



Quantitative a posteriori error estimators in Finite Element-based shape optimization

Matteo Giacomini

► To cite this version:

Matteo Giacomini. Quantitative a posteriori error estimators in Finite Element-based shape optimization. Numerical Analysis [cs.NA]. Université Paris Saclay (COMUE), 2016. English. NNT : 2016SACLX070 . tel-01418841v2

HAL Id: tel-01418841

<https://theses.hal.science/tel-01418841v2>

Submitted on 29 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2016SACLX070



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PARIS-SACLAY
PRÉPARÉE À L'ÉCOLE POLYTECHNIQUE

Ecole doctorale n°574
Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées
par

M. MATTEO GIACOMINI

Estimations d'erreur a posteriori quantitatives
pour l'approximation des problèmes d'optimisation de forme
par la méthode des éléments finis

Thèse présentée et soutenue à Palaiseau, le 9 Décembre 2016.

Composition du Jury :

M. ANTONIN CHAMBOLLE	Directeur de Recherche CNRS École polytechnique	(Examineur)
M. MARC DAMBRINE	Professeur des Universités Université de Pau et des Pays de l'Adour	(Rapporteur)
M. PIERRE DUYSINX	Professeur Université de Liège	(Rapporteur)
M. ALEXANDRE ERN	Professeur des Universités Université Paris-Est	(Président du jury)
M. OLIVIER PANTZ	Professeur des Universités Université Nice-Sophia Antipolis	(Directeur de thèse)
M. GIANLUIGI ROZZA	Professeur Associé SISSA International School for Advanced Studies	(Examineur)
M. KARIM TRABELSI	Enseignant-Chercheur Institut Polytechnique des Sciences Avancées	(Codirecteur de thèse)

THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription : École polytechnique

Laboratoire d'accueil : Centre de Mathématiques Appliquées de Polytechnique,
UMR 7641 CNRS

Spécialité de doctorat : Mathématiques appliquées

Matteo GIACOMINI

Quantitative a posteriori error estimators
in Finite Element-based shape optimization

Date de soutenance : 9 Décembre 2016

Après avis des rapporteurs : MARC DAMBRINE (Université de Pau et des Pays de l'Adour)
PIERRE DUYSINX (Université de Liège)

	ANTONIN CHAMBOLLE	(CNRS École polytechnique) Examineur
	MARC DAMBRINE	(Université de Pau) Rapporteur
	PIERRE DUYSINX	(Université de Liège) Rapporteur
<i>Jury de soutenance</i> :	ALEXANDRE ERN	(Université Paris-Est) Président du jury
	OLIVIER PANTZ	(UNS) Directeur de thèse
	GIANLUIGI ROZZA	(SISSA) Examineur
	KARIM TRABELSI	(IPSA) Codirecteur de thèse



Thèse financée par la **Direction de la Recherche et de l'Innovation** de l'**Institut Polytechnique des Sciences Avancées**.

63 Boulevard de Brandebourg
94200 Ivry-sur-Seine



Thèse préparée au sein du **Centre de Mathématiques Appliquées** de l'**École polytechnique** et de l'équipe-projet DeFI du Centre de Recherche **Inria Saclay**.

Route de Saclay
91128 Palaiseau CEDEX

Remerciements

Il y a un peu plus de trois ans, j'étais en train de clôturer ma thèse de Master avec une citation de Daniel Pennac : *J'ai toujours pensé que l'école, c'était d'abord les professeurs*. Après trois autres années passées à l'École, je suis encore plus convaincu de l'importance des professeurs à la fois comme enseignants et comme personnes. C'est pour ça que je tiens premièrement à remercier mes directeurs de thèse, Olivier Pantz et Karim Trabelsi. Ils m'ont donné la liberté de poursuivre mes idées, tout en restant toujours disponibles à discuter et à me fournir les moyens dont j'avais besoin pour m'orienter. Ils m'ont appris que simplement parce qu'une méthode donne des bons résultats n'implique pas qu'elle soit mathématiquement correcte ; et que si une méthode est mathématiquement correcte n'implique pas qu'elle donne des bons résultats dans des applications réelles. Mais surtout, ils m'ont transmis leur enthousiasme et leur passion pour la recherche et pour la vie (même en dehors du labo, cette dernière étant parfois un peu négligée dans le monde académique...).

Cette thèse a bénéficié aussi de deux collaborations. Juan-Antonio Escareno m'a introduit au fascinant monde des drones et m'a poussé à trouver une réponse simple à un problème compliqué. Je le remercie pour la confiance qu'il m'a témoignée lors de ce travail et pour me permettre de l'aider dans l'encadrement de Jesus Vazquez. Pendant l'édition 2015 du CEMRACS, j'ai eu l'opportunité de travailler sur la modélisation d'un phénomène industriel avec Guillaume Dollé, Omar Duran, Nelson Feyeux, Emmanuel Frénod et Christophe Prud'homme. Je suis gré à la Société des Mathématiques Appliquées et Industrielles (SMAI) pour le soutien et je tiens à remercier tous les membres du projet SiViBiR++ pour les discussions fructueuses et la bonne ambiance.

Un grand merci va à Marc Dambrine et à Pierre Duysinx pour avoir accepté de rapporter sur mon travail de thèse, pour leurs remarques pertinentes et pour me faire l'honneur de participer au jury de thèse. Je sais profondément gré à Antonin Chambolle d'avoir accepté de faire partie de mon jury de soutenance et surtout de m'avoir mis en contact avec Olivier et Karim quand j'étais en train de chercher un sujet de thèse. C'est là que cette aventure a commencé, merci. Last but not least, je remercie Alexandre Ern pour les très nombreuses discussions enrichissantes et Gianluigi Rozza pour l'intérêt qu'il a toujours montré vers ma recherche. Je les remercie aussi d'avoir accepté de participer comme examinateurs à ma soutenance de thèse.

Pendant ma thèse j'ai eu la chance de travailler dans plusieurs environnements, à partir du CMAP et de la DRI IPSA jusqu'au Laboratoire J.A. Dieudonné de l'Université de Nice-Sophia Antipolis qui m'a reçu pendant mon séjour niçois. Je remercie tous les membres de ces laboratoires pour m'avoir si bien accueilli. La liste des gens avec lesquelles j'ai pu discuter ou tout simplement passer une bonne pause-café étant incroyablement longue, parmi tous je me limite à remercier ici les membres de l'équipe DeFI (élargie) et mes riverains des bureaux 2015 au CMAP, D05 à l'IPSA et 709 au LJAD pour la bonne ambiance et la bonne humeur qu'ils ont su donner aux longues journées aux labos.

Un remerciement particulier est destiné aux équipes administratives, pour leur efficacité, gentillesse et pour être toujours disponibles à m'accueillir avec un sourire : merci Nasséra, Alex, Jessica, Manoella, Vincent et Wilfried (CMAP), Eric et Manon (IPSA), Julia et Chiara (LJAD). Je sais aussi extrêmement gré à Jean-Marc et Roland pour le soutien informatique et l'aide avec les machines de calcul de Nice. Un grand merci à toutes et tous ceux qui sont investis dans la vie des labos, de la gestion du café à l'organisation des séminaires des labos, des équipes, des thésards, 2.0, ... Une mention spéciale va aux directeurs des labos - Antonin, Anne (CMAP), Karim (IPSA) et Sorin (LJAD) - pour leur activité presque invisible qui nous permet quotidiennement de nous consacrer à notre recherche dans une ambiance très stimulante.

Infine, un grazie di cuore alla mia famiglia, alla mia piccola comunità italiana a Parigi e agli amici di una vita, quelli che ormai sono sparsi per il mondo ma che ci sono comunque sempre.

Matteo

Abstract

Gradient-based shape optimization strategies rely on the computation of the so-called shape gradient. In many applications, the objective functional depends on the solution of a PDE which can only be solved approximately (e.g. via the Finite Element Method). This is also the case for the shape gradient. Hence, the direction computed using the discretized shape gradient may not be a genuine descent direction for the objective functional. This Ph.D. thesis is devoted to the construction of a certification procedure for the descent direction in gradient-based shape optimization methods using *a posteriori* estimators of the error due to the Finite Element approximation of the shape gradient.

By means of a goal-oriented procedure, we derive a fully computable certified upper bound of the aforementioned error. The resulting Certified Descent Algorithm (CDA) for shape optimization is able to identify a genuine descent direction at each iteration and features a reliable stopping criterion based on the norm of the shape gradient.

Two main applications are tackled in the thesis. First, we consider the scalar inverse identification problem of Electrical Impedance Tomography and we investigate several *a posteriori* estimators. A first procedure is inspired by the complementary energy principle and involves the solution of additional global problems. In order to reduce the computational cost of the certification step, an estimator which depends solely on local quantities is derived via an equilibrated fluxes approach. The estimators are validated for a two-dimensional case and some numerical simulations are presented to test the discussed methods. A second application focuses on the vectorial problem of optimal design of elastic structures. Within this framework, we derive the volumetric expression of the shape gradient of the compliance using both the pure displacement and the dual mixed variational formulations of the linear elasticity equation. Some preliminary numerical tests are performed to minimize the compliance under a volume constraint in 2D using the Boundary Variation Algorithm and an *a posteriori* estimator of the error in the shape gradient is obtained via the complementary energy principle.

Keywords : Shape optimization, *a posteriori* error estimators, Certified Descent Algorithm (CDA), volumetric shape gradient, Electrical Impedance Tomography, compliance minimization

Résumé

Les méthodes de gradient en optimisation de forme reposent sur le calcul de la dérivée de forme. Dans beaucoup d'applications, la fonctionnelle coût dépend de la solution d'une EDP. Il s'en suit qu'elle ne peut être résolue exactement et que seule une approximation de celle-ci peut être calculée, par exemple par la méthode des éléments finis. Il en est de même pour la dérivée de forme. Ainsi, les méthodes de gradient en optimisation de forme - basées sur des approximations du gradient - ne garantissent pas a priori que la direction calculée à chaque itération soit effectivement une direction de descente pour la fonctionnelle coût. Cette thèse est consacrée à la construction d'une procédure de certification de la direction de descente dans des algorithmes de gradient en optimisation de forme grâce à des estimations a posteriori de l'erreur introduite par l'approximation de la dérivée de forme par la méthode des éléments finis.

On présente une procédure pour estimer l'erreur dans une Quantité d'Intérêt et on obtient une borne supérieure certifiée et explicitement calculable. L'Algorithme de Descente Certifiée (CDA) pour l'optimisation de forme identifie une véritable direction de descente à chaque itération et permet d'établir un critère d'arrêt fiable basé sur la norme de la dérivée de forme.

Deux applications principales sont abordées dans la thèse. Premièrement, on considère le problème scalaire d'identification de forme en tomographie par impédance électrique et on étudie différentes estimations d'erreur. Une première approche est basée sur le principe de l'énergie complémentaire et nécessite la résolution de problèmes globaux additionnels. Afin de réduire le coût de calcul de la procédure de certification, une estimation qui dépend seulement de quantités locales est dérivée par la reconstruction des flux équilibrés. Après avoir validé les estimations de l'erreur pour un cas bidimensionnel, des résultats numériques sont présentés pour tester les méthodes discutées. Une deuxième application est centrée sur le problème vectoriel de la conception optimale des structures élastiques. Dans ce contexte, on calcule l'expression volumique de la dérivée de forme de la compliance à partir de la formulation primale en déplacements et de la formulation duale mixte pour l'équation de l'élasticité linéaire. Quelques résultats numériques préliminaires pour la minimisation de la compliance sous une contrainte de volume en 2D sont obtenus à l'aide de l'Algorithme de Variation de la Frontière et une estimation a posteriori de l'erreur de la dérivée de forme basée sur le principe de l'énergie complémentaire est calculée.

Mots-clés : Optimisation de forme, estimations d'erreur a posteriori, Algorithme de Descente Certifiée (CDA), dérivée de forme volumique, Tomographie d'Impédance Electrique, minimisation de la compliance

Contents

Introduction (en français)	1
Introduction	11
The role of shape optimization in the aerospace and aeronautic industry	12
Non-destructive testing	12
Optimal design	13
Thesis outline	15
I Optimization of shape-dependent functionals and a posteriori error estimators	21
1 Shape optimization and shape identification problems	23
1.1 Abstract framework	23
1.1.1 A note on the ill-posedness of shape optimization problems	24
1.2 Geometrical description of the shape	25
1.3 Classification of shape optimization problems	26
1.4 A review of the numerical techniques to handle shape optimization problems	27
1.4.1 A moving mesh approach	27
1.4.2 The homogenization method	28
1.4.3 SIMP method: Simplified Isotropic Material with Penalization	29
1.4.4 A level-set approach	30
1.4.5 A phase-field model	32
1.4.6 Gradient-free strategies	32
1.5 Differentiation with respect to the shape	33
1.5.1 Hadamard's boundary variation method	33
1.5.2 Volumetric and surface expressions of the shape gradient	34
1.6 A posteriori error estimators for shape optimization	35
2 Certification procedure for a genuine descent direction	37
2.1 Optimize-then-Discretize and Discretize-then-Optimize	37
2.1.1 The Boundary Variation Algorithm	38
2.1.2 The discretized Boundary Variation Algorithm	39
2.1.3 Certification procedure for a genuine descent direction	40

2.2	Numerical error in the shape gradient	41
2.2.1	Bound for the error due to the approximation of a linear functional	41
2.2.2	Variational formulation of the error in the shape gradient	43
2.3	Estimate of the error in the shape gradient based on energy-norm error estimates . . .	43
2.3.1	Complementary energy principle	44
2.3.2	Explicit residual estimates	45
2.3.3	Implicit residual estimates	47
2.3.4	Equilibrated residual estimates	49
2.4	Improving the estimate of the error in the shape gradient	51
2.4.1	Dual Weighted Residual method	51
2.4.2	Energy-norm estimate via the parallelogram identity	52
2.4.3	Equilibrated fluxes and flux-free approaches	53
2.5	The Certified Descent Algorithm	54
2.5.1	CDA based on the complementary energy principle	55
2.5.2	CDA based on an equilibrated fluxes approach	56
2.5.3	The mesh adaptation procedure	57
 II Shape identification: an inverse problem in Electrical Impedance Tomography		59
3	Minimization of the Kohn-Vogelius functional	61
3.1	The problem of Electrical Impedance Tomography	61
3.2	A shape optimization approach	63
3.2.1	Shape gradient of the Kohn-Vogelius functional	63
4	Certified Descent Algorithm based on the complementary energy principle	65
4.1	Conforming Finite Element approximation	65
4.1.1	The state problems	66
4.1.2	The adjoint problems	66
4.2	Estimate of the error in the shape gradient via the complementary energy principle . .	66
4.2.1	Energy-norm error estimates for the state equations	67
4.2.2	Energy-norm error estimates for the adjoint equations	68
4.3	Numerical results	69
4.3.1	Numerical assessment of the goal-oriented estimator	69
4.3.2	1-mesh and 2-mesh reconstruction strategies	71
4.3.3	A more involved test case	74
4.3.4	The case of multiple boundary measurements	77
5	Certified Descent Algorithm based on an equilibrated fluxes approach	81
5.1	Estimate of the error in the shape gradient via the equilibrated fluxes	81
5.2	Equilibrated fluxes for a conforming Finite Element approximation	82
5.2.1	Equilibrated fluxes for the state equations	82
5.2.2	Equilibrated fluxes for the adjoint equations	84

5.2.3	Goal-oriented equilibrated fluxes error estimator	85
5.3	Discontinuous Galerkin approximation	86
5.3.1	Weak imposition of the essential boundary conditions	87
5.3.2	The state problems	88
5.3.3	The adjoint problems	90
5.4	Equilibrated fluxes for a Discontinuous Galerkin approximation	90
5.4.1	Equilibrated fluxes for the state equations	90
5.4.2	Equilibrated fluxes for the adjoint equations	91
5.4.3	Goal-oriented equilibrated fluxes error estimator	91
5.5	Numerical results	93
5.5.1	Numerical assessment of the goal-oriented estimator	93
5.5.2	Reconstruction of a single inclusion	95
5.5.3	The case of two inclusions featuring multiple boundary measurements	96
Summary of the results and prospective developments		99
 III Shape optimization: a forward problem in the design of elastic structures		
6	Minimization of the compliance under a volume constraint	103
6.1	The linear elasticity problem	103
6.1.1	The pure displacement variational formulation	105
6.1.2	Mixed variational formulations via the Hellinger-Reissner principle	106
6.1.3	A dual mixed variational formulation with weakly enforced symmetry of the stress tensor	108
6.2	Minimization of the compliance under a volume constraint	109
6.3	Volumetric shape gradient of the compliance via the pure displacement formulation	111
6.4	Volumetric shape gradient of the compliance via the dual mixed formulation	113
6.4.1	The case of strongly enforced symmetry of the stress tensor	116
6.4.2	The case of weakly enforced symmetry of the stress tensor	117
6.5	Qualitative assessment of the discretized shape gradients via numerical simulations	119
6.5.1	Experimental analysis of the convergence of the error in the shape gradient	120
6.5.2	Boundary Variation Algorithm using the pure displacement and the dual mixed formulations	121
6.5.3	A note on the moving mesh approach	126
7	Complementary energy-based estimate of the error in the shape gradient of the compliance	129
7.1	Discretization of the pure displacement formulation of the linear elasticity problem	129
7.1.1	The state problem	130
7.1.2	The adjoint problem	130
7.2	Estimate of the error in the shape gradient via the complementary energy principle	131
7.2.1	Energy-norm error estimates for the state equation	133
7.2.2	Energy-norm error estimates for the adjoint equation	134

7.3 A note on the Finite Element approximation of the complementary energy problem . .	135
Summary of the results and prospective developments	137
Conclusion	139
Conclusion (en français)	141
A Identification of optimal grasping points for an Unmanned Aerial Vehicle	145
B Modeling and simulation of a bioreactor landfill	153
Published and preprint articles	183
Bibliography	185

Introduction (en français)

L'optimisation de forme regroupe une classe des problèmes d'optimisation dont la fonctionnelle coût dépend de la forme du domaine et de la solution d'une équation aux dérivées partielles (EDP) qui y est formulée. Ainsi, on peut formuler ce problème en tant que problème d'optimisation d'une fonctionnelle qui dépend de la forme du domaine avec une contrainte représentée par une équation différentielle. De ce fait, le domaine peut être considéré comme la variable à optimiser dans le problème. Cette classe de problèmes a été abordée dans la littérature en utilisant à la fois des méthodes basées sur le gradient et des algorithmes d'optimisation sans dérivée (cf. [17]). Dans ce travail, on considère une stratégie qui fait partie du premier groupe et on calcule la dérivée de forme des fonctionnelles coût analysées.

Dans la plupart des applications, la différentielle de la fonctionnelle coût par rapport à la forme dépend de la solution d'une EDP qui ne peut qu'être résolue approximativement au moyen d'une stratégie de discrétisation telle que la méthode des éléments finis. L'approximation de l'équation qui décrit le phénomène analysé introduit une incertitude qui peut empêcher la dérivée de forme d'être strictement négative le long de la direction de descente identifiée. La direction calculée au moyen de la dérivée de forme discrétisée peut ne pas mener à une amélioration de la fonctionnelle à minimiser. De plus, à cause de l'approximation numérique, les critères d'arrêt basés sur la norme de la dérivée de forme ne peuvent pas être atteints si la tolérance fixée a priori est trop petite par rapport à la discrétisation choisie. Dans ce cas de figure, les estimations d'erreur a posteriori donnent des informations utiles pour améliorer les algorithmes d'optimisation de forme basés sur le gradient.

Un des aspects cruciaux dans les problèmes d'optimisation de forme est la définition d'un critère d'arrêt pour la procédure d'optimisation. A ce sujet, dans [253] Sigmund et Maute déclarent :

Beaucoup d'articles commentent les exemples par des déclarations du type : “la conception optimale est visuellement similaire aux résultats publiés dans la littérature”. Comme l'illustre l'exemple suivant, ceci n'est pas une conclusion satisfaisante. De plus, on devrait toujours éviter d'employer le terme “optimal” dans le contexte de l'optimisation des structures - à moins qu'on démontre que le problème est convexe. La figure 1 exhibe un exemple de conception d'un mécanisme tel que décrit dans [249, 250]. Les deux conceptions dans les figures 1a et 1b sont visuellement similaires, cependant, elles représentent la solution respectivement à l'itération 100 et 1562. Les fonctions coût correspondantes (déplacements en sortie) sont $-1,46$ et $-1,54$, respectivement, c'est-à-dire la conception finale est 5,5% meilleure que la conception obtenue après 100 itérations. La figure 1c montre la différence de densité entre les deux conceptions. Du point de vue des sciences de l'ingénieur, le 5,5% peut être sans intérêt pour une application spécifique. Cependant, du point de vue de l'optimisation, un

algorithme devrait toujours être capable de converger vers la meilleure conception. La différence entre les résultats est probablement due à l'utilisateur qui arrête la procédure prématurément à cause des contraintes de temps - pas aux failles de convergence de l'algorithme. D'où, la nécessité de mener des comparaisons quantitatives et non seulement qualitatives des résultats de la littérature lors de l'analyse d'un algorithme.

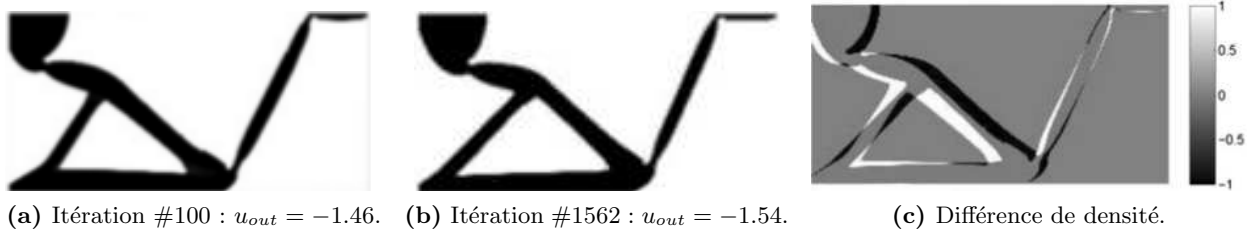



Figure 1 – Conception d'un mécanisme inverseur compliant à l'itération 100 et 1562. La fonctionnelle coût dans la conception finale (b) est 5,5% meilleure qu'à l'itération 100 (a). (c) Différence de densité entre les deux conceptions. Images extraites de [253].

Cette thèse est consacrée à l'étude de quelques aspects critiques dans la solution numérique des problèmes d'optimisation de forme, en particulier certains détails qui n'ont pas été pris en compte par la communauté scientifique jusqu'à présent. L'investigation des critères d'arrêt en optimisation de forme est motivée par la remarque que tous les algorithmes existants sont fortement dépendants des choix de l'utilisateur qui dans la plupart des cas choisit quand arrêter la procédure d'optimisation. Comme remarqué par Sigmund et Maute, beaucoup de résultats dans la littérature sont obtenus au moyen de stratégies d'optimisation arrêtées quand un résultat qualitativement raisonnable est obtenu. L'objectif principal de ce travail est de fournir un cadre rigoureux pour automatiser les algorithmes d'optimisation de forme basés sur le gradient. Plus précisément, on est intéressé par le calcul d'une véritable direction de descente pour la fonctionnelle coût à chaque itération de l'algorithme et par la définition d'un critère d'arrêt fiable pour la stratégie globale d'optimisation.

Le rôle de l'optimisation de forme dans l'industrie aéronautique et spatiale

Historiquement, les premiers exemples d'optimisation de forme se sont focalisés sur le domaine de l'optimisation de structures. Néanmoins, plusieurs applications industrielles peuvent être formulées dans le même cadre. En particulier, l'intérêt de l'industrie aéronautique et spatiale pour ce domaine couvre les problèmes classiques de conception optimale de structures décrite par l'équation de l'élasticité, la réduction de la traînée dans un contexte compressible/incompressible jusqu'au contrôle non-destructif et l'identification d'endommagement. A proprement parler, les deux premières applications sont des problèmes d'optimisation de forme tandis que les deux dernières sont des problèmes d'identification de forme. En général, des problèmes directs et des problèmes inverses peuvent à la fois être formulés comme des problèmes d'optimisation de fonctionnelles qui dépendent de la forme. Dans la suite de cette section, on présente des exemples des applications mentionnées ci-dessus en ingénierie, qu'on retrouve, en particulier, dans le domaine de l'industrie aéronautique et spatiale.



Figure 2 – Section transversale d'un fuselage d'Airbus A300. On remarque les trous circulaires dans la poutre. Image disponible sous licence Creative Commons 2.5 

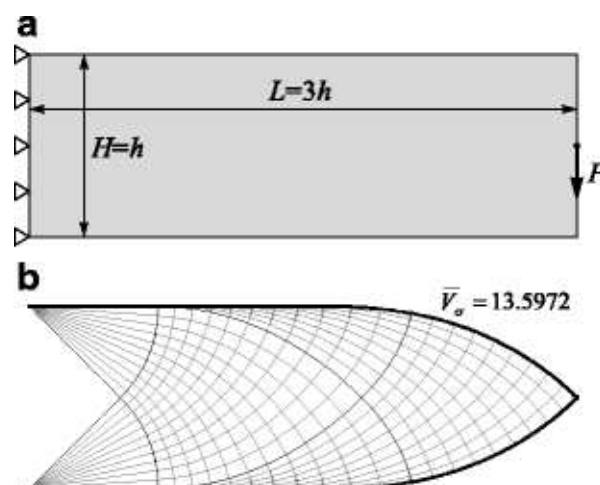


Figure 3 – Structure optimale exacte en treillis pour le problème d'une longue poutre. Solution calculée par Lewiński *et al.*. Image extraite de [190].

Contrôle non-destructif

Comme plusieurs autres secteurs, l'industrie aéronautique et spatiale utilise intensivement les techniques de contrôle non-destructif. Bien que la façon la plus simple et efficace pour évaluer la fiabilité d'une structure soit de la tester jusqu'à destruction, ceci n'est pas un choix réaliste quand on considère des composants de petite taille ou très chers. Outre l'étape de vérification pendant le processus de production, le contrôle non-destructif a aussi un rôle extrêmement important dans les procédures de maintenance des structures existantes. En effet, l'examen visuel n'est pas suffisant pour certifier sans risque la fiabilité d'un composant donné : d'une part, des défauts à la micro- ou nano-échelle sur la surface de la structure pourrait l'affaiblir sans être relevés par un observateur externe ; d'autre part, les examens visuels ne peuvent fournir aucune information sur des possibles dommages internes. Pour traiter ces difficultés, deux familles d'approches ont été développées au cours des années dans le domaine du contrôle non-destructif, qualifiées respectivement de surfaciques et sous-surfaciques.

Les méthodes surfaciques de contrôle non-destructif se concentrent sur l'analyse de la partie extérieure de la structure tandis que les techniques sous-surfaciques offrent la possibilité d'examiner des régions autrement inaccessibles sans démonter le composant analysé. La dernière approche est basée sur l'utilisation de mesures sur la surface extérieure de la structure pour récupérer des informations autour des phénomènes qui ont lieu à l'intérieur de l'objet. Un exemple de méthode de contrôle non-destructif est l'imagerie par ultrason pour la détection de petites fissures et pour l'évaluation de l'amincissement du matériau, par exemple à cause de la corrosion de la peau d'un avion [125]. D'autres méthodes exploitent les rayons X pour identifier des défauts dans un composant et pour détecter la présence anormale de fluides à l'intérieur d'une structure [155]. Dans cette thèse, on considère un cas particulier de technique sous-surfacique - notamment, la tomographie par impédance électrique - qui utilise un potentiel électrique pour identifier des inhomogénéités dans le coefficient de conductivité à l'intérieur du matériau. Cette technique a été récemment appliquée effi-

cacement au problème de la détection de fissures transversales dans un matériau composite stratifié [245].

Conception optimale de structures

La conception de la forme optimale d'une structure est un problème qui apparaît fréquemment dans l'industrie aéronautique et spatiale. L'intérêt pour ce sujet est notamment motivé par des enjeux économiques : une structure plus légère a besoin d'une moindre quantité de carburant pour sa propulsion réduisant conséquemment son coût d'exploitation. Ainsi, la conception de structures compliantes et légères a été un sujet majeur de recherche dans l'industrie pendant plusieurs années et représente encore une piste d'investigation très active.

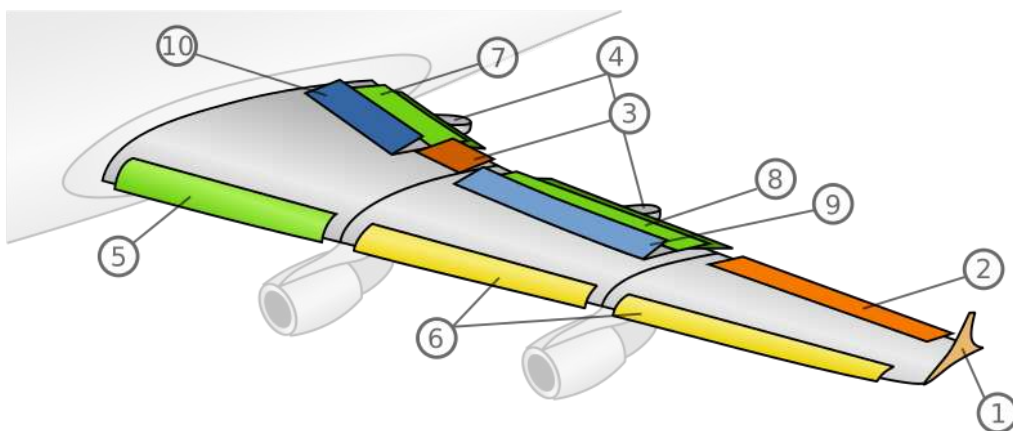


Figure 4 – Schéma de l'aile d'un avion. (1) Winglet ou ailerette. (2) Aileron basse vitesse. (3) Aileron haute vitesse. (4) Rail de glissement des volets. (5) Bec de bord d'attaque de type Krüger. (6) Bec de bord d'attaque de type slats. (7) Volet intérieur de type Fowler. (8) Volet extérieur de type Fowler. (9) Spoilers. (10) Spoilers (destructeurs de portance et aérofrein). Image disponible sous licence Creative Commons 3.0

Une première application d'optimisation de forme en aéronautique est représentée par la conception de la forme optimale minimisant la compliance d'une structure élastique soumise à une force extérieure et sous une contrainte de volume. D'un point de vue pratique, ce problème peut être appliqué à l'optimisation de certaines pièces à l'intérieur d'un avion, par exemple la structure longitudinale portante d'un avion. En fait, plusieurs experts sont d'accord sur le fait que le poids des conceptions actuellement construites par des industries leader du secteur n'est pas optimal et des estimations officielles rapportent un surpoids approximativement de 6000 ~ 11000 kg pour différents modèles d'avions. De plus, l'approche actuellement utilisée pour réduire le poids de la structure - quand appliquée - est basée sur l'insertion de trous circulaires dans un matériau en vrac (Fig. 2) alors qu'il est bien connu dans la littérature [198] que les conceptions optimales pour ce type de problème sont des structures en treillis (Fig. 3).

D'autres applications portent sur la conception optimale d'une aile. C'est un problème multiphysique extrêmement complexe impliquant la dynamique de l'air, le comportement de la structure élastique et de leur interaction - d'où, la notion de problème d'interaction fluide-structure. Le mouvement de l'air est décrit par les équations de Navier-Stokes compressibles instationnaires en cas de

régime turbulent, tandis que la déformation de la structure est régie par les équations de l'élasticité non-linéaire. L'ensemble des équations qu'on obtient est le point de départ pour la définition des contraintes dans la procédure d'optimisation. Le problème d'optimisation vise à minimiser la traînée ou, plus généralement, le coût de propulsion du véhicule. En général, lorsqu'on considère des problèmes multi-physiques complexes, on est intéressé par l'optimisation de plusieurs fonctionnelles qui peuvent être issues de disciplines différentes. Ce domaine de recherche est connu sous le nom d'optimisation multiobjectifs (MOO) et a été largement étudié dans la littérature. Plusieurs méthodes ont été développées pour gérer à la fois la phase coopérative et celle non-coopérative du processus d'optimisation. On fait référence à [119, 152] et aux références qu'ils contiennent pour une vue d'ensemble du sujet et pour la présentation de quelques techniques récentes pour la résolution de ce problème. En raison de la nature différente des phénomènes physiques à l'étude, le problème d'optimisation associé peut être aussi classé comme un problème d'optimisation multidisciplinaire (MDO) [197]. De plus, en raison des conditions variables auxquelles la structure est soumise pendant le décollage, la croisière et l'atterrissage, plusieurs études sont réalisées pour identifier la forme qui améliore le plus la fonctionnelle coût en même temps pour toutes les conditions évoquées ci-dessus. La solution de ce problème d'optimisation multidisciplinaire multipoint [173] a une application directe dans la gestion de différentes conceptions d'aile pendant les phases de décollage, de croisière et d'atterrissage au moyen de changements de la position des ailerons, lamelles et spoilers (Fig. 4 et 5).

Dans cette thèse, on se limite à l'analyse du premier exemple présenté dans cette section et on considère un problème purement mécanique dans lequel on vise à minimiser la compliance d'un composant sous une contrainte de volume.

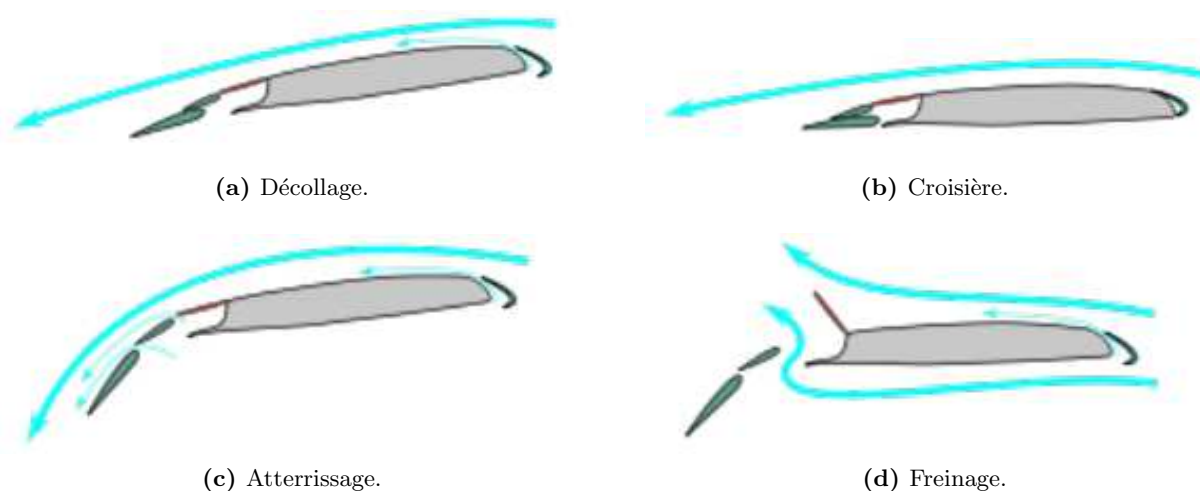





Figure 5 – Conception d'une aile pour différentes phases de vol. (a) Décollage - Surface de l'aile augmentée. (b) Croisière - Meilleure efficacité. (c) Atterrissage - Portance maximale et traînée élevée. (d) Freinage - Traînée maximale et portance réduite. Image disponible sous licence Creative Commons 3.0   

Plan de la thèse

Cette thèse est composée de sept chapitres - groupés en trois parties - et deux annexes.

À partir des remarques présentées au début de cette introduction, dans la première partie de la thèse (cf. partie I - chapitres 1 et 2), on développe une stratégie de résolution numérique des problèmes d'optimisation de forme en utilisant une procédure de certification. L'idée fondamentale de cette procédure consiste à calculer une borne supérieure de l'erreur introduite par l'approximation de la dérivée de forme pour vérifier à chaque itération que la direction calculée utilisant la dérivée de forme discrétisée est une véritable direction de descente pour la fonctionnelle coût. On néglige la contribution de l'approximation de la géométrie et on se focalise sur l'erreur due à la discrétisation des équations qui décrivent le phénomène analysé.

On introduit une nouvelle stratégie d'optimisation de forme - nommée Algorithme de Descente Certifiée (CDA) - qui génère une suite minimisante de formes au moyen d'une procédure qui certifie à chaque itération que la direction calculée utilisant la dérivée de forme est une véritable direction de descente pour la fonctionnelle coût et s'arrête automatiquement lorsqu'un critère d'arrêt fiable est vérifié. La nouveauté de cette approche repose sur la procédure de certification de la direction de descente, pour laquelle des estimations d'erreur a posteriori explicitement calculables sont requises. En effet, à ma connaissance tous les travaux dans la littérature qui utilisent des éléments finis adaptatifs pour l'optimisation de forme se focalisent sur l'information qualitative fournie par les estimations d'erreur pour effectuer de l'adaptation de maillage et ils n'exploitent pas l'information quantitative contenue dans les estimations pour améliorer et automatiser la stratégie globale d'optimisation.

Ensuite, on analyse une application qui peut être reformulée en tant que problème d'optimisation d'une fonctionnelle qui dépend de la forme sous une contrainte représentée par une EDP et on présente les détails de la procédure de certification mentionnée ci-dessus. En particulier, dans la deuxième partie (cf. partie II - chapitres 3, 4 et 5), on considère une application à la tomographie par impédance électrique, pour laquelle l'identification d'inclusions peut s'envisager comme un problème d'optimisation de forme. Pour construire l'estimation d'erreur a posteriori explicitement calculable requise par la procédure de certification, on propose deux approches. La première est basée sur le principe de l'énergie complémentaire et nécessite la résolution d'un problème dual de calcul des flux. Afin de réduire le coût de calcul de la procédure de certification, on introduit une deuxième stratégie qui n'utilise que des quantités locales pour reconstruire des flux équilibrés à la fois pour une discrétisation par éléments finis conformes et par Galerkin discontinu selon une procédure décrite par Ern et Vohralík (cf. [132]) pour la dérivation des estimations a posteriori. On met l'accent sur cette dernière approche pour l'étude du problème de la tomographie par impédance électrique à cause d'un intérêt croissant dans la communauté scientifique pour une classe particulière de discrétisation Galerkin discontinu - connues sous le nom de Galerkin discontinu symétrique avec pénalité intérieure pondérée - pour des problèmes où le tenseur de diffusion est inhomogène comme celui qui apparaît dans la tomographie par impédance électrique.

Finalement, dans la troisième et dernière partie (cf. partie III - chapitres 6 et 7), on considère une autre application et on introduit un problème direct en optimisation de forme, c'est-à-dire la minimisation de la compliance sous une contrainte de volume en mécanique des structures. En ce qui concerne cette application, on dérive l'expression volumique de la dérivée de forme à partir à la fois de la formulation primale en déplacements et de la formulation duale mixte pour l'élasticité

linéaire. Une comparaison qualitative des deux expressions de la dérivée de forme et de leur application à l'Algorithme de Variation de la Frontière est présentée à l'aide de simulations numériques. Enfin, on calcule une estimation a posteriori de l'erreur de discrétisation pour la dérivée de forme de la compliance au moyen du principe de l'énergie complémentaire pour la formulation primale en déplacements. Cette estimation représente le point de départ pour l'application de l'Algorithme de Descente Certifiée préalablement présenté à la conception optimale de structures élastiques qui fera l'objet du prochain article [5].

À la fin de la thèse (cf. annexe A et B), on présente deux travaux complètement indépendants du sujet principal de cette recherche. Ces travaux sont focalisés sur l'analyse de systèmes complexes qui décrivent des applications réelles par l'intermédiaire de modèles mathématiques. Les deux annexes traitent de problèmes extrêmement différents mais partagent la motivation de la modélisation de contraintes pratiques en ingénierie.

Dans la suite, on décrit brièvement le contenu de la thèse chapitre par chapitre.

Chapitre 1 - Dans ce chapitre introductif, on présente le cadre pour un problème d'optimisation de forme. Après avoir résumé les approches plus communes proposées dans la littérature pour traiter cette classe de problèmes d'un point de vue numérique, on se focalise sur des méthodes basées sur le gradient et on rappelle la notion de dérivée de forme. Finalement, on présente les stratégies obtenues à partir du couplage des algorithmes d'optimisation de forme basés sur le gradient avec les estimations a posteriori de l'erreur de discrétisation.

Chapitre 2 - On introduit l'approximation du problème d'optimisation de forme préalablement discuté par la méthode des éléments finis. Pour faire face au manque d'un critère d'arrêt et aux possibles problèmes de convergence des méthodes de gradient, on prend en compte l'erreur numérique introduite par la discrétisation de la dérivée de forme. On présente un résumé des différentes techniques proposées dans la littérature pour estimer l'erreur d'une Quantité d'Intérêt due à l'approximation d'une équation aux dérivées partielles par la méthode des éléments finis. Grâce à l'information fournie par cette estimation, on introduit une procédure de certification qui vérifie que la direction calculée utilisant la dérivée de forme discrétisée est une véritable direction de descente pour la fonctionnelle coût. Finalement, on définit une nouvelle stratégie d'optimisation de forme, nommée Algorithme de Descente Certifiée (CDA). Le contenu de ce chapitre est partiellement basé sur les travaux [2, 3, 4].

