DISTRIBUTED SHAPE DERIVATIVE *VIA* AVERAGED ADJOINT METHOD AND APPLICATIONS

Antoine Laurain[1] and Kevin Sturm[2]

Abstract. The structure theorem of Hadamard–Zolésio states that the derivative of a shape functional is a distribution on the boundary of the domain depending only on the normal perturbations of a smooth enough boundary. Actually the domain representation, also known as distributed shape derivative, is more general than the boundary expression as it is well-defined for shapes having a lower regularity. It is customary in the shape optimization literature to assume regularity of the domains and use the boundary expression of the shape derivative for numerical algorithms. In this paper we describe several advantages of the distributed shape derivative in terms of generality, easiness of computation and numerical implementation. We identify a tensor representation of the distributed shape derivative, study its properties and show how it allows to recover the boundary expression directly. We use a novel Lagrangian approach, which is applicable to a large class of shape optimization problems, to compute the distributed shape derivative. We also apply the technique to retrieve the distributed shape derivative for electrical impedance tomography. Finally we explain how to adapt the level set method to the distributed shape derivative framework and present numerical results.

1. Introduction

In his research on elastic plates [19] in 1907, Hadamard showed how to obtain the derivative of a shape functional $J(\Omega)$ by considering normal perturbations of the boundary $\partial\Omega$ of a smooth set Ω. This fundamental result of shape optimization was made rigorous later by Zolésio [13] in the so-called "structure theorem". When $J(\Omega)$ and the domain are smooth enough, one may also write the shape derivative as an integral over $\partial\Omega$, which is the canonical form in the shape optimization literature.

However, when Ω is less regular, the shape derivative can often be written as a domain integral even when the boundary expression is not available. The *domain expression* also known as *distributed shape derivative* has been generally ignored in the shape optimization literature for several reasons: firstly the boundary representation provides a straightforward way of determining an explicit descent direction since it depends linearly on the

[1] Technische Universität Berlin, Institut für Mathematik, Str. des 17. Juni 136, 10623 Berlin, Germany. `laurain@ime.usp.br`

[2] Universität Duisburg-Essen, Fakultät für Mathematik, Thea-Leymann-Straße 9, 45127 Essen, Germany. `kevin.sturm@uni-due.de`

boundary perturbation θ and not on its gradient, secondly this descent direction only needs to be defined on the boundary. When considering the domain expression, these two advantages disappear as the shape derivative is defined on Ω and depends on the gradient of θ, so that a partial differential equation needs to be solved to obtain a descent direction θ on Ω.

It seems that these drawbacks would definitely rule out the distributed shape derivative, however they turn out to be less dramatic than expected in many situations and the domain formulation has other less foreseeable advantages over the boundary representation. In this paper we advocate for the use of the distributed shape derivative and discuss the advantages of this formulation.

The boundary representation has the following drawbacks. First of all if the data is not smooth enough the integral representation does not exist so that the more general domain representation is the only rigorous alternative. Even when the boundary representation exists and has the form $\int_{\partial\Omega} g\,\theta\cdot n$, it is usually not legitimate to choose $\theta\cdot n = -g$ on $\partial\Omega$ for a descent direction if g is not smooth enough, for instance if $g \in L^1(\partial\Omega)$. Therefore, a smoother θ must be chosen, which requires to solve a partial differential equation on the boundary $\partial\Omega$. When taking $\theta \cdot n = -g$ is legitimate, it might still not be desirable as this may yield a θ with low regularity, in which case one needs to regularize θ on the boundary as well. In these cases the first advantage of the boundary representation disappears. The second advantage of the boundary representation is that the perturbation field only needs to be defined on the boundary instead of on the whole domain, reducing the cost of the computation. Actually, the distributed shape derivative also has its support on the boundary, and may be computed in a small neighborhood of the boundary so that the additional cost is minimal. In addition, in most shape optimization applications, g is the restriction of a function defined in a neighborhood of the boundary and not a quantity depending only on the boundary such as the curvature. Therefore from a practical point of view, g must be evaluated in a neighborhood of $\partial\Omega$ anyway. Also, in many numerical applications, θ must be extended to a neighborhood of Γ or even to the entire domain Ω. This is the case for level set methods for instance, where the level set function must be updated on Ω, or when one wishes to update the mesh along with the domain update, to avoid re-meshing the new domain. The distributed shape derivative then directly gives an extension of θ well-suited to the optimization problem.

Recent results have shown that the distributed shape derivative is also more accurate than the boundary representation from a numerical point of view; see [27] for a comparison. Indeed functions such as gradients of the state and adjoint state appearing in the distributed shape derivative only need to be defined at grid points and not on the interface. Therefore one avoids interpolation of these irregular terms. This is particularly useful for transmission problems where the boundary representation requires to compute the jump of a function over the interface, a delicate and error-prone operation from the numerical point of view.

Having considered these equivalent expressions of the shape derivative (*i.e.* boundary and domain expression) leads to a general form of the shape derivative using tensors. We introduce such a tensor representation in Section 4 which covers a large class of problems and in particular contains the boundary and domain expression. We show how this abstract form allows to identify simple relations between the domain and boundary expressions of the shape derivative.

In this paper we also extend and simplify the averaged adjoint method from [42], a Lagrangian-type method which is well-suited to compute the shape derivative of a cost function in an efficient way. Lagrangian methods are commonly used in shape optimization and have the advantage of providing the shape derivative without the need to compute the material derivative of the state; see [4, 9, 12, 37, 42, 43]. Compared to these known shape-Lagrangian methods, the averaged adjoint method is fairly general due to minimal required conditions. The assumptions are for instance less restrictive than those required for the theorem of Correa-Seeger [12], therefore it can be applied to more general situations such as non-convex functionals. As the direct approach our method can also be applied for problems depending on nonlinear partial differential equations. In this paper we give an example of application to a transmission problem (in electrical impedance tomography – see Sect. 6). Our method provides the domain expression of the shape derivative and the boundary expression can be computed easily from the tensor representation of the domain expression.

To complete the numerical implementation aspect, we also show how the domain expression of the shape derivative can be used in the level set method framework [3,15–18,23,24,34]. The level set method can be modified to use the domain expression which leads to a method which is actually easier to implement. Combining all these techniques, we obtain a straightforward and general way of solving the shape optimization problem, from the rigorous theoretical computation of the shape derivative to the numerical implementation.

In Section 2 we recall the concept of shape derivative and the structure theorem on an abstract level. In Section 3 a shape-Lagrangian method, the averaged adjoint method, is described. In Section 4 we identify a general tensor representation of the shape derivative, establish some of its properties, and give a few examples. In Section 5 we explain how to compute descent directions for the distributed shape derivative for use in gradient methods. In Section 6 we apply the results of Sections 3 and 4 to the inverse problem of electrical impedance tomography. In Section 7 we extend the level set method to the case of the distributed shape derivative and finally in Section 8 we show numerical results for various problems including the problem of electrical impedance tomography.

2. The structure theorem revisited

Our aim in this section is to describe properties of the shape derivative on an abstract level and to emphasize that all representations of the shape derivative satisfy the same *structure theorem.*

Let $\mathcal{P}(D)$ be the set of subsets of $D \subset \mathbf{R}^d$ compactly contained in D, where the so-called "universe" $D \subset \mathbf{R}^d$ is assumed to be open and bounded. Define for $k \geq 0$ and $0 \leq \alpha \leq 1$,

$$C_c^{k,\alpha}(D, \mathbf{R}^d) := \{\theta \in C^{k,\alpha}(D, \mathbf{R}^d)| \quad \theta \text{ has compact support in } D\}. \tag{2.1}$$

Also for given domain $\Omega \subset D$ with at least a C^1 boundary we introduce the space of vector field

$$C_{\partial\Omega}^{k,\alpha}(D, \mathbf{R}^d) := \{\theta \in C_c^{k,\alpha}(D, \mathbf{R}^d)| \quad \theta \cdot n = 0 \text{ on } \partial\Omega\} \tag{2.2}$$

where n is the outward unit normal vector to Ω.

Consider a vector field $\theta \in C_c^{0,1}(D, \mathbf{R}^d)$ and the associated flow $\Phi_t^\theta : \overline{D} \to \mathbf{R}^d$, $t \in [0, \tau]$ defined for each $x_0 \in \overline{D}$ as $\Phi_t^\theta(x_0) := x(t)$, where $x : [0, \tau] \to \mathbf{R}$ solves

$$\dot{x}(t) = \theta(x(t)) \quad \text{for } t \in (0, \tau), \quad x(0) = x_0. \tag{2.3}$$

We will sometimes use the simpler notation $\Phi_t = \Phi_t^\theta$ when no confusion is possible. Since $\theta \in C_c^{0,1}(D, \mathbf{R}^d)$ we have by Nagumo's theorem [32] that for fixed $t \in [0, \tau]$ the flow Φ_t is a homeomorphism from D into itself and maps boundary onto boundary and interior onto interior. Further, we consider the family

$$\Omega_t := \Phi_t^\theta(\Omega) \tag{2.4}$$

of perturbed domains.

In the following let $J : \mathfrak{P} \to \mathbf{R}$ be a shape function defined on some admissible set $\mathfrak{P} \subset \mathcal{P}(D)$.

Definition 2.1. The Eulerian semiderivative of J at Ω in direction $\theta \in C_c^{0,1}(D, \mathbf{R}^d)$, when the limit exists, is defined by

$$\mathrm{d}J(\Omega)(\theta) := \lim_{t \searrow 0} \frac{J(\Omega_t) - J(\Omega)}{t}. \tag{2.5}$$

(i) J is said to be *shape differentiable* at Ω if it has a Eulerian semiderivative at Ω for all $\theta \in C_c^\infty(D, \mathbf{R}^d)$ and the mapping

$$\mathrm{d}J(\Omega) : C_c^\infty(D, \mathbf{R}^d) \to \mathbf{R}, \ \theta \mapsto \mathrm{d}J(\Omega)(\theta)$$

is linear and continuous, in which case $\mathrm{d}J(\Omega)(\theta)$ is called the *shape derivative* at Ω.

(ii) The shape derivative $\mathrm{d}J(\Omega)$ is of finite order if there is an integer $l \geq 0$ and a constant $c > 0$ such that for each compact $K \subset D$

$$|\mathrm{d}J(\Omega)(\theta)| \leq c\|\theta\|_l \quad \forall \theta \in C_c^\infty(K, \mathbf{R}^d),$$

where $\|\theta\|_l := \sum_{|\alpha| \leq l} |D^\alpha \theta|_\infty$. The smallest such integer $l \geq 0$ is called order of $\mathrm{d}J(\Omega)$.

The shape derivative from Definition 2.1 has a particular structure. Intuitively, it is clear that the form functional stays constant for a transformation Φ that leaves Ω unchanged, that is $\Phi(\Omega) = \Omega$, even if some points inside Ω move and consequently the shape derivative is zero in this case. This property is valid when Ω is open or closed; cf. [13]. Mathematically, this is expressed in the following basic theorem proved in [44].

Theorem 2.2. *Let $\Omega \in \mathfrak{P}$ be open or closed. Let $\theta \in C_c^{0,1}(D, \mathbf{R}^d)$ be a vector field with compact support in Ω and denote by Φ_t its flow defined in (2.3). Then we have*

$$\mathrm{d}J(\Omega)(\theta) = 0.$$

Note that the shape derivative of $J(\Omega)$ always exists for vector fields with compact support in Ω, even if it does not exist for other vector fields. An important consequence of Theorem 2.2, also for numerical methods, is that *independently of the representation* of the shape derivative and the regularity of the domain Ω, the values of θ outside the boundary of Ω have no influence on the shape derivative.

Corollary 2.3. *Let $\Omega \in \mathfrak{P}$ be a set with C^1-boundary. Assume that J is shape differentiable on $\mathfrak{P}$. Let $\theta \in C_{\partial\Omega}^{0,1}(D, \mathbf{R}^d)$. Then we have*

$$\mathrm{d}J(\Omega)(\theta) = 0.$$

The previous discussion immediately yields the following fundamental result of shape optimization.

Theorem 2.4 (Structure theorem). *Assume $\Gamma := \partial\Omega$ is compact and J is shape differentiable. Denote the shape derivative by*

$$\mathrm{d}J(\Omega) : C_c^\infty(D, \mathbf{R}^d) \to \mathbf{R}, \quad \theta \mapsto \mathrm{d}J(\Omega)(\theta). \tag{2.6}$$

Assuming $\mathrm{d}J(\Omega)$ is of order $k \geq 0$ and Γ of class C^{k+1}, then there exists a linear and continuous functional $g : C^k(\Gamma) \to \mathbf{R}$ such that

$$\mathrm{d}J(\Omega)(\theta) = g(\theta_{|\Gamma} \cdot n), \tag{2.7}$$

Proof. See ([13], pp. 480–481).   □

3. Shape derivatives *via* averaged adjoint method

Lagrangian methods in shape optimization allow to compute the shape derivative of functions depending on the solution of partial differential equations without the need to compute the material derivative of the partial differential equations; see [13] for a description of such a method in the linear case. Here we extend and simplify the averaged adjoint method, a Lagrangian-type method introduced in [42]. With this approach the computation of the domain representation of the shape derivative is fast, the retrieval of the boundary form is convenient, and no saddle point assumptions is required unlike in [13].

Let two vector spaces $E = E(\Omega), F = F(\Omega)$ and $\tau > 0$ be given, and consider a parameterization $\Omega_t = \Phi_t(\Omega)$ for $t \in [0, \tau]$. Ultimately, our goal is to differentiate shape functions of the type $J(\Omega_t)$ which can be written using a Lagrangian as $J(\Omega_t) = \mathcal{L}(\Omega_t, u^t, \hat{\psi})$, where $u^t \in E(\Omega_t)$ and $\hat{\psi} \in F(\Omega_t)$. The main appeal of the Lagrangian is that we actually only need to compute the derivative with respect to t of $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$ to compute the derivative of $J(\Omega_t)$, indeed this is the main result of Theorem 3.1, but this requires a few explanations.

Since $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$, is often constituted of integrals on $\Phi_t(\Omega)$, using a change of variable we can rewrite these integrals to integrals on the fixed domain Ω, and consequently transfer the dependence on t to the integrand.

However, in the process appear the composed functions $\hat{\varphi} \circ \Phi_t \in E(\Omega)$ and $\hat{\psi} \circ \Phi_t \in F(\Omega)$, whose derivatives are not straightforward to compute since $\hat{\varphi}$ and $\hat{\psi}$ are defined on the moving spaces $E(\Omega_t)$ and $F(\Omega_t)$.

