepsilon \approx 1$. These results match those obtained in [4] by a level set method.

7.4. Concluding remarks

With regards to the tests performed, the proposed algorithm appears to be competitive in comparison with first order topology optimization meth-

Example	1	2	3(a)	3(b)
Nb. of d.o.f.	9801	9801	19701	19701
CPU time (s)	13.3	17.1	49.2	48.6

Table 1: Some computational data.

ods. In particular, it is still efficient when the topological derivative at the optimum is zero, which allows to be optimistic as to its application to reconstruction problems. It is also compatible with the use of primal-dual Newton type methods in topology optimization with constraints, which offers an alternative to Uzawa type methods [5].

[1] G. Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002.

[2] G. Allaire, F. de Gournay, F. Jouve, and A.-M. Toader. Structural optimization using topological and shape sensitivity via a level set method. *Control Cybern.*, 34:59–80, 2005.

[3] G. Allaire, F. Jouve, and A.-M. Toader. Structural optimization using sensitivity analysis and a level-set method. *J. Comput. Phys.*, 194(1):363–393, 2004.

[4] S. Amstutz. Analysis of a level-set method for topology optimization. Technical Report 94, Laboratoire d'analyse non linéaire et géométrie, 2009.

[5] S. Amstutz. Augmented lagrangian for cone constrained topology optimization. *Comput. Optim. Appl.*, To appear.

[6] S. Amstutz and H. Andrä. A new algorithm for topology optimization using a level-set method. *J. Comput. Phys.*, 216(2):573–588, 2006.

[7] P. M. Anselone. *Collectively compact operator approximation theory and applications to integral equations*. Prentice-Hall Inc., Englewood Cliffs, N. J., 1971. With an appendix by Joel Davis, Prentice-Hall Series in Automatic Computation.

[8] M. P. Bendsøe and O. Sigmund. *Topology optimization*. Springer-Verlag, Berlin, 2003. Theory, methods and applications.

[9] M. Bonnet. Inverse acoustic scattering by small-obstacle expansion of a misfit function. *Inverse Problems*, 24(3):035022, 27, 2008.

[10] M. Bonnet. Higher-order topological sensitivity for 2-D potential problems. Application to fast identification of inclusions. *Internat. J. Solids Structures*, 46(11-12):2275–2292, 2009.

[11] M. Burger, B. Hackl, and W. Ring. Incorporating topological derivatives into level set methods. *J. Comput. Phys.*, 194(1):344–362, 2004.

[12] S. N. Chandler-Wilde and B. Zhang. A generalized collectively compact operator theory with an application to integral equations on unbounded domains. *J. Integral Equations Appl.*, 14(1):11–52, 2002.

[13] J. R. de Faria, A. A. Novotny, R. A. Feijóo, E. Taroco, and C. Padra. Second order topological sensitivity analysis. *Internat. J. Solids Structures*, 44(14-15):4958–4977, 2007.

[14] H. Eschenauer, V. V. Kobelev, and A. Schumacher. Bubble method for topology and shape optimization of structures. *Structural optimization*, 8:42–51, 1994.

[15] J. Fehrenbach and M. Masmoudi. Coupling topological gradient and gauss-newton method. In M. Bendsoe, N. Olhoff, and O. Sigmund, editors, *IUTAM Symposium on Topological Design Optimization of Structures, Machines and Materials: Status and Perspectives*, 2006.

[16] S. Garreau, P. Guillaume, and M. Masmoudi. The topological asymptotic for PDE systems: the elasticity case. *SIAM J. Control Optim.*, 39(6):1756–1778 (electronic), 2001.

[17] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888 (electronic) (2003), 2002.

[18] F. Murat and J. Simon. Quelques résultats sur le contrôle par un domaine géométrique. *Publ. Laboratoire d'Analyse Numérique, Université Paris 6*, pages 1–46, 1974.

[19] G. Savaré. Regularity results for elliptic equations in Lipschitz domains. *J. Funct. Anal.*, 152(1):176–201, 1998.

[20] J. Sokołowski and A. Żochowski. On the topological derivative in shape optimization. *SIAM J. Control Optim.*, 37(4):1251–1272 (electronic), 1999.

[21] J. Sokołowski and J.-P. Zolésio. *Introduction to shape optimization*, volume 16 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.

[22] M. Y. Wang, X. Wang, and D. Guo. A level set method for structural topology optimization. *Comput. Methods Appl. Mech. Engrg.*, 192(1-2):227–246, 2003.