Chapitre 3 - On considère un premier exemple de problème d'optimisation d'une fonctionnelle qui dépend de la forme sous une contrainte représentée par une EDP. Plus précisément, on étudie un problème issu de la tomographie par impédance électrique pour laquelle l'identification d'inclusions peut s'envisager comme un problème d'optimisation de forme. On analyse un solide qui présente une inhomogénéité interne du coefficient de conductivité. Le but du problème est l'identification de la forme et de la position de cette inclusion au moyen de mesures effectuées sur le bord extérieur. On suppose que les données sont continues sur la frontière et on considère le modèle classique caractérisé par des électrodes ponctuelles. Grâce à la fonctionnelle de Kohn-Vogelius, le problème d'identification de l'inclusion à partir des données au bord peut s'assimiler à un problème d'optimisation de forme. Finalement, on rappelle les expressions volumiques et surfaciques de la dérivée de forme de la fonctionnelle susdite. Le contenu de ce chapitre est basé sur les travaux [2, 3, 4].

Chapitre 4 - On applique l’Algorithme de Descente Certifiée à la minimisation de la fonctionnelle de Kohn-Vogelius. On introduit une approximation du problème de tomographie par impédance électrique au moyen d’une discrétisation par éléments finis conformes. Une stratégie pour construire une estimation a posteriori explicitement calculable et sans constantes inconnues de l’erreur de la dérivée de forme est discutée. À partir du principe de l’énergie complémentaire, on définit des problèmes variationnels additionnels pour évaluer précisément l’erreur dans le calcul des flux numériques, à partir desquels on obtient une estimation de l’erreur pour notre Quantité d’Intérêt. On valide l’estimation à l’aide d’un cas test pour lequel la solution analytique est connue et on applique la stratégie d’optimisation de forme résultante à un problème bidimensionnel issu de la tomographie par impédance électrique. Une discussion des résultats numériques est présentée. Le contenu de ce chapitre est basé sur les travaux [3, 4].

Chapitre 5 - On propose une variante de l’Algorithme de Descente Certifiée qui utilise uniquement des quantités locales pour estimer l’erreur d’approximation de la dérivée de forme. En particulier, on construit une estimation dans une Quantité d’Intérêt basée sur la reconstruction des flux équilibrés. Un cadre générique valide pour une discrétisation par éléments finis conformes et Galerkin discontinu est présenté. L’estimation a posteriori obtenue est explicitement calculable et donne lieu à une borne supérieure certifiée de l’erreur de notre Quantité d’Intérêt. Les avantages principaux de cette approche sont le coût de calcul extrêmement faible de la procédure et la possibilité de paralléliser facilement l’algorithme. Après avoir validé les estimations d’erreurs, cette variante de l’Algorithme de Descente Certifiée est testée sur le problème d’identification de forme issu de la tomographie par impédance électrique préalablement discuté. Le contenu de ce chapitre est basé sur l’article [2].

Chapitre 6 - On considère un deuxième exemple de problème d’optimisation d’une fonctionnelle qui dépend de la forme sous une contrainte représentée par une EDP. On introduit un problème direct issu de la conception optimale d’une structure élastique décrit par les équations de l’élasticité linéaire. En particulier, on considère la minimisation de la compliance sous une contrainte de volume et on étudie deux formulations différentes. Premièrement, on analyse la formulation primale en déplacements pour l’élasticité linéaire et on calcule l’expression volumique de la dérivée de forme de la compliance. Ensuite, on considère une formulation duale mixte de l’équation de l’élasticité et on calcule l’expression correspondante de la dérivée de forme pour le problème étudié. Les deux expressions de la dérivée de forme sont examinées en utilisant l’Algorithme de Variation de la Frontière et une comparaison qualitative pour évaluer leur validité est effectuée à l’aide des simulations numériques. Le contenu de ce chapitre est basé sur l’article [6].

Chapitre 7 - À partir du cadre présenté dans le chapitre précédent pour la formulation primale en déplacements du problème de l’élasticité linéaire, on construit une estimation a posteriori de l’erreur dans la dérivée de forme. On utilise la stratégie discutée dans le chapitre 4 pour le problème de la tomographie par impédance électrique et on applique le principe de l’énergie complémentaire pour obtenir une borne supérieure certifiée et explicitement calculable de l’erreur effectuée sur le calcul de notre Quantité d’Intérêt. On introduit un problème variationnel additionnel pour évaluer précisément l’erreur dans le calcul du tenseur des déformations, c’est-à-dire la partie symétrique du gradient du champ de déplacement. Grâce à la solution du problème dual, on obtient une estimation a posteriori

sans constante inconnue de l'erreur dans la dérivée de forme. Le contenu de ce chapitre est basé sur l'article [5].

Un résumé des résultats obtenus et une présentation des possibles perspectives de développements futurs clôturent cette thèse. Les parties II et III de cette thèse sont basées sur des notions introduites dans la partie I. À l'exception de cette contrainte, les deux dernières parties peuvent être lues de façon indépendante mais on recommande au lecteur de suivre l'ordre proposé par l'auteur.

Annexe A - Ce travail est consacré à l'application des techniques d'optimisation au problème d'identification des points optimaux pour agripper un objet au moyen d'une pince embarquée sur un drone. À partir des images acquises par une caméra embarquée, on utilise les informations géométriques sur le bord de l'objet à saisir pour identifier les points de contact optimaux entre la pince et la cible. Dans cette annexe, on joint l'acte de conférence [7].

Annexe B - Ce travail concerne la modélisation mathématique et la simulation numérique d'un bioréacteur, c'est-à-dire un site pour la gestion des déchets où on exploite des matériaux biodégradables pour produire du méthane. Après avoir analysé les phénomènes chimiques ayant lieu à l'intérieur du bioréacteur et la dynamique des fluides impliqués, on propose un modèle qui inclut une équation de la chaleur couplée avec une équation de réaction non-linéaire pour décrire le phénomène de méthanogenèse étudié et des équations d'advection et d'advection-diffusion qui représentent les écoulements multiphasiques dans le milieu poreux composé par les déchets. Quelques simulations numériques préliminaires obtenues par l'approximation du modèle décrit ci-dessus par la méthode des éléments finis sont présentées. Dans cette annexe, on joint l'article [1].

Articles publiés et preprint

Le contenu principal de cette thèse a été développé en collaboration avec Olivier Pantz (Professeur de Mathématiques Appliquées à l'Université Nice-Sophia Antipolis) et Karim Trabelsi (Enseignant-Chercheur en Mathématiques Appliquées à l'Institut Polytechnique des Sciences Avancées). Les résultats ont été publiés dans l'acte de conférence avec comité de lecture [3] et dans les articles suivants publiés ou preprints [2, 4, 5, 6].

L'annexe A est un travail en collaboration avec Juan-Antonio Escareno (Enseignant-Chercheur en Systèmes Robotiques Aériens à l'Institut Polytechnique des Sciences Avancées) et ses collaborateurs et a été publié dans l'acte de conférence avec comité de lecture [7].

L'annexe B est un travail en collaboration avec Guillaume Dollé (Doctorant à l'Université de Strasbourg), Omar Duran (Doctorant à l'Universidade Estadual de Campinas), Nelson Feyeux (Doctorant à l'Université Joseph Fourier), Emmanuel Frénod (Professeur de Mathématiques Appliquées à l'Université de Bretagne-Sud) et Christophe Prud'homme (Professeur de Mathématiques Appliquées à l'Université de Strasbourg). L'article associé à ce travail a été accepté pour publication [1].

Introduction

The term shape optimization refers to a class of optimization problems in which the objective functional depends on the shape of the domain in which a Partial Differential Equation (PDE) is formulated and on the solution of the PDE itself. Thus, we may view these problems as PDE-constrained optimization problems of a shape-dependent functional, the domain being the optimization variable and the PDE being the constraint. This class of problems has been tackled in the literature using both gradient-based and gradient-free methods (cf. [17]) and in this work we consider a strategy issue of the former group by computing the so-called shape gradient.

In most applications, the differential form of the objective functional with respect to the shape depends on the solution of a PDE which usually can only be solved approximately by means of a discretization strategy like the Finite Element Method. The approximation of the governing equation for the phenomenon under analysis introduces an error which may prevent the shape gradient from being strictly negative along the computed descent direction, that is, the discretized direction may not lead to any improvement of the objective functional we are trying to optimize. Moreover, due to the aforementioned approximation, stopping criteria based on the norm of the shape gradient may never be fulfilled if the *a priori* given tolerance is too small with respect to the chosen discretization.

One of the crucial aspects in shape optimization problems is the definition of a criterion to stop the optimization procedure. On this subject, we report here a statement by Sigmund and Maute [253]:

Many papers conclude examples by statements like: “the optimal design is visually similar to results published in the literature”. As the following example will show this is not a satisfactory conclusion. Also, one should always avoid using the term “optimal” in the field of structural optimization - unless one can prove that the problem is convex. Figure 6 shows an example of mechanism design as described in [249, 250]. The two designs in Fig. 6a and 6b are visually similar, however, they represent snapshots at the 100s and the 1562nd iteration. The corresponding objective functions (output displacements) are -1.46 and -1.54 , respectively, i.e. the final design is 5.5% better than the design obtained after 100 iterations. Figure 6c shows the difference in densities between the two designs. Seen from an engineering perspective the 5.5% may be unimportant for a specific application. However, seen from an optimization perspective, an algorithm should always be able to converge to the better design. It must be the user who stops the procedure prematurely due to time constraints - not convergence flaws of the algorithm. Hence, when testing an algorithm, quantitative and not only qualitative comparisons to results from the literature should be performed.

This thesis is devoted to the study of some critical issues in the numerical solution of shape optimization problems, that have been so far overlooked by the scientific community. Our interest is

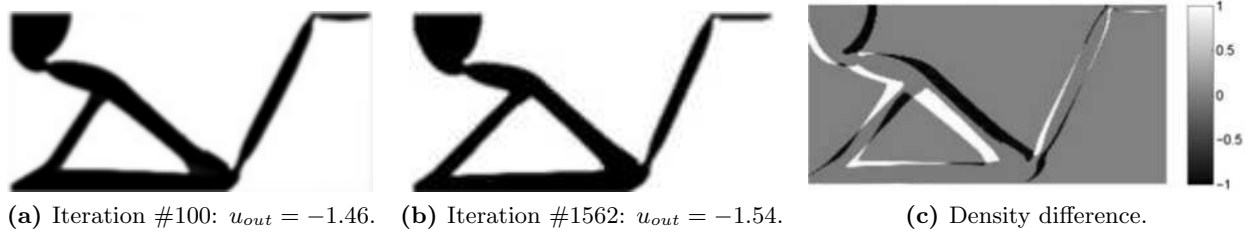


Figure 6 – Snapshots of compliant inverter design at iteration 100 and 1562. The objective function of the final design (b) is 5.5% better than at 100 iterations (a). (c) Density difference between the two designs. Images extracted from [253].

driven by the remark that all existing algorithms to perform shape optimization are highly-dependent on the end user who is, in most cases, the one to choose when to stop the optimization procedure. As remarked by Sigmund and Maute, many results in the literature are obtained via optimization strategies stopped when a qualitatively reasonable result is achieved. The main goal of this work is to provide a rigorous framework to automatize gradient-based shape optimization procedures. In particular, we are interested in the computation of a genuine descent direction for the objective functional at each iteration of the algorithm and in the definition of a reliable stopping criterion for the overall optimization strategy. Within this framework, *a posteriori* estimators of the error introduced by the Finite Element approximation of the problem provide useful information to improve gradient-based algorithms for shape optimization.

The role of shape optimization in the aerospace and aeronautic industry

Though the original examples studied within the shape optimization community dealt with structural optimization, several other industrial applications may be formulated within the same framework. In particular, the interest in the aerospace and the aeronautic industry spans classical optimal design problems governed by the elasticity equations, drag reduction in a compressible/incompressible context as well as non-destructive testing and damage identification. The two former applications are strictly-speaking shape optimization problems whereas the latter ones are shape identification problems, that is both forward and inverse problems may be modeled as optimization problems featuring shape-dependent functionals. In the following, we present some examples of the aforementioned engineering applications arising in the aerospace and aeronautic industry.

Non-destructive testing

As many other sectors, the aerospace and aeronautic industry makes extensive use of non-destructive testing techniques. Though the simplest and most effective way to assess the reliability of a structure is to test it to destruction, this is not a feasible option when dealing with low-volume or very expensive components. Besides the preliminary tests during the production process, non-destructive techniques are extremely important in the maintenance of the existing structures. As a matter of fact, visual



Figure 7 – Cross-section of an Airbus A300 fuselage. We remark the circular holes in the cantilever. Image released under Creative Commons license 2.5

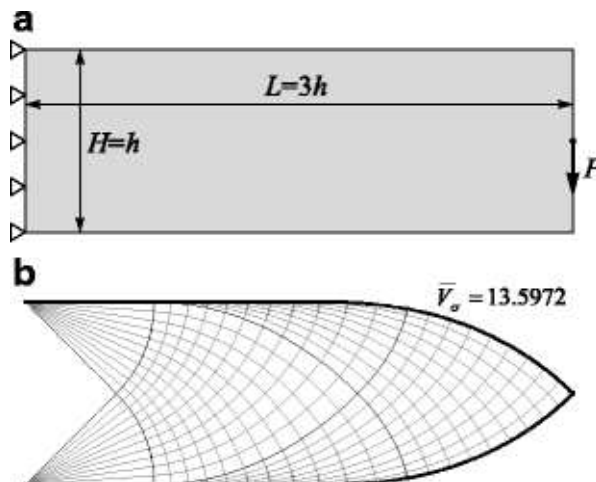


Figure 8 – Exact optimal truss-like structure for a long cantilever problem. Solution computed by Lewiński *et al.*. Image extracted from [190].

inspection is not sufficient to safely confirm the reliability of a given component: on the one hand, micro- or nano-defects on the surface of the structure may weaken it without being remarked by an external observer; on the other hand, visual inspections are not able to provide any insight on possible internal damages. To handle these issues, two families of approaches have been developed over the years within the field of non-destructive testing, that is surface and sub-surface techniques.

As their names suggest, surface non-destructive testing methods focus on the analysis of the external part of the structure, whereas sub-surface techniques allow to inspect regions otherwise unreachable without disassembling the component under analysis. The latter approach relies on the use of specific measurements on the external surface of the structure to retrieve information on phenomena taking place inside it. An example of non-destructive testing technique is the ultrasound measurement for the detection of small cracks and for the estimation of material thinning, e.g. due to the corrosion of the material skin of an aircraft [125]. Other methods exploit X-Ray techniques to identify defects buried within a component and to detect the anomalous presence of fluid inside a structure [155]. In this thesis, we consider a specific case of sub-surface technique - namely Electrical Impedance Tomography (EIT) - which relies on the use of the electrical potential to identify conductivity inhomogeneities inside the material. This technique has been recently proved to be extremely efficient in the very interesting problem of detecting transverse cracks in laminated composites [245].

Optimal design

The design of the optimal shape of a structure is a problem arising frequently in the aerospace and aeronautic industry. By intuition, the interest in this subject is motivated by an economic purpose: the lighter the structure, the smaller the quantity of fuel required for its propulsion and consequently the cost of moving it. Hence, the ability of constructing compliant elastic structures concurrently minimizing their weight has been a key topic of research in the industry for years.

A first application of shape optimization in aeronautics arises within the field of structural me-

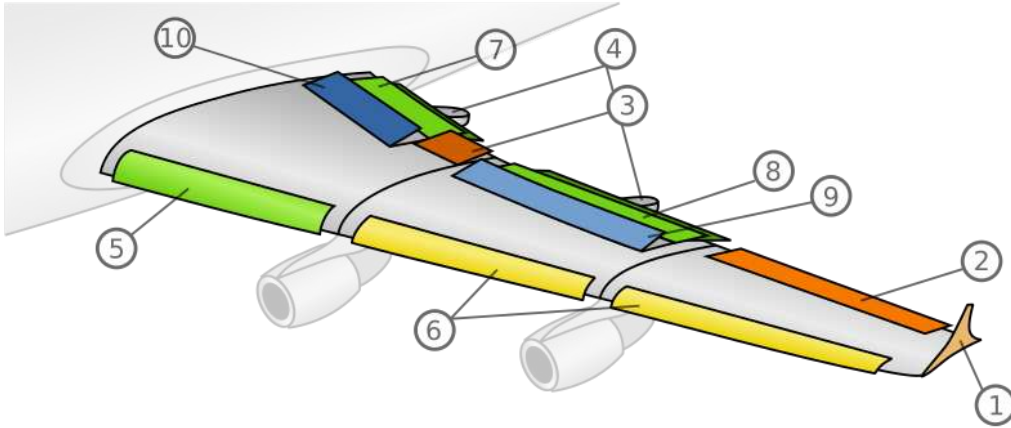


Figure 9 – Schematics of an aircraft wing. (1) Wingtip. (2) Low Speed Aileron. (3) High Speed Aileron. (4) Flap track fairing. (5) Krüger flaps. (6) Slats. (7) Three slotted inner flaps. (8) Three slotted outer flaps. (9) Spoilers. (10) Spoilers Air-brakes. Image released under Creative Commons license 3.0

chanics and focuses on the design of a shape that minimizes the compliance of an elastic component subject to a given external load under a volume constraint. From a practical point of view, this problem may be applied to the optimization of the internal parts of an aircraft, e.g. its longitudinal supporting structure. As a matter of fact, several experts in the field agree in stating that the weight of designs currently manufactured by world-leading industries is not optimal and unofficial estimates report an overweight by approximately 6000 ~ 11000 kg for different aircraft models. Moreover, the existing approach to reduce the weight of the structure - when applied - relies on the insertion of circular holes within a bulk material (Fig. 7) whereas it is well-known in the literature [198] that the optimal configurations are truss-like structures as the one in figure 8.

Another application focuses on the optimal design of a wing. This is an extremely complex multi-physics problem involving the dynamics of air, the behavior of the elastic structure and their interplay - whence, the notion of Fluid-Structure Interaction problem. The movement of air is described by unsteady compressible Navier-Stokes equations in a turbulent regime, whereas the deformation of the structure is governed by the non-linear elasticity equations. The resulting problem is the starting point for the definition of the constraints of an optimization procedure in which we aim to minimize the drag or, more generally, the cost of the propulsion. In general, when considering complex multiphysics problems, several cost functionals - possibly arising from different disciplines - may be of interest in the optimization process. This research field is known as Multi-Objective Optimization (MOO) and has been extensively studied in the literature via the development of methods to handle both the cooperative and the competitive phases of the optimization procedure. We refer to [119, 152] and references therein for an overview of the subject and for the discussion of some state-of-the-art approaches to this problem. Owing to the different nature of the physical phenomena under analysis, the corresponding optimization problem is also said to belong to the field of Multi-Disciplinary Optimization (MDO) [197]. Moreover, due to the variable conditions to which the structure is subject during take-off, cruise and landing, several studies have been performed to identify the shape that most improves the objective functional in all the aforementioned configurations at the same time. The solution of this so-called Multi-Point Multi-Disciplinary Optimization problem [173] has a straightforward application in the

management of different wing configurations during take-off, cruise and landing phases by means of appropriate changes in the position of flaps, slats and spoilers (Fig. 9 and 10).

In this thesis, we restrict to the former example discussed in this section and we consider a purely mechanical problem in which we aim to minimize the compliance of a component under a volume constraint.

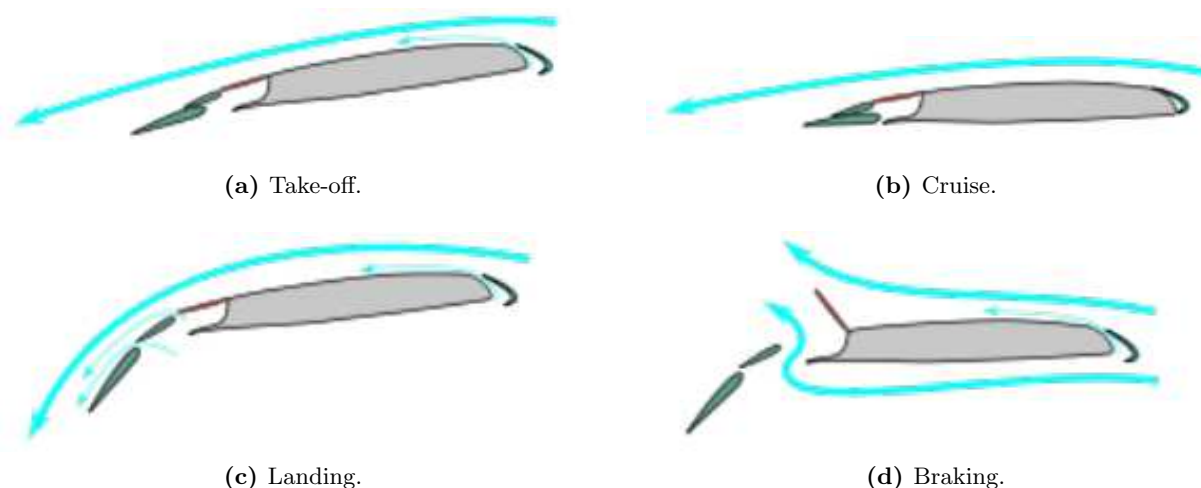





Figure 10 – Wing configurations for different flight phases. (a) Take-off - Increased wing area. (b) Cruise - Best efficiency. (c) Landing - Maximum lift and high drag. (d) Braking - Maximum drag and reduced lift. Image released under Creative Commons license 3.0   

Outline of the thesis

This thesis is composed of seven chapters - grouped into three parts - and two appendices.

Owing to the remarks presented at the beginning of this introduction, in the first part of this thesis (cf. part I - chapters 1 and 2) we develop a strategy to numerically solve shape optimization problems based on a certification procedure. Basic idea relies on the derivation of an upper bound of the error due to the approximation of the shape gradient to verify at each iteration whether the direction computed using the discretized shape gradient is a genuine descent direction for the objective functional or not. We neglect the contribution of the approximation of the geometry and we focus on the error arising from the discretization of the governing equations.

We propose a novel shape optimization strategy - named Certified Descent Algorithm (CDA) - that generates a minimizing sequence of shapes by certifying at each iteration the descent direction to be genuine and automatically stops when a reliable stopping criterion is fulfilled. The novelty of this approach resides in the certification procedure for the descent direction, for which fully computable *a posteriori* error estimators are required. As a matter of fact, to the best of our knowledge all the works in the literature on Adaptive Finite Element for shape optimization focus on the qualitative information provided by the error estimators to drive mesh adaptation and do not exploit the quantitative information to improve and automatize the overall optimization strategy.

We subsequently analyze an application that can be recast as a PDE-constrained optimization problem featuring a shape-dependent functional and we provide the details associated with the aforementioned certification procedure. In particular, in the second part (cf. part II - chapters 3, 4 and 5), we consider an application to Electrical Impedance Tomography: as a matter of fact, the identification of the inclusion in a given material starting from boundary measurements may be formulated as a shape optimization problem. To construct the required fully computable *a posteriori* error estimator, we follow two approaches. First, we develop an estimator based on the complementary energy principle to retrieve a good approximation of the numerical fluxes by solving a dual problem. In order to reduce the computational cost of the certification procedure, we introduce a strategy which relies solely on local quantities to construct the equilibrated fluxes for both conforming Finite Element and Discontinuous Galerkin discretizations, as proposed by Ern and Vohralík in [132]. We especially highlight the interest of this latter approach for the study of the problem of Electrical Impedance Tomography owing to the growing interest in the scientific community towards a special class of Discontinuous Galerkin methods - known as Symmetric Weighted Interior Penalty Discontinuous Galerkin - for problems featuring an inhomogeneous diffusion tensor as the one appearing in the EIT.

Eventually in the third and last part (cf. part III - chapters 6 and 7), we consider another application by introducing a forward problem in shape optimization, that is the minimization of the compliance under a volume constraint in structural mechanics. Concerning this application, we derive the volumetric expression of the shape gradient using both the pure displacement and the dual mixed variational formulations of the linear elasticity problem. A qualitative comparison of the results obtained by the application of the Boundary Variation Algorithm to the minimization of the compliance using the two expressions of the shape gradient is presented via some numerical tests. Eventually, we derive an *a posteriori* estimator of the discretization error for the shape gradient via the complementary energy principle for the pure displacement formulation. This estimate is the starting point for the application of the previously described Certified Descent Algorithm to the optimal design of elastic structures which will be the subject of the future work [5].

At the end of this thesis (cf. appendix A and B), two works completely independent from the main subject are reported. The two appendices tackle different engineering problems but they share the motivating aspect of accounting for practical constraints in complex systems that describe real-world applications via mathematical models.

In the following, we briefly sketch the content of the thesis chapter by chapter.

Chapter 1 - In this introductory chapter, we present the abstract framework of a shape optimization problem. After reviewing the most common approaches described in the literature to tackle this class of problems from a numerical point of view, we focus on gradient-based methods and we recall the notion of shape gradient. Eventually, an overview of the strategies obtained by coupling shape optimization algorithms and *a posteriori* estimators of the discretization error is presented.

Chapter 2 - We introduce the approximation of the previously discussed shape optimization problem via the Finite Element Method. In order to remedy the lack of a stopping criterion and the convergence issues of gradient methods, we account for the numerical error introduced by the discretization of the shape gradient. We review several techniques proposed in the literature to estimate the error due to the Finite Element approximation of a Quantity of Interest depending on the solution

of a Partial Differential Equation. Owing to this information, we construct an *a posteriori* estimator for the error in the shape gradient and we propose a certification procedure to verify whether the direction computed using the discretized shape gradient is a genuine descent direction for the objective functional or not. Eventually, the framework for a novel guaranteed shape optimization strategy - named Certified Descent Algorithm (CDA) - is presented. The content of this chapter is partially based on the works [2, 3, 4].

Chapter 3 - We consider a first example of PDE-constrained optimization problem of a shape-dependent functional. We study an inverse identification problem arising from Electrical Impedance Tomography and we formulate it as a shape optimization problem. Given a material featuring an internal conductivity inhomogeneity, the aim of the problem is to identify the shape and the location of the inclusion by means of boundary measurements. We assume that the boundary data are continuous and we consider the classical Point Electrode Model. After introducing the Kohn-Vogelius functional, we recast the problem within a shape optimization framework and we recall the volumetric and the boundary expressions of the shape gradient. The content of this chapter is based on the works [2, 3, 4].

Chapter 4 - We apply the Certified Descent Algorithm to the optimization of the Kohn-Vogelius functional. We consider an approximation of the previously introduced problem of Electrical Impedance Tomography by means of a conforming Finite Element discretization. A strategy to construct a fully computable, constant-free *a posteriori* estimator of the error in the shape gradient is discussed. Starting from the complementary energy principle, we define additional variational problems to compute an accurate approximation of the error in the numerical fluxes and we derive an estimator for our Quantity of Interest. We validate the error estimator on a test case for which the analytical solution may be derived. Eventually, the overall shape optimization strategy is applied the EIT problem in 2D and a discussion of the numerical results is performed. The content of this chapter is based on the works [3, 4].

Chapter 5 - We propose a variant of the Certified Descent Algorithm which relies solely on local quantities to estimate the error in the shape gradient. In particular, we construct a goal-oriented estimator based on an equilibrated fluxes approach. A unified framework valid for both conforming Finite Element and Discontinuous Galerkin discretizations is introduced. The resulting *a posteriori* estimator is fully computable and provides a reliable bound for the error in our Quantity of Interest which is computationally inexpensive and easy to parallelize. After validating the novel error estimators, this variant of the Certified Descent Algorithm is tested on the previously discussed inverse identification problem of Electrical Impedance Tomography. The content of this chapter is based on the work [2].

Chapter 6 - We consider a second example of PDE-constrained optimization problem of a shape-dependent functional. We introduce the forward problem of the optimal design of an elastic structure governed by the linear elasticity equations. In particular, we analyze the problem of minimizing the compliance under a volume constraint for which we compare two different formulations. First, we investigate the pure displacement formulation of the linear elasticity problem and we compute the volumetric expression of the shape gradient of the compliance by means of the velocity method. Then,

we consider the dual mixed formulation of the elasticity equation and we derive the corresponding expression of the shape gradient for the problem under analysis. The two expressions of the shape gradient are tested using the classical Boundary Variation Algorithm and a qualitative comparison to assess their validity is performed through some numerical simulations. The content of this chapter is based on the work [6].

Chapter 7 - Starting from the framework introduced in the previous chapter for the pure displacement formulation of the linear elasticity problem, we construct an *a posteriori* estimator of the error in the shape gradient. We follow the strategy used in chapter 4 for the problem of Electrical Impedance Tomography and we apply the complementary energy principle to derive a fully computable bound for the error in our Quantity of Interest. We introduce an additional variational problem to compute an accurate approximation of the error in the strain tensor, that is the symmetric part of the gradient of the displacement field. Exploiting the solution of the dual problem, a constant-free *a posteriori* estimator of the error in the shape gradient is derived. The content of this chapter is based on the work [5].

A summary of the results and an overview of possible future developments end this thesis. Both part II and III of this thesis require the notions introduced in part I to be understood. Besides this constraint, the two parts may be read independently but we recommend the reader to follow the order proposed by the author.

Appendix A - This work focuses on the application of optimization techniques to the identification of the optimal grasping points of an aerogripper. Starting from the images acquired by a camera on-board of a Micro Air Vehicle, we exploit the geometrical information of the boundary of the object to be grasped to identify the optimal contact points between the aerogripper and the target. In this appendix we attach the work [7].

Appendix B - This work is devoted to the mathematical modeling and the numerical simulation of a bioreactor landfill, that is a waste management facility in which biodegradable waste is used to generate methane. After analyzing the chemical phenomena taking place inside the bioreactor and the dynamics of the involved fluids, we propose a model that features a heat equation coupled with a non-linear reaction equation that describes the methanogenic chemical phenomena under analysis. Moreover, several advection and advection-diffusion equations model multiphase flows inside a porous environment representing the biodegradable waste. Some preliminary numerical simulations obtained by the approximation of the aforementioned model via the Finite Element Method are presented. In this appendix we attach the work [1].

Published and preprint articles

The main content of this thesis has been developed in collaboration with Olivier Pantz (Full Professor in Applied Mathematics at Université de Nice-Sophia Antipolis) and Karim Trabelsi (Associate Professor in Applied Mathematics at Institut Polytechnique des Sciences Avancées) and resulted in the published conference proceeding [3] and in the published or preprint articles [2, 4, 5, 6].

Appendix A is a work in collaboration with Juan-Antonio Escareno (Associate Professor in Aerial Robotic Systems at Institut Polytechnique des Sciences Avancées) and his collaborators and originated the published conference proceeding [7].

Appendix B is a work in collaboration with Guillaume Dollé (Ph.D. student at Université de Strasbourg), Omar Duran (Ph.D. student at Universidade Estadual de Campinas), Nelson Feyeux (Ph.D. student at Université Joseph Fourier), Emmanuel Frénod (Full Professor in Applied Mathematics at Université de Bretagne-Sud) and Christophe Prud'homme (Full Professor in Applied Mathematics at Université de Strasbourg). The associated paper has been accepted for publication [1].

Part I

Optimization of shape-dependent functionals and a posteriori error estimators

Chapter 1

Shape optimization and shape identification problems

We introduce the abstract framework of a shape optimization problem, that is an optimization problem in which the objective functional depends on the shape of the domain in which a Partial Differential Equation (PDE) is formulated and on the solution of the PDE itself. We review the most common approaches described in the literature to tackle this class of problems and we focus on gradient-based methods which rely on the computation of the so-called shape gradient, that is the differential form of the objective functional with respect to the shape. Eventually, a review of the results on a posteriori estimators of the discretization error for this class of problems is presented.

1.1 Abstract framework

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

$$a_\Omega(u_\Omega, \delta u) = F_\Omega(\delta u) \quad \forall \delta u \in V_\Omega \quad (1.1)$$

where $a_\Omega(\cdot, \cdot) : V_\Omega \times V_\Omega \rightarrow \mathbb{R}$ is a continuous bilinear form satisfying the inf-sup condition

$$\inf_{w \in V_\Omega} \sup_{v \in V_\Omega} \frac{a_\Omega(v, w)}{\|v\| \|w\|} = \inf_{v \in V_\Omega} \sup_{w \in V_\Omega} \frac{a_\Omega(v, w)}{\|v\| \|w\|} > 0$$

and $F_\Omega(\cdot)$ is a continuous linear form on V_Ω , both of them depending on Ω . Under these assumptions, problem (1.1) has a unique solution u_Ω .

We introduce a cost functional $J(\Omega)$ which depends on the domain Ω . We consider the following shape optimization problem

$$\inf_{\Omega \in \mathcal{U}_{\text{ad}}} J(\Omega) \quad (1.2)$$

where \mathcal{U}_{ad} is the set of admissible domains in \mathbb{R}^d . Several examples of this class of problems arise from applications. A very classical case of shape optimization problem is the minimal surface problem

[116] in which we aim to minimize the $(d - 1)$ -dimensional Hausdorff measure \mathcal{H}^{d-1} with respect to all variations with compact support:

$$J(\Omega) := \mathcal{H}^{d-1}(\Omega) \quad , \quad \mathcal{U}_{ad} := \{\Omega \subset \mathbb{R}^d : K \subset \partial\Omega\}.$$

Another well-known example [95, 161] of shape optimization problem is represented by the minimization of the Willmore functional

$$J(\Omega) := \int_S |H|^2 dS \quad , \quad \mathcal{U}_{ad} := \{\Omega \subset \mathbb{R}^d : \mathcal{H}^{d-1}(\Omega) = A\}$$

where H is the mean curvature and $A \in \mathbb{R}^+$ is given. In a more general context, the criterion J depends both on the shape of the domain Ω and on the solution u_Ω of the state equation and the objective functional and the set of the admissible shapes read as follows:

$$J(\Omega) := j(\Omega, u_\Omega), \tag{1.3}$$

$$\mathcal{U}_{ad} := \{\Omega \subset \mathbb{R}^d \text{ s.t. } u_\Omega \in V_\Omega \text{ and } a_\Omega(u_\Omega, \delta u) = F_\Omega(\delta u) \text{ , } \forall \delta u \in V_\Omega\}. \tag{1.4}$$

Within this framework, problem (1.2) may be viewed as a PDE-constrained optimization problem, in which we aim to minimize the functional $j(\Omega, u)$ under the constraint $u = u_\Omega$, that is the minimizer u is solution of the state equation (1.1).

Several industrial problems may be formulated as shape optimization problems in which the objective functional and the set of admissible shapes respectively have the forms (1.3) and (1.4). A non-exhaustive list of applications spans the optimal design of elastic structures [17], the drag minimization in aerodynamics [200] and the identification of physical inhomogeneities in Electrical Impedance Tomography [128]. A vast literature has been developed in this field during the last forty years and for a complete introduction to the subject we refer the reader to [53, 60, 178, 207, 229, 238, 242, 256] and references therein.

1.1.1 A note on the ill-posedness of shape optimization problems

It is well-known in the literature that there may not exist an optimal solution in shape optimization problems if additional constraints on the nature of the shape are not imposed [88, 106, 162]. For example, in a wide number of shape optimization problems porous structures are known to be more efficient than bulky ones [17] but this class of domains does not belong to the set of admissible shapes \mathcal{U}_{ad} . From a practical point of view, this theoretical issue is responsible for many limitations experienced by the strategies proposed in the literature to numerically solve shape optimization problems. Among them, we highlight the high sensitivity of the algorithms to the initialization of the guessed shape - and the consequent risk of getting trapped into local minima - and their mesh dependency. This latter aspect is especially critical since the finer the computational mesh under analysis is, the closer to a porous shape whose micro-structure is of infinitesimal size the optimization procedure is allowed to go.

A first strategy to circumvent the issue of non-existence of optimal solutions relies on relaxing the original problem (1.2) by enlarging the set of admissible shapes to include homogenized - that is, porous - structures. This approach finds a mathematically-sound justification in the homogenization theory and for more details we refer the interested reader to [16, 208]. An alternative idea is based on

the introduction of additional constraints in the set of admissible shapes in order to impose restrictions on the topology of the shape and avoid highly oscillatory phenomena near the boundaries. Here, we briefly recall some possibilities investigated in the literature and we refer to the corresponding works for additional details. In [31], Ambrosio and Buttazzo proved the existence of optimal shapes by adding a constraint on the perimeter of the shape. In [106], Chenais proved that the shape optimization problem (1.2) is well-posed owing the admissible shapes are uniformly Lipschitz. The well-posedness of the shape optimization problem may be alternatively retrieved by adding a constraint on the topology of the shapes, e.g. by setting an upper bound on the number of connected components in the domain. We refer to [105, 266] for additional details on the subject.

For the rest of this thesis, we will neglect the problem of the existence of an optimal shape for the shape optimization problem (1.2). On the contrary, assuming the existence of a local minimizer, (1.2) reduces to

$$\min_{\Omega \in \mathcal{U}_{\text{ad}}} J(\Omega). \quad (1.5)$$

In this thesis, we investigate some numerical strategies to efficiently solve (1.5) by tackling the crucial aspect of defining a reliable stopping criterion for the overall optimization procedure.

1.2 Geometrical description of the shape

A key aspect in shape optimization problems is the efficient description of the geometry of the domain under analysis. In particular, we recall that the aforementioned shape acts both as the domain on which the state equation (1.1) is posed and as the variable in the optimization problem (1.5). There exist two main kinds of framework that one can use to describe the geometry of the shape: an explicit one and an implicit one.

Explicit treatment of the geometry

The most intuitive way to describe the shape of the domain is by means of an explicit treatment of the geometry, e.g. through a polynomial approximation of the boundaries [67], using Bézier curves or Non-Uniform Rational B-Splines (NURBS) [79, 244] or via a FreeForm Deformation (FFD) map [51]. In all these cases, a set of Degrees of Freedom explicitly provides all the information for the construction of the initial domain and the deformation of the shape during the evolution of the algorithm. When considering a polygonal representation of the domain, its Degrees of Freedom are the physical coordinates of the nodes in the triangulation, whereas for the case of Bézier curves, NURBS and FFD maps their control points act as Degrees of Freedom for the description of the geometry. We remark that the aforementioned dependency is said to be explicit, since a change in the set of the Degrees of Freedom is automatically responsible for a modification of the corresponding shape.

Implicit treatment of the geometry

In their seminal work [217], Osher and Sethian introduced a novel paradigm for the description of the motion of a domain. Basic idea relies on using an implicit representation of the domain $\Omega \subset \mathbb{R}^d$

as the subdomain identified by the negative values of an auxiliary function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ - also known as the level-set function - defined as follows:

$$\begin{cases} \phi(x) < 0 & , \ x \in \Omega \\ \phi(x) = 0 & , \ x \in \partial\Omega \\ \phi(x) > 0 & , \ x \in \overline{\Omega}^c := \mathbb{R}^d \setminus \overline{\Omega}. \end{cases} \quad (1.6)$$

Owing to (1.6), the boundary of the design domain is identified by the zero-level contour of the function $\phi(x)$ and not by an explicit set of parameters as in the previously described framework. Moreover, we may express local geometric quantities of the shape in terms of the level-set function [283]. Let Ω be of class \mathcal{C}^2 , for any point x on the boundary $\partial\Omega$ such that $\nabla\phi(x) \neq 0$ we may write the outward normal vector $n(x)$ to $\partial\Omega$ and its mean curvature $\kappa(x)$ as follows

$$n(x) := \frac{\nabla\phi(x)}{|\nabla\phi(x)|} \quad , \quad \kappa(x) := \operatorname{div} \left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|} \right). \quad (1.7)$$

1.3 Classification of shape optimization problems

The goal of shape optimization is to find the shape of a domain that optimizes a given criterion which may depend on the solution of a Boundary Variation Problem as discussed in section 1.1. When dealing with shape optimization problems, three main aspects have to be accounted for: the number of connected regions inside the domain, their shape and their size. Owing to this observation, the following three classes of optimization problems may be identified.

Parametric optimization

Also known as size optimization, parametric optimization aims to improve a given objective functional by modifying a set of parameters describing the domain under analysis [110]. In particular, the number of connected regions and their shape is set *a priori* and only minor variations of the existing configuration are allowed [17], e.g. via the modification of the size of the different regions. Despite the limited improvements it offers, this optimization strategy is widely used in industrial applications owing to its simplicity and to the possibility of exploiting physical quantities as the size and the thickness of the structural elements as design variables in the problem.

Geometric optimization

By performing geometric optimization, one is interested in changing the shape of the domain under analysis to improve the objective functional [17]. Within this framework, the optimization variable of the problem is the boundary of the domain which may be represented either using an explicit or an implicit description as discussed in section 1.2. No *a priori* restrictions on the shape or the size of the domain is posed but the number of connected regions is not allowed to change. In many published works, geometric optimization is also referred to as shape optimization, whereas the latter term is more accurate to describe the ensemble of the problems discussed in this section.

Topology optimization

A key aspect in the engineering community is the identification of the optimal topology (or layout) of the domain under analysis in order to improve a given criterion. This problem is known as topology optimization or layout optimization and aims to identify the best configuration of the domain without imposing any constraint on the number of connected regions inside it [17].

Remark 1.1. Though one may think that topology and geometric optimization problems are extremely different, they actually represent complementary aspects of a unique problem. On the one hand, topology optimization focuses on the identification of the optimal number of connected regions inside the domain. On the other hand, geometric optimization is concerned with the optimization of the shape of the existing regions, whose number is *a priori* set. With an abuse of notation, in the rest of this thesis the term shape optimization will be used to indicate both the geometric optimization paradigm and the ensemble of parametric, geometric and topology optimization. In case of ambiguity, a remark will be provided, e.g. when a given technique requires an additional restriction as the knowledge of the number of connected regions inside the domain.