Fortunately, and this is the crucial point of the shape-Lagrangian approach, to compute the shape derivative we can reparameterize the problem by considering $\mathcal{L}(\Omega_t, \Psi_t \circ \varphi, \Psi_t \circ \psi)$ instead of $\mathcal{L}(\Omega_t, \hat{\varphi}, \hat{\psi})$, where Ψ_t is an appropriate bijection between $E(\Omega)$ and $E(\Omega_t)$, and $\varphi \in E(\Omega)$, $\psi \in F(\Omega)$. Now the change of variable in the integrals yields functions φ and ψ in the integrand, which are defined on fixed spaces. In this paper E and F are H^1-spaces, and in this case we may consider the particular reparameterization $\mathcal{L}(\Omega_t, \varphi \circ \Phi_t^{-1}, \psi \circ \Phi_t^{-1})$. For spaces such as $H(\mathrm{curl}; \Omega)$, other transformations Ψ_t can be used; see [21, 26, 28].

Thus we are led to consider general functions of the type $G : [0, \tau] \times E \times F \to \mathbf{R}$ with

$$G(t, \varphi, \psi) := \mathcal{L}(\Phi_t(\Omega), \varphi \circ \Phi_t^{-1}, \psi \circ \Phi_t^{-1}).$$

This is precisely what we do in (6.29) when showcasing an application of the method. The main result of this section, Theorem 3.1, shows that to obtain the shape derivative of $\mathcal{L}$, it is enough to compute the derivative with respect to t of G while assigning the proper values to φ and ψ. The main ingredient is the introduction of the averaged adjoint equation.

In addition, in this paper we consider the following specific form

$$G(t, \varphi, \psi) := a(t, \varphi, \psi) + b(t, \varphi), \tag{3.1}$$

where

$$a : [0, \tau] \times E \times F \to \mathbf{R}, \qquad b : [0, \tau] \times E \to \mathbf{R},$$

are functions such that $\psi \mapsto a(t, \varphi, \psi)$ is linear for all $t \in [0, \tau]$ and $\varphi \in E$. The function G is commonly called *Lagrangian*, hence the name of the method. In the applications we have in mind, the function b arises from the objective function while a corresponds to the constraint, after transporting back to the fixed domain Ω.

Throughout the paper, the Greek letters φ and ψ are used for variables, while the roman letters u, p are used for the solutions of the state and adjoint states, respectively.

Let us assume that for each $t \in [0, \tau]$ the equation

$$d_\psi G(t, u^t, 0; \hat{\psi}) = a(t, u^t, \hat{\psi}) = 0 \ \text{ for all } \hat{\psi} \in F. \tag{3.2}$$

admits a unique solution $u^t \in E$. Further we make the following assumptions for G.

Assumption (H0). *For every $(t, \psi) \in [0, \tau] \times F$*

(i) $[0, 1] \ni s \mapsto G(t, su^t + s(u^t - u^0), \psi)$ *is absolutely continuous.*
(ii) $[0, 1] \ni s \mapsto d_\varphi G(t, su^t + (1 - s)u^0, \psi; \hat{\varphi})$ *belongs to $L^1(0, 1)$ for all $\hat{\varphi} \in E$.*

When Assumption (H0) is satisfied, for $t \in [0, \tau]$ we introduce the *averaged adjoint equation* associated with u^t and u^0: Find $p^t \in F$ such that

$$\int_0^1 d_\varphi G(t, su^t + (1 - s)u^0, p^t; \hat{\varphi}) \, \mathrm{d}s = 0 \quad \text{ for all } \hat{\varphi} \in E. \tag{3.3}$$

Notice that, in view Assumption (H0), for all $t \in [0, \tau]$,

$$G(t, u^t, p^t) - G(t, u^0, p^t) = \int_0^1 d_\varphi G(t, su^t + (1 - s)u^0, p^t; u^t - u^0) \, \mathrm{d}s = 0. \tag{3.4}$$

We can now state the main result of this section.

Assumption (H1). *We assume that*

$$\lim_{t \searrow 0} \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} = \partial_t G(0, u^0, p^0).$$

Theorem 3.1. *Let* (H0) *and* (H1) *be satisfied and assume there exists a unique solution p^t of the averaged adjoint equation* (3.3). *Then for $\psi \in F$ we obtain*

$$\frac{\mathrm{d}}{\mathrm{d}t}b(t,u^t)|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(G(t,u^t,\psi))|_{t=0} = \partial_t G(0,u^0,p^0). \tag{3.5}$$

Proof. Put $g(t) := G(t,u^t,0) - G(0,u^0,0)$, and note that $g(t) = G(t,u^t,\psi) - G(0,u^0,\psi)$ for all $\psi \in F$ and $g(0) = 0$. We have to show that

$$g'(0) := \lim_{t \searrow 0} \frac{G(t,u^t,0) - G(0,u^0,0)}{t} \quad \text{exists.}$$

Thanks to Assumption (H0) we can define the averaged adjoint p^t and using that G is affine with respect to the third argument, we obtain

$$g(t) = \underbrace{G(t,u^t,p^t) - G(t,u^0,p^t)}_{=0 \text{ in view of } (3.4)} + G(t,u^0,p^t) - G(0,u^0,p^t).$$

Dividing by $t > 0$ and using Assumption (H1) yields

$$g'(0) = \lim_{t \searrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \searrow 0} \frac{G(t,u^0,p^t) - G(0,u^0,p^t)}{t} = \partial_t G(0,u^0,p^0)$$

which concludes the proof. $\square$

Remark 3.2. In terms of a and b, equation (3.3) reads:

$$\int_0^1 d_\varphi a(t, su^t + (1-s)u^0, p^t; \hat{\varphi})\,\mathrm{d}s = -\int_0^1 d_\varphi b(t, su^t + (1-s)u^0; \hat{\varphi})\,\mathrm{d}s$$

for all $\hat{\varphi} \in E$. If $\varphi \mapsto a(t,\varphi,\psi)$ is in addition linear, then (3.3) becomes

$$a(t,\hat{\varphi},p^t) = -\int_0^1 d_\varphi b(t, su^t + (1-s)u^0; \hat{\varphi})\,\mathrm{d}s$$

for all $\hat{\varphi} \in E$.

4. TENSOR REPRESENTATION OF THE SHAPE DERIVATIVE

In this section we identify tensor representations of the shape derivative that correspond to a large class of problems studied in the literature for PDE-constrained shape optimization. This tensor representation has several interesting properties that we investigate. In particular we exhibit the link between this tensor representation and the usual boundary expression of the shape derivative.

4.1. Definition and properties

Definition 4.1. Let $\Omega \in \mathfrak{P}$ be a set with C^k-boundary, $k \geq 1$. A shape differentiable function J of order k is said to admit a tensor representation if there exist tensors $\mathbf{S}_l \in L^1(D, \mathcal{L}^l(\mathbf{R}^d, \mathbf{R}^d))$ and $\mathfrak{S}_l \in L^1(\partial\Omega; \mathcal{L}^l(\mathbf{R}^d, \mathbf{R}^d))$, $l = 0,\ldots,k$, such that

$$\mathrm{d}J(\Omega)(\theta) = \sum_{l=0}^k \int_D \mathbf{S}_l \cdot D^l\theta\,\mathrm{d}x + \int_{\partial\Omega} \mathfrak{S}_l \cdot D_\Gamma^l\theta\,\mathrm{d}s \quad \text{for all } \theta \in C_c^k(D, \mathbf{R}^d), \tag{4.1}$$

where $D_\Gamma\theta := D\theta - (D\theta n) \otimes n$ is the tangential derivative of θ along $\partial\Omega$. Here $\mathcal{L}^l(\mathbf{R}^d, \mathbf{R}^d)$ denotes the space of multilinear maps from $\mathbf{R}^d \times \cdots \times \mathbf{R}^d$ to $\mathbf{R}^d$.

Most if not all examples involving PDEs from classical textbooks [13, 20, 41] can be written in the form (4.1).

Remark 4.2.

(a) A particular case of the tensor representation (4.1) is the Eshelby energy momentum tensor in continuum mechanics introduced in [14]; see also [33]. In this case only $\mathbf{S}_1$ is not zero.
(b) When J is is shape differentiable in Ω then by definition $\theta \mapsto \mathrm{d}J(\Omega)(\theta)$ is a distribution, and if $\partial\Omega$ is compact, the distribution $\theta \mapsto \mathrm{d}J(\Omega)(\theta)$ is of finite order.
(c) If $\mathrm{d}J(\Omega)$ is of order $k = 1$ and $|\mathrm{d}J(\Omega)(\theta)| \leq C\|\theta\|_{H^1(D,\mathbf{R}^d)}$ for all $\theta \in C_c^\infty(D, \mathbf{R}^d)$ then by density of $C_c^\infty(D, \mathbf{R}^d)$ in $H_0^1(D, \mathbf{R}^d)$ the derivative $\mathrm{d}J(\Omega)$ extends to a continuous functional on $H_0^1(D, \mathbf{R}^d)$, that is,

$$|\widehat{\mathrm{d}J(\Omega)}(\theta)| \leq c\|\theta\|_{H^1(D,\mathbf{R}^d)} \quad \text{for all } \theta \in H_0^1(D, \mathbf{R}^d).$$

Therefore by the theorem of Riesz, we obtain a vector field W in $H_0^1(D, \mathbf{R}^d)$ such that

$$\forall \theta \in H_0^1(D, \mathbf{R}^d), \quad \widehat{\mathrm{d}J(\Omega)}(\theta) = \int_D DW \cdot D\theta + W \cdot \theta \, \mathrm{d}x$$

and this defines a tensor representation with $\mathbf{S}_1 = DW$, $\mathbf{S}_0 = W$, $\mathfrak{S}_1 = 0$ and $\mathfrak{S}_0 = 0$.
(d) The assumption that Ω be a set of class C^k can be reduced if $\mathfrak{S}_l \equiv 0$ for all $0 \leq k_0 \leq l \leq k$.

The tensor representation (4.1) is not unique in the sense that there might be several ways to choose the tensors $\mathbf{S}_l$ and $\mathfrak{S}_l$. This is expressed by the fact that these tensors are correlated. We describe these relations below in the case $k = 1$ in Proposition 4.3, which also describes the link between the tensor representation and the usual boundary representation (2.7) of the shape derivative.

Proposition 4.3. *Let Ω be a subset of D with C^1-boundary. Suppose that the derivative $\mathrm{d}J(\Omega)$ has the representation*

$$\mathrm{d}J(\Omega)(\theta) = \int_D \mathbf{S}_1 \cdot D\theta + \mathbf{S}_0 \cdot \theta \, \mathrm{d}x + \int_{\partial\Omega} \mathfrak{S}_1 \cdot D_\Gamma\theta + \mathfrak{S}_0 \cdot \theta \, \mathrm{d}s. \tag{4.2}$$

If $\mathbf{S}_l$ is of class $W^{1,1}$ in Ω and $D \setminus \overline{\Omega}$ then indicating by $+$ and $-$ the restrictions of the tensors to Ω and $D \setminus \overline{\Omega}$, respectively, we get

$$\begin{aligned} -\operatorname{div}(\mathbf{S}_1^+) + \mathbf{S}_0^+ &= 0 \quad \text{in } \Omega \\ -\operatorname{div}(\mathbf{S}_1^-) + \mathbf{S}_0^- &= 0 \quad \text{in } D \setminus \overline{\Omega}. \end{aligned} \tag{4.3}$$

Moreover, we can rewrite the tensor representation as a distribution on the boundary:

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} [(\mathbf{S}_1^+ - \mathbf{S}_1^-)n] \cdot \theta + \mathfrak{S}_1 \cdot D_\Gamma\theta + \mathfrak{S}_0 \cdot \theta \, \mathrm{d}s$$

where n denotes the outward unit normal vector to Ω.

If the boundary $\partial\Omega$ is C^2 and $\mathfrak{S}_1 \in W^{1,1}(\partial\Omega; \mathcal{L}^1(\mathbf{R}^d, \mathbf{R}^d))$, then we obtain a more regular distribution, the so-called boundary expression of the shape derivative:

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} g_1 \, \theta \cdot n \, \mathrm{d}s, \tag{4.4}$$

where

$$g_1 := [(\mathbf{S}_1^+ - \mathbf{S}_1^-)n] \cdot n + \mathfrak{S}_0 \cdot n + \mathfrak{S}_1 \cdot D_\Gamma n - \operatorname{div}_\Gamma(\mathfrak{S}_1^T n) + \mathcal{H}(\mathfrak{S}_1^T n \cdot n). \tag{4.5}$$

and $\mathcal{H} = \operatorname{div}_\Gamma n$ denotes the mean curvature[3] of $\partial\Omega$ while $\operatorname{div}_\Gamma := \operatorname{tr}(D_\Gamma)$ is the tangential divergence.

---

[3] We define the mean curvature as the sum of the principal curvatures κ_i, that is, $\mathcal{H} := \sum_{i=1}^d \kappa_i$.

Proof. Applying Theorem 2.2 we have

$$\mathrm{d}J(\Omega)(\theta) = \int_D \mathbf{S}_1 \cdot D\theta + \mathbf{S}_0 \cdot \theta \,\mathrm{d}x + \int_{\partial\Omega} \mathfrak{S}_1 \cdot D_\Gamma \theta + \mathfrak{S}_0 \cdot \theta \,\mathrm{d}s = 0 \quad \text{for all } \theta \in C_c^1(\Omega \cup (D \setminus \overline{\Omega}), \mathbf{R}^d).$$

An integration by parts shows (4.3).

