An *a posteriori* error estimator for shape optimization: application to EIT

M Giacomini$^{1,2}$, O Pantz$^1$ and K Trabelsi$^2$

$^1$ CMAP Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France (M. Giacomini and O. Pantz are members of the DEFI project at INRIA Saclay Ile-de-France)

$^2$ DRI Institut Polytechnique des Sciences Avancées, 7-9 rue M. Grandcoing, 94200 Ivry-sur-Seine, France

E-mail: matteo.giacomini@polytechnique.edu

Abstract. In this paper we account for the numerical error introduced by the Finite Element approximation of the shape gradient to construct a guaranteed shape optimization method. We present a goal-oriented strategy inspired by the complementary energy principle to construct a constant-free, fully-computable *a posteriori* error estimator and to derive a certified upper bound of the error in the shape gradient. The resulting Adaptive Boundary Variation Algorithm (ABVA) is able to identify a genuine descent direction at each iteration and features a reliable stopping criterion for the optimization loop. Some preliminary numerical results for the inverse identification problem of Electrical Impedance Tomography are presented.

1. Introduction

A classical approach to shape optimization and shape identification problems relies on the application of the Steepest Descent Method using the so-called shape gradient. In most applications, the shape gradient depends on the solution of a PDE which can only be solved approximately. Thus stopping criteria based on its norm are never fulfilled if the *a priori* given tolerance is too small. In this paper, we evaluate the discretization error in the shape gradient relying on the well-established results on goal-oriented error estimators [1] and we propose a guaranteed shape optimization strategy. Several works in the literature have focused on Adaptive Finite Element Method for shape optimization problems ([2], [3], [4], [5]) but to the best of our knowledge, all the proposed approaches used error estimators only to drive mesh adaptivity without accounting for the error in the shape gradient itself.

After briefly recalling the framework of shape optimization and shape identification problems in section 2 and the notion of shape gradient in section 3, we introduce the Adaptive Boundary Variation Algorithm for shape optimization problems (Section 4). In section 5 we present some preliminary numerical simulations for the inverse problem of Electrical Impedance Tomography and a summary of our results ends the paper.

2. Shape optimization and shape identification problems

Let us consider an open domain $\Omega \subset \mathbb{R}^d$ ($d \geq 2$) with Lipschitz boundary $\partial \Omega$. Let $V_\Omega$ be a separable Hilbert space depending on $\Omega$, we define $u_\Omega \in V_\Omega$ to be the solution of a state equation

$$a_{\Omega}(u_\Omega, \delta u) = F_\Omega(\delta u) \quad \forall \delta u \in V_\Omega$$

(1)
where $a_{\Omega}(\cdot,\cdot)$ is a continuous and coercive bilinear form on $V_\Omega$ and $F_\Omega(\cdot)$ is a continuous linear form on $V_\Omega$, both of them depending on $\Omega$. Under these assumptions, problem (1) has a unique solution.

We introduce a cost functional $J(\Omega) = j(\Omega, u_\Omega)$ which depends on the domain $\Omega$ itself and on the solution $u_\Omega$ of the state equation. We consider the following shape optimization problem

$$\min_{\Omega \in \mathcal{U}_{\text{ad}}} J(\Omega)$$

(2)

where $\mathcal{U}_{\text{ad}}$ is the set of admissible domains in $\mathbb{R}^d$. Within this framework, problem (2) may be viewed as a PDE-constrained optimization problem where we aim at minimizing $j(\Omega, u)$ under the constraint $u = u_\Omega$.