1.4 A review of the numerical techniques to handle shape optimization problems

The interest in the development of techniques to solve shape optimization problems originally arose in the field of engineering and computational mechanics to answer the question of *where to place material within a given domain in order to obtain the best structure for a specific target objective*. In this section, we provide an overview of the most important methods developed over the years to handle shape optimization problems. For a more detailed comparison of these techniques, we refer to the review paper [253] by Sigmund and Maute and to the works [239] and [140, 269] which focus on more specific aspects, respectively methods that have achieved the stage of application in industrial environments and level-set approaches to structural topology optimization.

The main idea of shape optimization goes back to the seminal work of Hadamard [153] and has been later developed by many authors over several decades [53, 106, 207, 229, 254, 256]. Starting from the original idea of deforming the computational domain according to a given velocity field to improve the value of the objective functional (cf. subsection 1.4.1), we discuss the subsequent developments using the homogenization theory (cf. subsection 1.4.2), the SIMP method (cf. subsection 1.4.3) until the most recent results obtained via the level-set approach (cf. subsection 1.4.4) and the phase-field model (cf. subsection 1.4.5). Eventually, in subsection 1.4.6, we provide a short commentary on the evolutionary strategies that part of the engineering community has been extensively investigating for problems of shape optimization, highlighting the reasons why in our opinion they do not represent a suitable approach for this kind of problems.

1.4.1 A moving mesh approach

Given an initial guess for the shape, a first idea to handle shape optimization problems consists in deforming the boundary of the shape along its normal direction. In a more general framework, the deformation may be driven by any smooth velocity field which motivates the name of this method, known in the literature as the velocity approach [118]. From a practical point of view, the most

intuitive implementation of the velocity approach is the moving mesh method which exploits the geometrical information of the computational mesh to deform the shape by moving the nodes of the triangulation [83].

The main advantages of this approach are the convenient framework based on the calculus of variations that can be used to handle its theoretical analysis and its straightforward numerical implementation. Nevertheless, the method suffers from two main drawbacks. On the one hand, it requires a smooth parameterization of the boundary. On the other one, no topological changes are allowed, that is the number of connected regions inside the domain is *a priori* set. As pointed out in [16], the latter aspect represents a serious issue when dealing with applications in structural design since it is well-known that for a given weight, porous structures are extremely more efficient than bulky ones. In order to remedy the issues due to the smoothness properties required for the boundary and the constraint on topological changes, an approach to shape optimization based on the homogenization theory was developed (cf. e.g. [16]). In next subsections, we will discuss the theoretical improvements due to the use of homogenization as a tool to solve shape optimization problems and we will describe the numerical strategies arising from this approach.

1.4.2 The homogenization method

The application of homogenization theory to the problem of optimal design of structures was developed by several authors starting from the late 1970's. In particular, we refer the interested reader to the works of Murat and Tartar [208, 209], Cherkaev, Lurie and co-workers [192, 193, 194] and Kohn and Strang [181, 182, 183]. During the same years, only few works [144, 146, 147] have dealt with the construction of numerical methods based on the homogenization theory and the authors mainly focused on academic problems with limited impact on real-life applications. It is with the seminal paper [61] by Bendsøe and Kikuchi that homogenization methods were used for the first time to tackle real-life problems in the field of structural shape optimization. After this pioneering contribution, several works have proved the efficiency of this method in shape optimization [22, 27, 122, 150]. For a complete and rigorous introduction to the subject, we refer the reader to [16].

Starting from the notion of relaxed or homogenized formulation, two main strategies may be pursued to solve shape optimization problems (cf. [16]). On the one hand, the so-called optimality criteria method relies on iteratively using the optimality conditions for the relaxed formulation to update the design parameters. On the other hand, by deriving the objective functional with respect to the design variable, gradient-based strategies may be developed. In both cases, a key aspect of the application of the homogenization method to shape optimization is represented by the choice of a distribution of intermediate densities going from 0 to 1 as design variable. Hence, the geometrical description of the structure as a shape with sharp boundaries - one may visualize it as a black-and-white image in which the white regions are void whereas the black ones feature a given material - is replaced by a graded structure representation, that is a grayscale image. Within this framework, the regions in which the density is small are filled with a *soft* material mimicking void, whereas a *hard* material is used where the density tends to 1.

From a practical point of view, the optimal shapes obtained via the homogenization method have the major drawback of not being manufacturable using the machining and molding procedures classically employed in industrial applications. The recent developments in 3D printing technologies may overcome this issue but since the original work [61] of Bendsøe and Kikuchi in 1988, the scientific

community has been investigating several techniques to retrieve classical shapes starting from their homogenized optima. A widely used approach relies on the penalization of the intermediate densities to produce 0 – 1 structures [16] and led to the development of SIMP method [60]. Recently, an alternative post-processing technique for the optimal homogenized shapes by means of a projection strategy has been proposed in [221].

1.4.3 SIMP method: Simplified Isotropic Material with Penalization

Inspired by the homogenization method, in [59] Bendsøe proposed the strategy known as Simplified Isotropic Material with Penalization (SIMP) or power-law approach. The name is due to the relationship posed between the density design variable ρ and the mechanical properties of the material, i.e. the Young's modulus E :

$$E(\rho) = \rho^q E_0. \quad (1.8)$$

In (1.8), E_0 stands for the Young's modulus of the solid material and q is the penalization parameter. The SIMP was driven by the will of reducing the complexity of the homogenization method while improving at the same time the convergence of the algorithm towards 0 – 1 shapes. Several other density approaches have been proposed in the literature over the years. Among them, we cite RAMP (Rational Approximation of Material Properties) [264] but we restrict our description to SIMP since all these strategies are based on the idea of providing a continuous interpolation between solid and void through a penalization of the intermediate densities.

Though extremely efficient and easy to implement, SIMP had no rigorous justification when it first appeared in the literature, especially concerning the choice of the penalization parameter q . Prompted by the promising results obtained with the power-law approach, several authors investigated it in the following years. For $q = 1$, the minimization of the compliance using SIMP corresponds to the well-known Variable Thickness Sheet problem [110] and has been studied in [63, 227]. For $q > 1$, a penalization of the intermediate densities is introduced: within this framework, small values of q are responsible for grayscale areas whereas big q 's cause the algorithm to quickly converge towards a local minimum. A physical justification of SIMP was provided in [62], where the authors showed that choosing $q = 3$ the method converges to almost 0 – 1 shapes and the physical realizability of the elements featuring intermediate densities is ensured. Eventually, in [34] the author proved that $q = 3$ is the exponent that makes density gradients equal to topological derivatives for the case of linear elasticity.

It is well-known in the literature that without smoothing or filtering, SIMP typically converges to local minima with poor performance. Moreover, a major issue of the computed solutions is their mesh dependency, as remarked by the extensive documentation on the checkerboard problem, that is the phenomenon of solutions featuring patches of alternating black and white elements [228]. To circumvent the issue of rapid oscillations of the density distribution, several techniques have been studied over the years. In the rest of this subsection, we briefly recall some ideas to improve SIMP and we refer to [253] for a detailed review of these techniques. A first improvement of SIMP may be obtained by introducing additional constraints on the length of the perimeter or explicit penalizations which depend on the gradient of the density in order to regularize the solution. A possible alternative is represented by sensitivity filters which modify the value of the sensitivity in a mesh element using a weighted average of the values in the neighboring elements within a mesh-independent radius [249]. In [74], Bourdin proposed a density filter which defines the density inside a given mesh element

as a weighted average of the design variable in the neighboring elements contained within a mesh-independent radius. More recently, novel techniques have been proposed to overcome the issue of regions featuring grayscale transitions. Starting from the aforementioned filters, in [151] the authors proposed to project the resulting shape in a binary space by means of a smoothed Heaviside function.

Due to its simplicity and the good quality of the resulting shapes, SIMP has encountered a huge success within the engineering community and is currently the basic strategy implemented in most commercial softwares dedicated to topology optimization. Nevertheless, the dependency of the method on multiple parameters and the need of several post-processing techniques to retrieve a 0 – 1 shape starting from the optimum obtained through SIMP limit the robustness of the overall approach. As a matter of fact, a precise tuning of the involved parameters is required to obtain an efficient optimization algorithm. Moreover, the problem-dependent nature of the aforementioned quantities to be estimated makes the resulting strategy extremely heuristic.

1.4.4 A level-set approach

Starting from the implicit representation of the domain discussed in section 1.2, Allaire *et al.* [25, 26] and Wang *et al.* [278] have developed a level-set approach to handle shape optimization problems. The key aspect of this method relies on the definition of the equation describing the evolution of the boundary of the shape.

We consider a vector field $\theta \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d)$ and we introduce a transformation $X_\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$. The open subset $\Omega_\theta \subset \mathbb{R}^d$ is such that $\Omega_\theta = X_\theta(\Omega)$, where Ω is the initial reference domain. The displacement of a point from its initial configuration $x \in \Omega$ is governed by the following differential equation:

$$\begin{cases} \frac{dx_\theta}{dt}(t) = \theta(x_\theta(t)) \\ x_\theta(0) = x \end{cases} \quad (1.9)$$

which admits a unique solution $t \mapsto x_\theta(t, x)$ in $C^1(\mathbb{R}; \mathbb{R}^d)$. Owing to (1.9), the initial point $x \in \Omega$ is transported by the field θ to the point $x_\theta = X_\theta(x)$ which belongs to the deformed domain Ω_θ . We remark that the time variable in (1.9) is an artificial pseudo-time introduced to describe the evolution of the shape of the domain.

Let us now consider a point $x_\theta \in \partial\Omega_\theta$. The boundary of the evolving domain Ω_θ may be represented by the following zero-level contour of a level-set function:

$$\phi(t, x_\theta) = 0 \quad , \quad \forall t \in \mathbb{R}. \quad (1.10)$$

Since this is true $\forall t \in \mathbb{R}$ and for each initial point $x \in \mathbb{R}^d$, owing to the chain-rule (1.10) yields to the following advection equation for the level-set function:

$$\frac{\partial \phi}{\partial t}(t, x) + \theta(t, x) \cdot \nabla \phi(t, x) = 0 \quad , \quad \forall (t, x) \in \mathbb{R} \times \mathbb{R}^d. \quad (1.11)$$

We remark that the velocity field $\theta(t, x)$ is often oriented along the normal direction to the moving boundary, that is $\theta(t, x) = \Theta(t, x)n(t, x)$. As a matter of fact, within the framework of shape optimization the velocity field driving the evolution of the domain accounts for the sensitivity of the objective functional with respect to perturbations of the interface along the normal direction. Owing to the

expression (1.7) of the unit normal vector in terms of the level-set function, (1.11) may be rewritten as the Hamilton-Jacobi equation:

$$\frac{\partial \phi}{\partial t}(t, x) + \Theta(t, x)|\nabla \phi(t, x)| = 0 \quad , \quad \forall (t, x) \in \mathbb{R} \times \mathbb{R}^d. \quad (1.12)$$

When tackling the problem of discretizing equation (1.12), several issues may appear. Though these issues are mainly related to the level-set technique rather than to the shape optimization problem under analysis, we briefly recall them in order to provide the reader a global presentation of the method and the corrections proposed in the literature over the years. A first improvement introduced a re-initialization step for the level-set function in order to prevent the gradient $\nabla \phi$ from becoming too steep or too flat along the interface [278]. Moreover, the numerical approximation of the Hamilton-Jacobi equation has to be carefully handled in order to prevent rapidly oscillating level-set fields at the interface between solid and void [90]. This issue is mainly due to the purely convective nature of the problem under analysis. It is well-known in the literature that highly accurate and stabilized numerical schemes as the Essentially Non-Oscillatory (ENO) and Weighted Essentially Non-Oscillatory (WENO) are required to solve this class of PDE's [156, 218]. Moreover, the time evolution is performed by means of explicit or semi-implicit time-marching schemes for which additional stability issues may appear [176, 246].

In general, the level-set approach by Allaire *et al.* [25] describes the evolution of the interfaces existing in the initial design, thus solving equation (1.12) does not allow the nucleation of new holes. In order to circumvent this issue, several strategies have been proposed. In [26, 277], the authors proposed to augment the Hamilton-Jacobi equation via a diffusive term to smooth and regularize the solution and a reactive term to nucleate holes. Starting from the aforementioned approach, alternative models neglecting the convective term have been investigated, driving the evolution of the level-set function solely by means of the reaction term [281]. An alternative way to introduce holes into the level-set function takes advantage of the notion of topological gradient [255]. The use of topological derivatives in shape optimization was first proposed by Eschenauer *et al.* [133] with the so-called bubble method: basic idea relies on the evaluation of the effect of an infinitesimal bubble at a given point to consequently drive the insertion of a hole in the domain under analysis. This idea has been later developed by other authors within the framework of level-set formulations [21, 89, 104]: in these works, the shape of the domain is modified according to the shape gradient obtained via the Hadamard's method (cf. section 1.5) and from time to time the underlying layout is updated via a topological sensitivity analysis. In a similar fashion, in [220] Pantz and Trabelsi added a pre-processing stage to the aforementioned strategy in order to retrieve a clever initialization for the shape optimization problem via the homogenization method.

Since their first appearance in 2002, level-set techniques for shape optimization problems have gained in popularity within both the mathematical and the engineering community. This is mainly due to the observation that these methods operate on the boundaries of the shape throughout the optimization process and they efficiently provide extremely interesting results. Nevertheless, as all methods based on the information carried by gradients (i.e. shape sensitivities, shape gradient or topological gradients), their results are strongly dependent on the initial guess for the shape and may be trapped into local minima.

1.4.5 A phase-field model

An alternative to the implicit penalization scheme in SIMP is the use of explicit penalization terms. This is the founding idea of the phase-field models which aim to minimize a functional with respect to the density variable. Starting from the criterion $J(\Omega)$ of the optimization problem (e.g. the compliance in structural mechanics), the aforementioned functional is obtained by adding a term to penalize the intermediate densities:

$$\bar{J}(\Omega) := J(\Omega) + \int_{\Omega} \left(\varepsilon |\nabla \rho(x)|^2 + \frac{1}{\varepsilon} W(\rho(x)) \right) dx. \quad (1.13)$$

In (1.13), the density ρ acts as design variable whereas $W(\cdot)$ is an energy term that accounts for the phenomena taking place near the interface, that is the region in which the transition between solid and void is located. A key aspect of this approach is the explicit dependency of the penalty term on the parameter $\varepsilon > 0$ that weights the thickness of the aforementioned interface. The second term on the right-hand side of (1.13) was originally introduced by Cahn and Hilliard [93] to model the interfacial energy of a mixture of fluids and has been extensively studied in the literature in the last fifty years. In the last decade, several methods to directly minimize (1.13) by adding a volume constraint and without introducing any auxiliary field have been proposed [75, 91, 279]. More recently, an alternative approach based on the splitting of the fourth-order Cahn-Hilliard equation arising from the minimization of (1.13) into two second-order equations have been investigated [276].

Though the phase-field models appear to be interesting tools for the treatment of shape optimization problems, the results available so far in the literature have not achieved the same degree of maturity as the density or the level-set methods and further investigations may be necessary, especially concerning their efficiency and optimization.

1.4.6 Gradient-free strategies

Since the design variables involved in the original formulations of shape optimization problems are discrete, several approaches based on discrete optimization techniques have been proposed in the literature. We refer the reader to the review paper [206].

As Sigmund argued in [251], the use of non-gradient techniques is not a viable approach for shape optimization problems. As a matter of fact, to the best of our knowledge all the published results using Genetic Algorithms are restricted to small-size problems for which the computational cost is already enormous. Though some counterexamples may be encountered in the literature, the corresponding methods usually are extremely sensitive to variations of the parameters and may result in oscillating solutions. Moreover, it is unclear how to construct update strategies and stopping criteria that guarantee convergence to physically admissible designs in an efficient and robust way. Eventually, a major drawback of the aforementioned strategies is the high number of evaluations of the objective functional - which usually depends on the numerical solution of a PDE - that explodes quite rapidly making Genetic Algorithms substantially unfeasible in practical shape optimization applications.

Owing to the previous remarks, we support the cited statement of Sigmund on the limited usefulness of gradient-free approaches in shape optimization and for the rest of this thesis we will solely focus on gradient-based methods.

1.5 Differentiation with respect to the shape

All the gradient-based methods presented in the previous section rely on the ability of computing the sensitivity of the objective functional $J(\Omega)$ with respect to perturbations of the shape. In a first approach - mainly applied in the framework of density methods like SIMP -, the sensitivity of the objective functional with respect to variations of the shape may be interpreted as a parameter sensitivity, being the parameter under analysis the presence of material or void in a given computational cell. Another possibility is to perform the sensitivity analysis via the so-called shape gradient, that is by computing the variation of the objective functional with respect to perturbations of the boundary of the shape. This method is also known as Hadamard's method or velocity approach [256]. Alternatively, one may perform a topological sensitivity analysis by evaluating the sensitivity of the objective functional with respect to the nucleation of infinitesimally small holes inside the design configuration under analysis [255]. In this thesis, we focus on the approach based on the shape gradient whereas we think that using the information arising from the topological gradient definitely represents an interesting research line and promising results were recently obtained within this framework (cf. e.g. [20]).

1.5.1 Hadamard's boundary variation method

In this subsection, we recall the notion of shape gradient of $J(\Omega)$ in the direction θ . Let $X \subset W^{1,\infty}(\Omega; \mathbb{R}^d)$ be a Banach space and $\theta \in X$ be an admissible smooth deformation of Ω . The cost functional $J(\Omega)$ 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 $\forall \theta \in X$

$$J((\text{Id} + \theta)\Omega) = J(\Omega) + \langle dJ(\Omega), \theta \rangle + o(\theta).$$

Several approaches are feasible to compute the shape gradient. Here we briefly recall the fast derivation method by C  a [103] and the material derivative approach [256].

Fast derivation method by C  a

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

$$\mathcal{L}(\Omega, u, p) = j(\Omega, u) + a_\Omega(u, p) - F_\Omega(p). \quad (1.14)$$

Let $p_\Omega \in V_\Omega$ be the solution of the so-called adjoint problem, that is

$$a_\Omega(\delta p, p_\Omega) + \left\langle \frac{\partial j}{\partial u}(\Omega, u_\Omega), \delta p \right\rangle = 0 \quad \forall \delta p \in V_\Omega. \quad (1.15)$$

By applying the fast derivation method by C  a, we get the following expression of the shape gradient:

$$\langle dJ(\Omega), \theta \rangle = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}(\Omega, u_\Omega, p_\Omega), \theta \right\rangle. \quad (1.16)$$

A material derivative approach

An alternative procedure to compute the shape gradient relies on the definition of a diffeomorphism $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that for every admissible set Ω

$$\Omega_\varphi := \varphi(\Omega).$$

Moreover, all functions $u, p \in V_\Omega$ defined on the reference domain Ω may be mapped to the deformed domain Ω_φ by

$$u_\varphi := u \circ \varphi^{-1} \quad \text{and} \quad p_\varphi := p \circ \varphi^{-1}.$$

We admit that $u \mapsto u_\varphi$ is a one-to-one map between V_Ω and V_{Ω_φ} . The Lagrangian (1.14) is said to admit a material derivative if there exists a linear form $\frac{\partial \mathcal{L}}{\partial \varphi}$ such that

$$\mathcal{L}(\Omega_\varphi, u_\varphi, p_\varphi) = \mathcal{L}(\Omega, u, p) + \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u, p), \theta \right\rangle + o(\theta)$$

where $\varphi = \text{Id} + \theta$. Provided that u_φ is differentiable with respect to φ at $\varphi = \text{Id}$ in V_{Ω_φ} , from the fast derivation method of C ea we obtain the following expression for the shape gradient:

$$\langle dJ(\Omega), \theta \rangle = \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega, p_\Omega), \theta \right\rangle. \quad (1.17)$$

A variant of the latter method consists in computing the shape gradient via the Lagrangian formulation without explicitly constructing the material derivative of the state and adjoint solutions. We refer to [189] for additional information about this approach.

1.5.2 Volumetric and surface expressions of the shape gradient

The most common approach in the literature to compute the shape gradient is based on an Eulerian point of view and leads to a surface expression of the shape gradient.

The main advantage of this method relies on the fact that the boundary representation intuitively provides an explicit expression for the descent direction. Let us assume that the shape gradient has the following form

$$\langle dJ(\Omega), \theta \rangle = \int_{\partial\Omega} h\theta \cdot n \, ds$$

then $\theta = -hn$ on $\partial\Omega$ is a descent direction. Moreover, by the Hadamard-Zol sio structure theorem it is well-known that the shape gradient is carried on the boundary of the shape [118] and using this approach the descent direction has to be defined solely on $\partial\Omega$. Nevertheless, if the boundary datum of the state problem is not sufficiently smooth, the surface expression of the shape gradient may not exist or the descent direction θ may suffer from poor regularity.

Starting from the surface representation of the shape gradient, it is possible to derive a volumetric expression as well. Though the two expressions are equivalent in a continuous framework, they usually are not when considering their numerical counterparts, e.g. their Finite Element approximations: as a matter of fact, in [166] Hiptmair *et al.* prove that the volumetric formulation generally provides better accuracy when using the Finite Element Method. Moreover, we may be able to derive the volumetric expression of the shape gradient even when its boundary representation fails to exist.

In this work, we consider the volumetric expression of the shape gradient in order to take advantage of the better accuracy it provides from a numerical point of view and to construct an estimator of the error in a Quantity of Interest using the procedure described by Oden and Prudhomme in [215].

We remark that in order to derive a descent direction θ on Ω from the volumetric expression of the shape gradient, an additional variational problem has to be solved, as described in next chapter.

1.6 A posteriori error estimators for shape optimization

For a long time, the numerical approaches discussed in section 1.4 were developed exclusively for fixed cartesian computational grids. The main drawback of this choice was the irregularity of the optimized shapes featuring jagged edges that may contradict the underlying physics of the problem. To circumvent this issue, in recent years several works have been dealing with generalized Finite Element techniques, e.g. the eXtended Finite Element Method (XFEM) [184, 270] and the Adaptive Finite Element Method (AFEM) by means of local remeshing strategies [282] and mesh evolution techniques [19, 20]. In this thesis, we are interested in the latter strategy, that is in using the Adaptive Finite Element Method for shape optimization problems. Within this framework, a key question is how to define a procedure to modify the mesh. It is a widely accepted approach in the Finite Element community to drive mesh adaptation strategies according to the information arising from *a posteriori* error estimators.

Several works in the literature have highlighted the great potential of coupling *a posteriori* error estimators to shape optimization algorithms. In the pioneering work [54], the authors identified two different sources for the numerical error: on the one hand, the error arising from the discretization of the differential problem and on the other hand, the error due to the approximation of the geometry. Starting from this observation, Banichuk *et al.* presented a first attempt to exploit the information on the discretization of the differential problem provided by a recovery-based estimator and the error arising from the approximation of the geometry to develop an adaptive shape optimization strategy. This work has been later extended by Morin *et al.* in [201], where the adaptive discretization of the governing equations by means of the Adaptive Finite Element Method is linked to an adaptive strategy for the approximation of the geometry. The authors derive estimators of the numerical error that are later used to drive an Adaptive Sequential Quadratic Programming algorithm to appropriately refine and coarsen the computational mesh. Several other authors have used adaptive techniques for the approximation of PDE's in order to improve the accuracy of the solution and obtain better final configurations in problems of design of optimal structures. We refer to [15, 174, 236, 243] for some examples.

We remark that in all the aforementioned works, *a posteriori* estimators only provide qualitative information about the numerical error due to the discretization of the problems and are essentially used to drive mesh adaptation procedures. Though these approaches are certainly interesting, they do not allow the optimization procedure to be automatized and remain highly dependent on the choices of the end user concerning the conditions to stop the algorithm. To the best of our knowledge, the problem of automatizing shape optimization strategies is still open and no rigorous solution has been proposed in the literature. In particular, no guaranteed fully computable estimate has been investigated in the context of shape optimization and the error in the shape gradient itself is not accounted for, thus preventing reliable stopping criteria to be derived. In the next chapter, we tackle

the issue of estimating the error due to the numerical approximation of the shape gradient and we employ this information to introduce a certification procedure that allows us to automatize the shape optimization strategy and derive a reliable stopping criterion.

Chapter 2

Certification procedure for a genuine descent direction

We account for the numerical error introduced by the Finite Element approximation of the shape gradient and we introduce a certification procedure to verify whether the direction computed using the discretized shape gradient is a genuine descent direction for the objective functional or not. Several techniques for estimating the error in the shape gradient are discussed and a novel guaranteed shape optimization strategy - named Certified Descent Algorithm (CDA) - is presented.

2.1 Optimize-then-Discretize and Discretize-then-Optimize

Two main approaches have been proposed in the literature to handle PDE-constrained optimization problems - thus shape optimization problems of the form (1.5) as well - via gradient-based methods. On the one hand, the Optimize-then-Discretize strategy [65] relies on the analytical computation of the gradient of the cost functional which is then discretized to run the optimization loop. On the other hand, the Discretize-then-Optimize procedure [66] works the other way around, by first computing a discretized version of the objective functional and subsequently constructing its gradient to perform the optimization strategy. The Discretize-then-Optimize strategy has two main drawbacks: on the one hand, the discretized functional may not be differentiable, thus limiting the possibility of using a gradient method; on the other hand, this approach may suffer from severe mesh dependency. In the rest of this thesis, we consider an Optimize-then-Discretize approach for problem (1.5) by studying a variant of the Boundary Variation Algorithm (BVA). In subsection 2.1.1, we recall the basic idea of the Boundary Variation Algorithm described by Allaire and Pantz in [28]: this method relies on the computation of the so-called shape gradient which arises from the differentiation of the functional with respect to the shape (cf. section 1.5). Then, in subsection 2.1.2 we highlight the issues arising when moving from the continuous framework to the discrete one and eventually we propose a certification procedure to remedy them (cf. subsection 2.1.3).

2.1.1 The Boundary Variation Algorithm

From now on, we consider X to be a Hilbert space. Starting from the formulation (1.17), we seek a descent direction for the functional $J(\Omega)$. For this purpose, we solve an additional variational problem and we seek $\theta \in X$ such that

$$(\theta, \delta\theta)_X + \langle dJ(\Omega), \delta\theta \rangle = 0 \quad \forall \delta\theta \in X. \quad (2.1)$$

Remark 2.1. The choice of the scalar product in (2.1) is a key point for the development of an efficient shape optimization method. In [44], the authors proposed the so-called traction method to get rid of some irregularity issues in shape optimization problems. This approach is based on the regularization of the descent direction by means of a scalar product inspired by the linear elasticity equation. In recent years, a comparison of the L^2 , H^1 and H^{-1} scalar products defined on a surface was presented (cf. [123]) but, as the authors state, the best choice is strongly dependent on the application of interest.

In this subsection, we present a gradient method for shape optimization problems. After computing the solution of the state equation (1.1), we solve the adjoint problem (1.15) to derive the expression (1.17) of the shape gradient. Then, a descent direction is identified through (2.1) and is used to deform the domain via a perturbation of the identity map $\text{Id} + \theta$. The resulting shape optimization strategy is known in the literature as Boundary Variation Algorithm [28] and is sketched in script 2.1.

Algorithm 2.1 – The Boundary Variation Algorithm

Given the domain Ω_0 , set $\text{tol} > 0$, $j = 0$ and iterate:

1. Compute the solution u_{Ω_j} of the state equation;
2. Compute the solution p_{Ω_j} of the adjoint equation;
3. Compute a descent direction $\theta_j \in X$ solving

$$(\theta_j, \delta\theta)_X + \langle dJ(\Omega_j), \delta\theta \rangle = 0 \quad \forall \delta\theta \in X ;$$

4. Identify an admissible step μ_j ;
 5. Update the domain $\Omega_{j+1} = (\text{Id} + \mu_j \theta_j) \Omega_j$;
 6. While $|\langle dJ(\Omega_j), \theta_j \rangle| > \text{tol}$, $j = j + 1$ and repeat.
-

We recall the following definition:

Definition 2.2. A direction θ is said to be a genuine descent direction for the functional $J(\Omega)$ if

$$\langle dJ(\Omega), \theta \rangle < 0. \quad (2.2)$$

It is straightforward to observe that a direction θ fulfilling (2.2) is such that $J(\Omega)$ decreases along θ , that is $J((\text{Id} + \theta)\Omega) < J(\Omega)$. We remark that algorithm 2.1 relies on the computation of a genuine descent direction for the functional $J(\Omega)$ at each iteration. In next subsections, we discuss the modifications that occur when moving from the continuous to the discrete formulation of the problems and consequently the condition that the discretization of θ has to fulfill in order to be a genuine descent direction for the functional $J(\Omega)$.

2.1.2 The discretized Boundary Variation Algorithm

We introduce the discretization of the state and adjoint problems described in the previous chapter. Let $V_\Omega^{h,\ell}$ be an appropriate Finite Element or Discontinuous Galerkin approximation space featuring basis functions of degree ℓ . We define the bilinear form $a_\Omega^h(\cdot, \cdot)$ and the linear form $F_\Omega^h(\cdot)$ associated with the following discrete state problem: we seek $u_\Omega^h \in V_\Omega^{h,\ell}$ such that

$$a_\Omega^h(u_\Omega^h, \delta u^h) = F_\Omega^h(\delta u^h) \quad \forall \delta u^h \in V_\Omega^{h,\ell}. \quad (2.3)$$

Following the same procedure, we introduce the discrete adjoint problem which consists in seeking $p_\Omega^h \in V_\Omega^{h,\ell}$ such that

$$a_\Omega^h(\delta p^h, p_\Omega^h) + \left\langle \frac{\partial j}{\partial u}(\Omega, u_\Omega^h), \delta p^h \right\rangle = 0 \quad \forall \delta p^h \in V_\Omega^{h,\ell}. \quad (2.4)$$

A discussion on the degree of the approximating functions in the space $V_\Omega^{h,\ell}$ will be detailed in chapter 4 and 5 for the problem of Electrical Impedance Tomography and in chapter 6 for the shape optimization problem in linear elasticity.

The discretized direction $\theta^h \in X$ is obtained by solving the problem

$$(\theta^h, \delta \theta)_X + \langle d_h J(\Omega), \delta \theta \rangle = 0 \quad \forall \delta \theta \in X \quad (2.5)$$

where $\langle d_h J(\Omega), \delta \theta \rangle$ reads as follows:

$$\langle d_h J(\Omega), \delta \theta \rangle := \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega^h, p_\Omega^h), \delta \theta \right\rangle. \quad (2.6)$$

The discretized version of the Boundary Variation Algorithm is derived by substituting the continuous functions u_Ω , p_Ω with their approximations u_Ω^h , p_Ω^h and θ with θ^h :

Algorithm 2.2 – The discretized Boundary Variation Algorithm

Given the domain Ω_0 , set $\text{tol} > 0$, $j = 0$ and iterate:

1. Compute the solution $u_{\Omega_j}^h$ of the state equation;
2. Compute the solution $p_{\Omega_j}^h$ of the adjoint equation;
3. Compute a descent direction $\theta_j^h \in X$ solving

$$(\theta_j^h, \delta \theta)_X + \langle d_h J(\Omega_j), \delta \theta \rangle = 0 \quad \forall \delta \theta \in X ;$$

4. Identify an admissible step μ_j ;
 5. Update the domain $\Omega_{j+1} = (\text{Id} + \mu_j \theta_j^h) \Omega_j$;
 6. While $|\langle d_h J(\Omega_j), \theta_j^h \rangle| > \text{tol}$, $j = j + 1$ and repeat.
-

We remark that due to the numerical error introduced by the discretization, even though $\langle d_h J(\Omega), \theta^h \rangle < 0$, θ^h is not necessarily a genuine descent direction for the functional $J(\Omega)$. Moreover, it is important to notice that the stopping criterion (Algorithm 2.2 - step 6) will usually not be fulfilled if the required tolerance tol is too sharp with respect to the chosen discretization.

In order to bypass these issues, in the following subsection we present a strategy to account for the error introduced by the approximation of the shape gradient. This results in a certification procedure that allows to verify whether a given direction is a genuine descent direction for the functional $J(\Omega)$ or not.

2.1.3 Certification procedure for a genuine descent direction

We introduce the notion of certified descent direction, that is a direction which is verified to be a genuine descent direction for the functional $J(\Omega)$. As previously stated, a genuine descent direction θ is such that $\langle dJ(\Omega), \theta \rangle < 0$. When moving from the continuous to the discrete framework, an approximation of the shape gradient is introduced (cf. definition (2.6)) and consequently a numerical error appears. We define the error in the shape gradient E^h as follows:

$$E^h := \langle dJ(\Omega) - d_h J(\Omega), \theta^h \rangle. \quad (2.7)$$

From (2.7), it follows that

$$\langle dJ(\Omega), \theta^h \rangle = \langle d_h J(\Omega), \theta^h \rangle + E^h. \quad (2.8)$$

In general, θ^h is constructed as the solution of (2.5), where (2.6) is the discretized shape gradient. It follows that θ^h is such that $\langle d_h J(\Omega), \theta^h \rangle < 0$. In order for θ^h to be a descent direction for the objective functional $J(\Omega)$, condition (2.2) has to be fulfilled and the quantity E^h in (2.8) has to be accounted for. Within this framework, we obtain the following condition on θ^h :

$$\langle d_h J(\Omega), \theta^h \rangle + E^h < 0. \quad (2.9)$$

Nevertheless, this condition does not imply that θ^h is a genuine descent direction for $J(\Omega)$ since the quantity E^h in (2.9) may be either positive or negative. In order to derive a relationship that stands independently of the sign of E^h and since no *a priori* information on the aforementioned sign is available, we modify (2.9) by introducing the absolute value of E^h :

$$\langle d_h J(\Omega), \theta^h \rangle + E^h \leq \langle d_h J(\Omega), \theta^h \rangle + |E^h| < 0. \quad (2.10)$$

We may now introduce the following definition:

Definition 2.3. Let \bar{E} be the upper bound of the error $|E^h|$ in the shape gradient. A direction θ^h is said to be a certified descent direction for the functional $J(\Omega)$ if

$$\langle d_h J(\Omega), \theta^h \rangle + \bar{E} < 0. \quad (2.11)$$

The expression *certified* is due to the fact that a direction constructed within this framework is verified to be a genuine descent direction for the functional $J(\Omega)$. As a matter of fact, it is straightforward to observe that if θ^h fulfills (2.11), then it verifies (2.2) as well, that is, it is guaranteed that the functional $J(\Omega)$ decreases along θ^h .

Remark 2.4. It is important to observe that a direction fulfilling (2.11) is a genuine descent direction for $J(\Omega)$, whether it is the solution of equation (2.1) or not. This is extremely important since the computation of the descent direction is done through the discretization of the aforementioned variational problem, that is θ^h is only an approximation of the direction θ solution of (2.1).

In the following section, we present a strategy to construct a fully computable guaranteed upper bound of the error in the approximation of the shape gradient in order to practically implement the certification procedure described above.

2.2 Numerical error in the shape gradient

In this section, we provide the details of the technique used to derive an upper bound \bar{E} of the error $|E^h|$ in the shape gradient. The strategy to estimate the error in a Quantity of Interest (QoI) - namely, the shape gradient - is derived from the work [215] by Oden and Prudhomme. Basic idea relies on the definition of an adjoint problem whose right-hand side is the quantity whose error estimate is sought. An alternative approach which has found success especially in the computational mechanics community is the Constitutive Relation Error (CRE) method. We refer to [185] for a review of the basics and the most recent developments of this technique.

2.2.1 Bound for the error due to the approximation of a linear functional

Here we briefly recall the aforementioned strategy by Oden and Prudhomme for the derivation of an *a posteriori* error estimate for a bounded linear functional $Q : V_\Omega \rightarrow \mathbb{R}$, also known as Quantity of Interest. We consider a quantity Q which we aim to evaluate for the function u_Ω , solution of the state problem (1.1). We introduce the solution u_Ω^h of the corresponding discrete problem (2.3) and we seek an estimate of the error in the target functional Q :

$$E_Q := Q(u_\Omega) - Q(u_\Omega^h) = Q(u_\Omega - u_\Omega^h) \quad (2.12)$$

where the equality follows from the linearity of Q . We introduce an adjoint problem featuring the quantity Q as right-hand side, that is we seek an influence function $r_\Omega \in V_\Omega$ such that

$$a_\Omega(\delta r, r_\Omega) = Q(\delta r) \quad \forall \delta r \in V_\Omega. \quad (2.13)$$

Existence and uniqueness of the solution r_Ω of the adjoint problem follow from the Lax-Milgram theorem. From a practical point of view, the solution of the aforementioned adjoint problem is only approximated and we consider the same method used for the state problem. Let $V_\Omega^{h,m}$ be the space of Finite Element (respectively Discontinuous Galerkin) functions of degree $m \geq \ell$ and $a_\Omega^h(\cdot, \cdot)$ be the discrete bilinear form associated with problem (2.13). We seek $r_\Omega^h \in V_\Omega^{h,m}$ such that

$$a_\Omega^h(\delta r^h, r_\Omega^h) = Q(\delta r^h) \quad \forall \delta r^h \in V_\Omega^{h,m}. \quad (2.14)$$

From (2.13) and (1.1) it is straightforward to observe that

$$F_\Omega(r_\Omega) = a_\Omega(u_\Omega, r_\Omega) = Q(u_\Omega). \quad (2.15)$$

Thus, the following relationship between the error in the approximation of the target functional and the solutions of the state and adjoint problems is derived:

$$E_Q = Q(u_\Omega) - Q(u_\Omega^h) = F_\Omega(r_\Omega) - a_\Omega^h(u_\Omega^h, r_\Omega^h) = F_\Omega(r_\Omega) - a_\Omega^h(u_\Omega^h, r_\Omega) \quad (2.16)$$

where the first equality follows from the approximation (2.14) of the adjoint problem (2.13) whereas the justification of the last one exploits different properties when dealing with conforming Finite Element or Discontinuous Galerkin approximations.

Conforming Finite Element approximation

Let $V_\Omega^{h,\kappa}$ be the space of Lagrangian Finite Element functions of degree κ , $\kappa = \ell, m$. The last equality in (2.16) follows from the fact that $a_\Omega^h(\cdot, \cdot) = a_\Omega(\cdot, \cdot)$ and $a_\Omega(\delta r^h, r_\Omega) = Q(\delta r^h)$ stands $\forall \delta r^h \in V_\Omega^{h,m}$ owing to (2.13) and $V_\Omega^{h,m}$ being a subspace of V_Ω .

Let us introduce the residue associated with the approximation of the state problem (1.1):

$$R_\Omega^u(\delta u) := F_\Omega(\delta u) - a_\Omega(u_\Omega^h, \delta u). \quad (2.17)$$

We recall that the error $e^h := u_\Omega - u_\Omega^h$ is solution of the so-called residual equation that reads

$$a_\Omega(e^h, \delta u) = R_\Omega^u(\delta u) \quad \forall \delta u \in V_\Omega. \quad (2.18)$$

Hence, from (2.16) we retrieve the expression of the residue of the state equation applied to the function r_Ω :

$$Q(e^h) = R_\Omega^u(r_\Omega) = a_\Omega(e^h, r_\Omega).$$

Within this framework, r_Ω is known as *influence function* since it carries the information on the effect of the quantity $u_\Omega - u_\Omega^h$ on the error E_Q on the functional Q . Owing to the Galerkin orthogonality, the evaluation of $Q(e^h)$ reads as

$$Q(e^h) = a_\Omega(e^h, r_\Omega) = a_\Omega(e^h, r_\Omega - r_\Omega^h) = a_\Omega(e^h, \epsilon^h) \quad (2.19)$$

where $\epsilon^h := r_\Omega - r_\Omega^h$ is the error due to the discretization of the adjoint problem.

Discontinuous Galerkin approximation

Let $V_\Omega^{h,\kappa}$ be the space of discontinuous functions of degree κ , $\kappa = \ell, m$ for the Discontinuous Galerkin (DG) approximation of the state and adjoint problems. When dealing with DG formulations, the expression of the discrete bilinear form features the terms associated with the jumps of the discontinuous functions across the edges (respectively the faces in 3D) of the triangulation. Thus, it is straightforward to observe that the discrete bilinear form cannot be identified with its continuous version. Within this framework, the last equality in (2.16) stands if the numerical method used to discretize the adjoint problem is consistent, that is if

$$a_\Omega^h(\delta r^h, r_\Omega) = Q(\delta r^h) \quad \forall \delta r^h \in V_\Omega^{h,m}.$$

This property is known as adjoint consistency and is equivalent to the usual Galerkin orthogonality property in Finite Element (cf. [121]). We refer to [157] for more details on the role of adjoint consistency in the construction of discretizations of optimal order in terms of target functionals. The explicit formulation of the error estimator in the Quantity of Interest Q is problem-dependent and may be derived starting from (2.16) and substituting the expressions of the linear form of the state problem and the discrete bilinear form. Hence, we refer to section 5.3 for the details on the consistency of the Discontinuous Galerkin strategy chosen for the test case of Electrical Impedance Tomography and for the explicit derivation of the error in the shape gradient. For a general introduction to the mathematical aspects of Discontinuous Galerkin methods we refer the interested reader to [121].