Then, when $\partial\Omega$ is C^1, replacing (4.3) in the expression of the shape derivative and using Green's formula we obtain

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} \mathfrak{S}_1 \cdot D_\Gamma \theta + \mathfrak{S}_0 \cdot \theta \,\mathrm{d}s + \int_{\partial\Omega} [(\mathbf{S}_1^+ - \mathbf{S}_1^-)n] \cdot \theta \,\mathrm{d}s$$

$$+ \int_\Omega \underbrace{(-\operatorname{div}(\mathbf{S}_1^+) + \mathbf{S}_0^+)}_{=0} \cdot \theta \,\mathrm{d}x + \int_{D\setminus\overline{\Omega}} \underbrace{(-\operatorname{div}(\mathbf{S}_1^-) + \mathbf{S}_0^-)}_{=0} \cdot \theta \,\mathrm{d}x$$

$$\overset{(4.3)}{=} \int_{\partial\Omega} [(\mathbf{S}_1^+ - \mathbf{S}_1^-)n] \cdot \theta + \mathfrak{S}_1 \cdot D_\Gamma \theta + \mathfrak{S}_0 \cdot \theta \,\mathrm{d}s \quad \text{for all } \theta \in C_c^1(D, \mathbf{R}^d). \tag{4.6}$$

With a slight abuse of notation we keep the same notation n for the extension of the normal to a neighborhood of $\partial\Omega$. Let $\theta \in C^1(\overline{D}, \mathbf{R}^d)$ and define $\theta_\tau := \theta - (\theta \cdot n)n$ the tangential part of θ. Then $\theta_\tau \cdot n = 0$ on $\partial\Omega$ and hence from the structure theorem we get $\mathrm{d}J(\Omega)(\theta_\tau) = 0$ which yields in view of (4.6):

$$\mathrm{d}J(\Omega)(\theta) = \mathrm{d}J(\Omega)((\theta \cdot n)n)$$

$$= \int_{\partial\Omega} ((\mathbf{S}_1^+ - \mathbf{S}_1^-)n \cdot n)(\theta \cdot n) + \mathfrak{S}_1 \cdot D_\Gamma(n(\theta \cdot n)) + (\mathfrak{S}_0 \cdot n)(\theta \cdot n) \,\mathrm{d}s$$

$$= \int_{\partial\Omega} ((\mathbf{S}_1^+ - \mathbf{S}_1^-)n \cdot n)(\theta \cdot n) + (\mathfrak{S}_0 \cdot n)(\theta \cdot n) \,\mathrm{d}s$$

$$+ \int_{\partial\Omega} \mathfrak{S}_1 \cdot D_\Gamma n(\theta \cdot n) + n \cdot \mathfrak{S}_1 \nabla_\Gamma(\theta \cdot n) \,\mathrm{d}s, \tag{4.7}$$

where we used that for all functions $f \in C^1(\mathbf{R}^d, \mathbf{R}^d)$ and $g \in C^1(\mathbf{R}^d)$ we have

$$D(gf) = gDf + f \otimes \nabla g. \tag{4.8}$$

Finally using $\mathfrak{S}_1 \in W^{1,1}(\partial\Omega; \mathcal{L}^1(\mathbf{R}^d, \mathbf{R}^d))$ we integrate by parts on the boundary $\partial\Omega$ to transform the last term in (4.7)

$$\int_{\partial\Omega} n \cdot \mathfrak{S}_1 \nabla_\Gamma(\theta \cdot n) \,\mathrm{d}s = \int_{\partial\Omega} \left(-\operatorname{div}_\Gamma\left(\mathfrak{S}_1^T n\right) + \mathcal{H}\left(\mathfrak{S}_1^T n \cdot n\right) \right)(\theta \cdot n)\mathrm{d}s.$$

Therefore (4.7) reads

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} ((\mathbf{S}_1^+ - \mathbf{S}_1^-)n \cdot n)(\theta \cdot n) + (\mathfrak{S}_0 \cdot n)(\theta \cdot n) \,\mathrm{d}s$$

$$+ \int_{\partial\Omega} \mathfrak{S}_1 \cdot D_\Gamma n(\theta \cdot n) + (-\operatorname{div}_\Gamma(\mathfrak{S}_1^T n) + \mathcal{H}(\mathfrak{S}_1^T n \cdot n))(\theta \cdot n) \,\mathrm{d}s, \tag{4.9}$$

which we can rewrite as (4.4). □

Remark 4.4. In Proposition 4.3, if $\mathfrak{S}_1 \equiv 0$, one can still obtain (4.3) when Ω is only Lipschitz instead of C^1.

Corollary 4.5. *Let the assumptions of Proposition 4.3 be satisfied. Suppose that the tensor $\mathfrak{S}_1 \colon \partial\Omega \to \mathcal{L}(\mathbf{R}^d, \mathbf{R}^d)$ has the form $\mathfrak{S}_1 = \alpha(I - n \otimes n)$, where $\alpha \in C^0(\partial\Omega)$. Then (4.4) simplifies to*

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} g_1 \,\theta \cdot n \,\mathrm{d}s, \tag{4.10}$$

where g_1 is given by

$$g_1 := [(\mathbf{S}_1^+ - \mathbf{S}_1^-)n] \cdot n + \mathfrak{S}_0 \cdot n + \alpha\mathcal{H}.$$

Proof. First $\mathfrak{S}_1^T = \alpha(I - n \otimes n)^T = \mathfrak{S}_1$ and $\mathfrak{S}_1^T n = \alpha(I - n \otimes n)n = \alpha(n - (n \cdot n)n) = 0$, thus the two last terms in (4.5) vanish. Concerning the third term in (4.5) we write

$$\mathfrak{S}_1 \cdot D_\Gamma n = \alpha(I - n \otimes n) \cdot D_\Gamma n = \alpha(\mathrm{tr}(D_\Gamma n) - (n \otimes n) \cdot D_\Gamma n) = \alpha(\mathrm{div}_\Gamma n - (D_\Gamma nn) \cdot n) = \alpha \mathcal{H}.$$

where we have used $(D_\Gamma nn) \cdot n = 0$. $\qquad\square$

Remark 4.6. The particular tensor $\mathfrak{S}_1 = \alpha(I - n \otimes n)$ in Corollary 4.5 is commonly encountered in shape optimization problems. In fact, (4.10) corresponds to a standard formula that can be found in most textbooks on shape optimization.

Remark 4.7. Recall that for given vector fields θ, ζ, the second order shape derivative is defined by

$$\mathrm{d}^2 J(\Omega)(\theta)(\zeta) := \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{d}J(\Phi_t^\zeta(\Omega))(\theta)|_{t=0}.$$

Once we have identified a tensor representation (4.1) for the shape derivative $\mathrm{d}J(\Omega)(\theta)$ for fixed θ, it is convenient to differentiate it to also obtain a tensor representation for the second-order shape derivative. Further, Proposition 4.3 or Corollary 4.5 can also be applied to obtain a boundary expression for the second order shape derivative.

Similar relations as in Proposition 4.3 could be obtained for any tensor representation of order k. For instance in the case $k = 2$ we obtain the relations

$$\begin{aligned}
\mathcal{A}\mathbf{S}_2^+ - \mathrm{div}(\mathbf{S}_1^+) + \mathbf{S}_0^+ = 0 \quad &\text{in } \Omega, \\
\mathcal{A}\mathbf{S}_2^- - \mathrm{div}(\mathbf{S}_1^-) + \mathbf{S}_0^- = 0 \quad &\text{in } D \setminus \overline{\Omega},
\end{aligned} \tag{4.11}$$

where $(\mathcal{A}\mathbf{S}_2)_l = \sum_{i,j=1}^d \partial_{x_i x_j}(\mathbf{S}_2)_{ijl}$.

Using the averaged adjoint approach from Theorem 3.1 yields the tensor representation (4.1) of the shape derivative. Then Proposition 4.3 can be used to immediately derive the standard boundary expression of the shape gradient from this tensor representation.

4.2. Examples of tensor representations

In this section we present several examples of representations corresponding to Definition 4.1 and apply the observations from Section 4.1.

First order tensor representation

A basic example of a first order tensor representation of the shape derivative is for

$$J(\Omega) = \int_\Omega f \, \mathrm{d}x + \int_{\partial\Omega} g \, \mathrm{d}s$$

with $f, g \in C^2(\mathbf{R}^d)$. Then one easily computes

$$\mathrm{d}J(\Omega)(\theta) = \int_\Omega \nabla f \cdot \theta + f \, \mathrm{div}(\theta) \, \mathrm{d}x + \int_{\partial\Omega} \nabla g \cdot \theta + g \, \mathrm{div}_\Gamma \theta \, \mathrm{d}s.$$

The corresponding tensor representation (4.1) is

$$\mathbf{S}_1^+ := fI, \quad \mathbf{S}_1^- := 0, \quad \mathbf{S}_0^+ := \nabla f, \quad \mathbf{S}_0^- := 0, \quad \mathfrak{S}_1 := g(I - n \otimes n), \quad \mathfrak{S}_0 := \nabla g.$$

Note that $\mathfrak{S}_1$ has the form assumed in Corollary 4.5. Applying this Corollary, assuming the domain has enough regularity, we obtain in view of (4.10) the classical formula:

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} g_1 \, \theta \cdot n \, \mathrm{d}s,$$

where g_1 is given by

$$g_1 := f + \partial_n g + g\mathcal{H}.$$

Note that in the particular case $f = 0$ we have obtained as a byproduct the formula

$$\int_{\partial\Omega} \nabla g \cdot \theta + g \operatorname{div}_\Gamma \theta \, \mathrm{d}s = \int_{\partial\Omega} (\partial_n g + g\mathcal{H}) \, \theta \cdot n \, \mathrm{d}s, \tag{4.12}$$

and when in addition $\partial_n g = 0$ or g is defined only on $\partial\Omega$, (4.12) becomes the classical *tangential Green's formula*; see for instance ([20], Prop. 5.4.9).

Non-homogeneous Dirichlet problem

The following problem was already considered for instance in [13]. Here we present a fairly easy way to compute the shape derivative. Let Ω and D be two open and bounded subset of $\mathbf{R}^d$ with $\Omega \subset D$. Consider

$$-\Delta v = f \text{ in } \Omega, \tag{4.13}$$
$$v = g \text{ on } \partial\Omega, \tag{4.14}$$

where $f \in L^2(D)$ and $g \in H^2(D)$. Introducing the variable $u := v - g$, we observe that (4.13)–(4.14) is equivalent to the homogeneous Dirichlet problem

$$-\Delta u = \Delta g + f \text{ in } \Omega, \tag{4.15}$$
$$u = 0 \text{ on } \partial\Omega. \tag{4.16}$$

Consider the cost function

$$J(\Omega) = \int_\Omega |v - u_d|^2 \, \mathrm{d}x = \int_\Omega |u + g - u_d|^2 \, \mathrm{d}x. \tag{4.17}$$

The weak formulation of (4.15), (4.16) reads:

$$\text{Find } u \in H_0^1(\Omega) : \int_\Omega \nabla u \cdot \nabla \psi \, \mathrm{d}x = \int_\Omega -\nabla g \cdot \nabla \psi + f\psi \, \mathrm{d}x \quad \text{for all } \psi \in H_0^1(\Omega). \tag{4.18}$$

Note that the previous weak formulation is already well-defined for arbitrary open and bounded set Ω. We do not need to impose any regularity on Ω. The perturbed problem of the previous equation, which is obtained by considering (4.18) on $\Phi_t(\Omega)$ and performing a change of variables, reads: find $u^t \in H_0^1(\Omega)$ such that

$$\int_\Omega A(t)\nabla u^t \cdot \nabla \psi \, \mathrm{d}x = \int_\Omega -A(t)\nabla g^t \cdot \nabla \psi + \xi(t)f^t\psi \, \mathrm{d}x \quad \text{for all } \psi \in H_0^1(\Omega), \tag{4.19}$$

where $\xi(t) := \det(D\Phi_t)$ and $A(t) := \xi(t)D\Phi_t^{-1}D\Phi_t^{-T}$. The following continuity result is standard:

Lemma 4.8. *There exists a constant $c > 0$ such that $\|u^t - u^0\|_{H_0^1(\Omega)} \leq ct$ for all $t \in [0, \tau]$.*

Introduce

$$a(t, \varphi, \psi) := \int_\Omega A(t)\nabla\varphi \cdot \nabla\psi \, \mathrm{d}x + \int_\Omega A(t)\nabla g^t \cdot \nabla\psi - \xi(t)f^t\psi \, \mathrm{d}x$$

$$b(t, \varphi) := \int_\Omega \xi(t)|\varphi + g^t - u_d^t|^2 \, \mathrm{d}x.$$

where $g^t := g \circ \Phi_t$ and $u_d^t := u_d \circ \Phi_t$.

Recall that the associated Lagrangian (3.1) is $G(t, \varphi, \psi) = a(t, \varphi, \psi) + b(t, \varphi)$. The averaged adjoint equation (3.3) reads

$$\int_\Omega A(t)\nabla\varphi \cdot \nabla p^t \, \mathrm{d}x = \int_\Omega (u^t + u^0 + 2g^t - 2u_d^t)\varphi \, \mathrm{d}x \quad \text{for all } \varphi \in H_0^1(\Omega).$$

The following continuity result for the adjoint is standard:

Lemma 4.9. *There exists a constant $c > 0$ such that*

$$\|p^t - p^0\|_{H^1(\Omega)} \leq ct \quad \text{for all } t \in [0, \tau].$$

One readily verifies that all assumptions of Theorem 3.1 are satisfied, except for (H1) which we now prove. Indeed using $p^t \to p^0$ in $H_0^1(\Omega)$ as t goes to zero and the strong differentiability of $t \mapsto A(t)$ and $t \mapsto \xi(t)$, we get

$$
\begin{aligned}
\lim_{t \searrow 0} \frac{G(t, u^0, p^t) - G(0, u^0, p^t)}{t} &= \lim_{t \searrow 0} \left(\int_\Omega \left(\frac{A(t) - I}{t} \right) \nabla u^0 \cdot \nabla p^t \, \mathrm{d}x + \int_\Omega \left(\frac{A(t)\nabla g^t - \nabla g}{t} \right) \cdot \nabla p^t \right. \\
&\quad \left. - \left(\frac{\xi(t)f^t - f}{t} \right) p^t \, \mathrm{d}x + \int_\Omega \frac{\xi(t)|u^0 + g^t - u_d^t|^2 - |u^0 + g - u_d|^2}{t} \mathrm{d}x \right) \\
&= \partial_t G(0, u^0, p^0),
\end{aligned}
$$

which shows that (H1) is satisfied.

Hence, applying Theorem 3.1 yields

$$\mathrm{d}J(\Omega)(\theta) = \partial_t a(0, u, p) + \partial_t b(0, u),$$

which is by definition equivalent to

$$
\begin{aligned}
\mathrm{d}J(\Omega)(\theta) = \int_\Omega A'(0)(\nabla u \cdot \nabla p + \nabla g \cdot \nabla p) \, \mathrm{d}x + \int_\Omega \nabla(\nabla g \cdot \theta) \cdot \nabla p - \mathrm{div}(f\theta)p \, \mathrm{d}x \\
+ \int_\Omega \mathrm{div}(\theta)|u + g - u_d|^2 + (\nabla(g - u_d) \cdot \theta))(u + g - u_d) \, \mathrm{d}x. \quad (4.20)
\end{aligned}
$$

Since $A'(0) = (\mathrm{div}\,\theta)I - D\theta^T - D\theta$ we obtain the tensor representation (4.2) with:

$$
\begin{aligned}
\mathbf{S}_1 &= I(\nabla u \cdot \nabla p + \nabla g \cdot \nabla p - fp + |u + g - u_d|^2) - \nabla u \otimes \nabla p - \nabla p \otimes \nabla u - \nabla p \otimes \nabla g, \\
\mathbf{S}_0 &= D^2 g \nabla p - p\nabla f + (u + g - u_d)\nabla(g - u_d), \\
\mathfrak{S}_1 &= 0, \quad \mathfrak{S}_0 = 0.
\end{aligned}
$$

Now applying (4.10) we get immediately

$$g_1 = \nabla u \cdot \nabla p + \nabla g \cdot \nabla p - fp + |u + g - u_d|^2 - 2\partial_n u \partial_n p - \partial_n p \partial_n g.$$

Using the definition of the tangential gradient and $p = 0, u = 0$ on Γ implies $\nabla_\Gamma u = \nabla_\Gamma p = 0$, so we obtain the simpler expression

$$g_1 = -\partial_n u \partial_n p + |u + g - u_d|^2 = -\partial_n(v - g)\partial_n p + |v - u_d|^2.$$

Finally, substituting back $u = v - g$ we obtain the formula

$$\mathrm{d}J(\Omega)(\theta) = \int_{\partial\Omega} (-\partial_n(v - g)\partial_n p + |v - u_d|^2)\, \theta \cdot n \, \mathrm{d}s.$$

This formula can be found for instance in ([13], p. 566, Formula 6.38), where the adjoint has the sign opposite to our case.