The case of Electrical Impedance Tomography

Let $D$ be a domain featuring an inclusion $\Omega$. We define the conductivity $k_{\Omega}$ as a piecewise constant function equal to $k_I$ inside the inclusion $\Omega$ and $k_{\Omega}$ otherwise $(k_I, k_{\Omega} > 0)$. We are interested in identifying the shape and the location of the inclusion, fitting given boundary measurements $g$ and $U_D$ of the flux and the potential. Thus we introduce $u_{\Omega,N}, u_{\Omega,D} \in H^1(D)$ such that $u_{\Omega,D} = U_D$ on $\partial D$, solutions of respectively a Neumann and a Dirichlet Boundary Value Problem:

$$\int_D \left( k_{\Omega} \nabla u_{\Omega,i} \cdot \nabla \delta u_i + u_{\Omega,i} \delta u_i \right) dx = \begin{cases} \int_{\partial D} g \delta u_i ds & \forall \delta u_i \in H^1(D), \ i = N \\ 0 & \forall \delta u_i \in H^1_0(D), \ i = D \end{cases}$$

(3)

The inverse identification problem of Electrical Impedance Tomography may be written as in (2) and we seek the open subset that minimizes the Kohn-Vogelius functional [6]:

$$J(\Omega) = \frac{1}{2} \int_D \left( k_{\Omega} |\nabla(u_{\Omega,N} - u_{\Omega,D})|^2 + |u_{\Omega,N} - u_{\Omega,D}|^2 \right) dx.$$  

(4)

3. Differentiation with respect to the shape

Let $X$ be a Banach space and $\theta \in X$ be an admissible smooth deformation of $\Omega$. The objective functional $J$ is said to be $X$-differentiable at $\Omega \in \mathcal{U}_{\text{ad}}$ if there exists a continuous linear form $dJ(\Omega)$ on $X$ such that $\forall \theta \in X$ we have $J((\text{Id} + \theta)\Omega) = J(\Omega) + \langle dJ(\Omega), \theta \rangle + o(\theta)$.

Let us introduce the Lagrangian functional, defined for every admissible open set $\Omega$ and every $u, p \in V_\Omega$ by

$$\mathcal{L}(\Omega, u, p) = j(\Omega, u) + a_{\Omega}(u, p) - F_\Omega(p),$$

(5)

and the adjoint state $p_\Omega \in V_\Omega$ such that $\forall \delta p \in V_\Omega$

$$a^*(p_\Omega, \delta p) + \left\langle \frac{\partial j}{\partial u}(\Omega, u_\Omega), \delta u \right\rangle = 0$$

(6)

Let us now introduce a diffeomorphism $\varphi : \mathbb{R}^d \to \mathbb{R}^d$. For every admissible set $\Omega$, and for all $u, p \in V_\Omega$, we set $\Omega_\varphi := \varphi(\Omega)$, $u_\varphi := u \circ \varphi^{-1}$ and $p_\varphi := p \circ \varphi^{-1}$. We admit that $u \mapsto u_\varphi$ is a one-to-one map between $V_\Omega$ and $V_{\varphi(\Omega)}$. We define the material derivative $\frac{\partial \mathcal{L}}{\partial \varphi}$ of $\mathcal{L}$ as the derivative of the map $\varphi \mapsto \mathcal{L}(\Omega_\varphi, u_\varphi, p_\varphi)$ evaluated in $\varphi = \text{Id}$. Hence, from the fast derivation method of Céa [7] the shape gradient reads as

$$\langle dJ(\Omega), \theta \rangle = \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega, p_\Omega), \theta \right\rangle.$$  

(7)
The case of Electrical Impedance Tomography

In order to compute the shape gradient of (4), we need to determine the adjoint solutions $p_{\Omega,N}$ and $p_{\Omega,D}$ associated with the states $u_{\Omega,N}$ and $u_{\Omega,D}$: the Kohn-Vogelius problem is self-adjoint and we get that $p_{\Omega,N} = u_{\Omega,N} - u_{\Omega,D}$ and $p_{\Omega,D} = 0$. The most common approach in the literature is based on an Eulerian viewpoint and leads to a surface expression of the shape gradient

$$\langle dJ(\Omega), \theta \rangle = \frac{1}{2} \int_{\partial \Omega} \left( \|k_1\|^2 \left( \frac{\partial u_{\Omega,N}}{\partial \tau} \right)^2 - \|k_2\|^2 \left( \frac{\partial u_{\Omega,D}}{\partial \tau} \right)^2 \right) (\theta \cdot n) ds$$