2.2.2 Variational formulation of the error in the shape gradient

Several works in the literature [149, 240] have dealt with an extension of the aforementioned framework to compute a bound of the approximation error in a target functional to the case of non-linear Quantities of Interest. To account for the error (2.7), we follow the approach proposed by these authors and we perform a linearization of the functional whose error estimate is sought. Thus we rewrite the numerical error (2.7) in the shape gradient by introducing the linearized error \tilde{E}^h :

$$\begin{aligned} E^h &= \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega, p_\Omega) - \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega^h, p_\Omega^h), \theta^h \right\rangle \\ &\simeq \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, u_\Omega - u_\Omega^h] + \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, p_\Omega - p_\Omega^h] =: \tilde{E}^h. \end{aligned} \quad (2.20)$$

In order to compute an upper bound of the error (2.20), we introduce two adjoint problems, each of which is associated with one term on the right-hand side of (2.20). Thus, we seek $r_\Omega, s_\Omega \in V_\Omega$ such that

$$\begin{aligned} a_\Omega(\delta r, r_\Omega) &= \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, \delta r] \quad \forall \delta r \in V_\Omega, \\ a_\Omega(\delta s, s_\Omega) &= \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, \delta s] \quad \forall \delta s \in V_\Omega. \end{aligned} \quad (2.21)$$

We remark that in order for the aforementioned adjoint problems to be well-posed, their right-hand sides have to be linear and continuous forms on V_Ω and this motivates the linearization introduced in (2.20). In the following sections, we present different strategies to exploit the information in (2.21) to construct an *a posteriori* error estimator for the quantity E^h in (2.20).

2.3 Estimate of the error in the shape gradient based on energy-norm error estimates

Let us denote by r_Ω^h, s_Ω^h the approximations of the solutions r_Ω, s_Ω of equations (2.21) arising from a conforming Finite Element discretization. By plugging (2.21) into (2.20) and owing to the Galerkin orthogonality, the formulation (2.19) of the quantity \tilde{E}^h is obtained:

$$\tilde{E}^h = a_\Omega(u_\Omega - u_\Omega^h, r_\Omega - r_\Omega^h) + a_\Omega(p_\Omega - p_\Omega^h, s_\Omega - s_\Omega^h).$$

Moreover, we derive the following upper bound \bar{E} for the numerical error in the shape gradient:

$$\begin{aligned} |E^h| &\simeq |\tilde{E}^h| \leq |a_\Omega(u_\Omega - u_\Omega^h, r_\Omega - r_\Omega^h)| + |a_\Omega(p_\Omega - p_\Omega^h, s_\Omega - s_\Omega^h)| \\ &\leq \|u_\Omega - u_\Omega^h\|_\Omega \|r_\Omega - r_\Omega^h\|_\Omega + \|p_\Omega - p_\Omega^h\|_\Omega \|s_\Omega - s_\Omega^h\|_\Omega =: \bar{E} \end{aligned} \quad (2.22)$$

where $\|\cdot\|_\Omega$ is the energy-norm induced by the bilinear form $a_\Omega(\cdot, \cdot)$. The first inequality follows from triangle inequality whereas the upper bound \bar{E} is derived exploiting Cauchy-Schwarz inequality.

Remark 2.5. In (2.22) we derived an upper bound of the numerical error in the linearized shape gradient and not in the shape gradient itself. In this thesis, we follow the framework in [149] by assuming the linearization error to be negligible and \bar{E} to be an upper bound of the numerical error E^h itself and

not of its linearized version \tilde{E}^h . In subsection 4.3.1, a validation of the error estimator is presented for the case of Electrical Impedance Tomography: in particular, we verify that the linearization error is indeed negligible with respect to the error due to the Finite Element discretization and thus the previous assumption stands.

In order to fully compute the error estimator (2.22), we have to estimate the energy-norm of the error for:

- the state equation;
- the adjoint equation used to compute the shape gradient;
- the two adjoint equations associated with the Quantity of Interest.

Several strategies are possible to tackle these problems. In the following subsections, we present a brief overview of the existing techniques to compute *a posteriori* estimates in the energy-norm of the error due to the Finite Element approximation of a Boundary Value Problem. Several works in the literature have been dealing with this subject since the seminal paper of Babuška and Rheinboldt [46]. For a detailed introduction to the subject, we refer to [13, 213, 274]. We remark that these estimates are problem-dependent, thus in this chapter we only present the abstract framework for their derivation highlighting advantages and disadvantages of each strategy. A detailed description will be provided in chapter 4 for the case of Electrical Impedance Tomography and in chapter 7 for the problem of shape optimization in linear elasticity.

Remark 2.6. In order to verify the certification condition in (2.11), a key aspect is represented by the ability of computing the *a posteriori* error estimates as precisely as possible. Thus, from now on we focus on deriving fully computable constant-free estimates of the error in the energy-norm.

2.3.1 Complementary energy principle

In this subsection, we describe a method inspired by the complementary energy principle which allows to derive fully computable, constant-free estimators of the error in the energy-norm by solving an additional variational problem to retrieve an accurate approximation of the numerical flux.

In order to present the main idea for the construction of the aforementioned error estimator, we introduce the following Boundary Value Problem as test case:

$$\begin{cases} -\Delta u_\Omega = f & \text{in } \Omega \\ u_\Omega = 0 & \text{on } \partial\Omega \end{cases} \quad (2.23)$$

We seek $u_\Omega \in V_\Omega := H_0^1(\Omega)$ such that

$$a_\Omega(u_\Omega, \delta u) := \int_\Omega \nabla u_\Omega \cdot \nabla \delta u \, dx = \int_\Omega f \delta u \, dx =: F_\Omega(\delta u) \quad \forall \delta u \in V_\Omega. \quad (2.24)$$

Let $u_\Omega^h \in V_\Omega^h$ be the Finite Element approximation of problem (2.24) and e^h the corresponding numerical error. Owing to the Mikhlin's identity for quadratic functionals (cf. [199]), the residual equation (2.18) may be rewritten as the following minimum potential energy principle:

$$E_\Omega(e^h) = \inf_{w \in V_\Omega} E_\Omega(w) \quad , \quad E_\Omega(w) := \frac{1}{2} \int_\Omega |\nabla w|^2 \, dx + \int_\Omega \nabla u_\Omega^h \cdot \nabla w \, dx - F_\Omega(w). \quad (2.25)$$

We remark that $a_\Omega(\cdot, \cdot)$ and $F_\Omega(\cdot)$ being bounded on V_Ω , there exists a unique solution to the minimum potential energy problem. A duality argument (cf. [158]) leads to the definition of the following dual complementary energy principle:

$$C_\Omega(\sigma_\Omega) = \sup_{\substack{\sigma \in H(\text{div}) \\ \nabla \cdot \sigma + f = 0 \text{ in } \Omega}} C_\Omega(\sigma) \quad , \quad C_\Omega(\sigma) := -\frac{1}{2} \int_\Omega |\sigma - \nabla u_\Omega^h|^2 dx \quad (2.26)$$

where $H(\text{div}) := \{\tau \in L^2(\Omega; \mathbb{R}^d) : \nabla \cdot \tau \in L^2(\Omega)\}$. The unique solution of the complementary problem is $\sigma_\Omega = \nabla u_\Omega$ and the following estimate may be derived (cf. [235]):

$$\|e^h\|_\Omega \leq \sqrt{-2C_\Omega(\sigma_\Omega)} \quad \forall \sigma_\Omega \in H(\text{div}) \text{ s.t. } \nabla \cdot \sigma_\Omega + f = 0 \text{ in } \Omega. \quad (2.27)$$

A computable upper bound of the error in the energy-norm may be derived via a suitable choice of the dual variable σ_Ω under the constraint $\nabla \cdot \sigma_\Omega + f = 0$ in Ω . We remark that the aforementioned condition has to be fulfilled exactly. One possibility to obtain a suitable σ_Ω was suggested by Aubin and Burchard [43] and Fraeijs de Veubeke [137] through the introduction of a Finite Element approximation of the dual complementary problem. An alternative procedure based on the Aubin-Nitsche trick has been proposed by Babuška and Rheinboldt in [48] and relies on the relaxation of the constraint by means of the solution of a perturbed problem.

The discussed method requires the solution of an additional global problem. Though this approach is known to provide extremely precise upper bounds of the error in the energy-norm, the computational effort required by the solution of a global problem is comparable with the one needed to compute the Finite Element approximation of problem (2.24). Nevertheless, within the framework of the certification procedure described in subsection 2.1.3, it appears more important retrieving a sharp fully computable estimate than reducing the computational cost of the overall algorithm. Hence, the complementary energy principle represents a good starting point for the derivation of a reliable estimator of the error in the shape gradient using the energy-norm estimates of the errors for the state and adjoint equations.

2.3.2 Explicit residual estimates

A first approach to the construction of *a posteriori* error estimates based on local quantities relies on the direct computation of the numerical residues on mesh elements and at the interface between elements. This class of estimators was first derived by Babuška and Rheinboldt [46, 47] and later developed by Babuška and co-workers in [45, 172].

Let us consider the model problem (2.23) and its corresponding variational formulation (2.24). The starting point is the representation (2.18) for the error e^h . We split the terms in (2.18) into their local contributions on all the elements T of the triangulation \mathcal{T}_h approximating the domain Ω :

$$a_\Omega(e^h, \delta e) = \sum_{T \in \mathcal{T}_h} \left[\int_T f \delta e \, dx - \int_T \nabla u_\Omega^h \cdot \nabla \delta e \, dx \right] \quad \forall \delta e \in V_\Omega. \quad (2.28)$$

Integrating by parts the last term in (2.28) and introducing the set of all the edges \mathcal{E}_h of the triangulation \mathcal{T}_h , we obtain that

$$a_\Omega(e^h, \delta e) = \sum_{T \in \mathcal{T}_h} \int_T \rho_T \delta e \, dx + \sum_{e \in \mathcal{E}_h} \int_e \rho_e \delta e \, ds \quad \forall \delta e \in V_\Omega \quad (2.29)$$

where ρ_T and ρ_e respectively represent the interior element residual and the boundary residual terms:

$$\rho_T := f + \Delta u_\Omega^h \quad , \quad \rho_e := \begin{cases} (\nabla u_\Omega^h|_T \cdot n_T + \nabla u_\Omega^h|_{T'} \cdot n_{T'}) & , \partial\Omega \not\supset e = T \cap T' \\ 0 & , e \subset \partial\Omega \end{cases} . \quad (2.30)$$

In (2.30), n_T stands for the outward normal vector to the boundary ∂T of the element T . For the edges not belonging to the external boundary of the domain, the boundary residual ρ_e accounts for the interelement jump of the gradient across the edge e separating elements T and T' .

Let us introduce the interpolant $\pi^h \delta u$ to δu in the subspace $V_\Omega^h \subset V_\Omega$. From (2.29) and owing to the Galerkin orthogonality, it follows that

$$a_\Omega(e^h, \delta e) = \sum_{T \in \mathcal{T}_h} \int_T \rho_T (\delta e - \pi^h \delta e) \, dx + \sum_{e \in \mathcal{E}_h} \int_e \rho_e (\delta e - \pi^h \delta e) \, ds \quad \forall \delta e \in V_\Omega. \quad (2.31)$$

Applying the Cauchy-Schwarz inequality to (2.31) and exploiting the interpolation estimates for the Cl  ment operator (cf. [113]), we obtain

$$\begin{aligned} a_\Omega(e^h, \delta e) &\leq \sum_{T \in \mathcal{T}_h} \|\rho_T\|_{L^2(T)} \|\delta e - \pi^h \delta e\|_{L^2(T)} + \sum_{e \in \mathcal{E}_h} \|\rho_e\|_{L^2(e)} \|\delta e - \pi^h \delta e\|_{L^2(e)} \\ &\leq C |\delta e|_{H^1(\Omega)} \left[\sum_{T \in \mathcal{T}_h} h_T^2 \|\rho_T\|_{L^2(T)}^2 + \sum_{e \in \mathcal{E}_h} h_T \|\rho_e\|_{L^2(e)}^2 \right]^{\frac{1}{2}} \end{aligned} \quad (2.32)$$

where C is a constant independent of the δe and h_T . Noting that $|\delta e|_{H^1(\Omega)} \leq \|\delta e\|_\Omega$ and replacing δe by e^h , we obtain the following explicit residual estimate of the error in the energy-norm:

$$\|e^h\|_\Omega^2 \leq C \left[\sum_{T \in \mathcal{T}_h} h_T^2 \|\rho_T\|_{L^2(T)}^2 + \sum_{e \in \mathcal{E}_h} h_T \|\rho_e\|_{L^2(e)}^2 \right] \quad (2.33)$$

which can be equivalently rewritten in terms of the local contributions η_T^e 's

$$\|e^h\|_\Omega^2 \leq C \sum_{T \in \mathcal{T}_h} \left[h_T^2 \|\rho_T\|_{L^2(T)}^2 + \frac{1}{2} h_T \|\rho_e\|_{L^2(\partial T)}^2 \right] =: C \sum_{T \in \mathcal{T}_h} (\eta_T^e)^2. \quad (2.34)$$

The quantities on the right-hand side of (2.34) can be explicitly evaluated starting from the data of the problem and the computed Finite Element approximation. Nevertheless, the constant C is in general unknown. As a matter of fact, in order to fully compute the constant C , several contributions should be accounted for. In particular, the constants arising from the interpolation estimates, from the Poincar   inequality, from the trace inequality and from the inequality to control the H^1 -seminorm by means of the energy-norm. Some investigations have been conducted in the literature to evaluate these constants but this approach requires a high computational effort and does not lead to the construction of estimators useful for practical applications. As a matter of fact, the values computed for these constants (cf. e.g. [56, 99, 171]) are usually related to worst-case scenarios and generate rather pessimistic error bounds. Moreover, the use of the Cauchy-Schwarz inequality to move from

(2.31) to (2.32) is responsible for the degradation of the final estimate since the cancellation of errors over the domain is lost and this may result in a severe inaccuracy of the computed bound. Hence, though estimate (2.34) may provide interesting information to drive a mesh adaptation procedure, its qualitative nature makes it of no practical use when dealing with the certification procedure described in subsection 2.1.3.

2.3.3 Implicit residual estimates

In order to circumvent the aforementioned issues of explicit error estimates, several authors investigated the so-called implicit estimators. This class of *a posteriori* error estimates deals with the solution of an additional local Boundary Value Problem (BVP) approximating the residual equation (2.18). These problems are local in the sense that they are defined either over a single element (element residual method) or on a small subdomain (subdomain residual method) and exploit the residuals introduced in the previous subsection as data. No unknown constants arise from this approach but the number of additional BVP's may not be negligible. Nevertheless, the local nature of these problems allows for a straightforward parallelization of the procedure to approximate them.

Element residual method

From (2.18), it is straightforward to observe that on each element T of the triangulation the error e^h satisfies

$$a_T(e^h, \delta e) = F_T(\delta e) - a_T(u_\Omega^h, \delta e) + \int_{\partial T} \nabla u_\Omega \cdot n_T \delta e \, ds \quad \forall \delta e \in V_T \quad (2.35)$$

where $V_T := \{u \in H^1(T) : u = 0 \text{ on } \partial T \cap \partial\Omega\}$ and $a_T(\cdot, \cdot)$ (respectively $F_T(\cdot)$) is the restriction of the bilinear (respectively linear) form to the local element T .

We remark that if the boundary ∂T of the element T matches an interelement boundary, the flux $\nabla u_\Omega \cdot n_T$ is unknown and it has to be replaced by a numerical approximation. A common choice in the literature (cf. [55]) is represented by the following average of the discontinuous Finite Element flux to retrieve an approximation of the true flux:

$$\nabla u_\Omega \cdot n_T \approx \{\{\nabla u_\Omega^h \cdot n_T\}\} := \frac{1}{2} \left(\nabla u_\Omega^h|_T + \nabla u_\Omega^h|_{T'} \right) \cdot n_T \quad \text{on } e \subset \partial T \quad (2.36)$$

where T' is the element of the triangulation sharing the edge e with the element T .

The formulation of the element-wise error residual problem leads to the definition of one variational problem on each element of the mesh. The approximation of these problems is performed in appropriate discrete spaces, constructed starting from the Finite Element space used for the solution of equation (2.24). Several strategies [55, 273] have been proposed in the literature for the definition of the aforementioned spaces and they mainly rely on the idea of increasing the order of the original Finite Element space V_Ω^h by introducing bubble functions. Let W_T^h be the restriction to the mesh element T of the discrete space obtained from the addition of bubble functions to V_Ω^h . We seek $\varsigma_T \in W_T^h$ such that

$$a_T(\varsigma_T, \delta e) = F_T(\delta e) - a_T(u_\Omega^h, \delta e) + \int_{\partial T} \{\{\nabla u_\Omega^h \cdot n_T\}\} \delta e \, ds \quad \forall \delta e \in W_T^h. \quad (2.37)$$

A key issue of problem (2.37) is the existence of a solution which is guaranteed only if the data satisfy appropriate compatibility conditions. In order to ensure the well-posedness of problem (2.37), a careful

choice of the boundary data has to be performed and we refer to [13] for a more detailed discussion on this topic.

From the solution ς_T of problem (2.37), the following implicit estimator may be derived for the element residual method [49, 50]:

$$\|e^h\|_\Omega^2 \leq \sum_{T \in \mathcal{T}_h} (\eta_T^i)^2 \quad , \quad \eta_T^i := \|\varsigma_T\|_T. \quad (2.38)$$

Subdomain residual method

An alternative method to derive implicit *a posteriori* error estimators relies on decomposing equation (2.18) into local contributions on small element patches. Following [45, 46], we define local residual problems featuring homogeneous essential boundary conditions over each subdomain of Ω . In recent years, several authors [98, 195, 203] have investigated variants of the subdomain residual method by proposing alternative choices of boundary conditions.

Here, we briefly sketch the variant proposed by Prudhomme *et al.* in [230], since it will be at the base of the equilibrated fluxes method of subsection 2.4.3. We consider a number of subdomains equal to the number N_v of vertices x_v 's of the mesh. Each patch - denoted by ω_v - features all the elements $T \in \mathcal{T}_h$ sharing the node x_v . For each vertex x_v , $v = 1, \dots, N_v$ we introduce a linear shape function ψ_v such that $\psi_v(x_w) = \delta_{vw}$, δ being the classical Kronecker delta. Moreover, the family of functions ψ_v 's fulfills the condition known as partition of the unity, that is

$$\sum_{v=1}^{N_v} \psi_v = 1.$$

Let us rewrite (2.18) owing to the splitting introduced above:

$$a_\Omega(e^h, \delta e) = \sum_{v=1}^{N_v} a_\Omega(e^h, \psi_v \delta e) = \sum_{v=1}^{N_v} \left[F_\Omega(\psi_v \delta e) - a_\Omega(u_\Omega^h, \psi_v \delta e) \right] \quad \forall \delta e \in V_\Omega. \quad (2.39)$$

By introducing the local bilinear form $a_v : H_0^1(\omega_v) \times H_0^1(\omega_v) \rightarrow \mathbb{R}$ and the local linear form $F_v : H_0^1(\omega_v) \rightarrow \mathbb{R}$ as the restrictions to ω_v of the global forms $a_\Omega(\cdot, \cdot)$ and $F_\Omega(\cdot)$, we obtain the following variational problem on each subpatch: we seek $\varsigma_v \in H_0^1(\omega_v)$ such that

$$a_v(\varsigma_v, \delta e) = F_v(\delta e) - a_v(u_\Omega^h, \delta e) \quad \forall \delta e \in H_0^1(\omega_v). \quad (2.40)$$

Plugging (2.40) into (2.39) and considering $\delta e = e^h$, we obtain

$$\|e^h\|_\Omega = \sqrt{a_\Omega(e^h, e^h)} \leq \sqrt{\sum_{v=1}^{N_v} a_v(\varsigma_v, \varsigma_v)}. \quad (2.41)$$

We remark that in general the exact solutions ς_v 's are unknown. Thus, to fully compute (2.41) we introduce an high-order Finite Element space W_v^h on each subpatch ω_v and we seek $\varsigma_v^h \in W_v^h$ such that it approximates problem (2.40) for $v = 1, \dots, N_v$. The following error estimate may be derived (cf. [230]):

$$\|e^h\|_\Omega \leq \sqrt{\sum_{v=1}^{N_v} a_v(\varsigma_v^h, \varsigma_v^h)} + 2 \inf_{w \in W_\Omega^h} \|u_\Omega - w\|_\Omega \quad , \quad \forall w \in W_\Omega^h \quad (2.42)$$

where W_Ω^h is the high-order Finite Element space defined on the global domain Ω . The choice of high-order spaces for the approximation of the local problems on the subpatches is needed for the derivation of a sharp estimate (2.42). We highlight that the second term on the right-hand side of (2.42) decreases as the polynomial degree of the functions in W_Ω^h increases (cf. [230]). Within this framework, in many practical applications the second term in (2.42) may be neglected and the following implicit estimator may be derived for the subdomain residual method:

$$\|e^h\|_\Omega^2 \leq \sum_{v=1}^{N_v} (\eta_v^i)^2, \quad \eta_v^i := \|\varsigma_v^h\|_{\omega_v}. \quad (2.43)$$

The estimators (2.38) and (2.43) localize the source of the numerical error inside the domain, thus they provide useful information to drive a mesh adaptation procedure. Moreover, they are fully computable and do not feature unknown constants or quantities, making them useful also on a quantitative level. Hence, both the estimators may be used to perform the certification procedure described in subsection 2.1.3. Nevertheless, these two strategies present some drawbacks that limit their interest when dealing with practical applications. On the one hand, the element residual method requires a careful choice of the boundary data in order for the local problems to be well-posed. This may limit its applicability to simple BVP's, restricting its interest to academic problems. On the other hand, the subdomain residual method localizes the bilinear form and allows a straightforward way to impose the boundary conditions for the local problems by implicitly accounting for the effect of the interelement jump of the flux. However, the computational cost associated with the solutions of the local problems may rapidly become unfeasible to manage, especially when dealing with 3D applications.

2.3.4 Equilibrated residual estimates

The methods discussed in the previous subsection remain unsatisfactory from a practical point of view. As a matter of fact, the local problems defined by the element residual method may not have a solution, whereas the computational cost associated with the subdomain residual method may quickly explode making this strategy unfeasible. In order to remedy these issues, a procedure to improve the element residual method by constructing boundary data for the local problems in equilibrium with the interior residual has been proposed by Ladevèze and Leguillon [186]. Starting from this paper, several other authors have been working on this subject and we refer the interested reader to [12, 55].

The starting point for the development of the equilibrated residual method is the error residual equation (2.18) and its formulation (2.25) as an optimization problem. Contrary to the solution of the global problem discussed in subsection 2.3.1, here we consider a strategy to reduce equation (2.18) to a sequence of independent problems posed locally over each mesh element. Owing to the local nature of these problems and to their small size, their approximation is computationally inexpensive and easy to parallelize. As per the element residual method, a key role in this approach is played by the approximation of the unknown true flux on the interelement boundaries. Let $\{g_T : T \in \mathcal{T}_h\}$ be a set of boundary fluxes that approximate the flux of the true solution on all the elements T of the triangulation:

$$g_T \approx \nabla u_\Omega|_T \cdot n_T \quad \text{on} \quad \partial T. \quad (2.44)$$

Owing to the continuity of the interelement fluxes, we get

$$g_T + g_{T'} = 0 \quad \text{on} \quad \partial T \cap \partial T'. \quad (2.45)$$

As in (2.35), we decompose the residual error into contributions over individual elements and we exploit the definition (2.44) of the interelement fluxes:

$$a_T(e^h, \delta e) = F_T(\delta e) - a_T(u_\Omega^h, \delta e) + \int_{\partial T} g_T \delta e \, ds \quad \forall \delta e \in V_T. \quad (2.46)$$

The local residue may be represented in terms of the solution $\varsigma_f \in V_T$ of the following local problem

$$a_T(\varsigma_f, \delta e) = F_T(\delta e) - a_T(u_\Omega^h, \delta e) + \int_{\partial T} g_T \delta e \, ds \quad \forall \delta e \in V_T \quad (2.47)$$

where V_T is the space previously introduced for the element residual method. We remark that the local Neumann Boundary Value Problem (2.47) has a unique solution if and only if the following compatibility condition - also known as equilibration condition - is fulfilled:

$$F_T(1) - a_T(u_\Omega^h, 1) + \int_{\partial T} g_T \, ds = 0. \quad (2.48)$$

Remark 2.7. The condition (2.48) expresses the fact that the boundary flux g_T is in equilibrium with the interior load given by the restriction of the element residual to T . Though condition (2.48) may not be strictly necessary for the well-posedness of the local problems when dealing with equations different from the Laplace problem under analysis, it still has a non-negligible effect on the quality of the computed upper bounds and must be satisfied exactly as the mesh size tends to zero (cf. [13]).

Following the same procedure described for the element residual method, we may derive the implicit estimator for the equilibrated residual method starting from the solution ς_f of problem (2.47):

$$\|e^h\|_\Omega^2 \leq \sum_{T \in \mathcal{T}_h} (\eta_T^f)^2 \quad , \quad \eta_T^f := \|\varsigma_f\|_T. \quad (2.49)$$

The equilibrated residual method allows for an inexpensive evaluation of the local contributions of the global error in the energy-norm owing to the appropriate construction of the local flux g_T for each element T . In order for the flux g_T to be suitable for the accurate estimate of the error, conditions (2.45) and (2.48) have to be fulfilled. These conditions are known as zeroth-order equilibration conditions since the reconstructed fluxes are compatible with respect to constant test functions in (2.48). An extension to p th-order equilibration conditions is possible and we refer the interested reader to [13]. A key aspect of the discussed strategy is the practical construction of the equilibrated fluxes. In [13], the authors provide an extensive description of the technical details for this procedure and discuss several examples on different computational meshes. The main disadvantage of this approach is related to the necessity of concurrently handling information arising from both the geometry of the triangulation and the discretization of the Boundary Value Problem. Hence, its integration into existing Finite Element codes may be extremely difficult. In next section, we discuss an alternative strategy to circumvent these issues, namely the so-called equilibrated fluxes method (cf. subsection 2.4.3).

2.4 Improving the estimate of the error in the shape gradient

The main drawback of estimate (2.22) lies in the use of the triangle and the Cauchy-Schwarz inequalities for the derivation of the upper bound of the error in the shape gradient. In order to bypass the loss of accuracy due to the aforementioned inequalities, we derive a goal-oriented estimator starting from the variational formulation of the error (cf. [145]). Following the framework discussed in section 2.3, we perform the linearization of the quantity whose error estimate is sought as in (2.20) and we consider the adjoint problems (2.21). For the corresponding conforming Finite Element (respectively Discontinuous Galerkin) discretizations of problems (2.21), we seek $r_\Omega^h, s_\Omega^h \in V_\Omega^{h,m}$ such that

$$\begin{aligned} a_\Omega^h(\delta r^h, r_\Omega^h) &= \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, \delta r^h] \quad \forall \delta r^h \in V_\Omega^{h,m}, \\ a_\Omega^h(\delta s^h, s_\Omega^h) &= \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, \delta s^h] \quad \forall \delta s^h \in V_\Omega^{h,m}. \end{aligned} \quad (2.50)$$

Let us define two quantities associated respectively with the contribution of the state error $u_\Omega - u_\Omega^h$ and the adjoint error $p_\Omega - p_\Omega^h$ to the error in the shape gradient E^h . By exploiting (2.16), we get:

$$E_u^h = \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, u_\Omega - u_\Omega^h] = F_\Omega(r_\Omega) - a_\Omega^h(u_\Omega^h, r_\Omega), \quad (2.51)$$

$$E_p^h = \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial p}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, p_\Omega - p_\Omega^h] = - \left\langle \frac{\partial j}{\partial u}(\Omega, u_\Omega), s_\Omega \right\rangle - a_\Omega^h(p_\Omega^h, s_\Omega). \quad (2.52)$$

It is straightforward to observe that $E^h = E_u^h + E_p^h$. In order to derive a practical strategy to perform the certification procedure in subsection 2.1.3 by verifying condition (2.11), an upper bound \bar{E} of the error $|E^h|$ is required.

Remark 2.8. It is well-known in the literature (cf. e.g. [14]) that in order to retrieve a sharp upper bound for the error in a Quantity of Interest, high-order approximations should be employed for the solution of the adjoint problem. Hence, from now on we consider $m > \ell$.

As previously done for the *a posteriori* error estimates in the energy-norm, in the following subsections we review some techniques discussed in the literature to construct the estimates of the error in a given functional of interest. As per the case of the error in the energy-norm, we are interested in constructing estimators in a Quantity of Interest which are fully computable and possibly constant-free. Having this in mind, in the rest of this chapter we focus on the abstract framework for the derivation of the goal-oriented error estimators by highlighting their upsides and downsides. The details of the construction of the estimator for the error in the shape gradient - that is the procedure to estimate the error in E_u^h and E_p^h - will be presented for the case of Electrical Impedance Tomography in chapter 5.

2.4.1 Dual Weighted Residual method

Following the previously discussed approach for the construction of explicit estimates of the error in the energy-norm, Rannacher and co-workers [52] proposed an analogous estimator for the error in a Quantity of Interest. Starting from (2.19) and exploiting the information on the interior element

residual and the boundary residual (2.30), we obtain

$$Q(e^h) = a_\Omega(e^h, e^h) = \sum_{T \in \mathcal{T}_h} \left[\int_T \rho_T(r_\Omega - r_\Omega^h) \, dx + \frac{1}{2} \int_{\partial T} \rho_e(r_\Omega - r_\Omega^h) \, ds \right]. \quad (2.53)$$

This method is known as Dual Weighted Residual since it constructs an estimate of the error in a Quantity of Interest featuring the residual terms (2.30) of the state problem weighted by the error of the adjoint problem. Hence, it is confirmed the role of the influence function r_Ω which carries the information of the effect of the error due to the approximation of the state problem on the error in the target functional. We remark that (2.53) is not fully computable since the exact solution of the adjoint problem is unknown. A possible workaround for this issue relies on the solution of the adjoint problem by means of a higher-order approximation: within this framework, the unknown solution r_Ω is replaced by its high-order numerical counterpart whereas r_Ω^h is substituted by the projection of the higher-order approximation onto the space of low-order Finite Element functions used to compute the solution of the state problem. Let $I_m^\ell : V_\Omega^{h,m} \rightarrow V_\Omega^{h,\ell}$ be the projection operator from the space of high-order functions to the low-order one. The resulting estimator for the Quantity of Interest reads as follows:

$$Q(e^h) \approx \sum_{T \in \mathcal{T}_h} \left[\int_T \rho_T(r_\Omega^h - I_m^\ell r_\Omega^h) \, dx + \frac{1}{2} \int_{\partial T} \rho_e(r_\Omega^h - I_m^\ell r_\Omega^h) \, ds \right]. \quad (2.54)$$

An alternative approach relies on the approximation of the adjoint problem by means of the same numerical scheme used for the state problem. A post-processing technique is then applied to the numerical solution r_Ω^h to retrieve a higher-order interpolated function. Let $I_\ell^m : V_\Omega^{h,\ell} \rightarrow V_\Omega^{h,m}$ be the operator that starting from a low-order function constructs its high-order interpolation. The resulting estimator of the error in the Quantity of Interest has the form

$$Q(e^h) \approx \sum_{T \in \mathcal{T}_h} \left[\int_T \rho_T(I_\ell^m r_\Omega^h - r_\Omega^h) \, dx + \frac{1}{2} \int_{\partial T} \rho_e(I_\ell^m r_\Omega^h - r_\Omega^h) \, ds \right]. \quad (2.55)$$

Both strategies provide fully computable goal-oriented error estimates and have been successfully tested on various problems over the years, showing effectivity indices close to 1.0 (cf. [57]). A major drawback of the Dual Weighted Residual (DWR) approach is represented by the fact that both (2.54) and (2.55) are asymptotically exact but may underestimate the error on coarse meshes since the solution of the adjoint problem is unknown. Several works in the literature have dealt with the analysis of this method, especially on the possible choices to derive a computable version of (2.53). In [214], the authors propose a safeguarded version of the DWR introducing an additional term to estimate the error due to the approximation of the adjoint problem. In a similar fashion, in [14] a detailed discussion on the different choices for the discretization of the adjoint problem is presented and extensive numerical tests on the optimal degree of the approximating functions are reported.

2.4.2 Energy-norm estimate via the parallelogram identity

A technique to provide sharper bounds of the error in a Quantity of Interest has been proposed by Prudhomme and Oden in a series of works starting from late 1990's. In [215, 231], the authors

propose to employ the parallelogram identity instead of the Cauchy-Schwarz inequality for problems featuring a symmetric bilinear form. Owing to this idea, we may rewrite (2.19) as

$$Q(e^h) = a_\Omega(e^h, e^h) = \frac{1}{4} \left\| \alpha e^h + \frac{1}{\alpha} \epsilon^h \right\|_\Omega^2 - \frac{1}{4} \left\| \alpha e^h - \frac{1}{\alpha} \epsilon^h \right\|_\Omega^2 \quad (2.56)$$

where the scaling factor $\alpha \in \mathbb{R}$ is such that $\alpha := \sqrt{\|\epsilon^h\|_\Omega \|\epsilon^h\|_\Omega^{-1}}$. Let $\eta_U^+, \eta_U^- \in \mathbb{R}$ (respectively $\eta_L^+, \eta_L^- \in \mathbb{R}$) be the following upper (respectively lower) bounds of the quantities on the right-hand side of (2.56):

$$\eta_L^+ \leq \left\| \alpha e^h + \frac{1}{\alpha} \epsilon^h \right\|_\Omega \leq \eta_U^+ \quad , \quad \eta_L^- \leq \left\| \alpha e^h - \frac{1}{\alpha} \epsilon^h \right\|_\Omega \leq \eta_U^- \quad (2.57)$$

By plugging (2.57) into (2.56), we obtain the following bounds for the error in the Quantity of Interest:

$$\frac{1}{4}(\eta_L^+)^2 - \frac{1}{4}(\eta_U^-)^2 \leq Q(e^h) \leq \frac{1}{4}(\eta_U^+)^2 - \frac{1}{4}(\eta_L^-)^2 \quad (2.58)$$

It is straightforward to observe that an upper bound of the error in the Quantity of Interest may be also obtained by neglecting the term η_L^- in (2.58): though the sharpness of the resulting estimator is degraded, this bound is less expensive to evaluate, proving to be useful in practical applications [216, 222].

From a practical point of view, the upper and lower bounds in (2.58) are computed via the implicit *a posteriori* error estimates presented in subsections 2.3.3 and 2.3.4. Nevertheless, as the aforementioned techniques to derive fully computable estimates of the error in the energy-norm are either computationally unfeasible or extremely difficult to implement in existing Finite Element codes, the advantages of this approach remain limited.

2.4.3 Equilibrated fluxes and flux-free approaches

As stated in section 2.3, a key aspect of a *good* estimator is the low computational cost associated with its derivation whence the great interest in estimators constructed using solely local quantities. In the last decade several authors have been working on efficient techniques to derive fully computable *a posteriori* estimators of the error in a Quantity of Interest. Within this framework, two main strategies have emerged. On the one hand, the equilibrated fluxes approach inspired by the equilibrated residual method. On the other hand, the flux-free approach originated by the subdomain residual method.

The equilibrated fluxes approach aims to reconstruct a suitable approximation of the unknown exact fluxes in the space $H(\text{div})$ in order for the discretized solution to fulfill the problem under analysis exactly. This method consists of two steps: first, a precise and efficient element-wise reconstruction of the numerical flux is performed in order to mimic the behavior of the unknown exact flux for both the state and adjoint problems; then, starting from the framework introduced at the beginning of section 2.4 the error in the target functional is locally evaluated. For the case of conforming Finite Element, local $H(\text{div})$ -reconstructions leading to fully computable upper bounds have been studied by several authors [78, 120, 191, 210, 275]. Moreover, we refer to [11, 114, 130, 175] for the main results obtained in recent years on equilibrated fluxes for Discontinuous Galerkin formulations.

The goal-oriented estimator associated with the flux-free method is derived from the upper bound obtained by neglecting the second term on the right-hand side of (2.56). In order to efficiently approximate this quantity, the error residual equations for both the state and the adjoint problems have to be

solved. Basic idea of the flux-free method is to circumvent the construction of the equilibrated fluxes required for the element-wise solution of these equations by posing the local problems on patches of elements, the interelement fluxes being consequently naturally equilibrated. This technique has been studied in several works [223, 224] in the literature and in [225, 241] the authors propose a variant to retrieve a more precise approximation of the numerical fluxes using a dual formulation in which the complementary energy is minimized.

One of the main advantages of the equilibrated fluxes method is the unified framework it provides for multiple Finite Element approximations. In chapter 5, we will detail this approach for conforming Finite Element and Discontinuous Galerkin discretizations applied to the problem of Electrical Impedance Tomography and we will highlight advantages and disadvantages of both choices. The flux-free approach - especially in its most recent variants - is a very promising technique for the derivation of accurate estimates of the error in a Quantity of Interest but further investigation are needed within the framework of the certification procedure described in subsection 2.1.3. We remark that despite being extremely different, the methods discussed in this subsection also present several points in common. Hence, we refer the interested reader to the paper [109] by Choi and Paraschivoiu for a detailed comparison of the equilibrated fluxes - also known as hybrid-flux - and the flux-free approaches.

2.5 The Certified Descent Algorithm

We are now ready to introduce the novel Certified Descent Algorithm (CDA), arising from the coupling of the Boundary Variation Algorithm for shape optimization (cf. subsection 2.1.2) and the goal-oriented estimator for the error in the shape gradient (cf. section 2.3 and 2.4). In script 2.3, we present a variant of algorithm 2.2 that takes advantage of the previously introduced *a posteriori* estimator of the error in the shape gradient in order to bypass the issues in the identification of a genuine descent direction due to the discretization error.

First, the procedure constructs an initial computational domain. At each iteration, the algorithm solves the state and adjoint problems [steps 1 and 2] and computes a descent direction θ^h solving equation (2.5) [step 3]. Then, the adjoint problems (2.21) are solved and an upper bound of the numerical error in the shape gradient along the direction θ^h is computed [step 4]. If condition (2.11) is not fulfilled, the mesh is adapted in order to improve the error estimate. This procedure is iterated until the direction θ^h is a certified descent direction for $J(\Omega)$ [step 5]. Once a certified descent direction has been identified, we compute a step [step 6] such that the following Armijo condition is fulfilled: let us consider the iteration j , given $0 < \alpha < 1$ we use a backtracking strategy to identify the step $\mu_j \in \mathbb{R}^+$ such that

$$J\left(\left(\text{Id} + \mu_j \theta_j^h\right) \Omega_j\right) \leq J(\Omega_j) + \alpha \mu_j \langle d_h J(\Omega_j), \theta_j^h \rangle.$$

An alternative bisection-based line search technique has been proposed by Morin *et al.* in [201].

Then the shape of the domain is updated according to the computed perturbation of the identity $\text{Id} + \mu_j \theta_j^h$ [step 7]. Eventually, a novel stopping criterion is proposed [step 8] in order to use the information embedded in the error bound \overline{E} to derive a reliable condition to end the evolution of the algorithm.

Algorithm 2.3 – The Certified Descent Algorithm

Given the domain Ω_0 , set $\text{tol} > 0$, $j = 0$ and iterate:

1. Compute the solution $u_{\Omega_j}^h$ of the state equation;
2. Compute the solution $p_{\Omega_j}^h$ of the adjoint equation;
3. Compute a descent direction $\theta_j^h \in X$ solving

$$(\theta_j^h, \delta\theta)_X + \langle d_h J(\Omega_j), \delta\theta \rangle = 0 \quad \forall \delta\theta \in X ;$$

4. Compute an upper bound \overline{E} of the numerical error E^h ;
 5. If $\langle d_h J(\Omega_j), \theta_j^h \rangle + \overline{E} \geq 0$, refine the mesh and go to 1;
 6. Identify an admissible step μ_j ;
 7. Update the shape $\Omega_{j+1} = (\text{Id} + \mu_j \theta_j^h) \Omega_j$;
 8. While $|\langle d_h J(\Omega_j), \theta_j^h \rangle| + \overline{E} > \text{tol}$, $j = j + 1$ and repeat.
-

The advantage of the Certified Descent Algorithm is twofold. On the one hand, the computation of the upper bound of the numerical error in the shape gradient provides useful information to identify a certified descent direction at each iteration of the optimization algorithm and to construct a guaranteed shape optimization strategy. Hence, the resulting objective functional improves at each iteration of the optimization loop. On the other hand, the fully computable and constant-free error estimator provides quantitative information to derive a reliable stopping criterion for the overall optimization procedure.

The novelty of algorithm 2.3 is the certification procedure that plays a crucial role in steps 4, 5 and 8. A key aspect of the procedure is the mesh adaptation routine that has to be run if condition (2.11) is not fulfilled by the current configuration. From a practical point of view, in order to guarantee that condition (2.11) is fulfilled in a reasonable number of iterations of the adaptation routine, we construct a refinement strategy to explicitly reduce the error \overline{E} at each iteration. In particular, we perform a goal-oriented mesh adaptation procedure as suggested in [215]: at each iteration, we construct the upper bound \overline{E} and an indicator based on the estimator of the error in the shape gradient. This approach exploits the previously constructed estimator to localize the regions of the domain that are mainly responsible for the error in the Quantity of Interest and performs a targeted refinement in order to concurrently reduce the error in the shape gradient and limit the number of newly inserted Degrees of Freedom. The efficiency of this strategy has been extensively studied in the literature (cf. [135, 232]) and the numerical results in chapters 4 and 5 confirm the ability of this method to reduce the target error.