Elliptic problem: first order tensor representation

Suppose that $\Omega \subset D \subset \mathbf{R}^d$ is a smooth bounded domain, where $D \subset \mathbf{R}^d$ is the smooth "universe". Let us consider the Dirichlet problem:

$$-\operatorname{div}(M\nabla u) + u = f \text{ in } \Omega,$$
$$u = 0 \text{ on } \partial\Omega, \tag{4.21}$$

where $M \in \mathbf{R}^{d,d}$ is a positive definite matrix. Consider the cost function

$$J(\Omega) = \int_\Omega |u - u_d|^2 \, \mathrm{d}x, \tag{4.22}$$

where $u_d \in C^1(\mathbf{R}^d)$. Let us introduce

$$a(t, \varphi, \psi) := \int_\Omega (MQ^t\nabla\varphi \cdot Q^t\nabla\psi + \varphi\psi)\xi(t)\,\mathrm{d}x - \int_\Omega \xi(t)f^t\psi \,\mathrm{d}x$$
$$b(t, \varphi) := \int_\Omega \xi(t)|\varphi - u_d^t|^2 \,\mathrm{d}x,$$

where $Q^t := D\Phi_t^{-T}$ and $\xi(t) := \det(D\Phi_t)$. Then the weak formulation of (4.21) on the perturbed domain Ω_t, once transported back to Ω is

$$a(t, u^t, \psi) = 0 \quad \text{for all } \psi \in H_0^1(\Omega).$$

The Lagrangian corresponding to the minimization of $J(\Omega)$ and the PDE constraint (4.21) is

$$G(t, \varphi, \psi) = b(t, \varphi) + a(t, \varphi, \psi). \tag{4.23}$$

It can be shown using Theorem 3.1 that $\mathrm{d}J(\Omega)(\theta) = \partial_t G(0, u, p)$, where $p \in H_0^1(\Omega)$ denotes the adjoint state:

$$\int_\Omega M\nabla\psi \cdot \nabla p + p\psi \,\mathrm{d}x = -\int_\Omega 2(u - u_d)\psi \,\mathrm{d}x \quad \text{for all } \psi \in H_0^1(\Omega). \tag{4.24}$$

The tensor representation (4.1) of the shape derivative of $J(\Omega)$ in direction $\theta \in C_c^2(D, \mathbf{R}^d)$ is given by

$$\mathrm{d}J(\Omega)(\theta) = \int_\Omega \mathbf{S}_1 \cdot D\theta + \mathbf{S}_0 \cdot \theta \,\mathrm{d}x. \tag{4.25}$$

where we use the relation $(\nabla p \otimes \nabla u) \cdot D\theta = D\theta\nabla u \cdot \nabla p$ to get the tensors

$$\mathbf{S}_0 = -2(u - u_d)\nabla u_d - p\nabla f, \tag{4.26}$$
$$\mathbf{S}_1 = -\nabla p \otimes M\nabla u - \nabla u \otimes M^T\nabla p + (M\nabla u \cdot \nabla p + up - fp + (u - u_d)^2)I. \tag{4.27}$$

In the simple case where $M = I$, assuming $u, p \in C^2(\overline{\Omega})$, we know from the previous discussion that (4.3) is satisfied. Noting that

$$\operatorname{div}(\nabla p \otimes \nabla u) = \Delta u\nabla p + (D^2p)^T\nabla u,$$
$$\operatorname{div}(\nabla u \otimes \nabla p) = \Delta p\nabla u + (D^2u)^T\nabla p,$$
$$\nabla(\nabla u \cdot \nabla p) = D^2u\nabla p + D^2p\nabla u,$$

the relation

$$-\operatorname{div}(\mathbf{S}_1) + \mathbf{S}_0 = 0 \quad \text{in } \Omega \tag{4.28}$$

is equivalent to

$$(-\Delta u + u - f)\nabla p + (-\Delta p + p + 2(u - u_d))\nabla u = 0 \quad \text{in } \Omega.$$

Therefore, we observe that the fundamental relation (4.3) between the tensors $\mathbf{S}_1$ and $\mathbf{S}_0$ corresponds to the strong solvability of the state and adjoint state equation.

5. Descent directions

In this paper we are interested in numerical methods for shape optimization problems of the type

$$\min_{\Omega \in \mathfrak{P}} J(\Omega), \tag{5.1}$$

where $\mathfrak{P} \subset \mathcal{P}(D)$ is the admissible set. Assume $J : \mathfrak{P} \to \mathbf{R}$ is shape differentiable at $\Omega \subset D \subset \mathbf{R}^d$.

Definition 5.1 (Descent direction)**.** The vector field $\theta \in C_c^{0,1}(D, \mathbf{R}^d)$ is called a *descent direction* for J at Ω if there exists an ε such that

$$J(\Phi_t^\theta(\Omega)) < J(\Omega) \text{ for all } t \in (0, \varepsilon).$$

If the Eulerian semiderivative of J at Ω in direction θ exists and if it is a descent direction then by definition

$$\mathrm{d}J(\Omega)(\theta) < 0. \tag{5.2}$$

Descent directions are used in iterative methods for finding approximate (possibly local) minimizers of $J(\Omega)$. Typically, at a given starting point Ω, one determines a descent direction θ and proceeds along this direction as long as the cost functional J reduces sufficiently using a step size strategy. In this section we give a general setting for computing descent directions in the framework of gradient methods using the domain and boundary representations of the shape derivative according to Theorem 2.4. We show how a descent direction θ with any regularity H^s, $s \geq 1$ can be obtained by solving an appropriate partial differential equation. We also show how to deal with bound constraints on θ. In order to develop a setting allowing to define general descent directions, we recall sufficient conditions for the solvability of the following operator equation

$$A\theta = f,$$

where $A : \mathbf{E} \to \mathbf{E}'$ is an operator between a Banach space $\mathbf{E}$ and its dual $\mathbf{E}'$. Sufficient conditions for the bijectivity of A are given by the theorem of Minty–Browder ([39], p. 364, Thm. 10.49).

Theorem 5.2 (Minty–Browder)**.** *Let $(\boldsymbol{E}; \|\cdot\|_{\boldsymbol{E}})$ be a reflexive separable Banach space and $A : \boldsymbol{E} \to \boldsymbol{E}'$ a bounded, hemi-continuous, monotone and coercive operator. Then A is surjective,* i.e. *for each $f \in \boldsymbol{E}'$ there exists $\theta \in \boldsymbol{E}$ such that $A\theta = f$. Moreover if A is strictly monotone then it is bijective.*

Let $A : \mathbf{E} \to \mathbf{E}'$ be an operator on a reflexive, separable Banach space $\mathbf{E}$ satisfying the assumptions of Theorem 5.2 with $A(0)\theta \geq 0$ for $v \in \mathbf{E}$. Assume $\mathrm{d}J(\Omega)$ can be extended to $\mathbf{E}'$ if necessary; for simplicity we keep the same notation for the extension. Introduce the bilinear form

$$\mathcal{B} : \mathbf{E} \times \mathbf{E} \to \mathbf{R}, \qquad \mathcal{B}(\theta, \zeta) := \langle A\theta, \zeta \rangle_{\mathbf{E}', \mathbf{E}}. \tag{5.3}$$

Consider the variational problem:

$$(\text{VP}) \qquad \text{Find } \theta_1 \in \mathbf{E} \text{ such that } \mathcal{B}(\theta_1, \zeta) = -\mathrm{d}J(\Omega)(\zeta) \text{ for all } \zeta \in \mathbf{E}, \tag{5.4}$$

Then the solution θ_1 of (VP) is a descent direction since $\mathrm{d}J(\Omega)(\theta_1) = -\mathcal{B}(\theta_1, \theta_1) \leq 0$.

In certain situations it is desirable to have bound constraints on the shape perturbation. This may be handled by considering the more general case of a variational inequality. Given a subset $K \subset \mathbf{E}$ with $0 \in K$, consider the variational inequality:

$$(\text{VI}) \qquad \text{Find } \theta_2 \in K \text{ such that } \mathcal{B}(\theta_2, \theta_2 - \zeta) \leq \mathrm{d}J(\Omega)(\zeta - \theta_2) \text{ for all } \zeta \in K.$$

The solution θ_2 of (VI) yields a descent direction for J at Ω since taking $\zeta = 0 \in K$ we get

$$\mathrm{d}J(\Omega)(\theta_2) \leq -\mathcal{B}(\theta_2, \theta_2) \leq 0.$$

In view of Theorem 2.4, we choose $\mathbf{E} \subset H^s(D)$ where s is such that $\mathrm{d}J(\Omega) : H^s(D, \mathbf{R}^d) \to \mathbf{R}^d$ is continuous. When $\mathbf{E}$ is a Hilbert space, one may identify $\mathbf{E}'$ with $\mathbf{E}$. Therefore if $\mathcal{B}$ is bilinear, coercive, and continuous, then Lax Milgram's lemma ensures that (VP) has a unique solution. For all other cases we may have to use Theorem 5.2 or similar results.

Remark 5.3.

(a) Let $\mathbf{E} := H_0^1(D, \mathbf{R}^d)$, $\mathcal{B}(\theta, \zeta) := \int_D D\theta : D\zeta \, \mathrm{d}x$ and $\Omega^+ \Subset D$. Then (5.4) reads: find $\theta \in H_0^1(D, \mathbf{R}^d)$, such that $\mathcal{B}(\theta, \zeta) = -\mathrm{d}J(\Omega^+)(\zeta)$ for all $\zeta \in H_0^1(D, \mathbf{R}^d)$. Under the assumption that $\partial\Omega^+ \in C^2$, $\theta|_{\Omega^+} \in H^2(\Omega^+)$, and $\theta|_{D \setminus \overline{\Omega^+}} \in H^2(D \setminus \overline{\Omega^+})$, Proposition 4.3 yields

$$\int_{\partial\Omega^+} g \, \zeta \cdot n \, \mathrm{d}s = \mathrm{d}J(\Omega^+)(\zeta) \quad \text{for all } \zeta \in H_0^1(D, \mathbf{R}^d), \quad \text{where} \quad g = -(D\theta^+ n - D\theta^- n) \cdot n.$$

This shows that the restriction to $\partial\Omega^+$ of the obtained descent direction θ is more regular than the function g.

(b) Let n be an extension of the unit normal to Ω^+ in D. If θ defined on D is a descent direction then $(\theta \cdot n)n|_{\partial\Omega^+}$ is also a descent direction, for the tangential part of θ does not influence the derivative. Indeed define $\theta_\tau := \theta - (\theta \cdot n)n$, then by Nagumo's theorem $\mathrm{d}J(\Omega^+)(\theta_\tau) = 0$ and thus $\mathrm{d}J(\Omega^+)(\theta) = \mathrm{d}J(\Omega^+)((\theta \cdot n)n)$. However, θ and $(\theta \cdot n)n$ lead to different transformations of the domains in general, indeed the tangential term actually has an influence for large deformations, which means $\Phi_t^\theta(\Omega^+) \neq \Phi_t^{(\theta \cdot n)n}(\Omega^+)$. This influence appears for instance in the shape Hessian.

6. ELECTRICAL IMPEDANCE TOMOGRAPHY

We consider an application of the results above to a typical and important interface problem: the inverse problem of electrical impedance tomography (EIT) also known as the inverse conductivity or Calderón's problem [6] in the mathematical literature. It is an active field of research with an extensive literature; for further details we point the reader toward the survey papers [5, 10] as well as [31] and the references therein. We consider the particular case where the objective is to reconstruct a piecewise constant conductivity σ which amounts to determine an interface Γ^+ between some inclusions and the background. We refer the reader to [2, 7, 8, 11, 22, 25, 29] for more details on this approach.

The main interest of studying EIT is to apply the approach developed in this paper to a problem which epitomizes general interface problems and simultaneously covers the entire spectrum of difficulties encountered with severely ill-posed inverse problem.

6.1. Problem statement

Let $D \subset \mathbf{R}^d$ be a Lipschitz domain, and $\Omega^+, \Omega^- \subset D$ open sets such that $D = \Omega^+ \cup \Omega^- \cup \Gamma^+$, where $\Gamma^+ = \partial\Omega^+ = \overline{\Omega^+} \cap \overline{\Omega^-}$ and $\Gamma = \partial D = \partial\Omega^- \setminus \Gamma^+$; see Figure 1. In this section n denotes either the outward unit normal vector to D or the outward unit normal vector to Ω^+. Decompose Γ as $\Gamma = \Gamma_d \cup \Gamma_n$. Let $\sigma = \sigma^+ \chi_{\Omega^+} + \sigma^- \chi_{\Omega^-}$ where $\sigma^\pm$ are scalars and $f = f^+ \chi_{\Omega^+} + f^- \chi_{\Omega^-}$ where $f^\pm \in H^1(D)$. Consider the following problems: find $u_n \in H_d^1(D)$ such that

$$\int_D \sigma \nabla u_n \cdot \nabla z = \int_D fz + \int_{\Gamma_n} gz \quad \text{for all } z \in H_d^1(D) \tag{6.1}$$

and find $u_d \in H_{dn}^1(D)$ such that

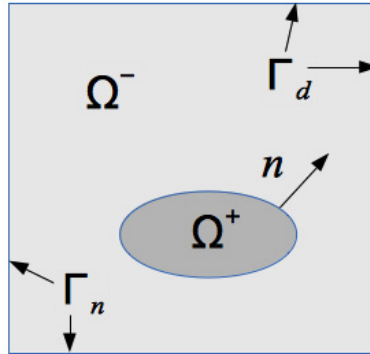$$\int_D \sigma \nabla u_d \cdot \nabla z = \int_D fz \quad \text{for all } z \in H_0^1(D) \tag{6.2}$$

FIGURE 1. Partition $D = \Omega^+ \cup \Omega^- \cup \Gamma$.

where

$$H^1_d(D) := \{v \in H^1(D) \mid v = 0 \text{ on } \Gamma_d\},$$
$$H^1_{dn}(D) := \{v \in H^1(D) \mid v = 0 \text{ on } \Gamma_d, v = h \text{ on } \Gamma_n\},$$
$$H^1_0(D) := \{v \in H^1(D) \mid v = 0 \text{ on } \Gamma\}$$

and $g \in H^{-1/2}(\Gamma_n)$ represents the input, in this case the electric current applied on the boundary and $h \in H^{1/2}(\Gamma_n)$ is the measurement of the potential on Γ_n, or the other way around, *i.e.* h can be the input and g the measurement. Define also the space

$$PH^k(D) := \{u = u^+ \chi_{\Omega^+} + u^- \chi_{\Omega^-} \mid u^+ \in H^k(\Omega^+), \ u^- \in H^k(\Omega^-)\}.$$

Consider the following assumption which will be used only to derive the boundary expression of the shape derivative but is not required for the domain expression:

Assumption 6.1. The domains D, Ω^+, Ω^- are of class C^k, $f \in PH^{\max(k-2,1)}(D)$, $g \in H^{k-\frac{3}{2}}(D)$ and $h \in H^{k-\frac{1}{2}}(D)$ for $k \geq 2$.