where $n$ is the outward normal to $\partial \Omega$, $\tau$ is the tangential direction to $\partial \Omega$ and $\|k_1\| = k_E - k_I$ and $\|k_2\| = k_I^{-1} - k_I^{-1}$ are the jumps across $\partial \Omega$. Let us now introduce the following operator:

$$\langle G(\Omega, u), \theta \rangle = \frac{1}{2} \int_{\Omega} \left( k_1 M(\theta) \nabla u \cdot \nabla \theta - \nabla \cdot \theta^2 u^2 \right) dx$$

where $M(\theta) = \nabla \theta + \nabla \theta^T - \left( \nabla \cdot \theta \right)$ Id. From (7), we get a volumetric expression of the shape gradient

$$\langle dJ(\Omega), \theta \rangle = \langle G(\Omega, u_{\Omega,N}) - G(\Omega, u_{\Omega,D}), \theta \rangle$$

We refer to [8] for more details on the differentiation of the Kohn-Vogelius functional and its application to the identification of discontinuities of the conductivity parameter.

4. The Adaptive Boundary Variation Algorithm for shape optimization

In algorithm 1, we present the Boundary Variation Algorithm, that is the application of the classical Steepest Descent Method to a shape optimization problem. After computing the solution of the state equation, we derive the expression of the shape gradient and we compute the corresponding descent direction to deform the computational domain.

Algorithm 1: Continuous gradient method - The Boundary Variation Algorithm

Given the domain $\Omega_0$, set $\ell = 0$ and iterate:

1. Compute the solutions $u_{\Omega_\ell}$ and $p_{\Omega_\ell}$ of the state and adjoint equations;
2. Compute a descent direction $\theta_\ell \in X$ solving $(\theta_\ell, \delta \theta)_X + \langle dJ(\Omega_\ell), \delta \theta \rangle = 0 \ \forall \delta \theta \in X$;
3. Identify an admissible step $\mu_\ell$ and update the domain $\Omega_{\ell+1} = (\text{Id} + \mu_\ell \theta_\ell) \Omega_\ell$;
4. While $\langle \langle dJ(\Omega_\ell), \theta_\ell \rangle \rangle > \text{tol}$, $\ell = \ell + 1$ and repeat.

This method relies on the computation at each iteration of a direction $\theta$ such that the shape gradient of the objective functional in this direction is strictly negative, that is we seek $\theta$ such that $\langle dJ(\Omega), \theta \rangle < 0$.

Let us now introduce the discretized solutions $u_{\Omega_\ell}^h$ and $p_{\Omega_\ell}^h$ of the state and adjoint equations using the Finite Element Method. We consider for both $u_{\Omega_\ell}^h$ and $p_{\Omega_\ell}^h$ the approximation in the space of linear Lagrangian Finite Element functions. In the same fashion, we define the Finite Element discretization of the shape gradient as

$$\langle d_hJ(\Omega), \theta \rangle = \left\langle \frac{\partial L}{\partial \varphi}(\Omega, u_{\Omega_\ell}^h, p_{\Omega_\ell}^h), \theta \right\rangle$$

(11)

The discretized descent direction $\theta^h \in X$ is computed by solving problem (12) using linear Lagrangian Finite Elements.

$$\langle \theta^h, \delta \theta^h \rangle_X + \langle d_hJ(\Omega), \delta \theta^h \rangle = 0 \ \forall \delta \theta^h \in X$$

(12)

We remark that due to the numerical approximation, it is not sufficient that $\langle d_hJ(\Omega), \theta^h \rangle < 0$ for $J(\Omega)$ to decrease along the direction $\theta^h$. Moreover, if the approximation error is too large the stopping criterion (Algorithm 1 - step 4) will usually not be fulfilled. In this paper we present a strategy to bypass this issue by providing a computable upper bound of the error $E^h$ in the approximation of the shape gradient and by seeking a certified direction $\theta^h$ such that $\langle d_hJ(\Omega), \theta^h \rangle + |E^h| < 0$. 