2.5.1 CDA based on the complementary energy principle

A first version of the Certified Descent Algorithm is obtained from the estimate of the error in the shape gradient via the complementary energy principle (cf. subsection 2.3.1) applied to the energy-norm error estimates of the solutions of the state and adjoint problems. Owing to (2.22), step 4 of algorithm 2.3 reads as follows:

Algorithm 2.4 – Estimating the error in the shape gradient via the complementary energy principle

4. Compute an upper bound \overline{E} of the numerical error $|E^h|$:
 - (a) Compute the solutions $r_{\Omega_j}^h$ and $s_{\Omega_j}^h$ of the adjoint problems to estimate the error in the shape gradient;
 - (b) Compute the dual flux approximations for $u_{\Omega_j}^h$, $p_{\Omega_j}^h$, $r_{\Omega_j}^h$ and $s_{\Omega_j}^h$;
 - (c) Compute $\overline{E} = \|u_{\Omega_j} - u_{\Omega_j}^h\|_{\Omega} \|r_{\Omega_j} - r_{\Omega_j}^h\|_{\Omega} + \|p_{\Omega_j} - p_{\Omega_j}^h\|_{\Omega} \|s_{\Omega_j} - s_{\Omega_j}^h\|_{\Omega}$;
-

In script 2.4, step 4c is performed via the evaluation of the dual complementary energy as in (2.27) for the model problem of the Laplace equation (cf. subsection 2.3.1). Owing to the subsequent refinements of the mesh (cf. algorithm 2.3 - step 5), the approximated solution u_{Ω}^h tends to the exact solution u_{Ω} and analogously do the solutions p_{Ω}^h , r_{Ω}^h and s_{Ω}^h of the adjoint problems. Thus the term \overline{E} tends to zero as the mesh size tends to zero, assuring that the mesh adaptation procedure performed by the Certified Descent Algorithm eventually leads to the fulfillment of condition (2.11).

In chapter 4, we will present the application of this version of the Certified Descent Algorithm to the identification problem of Electrical Impedance Tomography. In chapter 7, we will provide the details for the construction of the *a posteriori* estimator of the error in the shape gradient of the compliance. The solution of the corresponding shape optimization problem in linear elasticity via the Certified Descent Algorithm will be the subject of the future work [5].

2.5.2 CDA based on an equilibrated fluxes approach

We construct an *a posteriori* error estimator in the shape gradient which relies solely on local quantities. In particular, we investigate a variant of the CDA based on the equilibrated fluxes approach (cf. subsection 2.4.3). Following the procedure described in section 2.4, we sketch the construction of the goal-oriented equilibrated fluxes estimator of the error in the shape gradient:

Algorithm 2.5 – Estimating the error in the shape gradient via an equilibrated fluxes approach

4. Compute an upper bound \overline{E} of the numerical error E^h :
 - (a) Compute the solutions $r_{\Omega_j}^h$ and $s_{\Omega_j}^h$ of the adjoint problems to estimate the error in the shape gradient;
 - (b) Compute the equilibrated fluxes for $u_{\Omega_j}^h$, $p_{\Omega_j}^h$, $r_{\Omega_j}^h$ and $s_{\Omega_j}^h$;
 - (c) Compute $\overline{E} = E_u^h + E_p^h$;
-

The *a posteriori* estimator of the error in the shape gradient is obtained using the information carried by the $H(\text{div})$ -reconstruction of the numerical fluxes as stated in subsection 2.4.3. Owing to the well-known asymptotic exactness of the goal-oriented error estimators constructed using the equilibrated fluxes approach (cf. [14, 205]), the term \overline{E} tends to zero as the mesh size tends to zero. This property plays a crucial role since it guarantees that the mesh adaptation routine performed to certify the descent direction (cf. algorithm 2.3 - step 5) eventually leads to the fulfillment of condition (2.11).

In chapter 5, we will detail the construction of the equilibrated fluxes and the corresponding goal-oriented estimator for the error in the shape gradient for the test problem of Electrical Impedance Tomography.

2.5.3 The mesh adaptation procedure

A key aspect of algorithm 2.3 is the mesh adaptation procedure on step 5. Starting from the upper bound of the error \overline{E} , computed as discussed in subsections 2.5.1 and 2.5.2, we construct an error indicator ι_T that accounts for the local contribution of the element T of the triangulation \mathcal{T}_h to the global error in the domain Ω .

In this thesis, all the numerical simulations are performed using FreeFem++ [160] which relies on BAMG [159] for mesh generation and mesh adaptation procedures. In particular, in order to perform step 5 of algorithm 2.3 we consider a greedy approach that subdivides all the elements T 's of the triangulation such that the corresponding indicator ι_T is greater than a user-defined positive constant C . Thus, the algorithm relies on the construction of a metric guaranteeing that the error in each element is lower than a fixed threshold C . From a practical point of view, the computational mesh is modified by means of the `adaptmesh` command in FreeFem++ which allows both the refinement and the coarsening of the elements owing to the information carried by the aforementioned defined metric. A qualitative and quantitative discussion on the meshes resulting from this adaptation procedure is provided in chapters 4, 5 and 6 where some numerical simulations are presented.

Remark 2.9. The discussed adaptation procedure may tend to overrefine the computational mesh, consequently increasing the number of mesh elements and the number of Degrees of Freedom of the problem. Several strategies have been proposed in the literature to limit the growth of the cardinality of the mesh in the Adaptive Finite Element Method. A common idea relies on performing local modifications by refining solely the elements T 's in which the contribution ι_T to the global error is more significant. This results in the well-known Dörfler marking strategy [124], in which we seek the set $\mathcal{M}_h \subset \mathcal{T}_h$ of minimal size such that

$$\sum_{T \in \mathcal{M}_h} \iota_T^2 \geq \zeta \sum_{T \in \mathcal{T}_h} \iota_T^2 \quad , \quad \zeta \in (0, 1).$$

Quasi-optimal and optimal convergence rates for the Adaptive Finite Element Method using Dörfler marking have been proved by several authors, including Morin *et al.* [202], Binev *et al.* [68], Stevenson [262, 263] and Cascon *et al.* [102].

Henceforth, we consider only the metric-based adaptation procedure discussed at the beginning of this subsection, whereas the Dörfler marking strategy will be investigated in future works.

Part II

Shape identification: an inverse problem in Electrical Impedance Tomography

Chapter 3

Minimization of the Kohn-Vogelius functional

We describe the inverse problem of Electrical Impedance Tomography in the classical Point Electrode Model. We introduce its formulation as a shape optimization problem and we present the boundary and the volumetric expressions of the shape gradient for the Kohn-Vogelius functional.

3.1 The problem of Electrical Impedance Tomography

In this part, we present the application of the Certified Descent Algorithm to a non-destructive testing problem, namely the inverse identification problem of Electrical Impedance Tomography (EIT). The choice of EIT as test case has to be interpreted as a proof of concept to preliminarily assess the validity of the discussed method on a non-trivial scalar problem before studying the vectorial case.

It is well-known in the literature that the problem of Electrical Impedance Tomography is severely ill-posed. Moreover, classical shape optimization methods proved to be highly unsatisfactory by remaining trapped into local minima and consequently providing fairly poor reconstructions of the inclusions. The Certified Descent Algorithm does not aim to solve the known issues of gradient-based strategies when dealing with ill-posed problems. Nevertheless, the interest in the EIT problem is twofold. On the one hand, the EIT is a non-trivial scalar problem that may guide towards the establishment of some properties of this new shape optimization method. On the other hand, the CDA proves to be a viable example to confirm the aforementioned limitations of gradient-based methods applied to inverse problems, especially through the observation of the rapidly increasing number of Degrees of Freedom required for the certification procedure which highlights the severe ill-posedness of the EIT problem. The reader interested in an overview of the methods investigated in the literature for the EIT problem is invited to refer to [10, 58, 127, 189] for shape optimization approaches, to [96, 163, 164] for strategies based on topological derivatives and to [111, 167, 170] for regularization techniques.

We are now ready to introduce the formulation of the Electrical Impedance Tomography problem. Let us consider an open domain $\mathcal{D} \subset \mathbb{R}^d$. We suppose that there exists an open subdomain $\Omega \subset \subset \mathcal{D}$ such that some given physical properties of the problem under analysis are discontinuous along the interface $\partial\Omega$ between the inclusion Ω and the complementary set $\mathcal{D} \setminus \Omega$. The location and the shape of

the inclusion are to be determined, thus Ω acts as unknown parameter in the state equations and in the inversion procedure. Our aim is to identify the inclusion Ω by performing non-invasive measurements on the boundary $\partial\mathcal{D}$ of the domain \mathcal{D} . This problem is well-known in the literature and is often referred to as Calderón's problem. Several review papers on Electrical Impedance Tomography have been published in the literature over the years. We refer to [72, 94, 107] for more details on the physical problem, its mathematical formulation and its numerical approximation.

Let χ_Ω be the characteristic function of the open set Ω , we define the conductivity k_Ω as a piecewise constant function such that $k_\Omega := k_I\chi_\Omega + k_E(1 - \chi_\Omega)$, $k_I, k_E > 0$. We introduce two Boundary Value Problems on the domain \mathcal{D} , respectively with Neumann and Dirichlet boundary conditions on $\partial\mathcal{D}$:

$$\begin{cases} -k_\Omega \Delta u_{\Omega,N} + u_{\Omega,N} = 0 & \text{in } \mathcal{D} \setminus \partial\Omega \\ \llbracket u_{\Omega,N} \rrbracket = 0 & \text{on } \partial\Omega \\ \llbracket k_\Omega \nabla u_{\Omega,N} \cdot n \rrbracket = 0 & \text{on } \partial\Omega \\ k_E \nabla u_{\Omega,N} \cdot n = g & \text{on } \partial\mathcal{D} \end{cases} \quad (3.1)$$

$$\begin{cases} -k_\Omega \Delta u_{\Omega,D} + u_{\Omega,D} = 0 & \text{in } \mathcal{D} \setminus \partial\Omega \\ \llbracket u_{\Omega,D} \rrbracket = 0 & \text{on } \partial\Omega \\ \llbracket k_\Omega \nabla u_{\Omega,D} \cdot n \rrbracket = 0 & \text{on } \partial\Omega \\ u_{\Omega,D} = U_D & \text{on } \partial\mathcal{D} \end{cases} \quad (3.2)$$

where the boundary data $g \in L^2(\partial\mathcal{D})$ and $U_D \in H^{\frac{1}{2}}(\partial\mathcal{D})$ arise from the performed physical measurements. As previously stated, 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.

Remark 3.1. It is well-known in the literature that the most accurate mathematical model for EIT is the so-called Complete Electrode Model (CEM). The CEM accounts for both the shunting effect experienced by the electrodes and the contact impedance between the object under analysis and the electrodes themselves. In [108, 257], the authors prove that the CEM is able to predict real-life measurements up to machine precision. Nevertheless, in many applications the size of the electrodes is negligible with respect to the size of the object under analysis and to the uncertainty due to the their unknown exact position on $\partial\mathcal{D}$. Within this framework, the Point Electrode Model describes small electrodes as delta distributions introducing a continuum model of the Electrical Impedance Tomography problem (cf. [72]). The connection between the two aforementioned models have been studied in [154, 169] and references therein. In particular, in these works the authors prove that the Point Electrode Model may be derived as the asymptotic limit of the Complete Electrode Model when the size of the electrodes tends to zero. Owing to the fact of considering the EIT problem as a proof of concept to preliminarily assess the validity of the Certified Descent Algorithm, we restrict our study to the classical Point Electrode Model though it is known to be less accurate than the Complete Electrode Model.

3.2 A shape optimization approach

Let us consider the Kohn-Vogelius functional first introduced in [280] and later investigated by Kohn and Vogelius in [180]:

$$J(\Omega) = \frac{1}{2} \int_{\mathcal{D}} \left(k_{\Omega} |\nabla(u_{\Omega,N} - u_{\Omega,D})|^2 + |u_{\Omega,N} - u_{\Omega,D}|^2 \right) dx. \quad (3.3)$$

In (3.3), $u_{\Omega,N}$ and $u_{\Omega,D}$ respectively stand for the solutions of the state problems (3.1) and (3.2). The inverse identification problem of Electrical Impedance Tomography may be written as the PDE-constrained optimization problem (1.5) in which we seek the open subset Ω that minimizes (3.3) under the constraint of $u_{\Omega,N}$ and $u_{\Omega,D}$ respectively being the solutions of equations (3.1) and (3.2).

3.2.1 Shape gradient of the Kohn-Vogelius functional

As stated in section 1.5, in order to differentiate the Kohn-Vogelius functional with respect to the shape, we introduce an adjoint problem for each state variable. First, we need to determine the adjoint solutions $p_{\Omega,N}$ and $p_{\Omega,D}$ associated with the states $u_{\Omega,N}$ and $u_{\Omega,D}$. Owing to the fact that the Kohn-Vogelius problem is self-adjoint, we get that $p_{\Omega,N} = u_{\Omega,N} - u_{\Omega,D}$ and $p_{\Omega,D} = 0$.

In order to rigorously compute the expression of the shape gradient of the functional (3.3), we refer to the procedure described in [219]. In the aforementioned article, the author provides the details and the technical difficulties associated with the differentiation of the functional under analysis and describes its application to the identification of discontinuities of the conductivity parameter. In the rest of this subsection, we recall the results from [219]. In particular, let $\theta \in W^{1,\infty}(\mathcal{D}; \mathbb{R}^d)$ be an admissible deformation of the domain such that $\theta = 0$ on $\partial\mathcal{D}$. The most common approach in the literature to compute the shape gradient leads to the surface expression

$$\langle dJ(\Omega), \theta \rangle = \frac{1}{2} \int_{\partial\Omega} \left(\llbracket k_{\Omega} \rrbracket \left(\left| \frac{\partial u_{\Omega,N}}{\partial \tau} \right|^2 - \left| \frac{\partial u_{\Omega,D}}{\partial \tau} \right|^2 \right) - \llbracket k_{\Omega}^{-1} \rrbracket \left(\left| k_{\Omega} \frac{\partial u_{\Omega,N}}{\partial n} \right|^2 - \left| k_{\Omega} \frac{\partial u_{\Omega,D}}{\partial n} \right|^2 \right) \right) (\theta \cdot n) ds \quad (3.4)$$

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

$$\langle G(\Omega, u), \theta \rangle = \frac{1}{2} \int_{\mathcal{D}} \left(k_{\Omega} M(\theta) \nabla u \cdot \nabla u - \nabla \cdot \theta u^2 \right) dx \quad (3.5)$$

where $M(\theta) = \nabla \theta + \nabla \theta^T - (\nabla \cdot \theta) \text{Id}$. The volumetric expression of the shape gradient of (3.3) has the following form:

$$\langle dJ(\Omega), \theta \rangle = \langle G(\Omega, u_{\Omega,N}) - G(\Omega, u_{\Omega,D}), \theta \rangle. \quad (3.6)$$

Remark 3.2. As highlighted in subsection 1.5.2, in this work we foster the volumetric formulation of the shape gradient over its surface representation, the former expression providing a better accuracy from a numerical point of view. We refer to [166] for a detailed comparison of the volumetric and surface expression of the shape gradient for elliptic problems. In particular, in this work the authors prove that within the framework of Finite Element discretizations a better accuracy and a higher convergence rate for the shape gradient are achieved when using the volumetric formulation.

Chapter 4

Certified Descent Algorithm based on the complementary energy principle

We introduce an approximation of the problem of Electrical Impedance Tomography by means of a conforming Finite Element discretization. A strategy to construct a fully computable, constant-free a posteriori estimator of the error in the shape gradient is discussed. Starting from the complementary energy principle, we define additional variational problems to compute an accurate approximation of the error in the numerical fluxes and we derive an estimator for our Quantity of Interest. After validating the error estimator, some numerical results of the application of the CDA to a 2D example of EIT are presented.

4.1 Conforming Finite Element approximation

In this section, we introduce a discretization of the EIT problem based on conforming Finite Element functions. Let $\{\mathcal{T}_h\}_{h>0}$ be a family of triangulations of the domain \mathcal{D} with no hanging nodes. Having in mind that $d = 2$, each element $T \in \mathcal{T}_h$ is a triangle and for each couple $T, T' \in \mathcal{T}_h$ such that $T \neq T'$, the intersection of the two elements is either an empty set or a vertex or an edge. An edge e is said to be an interior edge of the triangulation \mathcal{T}_h if there exist two elements $T^-(e), T^+(e) \in \mathcal{T}_h$ such that $e = T^-(e) \cap T^+(e)$, whereas is a boundary edge if there exists $T(e) \in \mathcal{T}_h$ such that $e = T(e) \cap \partial\mathcal{D}$. In the former case, the unit normal vector to e is denoted by n_e and goes from $T^-(e)$ towards $T^+(e)$. In the latter one, n is the classical outward normal to $\partial\mathcal{D}$. The set of the internal edges is noted as $\mathcal{E}_h^{\mathcal{I}}$, the boundary edges are collected into $\mathcal{E}_h^{\mathcal{B}}$ and we set $\mathcal{E}_h := \mathcal{E}_h^{\mathcal{I}} \cup \mathcal{E}_h^{\mathcal{B}}$. The state and adjoint problems are solved using the following Lagrangian Finite Element space

$$V_\Omega^{h,\kappa} := \{u^h \in \mathcal{C}^0(\overline{\mathcal{D}}) : u^h|_T \in \mathbb{P}^\kappa(T) \ \forall T \in \mathcal{T}_h\}$$

where $\mathbb{P}^\kappa(T)$ is the set of polynomials of degree less than or equal to κ on an element T .

4.1.1 The state problems

Let $a_\Omega(\cdot, \cdot)$ be the bilinear form associated with both the problems and $F_{\Omega,i}(\cdot)$, $i = N, D$ the linear forms respectively for the Neumann and the Dirichlet problem:

$$a_\Omega(u, \delta u) = \int_{\mathcal{D}} \left(k_\Omega \nabla u \cdot \nabla \delta u + u \delta u \right) dx, \quad (4.1)$$

$$F_{\Omega,N}(\delta u) = \int_{\partial \mathcal{D}} g \delta u \, ds \quad \text{and} \quad F_{\Omega,D}(\delta u) = 0. \quad (4.2)$$

We consider $u_{\Omega,N}, u_{\Omega,D} \in H^1(\mathcal{D})$ such that $u_{\Omega,D} = U_D$ on $\partial \mathcal{D}$, solutions of the following Neumann and Dirichlet variational problems $\forall \delta u_N \in H^1(\mathcal{D})$ and $\forall \delta u_D \in H_0^1(\mathcal{D})$:

$$a_\Omega(u_{\Omega,i}, \delta u_i) = F_{\Omega,i}(\delta u_i) \quad , \quad i = N, D. \quad (4.3)$$

For the Dirichlet problem, the non-homogeneous boundary datum is taken care of by means of a classical substitution technique. We remark that within the framework of conforming Finite Element discretizations, the continuous and discrete bilinear (respectively linear) forms have the same expressions. Hence, the corresponding discrete formulations of the state problems (4.3) may be derived by replacing the analytical solutions $u_{\Omega,N}$ and $u_{\Omega,D}$ with their approximations $u_{\Omega,N}^h$ and $u_{\Omega,D}^h$ which belong to the space $V_\Omega^{h,\ell}$ of Lagrangian Finite Element functions of degree ℓ . In a similar fashion, θ^h is the solution of equation (2.5) computed using a Lagrangian Finite Element space and substituting the expression of the discretized shape gradient (2.6) to its analytical counterpart (1.17). The degree chosen for the Finite Element basis functions will be discussed in section 4.3.

4.1.2 The adjoint problems

Let $r_{\Omega,N}$ and $r_{\Omega,D}$ be the solutions of the adjoint problems (2.21) introduced to evaluate the contributions of the Neumann and Dirichlet state problems to the error in the Quantity of Interest: we seek $r_{\Omega,N} \in H^1(\mathcal{D})$ and $r_{\Omega,D} \in H_0^1(\mathcal{D})$ such that respectively $\forall \delta r_N \in H^1(\mathcal{D})$ and $\forall \delta r_D \in H_0^1(\mathcal{D})$

$$a_\Omega(\delta r_i, r_{\Omega,i}) = H_{\Omega,i}(\delta r_i) \quad , \quad i = N, D \quad (4.4)$$

where the linear forms $H_{\Omega,i}(\delta r)$, $i = N, D$ read as

$$H_{\Omega,i}(\delta r) := \frac{\partial G}{\partial u}(\Omega, u_{\Omega,i}^h)[\theta^h, \delta r] = \int_{\mathcal{D}} \left(k_\Omega M(\theta^h) \nabla u_{\Omega,i}^h \cdot \nabla \delta r - \nabla \cdot \theta^h u_{\Omega,i}^h \delta r \right) dx. \quad (4.5)$$

As per the state problems, the discretized solutions $r_{\Omega,N}^h$ and $r_{\Omega,D}^h$ are obtained solving the adjoint equations (4.4) within an appropriate space of Lagrangian Finite Element functions, that is the space $V_\Omega^{h,m}$ of degree m .

4.2 Estimate of the error in the shape gradient via the complementary energy principle

We consider the framework introduced in section 2.3. Since the Kohn-Vogelius problem is self-adjoint, (2.22) reduces to

$$\bar{E} = \|u_{\Omega,N} - u_{\Omega,N}^h\|_\Omega \|r_{\Omega,N} - r_{\Omega,N}^h\|_\Omega + \|u_{\Omega,D} - u_{\Omega,D}^h\|_\Omega \|r_{\Omega,D} - r_{\Omega,D}^h\|_\Omega. \quad (4.6)$$

In order to obtain a computable upper bound for the error in the shape gradient, we seek an estimate of the energy-norm of the error for the state and adjoint solutions in (4.6).

Let $e_{\Omega,i} = u_{\Omega,i} - u_{\Omega,i}^h$ and $\epsilon_{\Omega,i} = r_{\Omega,i} - r_{\Omega,i}^h$ for $i = N, D$. In the following subsections, we derive the estimates of the energy-norm of the $e_{\Omega,i}$'s and the $\epsilon_{\Omega,i}$'s using a strategy inspired by the so-called complementary energy principle (cf. [235]). In practice, we introduce a dual flux variable for every problem and each bound is computed by solving an additional adjoint problem thus leading to a better approximation of the numerical fluxes $\nabla e_{\Omega,i}$'s and $\nabla \epsilon_{\Omega,i}$'s. For additional information on this approach, we refer to [272].

4.2.1 Energy-norm error estimates for the state equations

For the sake of readability, let us rename $H^1(\mathcal{D})$ as V_N and $H_0^1(\mathcal{D})$ as V_D . We recall the previously mentioned residual equations such that $\forall \delta u_N \in V_N$ and $\forall \delta u_D \in V_D$

$$a_{\Omega}(e_{\Omega,i}, \delta u_i) = F_{\Omega,i}(\delta u_i) - a_{\Omega}(u_{\Omega,i}^h, \delta u_i) \quad , \quad i = N, D. \quad (4.7)$$

We recall that solving equation (4.7) is equivalent to the following minimization problem, that is we seek $w \in V_i$ such that

$$E_{\Omega,i}(e_{\Omega,i}) = \min_{w \in V_i} E_{\Omega,i}(w) \quad , \quad i = N, D \quad (4.8)$$

where the global energy functional associated with the Neumann and Dirichlet problems reads

$$E_{\Omega,i}(w) = \frac{1}{2} \int_{\mathcal{D}} \left(k_{\Omega} |\nabla w|^2 + |w|^2 \right) dx + \int_{\mathcal{D}} \left(k_{\Omega} \nabla u_{\Omega,i}^h \cdot \nabla w + u_{\Omega,i}^h w \right) dx - F_{\Omega,i}(w). \quad (4.9)$$

We consider the space $H(\text{div}) = \{ \tau \in L^2(\mathcal{D}; \mathbb{R}^d) : \nabla \cdot \tau \in L^2(\mathcal{D}) \}$. By introducing an additional variable $z = \nabla w$ and a dual variable $\sigma \in H(\text{div})$, we may construct the Lagrangian functional $L_{\Omega,i} : V_i \times L^2(\mathcal{D}) \times H(\text{div}) \rightarrow \mathbb{R}$ which has the following form

$$\begin{aligned} L_{\Omega,i}(w, z, \sigma) = & \frac{1}{2} \int_{\mathcal{D}} \left(k_{\Omega} |z|^2 + |w|^2 \right) dx + \int_{\mathcal{D}} \left(k_{\Omega} \nabla u_{\Omega,i}^h \cdot z + u_{\Omega,i}^h w \right) dx - F_{\Omega,i}(w) \\ & + \int_{\mathcal{D}} \sigma \cdot (\nabla w - z) dx. \end{aligned} \quad (4.10)$$

Thus the minimization problem (4.8) may be rewritten as a min-max problem and owing to the Lagrange duality, we get

$$\begin{aligned} \min_{w \in V_i} E_{\Omega,i}(w) &= \min_{\substack{w \in V_i \\ z = \nabla w}} \max_{\sigma \in H(\text{div})} L_{\Omega,i}(w, z, \sigma) \\ &= \max_{\sigma \in H(\text{div})} \min_{\substack{w \in V_i \\ z = \nabla w}} L_{\Omega,i}(w, z, \sigma). \end{aligned} \quad (4.11)$$

From the system of first-order optimality conditions for $L_{\Omega,i}(w, z, \sigma)$ we derive the following relationships among the variables:

$$\begin{cases} z = k_{\Omega}^{-1} \sigma - \nabla u_{\Omega,i}^h & \text{in } \mathcal{D}, \\ w = \nabla \cdot \sigma - u_{\Omega,i}^h & \text{in } \mathcal{D} \end{cases} \quad (4.12)$$

and for the Neumann case, we get the additional boundary condition:

$$\sigma \cdot n = g \text{ on } \partial\mathcal{D}. \quad (4.13)$$

Hence, by plugging (4.12) into (4.11), we get the following maximization problems for $i = N, D$

$$\max_{\substack{\sigma \in H(\text{div}) \\ (\sigma \cdot n = g \text{ on } \partial\mathcal{D}, i=N)}} C_\Omega(\sigma) \quad , \quad C_\Omega(\sigma) := -\frac{1}{2} \int_{\mathcal{D}} \left(k_\Omega^{-1} |\sigma - k_\Omega \nabla u_{\Omega,i}^h|^2 + |\nabla \cdot \sigma - u_{\Omega,i}^h|^2 \right) dx. \quad (4.14)$$

In (4.14), the objective functional $C_\Omega(\sigma)$ - known as the dual complementary energy associated with the problems - is maximized over the set of all vectorfields in $\sigma \in H(\text{div})$ for the Dirichlet case and for the vectorfields $\sigma \in H(\text{div})$ such that $\sigma \cdot n = g$ on $\partial\mathcal{D}$ for the Neumann one.

In order to compute the dual flux variables, we derive the first-order optimality conditions for the dual complementary energy functional in (4.14). Thus, we seek $\sigma_{\Omega,N}, \sigma_{\Omega,D} \in H(\text{div})$ such that $\sigma_{\Omega,N} \cdot n = g$ on $\partial\mathcal{D}$ which satisfy $\forall \delta\sigma_N, \delta\sigma_D \in H(\text{div})$ such that $\delta\sigma_N \cdot n = 0$ on $\partial\mathcal{D}$

$$\int_{\mathcal{D}} \left(k_\Omega^{-1} \sigma_{\Omega,i} \cdot \delta\sigma_i + (\nabla \cdot \sigma_{\Omega,i})(\nabla \cdot \delta\sigma_i) \right) dx = \begin{cases} 0 & , i = N. \\ \int_{\partial\mathcal{D}} U_D(\delta\sigma_i \cdot n) ds & , i = D. \end{cases} \quad (4.15)$$

Let $\sigma_{\Omega,N}^h$ and $\sigma_{\Omega,D}^h$ be the dual fluxes discretized using Raviart-Thomas Finite Element functions. By combining the definition of energy-norm induced by the bilinear form (4.1) with the information in (4.12) and (4.15), we get the following upper bound for the energy-norm of the error in the state equations:

$$\|u_{\Omega,i} - u_{\Omega,i}^h\|_\Omega^2 \leq \int_{\mathcal{D}} \left(k_\Omega^{-1} |\sigma_{\Omega,i}^h - k_\Omega \nabla u_{\Omega,i}^h|^2 + |\nabla \cdot \sigma_{\Omega,i}^h - u_{\Omega,i}^h|^2 \right) dx. \quad (4.16)$$

4.2.2 Energy-norm error estimates for the adjoint equations

As in the previous subsection, we present the formulation of the dual complementary energy associated with the discretization error of the adjoint problems (4.4). In a similar fashion, we introduce the dual variable $\xi \in H(\text{div})$ and we retrieve the following relationships

$$\begin{cases} z = k_\Omega^{-1} \xi + M(\theta^h) \nabla u_{\Omega,i}^h - \nabla r_{\Omega,i}^h & \text{in } \mathcal{D}, \\ w = \nabla \cdot \xi - (\nabla \cdot \theta^h) u_{\Omega,i}^h - r_{\Omega,i}^h & \text{in } \mathcal{D} \end{cases} \quad (4.17)$$

coupled with a boundary condition for the Neumann problem:

$$\xi \cdot n = 0 \text{ on } \partial\mathcal{D}. \quad (4.18)$$

The maximization problem for the dual complementary energy associated with the adjoint problems for $i = N, D$ reads as

$$\max_{\substack{\xi \in H(\text{div}) \\ (\xi \cdot n = 0 \text{ on } \partial\mathcal{D}, i=N)}} C_\Omega(\xi), \quad (4.19)$$

$$C_\Omega(\xi) := -\frac{1}{2} \int_{\mathcal{D}} \left(k_\Omega^{-1} |\xi + k_\Omega M(\theta^h) \nabla u_{\Omega,i}^h - k_\Omega \nabla r_{\Omega,i}^h|^2 + |\nabla \cdot \xi - (\nabla \cdot \theta^h) u_{\Omega,i}^h - r_{\Omega,i}^h|^2 \right) dx. \quad (4.20)$$

In order to compute the dual flux variables, we seek $\xi_{\Omega,N}, \xi_{\Omega,D} \in H(\text{div})$ such that $\xi_{\Omega,N} \cdot n = 0$ on $\partial\mathcal{D}$ satisfying $\forall \delta\xi_N, \delta\xi_D \in H(\text{div})$ such that $\delta\xi_N \cdot n = 0$ on $\partial\mathcal{D}$

$$\int_{\mathcal{D}} \left(k_{\Omega}^{-1} \xi_{\Omega,i} \cdot \delta\xi_i + (\nabla \cdot \xi_{\Omega,i})(\nabla \cdot \delta\xi_i) \right) dx = \int_{\mathcal{D}} \left((\nabla \cdot \theta^h) u_{\Omega,i}^h \nabla \cdot \delta\xi_i - M(\theta^h) \nabla u_{\Omega,i}^h \cdot \delta\xi_i \right) dx \quad (4.21)$$

and via their Raviart-Thomas Finite Element approximations $\xi_{\Omega,i}^h$'s, we derive the upper bound of the adjoint errors in the energy-norm

$$\|r_{\Omega,i} - r_{\Omega,i}^h\|_{\Omega}^2 \leq \int_{\mathcal{D}} \left(k_{\Omega}^{-1} |\xi_{\Omega,i}^h + k_{\Omega} M(\theta^h) \nabla u_{\Omega,i}^h - k_{\Omega} \nabla r_{\Omega,i}^h|^2 + |\nabla \cdot \xi_{\Omega,i}^h - (\nabla \cdot \theta^h) u_i^h - r_{\Omega,i}^h|^2 \right) dx. \quad (4.22)$$

Eventually, by plugging (4.16) and (4.22) into (4.6), we are able to explicitly compute the upper bound \bar{E} of the error in the shape gradient.

4.3 Numerical results

We present some numerical results of the application of the Certified Descent Algorithm (CDA) to the problem of Electrical Impedance Tomography. We remark that the simulations presented in this paper are based on a mesh moving approach for the deformation of the domain (Algorithm 2.3 - step 7). Thus, the procedure does not allow for topological changes and the correct number of inclusions within the material has to be set at the beginning of the algorithm. In order to account for the nucleation of new inclusions or the merging of existing ones, a mixed approach based on topological and shape gradients may be followed as suggested in [163]. For the rest of this section, we will consider several examples where the shape and the location of the inclusions evolve under the assumption that the number of subregions inside \mathcal{D} is *a priori* set and known.

Before running the shape optimization algorithm, we identify a set of consistent boundary conditions (g, U_D) for the state problems (4.3). First, we set a Neumann boundary condition g on $\partial\mathcal{D}$ for the flux; in order to impose the Dirichlet condition on the potential, we iteratively solve the Neumann state problem and subsequently refine the mesh until a very fine error estimate in the energy-norm is achieved. The trace of the resulting solution $u_{\Omega,N}$ on $\partial\mathcal{D}$ is eventually picked as boundary datum U_D for the Dirichlet state problem.

All the numerical results are obtained using FreeFem++ [160].

4.3.1 Numerical assessment of the goal-oriented estimator

We consider the configuration in figure 4.1, where $\mathcal{D} := \{(x, y) \mid x^2 + y^2 \leq \rho_E^2\}$ and $\Omega := \{(x, y) \mid x^2 + y^2 \leq \rho_I^2\}$. The values for the physical parameters are $\rho_E = 5$, $\rho_I = 4$, $k_E = 1$ and $k_I = 10$. The boundary datum for the Neumann problem (3.1) reads as $g = \cos(M\vartheta)$, $M = 5$. Using a polar coordinate system (ρ, ϑ) , we can compute the following analytical solution:

$$u_{\Omega,N} = \begin{cases} C_0 J_M \left(-i\rho k_I^{-\frac{1}{2}} \right) \cos(M\vartheta) & , \rho \in [0, \rho_I] \\ \left[C_1 J_M \left(-i\rho k_E^{-\frac{1}{2}} \right) + C_2 Y_M \left(-i\rho k_E^{-\frac{1}{2}} \right) \right] \cos(M\vartheta) & , \rho \in (\rho_I, \rho_E] \end{cases}$$

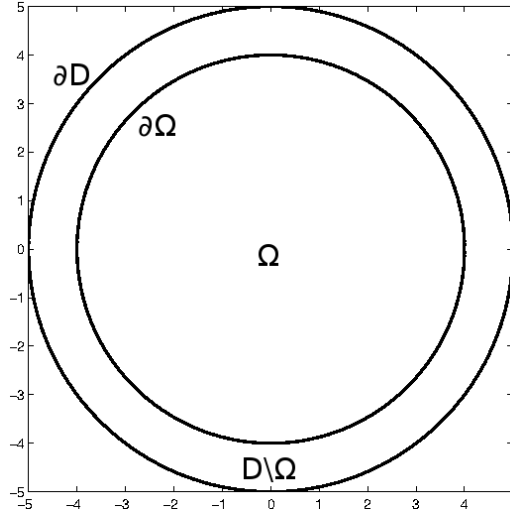


Figure 4.1 – Test case for the Electrical Impedance Tomography. Circular inclusion Ω inside the circular body \mathcal{D} .

Constant	$\Re[C_i]$	$\Im[C_i]$
C_0	$-6.3 \cdot 10^{-9}$	$+40.395$
C_1	$+1.3015$	$+0.3255$
C_2	$+1.5 \cdot 10^{-11}$	-1.3015

Table 4.1 – Constants for the analytical solution.

where $J_M(\cdot)$ and $Y_M(\cdot)$ respectively represent the first- and second-kind Bessel functions of order M . The constants C_0, \dots, C_2 are detailed in table 4.1.

For the approximation of the state equations (4.3), we consider both \mathbb{P}^1 and \mathbb{P}^2 Lagrangian Finite Element functions. In figures 4.2a and 4.2b, we present a comparison between the analytical error due to the discretization and the corresponding estimates arising from the complementary energy principle (cf. subsection 4.2.1) under uniform mesh refinements. We remark that using \mathbb{P}^1 Finite Elements, the estimated convergence rate is nearly 1, whereas using \mathbb{P}^2 basis functions for the Finite Element space leads to a convergence rate slightly lower than 2.

In order to construct the estimator for the error in the shape gradient, first we approximate equation (2.5) using $\mathbb{P}^1 \times \mathbb{P}^1$ Lagrangian Finite Element functions. For the discretization of the adjoint equations (4.4), we use the same Finite Element space as for the state problems, whereas the dual fluxes in equations (4.15) and (4.21) are approximated using Raviart-Thomas Finite Element functions. In particular, we choose the space of RT_0 (respectively RT_1) functions when the state and adjoint equations are solved using \mathbb{P}^1 (respectively \mathbb{P}^2) Finite Elements.

In figure 4.3, we present the convergence history of the discretization error in the shape gradient, the error in the Quantity of Interest arising from its linearization and the corresponding complementary energy estimates provided in (4.6) using both \mathbb{P}^1 (Fig. 4.3a) and \mathbb{P}^2 (Fig. 4.3b) Finite Element functions. In both figure 4.3a and figure 4.3b, we remark that the error in the linearized Quantity of Interest is very similar to the one in the shape gradient. This confirms that the linearization error

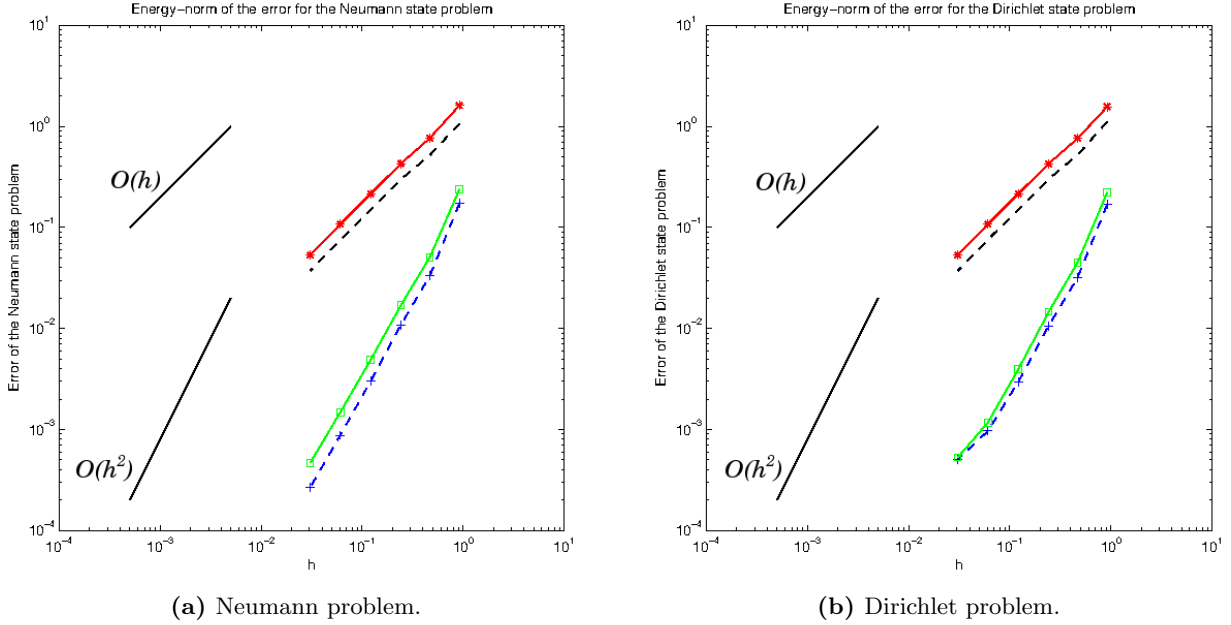


Figure 4.2 – Comparison of the convergence rates with respect to the mesh size h for the (a) Neumann and (b) Dirichlet state equations. Analytical errors using \mathbb{P}^1 (black dash) and \mathbb{P}^2 (blue cross) Lagrangian Finite Element functions; error estimates based on the complementary energy principle using \mathbb{P}^1 (red star) and \mathbb{P}^2 (green square) Lagrangian Finite Element functions.

introduced in (2.20) is negligible with respect to the discretization error due to the Finite Element approximation and the estimator constructed from the linearized Quantity of Interest provides reliable information on the error in the shape gradient itself.

Figure 4.3a shows the evolution of the error in the Quantity of Interest with respect to the number of Degrees of Freedom of the problem under uniform mesh refinements. The evolution of the error estimator is analogous to the one of the analytical error in the shape gradient, thus we verify that an upper bound for the error in the Quantity of Interest is derived. Nevertheless, when dealing with \mathbb{P}^2 Lagrangian Finite Element functions (Fig. 4.3b), the resulting error estimator for the Quantity of Interest underestimates the error in the shape gradient. This phenomenon may be caused by the error due to the approximation of the geometry that has not been accounted for in this thesis. As a matter of fact, in [201] Morin *et al.* observe that increasing the accuracy of the PDE approximation is useless if the expected geometrical error is higher than the one due to the discretization of the state problem. For this reason, in the following simulations, we stick to low-order Finite Element approximations ($\mathbb{P}^1 - RT_0$) since using higher-order elements would prevent from getting a certified upper bound of the error in the shape gradient which is crucial for the application of the Certified Descent Algorithm.