Applying Green's formula under Assumption 6.1, equations (6.1) and (6.2) are equivalent to the following transmission problems where $u_n = u_n^+ \chi_{\Omega^+} + u_n^- \chi_{\Omega^-}$ and $u_d = u_d^+ \chi_{\Omega^+} + u_d^- \chi_{\Omega^-}$:

$$-\sigma^+ \Delta u_n^+ = f \text{ in } \Omega^+, \quad -\sigma^- \Delta u_n^- = f \text{ in } \Omega^-, \tag{6.3}$$
$$u_n^- = 0 \text{ on } \Gamma_d, \tag{6.4}$$
$$\sigma^- \partial_n u_n^- = g \text{ on } \Gamma_n, \tag{6.5}$$

$$-\sigma^+ \Delta u_d^+ = f \text{ in } \Omega^+, \quad -\sigma^- \Delta u_d^- = f \text{ in } \Omega^-, \tag{6.6}$$
$$u_d^- = 0 \text{ on } \Gamma_d, \tag{6.7}$$
$$u_d^- = h \text{ on } \Gamma_n, \tag{6.8}$$

with the transmission conditions

$$\sigma^+ \partial_n u_n^+ = \sigma^- \partial_n u_n^-, \quad \sigma^+ \partial_n u_d^+ = \sigma^- \partial_n u_d^- \quad \text{on } \Gamma^+,$$
$$u_n^+ = u_n^-, \quad u_d^+ = u_d^- \quad \text{on } \Gamma^+. \tag{6.9}$$

On Γ_d we impose homogeneous Dirichlet conditions, meaning that the voltage is fixed and no measurement is performed. One may take $\Gamma_d = \emptyset$, in which case (6.1) becomes a pure Neumann problem and additional

care must be taken for the uniqueness and existence of a solution. The situation $\Gamma_d \neq \emptyset$ corresponds to partial measurements. Alternatively, it is also possible to consider a slightly different problem where each function u_n and u_d has both the boundary conditions (6.5) and (6.8) on different parts of the boundary.

Several measurements can be made by choosing sets of functions $\{g_i\}_{i=1}^I$ and $\{h_i\}_{i=1}^I$. Writing $u_{n,i}$ and $u_{d,i}$ for the corresponding states, the problem of electrical impedance tomography is:

$$(\text{EIT}): \text{Given } \{g_i\}_{i=1}^I \text{ and } \{h_i\}_{i=1}^I, \text{ find } \sigma \text{ such that } u_{n,i} = u_{d,i} \text{ in } D \text{ for } i = 1, \ldots, I. \tag{6.10}$$

Note that $u_{n,i} = u_{n,i}(\Omega^+)$ and $u_{d,i} = u_{d,i}(\Omega^+)$ actually depend on Ω^+ through $\sigma = \sigma(\Omega^+)$, however we often write $u_{n,i}$ and $u_{d,i}$ for simplicity. In this section, we assume that the conductivities (σ^+, σ^-) are known, therefore the EIT problem (6.10) reduces to the following shape optimization problem where Ω^+ is the unknown.

$$\text{Given } \{g_i\}_{i=1}^I, \{h_i\}_{i=1}^I \text{ and } (\sigma^+, \sigma^-) \text{ with } \sigma = \sigma^+ \chi_{\Omega^+} + \sigma^- \chi_{\Omega^-},$$
$$\text{find } \Omega^+ \text{ such that } u_{n,i} = u_{d,i} \text{ in } D \text{ for } i = 1, \ldots, I. \tag{6.11}$$

Actually, the result for several measurements can be straightforwardly deduced from the case of one measurement by summing the cost functionals corresponding to each measurement, therefore in this section we stick to the case $I = 1$ of one measurement g for simplicity of presentation. In section 8 we consider several measurements for the numerics.

The notion of well-posedness due to Hadamard requires the existence and uniqueness of a solution and the continuity of the inverse mapping. The severe ill-posedness of EIT is well-known: uniqueness and continuity of the inverse mapping depend on the regularity of σ, the latter being responsible for the instability of the reconstruction process. Additionally, partial measurements often encountered in practice render the inverse problem even more ill-posed. We refer to the reviews [5, 10, 31] and the references therein for more details. A standard cure against the ill-posedness is to regularize the inverse mapping. In this paper the regularization is achieved by considering smooth perturbations of the domains Ω^+.

To solve problem (6.11), we use an optimization approach by minimizing the shape functionals

$$J_1(\Omega^+) = \frac{1}{2} \int_D (u_d(\Omega^+) - u_n(\Omega^+))^2, \tag{6.12}$$

$$J_2(\Omega^+) = \frac{1}{2} \int_{\Gamma_n} (u_n(\Omega^+) - h)^2. \tag{6.13}$$

Since $u_d, u_n \in H^1(D)$ and $h \in H^{1/2}(\Gamma_n)$, J_1 and J_2 are well-defined. Note that J_1 and J_2 are redundant for the purpose of the reconstruction but our aim is to provide an efficient way of computing the shape derivative of two functions which are often encountered in the literature. To compute these derivatives we use the approach described in Section 3. First of all introduce

$$F_1(\varphi_d, \varphi_n) := \frac{1}{2} \int_D (\varphi_d - \varphi_n)^2, \tag{6.14}$$

$$F_2(\varphi_n) := \frac{1}{2} \int_{\Gamma_n} (\varphi_n - h)^2. \tag{6.15}$$

Note that $J_1(\Omega^+) = F_1(u_d(\Omega^+), u_n(\Omega^+))$ and $J_2(\Omega^+) = F_2(u_n(\Omega^+))$. Next consider $\mathfrak{P}$ a subset of $\mathcal{P}(D)$ and the Lagrangian $\mathcal{L} : \mathfrak{P} \times H_d^1(D) \times H_d^1(D) \times H_0^1(D) \times H_d^1(D) \to \mathbf{R}$:

$$\mathcal{L}(\Omega^+, \boldsymbol{\varphi}, \boldsymbol{\psi}) := \alpha_1 F_1(\varphi_d, \varphi_n) + \alpha_2 F_2(\varphi_n)$$
$$+ \int_D \sigma \nabla \varphi_d \cdot \nabla \psi_d - f \psi_d + \int_{\Gamma_n} -\sigma^- \partial_n \psi_d (\varphi_d - h)$$
$$+ \int_D \sigma \nabla \varphi_n \cdot \nabla \psi_n - f \psi_n - \int_{\Gamma_n} g \psi_n, \tag{6.16}$$

where $\boldsymbol{\varphi} := (\varphi_d, \varphi_n)$ and $\boldsymbol{\psi} := (\psi_d, \psi_n)$. The term $-\sigma^- \partial_n \psi_d$ in the second integral of (6.16) is used to enforce the boundary condition (6.8). Introduce the objective functional

$$J(\Omega^+) := \alpha_1 J_1(\Omega^+) + \alpha_2 J_2(\Omega^+).$$

To compute the derivative of the Lagrangian depending on (6.3)–(6.8) we apply the averaged adjoint method from Section 3.

6.2. State and adjoint equations

The state $\mathbf{u} := (u_d, u_n)$ and adjoint state $\boldsymbol{p} := (p_d, p_n)$ are solutions of the equations:

$$\partial_{\boldsymbol{\psi}} \mathcal{L}(\Omega^+, \mathbf{u}, \boldsymbol{p})(\hat{\boldsymbol{\psi}}) = 0 \text{ for all } \hat{\boldsymbol{\psi}} \in H_0^1(D) \times H_d^1(D), \tag{6.17}$$

$$\partial_{\boldsymbol{\varphi}} \mathcal{L}(\Omega^+, \mathbf{u}, \boldsymbol{p})(\hat{\boldsymbol{\varphi}}) = 0 \text{ for all } \hat{\boldsymbol{\varphi}} \in H_d^1(D) \times H_d^1(D). \tag{6.18}$$

Writing (6.17) explicitly, one can obtain easily the state equations (6.1) and (6.2). Then (6.18) yields the equation for the adjoint p_d:

$$\partial_{\varphi_d} \mathcal{L}(\Omega^+, \mathbf{u}, \boldsymbol{p})(\hat{\varphi}_d) = 0, \text{ for all } \hat{\varphi}_d \in H_d^1(D),$$

which leads to

$$\int_D \sigma \nabla p_d \cdot \nabla \hat{\varphi}_d \, \mathrm{d}x = -\alpha_1 \int_D (u_d - u_n) \hat{\varphi}_d \, \mathrm{d}x - \int_{\Gamma_n} -\sigma^- \partial_n p_d \hat{\varphi}_d \, \mathrm{d}s \quad \text{for all } \hat{\varphi}_d \in H_d^1(D) \tag{6.19}$$

which is the variational formulation for the adjoint state p_d. This yields the following variational formulation when test functions are restricted to $H_0^1(D)$:

$$\int_D \sigma \nabla p_d \cdot \nabla \widetilde{\varphi} \, \mathrm{d}x = -\alpha_1 \int_D (u_d - u_n) \widetilde{\varphi} \, \mathrm{d}x \quad \text{for all } \widetilde{\varphi} \in H_0^1(D). \tag{6.20}$$

If we use Assumption 6.1, we get $p_d \in PH^k(D)$ and using Green's formula in Ω^+ and Ω^- with $\widetilde{\varphi} \in C_c^\infty(\Omega^+)$ and $\widetilde{\varphi} \in C_c^\infty(\Omega^-)$, respectively, we obtain the strong form

$$-\operatorname{div}(\sigma \nabla p_d) = -\alpha_1 (u_d - u_n) \text{ in } \Omega^+ \text{ and } \Omega^-. \tag{6.21}$$

Hence using now Green's formula in (6.19) and using (6.21) gives

$$\int_{\Gamma^+} [\sigma \partial_n p_d]_{\Gamma^+} \hat{\varphi}_d \, \mathrm{d}s + \int_{\Gamma_n} (\sigma \partial_n p_d - \sigma^- \partial_n p_d) \hat{\varphi}_d \, \mathrm{d}s = 0 \quad \text{for all } \hat{\varphi}_d \in H_d^1(D),$$

where $[\sigma \partial_n p_d]_{\Gamma^+} = \sigma^+ \partial_n p_d^+ - \sigma^- \partial_n p_d^-$ is the jump of $\sigma \partial_n p_d$ across Γ^+. Since the integral on Γ_n above vanishes and $p_d \in H_0^1(D)$, we obtain

$$p_d = 0 \text{ on } \Gamma, \tag{6.22}$$

$$\sigma^+ \partial_n p_d^+ = \sigma^- \partial_n p_d^- \text{ on } \Gamma^+. \tag{6.23}$$

Finally solving

$$\partial_{\varphi_n} \mathcal{L}(\Omega^+, \mathbf{u}, \boldsymbol{p})(\hat{\varphi}_n) = 0, \text{ for all } \hat{\varphi}_n \in H_d^1(D),$$

leads to the variational formulation

$$\int_D -\alpha_1 (u_d - u_n) \hat{\varphi}_n + \sigma \nabla p_n \cdot \nabla \hat{\varphi}_n + \int_{\Gamma_n} \alpha_2 (u_n - h) \hat{\varphi}_n = 0 \tag{6.24}$$

for all $\hat{\varphi}_n \in H_d^1(D)$.