3
4.1. Numerical error in the shape gradient
Let us write the numerical error \( E^h \) in the shape gradient as follows

\[
E^h = \langle dJ(\Omega) - d_hJ(\Omega), \theta^h \rangle = \left\langle \frac{\partial \mathcal{L}}{\partial \varphi} (\Omega, u^h_\Omega, p_\Omega) - \frac{\partial \mathcal{L}}{\partial \varphi} (\Omega, u^h_\Omega, p^h_\Omega), \theta^h \right\rangle
\]

\[
\simeq \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u} (\Omega, u^h_\Omega, p^h_\Omega)[\theta^h, u^h_\Omega - u^h_\Omega] + \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p} (\Omega, u^h_\Omega, p^h_\Omega)[\theta^h, p_\Omega - p^h_\Omega]
\]  

(13)

In order to compute an upper bound of the error (13), we use the well-established framework of a posteriori error estimators for a Quantity of Interest [1]. For this purpose we introduce two adjoint problems, each of which is associated with one term on the right-hand side of (13) and we seek \( r_\Omega, s_\Omega \in V_\Omega \) such that \( \forall \delta r, \delta s \in V_\Omega \)

\[
a^*_\Omega(r_\Omega, \delta r) = \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u} (\Omega, u^h_\Omega, p^h_\Omega)[\theta^h, \delta r] \quad \text{and} \quad a^*_\Omega(s_\Omega, \delta s) = \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p} (\Omega, u^h_\Omega, p^h_\Omega)[\theta^h, \delta s]
\]

(14)

By plugging (14) into (13) and exploiting Galerkin orthogonality and Cauchy-Schwarz inequality, we may derive the upper bound \( \bar{E} \) for the numerical error in the shape gradient:

\[
|E^h| \leq \bar{E} := \|r_\Omega - r^h_\Omega\|_{\Omega} \|u^h_\Omega - u^h_\Omega\|_{\Omega} + \|s_\Omega - s^h_\Omega\|_{\Omega} \|p_\Omega - p^h_\Omega\|_{\Omega}
\]

(15)

where \( \cdot \| \cdot \) is the energy-norm induced by the bilinear form \( a(\cdot, \cdot) \).

We remark that the criterion \( \langle d_hJ(\Omega), \theta^h \rangle + \bar{E} < 0 \) ensures that \( \theta^h \) is a certified genuine descent direction, whether it is the solution of equation (12) or not. In particular, this also stands when the latter problem is only solved approximately.

4.2. Certified Steepest Descent Method
We may now present the resulting Adaptive Boundary Variation Algorithm arising from the coupling of the classical Boundary Variation Algorithm for shape optimization with the goal-oriented estimator of the error in the shape gradient.

Algorithm 2: Discretized gradient method - The Adaptive Boundary Variation Algorithm

---

**Given the domain** \( \Omega_0 \), **set** \( \ell = 0 \) and **iterate**

1. **Construct a coarse mesh** for \( \Omega_\ell \);
2. **Compute the solutions** \( u^h_\ell \) and \( p^h_\ell \) of the state and adjoint equations;
3. **Compute a descent direction** \( \theta^\ell \in X \) solving \( \langle \theta^\ell, \delta \theta^\ell \rangle_X + \langle d_hJ(\Omega_\ell), \delta \theta^\ell \rangle = 0 \) \( \forall \delta \theta^\ell \in X \);
4. **Compute an upper bound \( \bar{E} \)** of the numerical error \( |E^h| \):
   (a) **Compute the solutions** \( r^h_\ell \) and \( s^h_\ell \) to estimate the error in the Quantity of Interest;
   (b) **Compute** \( \bar{E} = \|r_\Omega - r^h_\Omega\|_{\Omega} \|u^h_\Omega - u^h_\Omega\|_{\Omega} + \|s_\Omega - s^h_\Omega\|_{\Omega} \|p_\Omega - p^h_\Omega\|_{\Omega} \);
5. If \( \langle d_hJ(\Omega_\ell), \theta^\ell \rangle + \bar{E} \geq 0 \), **refine the mesh** and go to 2;
6. **Identify an admissible step** \( \mu_\ell \) and update the domain \( \Omega_{\ell+1} = (\text{Id} + \mu_\ell \theta^\ell)\Omega_\ell \);
7. **While** \( \|d_hJ(\Omega_{\ell+1}), \theta^\ell \| + \bar{E} > \text{tol} \), \( \ell = \ell + 1 \) and **repeat**