4.3.2 1-mesh and 2-mesh reconstruction strategies

We may now apply the CDA to identify the unknown inclusion Ω inside the circular domain \mathcal{D} using one boundary measurement. The initial inclusion is a circle of radius $\rho_{ini} = 2$ and the corresponding

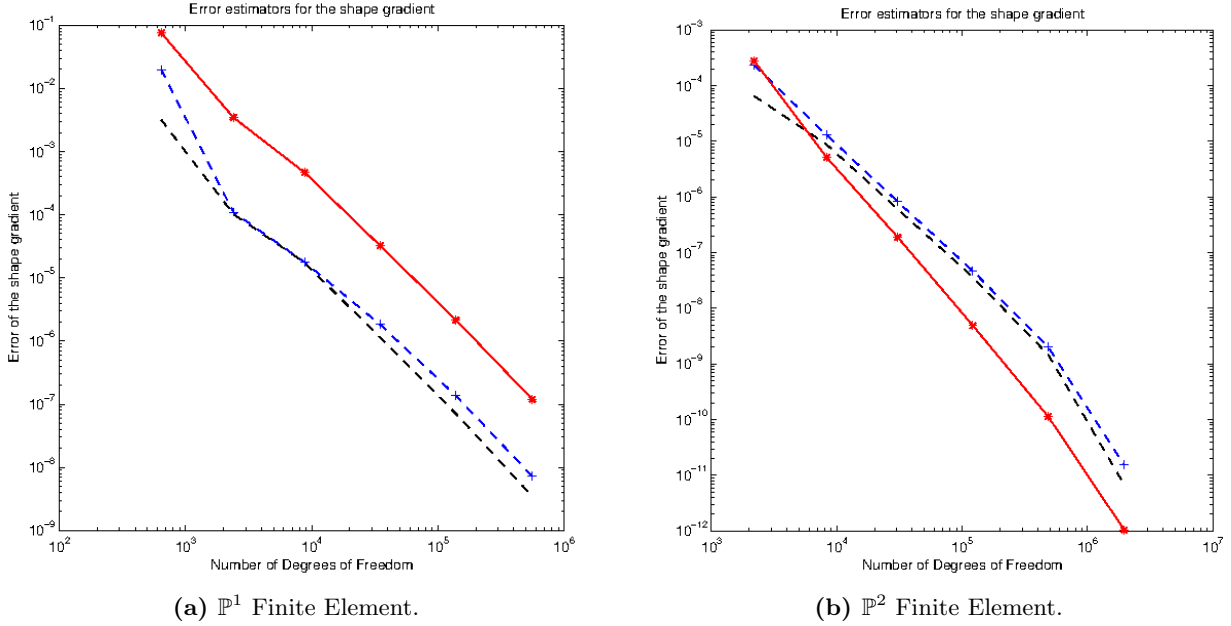


Figure 4.3 – Comparison of the convergence rates with respect to the number of Degrees of Freedom for the error in the Quantity of Interest using (a) \mathbb{P}^1 and (b) \mathbb{P}^2 Lagrangian Finite Element functions. Analytical error in the shape gradient (black dash); error in the linearized Quantity of Interest (blue cross); error estimator for the Quantity of Interest using the complementary energy principle (red star).

computational mesh counting 472 triangles is displayed in figure 4.4a.

It is well-known in the literature (cf. [28]) that using the same computational domain for both solving the state problem and computing a descent direction may lead to poor optimized shapes. In figures 4.4b and 4.4c we present the computational domains obtained using respectively a 1-mesh strategy to compute both the solutions $u_{\Omega,i}^h$'s and the descent direction θ^h and a 2-mesh approach in which the state equations are solved on a fine mesh whereas θ^h is computed on a coarser triangulation. A comparison of the reconstructed interfaces after 24 and 25 iterations is reported in figure 4.4d and as expected we observe that the 1-mesh algorithm provides a poor approximation of the inclusion whereas the 2-mesh strategy is able to precisely retrieve the boundary along which the conductivity k_Ω is discontinuous. Figures 4.4e and 4.4f confirm what was already observed by zooming on the local behavior of the interfaces and highlighting the oscillatory nature of the 1-mesh reconstruction. In figure 4.5a, we report the evolution of the objective functional with respect to the number of iterations using the two discussed approaches. It is straightforward to observe that the CDA identifies a genuine descent direction at each iteration, generating a minimizing sequence of shapes such that the objective functional $J(\Omega)$ is monotonically decreasing. Moreover, the error estimate in the shape gradient is also used to construct a guaranteed stopping criterion for the overall optimization strategy which automatically ends when $|\langle d_h J(\Omega), \theta^h \rangle| + \overline{E} < \text{tol}$ for an *a priori* set tolerance. Even though both versions of the algorithm generate shapes for which the functional is monotonically decreasing, only the 2-mesh strategy is able to precisely identify the target inclusion. For this reason, in the rest of this thesis we will focus solely on the 2-mesh approach.

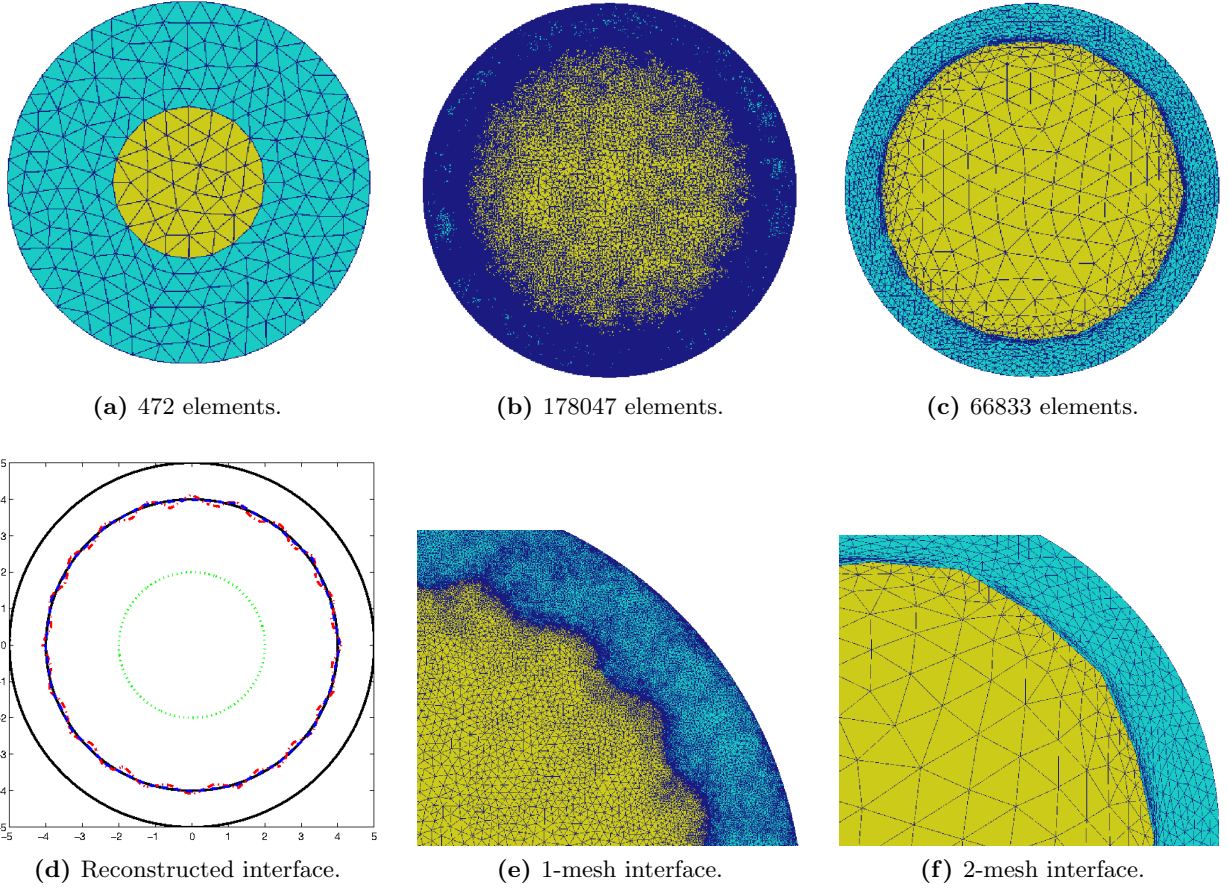


Figure 4.4 – Comparison of 1-mesh and 2-mesh reconstruction strategies. (a) Initial mesh. (b) Final mesh using 1-mesh strategy. (c) Final mesh using 2-mesh strategy. (d) Initial configuration (dotted green), target inclusion (solid black) and reconstructed interfaces using 1-mesh (dot-dashed red) and 2-mesh (dashed blue) strategies. (e) Zoom of the reconstructed interface using 1-mesh strategy. (f) Zoom of the reconstructed interface using 2-mesh strategy.

Besides the theoretical improvement of the Boundary Variation Algorithm discussed so far, an advantage of the CDA lies in the possibility of using relatively coarse meshes to identify certified descent directions. In figure 4.5b we observe that the number of Degrees of Freedom remains small until the reconstructed interface approaches the real inclusion, that is coarse meshes prove to be reliable during the initial iterations of the algorithm. Thus, another important feature of the Certified Descent Algorithm is the ability of certifying the reliability of coarse meshes for the identification of genuine descent directions for a shape functional, reducing the overall computational effort of the algorithm coupled with the *a posteriori* estimators during the initial phase of computation.

4.3.3 A more involved test case

In the previous subsection, we applied the Certified Descent Algorithm to a simple test case and we were able to retrieve a precise description of the interface $\partial\Omega$. This was mainly due to the fact that the inclusion Ω was located near the external boundary $\partial\mathcal{D}$ where the measurements are performed. In this subsection, we consider a more involved test case: on the one hand, we break the symmetry of the problem by considering the initial and target configurations in figure 4.6a; on the other hand, we highlight the difficulties of precisely identifying the boundary of an inclusion when its position is far away from $\partial\mathcal{D}$.

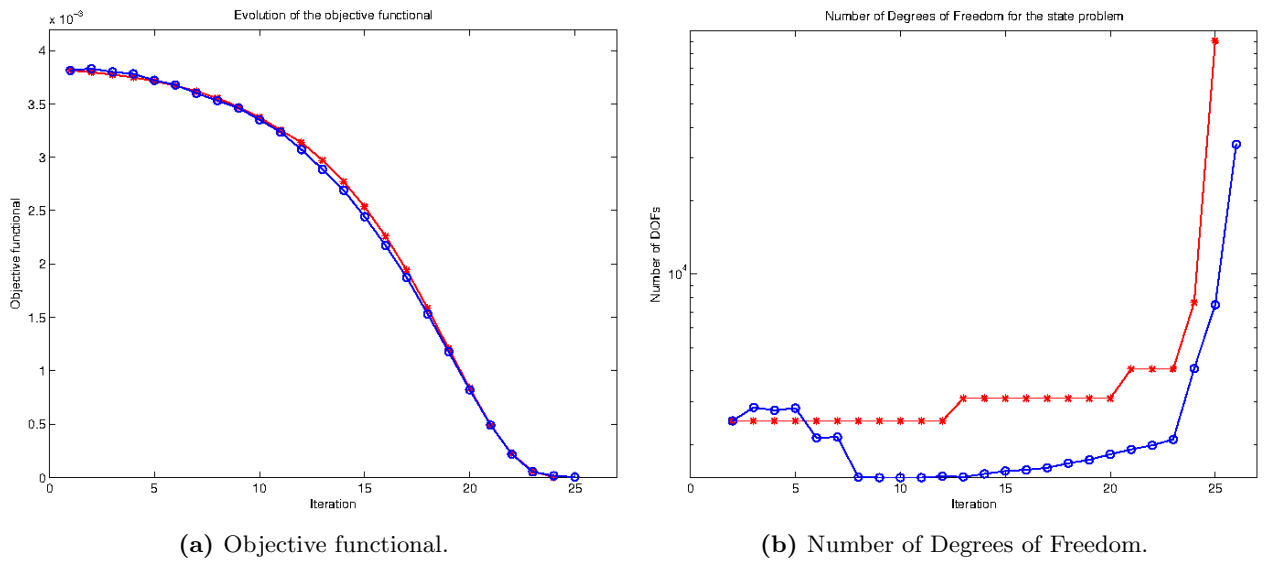


Figure 4.5 – Comparison of 1-mesh (red star) and 2-mesh (blue circle) reconstruction strategies. (a) Evolution of the objective functional. (b) Number of Degrees of Freedom.

As observed in the previous subsection, the evolution of the objective functional is monotonically decreasing (Fig. 4.6b - red curve), meaning a genuine descent direction is identified at each iteration of the optimization procedure. The final value of the approximated objective functional is $\mathcal{O}(10^{-4})$, in agreement with the zero value in the analytical optimal configuration of the inclusion. Moreover, the evolution of the number of Degrees of Freedom (Fig. 4.6c - red curve) shows that coarse meshes prove to be reliable during initial iterations. The size of the problem remains small for several successive iterations but after few tens of iterations the CDA performs multiple mesh refinements in order to identify a genuine descent direction. This results in a high number of Degrees of Freedom which rapidly increases when approaching the configuration for which the criterion $|\langle d_h J(\Omega), \theta^h \rangle| + \overline{E}$ fulfills a given tolerance $\text{tol} = 5 \cdot 10^{-8}$.

In figure 4.6a, we observe that the part of the interface closest to $\partial\mathcal{D}$ is well identified using the Certified Descent Algorithm with one boundary measurement. Nevertheless when moving away from the external boundary, the precision of the reconstructed interface decreases and the algorithm is not able to precisely retrieve the whole shape of the inclusion. The uncertainty of the reconstruction in the central region of \mathcal{D} is mainly due to the ill-posedness of the inverse problem. As a matter of fact, state

problems (3.1) and (3.2) are elliptic equations thus the effect of the boundary conditions becomes less and less important moving away from $\partial\mathcal{D}$.

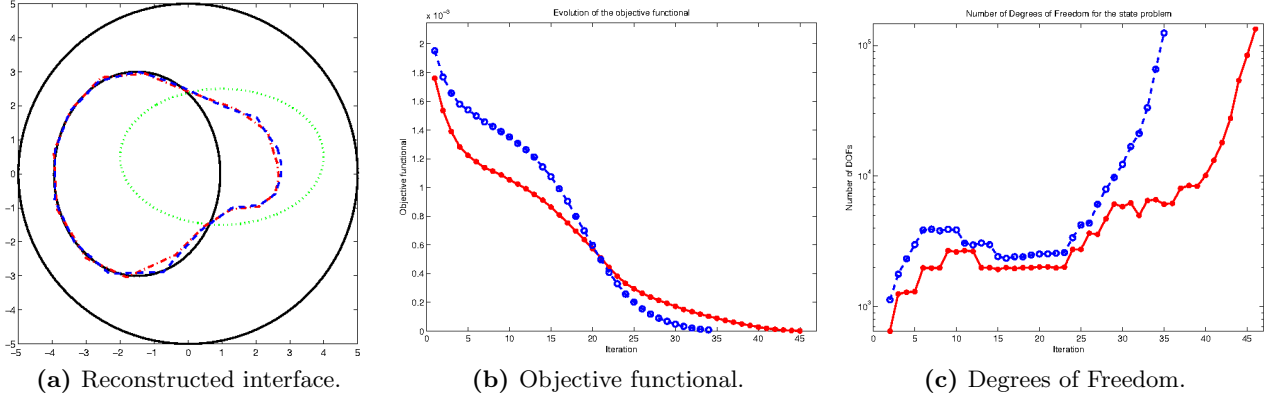


Figure 4.6 – Certified Descent Algorithm using one measurement (red stars) and ten measurements (blue circles). (a) Initial configuration (dotted green), target inclusion (solid black) and reconstructed interface (dot-dashed red and dashed blue). (b) Evolution of the objective functional. (c) Number of Degrees of Freedom.

The case of multiple boundary measurements

A strategy to improve the quality of the reconstruction via the CDA relies on the use of several boundary measurements to retrieve a better approximation of the inclusion in a smaller number of iterations. The procedure to construct the set of boundary conditions to be used by the CDA is detailed in next subsection. Here, we present the outcome of the Certified Descent Algorithm using ten measurements for the test case featuring the circular domain $\mathcal{D} := \{(x, y) \mid x^2 + y^2 \leq 25\}$ with the inclusion represented by the solid black line in figure 4.6a. We observe that using several boundary measurements the algorithm has access to more information to better identify the interface of the inclusion. First of all, the blue curve in figure 4.6a confirms the ability of the CDA to exactly identify the inclusion near the boundary $\partial\mathcal{D}$ and it highlights some minor improvements in the reconstruction of the interface with respect to the test case in red featuring one measurement (cf. the upper and lower parts of the ellipse). Nevertheless, the result is still degraded when moving towards the center of the domain. A possible workaround to this issue and to the low resolution of the reconstruction in the center of the computational domain was proposed by Ammari *et al.* in [32], where a hybrid imaging method arising from the coupling of electromagnetic tomography with acoustic waves is described. The observed phenomenon is due to the well-known ill-posedness of the problem. Though considering several measurements improve the overall outcome of the algorithm, the final result is far from being satisfactory. Nevertheless, these limitations are related to the nature of the problem and we cannot expect gradient-based strategies to successfully overcome this issue. These remarks are confirmed by figure 4.6c. As a matter of fact, the number of Degrees of Freedom rapidly increases in both test cases, reaching 10^5 and making the certification procedure unfeasible. Besides the improvement in the reconstructed interface, the use of several measurements is responsible for reducing the number of iterations required by the CDA to identify the inclusion (Fig. 4.6b). Moreover we remark that the tolerance that the quantity $|\langle d_h J(\Omega), \theta^h \rangle| + \bar{E}$ has to fulfill in this case drops to $\text{tol} = 10^{-6}$, that is

finer results are obtained in a smaller number of iterations and using lower precision in the case of multiple boundary measurements.

Sensitivity of the CDA to different initializations

It is well-known in the literature that gradient-based strategies are local, that is they are solely able to identify local minima. Thus, a key aspect for the success of these optimization procedures is represented by the initialization of the unknown design variable. For shape optimization problems, this reduces to the choice of the initial shape. In figure 4.7, we present three different initializations for the inclusion Ω and the corresponding final interfaces reconstructed by the Certified Descent Algorithm. The three test cases confirm the ability of the method to correctly retrieve the portion of the inclusion close to $\partial\mathcal{D}$ whereas the regions in the center of the domain suffer from a degraded reconstruction. Concerning the objective functional, the final values obtained using the proposed initializations are comparable (Fig. 4.8a), as the rapidly exploding number of Degrees of Freedom (Fig. 4.8b) which testifies again the ill-posedness of the inverse problem.

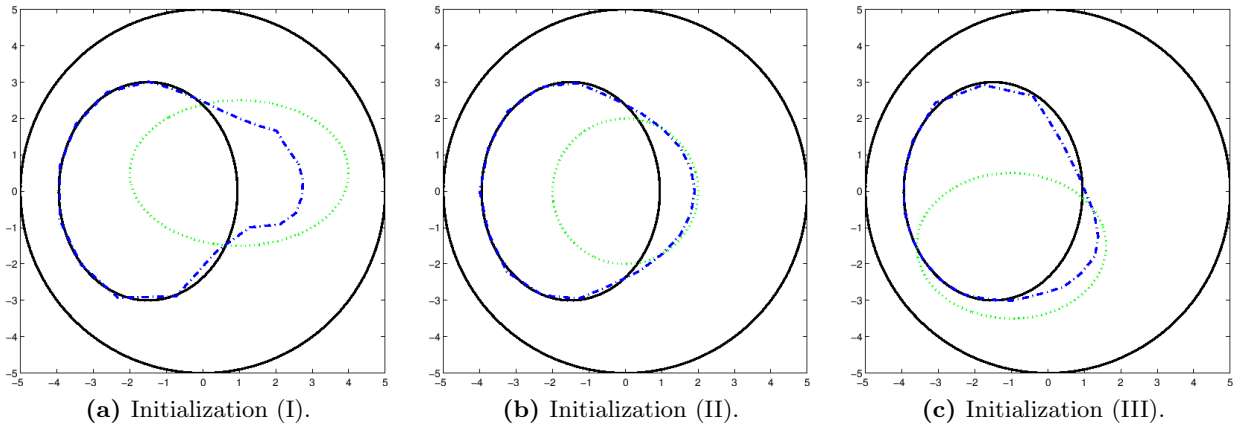


Figure 4.7 – Certified Descent Algorithm using multiple measurements for different initial configurations. Initial configuration (dotted green), target inclusion (solid black) and reconstructed interface (dashed blue).

Remark 4.1. In the literature concerning inverse problems, a key issue when discussing a new method is represented by its robustness to noise and data perturbations. It is straightforward to observe that the construction of the sets of boundary data (g, U_D) described at the beginning of this section is responsible for an additional contribution to the error. As previously observed, in order to retrieve reliable information to optimize the objective functional, the Certified Descent Algorithm requires an extremely high precision after few tens of iterations. In real-world applications, this results to be completely unfeasible since the additional information arising from the high precision would be lost due to the noisy nature of the boundary data. Hence, for the purpose of this work we neglect the contribution of the uncertainty due to the measurements. An interesting extension of the current CDA may focus on the role of the error on the boundary measurements: in particular, an additional criterion may be integrated into the method in order to stop the algorithm if the error on the solutions of the state problems is smaller than the error on the data, that is if the effect of data fluctuations

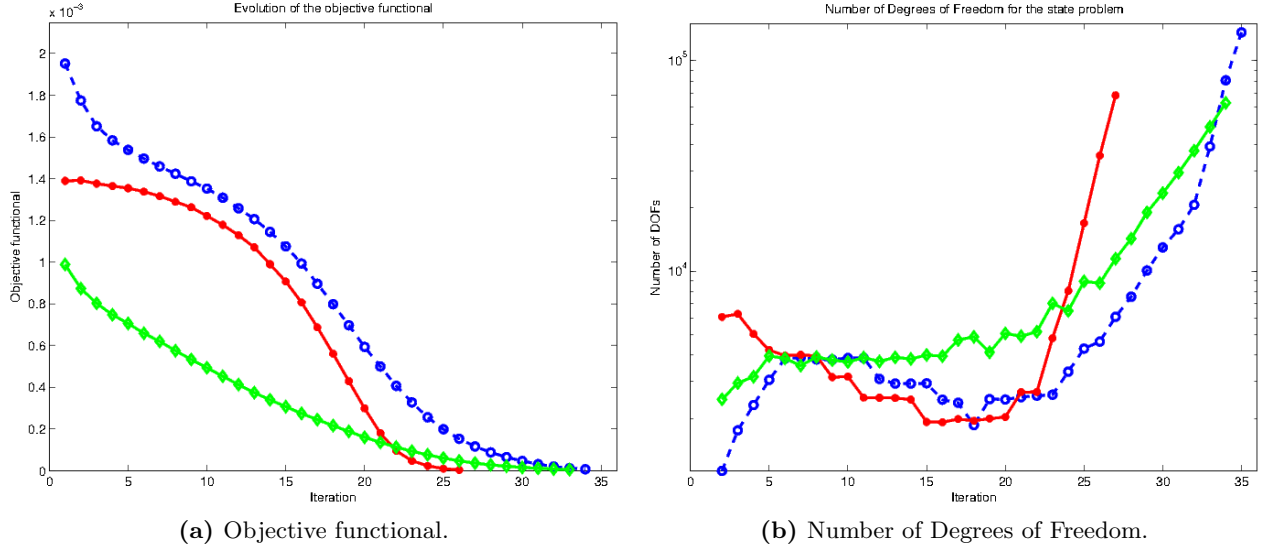


Figure 4.8 – Certified Descent Algorithm using multiple measurements for initial guess (I) blue circles; (II) red stars; (III) green diamonds. (a) Evolution of the objective functional. (b) Number of Degrees of Freedom.

becomes predominant in the certification procedure.

4.3.4 The case of multiple boundary measurements

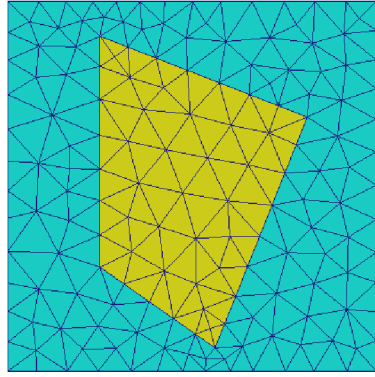
In [267], the authors proved that from an analytical point of view, one measurement is sufficient to uniquely reconstruct the inclusion within the Calderón’s problem. Other analytical results on this topic are presented in [9]. Nevertheless, from a numerical point of view, it is known that multiple measurements are required to have a correct approximation of the Electrical Impedance Tomography identification problem. In this subsection, we present several tests of the previously described algorithm using multiple boundary measurements. In particular, we consider $D = 10$ measurements such that $\forall j = 0, \dots, D - 1$

$$g_j(x, y) = (x + a_j y)^{b_j} a_j^{c_j} \quad , \quad a_j = 1 + 0.1j \quad , \quad b_j = \frac{j+1}{2} \quad , \quad c_j = j - 2 \left\lfloor \frac{j}{2} \right\rfloor$$

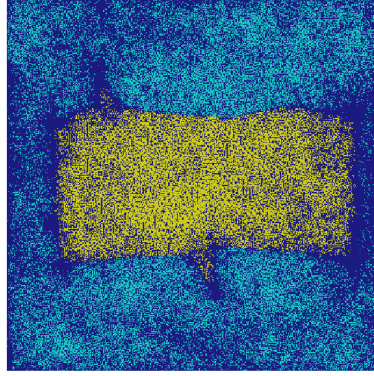
and we use them to test the following cases:

- (i) one inclusion in a square domain (Fig. 4.9a-4.9c);
- (ii) two inclusions in a circular domain (Fig. 4.9d-4.9f);
- (iii) two inclusions in a square domain (Fig. 4.9g-4.9i).

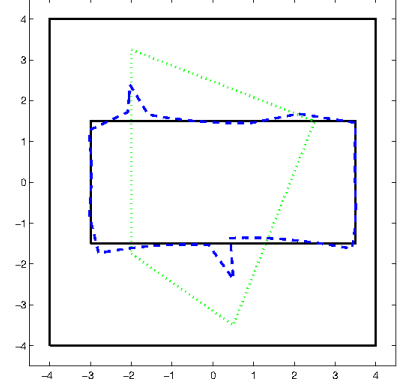
First, we present a simulation in which the body is the square $\mathcal{D} := [-4, 4]^2$ featuring a single polygonal inclusion (Fig. 4.9c). Then, we propose two cases with multiple inclusions (Fig. 4.9f and 4.9i): in both simulations, we assume that the number of inclusions is known *a priori* and equal to two and that the conductivity k_Ω has only two values, one inside the inclusions and one for the background.



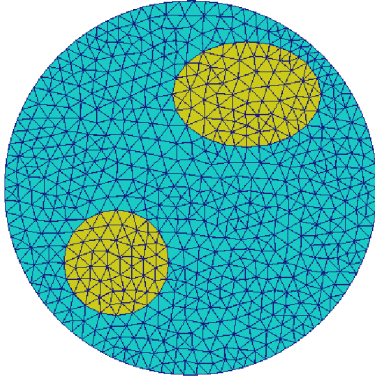
(a) 194 elements.



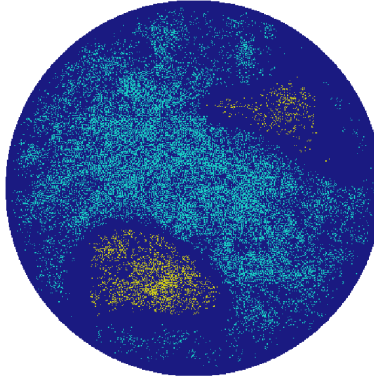
(b) 132040 elements.



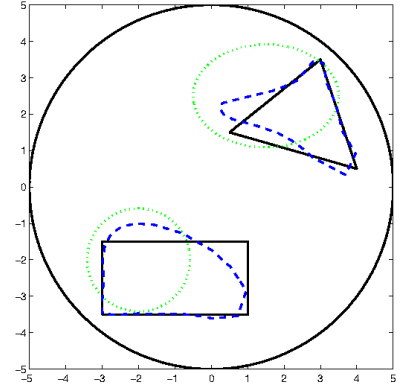
(c) Reconstructed interface.



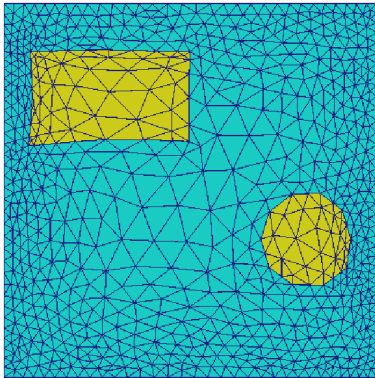
(d) 939 elements.



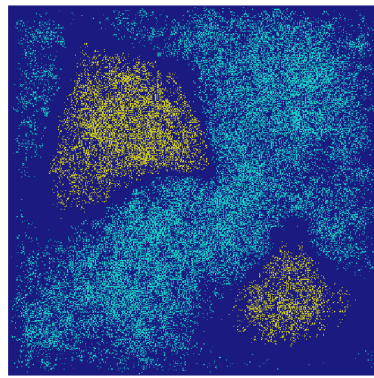
(e) 428251 elements.



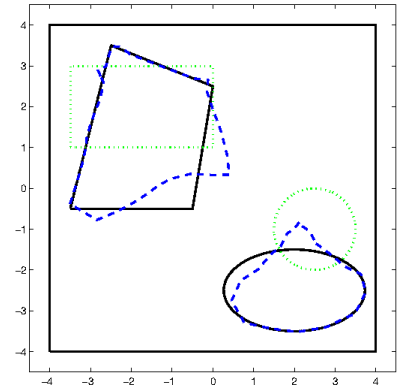
(f) Reconstructed interface.



(g) 1616 elements.



(h) 475744 elements.



(i) Reconstructed interface.

Figure 4.9 – Certified Descent Algorithm using multiple measurements: test cases (i)-(iii). Left: Initial mesh. Center: Final mesh. Right: Initial configuration (dotted green), target inclusion (solid black) and reconstructed interface (dashed blue).

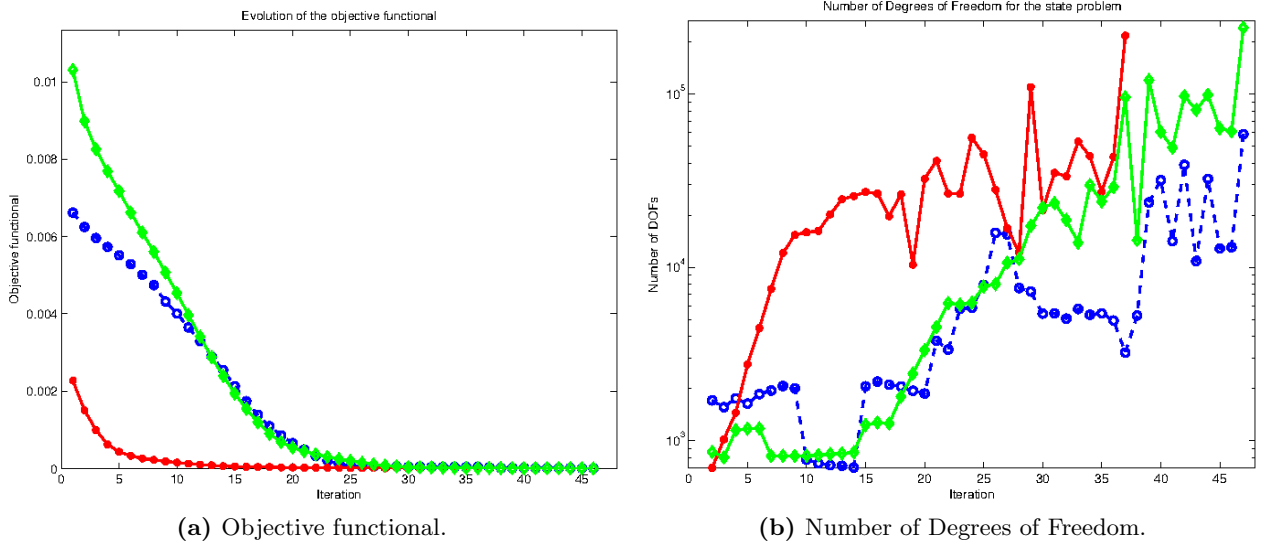


Figure 4.10 – Certified Descent Algorithm using multiple measurements: test case (i) blue circles; (ii) red stars; (iii) green diamonds. (a) Evolution of the objective functional. (b) Number of Degrees of Freedom.

As expected, the use of multiple measurements provides sufficient information to reconstruct the target interfaces near the boundary $\partial\mathcal{D}$ (Fig. 4.9c, 4.9f and 4.9i). On the contrary, the identification of the inner interfaces appears more difficult and with a less precise outcome. This phenomenon is due to the severe ill-posedness of the problem and the diffusive nature of the state equations is responsible for the loss of information in the center of \mathcal{D} , even when using several measurements. In particular, we remark that the final interface in figure 4.9c still presents two kinks from the initial configuration: this is mainly due to the aforementioned phenomenon and may be bypassed by choosing a regularizing scalar product. However, this approach would lead to a global smoothing of the reconstructed interface, including a loss of information about the potential sharp physical corners of the polygonal inclusion.

As previously remarked, figure 4.10a confirms the monotonically decreasing behavior of the objective functional with respect to the iterations of the algorithm. Moreover, the quantitative information associated with the estimator of the error in the shape gradient allows to derive a reliable stopping criterion for the optimization procedure which results to be fully automatic.

Eventually, coarse meshes are proved to be reliable for the computation during the initial iterations when the guessed position and shape of the inclusion are very unlikely to be precise. Within this context, even few Degrees of Freedom provide enough information to identify a genuine descent direction for the objective functional which we later certify using the discussed procedure based on the goal-oriented estimator for the error in the shape gradient. Thus, the same meshes may be used for several iterations increasing the number of Degrees of Freedom only when the descent direction is no more validated (Fig. 4.10b).

Both the inability of the method to reconstruct the interface far from the external boundary and the rapidly increasing number of Degrees of Freedom required to certify the descent direction clearly testifies the limitations of gradient-based approaches when dealing with the problem of Electrical Impedance Tomography. The ill-posedness of the problem is confirmed by the fact that after few

tens of iterations, we are unable to identify a genuine descent direction at a small computational cost since the number of Degrees of Freedom required by the certification procedure rapidly reaches 10^5 . Nevertheless, the main novelty of the Certified Descent Algorithm - that is its certification procedure - provides an heuristic criterion to stop the optimization routine when the number of Degrees of Freedom tends to explode, being the improvement of the solution negligible with respect to the huge precision the computation would require.

Chapter 5

Certified Descent Algorithm based on an equilibrated fluxes approach

We present a goal-oriented estimator of the error in the shape gradient which is fully computable and depends solely on local quantities. By means of the equilibrated fluxes approach, we construct a unified strategy valid for both conforming Finite Element and Discontinuous Galerkin discretizations. The estimated bounds for the error in the Quantity of Interest are validated and the new variant of the CDA is tested on the inverse problem of Electrical Impedance Tomography.

5.1 Estimate of the error in the shape gradient via the equilibrated fluxes

Starting from the framework described in section 2.4, we construct a goal-oriented estimator for the error in the shape gradient by evaluating the quantities E_u^h and E_p^h in (2.51)-(2.52). First of all, we observe that owing to the Kohn-Vogelius problem being self-adjoint, this reduces to estimating the quantity E_u^h for the Neumann and the Dirichlet cases. By recalling the expression (3.6) of the shape gradient of the Kohn-Vogelius functional, we may rewrite the error in the shape gradient as follows:

$$\begin{aligned} E^h &= \langle dJ(\Omega) - d_h J(\Omega), \theta^h \rangle \\ &= \langle G(\Omega, u_{\Omega,N}) - G(\Omega, u_{\Omega,N}^h), \theta^h \rangle - \langle G(\Omega, u_{\Omega,D}) - G(\Omega, u_{\Omega,D}^h), \theta^h \rangle \\ &\simeq H_{\Omega,N}(u_{\Omega,N} - u_{\Omega,N}^h) - H_{\Omega,D}(u_{\Omega,D} - u_{\Omega,D}^h) =: E_{u,N}^h - E_{u,D}^h. \end{aligned} \tag{5.1}$$

Before constructing the components of the estimator for the error in the shape gradient in (5.1), we recall the notion of equilibrated fluxes. In order to do so, we introduce the space of vector-valued functions $H(\text{div}) = \{\tau \in L^2(\mathcal{D}; \mathbb{R}^d) : \nabla \cdot \tau \in L^2(\mathcal{D})\}$ and the discrete space $W_\Omega^{h,\kappa}$ of the functions that restricted to a single element of the triangulation are Raviart-Thomas Finite Element functions of degree κ :

$$W_\Omega^{h,\kappa} := \{\tau^h \in H(\text{div}) : \tau^h|_T \in [\mathbb{P}^\kappa(T)]^d + x\mathbb{P}^\kappa(T) \ \forall T \in \mathcal{T}_h\}.$$

Remark 5.1. A function $\tau^h \in W_\Omega^{h,\kappa}$ is such that $\nabla \cdot \tau^h \in \mathbb{P}^\kappa(T) \ \forall T \in \mathcal{T}_h$, $\tau^h \cdot n_e \in \mathbb{P}^\kappa(e) \ \forall e \subset \partial T$, $\forall T \in \mathcal{T}_h$ and its normal trace is continuous across all edges $e \subset \partial T$ (cf. [70]).

5.2 Equilibrated fluxes for a conforming Finite Element approximation

Let us consider the discretization using a conforming Finite Element space introduced in section 4.1. The procedure to construct the equilibrated fluxes is performed via the solution of local subproblems defined on patches of elements using mixed Finite Element formulations. A key aspect of this approach - which will be precisely detailed in the following subsections - is the choice of the degree of the approximating functions for both the solution of the problems and the equilibrated fluxes.

In particular, it is well-known in the literature that in order to retrieve a good approximation of the error in a Quantity of Interest, a more precise approximation for the adjoint problems is required. Thus, we consider the space $V_\Omega^{h,\kappa}$ of the Lagrangian Finite Element functions of degree κ , that is the set of polynomials of degree less than or equal to κ on an element T , being $\kappa = \ell$ and $\kappa = m$, $m > \ell$ respectively the order of approximation for the state and adjoint equations.

For both the state problems and the computation of the descent direction, we consider a low-order approximation respectively using \mathbb{P}^1 and $\mathbb{P}^1 \times \mathbb{P}^1$ Lagrangian Finite Element functions. According to the requirement of higher-order methods to solve the adjoint problems, we consider a \mathbb{P}^2 Lagrangian Finite Element space for the discretization of (4.4).

5.2.1 Equilibrated fluxes for the state equations

The discretized solutions $u_{\Omega,i}^h$'s of the state problems are usually such that $-k_\Omega \nabla u_{\Omega,i}^h \notin H(\text{div})$ or $\nabla \cdot (-k_\Omega \nabla u_{\Omega,i}^h) + u_{\Omega,i}^h \neq 0$. On the contrary, the weak solutions $u_{\Omega,i}$'s - and their fluxes $\sigma_{\Omega,i} := -k_\Omega \nabla u_{\Omega,i}$ - fulfill $\sigma_{\Omega,i} \in H(\text{div})$ and $\nabla \cdot \sigma_{\Omega,i} + u_{\Omega,i} = 0$. In order to retrieve the aforementioned properties, we construct some discrete quantities known as equilibrated fluxes (cf. [78]):

Definition 5.2. Let $u_{\Omega,i}^h \in V_\Omega^{h,\ell}$ be the solution of a state problem (4.3) computed using Lagrangian Finite Element functions of degree ℓ . Let $\kappa = \max\{0, \ell - 1\}$, we define $\pi_Z^\kappa : L^2(\mathcal{D}) \rightarrow Z_\Omega^{h,\kappa}$ the L^2 -orthogonal projection operator onto the space $Z_\Omega^{h,\kappa}$ of the piecewise discontinuous Finite Element functions of degree κ . A function $\sigma_{\Omega,i}^h \in W_\Omega^{h,\kappa}$ is said to be an equilibrated flux for the problem (4.3) if

$$\nabla \cdot \sigma_{\Omega,i}^h + \pi_Z^\kappa u_{\Omega,i}^h = 0. \quad (5.2)$$

Under the previously introduced assumptions on the degree of the discretization spaces, we get that $\ell = 1$ and $\kappa = 0$, that is the equilibrated flux is sought in the lowest-order Raviart-Thomas space RT_0 and the projection operator returns \mathbb{P}^0 piecewise constant functions.