Similarly as for p_d we get, under Assumption 6.1, $p_n \in PH^k(D)$ and the strong form

$$-\operatorname{div}(\sigma \nabla p_n) = \alpha_1(u_d - u_n) \text{ in } \Omega^+ \text{ and } \Omega^-, \tag{6.25}$$

$$\sigma \partial_n p_n = -\alpha_2(u_n - h) \text{ on } \Gamma_n, \tag{6.26}$$

$$p_n = 0 \text{ on } \Gamma_d, \tag{6.27}$$

$$\sigma^+ \partial_n p_n^+ = \sigma^- \partial_n p_n^- \text{ on } \Gamma^+, \quad p_n^+ = p_n^- \text{ on } \Gamma^+. \tag{6.28}$$

6.3. Shape derivatives

Let us consider a transformation Φ_t^θ defined by (2.3) with $\theta \in C_c^1(D, \mathbf{R}^d)$. Note that $\Phi_t^\theta(D) = D$ but in general $\Phi_t^\theta(\Omega^+) \neq \Omega^+$. We use the notation $\Omega^+(t) := \Phi_t^\theta(\Omega^+)$. Our aim is to show the shape differentiability of $J(\Omega^+)$ with the help of Theorem 3.1. Following the methodology described in Section 3 we introduce

$$G(t, \varphi, \psi) := \mathcal{L}(\Omega^+(t), \varphi \circ \Phi_t^{-1}, \psi \circ \Phi_t^{-1}). \tag{6.29}$$

We proceed to the change of variables $\Phi_t(x) = y$ in (6.29) to get the canonical form (3.1). First of all let us denote $f_{\Omega^+(t)} = f^+ \chi_{\Omega^+(t)} + f^- \chi_{\Omega^-(t)}$ and $\sigma_{\Omega^+(t)} = \sigma^+ \chi_{\Omega^+(t)} + \sigma^- \chi_{\Omega^-(t)}$; recall that $\sigma^\pm$ are scalars but $f^\pm$ are functions. Then note that the change of variables $\Phi_t(x) = y$ leads to considering the following functions inside the integrals:

$$\sigma_{\Omega^+(t)} \circ \Phi_t = \sigma^+ \chi_{\Omega^+(t)} \circ \Phi_t + \sigma^- \chi_{\Omega^-(t)} \circ \Phi_t = \sigma^+ \chi_{\Omega^+} + \sigma^- \chi_{\Omega^-} = \sigma,$$

$$f_{\Omega^+(t)} \circ \Phi_t = f^+ \circ \Phi_t \, \chi_{\Omega^+(t)} \circ \Phi_t + f^- \circ \Phi_t \, \chi_{\Omega^-(t)} \circ \Phi_t = f^+ \circ \Phi_t \, \chi_{\Omega^+} + f^- \circ \Phi_t \, \chi_{\Omega^-}.$$

Thus we introduce the function $\tilde{f}_t := f^+ \circ \Phi_t \, \chi_{\Omega^+} + f^- \circ \Phi_t \, \chi_{\Omega^-}$. Now we obtain the canonical form (3.1) for the Lagrangian:

$$G(t, \varphi, \psi) = a(t, \varphi, \psi) + b(t, \varphi), \tag{6.30}$$

with

$$a(t, \varphi, \psi) := \int_D \sigma A(t) \nabla \varphi_d \cdot \nabla \psi_d - \tilde{f}_t \psi_d \xi(t) - \int_{\Gamma_n} \sigma^- \partial_n \psi_d (\varphi_d - h)$$

$$+ \int_D \sigma A(t) \nabla \varphi_n \cdot \nabla \psi_n - \tilde{f}_t \psi_n \xi(t) - \int_{\Gamma_n} g \psi_n,$$

$$b(t, \varphi) := \frac{\alpha_1}{2} \int_D (\varphi_d - \varphi_n)^2 \xi(t) + \frac{\alpha_2}{2} \int_{\Gamma_n} (\varphi_n - h)^2$$

where the Jacobian $\xi(t)$ and $A(t)$ are defined as $\xi(t) := \det(D\Phi_t)$ and $A(t) := \xi(t) D\Phi_t^{-1} D\Phi_t^{-T}$. In the previous expression (6.30), one should note that the integrals on subsets of Γ are unchanged since $\Phi_t^{-1} = I$ on Γ. Thus we have $\Phi_t^\theta(D) = D$, however the terms inside the integrals on D are modified by the change of variable since $\Phi_t^{-1} \neq I$ inside D. Note that

$$J(\Omega^+(t)) = G(t, \mathbf{u}^t, \psi), \text{ for all } \psi \in H_0^1(D) \times H_d^1(D),$$

where $\mathbf{u}^t = (u_n^t, u_d^t) := (u_{n,t} \circ \Phi_t, u_{d,t} \circ \Phi_t)$ and $u_{n,t}, u_{d,t}$ solve (6.1),(6.2), respectively, with the domains Ω^+ and Ω^- replaced by $\Omega^+(t)$ and $\Omega^-(t)$. As one can verify by applying a change of variables to (6.1) and (6.2) on the domain $\Omega^+(t)$ the functions u_n^t, u_d^t satisfy

$$\int_D \sigma A(t) \nabla u_n^t \cdot \nabla \hat{\psi}_n = \int_D \tilde{f}_t \hat{\psi}_n + \int_{\Gamma_n} g \hat{\psi}_n \text{ for all } \hat{\psi}_n \in H_d^1(D), \tag{6.31}$$

$$\int_D \sigma A(t) \nabla u_d^t \cdot \nabla \hat{\psi}_d = \int_D \tilde{f}_t \hat{\psi}_d \text{ for all } \hat{\psi}_d \in H_0^1(D). \tag{6.32}$$

Applying standards estimates for elliptic partial differential equations and the fact that $A(t)$ is uniformly bounded from below and above for t small enough, we infer from equations (6.31) and (6.32) the existence of constants $C_1, C_2 > 0$ independent of t and $\tau > 0$ such that for all $t \in [0, \tau]$:

$$\|u_d^t\|_{H^1(D)} \leq C_1, \quad \text{and} \quad \|u_n^t\|_{H^1(D)} \leq C_2. \tag{6.33}$$

From these estimates, we get $u_d^t \rightharpoonup w_d$ and $u_n^t \rightharpoonup w_n$ in $H^1(D)$ as $t \to 0$. Passing to the limit in (6.31) and (6.32) yields $w_d = u_d$ and $w_n = u_n$ by uniqueness.

Let us now check Assumption (H0) and the conditions of Theorem 3.1 for the function G given by (6.30) and the Banach spaces $E = H_d^1(D) \times H_d^1(D)$ and $F = H_0^1(D) \times H_d^1(D)$. First of all equation (3.2) admits a unique solution $\mathbf{u}^t := (u_n^t, u_d^t)$ for each $t \in [0, \tau]$. The conditions of Assumption (H0) are readily satisfied and also the function G is affine with respect to $\psi = (\psi_d, \psi_n)$.

Regarding the conditions of Theorem 3.1, first note that applying Lax–Milgram's lemma, we check that both equations (6.34) and (6.35) have indeed a unique solution in $F = H_0^1(D) \times H_d^1(D)$:

$$\int_D \sigma A(t) \nabla p_d^t \cdot \nabla \hat{\varphi}_d + \frac{\alpha_1}{2} \int_D \xi(t)(u_d^t + u_d - (u_n^t + u_n))\hat{\varphi}_d - \int_{\Gamma_n} \sigma^- \partial_n p_d^t \hat{\varphi}_d = 0, \tag{6.34}$$

$$\int_D \sigma A(t) \nabla p_n^t \cdot \nabla \hat{\varphi}_n - \frac{\alpha_1}{2} \int_D \xi(t)(u_d^t + u_d - (u_n^t + u_n))\hat{\varphi}_n + \frac{\alpha_2}{2} \int_{\Gamma_n} (u_n^t + u_n - 2h)\hat{\varphi}_n = 0, \tag{6.35}$$

for all $\hat{\varphi}_d, \hat{\varphi}_n$ in $H_d^1(D)$. Therefore there exists a unique solution $\mathbf{p}^t = (p_n^t, p_d^t)$ of the averaged adjoint equation (3.3).

Now we check Assumption (H1). Testing (6.34) with $\hat{\varphi}_d = p_d^t$ and (6.35) with $\hat{\varphi}_n = p_n^t$, we conclude by an application of Hölder's inequality together with (6.33) the existence of constants C_1, C_2 and $\tau > 0$ such that for all $t \in [0, \tau]$

$$\|p_d^t\|_{H^1(D)} \leq C_1, \quad \text{and} \quad \|p_n^t\|_{H^1(D)} \leq C_2.$$

We get that for each sequence t_k converging to zero, there exists a subsequence also denoted t_k such that $p_d^{t_k} \rightharpoonup q_d$ and $p_n^{t_k} \rightharpoonup q_n$ for two elements $q_d, q_n \in H^1(D)$. Passing to the limit in (6.34) and (6.35) yields $q_d = p_d$ and $q_n = p_n$ by uniqueness, where p_d and p_n are solutions of the adjoint equations. Since the limit is unique, we have in fact $p_d^t \rightharpoonup p_d$ and $p_n^t \rightharpoonup p_n$ as $t \to 0$. Finally, differentiating G with respect to t yields

$$\partial_t G(t, \boldsymbol{\varphi}, \boldsymbol{\psi}) = \frac{\alpha_1}{2} \int_D (\varphi_d - \varphi_n)^2 \xi(t) \, \text{tr}(D\theta_t D\Phi_t^{-1})$$

$$+ \int_D \sigma A'(t) \nabla \varphi_d \cdot \nabla \psi_d - \tilde{f}_t \psi_d \xi(t) \, \text{tr}(D\theta_t D\Phi_t^{-1}) - \psi_d \widetilde{\nabla} f_t \cdot \theta_t \xi(t)$$

$$+ \int_D \sigma A'(t) \nabla \varphi_n \cdot \nabla \psi_n - \tilde{f}_t \psi_n \xi(t) \, \text{tr}(D\theta_t D\Phi_t^{-1}) - \psi_n \widetilde{\nabla} f_t \cdot \theta_t \xi(t).$$

where

$$\widetilde{\nabla} f_t := \nabla f^+ \circ \Phi_t \chi_{\Omega^+} + \nabla f^- \circ \Phi_t \chi_{\Omega^-}, \qquad \theta_t = \theta \circ \Phi_t$$

$$A'(t) = \text{tr}(D\theta^t D\Phi_t^{-1}) A(t) - D\Phi_t^{-T} D\theta_t A(t) - (D\Phi_t^{-T} D\theta_t A(t))^T$$

and $D\theta_t$ is the Jacobian matrix of θ_t. In view of $\theta \in C_c^1(D, \mathbf{R}^d)$, the functions $t \mapsto D\theta_t$ and $t \mapsto \text{tr}(D\theta_t \Phi_t^{-1}) = \text{div}(\theta) \circ \Phi_t$ are continuous on $[0, T]$. Moreover $\varphi_d, \varphi_n, \psi_d, \psi_n$ are in $H^1(D)$, $f \in PH^1(D)$ so that $\partial_t G(t, \boldsymbol{\varphi}, \boldsymbol{\psi})$ is well-defined for all $t \in [0, T]$. Using the weak convergence of $\mathbf{p}^t$ and the strong differentiability of $t \mapsto A(t)$ and $t \mapsto \xi(t)$ it follows

$$\lim_{t \searrow 0} \frac{G(t, \mathbf{u}^0, \mathbf{p}^t) - G(0, \mathbf{u}^0, \mathbf{p}^t)}{t} = \partial_t G(0, \mathbf{u}^0, \mathbf{p}^0). \tag{6.36}$$

Thus we have verified all assumptions from Theorem 3.1. This yields

$$\mathrm{d}J(\Omega^+)(\theta) = \frac{\mathrm{d}}{\mathrm{d}t}\left(G(t, \mathbf{u}^t, \boldsymbol{\psi})\right)|_{t=0} = \partial_t G(0, \mathbf{u}^0, \boldsymbol{p}^0) \text{ for all } \boldsymbol{\psi} \in F = H_0^1(D) \times H_d^1(D),$$

and therefore we have proved the following result.

Proposition 6.2 (distributed shape derivative). *Let* $D \subset \mathbf{R}^d$ *be a Lipschitz domain,* $\theta \in C_c^1(D, \mathbf{R}^d)$, $f \in PH^1(D)$, $g \in H^{-1/2}(\Gamma_n)$, $h \in H^{1/2}(\Gamma_n)$, $\Omega^+ \subset D$ *is an open set, then the shape derivative of* $J(\Omega^+)$ *is given by*

$$\mathrm{d}J(\Omega^+)(\theta) = \int_D \left(\frac{\alpha_1}{2}(u_d - u_n)^2 - f(p_n + p_d)\right) \mathrm{div}\,\theta$$
$$+ \int_D -(p_d + p_n)\widetilde{\nabla}f \cdot \theta + \sigma A'(0)(\nabla u_d \cdot \nabla p_d + \nabla u_n \cdot \nabla p_n), \tag{6.37}$$

where $\widetilde{\nabla}f := \nabla f^+ \chi_{\Omega^+} + \nabla f^- \chi_{\Omega^-}$, $A'(0) = (\mathrm{div}\,\theta)I - D\theta^T - D\theta$, u_n, u_d *are solutions of* (6.1),(6.2) *and* p_n, p_d *of* (6.24), (6.19).

The shape derivative (6.37) *also has the tensor representation corresponding to* (4.1):

$$\mathrm{d}J(\Omega^+)(\theta) = \int_D \mathbf{S}_1 \cdot D\theta + \mathbf{S}_0 \cdot \theta, \tag{6.38}$$

where

$$\mathbf{S}_1 = -\sigma(\nabla u_d \otimes \nabla p_d + \nabla p_d \otimes \nabla u_d + \nabla u_n \otimes \nabla p_n + \nabla p_n \otimes \nabla u_n)$$
$$+ \sigma(\nabla u_d \cdot \nabla p_d + \nabla u_n \cdot \nabla p_n)I + \left(\frac{\alpha_1}{2}(u_d - u_n)^2 - f(p_n + p_d)\right)I,$$
$$\mathbf{S}_0 = -(p_d + p_n)\widetilde{\nabla}f.$$

Note that the volume expression of the shape gradient in Proposition 6.2 has been obtained without any regularity assumption on Ω^+. In order to obtain a boundary expression on the interface Γ^+ we need more regularity of Ω^+ provided by Assumption 6.1. If it is satisfied, we can apply Corollary 4.5 to obtain directly the boundary expression of the shape derivative, using mainly the standard tensor relation $(\nabla u_d \otimes \nabla p_d)n = (\nabla p_d \cdot n)\nabla u_d$, which yields Proposition 6.3.

Proposition 6.3 (Boundary expression). *Under Assumption* 6.1 *and* $\theta \in C_c^1(D, \mathbf{R}^d)$ *the shape derivative of* $J(\Omega^+)$ *is given by*

$$dJ(\Omega^+)(\theta) = \int_{\Gamma^+} [\sigma(-\partial_n u_d \partial_n p_d - \partial_n u_n \partial_n p_n)]_{\Gamma^+}\,\theta \cdot n$$
$$+ \int_{\Gamma^+} ([\sigma]_{\Gamma^+}(\nabla_{\Gamma^+} u_d \cdot \nabla_{\Gamma^+} p_d + \nabla_{\Gamma^+} u_n \cdot \nabla_{\Gamma^+} p_n) - [f]_{\Gamma^+}(p_n + p_d))\theta \cdot n.$$

Note that our results cover and generalize several results that can be found in the literature of shape optimization approaches for EIT, including [2,22]. For instance taking $\alpha_2 = 1$, $\alpha_1 = 0$ in Proposition 6.3 we get $p_D \equiv 0$ which yields the same formula as the one obtained in ([2], pp. 533).

Note also that from a numerical point of view, the boundary expression in Proposition 6.3 is delicate to compute compared to the domain expression in Proposition 6.2 for which the gradients of the state and adjoint states can be straightforwardly computed at grid points when using the finite element method for instance. The boundary expression, on the other hand, needs here the computation of the normal vector and the interpolation of the gradients on the interface Γ^+ which requires a precise description of the boundary and introduces an additional error.

7. LEVEL SET METHOD

The level set method, originally introduced in [34], gives a general framework for the computation of evolving interfaces using an implicit representation of these interfaces. The core idea of this method is to represent the boundary of the moving domain $\Omega^+(t) \subset D \in \mathbf{R}^N$ as the level set of a continuous function $\phi(\cdot, t) : D \to \mathbf{R}$.