---

The case of Electrical Impedance Tomography
Since the Kohn-Vogelius problem is self-adjoint, (15) reduces to

\[
\bar{E} = \|r_{\Omega,N} - r^h_{\Omega,N}\|_{\Omega} \|u_{\Omega,N} - u^h_{\Omega,N}\|_{\Omega} + \|r_{\Omega,D} - r^h_{\Omega,D}\|_{\Omega} \|u_{\Omega,D} - u^h_{\Omega,D}\|_{\Omega}
\]

(16)

where \( r_{\Omega,N} \in H^1(\mathcal{D}) \) and \( r_{\Omega,D} \in H^1_0(\mathcal{D}) \) are the solutions of the following adjoint problems introduced to construct the estimator in the Quantity of Interest:

\[
\int_{\mathcal{D}} \left( k_n \nabla r_{\Omega,i} \cdot \nabla \delta r_i + r_{\Omega,i} \delta r_i \right) dx = \left[ \frac{\partial G}{\partial u} (\Omega, u^h_{\Omega,i}) \right][\theta^h, \delta r_i] \quad \forall \delta r_i \in H^1(\mathcal{D}), \ i = N
\]

\[
\int_{\mathcal{D}} \left( k_n \nabla r_{\Omega,i} \cdot \nabla \delta r_i + r_{\Omega,i} \delta r_i \right) dx = \left[ \frac{\partial G}{\partial u} (\Omega, u^h_{\Omega,i}) \right][\theta^h, \delta r_i] \quad \forall \delta r_i \in H^1_0(\mathcal{D}), \ i = D
\]

(17)
To fully compute the estimate (16), we use the complementary energy principle. For the state equations, we introduce the dual flux variables $\sigma_{\Omega,N}, \sigma_{\Omega,D} \in H(\text{div}) = \{ \sigma \in L^2(D; \mathbb{R}^d) : \text{div} \sigma \in L^2(D) \}$ such that $\sigma_{\Omega,N} \cdot n = g$ on $\partial D$ which satisfy $\forall \delta\sigma_N, \delta\sigma_D \in H(\text{div})$ such that $\delta\sigma_N \cdot n = 0$ on $\partial D$

$$\int_D \left( k_\Omega^{-1} \sigma_{\Omega,i} \cdot \delta\sigma_i + (\nabla \cdot \sigma_{\Omega,i})(\nabla \cdot \delta\sigma_i) \right) dx = \begin{cases} 0 & \text{for } i = N \\ \int_{\partial D} (\delta\sigma_i \cdot n) ds & \text{for } i = D \end{cases} \tag{18}$$

Let $\sigma^h_{\Omega,N}$ and $\sigma^h_{\Omega,D}$ be the dual fluxes discretized using the lowest order Raviart-Thomas Finite Elements. We obtain the following bound for the energy-norm of the error in the state equations:

$$\|u_{\Omega,i} - u^h_{\Omega,i}\|_V^2 \leq \int_D \left( k_\Omega^{-1} |\sigma^h_{\Omega,i} - k_\Omega \nabla u^h_{\Omega,i}|^2 + |\nabla \cdot \sigma^h_{\Omega,i} - u^h_{\Omega,i}|^2 \right) dx \tag{19}$$

In a similar fashion, we introduce the dual fluxes $\xi_{\Omega,N}, \xi_{\Omega,D} \in H(\text{div})$ such that $\xi_{\Omega,N} \cdot n = 0$ on $\partial D$ satisfying $\forall \delta\xi_N, \delta\xi_D \in H(\text{div})$ such that $\delta\xi_N \cdot n = 0$

$$\int_D \left( k_\Omega^{-1} \xi_{\Omega,i} \cdot \delta\xi_i + (\nabla \cdot \xi_{\Omega,i})(\nabla \cdot \delta\xi_i) \right) dx = \int_D \left( (\nabla \cdot \theta^h) u^h_{\Omega,i} \nabla \cdot \delta\xi_i - M(\theta^h) \nabla u^h_{\Omega,i} \cdot \delta\xi_i \right) dx \tag{20}$$