To practically reconstruct the equilibrated fluxes $\sigma_{\Omega,i}^h$'s, we follow the approach proposed by Ern and Vohralík in [132] which is based on the work [78] by Braess and Schöberl. In particular, we consider a procedure that starting from the Finite Element functions $u_{\Omega,i}^h$, $i = N, D$ constructs the equilibrated fluxes locally on subpatches of elements. Thus, for each vertex x_v , $v = 1, \dots, N_v$ of the elements in the computational mesh we introduce a linear shape function ψ_v such that $\psi_v(x_w) = \delta_{vw}$, δ being the classical Kronecker delta. The support of ψ_v is the subpatch centered in x_v and is denoted by ω_v . We remark that the family of functions ψ_v 's fulfills the condition known as partition of the

unity, that is

$$\sum_{v=1}^{N_v} \psi_v = 1.$$

In order to retrieve a precise approximation of the fluxes, we consider a dual mixed Finite Element formulation of the aforementioned local problems. First, let us denote by $W_{\omega_v}^{h,\kappa}$ (respectively $Z_{\omega_v}^{h,\kappa}$) the restriction to ω_v of the space $W_{\Omega}^{h,\kappa}$ (respectively $Z_{\Omega}^{h,\kappa}$) defined at the beginning of the chapter. Moreover, we introduce the following Finite Element spaces:

$$W_{v,0}^{h,\kappa} := \{\tau^h \in W_{\omega_v}^{h,\kappa} : \tau^h \cdot n_e = 0 \text{ on } e \in \partial\omega_v\}, \quad (5.3)$$

$$W_{v,1}^{h,\kappa} := \{\tau^h \in W_{\omega_v}^{h,\kappa} : \tau^h \cdot n_e = 0 \text{ on } e \in \partial\omega_v \setminus \mathcal{E}_h^{\mathcal{B}}\}. \quad (5.4)$$

For each vertex x_v , $v = 1, \dots, N_v$ and for $i = N, D$, we prescribe $(\sigma_{i,v}^h, t_{i,v}^h) \in W_{i,v}^{h,\kappa} \times Z_{\omega_v}^{h,\kappa}$ such that $\forall (\delta\sigma_i^h, \delta t_i^h) \in W_v^{h,\kappa} \times Z_{\omega_v}^{h,\kappa}$

$$\begin{aligned} \int_{\omega_v} \nabla \cdot \sigma_{i,v}^h \delta t_i^h \, dx + \int_{\omega_v} t_{i,v}^h \delta t_i^h \, dx &= - \int_{\omega_v} \left(k_{\Omega} \nabla u_{\Omega,i}^h \cdot \nabla \psi_v + u_{\Omega,i}^h \psi_v \right) \delta t_i^h \, dx, \\ \int_{\omega_v} \sigma_{i,v}^h \cdot \delta \sigma_i^h \, dx - \int_{\omega_v} k_{\Omega} t_{i,v}^h \nabla \cdot \delta \sigma_i^h \, dx &= - \int_{\omega_v} k_{\Omega} \psi_v \nabla u_{\Omega,i}^h \cdot \delta \sigma_i^h \, dx. \end{aligned} \quad (5.5)$$

The spaces in which the trial and the test functions are sought are detailed below. It is important to highlight the different nature of problem (5.5) when the patch ω_v is centered on a vertex belonging to the interior of \mathcal{D} or to its boundary $\partial\mathcal{D}$. As Braess and Schöberl remark in [78], some caution has to be used when dealing with the corresponding boundary conditions: in particular, a flux-free condition is imposed on the whole boundary $\partial\omega_v$ of the patch for interior vertices, whereas it is limited to the edges in $\partial\omega_v \setminus \mathcal{E}_h^{\mathcal{B}}$ for points which belong to the external boundary of the global domain.

To construct the equilibrated fluxes for the Neumann state problem on ω_v centered in a vertex $x_v \in \partial\mathcal{D}$, equation (5.5) is solved using the spaces

$$W_{N,v}^{h,\kappa} := \{\tau^h \in W_{\omega_v}^{h,\kappa} : \tau^h \cdot n_e = 0 \text{ on } e \in \partial\omega_v \setminus \mathcal{E}_h^{\mathcal{B}} \text{ and } \tau^h \cdot n_e = \pi_{W \cdot n}^{\kappa}(\psi_v g) \text{ on } e \in \partial\omega_v \cap \mathcal{E}_h^{\mathcal{B}}\}, \quad (5.6)$$

$$W_v^{h,\kappa} = W_{v,0}^{h,\kappa}. \quad (5.7)$$

When considering the Dirichlet state problem on ω_v centered in $x_v \in \partial\mathcal{D}$, the trial and test spaces read as follows:

$$W_{D,v}^{h,\kappa} = W_v^{h,\kappa} = W_{v,1}^{h,\kappa}. \quad (5.8)$$

Eventually, for the vertices $x_v \in \text{int}(\mathcal{D})$, we solve (5.5) using the spaces

$$W_{N,v}^{h,\kappa} = W_{D,v}^{h,\kappa} = W_v^{h,\kappa} = W_{v,0}^{h,\kappa}. \quad (5.9)$$

In (5.6), $\pi_{W \cdot n}^{\kappa}$ stands for the L^2 -orthogonal projection operator from $L^2(\partial\mathcal{D})$ to the space $W_{\Omega}^{h,\kappa} \cdot n$ of polynomial functions of degree κ on the external boundary. For additional detail on the procedure to construct the equilibrated fluxes and on the properties of the resulting *a posteriori* error estimators, we refer to [132].

We now extend all the $\sigma_{i,v}^h$'s by zero in $\mathcal{D} \setminus \omega_v$. By combining the above information arising from all the subpatches, we may retrieve the global equilibrated fluxes for the state problems:

$$\sigma_{\Omega,i}^h = \sum_{v=1}^{N_v} \sigma_{i,v}^h, \quad i = N, D. \quad (5.10)$$

Lemma 5.3. *For the case of the Neumann state problem, there holds*

$$\sigma_{\Omega,N}^h \cdot n = \pi_{W,n}^\kappa(g) \quad \text{on} \quad \partial\mathcal{D}. \quad (5.11)$$

Proof. Let χ_v^e be equal to 1 if a given edge $e \in \mathcal{E}_h^\mathcal{B}$ belongs to the subpatch ω_v centered in x_v and 0 otherwise. Hence,

$$\sigma_{\Omega,N}^h|_e = \sum_{v=1}^{N_v} \chi_v^e \sigma_{N,v}^h.$$

Let $\delta u^h \in (W_\Omega^{h,\kappa} \cdot n)|_e$ be a polynomial function of degree κ on the edge $e \in \mathcal{E}_h^\mathcal{B}$. Owing to the condition on the normal trace $\sigma_{N,v}^h \cdot n_e$ in (5.6), we get

$$\langle \sigma_{\Omega,N}^h \cdot n_e, \delta u^h \rangle_e = \sum_{v=1}^{N_v} \chi_v^e \langle \sigma_{N,v}^h \cdot n_e, \delta u^h \rangle_e = \sum_{v=1}^{N_v} \chi_v^e \langle \psi_v g, \delta u^h \rangle_e = \langle g, \delta u^h \rangle_e,$$

where the last equality follows from the partition of the unity property fulfilled by the functions ψ_v 's. The result is inferred by observing that the previous chain of equality holds $\forall \delta u^h \in (W_\Omega^{h,\kappa} \cdot n)|_e$, $\forall e \in \mathcal{E}_h^\mathcal{B}$. \square

5.2.2 Equilibrated fluxes for the adjoint equations

Following the same approach discussed above for the state problems, we define the equilibrated fluxes for the adjoint problems:

Definition 5.4. Let $r_{\Omega,i}^h \in V_\Omega^{h,m}$ be the solution of an adjoint problem (4.4) computed using Lagrangian Finite Element functions of degree m . Let $\kappa = \max\{0, m - 1\}$ and $\pi_Z^\kappa : L^2(\mathcal{D}) \rightarrow Z_\Omega^{h,\kappa}$ the L^2 -orthogonal projection operator onto the space $Z_\Omega^{h,\kappa}$ defined in the previous section. A function $\xi_{\Omega,i}^h \in W_\Omega^{h,\kappa}$ is said to be an equilibrated flux for the problem (4.4) if

$$\nabla \cdot \xi_{\Omega,i}^h + \pi_Z^\kappa r_{\Omega,i}^h = -\pi_Z^\kappa \left(\nabla \cdot (k_\Omega M(\theta^h) \nabla u_{\Omega,i}^h) + \nabla \cdot \theta^h u_{\Omega,i}^h \right). \quad (5.12)$$

Having in mind that the adjoint equations are solved using \mathbb{P}^2 Lagrangian Finite Element functions - that is $m = 2$ - it follows that the equilibrated fluxes $\xi_{\Omega,i}^h$'s are constructed via RT_1 Raviart-Thomas functions of degree 1 and the operator π_Z^κ projects functions from $L^2(\mathcal{D})$ to the discrete space of piecewise discontinuous Finite Elements of degree 1.

The computation of the equilibrated fluxes for the adjoint problems is again performed via the solution of a mixed Finite Element problem. We consider the same discrete spaces introduced in definitions (5.7) to (5.9), whereas the space $W_{N,v}^{h,\kappa}$ associated with the Neumann adjoint problem featuring a patch centered on a boundary node is $W_{v,0}^{h,\kappa}$. Thus, for each subpatch ω_v , $v = 1, \dots, N_v$ and for $i = N, D$, we seek $(\xi_{i,v}^h, q_{i,v}^h) \in W_{i,v}^{h,\kappa} \times Z_{\omega_v}^{h,\kappa}$ such that $\forall (\delta \xi_i^h, \delta q_i^h) \in W_v^{h,\kappa} \times Z_{\omega_v}^{h,\kappa}$

$$\begin{aligned} \int_{\omega_v} \nabla \cdot \xi_{i,v}^h \delta q_i \, dx + \int_{\omega_v} q_{i,v}^h \delta q_i \, dx &= \int_{\omega_v} \left(k_\Omega M(\theta^h) \nabla u_{\Omega,i}^h \cdot \nabla \psi_v - \nabla \cdot \theta^h u_{\Omega,i}^h \psi_v \right) \delta q_i \, dx \\ &\quad - \int_{\omega_v} \left(k_\Omega \nabla r_{\Omega,i}^h \cdot \nabla \psi_v + r_{\Omega,i}^h \psi_v \right) \delta q_i \, dx, \\ \int_{\omega_v} \xi_{i,v}^h \cdot \delta \xi_i \, dx - \int_{\omega_v} k_\Omega q_{i,v}^h \nabla \cdot \delta \xi_i \, dx &= \int_{\omega_v} k_\Omega \psi_v M(\theta^h) \nabla u_{\Omega,i}^h \cdot \delta \xi_i \, dx - \int_{\omega_v} k_\Omega \psi_v \nabla r_{\Omega,i}^h \cdot \delta \xi_i \, dx. \end{aligned} \quad (5.13)$$

The corresponding equilibrated fluxes $\xi_{\Omega,i}^h$'s are obtained by extending the functions $\xi_{i,v}^h$'s by zero in $\mathcal{D} \setminus \omega_v$ and by combining the previously computed local information:

$$\xi_{\Omega,i}^h = \sum_{v=1}^{N_v} \xi_{i,v}^h, \quad i = N, D.$$

Remark 5.5. A key aspect of the discussed procedure is represented by the local nature of the problems to be solved for the construction of the equilibrated fluxes. The advantage of this approach is twofold. On the one hand, solving the local problems (5.5)-(5.13) is computationally inexpensive owing to the small size of the subpatches. On the other hand, every problem set on a subpatch is independent from the remaining ones thus it is straightforward to implement a version of the procedure that can efficiently exploit modern parallel architectures.

5.2.3 Goal-oriented equilibrated fluxes error estimator

As previously stated, the construction of the error estimator for the shape gradient for the case of the Electrical Impedance Tomography reduces to the evaluation of (2.51) for the Neumann and Dirichlet problems. For this purpose, we introduce respectively the quantities $E_{u,N}^h$ and $E_{u,D}^h$ and two parameters ζ_i 's such that $\zeta_N := 1$ and $\zeta_D := 0$. By exploiting the formulation of the bilinear and linear forms (4.1)-(4.2) and adding the expression of the equilibrated fluxes (5.2), $E_{u,i}^h$ reads as:

$$\begin{aligned} E_{u,i}^h &:= F_{\Omega,i}(r_{\Omega,i}) - a_\Omega(u_{\Omega,i}^h, r_{\Omega,i}) = \zeta_i \int_{\partial \mathcal{D}} g r_{\Omega,i} \, ds - \int_{\mathcal{D}} \left(k_\Omega \nabla u_{\Omega,i}^h \cdot \nabla r_{\Omega,i} + u_{\Omega,i}^h r_{\Omega,i} \right) dx \\ &\quad + \int_{\mathcal{D}} \left(\nabla \cdot \sigma_{\Omega,i}^h + \pi_Z^\kappa u_{\Omega,i}^h \right) r_{\Omega,i} \, dx. \end{aligned} \quad (5.14)$$

Integrating by parts the last integral and owing to lemma 5.3 and to $r_{\Omega,D} = 0$ on $\partial \mathcal{D}$, we obtain

$$E_{u,i}^h = \zeta_i \int_{\partial \mathcal{D}} \left(g - \pi_{W,n}^\kappa g \right) r_{\Omega,i} \, ds + \int_{\mathcal{D}} \left(\pi_Z^\kappa u_{\Omega,i}^h - u_{\Omega,i}^h \right) r_{\Omega,i} \, dx - \int_{\mathcal{D}} \left(\sigma_{\Omega,i}^h + k_\Omega \nabla u_{\Omega,i}^h \right) \cdot \nabla r_{\Omega,i} \, dx. \quad (5.15)$$

By adding and subtracting the corresponding terms featuring the Finite Element counterparts $r_{\Omega,i}^h$'s of the adjoint solutions and owing to definition 5.4 of the equilibrated fluxes $\xi_{\Omega,i}^h$'s, we are finally able to derive the expression of the errors $E_{u,i}^h$'s:

$$\begin{aligned} E_{u,i}^h &= \zeta_i \int_{\partial\mathcal{D}} \left(g - \pi_{W \cdot n}^\kappa g \right) r_{\Omega,i}^h \, ds + \zeta_i \int_{\partial\mathcal{D}} \left(g - \pi_{W \cdot n}^\kappa g \right) (r_{\Omega,i} - r_{\Omega,i}^h) \, ds \\ &\quad + \int_{\mathcal{D}} \left(\pi_Z^\kappa u_{\Omega,i}^h - u_{\Omega,i}^h \right) r_{\Omega,i}^h \, dx + \int_{\mathcal{D}} \left(\pi_Z^\kappa u_{\Omega,i}^h - u_{\Omega,i}^h \right) (r_{\Omega,i} - r_{\Omega,i}^h) \, dx \\ &\quad + \int_{\mathcal{D}} \left(\sigma_{\Omega,i}^h + k_\Omega \nabla u_{\Omega,i}^h \right) \cdot k_\Omega^{-1} \xi_{\Omega,i}^h \, dx - \int_{\mathcal{D}} \left(\sigma_{\Omega,i}^h + k_\Omega \nabla u_{\Omega,i}^h \right) \cdot \left(\nabla r_{\Omega,i} + k_\Omega^{-1} \xi_{\Omega,i}^h \right) \, dx. \end{aligned} \quad (5.16)$$

We remark that in (5.16) both the exact and the discretized solutions of the adjoint problems appear. From a practical point of view, in order to fully compute the quantity (5.16) we substitute the exact solutions with their Finite Element counterparts $r_{\Omega,i}^h \in V_\Omega^{h,m}$ obtained by the high-order approximation of (4.4). The corresponding approximated solutions are then replaced by the projection of the high-order approximations onto the space $V_\Omega^{h,\ell}$ of the low-order Finite Element functions used for the discretization of the state problems. Let $I_m^\ell : V_\Omega^{h,m} \rightarrow V_\Omega^{h,\ell}$ be the projection operator from the space of high-order approximations to the low-order one. The fully computable version of the estimator of the quantity $E_{u,i}^h$ is obtained by substituting $r_{\Omega,i} - r_{\Omega,i}^h$ with $r_{\Omega,i}^h - I_m^\ell r_{\Omega,i}^h$ and $\nabla r_{\Omega,i}$ with $\nabla r_{\Omega,i}^h$ in (5.16). By plugging the expressions of $E_{u,N}^h$ and $E_{u,D}^h$ arising from (5.16) into (5.1), we obtain a computable expression of the error in the shape gradient and the bound \bar{E} follows by considering its absolute value.

Remark 5.6. The goal-oriented error estimators constructed using the equilibrated fluxes approach are known to be asymptotically exact (cf. [14]). Owing to the aforementioned asymptotic exactness, the term \bar{E} tends to zero as the mesh size tends to zero. This property plays a crucial role since it guarantees that the mesh adaptation routine performed to certify the descent direction (cf. algorithm 2.3 - step 5) eventually leads to the fulfillment of condition (2.11).

5.3 Discontinuous Galerkin approximation

In this section, we present an alternative strategy for the approximation of the EIT problem based on the Symmetric Weighted Interior Penalty Discontinuous Galerkin (SWIP-DG) formulation. Let us consider the notations introduced in section 4.1 for the triangulation \mathcal{T}_h . The Discontinuous Galerkin (DG) problems are solved within the space

$$V_\Omega^{h,\kappa} := \{u^h \in L^2(\mathcal{D}) : u^h|_T \in \mathbb{P}^\kappa(T) \, \forall T \in \mathcal{T}_h\}$$

of the discontinuous functions whose restrictions to a single element are polynomials of degree less than or equal to κ . When dealing with DG formulations, discontinuous functions - as the ones of the aforementioned space $V_\Omega^{h,\kappa}$ - which are double-valued on $\mathcal{E}_h^\mathcal{T}$ and single-valued on $\mathcal{E}_h^\mathcal{B}$ have to be properly handled. We define the jump of u^h across the edge e shared by the elements $T^\pm(e)$ as

$$[[u^h]]_e := u^h|_{T^-(e)} - u^h|_{T^+(e)}. \quad (5.17)$$

In a similar fashion, the weighted average of u^h on $e \in \mathcal{E}_h^I$ reads as follows

$$\{\{u^h\}\}_\alpha := \alpha_{T^-(e),e} u^h|_{T^-(e)} + \alpha_{T^+(e),e} u^h|_{T^+(e)}. \quad (5.18)$$

where the weights are non-negative quantities such that $\alpha_{T^-(e),e} + \alpha_{T^+(e),e} = 1$. On boundary edges, we set $\llbracket u^h \rrbracket_e = u^h|_e$, $\alpha_{T^-(e),e} = 1$ and $\{\{u^h\}\}_\alpha = u^h$.

Classical Discontinuous Galerkin methods use arithmetic averages in (5.18), that is for all edges the weights are constant and equal $\alpha_{T^-(e),e} = \alpha_{T^+(e),e} = \frac{1}{2}$. As stated in the introduction, in recent years there has been a growing interest towards the so-called Symmetric Weighted Interior Penalty DG methods, especially when dealing with problems featuring inhomogeneous coefficients for the diffusion term (cf. [92, 131]). In particular, these methods rely on the definition of weights based on the information carried by the diffusion tensor. For the case of the Electrical Impedance Tomography under analysis, this results in the following weights based on the different values of the electrical conductivity:

$$\alpha_{T^-(e),e} := \frac{k_\Omega|_{T^+(e)}}{k_\Omega|_{T^+(e)} + k_\Omega|_{T^-(e)}}, \quad \alpha_{T^+(e),e} := \frac{k_\Omega|_{T^-(e)}}{k_\Omega|_{T^+(e)} + k_\Omega|_{T^-(e)}}.$$

It is well-known in the literature [121] that the bilinear form associated with Discontinuous Galerkin methods may suffer from lack of coercivity thus preventing the discrete problem from having a unique solution. A widely-spread workaround (cf. [248]) is represented by the interior penalty approach that introduces a *sufficiently large* penalization in order to retrieve the coercivity of the discrete bilinear form. Owing to the idea of exploiting the information carried by the diffusion tensor to construct the weights for the jump term, we define the stabilization parameter in a similar way [114]:

$$\gamma_e := \beta_e \frac{k_\Omega|_{T^+(e)} k_\Omega|_{T^-(e)}}{k_\Omega|_{T^+(e)} + k_\Omega|_{T^-(e)}}$$

where $\beta_e > 0$ is a user-dependent parameter.

As per the conforming Finite Element approximation described in the previous section, first we introduce the discrete state and adjoint problems and then we construct the equilibrated fluxes via a procedure relying solely on local quantities. As previously stated, a key aspect of this approach is represented by the choice of the degree of the approximating functions for both the solution of the problems and the equilibrated fluxes. The details of this choice will be discussed in the following subsections. For the sake of readability, from now on we will omit the subscript e associated with jumps, weights and averages if there is no risk of ambiguity.

5.3.1 Weak imposition of the essential boundary conditions

We present a formal derivation of the variational formulation of an elliptic problem featuring weakly-imposed Dirichlet boundary conditions. The idea of this approach dates back to the classical paper by Nitsche [211] and has been extensively studied in recent years by several authors (cf. e.g. [115] and references therein). We recall that the solution of a Boundary Value Problem may be interpreted as an optimization problem. Let us introduce the Lagrangian functional associated with the state problem (3.2) featuring Dirichlet boundary conditions:

$$\Lambda(w, \lambda) = \frac{1}{2} \int_{\mathcal{D}} (k_\Omega |\nabla w|^2 + |w|^2) dx - \int_{\partial \mathcal{D}} \lambda (w - U_D) ds. \quad (5.19)$$

The solution of the aforementioned PDE is equivalent to the following min-max problem:

$$\min_{w \in H^1(\mathcal{D})} \max_{\lambda \in H^{-\frac{1}{2}}(\mathcal{D})} \Lambda(w, \lambda).$$

The first-order optimality conditions for (5.19) read as

$$\begin{cases} \int_{\mathcal{D}} (k_{\Omega} \nabla w \cdot \nabla \delta w + w \delta w) dx - \int_{\partial \mathcal{D}} \lambda \delta w \, ds = 0, \\ \int_{\partial \mathcal{D}} (w - U_D) \delta \lambda \, ds = 0. \end{cases}$$

From the second condition, we retrieve the Dirichlet boundary condition on $\partial \mathcal{D}$. Integrating by parts the first condition and owing to the strong form of the problem, we obtain

$$\int_{\partial \mathcal{D}} (k_{\Omega} \nabla w \cdot n - \lambda) \delta w \, ds = 0.$$

By plugging $\lambda = k_{\Omega} \nabla w \cdot n$ on $\partial \mathcal{D}$ into (5.19) we may now derive the following dual variational problem by seeking $w \in H^1(\mathcal{D})$ such that $\forall \delta w \in H^1(\mathcal{D})$

$$\int_{\mathcal{D}} (k_{\Omega} \nabla w \cdot \nabla \delta w + w \delta w) dx - \int_{\partial \mathcal{D}} (k_{\Omega} \nabla w \cdot n \delta w + w k_{\Omega} \nabla \delta w \cdot n) ds = - \int_{\partial \mathcal{D}} U_D k_{\Omega} \nabla \delta w \cdot n \, ds. \quad (5.20)$$

We remark that the bilinear form on the left-hand side of (5.20) is not coercive thus we cannot establish the well-posedness of this problem. To bypass this issue, we consider the following augmented Lagrangian functional and we construct the corresponding dual variational formulation for the problem under analysis:

$$\Upsilon(w, \lambda, \gamma) = \Lambda(w, \lambda) + \frac{1}{2} \int_{\partial \mathcal{D}} \gamma (w - U_D)^2 ds. \quad (5.21)$$

Following the same procedure used to derive (5.20), we seek $w \in H^1(\mathcal{D})$ such that $\forall \delta w \in H^1(\mathcal{D})$

$$\begin{aligned} \int_{\mathcal{D}} (k_{\Omega} \nabla w \cdot \nabla \delta w + w \delta w) dx - \int_{\partial \mathcal{D}} (k_{\Omega} \nabla w \cdot n \delta w + w k_{\Omega} \nabla \delta w \cdot n) ds + \int_{\partial \mathcal{D}} \gamma w \delta w \, ds \\ = \int_{\partial \mathcal{D}} U_D (\gamma \delta w - k_{\Omega} \nabla \delta w \cdot n) ds. \end{aligned} \quad (5.22)$$

It is straightforward to observe that the bilinear form on the left-hand side of (5.22) is coercive owing a *sufficiently large* value of γ is chosen.

5.3.2 The state problems

In order to appropriately handle the terms involving the effect of the boundary data in the estimator for the error in the shape gradient, the boundary conditions have to imposed using the same strategy in both the weak and the discrete formulation. Owing to the fact that the essential boundary conditions are classically verified in a weak sense in Discontinuous Galerkin methods, we consider an alternative formulation of (4.1)-(4.2) to weakly impose the Dirichlet boundary condition on $\partial \mathcal{D}$ as discussed in

subsection 5.3.1. Let $\zeta_N := 1$ and $\zeta_D := 0$. The bilinear forms $a_{\Omega,i}(\cdot, \cdot)$ and the linear ones $F_{\Omega,i}(\cdot)$ associated with (3.1) and (3.2) respectively read as:

$$\begin{aligned} a_{\Omega,i}(u_{\Omega,i}, \delta u) &= \int_{\mathcal{D}} \left(k_{\Omega} \nabla u_{\Omega,i} \cdot \nabla \delta u + u_{\Omega,i} \delta u \right) dx \\ &\quad - (1 - \zeta_i) \int_{\partial \mathcal{D}} \left(k_{\Omega} \nabla u_{\Omega,i} \cdot n \delta u + u_{\Omega,i} k_{\Omega} \nabla \delta u \cdot n \right) ds \\ &\quad + (1 - \zeta_i) \int_{\partial \mathcal{D}} \gamma u_{\Omega,i} \delta u \, ds, \end{aligned} \quad (5.23)$$

$$F_{\Omega,N}(\delta u) = \int_{\partial \mathcal{D}} g \delta u \, ds \quad , \quad F_{\Omega,D}(\delta u) = \int_{\partial \mathcal{D}} U_D (\gamma \delta u - k_{\Omega} \nabla \delta u \cdot n) ds. \quad (5.24)$$

The variational formulation of the state equations (3.1) and (3.2) reads as follows: for $i = N, D$ we seek $u_{\Omega,i} \in H^1(\mathcal{D})$ such that

$$a_{\Omega,i}(u_{\Omega,i}, \delta u_i) = F_{\Omega,i}(\delta u_i) \quad \forall \delta u_i \in H^1(\mathcal{D}). \quad (5.25)$$

The corresponding discrete bilinear and linear forms arising from the Interior Penalty Discontinuous Galerkin method have the following expressions:

$$\begin{aligned} a_{\Omega,i}^h(u_{\Omega,i}^h, \delta u^h) &= \sum_{T \in \mathcal{T}_h} \int_T \left(k_{\Omega} \nabla u_{\Omega,i}^h \cdot \nabla \delta u^h + u_{\Omega,i}^h \delta u^h \right) dx \\ &\quad - \sum_{e \in \mathcal{E}_h^I} \int_e \left(n_e \cdot \{ \{ k_{\Omega} \nabla u_{\Omega,i}^h \} \}_{\alpha} [\![\delta u^h]\!] + [\![u_{\Omega,i}^h]\!] n_e \cdot \{ \{ k_{\Omega} \nabla \delta u^h \} \}_{\alpha} \right) ds \\ &\quad - (1 - \zeta_i) \sum_{e \in \mathcal{E}_h^B} \int_e \left(n_e \cdot \{ \{ k_{\Omega} \nabla u_{\Omega,i}^h \} \}_{\alpha} [\![\delta u^h]\!] + [\![u_{\Omega,i}^h]\!] n_e \cdot \{ \{ k_{\Omega} \nabla \delta u^h \} \}_{\alpha} \right) ds \\ &\quad + \sum_{e \in \mathcal{E}_h^I} \int_e \frac{\gamma_e}{|e|} [\![u_{\Omega,i}^h]\!] [\![\delta u^h]\!] ds + (1 - \zeta_i) \sum_{e \in \mathcal{E}_h^B} \int_e \frac{\gamma_e}{|e|} [\![u_{\Omega,i}^h]\!] [\![\delta u^h]\!] ds, \end{aligned} \quad (5.26)$$

$$F_{\Omega,N}^h(\delta u^h) = \int_{\partial \mathcal{D}} g \delta u^h \, ds \quad , \quad F_{\Omega,D}^h(\delta u^h) = \sum_{e \in \mathcal{E}_h^B} \int_e U_D \left(\frac{\gamma_e}{|e|} \delta u^h - k_{\Omega} \nabla \delta u^h \cdot n_e \right) ds. \quad (5.27)$$

Thus, according to the SWIP-DG problem we seek $u_{\Omega,N}^h, u_{\Omega,D}^h \in V_{\Omega}^{h,\ell}$ such that

$$a_{\Omega,i}^h(u_{\Omega,i}^h, \delta u_i^h) = F_{\Omega,i}^h(\delta u_i^h) \quad \forall \delta u_i^h \in V_{\Omega}^{h,\ell}. \quad (5.28)$$

Concerning the degree of the Discontinuous Galerkin approximating functions, we maintain the same choice previously presented for the conforming Finite Element discretization, that is a low-order approximation based on piecewise linear polynomials ($\ell = 1$). In a similar fashion, the computation of the descent direction θ^h is performed by means of the conforming discretization using the space of $\mathbb{P}^1 \times \mathbb{P}^1$ Lagrangian Finite Element functions discussed at the beginning of the chapter.

5.3.3 The adjoint problems

The Symmetric Weighted Interior Penalty Discontinuous Galerkin formulation of the adjoint problems may be derived following the same procedure used for the state problems. In particular, the bilinear forms in (5.26) also stand for the Neumann and Dirichlet adjoint problems. The corresponding linear forms for $i = N, D$ read as

$$\begin{aligned} H_{\Omega,i}^h(\delta r^h) &= \sum_{T \in \mathcal{T}_h} \int_T \left(k_{\Omega} M(\theta^h) \nabla u_{\Omega,i}^h \cdot \nabla \delta r^h - \nabla \cdot \theta^h u_{\Omega,i}^h \delta r^h \right) dx \\ &\quad - \sum_{e \in \mathcal{E}_h^i} \int_e \left(n_e \cdot \{ \{ k_{\Omega} M(\theta^h) \nabla u_{\Omega,i}^h \} \}_\alpha \llbracket \delta r^h \rrbracket + \llbracket k_{\Omega} M(\theta^h) \nabla u_{\Omega,i}^h \rrbracket n_e \cdot \{ \{ \delta r^h \} \}_\alpha \right) ds \\ &\quad - (1 - \zeta_i) \int_{\partial \mathcal{D}} k_{\Omega} M(\theta^h) \nabla u_{\Omega,i}^h \cdot n \delta r^h ds. \end{aligned} \quad (5.29)$$

The discretized solutions of the adjoint problems are the functions $r_{\Omega,i}^h \in V_{\Omega}^{h,m}$ such that $\forall \delta r_i^h \in V_{\Omega}^{h,m}$

$$a_{\Omega,i}^h(\delta r_i^h, r_{\Omega,i}^h) = H_{\Omega,i}^h(\delta r_i^h) \quad , \quad i = N, D. \quad (5.30)$$

It is straightforward to verify that the SWIP-DG formulation of the adjoint problems is consistent, that is (5.30) stands substituting the analytical solutions $r_{\Omega,i}$'s to their discretized counterparts $r_{\Omega,i}^h$'s (cf. [121]). As previously stated, this property plays a crucial role in the construction of discretizations of optimal order in terms of target functionals and we refer to [157] for a detailed presentation of this subject. In order to obtain a higher-order approximation of the adjoint problems, we consider $m = 2$, as per the case of the conforming Finite Element approximation in section 5.2.

5.4 Equilibrated fluxes for a Discontinuous Galerkin approximation

In this section we construct the equilibrated fluxes associated with the Discontinuous Galerkin approximations (5.28) and (5.30) and we derive the corresponding goal-oriented estimator for the error in the shape gradient. Following the procedure introduced for the case of a conforming Finite Element discretization, this problem reduces to estimating the quantity (5.1).

5.4.1 Equilibrated fluxes for the state equations

We introduced the notion of equilibrated fluxes for the state problems in definition 5.2. In particular, for each problem we aim to construct an $H(\text{div})$ -conforming flux $\sigma_{\Omega,i}^h \in W_{\Omega}^{h,\kappa}$ such that (5.2) stands. We recall that the state problems are approximated using Discontinuous Galerkin functions of degree $\ell = 1$, thus the fluxes are reconstructed using RT_0 Finite Element functions ($\kappa = 0$). Owing to the nature of the Degrees of Freedom of the lowest-order Raviart-Thomas Finite Element functions, the construction of the equilibrated fluxes is straightforward via the prescription of the normal fluxes

on all the edges:

$$\int_e \sigma_{\Omega,i}^h \cdot n_e \delta t^h ds = \int_e \left(\frac{\gamma_e}{|e|} \llbracket u_{\Omega,i}^h \rrbracket - n_e \cdot \{ \{ k_\Omega \nabla u_{\Omega,i}^h \} \}_\alpha \right) \delta t^h ds \quad , \quad \forall \delta t^h \in \mathbb{P}^\kappa(e) \quad \forall e \in \mathcal{E}_h^\mathcal{I} \quad (5.31)$$

$$\begin{aligned} \int_e \sigma_{\Omega,i}^h \cdot n_e \delta t^h ds &= (1 - \zeta_i) \int_e \left(\frac{\gamma_e}{|e|} (u_{\Omega,i}^h - U_D) - k_\Omega \nabla u_{\Omega,i}^h \cdot n_e \right) \delta t^h ds \\ &\quad - \zeta_i \int_e g \delta t^h ds \quad , \quad \forall \delta t^h \in \mathbb{P}^\kappa(e) \quad \forall e \in \mathcal{E}_h^\mathcal{B}. \end{aligned} \quad (5.32)$$

5.4.2 Equilibrated fluxes for the adjoint equations

In an analogous way, we may construct the equilibrated fluxes for the adjoint problems. We remark that owing to the higher-order approximation of (5.30) with respect to (5.28) - i.e. $m = 2$ -, the equilibrated fluxes $\xi_{\Omega,i}^h$'s in definition 5.4 are sought in the space $W_\Omega^{h,\kappa}$, $\kappa = 1$. The RT_1 reconstructed fluxes are such that

$$\int_e \xi_{\Omega,i}^h \cdot n_e \delta q_1^h ds = \int_e \left(\frac{\gamma_e}{|e|} \llbracket r_{\Omega,i}^h \rrbracket - n_e \cdot \{ \{ k_\Omega \nabla r_{\Omega,i}^h \} \}_\alpha \right) \delta q_1^h ds \quad , \quad \forall \delta q_1^h \in \mathbb{P}^\kappa(e) \quad \forall e \in \mathcal{E}_h^\mathcal{I} \quad (5.33)$$

$$\begin{aligned} \int_e \xi_{\Omega,i}^h \cdot n_e \delta q_1^h ds &= (1 - \zeta_i) \int_e \left(\frac{\gamma_e}{|e|} r_{\Omega,i}^h - k_\Omega \nabla r_{\Omega,i}^h \cdot n_e \right) \delta q_1^h ds \\ &\quad - \zeta_i \int_e k_\Omega M(\theta^h) \nabla u_{\Omega,i}^h \cdot n_e \delta q_1^h ds \quad , \quad \forall \delta q_1^h \in \mathbb{P}^\kappa(e) \quad \forall e \in \mathcal{E}_h^\mathcal{B} \end{aligned} \quad (5.34)$$

$$\begin{aligned} \int_T \xi_{\Omega,i}^h \cdot \delta q_2^h dx &= - \int_T k_\Omega \nabla r_{\Omega,i}^h \cdot \delta q_2^h dx + \sum_{e \subset \partial T \setminus \mathcal{E}_h^\mathcal{B}} \alpha_{T(e),e} \int_e k_\Omega \llbracket r_{\Omega,i}^h \rrbracket \delta q_2^h \cdot n_e ds \\ &\quad + (1 - \zeta_i) \sum_{e \subset \partial T \cap \mathcal{E}_h^\mathcal{B}} \int_e k_\Omega r_{\Omega,i}^h \delta q_2^h \cdot n_e ds \quad , \quad \forall \delta q_2^h \in [\mathbb{P}^{\kappa-1}(T)]^d \quad \forall T \in \mathcal{T}_h. \end{aligned} \quad (5.35)$$

Remark 5.7. The flux reconstruction procedure presented for both the state and adjoint equations relies solely on the computation of local quantities and is computationally inexpensive. A great advantage of the Discontinuous Galerkin framework is represented by the cheap algorithms to construct the equilibrated fluxes on an element-wise level as discussed by several authors, e.g. in [77, 101, 129, 175]. As previously remarked for the construction of the equilibrated fluxes in the case of conforming Finite Element discretizations, the local nature of the procedure allows the parallelization of the algorithm and the exploitation of modern parallel architectures.

5.4.3 Goal-oriented equilibrated fluxes error estimator

We may now evaluate the term (2.51) for the Neumann and Dirichlet problems by exploiting the information carried by (5.26) and (5.24). We recall that the Symmetric Weighted Interior Penalty Discontinuous Galerkin method under analysis is adjoint consistent (cf. [121]). Owing to the continuity of $r_{\Omega,i}$ and $k_\Omega \nabla r_{\Omega,i} \cdot n_e$ on all the edges e 's and adding the expression of the equilibrated fluxes (5.2),

we obtain:

$$\begin{aligned}
 E_{u,i}^h &:= F_{\Omega,i}(r_{\Omega,i}) - a_{\Omega,i}^h(u_{\Omega,i}^h, r_{\Omega,i}) = \\
 &= \zeta_i \int_{\partial\mathcal{D}} g r_{\Omega,i} \, ds + (1 - \zeta_i) \int_{\partial\mathcal{D}} U_D (\gamma r_{\Omega,i} - k_{\Omega} \nabla r_{\Omega,i} \cdot n) \, ds \\
 &\quad - \sum_{T \in \mathcal{T}_h} \int_T \left(k_{\Omega} \nabla u_{\Omega,i}^h \cdot \nabla r_{\Omega,i} + u_{\Omega,i}^h r_{\Omega,i} \right) dx + \sum_{e \in \mathcal{E}_h^I} \int_e \llbracket u_{\Omega,i}^h \rrbracket k_{\Omega} \nabla r_{\Omega,i} \cdot n_e \, ds \\
 &\quad + (1 - \zeta_i) \sum_{e \in \mathcal{E}_h^B} \int_e \llbracket u_{\Omega,i}^h \rrbracket k_{\Omega} \nabla r_{\Omega,i} \cdot n_e \, ds + \sum_{T \in \mathcal{T}_h} \int_T \left(\nabla \cdot \sigma_{\Omega,i}^h + \pi_Z^{\kappa} u_{\Omega,i}^h \right) r_{\Omega,i} \, dx.
 \end{aligned} \tag{5.36}$$

We integrate by parts the last integral and we plug in the expressions (5.31)-(5.32) of the equilibrated fluxes for the state problems. It follows from the homogeneous Dirichlet condition fulfilled by the adjoint solution $r_{\Omega,D}$ on $\partial\mathcal{D}$ that

$$\begin{aligned}
 E_{u,i}^h &= \zeta_i \int_{\partial\mathcal{D}} \left(g - \pi_W^{\kappa} n g \right) r_{\Omega,i} \, ds + (1 - \zeta_i) \int_{\partial\mathcal{D}} (u_{\Omega,i}^h - U_D) k_{\Omega} \nabla r_{\Omega,i} \cdot n \, ds \\
 &\quad + \sum_{e \in \mathcal{E}_h^I} \int_e \left(\llbracket u_{\Omega,i}^h \rrbracket k_{\Omega} \nabla r_{\Omega,i} \cdot n_e + \llbracket \sigma_{\Omega,i}^h \rrbracket n_e r_{\Omega,i} \right) ds \\
 &\quad + \sum_{T \in \mathcal{T}_h} \int_T \left(\pi_Z^{\kappa} u_{\Omega,i}^h - u_{\Omega,i}^h \right) r_{\Omega,i} \, dx - \sum_{T \in \mathcal{T}_h} \int_T \left(\sigma_{\Omega,i}^h + k_{\Omega} \nabla u_{\Omega,i}^h \right) \cdot \nabla r_{\Omega,i} \, dx.
 \end{aligned} \tag{5.37}$$

We remark that owing to the continuity of the normal traces of the fluxes, $\llbracket \sigma_{\Omega,i}^h \rrbracket \cdot n_e = 0$ for all the internal edges. By adding and subtracting the terms $r_{\Omega,i}^h$'s featuring the Discontinuous Galerkin approximations of the adjoint solutions and taking into account their equilibrated fluxes $\xi_{\Omega,i}^h$'s, the expressions of the errors $E_{u,i}^h$'s read as:

$$\begin{aligned}
 E_{u,i}^h &= \zeta_i \int_{\partial\mathcal{D}} \left(g - \pi_W^{\kappa} n g \right) r_{\Omega,i}^h \, ds + \zeta_i \int_{\partial\mathcal{D}} \left(g - \pi_W^{\kappa} n g \right) (r_{\Omega,i} - r_{\Omega,i}^h) \, ds \\
 &\quad - (1 - \zeta_i) \int_{\partial\mathcal{D}} (u_{\Omega,i}^h - U_D) \xi_{\Omega,i}^h \cdot n \, ds + (1 - \zeta_i) \int_{\partial\mathcal{D}} (u_{\Omega,i}^h - U_D) \left(k_{\Omega} \nabla r_{\Omega,i} + \xi_{\Omega,i}^h \right) \cdot n \, ds \\
 &\quad - \sum_{e \in \mathcal{E}_h^I} \int_e \llbracket u_{\Omega,i}^h \rrbracket \xi_{\Omega,i}^h \cdot n_e \, ds + \sum_{e \in \mathcal{E}_h^I} \int_e \llbracket u_{\Omega,i}^h \rrbracket \left(k_{\Omega} \nabla r_{\Omega,i} + \xi_{\Omega,i}^h \right) \cdot n_e \, ds \\
 &\quad + \sum_{T \in \mathcal{T}_h} \int_T \left(\pi_Z^{\kappa} u_{\Omega,i}^h - u_{\Omega,i}^h \right) r_{\Omega,i}^h \, dx + \sum_{T \in \mathcal{T}_h} \int_T \left(\pi_Z^{\kappa} u_{\Omega,i}^h - u_{\Omega,i}^h \right) (r_{\Omega,i} - r_{\Omega,i}^h) \, dx \\
 &\quad + \sum_{T \in \mathcal{T}_h} \int_T \left(\sigma_{\Omega,i}^h + k_{\Omega} \nabla u_{\Omega,i}^h \right) \cdot k_{\Omega}^{-1} \xi_{\Omega,i}^h \, dx - \sum_{T \in \mathcal{T}_h} \int_T \left(\sigma_{\Omega,i}^h + k_{\Omega} \nabla u_{\Omega,i}^h \right) \cdot \left(\nabla r_{\Omega,i} + k_{\Omega}^{-1} \xi_{\Omega,i}^h \right) \, dx.
 \end{aligned} \tag{5.38}$$

As already remarked in the estimator derived for the conforming Finite Element discretization, both the unknown exact solutions of the adjoint problems and their numerical counterparts appear in (5.38). Let $I_m^{\ell} : V_{\Omega}^{h,m} \rightarrow V_{\Omega}^{h,\ell}$ be the projection operator from the space of high-order Discontinuous Galerkin

approximations to the low-order one. The fully computable version of the estimator of the quantity $E_{u,i}^h$ is obtained by substituting $r_{\Omega,i} - r_{\Omega,i}^h$ with $r_{\Omega,i}^h - I_m^\ell r_{\Omega,i}^h$ and $\nabla r_{\Omega,i}$ with $\nabla r_{\Omega,i}^h$.