Let us consider the family of domains $\Omega^+(t) \subset D$ as defined in (2.4). Each domain $\Omega^+(t)$ can be defined as

$$\Omega^+(t) := \{x \in D, \ \phi(x, t) < 0\} \tag{7.1}$$

where $\phi : D \times \mathbf{R}^+ \to \mathbf{R}$ is continuous and called *level set function*. Indeed, if we assume $|\nabla\phi(\cdot, t)| \neq 0$ on the set $\{x \in D, \ \phi(x, t) = 0\}$ then we have

$$\partial\Omega^+(t) = \{x \in D, \ \phi(x, t) = 0\}, \tag{7.2}$$

i.e. the boundary $\partial\Omega^+(t)$ is the zero level set of $\phi(\cdot, t)$.

Let $x(t)$ be the position of a particle on the boundary $\partial\Omega^+(t)$ moving with velocity $\dot{x}(t) = \theta(x(t))$ according to (2.3). Differentiating the relation $\phi(x(t), t) = 0$ with respect to t yields the Hamilton–Jacobi equation:

$$\partial_t\phi(x(t), t) + \theta(x(t)) \cdot \nabla\phi(x(t), t) = 0 \quad \text{in } \partial\Omega^+(t) \times \mathbf{R}^+,$$

which is then extended to all of D *via* the equation

$$\partial_t\phi(x, t) + \theta(x) \cdot \nabla\phi(x, t) = 0 \quad \text{in } D \times \mathbf{R}^+, \tag{7.3}$$

or alternatively to $U \times \mathbf{R}^+$ where U is a neighbourhood of $\partial\Omega^+(t)$.

Traditionally, the level set method has been designed to track smooth interfaces moving along the normal direction to the boundary. Theoretically, this is supported by Theorem 2.4, *i.e.* if the domain $\Omega^+(t)$ and the shape gradient are smooth enough then the shape derivative only depends on $\theta \cdot n$ on $\partial\Omega^+(t)$. In this case, we may choose for the optimization a vector field $\theta = \vartheta_n n$ on $\partial\Omega^+(t)$. Then, noting that an extension to D of the unit outward normal vector n to $\Omega^+(t)$ is given by $n = \nabla\phi/|\nabla\phi|$, and extending ϑ_n to all of D, one obtains from (7.3) the level set equation

$$\partial_t\phi + \vartheta_n|\nabla\phi| = 0 \quad \text{in } D \times \mathbf{R}^+. \tag{7.4}$$

The initial data $\phi(x, 0) = \phi_0(x)$ accompanying the Hamilton–Jacobi equation (7.3) or (7.4) is chosen as the signed distance function to the initial boundary $\partial\Omega^+(0)$ in order to satisfy the condition $|\nabla u| \neq 0$ on $\partial\Omega^+$, *i.e.*

$$\phi_0(x) = \begin{cases} d(x, \partial\Omega^+(0)), & \text{if } x \in (\Omega^+(0))^c, \\ -d(x, \partial\Omega^+(0)), & \text{if } x \in \Omega^+(0). \end{cases} \tag{7.5}$$

7.1. Level set method and domain expression

In the case of the distributed shape derivative, for instance (6.37) or (6.38), ϕ is not governed by (7.4) but rather by the Hamilton–Jacobi equation (7.3). Indeed we obtain a descent direction θ defined in D by solving (5.4), where $dJ(\Omega^+)$ is given by Proposition 6.2 which can subsequently be used in (7.3) to compute the evolution of ϕ. On the other hand, in the usual level set method, one solves a PDE *on the boundary* $\partial\Omega^+$ in an analogous way as for (5.4) (for instance using a Laplace-Beltrami operator), and uses the boundary expression from Proposition 6.3 to obtain $\vartheta_n = \theta \cdot n$ on $\partial\Omega^+$.

Numerically it is actually more straightforward in many cases to use (7.3) instead of (7.4). Indeed, when using (7.4), ϑ_n is initially only given on $\partial\Omega^+(t)$ and must be extended to the entire domain D or at least to a narrow band around $\partial\Omega^+(t)$. Therefore it is convenient to use (7.3) with θ already defined in D as is the case of the distributed shape derivative, which provides an extension to D or to a narrow band around $\partial\Omega^+(t)$.

In shape optimization, ϑ_n usually depends on the solution of one or several PDEs and their gradient. Since the boundary $\partial\Omega^+(t)$ in general does not match the grid nodes where ϕ and the solutions of the partial differential

equations are defined in the numerical application, the computation of ϑ_n requires the interpolation on $\partial \Omega^+(t)$ of functions defined at the grid points only, complicating the numerical implementation and introducing an additional interpolation error. This is an issue in particular for interface problems where ϑ_n is the jump of a function across the interface, as in Proposition 6.3, which requires multiple interpolations and is error-prone. In the distributed shape derivative framework θ only needs to be defined at grid nodes.

7.2. Discretization of the level set equation

Let D be the unit square $D = (0,1) \times (0,1)$ to fix ideas. For the discretization of the Hamilton–Jacobi equation (7.3), we first define the mesh grid corresponding to D. We introduce the nodes P_{ij} whose coordinates are given by $(i\Delta x, j\Delta y)$, $1 \le i, j \le N$ where Δx and Δy are the steps discretization in the x and y directions respectively. Let us also write $t^k = k\Delta t$ the discrete time for $k \in \mathbb{N}$, where Δt is the time step. We are seeking for an approximation $\phi_{ij}^k \simeq \phi(P_{ij}, t^k)$.

In the usual level set method, the level set equation (7.4) is discretized using an explicit upwind scheme proposed by Osher and Sethian [34, 36, 40]. This scheme applies to the specific form (7.4) but is not suited to discretize (7.3) required for our application. Equation (7.3) is of the form

$$\partial_t \phi + H(\nabla \phi) = 0 \quad \text{in } D \times \mathbf{R}^+. \tag{7.6}$$

where $H(\nabla \phi) := \theta \cdot \nabla \phi$ is the so-called Hamiltonian. We use the Local Lax-Friedrichs flux originally conceived in [35] and which reduces in our case to:

$$\hat{H}^{LLF}(p^-, p^+, q^-, q^+) = H\left(\frac{p^- + p^+}{2}, \frac{q^- + q^+}{2}\right) - \frac{1}{2}(p^+ - p^-)\alpha^x - \frac{1}{2}(p^+ - p^-)\alpha^y$$

where $\alpha^x = |\theta_x|$, $\alpha^y = |\theta_y|$, $\theta = (\theta_x, \theta_y)$ and

$$p^- = D_x^- \phi_{ij} = \frac{\phi_{ij} - \phi_{i-1,j}}{\Delta x}, \ p^+ = D_x^+ \phi_{ij} = \frac{\phi_{i+1,j} - \phi_{ij}}{\Delta x},$$
$$q^- = D_y^- \phi_{ij} = \frac{\phi_{ij} - \phi_{i,j-1}}{\Delta y}, \ q^+ = D_y^+ \phi_{ij} = \frac{\phi_{i,j+1} - \phi_{ij}}{\Delta y}$$

are the backward and forward approximations of the x-derivative and y-derivative of ϕ at P_{ij}, respectively. Using a forward Euler time discretization, the numerical scheme corresponding to (7.3) is

$$\phi_{ij}^{k+1} = \phi_{ij}^k - \Delta t \ \hat{H}^{LLF}(p^-, p^+, q^-, q^+). \tag{7.7}$$

For numerical accuracy, the solution of the level set equation (7.3) should not be too flat or too steep. This is fulfilled for instance if ϕ is the distance function *i.e.* $|\nabla \phi| = 1$. Even if one initializes ϕ using a signed distance function, the solution ϕ of the level set equation (7.3) does not generally remain close to a distance function. We may occasionally perform a reinitialization of ϕ by solving a parabolic equation up to the stationary state; see [16, 17, 38]. Although in the level set method this reinitialization is standard, in the case of the distributed shape gradient, we observe experimentally that the level set function ϕ stays close to a distance function during the iterations and we do not need to reinitialize. The regularization of the shape gradient could explain this observed stability of the level set function.

The computational efficiency of the level set method can be improved by using the so-called "narrow band" approach introduced in [1], which consists in computing and updating the level set function only on a thin region around the interface. This allows to reduce the complexity of the problem to $N \log(N)$ instead of N^2 in two dimensions. In this paper we do not implement this approach but we mention that it could also be applied to the distributed shape derivative approach and equation (7.3) by taking θ with a support in a narrow band around the moving interface, which can be achieved by choosing the appropriate space $\mathbf{E}$ in (5.3).

8. Application and numerical results

8.1. Electrical impedance tomography

In this section we give numerical results for the problem of electrical impedance tomography presented in Section 6.1, precisely we look for an approximate solution of the shape optimization problem (6.11). Using the notations of Section 6.1 we take $D = (0,1) \times (0,1)$ and $\Gamma_d = \emptyset$, *i.e.* we have measurements on the entire boundary Γ. For easiness of implementation, we consider a slightly different problem than the one in Section 6.1. Denote Γ_t, Γ_b, Γ_l and Γ_r the four sides of the square, where the indices t,b,l,r stands for top, bottom, left and right, respectively. We consider the following problems: find $u_n \in H^1_{tb}(D)$

$$\int_D \sigma \nabla u_n \cdot \nabla \varphi = \int_D f\varphi + \int_{\Gamma_l \cup \Gamma_r} g\varphi \quad \text{for all} \quad \varphi \in H^1_{0,tb}(D) \tag{8.1}$$

and find $u_d \in H^1_{lr}(D)$ such that

$$\int_D \sigma \nabla u_d \cdot \nabla \varphi = \int_D f\varphi + \int_{\Gamma_t \cup \Gamma_b} g\varphi \quad \text{for all} \quad \varphi \in H^1_{0,lr}(D) \tag{8.2}$$

where

$$
\begin{aligned}
H^1_{tb}(D) &:= \{v \in H^1(D) \mid v = h \text{ on } \Gamma_t \cup \Gamma_b\}, \\
H^1_{lr}(D) &:= \{v \in H^1(D) \mid v = h \text{ on } \Gamma_l \cup \Gamma_r\}, \\
H^1_{0,tb}(D) &:= \{v \in H^1(D) \mid v = 0 \text{ on } \Gamma_t \cup \Gamma_b\}, \\
H^1_{0,lr}(D) &:= \{v \in H^1(D) \mid v = 0 \text{ on } \Gamma_l \cup \Gamma_r\}.
\end{aligned}
$$

In our experiments we choose $f \equiv 0$. The results of Section 6.1 can be straightforwardly adapted to equations (8.1) and (8.2).

We use the software package FEniCS for the implementation; see [30]. The domain D is meshed using a regular grid of 128×128 elements and we describe the evolution of the interface Γ^+ using the level set method from Section 7. The conductivity values are set to $\sigma_0 = 1$ and $\sigma_1 = 10$.

We obtain measurements h_i corresponding to fluxes g_i, $i = 1,..,I$, by taking the trace on Γ of the solution of a Neumann problem where the fluxes are equal to g_i. To simulate real noisy EIT data, the measurements h_i are corrupted by adding a normal Gaussian noise with mean zero and standard deviation $\delta * \|h_i\|_\infty$, where δ is a parameter. The noise level is computed as

$$noise = \frac{\sum_{i=1}^I \|h_i - \tilde{h}_i\|_{L^2(\Gamma)}}{\sum_{i=1}^I \|h_i\|_{L^2(\Gamma)}}, \tag{8.3}$$

where $\tilde{h}_i$ is the noisy measurement and h_i the synthetic measurement without noise on Γ.

We use a variation of the functional (6.12), *i.e.* in our context:

$$J(\Omega^+) = \sum_{i=1}^I \mu_i \int_\Omega \frac{1}{2}(u_{d,i}(\Omega^+) - u_{n,i}(\Omega^+))^2, \tag{8.4}$$

where $u_{d,i}$ and $u_{n,i}$ correspond to the different fluxes g_i. Here the coefficients μ_i are weights associated to the fluxes g_i. In our experiments we choose the weights μ_i such that each term of the sums in (8.4) are equal to 1 on initialization in order to have a well-distributed influence of each term. Practically, the μ_i are thus calculated during the first iteration. We use the distributed shape derivative $dJ(\Omega^+)$ from Proposition (6.2). We obtain a descent direction by solving (5.4) with $\mathbf{E}$ a finite dimensional subspace of $H^1_0(D)$ and $\mathcal{B}(v,w) = \int_D Dv \cdot Dw$. We choose $\mathbf{E}$ to be the space of linear Lagrange elements.
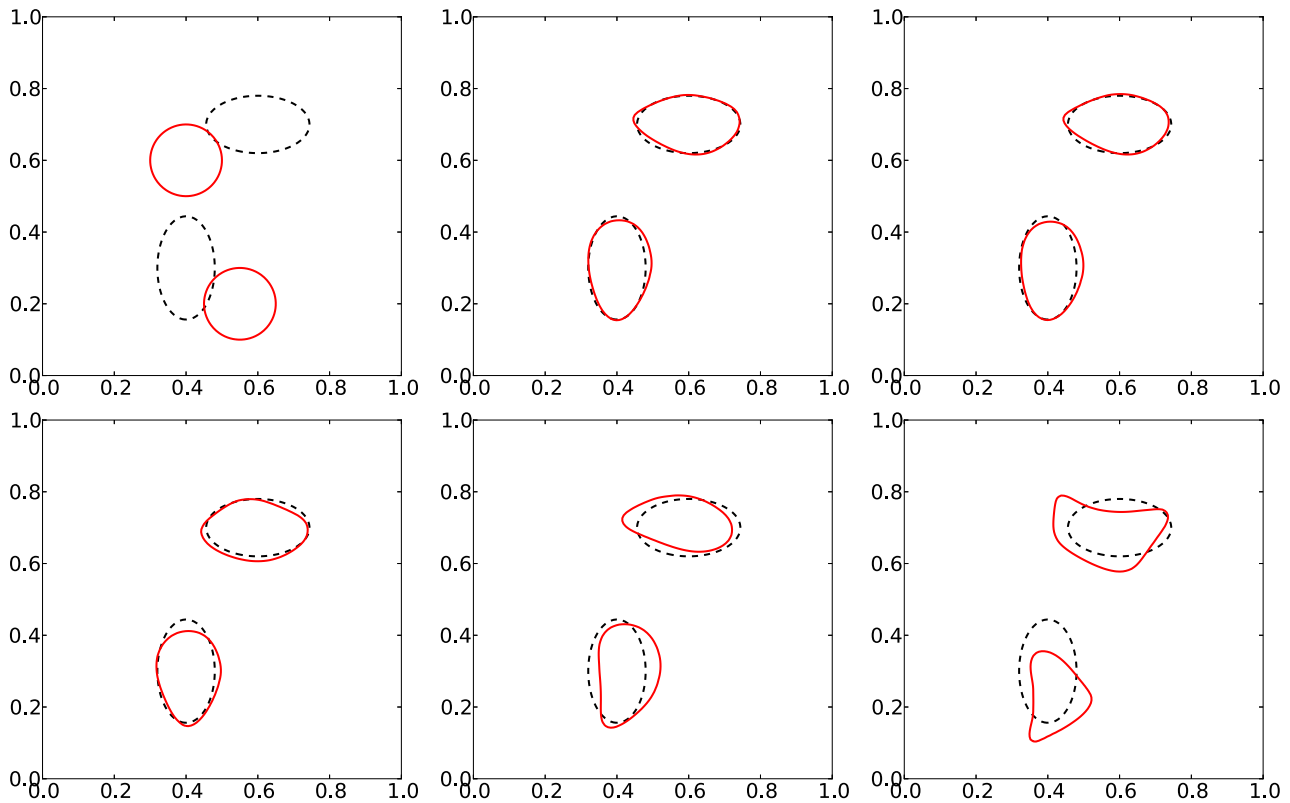
FIGURE 2. Reconstruction (*continuous contours*) of two ellipses (*dashed contours*) with different noise levels and using three measurements. From left to right and top to bottom: initialization (*continuous contours* - top left), 0% noise (367 iterations), 0.43% noise (338 iterations), 1.44% noise (334 iterations), 2.83% noise (310 iterations), 7% noise (356 iterations).