and via their $RT_0$ approximations $\xi^h_{\Omega,i}$’s, we derive the upper bound of the adjoint errors

$$\|r_{\Omega,i} - r^h_{\Omega,i}\|_V^2 \leq \int_D \left( k_\Omega^{-1} |\xi^h_{\Omega,i} + k_\Omega M(\theta^h) \nabla u^h_{\Omega,i} - k_\Omega \nabla r^h_{\Omega,i}|^2 + |\nabla \cdot \xi^h_{\Omega,i} - (\nabla \cdot \theta^h) u^h_{\Omega,i} - r^h_{\Omega,i}|^2 \right) dx \tag{21}$$

Eventually, by plugging (19) and (21) into (16), we are able to explicitly compute the bound $E$.

5. Numerical results

We apply the ABVA to the identification of an unknown inclusion $\Omega$ inside a circular domain $D$ using one boundary measurement. The deformation of the inclusion (Algorithm 2 - step 6) is performed by moving the whole computational mesh. In figure 1c, we present the reconstructed interface after 25 iterations of the procedure; figure 1d confirms that the ABVA identifies a genuine descent direction at each iteration, generating a sequence of minimizing shapes such that the objective functional is monotonically decreasing and provides a guaranteed stopping criterion. Moreover, in figure 1e we observe that the number of Degrees of Freedom remains small until the reconstructed interface approaches the real inclusion thus coarse meshes prove to be reliable during the initial iterations of the algorithm.

In figure 2, we report the results of a more complex case in which two inclusions are sought. We use ten boundary measurements and the ABVA is still able to identify a genuine descent direction at each iteration (Fig. 2b) but the reconstructed interface appears to be less precise (Fig. 2a). This phenomenon depends on the severe ill-posedness of the problem and the diffusive nature of the state equation is responsible for the loss of the information in the middle of the domain, even when using several measurements. This results in the ability of the algorithm to reconstruct the inclusions near the boundary of the domain whereas the inner interfaces appear more difficult to identify and with a less precise outcome. As previously observed, figure 2c confirms that the ABVA allows to use the same meshes for several iterations increasing the number of DOFs only when the descent direction is no more validated.

6. Conclusion

Introducing an a posteriori error estimator for the shape gradient provides quantitative information to define a reliable stopping criterion for the algorithm. Nevertheless, solving the
Figure 1: One inclusion in a circular domain. (a) Initial mesh. (b) Final mesh. (c) Initial configuration (dotted green), target inclusion (solid black) and reconstructed interface (dashed blue). (d) Objective functional. (e) Number of Degrees of Freedom.

Figure 2: Two inclusions in a square domain. (a) Initial (dotted green), target (solid black) and reconstructed configuration (dashed blue). (b) Objective functional. (c) Number of DOFs.

dual flux problems to derive an upper bound of the error in the shape gradient is computationally expensive hence the ABVA may result in higher computing times than the classical BVA on fine meshes. Prompted by the above promising results, we are now investigating alternative procedures to estimate the error in order to rely only on the computation of local quantities.

References
[1] Oden J and Prudhomme S 2001 Comput. Math. Appl. 41 735–756
[2] Banichuk N, Barthold F J, Falk A and Stein E 1995 Struct. Optim. 9 46–51
[3] Morin P, Nochetto R, Pauletti M and Verani M 2012 ESAIM: Control Optim. Calc. Var. 18(04) 1122–1149
[4] Kikuchi N, Chung K, Torigaki T and Taylor J 1986 Comput. Method Appl. Mech. Eng. 57 67–89
[5] Schleupen A, Maute K and Ramm E 2000 Struct. Multidisc. Optim. 19 282–302
[6] Kohn R V and Vogelius M 1987 Comm. Pure Appl. Math. 40 745–777
[7] Céa J 1986 Math. Model. Numer. Anal. 20 371–402
[8] Pantz O 2005 C. R. Acad. Sci. Paris Ser. I 333–337