Since we use a gradient-based method we implement an Armijo line search to adjust the time-stepping. The algorithm is stopped when the decrease of the functional becomes insignificant, practically when the following stopping criterion is repeatedly satisfied:

$$J(\Omega_k^+) - J(\Omega_{k+1}^+) < \gamma(J(\Omega_0^+) - J(\Omega_1^+))$$

where Ω_k^+ denotes the kth iterate of Ω^+. We take $\gamma = 5.10^{-5}$ in our tests.

In Figure 2 we compare the reconstruction for different noise levels computed using (8.3). We take in this example $I = 3$, *i.e.* we use three fluxes g_i, $i = 1, 2, 3$, defined as follows:

$$g_1 = 1 \text{ on } \Gamma_l \cup \Gamma_r \text{ and } g_1 = -1 \text{ on } \Gamma_t \cup \Gamma_b,$$
$$g_2 = 1 \text{ on } \Gamma_l \cup \Gamma_t \text{ and } g_2 = -1 \text{ on } \Gamma_r \cup \Gamma_b,$$
$$g_3 = 1 \text{ on } \Gamma_l \cup \Gamma_b \text{ and } g_3 = -1 \text{ on } \Gamma_r \cup \Gamma_t.$$

Without noise, the reconstruction is very close to the true object and degrades as the measurements become increasingly noisy, as is usually the case in EIT. However, the reconstruction is quite robust with respect to noise considering that the problem is severely ill-posed. We reconstruct two ellipses and initialize with two balls placed at the wrong location. The average number of iterations until convergence is around 340 iterations.
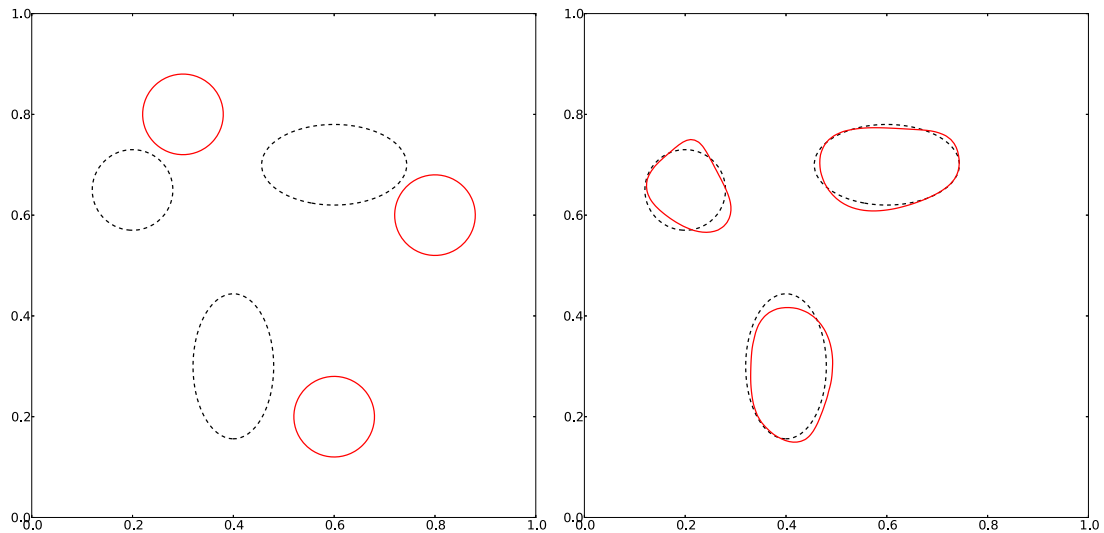
FIGURE 3. Initialization (*continuous contours* – left) and reconstruction (*continuous contours* – right) of two ellipses and a ball (*dashed contours*) with 1.55% noise (371 iterations) and using seven measurements.
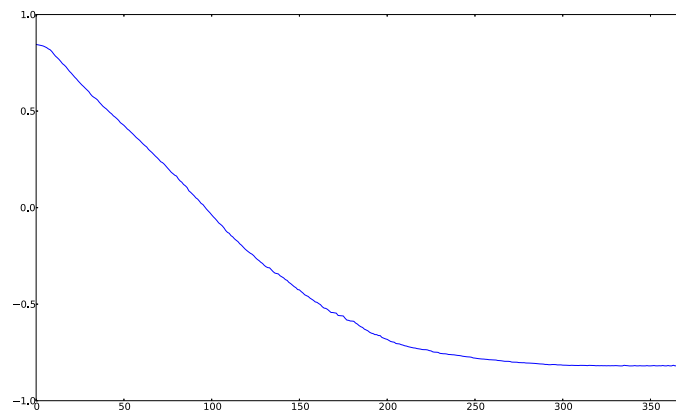

FIGURE 4. History of cost functional corresponding to Figure 3 in logarithmic scale.

In Figure 3 we reconstruct three inclusions this time using $I = 7$ different measurements, with 1.55% noise. The reconstruction is close to the true inclusion and is a bit degraded due to the noise. Figure 4 shows the convergence history of the cost functional in log scale for this example.

Our algorithm gives good results in comparison to existing results in the literature using level set methods to solve the EIT problem. In [11] the EIT problem has been treated numerically using a level set method, which is not based on the use of shape derivatives but on the differentiation of a smooth approximation of the Heaviside function to represent domains. In [22] the level set method using the boundary expression of the shape derivative is used based on equation (7.4).

Our algorithm converges fast in comparison to [11, 22]: convergence occurs after around 300 iterations. In [11] convergence occurs between 200 iterations for one inclusion and up to 50000 iterations for two inclusions. In [22] convergence occurs after 2000 or 10000 iterations on two examples with three inclusions. Concerning

measurements we obtain good reconstruction of two inclusions with $I = 3$ and three inclusions with $I = 7$, while in [11] sets of $4, 12, 28$ and 60 measurements are used but usually 60 measurements are required for complicated shapes such as two inclusions. In [22], 60 measurements are used. Nevertheless, our results are not directly comparable since the conductivities are unknown in [11,22], which makes the inverse problem harder and might explain the slower convergence. Also, the reconstructed shapes are not the same although the complexity of the unknown shapes is comparable since we also consider two and three inclusions as in [11,22]. Only an exact comparison using the same problem, test case, initialization, noise level, number and type of measurements could allow to conclude.

References

[1] D. Adalsteinsson and J.A. Sethian, A fast level set method for propagating interfaces. *J. Comput. Phys.* **118** (1995) 269–277.

[2] L. Afraites, M. Dambrine and D. Kateb, Shape methods for the transmission problem with a single measurement. *Numer. Funct. Anal. Optim.* **28** (2007) 519–551.

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

[4] Z. Belhachmi and H. Meftahi, Shape sensitivity analysis for an interface problem via minimax differentiability. *Appl. Math. Comput.* **219** (2013) 6828–6842.

[5] L. Borcea, Electrical impedance tomography. *Inverse Problems* **18** (2002) R99–R136.

[6] A.-P. Calderón, On an inverse boundary value problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro, 1980)*. Soc. Brasil. Mat., Rio de Janeiro (1980) 65–73.

[7] A. Canelas, A. Laurain and A.A. Novotny, A new reconstruction method for the inverse potential problem. *J. Comput. Phys.* **268** (2014) 417–431.

[8] A. Canelas, A. Laurain and A.A. Novotny, A new reconstruction method for the inverse source problem from partial boundary measurements. *Inverse Problems* **31** (2015) 075009.

[9] J. Céa, Conception optimale ou identification de formes: calcul rapide de la dérivée directionnelle de la fonction coût. *RAIRO M2AN* **20** (1986) 371–402.

[10] M. Cheney, D. Isaacson and J.C. Newell, Electrical impedance tomography. *SIAM Rev.* **41** (1999) 85–101 (electronic).

[11] E.T. Chung, T.F. Chan and X.-C. Tai, Electrical impedance tomography using level set representation and total variational regularization. *J. Comput. Phys.* **205** (2005) 357–372.

[12] M.C. Delfour and J.-P. Zolésio, Shape sensitivity analysis via min max differentiability. *SIAM J. Control Optim.* **26** (1988) 834–862.

[13] M.C. Delfour and J.-P. Zolésio, Shapes and geometries. Metrics, analysis, differential calculus, and optimization. Vol. 22 of *Advances in Design and Control*, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2011).

[14] J.D. Eshelby, The elastic energy-momentum tensor. Special issue dedicated to A.E. Green. *J. Elasticity* **5** (1975) 321–335.

[15] P. Fulmanski, A. Laurain and J.-F. Scheid, Level set method for shape optimization of Signorini problem. In *MMAR Proceedings* (2004) 71–75.

[16] P. Fulmański, A. Laurain, J.-F. Scheid and J. Sokołowski, A level set method in shape and topology optimization for variational inequalities. *Int. J. Appl. Math. Comput. Sci.* **17** (2007) 413–430.

[17] P. Fulmański, A. Laurain, J.-F. Scheid and J. Sokołowski, Level set method with topological derivatives in shape optimization. *Int. J. Comput. Math.* **85** (2008) 1491–1514.

[18] P. Fulmanski, A. Laurain, J.-F. Scheid and J. Sokolowski, Une méthode levelset en optimisation de formes. In *CANUM 2006 – Congrès National d'Analyse Numérique*. Vol. 22 of *ESAIM Proc. Survey* EDP Sciences, Les Ulis (2008) 162–168.

[19] J. Hadamard, Mémoire sur le probleme d'analyse relatif a l'équilibre des plaques élastiques. In Mémoire des savants étrangers, 33, 1907, Œuvres de Jacques Hadamard. Editions du C.N.R.S., Paris (1968) 515–641.

[20] A. Henrot and M. Pierre, Variation et optimisation de formes. Une analyse géométrique. [A geometric analysis]. Vol. 48 of *Math. Appl.* Springer, Berlin (2005).

[21] F. Hettlich, The domain derivative of time-harmonic electromagnetic waves at interfaces. *Math. Methods Appl. Sci.* **35** (2012) 1681–1689.

[22] M. Hintermüller and A. Laurain, Electrical impedance tomography: from topology to shape. *Control Cybernet.* **37** (2008) 913–933.

[23] M. Hintermüller and A. Laurain, Multiphase image segmentation and modulation recovery based on shape and topological sensitivity. *J. Math. Imaging Vision* **35** (2009) 1–22.

[24] M. Hintermüller and A. Laurain, Optimal shape design subject to elliptic variational inequalities. *SIAM J. Control Optim.* **49** (2011) 1015–1047.

[25] M. Hintermüller, A. Laurain and A.A. Novotny, Second-order topological expansion for electrical impedance tomography. *Adv. Comput. Math.* (2011) 1–31.

[26] M. Hintermüller, A. Laurain and I. Yousept, Shape sensitivities for an inverse problem in magnetic induction tomography based on the eddy current model. *Inverse Problems* **31** (2015) 065006.

[27] R. Hiptmair, A. Paganini and S. Sargheini, Comparison of approximate shape gradients. *BIT* **55** (2015) 459–485.

[28] D. Hömberg and J. Sokołowski, Optimal shape design of inductor coils for surface hardening. *SIAM J. Control Optim.* **42** (2003) 1087–1117 (electronic).

[29] R. Kress, Inverse problems and conformal mapping. *Complex Var. Elliptic Equ.* **57** (2012) 301–316.

[30] A. Logg, K.-A. Mardal and G.N. Wells, editors, Automated Solution of Differential Equations by the Finite Element Method. Vol. 84 of *Lecture Notes in Computational Science and Engineering*. Springer (2012).

[31] J.L. Mueller and S. Siltanen, Linear and nonlinear inverse problems with practical applications. Vol. 10 of *Computational Science & Engineering*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2012).

[32] M. Nagumo, Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen. *Proc. Phys. Math. Soc. Japan* **24** (1942) 551–559.

[33] A.A. Novotny and J. Sokołowski, Topological derivatives in shape optimization. *Interaction of Mechanics and Mathematics*. Springer, Heidelberg (2013).

[34] S. Osher and J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. *J. Comput. Phys.* **79** (1988) 12–49.

[35] S. Osher and C.-W. Shu, High-order essentially nonoscillatory schemes for Hamilton–Jacobi equations. *SIAM J. Numer. Anal.* **28** (1991) 907–922.

[36] S. Osher and R. Fedkiw, Level set methods and dynamic implicit surfaces. Vol. 153 of *Applied Mathematical Sciences*. Springer–Verlag, New York (2003).

[37] O. Pantz, Sensibilité de l'équation de la chaleur aux sauts de conductivité. *C. R. Math. Acad. Sci. Paris* **341** (2005) 333–337.

[38] D. Peng, B. Merriman, S. Osher, H. Zhao and M. Kang, A PDE-based fast local level set method. *J. Comput. Phys.* **155** (1999) 410–438.

[39] M. Renardy and R.C. Rogers. An introduction to partial differential equations. Vol. 13 of *Texts in Applied Mathematics*, 2nd edn. Springer–Verlag, New York (2004).

[40] J.A. Sethian, Level set methods and fast marching methods. Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science. Vol. 3 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, second edition (1999).

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

[42] K. Sturm, Minimax Lagrangian approach to the differentiability of non-linear PDE constrained shape functions without saddle point assumption. *SIAM J. Control Optim.* **53** (2015) 2017–2039.

[43] K. Sturm, D. Hömberg and M. Hintermüller, Shape optimization for a sharp interface model of distortion compensation. *WIAS-preprint* **4** (2013) 807–822.

[44] J.-P. Zolésio, *Identification de domaines par déformations*. Thèse de doctorat d'état, Université de Nice, France (1979).