Eventually, the upper bound \bar{E} of the error in the shape gradient is obtained by plugging the expressions of $E_{u,N}^h$ and $E_{u,D}^h$ arising from (5.38) into (5.1) and by considering its absolute value.

Remark 5.8. In [205], the authors prove that the contribution of the terms in (5.38) featuring the exact solution of the adjoint problems is negligible and the goal-oriented error estimator constructed using the previously described equilibrated fluxes approach is asymptotically exact. This property guarantees that the bound \bar{E} of the error in the shape gradient tends to zero by reducing the mesh size. Hence, the mesh adaptation procedure performed by the Certified Descent Algorithm eventually leads to the fulfillment of condition (2.11).

5.5 Numerical results

In this section we present some numerical results of the application of the variant of the Certified Descent Algorithm that exploits the equilibrated fluxes approach to estimate the error in the shape gradient. We consider the problem of Electrical Impedance Tomography as a proof of concept to establish some properties of this variant of the Certified Descent Algorithm on a non-trivial scalar test case. As previously stated in chapter 4, the Certified Descent Algorithm does not aim to remedy the known issues of gradient-based shape optimization methods when dealing with inverse ill-posed problems as the Electrical Impedance Tomography. The current work presents an improvement of the original Certified Descent Algorithm introduced in [4], in particular by using solely local quantities to compute the error in the shape gradient and to perform the certification procedure. The numerical results in this section focus on the quantitative bound \bar{E} obtained using the equilibrated fluxes approach for both conforming Finite Element and Discontinuous Galerkin discretizations.

The framework of the numerical simulations is the same introduced in chapter 4. In particular, the results are obtained using FreeFem++ [160] and are based on a mesh moving approach for the deformation of the domain. A 2-mesh strategy is adopted as proposed in [28]: a fine mesh is used to solve the state problems whereas the computation of the descent direction is performed on a coarser triangulation. Within this framework, the correct number of inclusions is set at the beginning of the algorithm and remains the same throughout its evolution. Techniques based on both topological and shape gradients to account for topological changes inside the domain have been investigated in [163].

5.5.1 Numerical assessment of the goal-oriented estimator

First, we consider the test case in subsection 4.3.1, for which the analytical solution of the state problems is known and we evaluate the goal-oriented error estimators derived in subsections 5.2.3 and 5.4.3. We introduce the polar coordinate system (ρ, ϑ) and we set $\mathcal{D} := \{x = (x, y) \mid x^2 + y^2 \leq \rho_E^2\}$ and $\Omega := \{x = (x, y) \mid x^2 + y^2 \leq \rho_I^2\}$ with $\rho_E = 5$ and $\rho_I = 4$. The value of the conductivity parameter is $k_I = 10$ inside Ω and $k_E = 1$ in $\mathcal{D} \setminus \Omega$. We consider the Neumann boundary condition $g = \cos(M\vartheta)$, $M = 5$ and the Dirichlet datum U_D is the trace of the following function which is the

analytical solution of problem (3.1):

$$u_{\Omega,N} = \begin{cases} C_0 J_M \left(-i\rho k_I^{-\frac{1}{2}} \right) \cos(M\vartheta) & , \rho \in [0, \rho_I] \\ \left[C_1 J_M \left(-i\rho k_E^{-\frac{1}{2}} \right) + C_2 Y_M \left(-i\rho k_E^{-\frac{1}{2}} \right) \right] \cos(M\vartheta) & , \rho \in (\rho_I, \rho_E] \end{cases}$$

where $J_M(\cdot)$ and $Y_M(\cdot)$ respectively represent the first- and second-kind Bessel functions of order M . The constants C_0, \dots, C_2 are detailed in table 4.1.

We recall that for both the conforming Finite Element and the Discontinuous Galerkin discretizations we have $\ell = 1$ and $m = 2$, that is the state problems are solved using functions of degree 1, whereas the adjoint solutions are approximated using functions of degree 2. The corresponding equilibrated fluxes are sought respectively in the space of RT_0 and RT_1 Finite Element functions.

Figure 5.1a presents the convergence history of the discretization error in the shape gradient and the goal-oriented estimator \bar{E} for the case of conforming Finite Element. The corresponding quantities for the case of Discontinuous Galerkin are depicted in figure 5.1b.

Eventually, in figure 5.1c we present the effectivity indices for the discussed discretizations. The effectivity index is the ratio between the estimator \bar{E} and the exact error E^h . If the effectivity index is bigger (respectively smaller) than one, we are overestimating (respectively underestimating) the error. The evolution of the effectivity indices in figure 5.1c confirms that the constructed estimators are guaranteed - that is they provide an upper bound of the error - and are asymptotically exact since the effectivity index tends to 1 as the mesh size tends to 0.

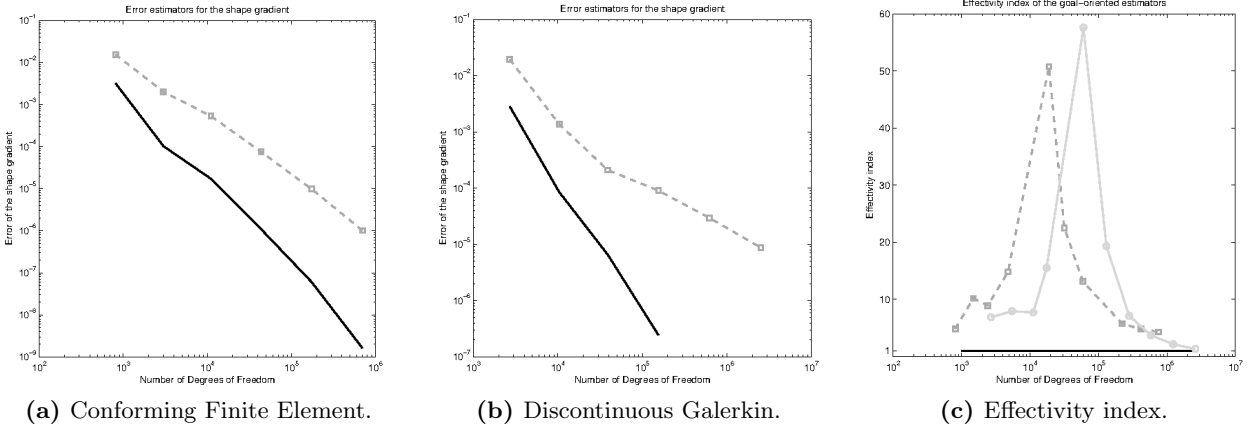


Figure 5.1 – Convergence rates and effectivity indices of the estimators of the error in the shape gradient with respect to the number of Degrees of Freedom. Analytical error in the shape gradient (solid black); goal-oriented estimator of the error based on the equilibrated fluxes (dashed gray squares) for the discretizations based on (a) conforming Finite Element and (b) Discontinuous Galerkin. (c) Effectivity indices for the conforming Finite Element (dark gray squares) and Discontinuous Galerkin (light gray circles).

5.5.2 Reconstruction of a single inclusion

We consider the problem of reconstructing an inclusion Ω in the domain \mathcal{D} by means of a couple of measurements (g, U_D) of the flux and the potential on the external boundary $\partial\mathcal{D}$. In all the following simulations, we consider a stopping criterion that combines the condition in step 8 of algorithm 2.3 and a bound on the number of admissible mesh elements - i.e. the size of the state and the adjoint problems. This choice is due to the ill-posed nature of the Electrical Impedance Tomography problem that we chose as test case for the Certified Descent Algorithm. As we will highlight throughout this section the ill-posedness of the problem represents an issue that prevents gradient-based strategies from efficiently solving the EIT problem since a huge precision is demanded after few iterations of the optimization procedure.

First, we consider the configuration described in figure 5.2a. The initial guess for the inclusion is represented by the circle of radius $\rho_{\text{ini}} = 2$. The Certified Descent Algorithm is able to correctly identify the interface along which the conductivity parameter k_Ω is discontinuous (Fig. 5.2a). Moreover, figure 5.2b shows that the objective functional $J(\Omega)$ is monotonically decreasing, meaning a genuine descent direction is computed at each iteration of the algorithm. In tables 5.1a-5.1b we present the specifics of the meshes used to certify the descent direction at several iterations of the CDA. In particular, we observe that coarse meshes are reliable during the initial iterations of the algorithm to identify a genuine descent direction, whereas the number of Degrees of Freedom increases when approaching a minimum of the functional $J(\Omega)$. This is also well-explained by figure 5.2c in which the evolution of the number of Degrees of Freedom is depicted. Eventually, we remark that figure 5.2c also highlights the ill-posed nature of the problem since a huge amount of Degrees of Freedom is rapidly required by the CDA to certify the descent direction, testifying the difficulties of gradient-based methods to handle inverse problems as the Electrical Impedance Tomography. By comparing the approximations arising from conforming Finite Element and Discontinuous Galerkin formulations, we remark that the latter provides sharper bounds of the error in the shape gradient thus allowing the algorithm to automatically stop for a given tolerance $\text{tol} = 10^{-6}$ (cf. table 5.1b). On the contrary, the certification in the case of conforming Finite Element is still able to identify a genuine descent direction at each iteration but rapidly requires a huge number of mesh elements making the computational cost explode.

The aforementioned issues are confirmed and highlighted by the test case in figure 5.2d. It is straightforward to observe that the Certified Descent Algorithm is able to identify a genuine descent direction at each iteration (Fig. 5.2e) and to reconstruct the interface in the region near the external boundary whereas the inner part is not correctly recovered. As previously stated, this phenomenon is due to the well-known ill-posedness of the problem and we cannot expect gradient-based strategies to successfully overcome this issue. These remarks are confirmed again by the rapidly exploding number of Degrees of Freedom required by the algorithm to certify the descent direction (Fig. 5.2f). A possible workaround is represented by the emerging field of hybrid imaging in which classical tomography techniques are coupled with acoustic or elastic waves (cf. [32]).

Though both the variant of the CDA based on the conforming Finite Element formulation and the one exploiting Discontinuous Galerkin functions is able to certify the descent direction at the beginning of the algorithm, the situation changes after few tens of iterations. In particular, the SWIP-DG formulation allows the computation of inexpensive and precise bounds of the error in the shape gradient, whereas using conforming Finite Element the computational cost rapidly becomes enormous making the certification procedure unfeasible (cf. table 5.2).

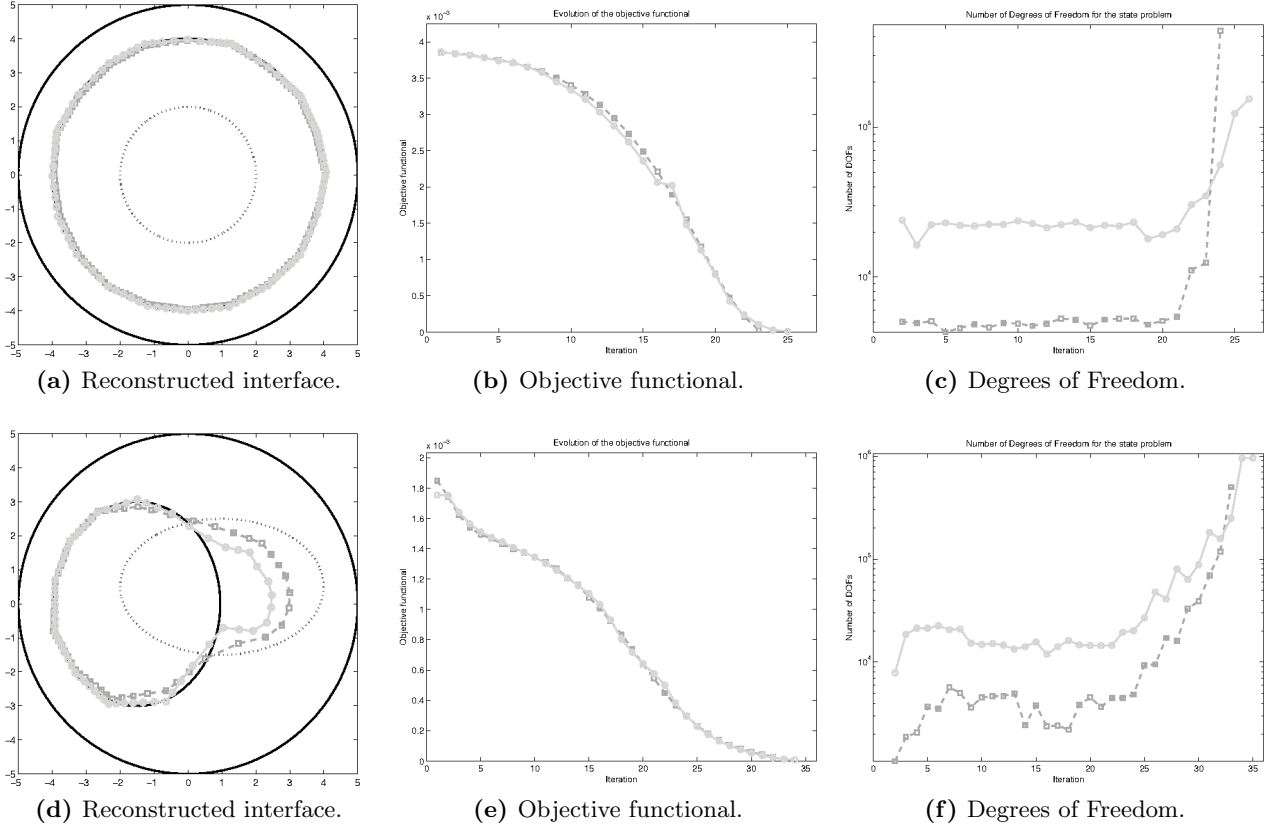


Figure 5.2 – Certified Descent Algorithm for the identification of one inclusion. Left: Initial configuration (dotted black), target inclusion (solid black) and reconstructed interface. Center: Evolution of the objective functional. Right: Number of Degrees of Freedom. Inversion performed using conforming Finite Element (dark gray squares) and Discontinuous Galerkin (light gray circles).

5.5.3 The case of two inclusions featuring multiple boundary measurements

In this subsection, we present a more involved test case in which the domain \mathcal{D} features two non-connected inclusions. As previously stated, we assume that the number of inclusions is set *a priori* and we restrict to the case of a two-valued conductivity parameter, that is we distinguish a value k_E for the background and a value k_I valid inside both the inclusions.

It is well-known in the literature that multiple boundary measurements are required to retrieve a correct approximation of the inclusion in Electrical Impedance Tomography. In this section, we consider $D = 10$ measurements such that $\forall j = 0, \dots, D - 1$

$$g_j(x, y) = (x + a_j y)^{b_j} a_j^{c_j} \quad , \quad a_j = 1 + 0.1j \quad , \quad b_j = \frac{j+1}{2} \quad , \quad c_j = j - 2 \left\lfloor \frac{j}{2} \right\rfloor .$$

As previously remarked, the Certified Descent Algorithm is able to identify the portions of the interfaces that lie near the external boundary $\partial\mathcal{D}$ whereas the inner parts suffer from a poor reconstruction (Fig. 5.3a). Moreover, also in this case after few tens of iterations, the certification procedure re-

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	8863	$-1.45 \cdot 10^{-6}$	$1.12 \cdot 10^{-6}$
5	8582	$-4.36 \cdot 10^{-6}$	$3.31 \cdot 10^{-6}$
10	8650	$-1.37 \cdot 10^{-5}$	$9.17 \cdot 10^{-6}$
15	9335	$-2.83 \cdot 10^{-5}$	$1.80 \cdot 10^{-5}$
20	19683	$-1.53 \cdot 10^{-5}$	$1.07 \cdot 10^{-5}$
22	864808	$-1.18 \cdot 10^{-6}$	$1.16 \cdot 10^{-6}$

(a) Conforming Finite Element.

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	5454	$-2.02 \cdot 10^{-6}$	$1.86 \cdot 10^{-6}$
5	7307	$-6.63 \cdot 10^{-6}$	$6.24 \cdot 10^{-6}$
10	7099	$-1.73 \cdot 10^{-5}$	$9.20 \cdot 10^{-6}$
15	7307	$-3.06 \cdot 10^{-5}$	$7.11 \cdot 10^{-6}$
20	10123	$-1.38 \cdot 10^{-5}$	$8.98 \cdot 10^{-6}$
24	51406	$-4.60 \cdot 10^{-7}$	$4.55 \cdot 10^{-7}$

(b) Discontinuous Galerkin.

Table 5.1 – Test case in figure 5.2a using (a) conforming Finite Element and (b) Discontinuous Galerkin. Approximated shape gradient and goal-oriented estimator for different meshes.

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	3366	$-2.29 \cdot 10^{-4}$	$1.79 \cdot 10^{-4}$
10	8312	$-7.63 \cdot 10^{-5}$	$5.49 \cdot 10^{-5}$
20	7893	$-1.29 \cdot 10^{-4}$	$1.03 \cdot 10^{-4}$
30	227847	$-6.16 \cdot 10^{-6}$	$6.15 \cdot 10^{-6}$
31	980555	$-3.62 \cdot 10^{-6}$	$3.60 \cdot 10^{-6}$

(a) Conforming Finite Element.

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	6189	$-2.21 \cdot 10^{-4}$	$1.34 \cdot 10^{-4}$
10	4868	$-7.80 \cdot 10^{-5}$	$4.64 \cdot 10^{-5}$
20	4842	$-1.62 \cdot 10^{-4}$	$1.30 \cdot 10^{-4}$
30	52595	$-4.71 \cdot 10^{-6}$	$4.60 \cdot 10^{-6}$
33	320137	$-2.54 \cdot 10^{-7}$	$2.25 \cdot 10^{-7}$

(b) Discontinuous Galerkin.

Table 5.2 – Test case in figure 5.2d using (a) conforming Finite Element and (b) Discontinuous Galerkin. Approximated shape gradient and goal-oriented estimator for different meshes.

quires a huge number of Degrees of Freedom to identify a genuine descent direction for the objective functional $J(\Omega)$ (Fig. 5.3c). The inability of the method to reconstruct the interface far from the external boundary and the rapidly increasing number of Degrees of Freedom required to certify the descent direction clearly testify the limitations of gradient-based approaches when dealing with Electrical Impedance Tomography.

Nevertheless, this new variant of the Certified Descent Algorithm proves to be able to certify the descent direction in order to construct a minimizing sequence of shapes for which the objective functional is monotonically decreasing (Fig. 5.3b). Moreover, the quantitative information carried by the error bound \bar{E} allows to derive a reliable stopping criterion that automatizes the overall optimization procedure.

Remark 5.9. The tables presented in this section show that the strategy based on a conforming Finite Element discretization rapidly requires a huge number of mesh elements to perform the certification of the descent direction. However, it is important to recall that the Discontinuous Galerkin formulations feature a higher number of Degrees of Freedom per mesh element, making the overall dimensions of the optimization problems comparable. Nevertheless, from a practical point of view the computation of the error bound \bar{E} in the framework of conforming Finite Element relies on the solution of a number of local subproblems on patches of element equal to the number of vertices of the triangulation \mathcal{T}_h . On the contrary, the Discontinuous Galerkin discretization is locally conservative and yields to a straightforward technique to construct the equilibrated fluxes based on an inexpensive and local post-

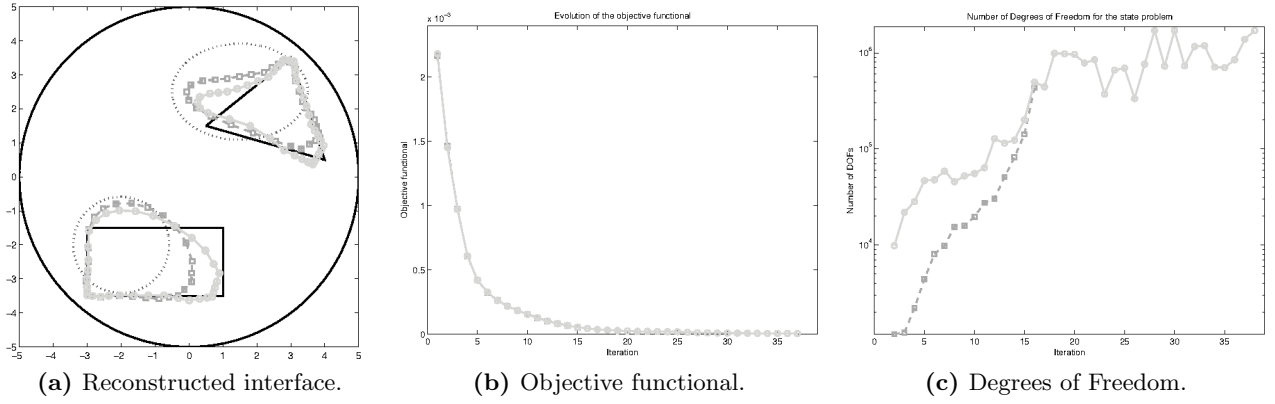


Figure 5.3 – Certified Descent Algorithm for the identification of two inclusions. (a) Initial configuration (dotted black), target inclusion (solid black) and reconstructed interface. (b) Evolution of the objective functional. (c) Number of Degrees of Freedom. Inversion performed using conforming Finite Element (dark gray squares) and Discontinuous Galerkin (light gray circles).

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	2221	$-1.62 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$
10	56487	$-1.09 \cdot 10^{-5}$	$1.04 \cdot 10^{-5}$
14	852782	$-3.36 \cdot 10^{-6}$	$3.21 \cdot 10^{-6}$

(a) Conforming Finite Element.

Iter. j	$\#\mathcal{T}_h$	$\langle d_h J(\Omega_j), \theta^h \rangle$	\bar{E}
1	7282	$-1.59 \cdot 10^{-3}$	$1.20 \cdot 10^{-3}$
10	42564	$-1.10 \cdot 10^{-5}$	$7.27 \cdot 10^{-6}$
20	282718	$-3.82 \cdot 10^{-6}$	$2.61 \cdot 10^{-6}$
30	389571	$-1.40 \cdot 10^{-6}$	$8.74 \cdot 10^{-7}$
36	568548	$-1.92 \cdot 10^{-7}$	$1.87 \cdot 10^{-7}$

(b) Discontinuous Galerkin.

Table 5.3 – Test case in figure 5.3a using (a) conforming Finite Element and (b) Discontinuous Galerkin. Approximated shape gradient and goal-oriented estimator for different meshes.

processing of the solutions of the state and adjoint problems. Thus, both approaches result valid and present an improvement of the original Certified Descent Algorithm introduced in [4] which required the solution of additional global problems to perform the certification procedure. Nevertheless, the computational cost of the version based on the Discontinuous Galerkin formulation appears more competitive, especially in view of future developments focusing on vectorial and three-dimensional problems.

Summary of the results and prospective developments

In this part, we presented the application of the Certified Descent Algorithm to a scalar shape optimization problem. In particular, we showed the advantages of coupling gradient-based shape optimization techniques with a goal-oriented estimator of the error in the shape gradient.

We used the problem of Electrical Impedance Tomography as a proof of concept to preliminarily establish some properties of the novel introduced method. First, a guaranteed bound for the error in the shape gradient has been derived by means of a certified *a posteriori* estimator. Then, coupling this information with the 2-mesh shape optimization strategy discussed in [28] allowed to formulate the novel Certified Descent Algorithm. Owing to the error estimator, a certified descent direction is identified at each iteration thus leading to a monotonically decreasing evolution of the objective functional. Moreover, we derived a reliable stopping criterion for the overall optimization strategy by exploiting the quantitative information carried by the goal-oriented estimator. Thus, the resulting CDA is a fully automatic procedure for certified shape optimization.

Two strategies to compute the error due to the numerical approximation of the shape gradient have been investigated. Starting from the complementary energy principle, we introduced an additional variational problem for each state and adjoint problem in order to retrieve an accurate approximation of the numerical fluxes and consequently a more precise estimate of the error in the energy-norm. The major drawback of the described procedure is the necessity of solving the dual flux problems to derive a fully computable upper bound of the error in the shape gradient. As a matter of fact, though the CDA is able to make coarse meshes reliable to identify a genuine descent direction for the objective functional during the initial iterations, the overall computational cost tends to remain high and this version of the Certified Descent Algorithm may result in higher computing times than the Boundary Variation Algorithm applied on fine meshes.

In order to circumvent the issue of solving additional global variational problems to construct the estimator via the complementary energy principle, a strategy based solely on the computation of local quantities has been investigated. In particular, we proposed an improved version of the CDA by deriving a goal-oriented estimator of the error in the shape gradient via the construction of equilibrated fluxes. This approach has been developed in a unified framework, valid for both conforming Finite Element and Discontinuous Galerkin discretizations. On the one hand, using a conforming Finite Element discretization, the number of Degrees of Freedom required by the approximation of the state and adjoint problems is small. Nevertheless, the construction of the equilibrated fluxes requires the solution of local subproblems defined on patches of elements, one for each vertex in the triangulation. On the other hand, though the Discontinuous Galerkin formulation of the problems features a higher

number of Degrees of Freedom per mesh element, the computation of the error estimator based on the equilibrated fluxes approach is straightforward via a post-processing procedure which involves solely local quantities. Both strategies proved to be valid but the bounds provided by the Discontinuous Galerkin approach appear more precise and computationally less expensive.

A validation of both approaches to the CDA has been presented by showing several test cases for the well-known inverse identification problem of Electrical Impedance Tomography. The discussed results highlight the importance of accounting for the error due to the discretization of the shape gradient in order to automatize the shape optimization procedure. The application of the Certified Descent Algorithm to the problem of Electrical Impedance Tomography provided preliminary information on the strategy and the promising numerical results led to begin the investigation of the vectorial problem of optimal design of elastic structures.

Several developments may be carried on within the field of Electrical Impedance Tomography. First of all, an improvement of the reconstruction should be sought. This may be achieved by coupling the electrical information with an acoustic (cf. [32]) or an elastic wave (cf. [33]) and by deriving the corresponding shape optimization problem. Other improvements may follow from the construction of an alternative objective functional to better exploit the information associated with multiple boundary measurements. Concerning the certification procedure via the construction of *a posteriori* error estimators, some interesting developments may focus on the comparison of the equilibrated fluxes framework with the flux-free approach described in [223, 225]. Moreover, accounting for anisotropic mesh adaptation as suggested in [135, 136] may lead to discretizations with a lower number of Degrees of Freedom and a better approximation of the physical problem under analysis.

Part III

Shape optimization: a forward problem in the design of elastic structures

Chapter 6

Minimization of the compliance under a volume constraint

We consider the problem of the optimal design of a compliant structure under a volume constraint, within the framework of linear elasticity. We introduce the pure displacement formulation of the linear elasticity problem and we compute the volumetric expression of the shape gradient of the compliance by means of the velocity method. In a similar fashion, we consider a dual mixed formulation of the elasticity problem and we derive the corresponding expression of the shape gradient of the compliance. A preliminary qualitative comparison of the two expressions of the shape gradient is performed through some numerical simulations using the Boundary Variation Algorithm.

6.1 The linear elasticity problem

In this part, we consider the vectorial problem of shape optimization in the framework of linear elasticity. In particular, we are interested in the optimal design of compliant structures, that is the construction of the shape that minimizes the compliance under a volume constraint. In order to tackle this problem, in this section we introduce the governing equations that describe the mechanical behavior of a domain Ω within the infinitesimal strain theory, that is under the assumption of small deformations and small displacements. For a complete introduction to this subject, we refer the interested reader to [112, 148, 196].

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be an open and connected domain representing the body under analysis. We provide a framework to describe the motion of Ω and to compute the corresponding deformation and stress field under the effect of given volume force f and surface force g . The full set of equations consists of two conservation laws - i.e. the conservation of mass and the balance of momentum - and additional material laws that describe the relationship among the variables at play and depend on the type of solid under analysis.

In this thesis we restrict to the case of elastostatics. Under this assumption, the conservation of mass reads as

$$\nabla \cdot (\rho u_\Omega) = 0 \quad , \quad \text{in } \Omega \quad (6.1)$$

where ρ is the mass density and u_Ω is the displacement due to the external forces. Moreover, by

setting the value of ρ equal to the initial value of the density, the material is said to be incompressible and (6.1) reduces to

$$\nabla \cdot u_\Omega = 0 \quad , \quad \text{in } \Omega. \quad (6.2)$$

From the balance of momentum, we derive the following relationship between the stress tensor σ_Ω and the volume forces f , also known as Cauchy's equation of motion or equilibrium equation:

$$-\nabla \cdot \sigma_\Omega = f \quad , \quad \text{in } \Omega. \quad (6.3)$$

Moreover, the balance of angular momentum implies the symmetry of the stress tensor, that is

$$\sigma_\Omega = \sigma_\Omega^T \quad , \quad \text{in } \Omega. \quad (6.4)$$

Whereas the balance of momentum (6.3)-(6.4) provides a relationship between the stress tensor and the volume force, the material laws relate the stress tensor to the strain tensor. We define the Green-St.Venant strain tensor as a measure of the local discrepancy of a given displacement u_Ω from a rigid-body displacement:

$$E(u_\Omega) := \frac{1}{2} (\nabla u_\Omega + \nabla u_\Omega^T + \nabla u_\Omega^T \nabla u_\Omega). \quad (6.5)$$

We remark that tensor (6.5) is non-linear with respect to the displacement u_Ω and a detailed analysis of its properties may be found in [196]. For the purpose of this thesis, we restrict ourselves to the case of linear elasticity which is based on two fundamental approximations. First, under the assumption of small deformations the Green-St.Venant strain tensor may be approximated by the linearized strain tensor $e(u_\Omega)$, that is

$$E(u_\Omega) \approx e(u_\Omega) := \frac{1}{2} (\nabla u_\Omega + \nabla u_\Omega^T). \quad (6.6)$$

Moreover, we consider a so-called linear elastic material and we prescribe a material law - known as Hooke's law - that establishes a linear dependency between the stress and the strain tensors:

$$\sigma_\Omega = A e(u_\Omega) \quad , \quad \text{in } \Omega \quad (6.7)$$

where $A = A(x)$ is a fourth-order tensor known as elasticity tensor. A material is said to be homogeneous if the elasticity tensor does not depend on x . Moreover, if A is independent of the direction of the main strains - that is, from the eigenvectors of the strain tensor $e(u_\Omega)$ - the material is isotropic. It is known (cf. e.g. [196]) that the mechanical properties of a linear elastic homogeneous and isotropic material are determined by a pair of constant values - namely (E, ν) -, E being the Young's modulus and ν the Poisson's ratio or the first and second Lamé constants (λ, μ) . Thus, within the range $E > 0$ and $\nu \in (0, \frac{1}{2})$ of physically admissible values of the constants, the relationship between stress tensor and strain tensor reads as follows

$$\sigma_\Omega = A e(u_\Omega) = \frac{E}{1 + \nu} \left(e(u_\Omega) + \frac{\nu}{1 - 2\nu} \text{tr}(e(u_\Omega)) \text{Id} \right) \quad (6.8)$$

where $\text{tr}(\cdot) := \cdot : \text{Id}$ is the trace operator and $:$ is the Frobenius product. Let us introduce the first and second Lamé constants:

$$\lambda := \frac{E\nu}{(1 + \nu)(1 - 2\nu)} \quad , \quad \mu := \frac{E}{2(1 + \nu)}. \quad (6.9)$$

Owing to (6.9), (6.8) may be rewritten as follows

$$\sigma_\Omega = Ae(u_\Omega) = 2\mu e(u_\Omega) + \lambda \operatorname{tr}(e(u_\Omega)) \operatorname{Id}. \quad (6.10)$$

We remark that the elasticity tensor exists and is invertible as long as $\nu < \frac{1}{2}$ or equivalently $\lambda < \infty$. Within this framework, we may introduce the so-called compliance tensor A^{-1} whose application to the stress tensor provides the strain tensor:

$$e(u_\Omega) = A^{-1}\sigma_\Omega = \frac{1}{2\mu} \operatorname{dev}(\sigma_\Omega) + \frac{1}{d(d\lambda + 2\mu)} \operatorname{tr}(\sigma_\Omega) \operatorname{Id} = \frac{1}{2\mu} \sigma_\Omega - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \operatorname{tr}(\sigma_\Omega) \operatorname{Id} \quad (6.11)$$

where $\operatorname{dev}(\cdot) := \cdot - \frac{1}{d} \operatorname{tr}(\cdot) \operatorname{Id}$ is the deviatoric part of the tensor under analysis.

Remark 6.1. Owing to $\operatorname{tr}(e(u_\Omega)) = \nabla \cdot u_\Omega$, it is straightforward to observe that when $\lambda \rightarrow \infty$ - or equivalently, $\nu = \frac{1}{2}$ - the divergence of the displacement field in (6.10) has to vanish, that is the material under analysis is incompressible (cf. condition (6.2)). Within this framework, the material is said to be nearly incompressible, the elasticity tensor does not exist and the compliance tensor is singular.

By combining the above information we may derive the governing equations for the problem of linear elasticity set in an open connected domain Ω . Let $\partial\Omega = \Gamma^N \cup \Gamma \cup \Gamma^D$ be such that the three parts of the boundary are disjoint and Γ^D has positive $(d-1)$ -dimensional Hausdorff measure. We denote by \mathbb{S}_d the space of $d \times d$ symmetric matrices and we consider the fourth-order tensor $A : \Omega \rightarrow \mathbb{S}_d$ modeling the Hooke's law (6.10) for a linear elastic material. We describe an elastic structure subject to a volume force f , a load g on the surface Γ^N , a free-boundary condition on Γ (Neumann boundary conditions on the surface traction) and clamped on Γ^D (Dirichlet boundary condition on the displacement):

$$\begin{cases} -\nabla \cdot \sigma_\Omega = f & \text{in } \Omega \\ \sigma_\Omega = Ae(u_\Omega) & \text{in } \Omega \\ \sigma_\Omega n = g & \text{on } \Gamma^N \\ \sigma_\Omega n = 0 & \text{on } \Gamma \\ u_\Omega = 0 & \text{on } \Gamma^D \end{cases} \quad (6.12)$$

Remark 6.2. The requirement of a non-trivial boundary subset Γ^D may be explained by intuition owing to the fact that the solution of the linear elasticity problem - i.e. the displacement and the stress fields - can be uniquely determined only if the domain is fixed along at least a part of its boundary, that is no rigid motion is allowed. On the contrary, if the whole boundary features Neumann conditions, the solution can be determined up to rigid-body motions that do not affect the computation of stress and strain tensors.

6.1.1 The pure displacement variational formulation

A classical formulation of the linear elasticity problem is the so-called pure displacement formulation in which we express the stress tensor σ_Ω in terms of u_Ω using (6.10) and we seek the displacement field within the Sobolev space $H^1(\Omega; \mathbb{R}^d)$. Let $f \in H^1(\mathbb{R}^d; \mathbb{R}^d)$ and $g \in H^2(\mathbb{R}^d; \mathbb{R}^d)$. We define the following space V_Ω

$$V_\Omega := H_{0,\Gamma^D}^1(\Omega; \mathbb{R}^d) = \{v \in H^1(\Omega; \mathbb{R}^d) : v = 0 \text{ on } \Gamma^D\} \quad (6.13)$$

and we seek a function $u_\Omega \in V_\Omega$ such that

$$a_\Omega(u_\Omega, \delta u) = F_\Omega(\delta u) \quad \forall \delta u \in V_\Omega \quad (6.14)$$

where the bilinear form $a_\Omega(\cdot, \cdot) : V_\Omega \times V_\Omega \rightarrow \mathbb{R}$ and the linear form $F_\Omega(\cdot) : V_\Omega \rightarrow \mathbb{R}$ read as follows

$$a_\Omega(u_\Omega, \delta u) := \int_\Omega A e(u_\Omega) : e(\delta u) \, dx \quad , \quad F_\Omega(\delta u) := \int_\Omega f \cdot \delta u \, dx + \int_{\Gamma^N} g \cdot \delta u \, ds. \quad (6.15)$$

Existence and uniqueness of the solution of problem (6.14) follow from the classical Lax-Milgram theorem. Within this framework, the crucial point is represented by the proof of the coercivity of the bilinear form on the left-hand side of (6.14), which relies on the use of Korn's inequality (cf. [168]).

Remark 6.3. The elasticity tensor A acts as a coefficient in the pure displacement formulation (6.14) of the linear elasticity problem. As previously stated, A deteriorates approaching condition (6.2) and does not exist in the incompressible limit (cf. remark 6.1). Hence, stability issues may arise in the nearly incompressible case, whereas in the incompressible limit the stress tensor cannot be expressed in terms of the displacement field and the pure displacement variational formulation cannot be posed.

Excluding the nearly incompressible case and the incompressible limit, the primal variational formulation (6.14) provides a good description of the mechanical phenomena involved in the linear elasticity problem. The corresponding numerical approximation has been extensively studied in the literature. A straightforward approach relies on the use of Lagrangian Finite Element functions to approximate the displacement field: within this framework, optimal convergence rate of the discretized solution to the continuous one is obtained and we refer to [76, 80] for a detailed discussion on the subject.

6.1.2 Mixed variational formulations via the Hellinger-Reissner principle

Besides the aforementioned stability issues, a major drawback of the pure displacement variational formulation is the indirect evaluation of the stress tensor which is not computed as part of the solution of the linear elasticity problem but may only be derived from (6.10) via a post-processing of the displacement field u_Ω .

A possible workaround for both these issues is represented by mixed variational formulations in which the target solution is the pair $(\sigma_\Omega, u_\Omega)$ representing respectively the stress and displacement fields. This family of approaches was first proposed by Reissner in his seminal work [234]. Since the aforementioned paper in 1950, mixed variational formulations of the linear elasticity problem have known a great success in the scientific community and it is now common practice to summarize these methods as Hellinger-Reissner methods. In this subsection we briefly recall the primal and dual mixed variational formulations of the linear elasticity problem as long as some basic results on the existence and uniqueness of the solution for these models. We refer to [35] for additional information on this subject whereas a detailed introduction to mixed Finite Element methods may be found in [70].

To state the primal mixed formulation, we consider the space $V_\Omega := H_{0,\Gamma^D}^1(\Omega; \mathbb{R}^d)$ introduced in (6.13) and we define the spaces $\Sigma_\Omega = \Sigma_{\Omega,0} := L^2(\Omega; \mathbb{S}_d)$. We seek $(\sigma_\Omega, u_\Omega) \in \Sigma_\Omega \times V_\Omega$ such that

$$\begin{aligned} a_\Omega(\sigma_\Omega, \delta \sigma) + b_\Omega(\delta \sigma, u_\Omega) &= 0 & \forall \delta \sigma \in \Sigma_{\Omega,0} \\ b_\Omega(\sigma_\Omega, \delta u) &= F_\Omega(\delta u) & \forall \delta u \in V_\Omega \end{aligned} \quad (6.16)$$

where the bilinear forms $a_\Omega(\cdot, \cdot) : \Sigma_\Omega \times \Sigma_\Omega \rightarrow \mathbb{R}$ and $b_\Omega(\cdot, \cdot) : \Sigma_\Omega \times V_\Omega \rightarrow \mathbb{R}$ and the linear form $F_\Omega(\cdot) : V_\Omega \rightarrow \mathbb{R}$ read as

$$a_\Omega(\sigma_\Omega, \delta\sigma) := \int_\Omega A^{-1}\sigma_\Omega : \delta\sigma \, dx \quad , \quad b_\Omega(\sigma_\Omega, \delta u) := - \int_\Omega \sigma_\Omega : e(\delta u) \, dx, \quad (6.17)$$

$$F_\Omega(\delta u) := - \int_\Omega f \cdot \delta u \, dx - \int_{\Gamma^N} g \cdot \delta u \, ds. \quad (6.18)$$

Existence and uniqueness of the solution of (6.16) may be proved following the framework introduced by Brezzi in [81].

Remark 6.4. This formulation is equivalent to the pure displacement formulation presented in subsection 6.1.1. In particular, as for the pure displacement formulation, the primal mixed formulation is known to suffer from instability issues when dealing with highly anisotropic domains, e.g. plates and shells and in the nearly incompressible case. From a numerical point of view, in these situations the discussed approach experiences respectively shear locking and volume locking phenomena. In both cases, the condition number of the resulting linear system deteriorates and the quality of the final discretized solution is extremely poor.

Let us now introduce the functional spaces for the dual mixed variational formulation. We recall the definition of the space $H(\text{div}, \Omega; \mathbb{S}_d) := \{\tau \in L^2(\Omega; \mathbb{S}_d) : \nabla \cdot \tau \in L^2(\Omega; \mathbb{R}^d)\}$ of the symmetric square-integrable tensors whose row-wise divergence is square-integrable. Thus, we define the spaces $V_\Omega := L^2(\Omega; \mathbb{R}^d)$, $\Sigma_\Omega := \{\tau \in H(\text{div}, \Omega; \mathbb{S}_d) : \tau n = g \text{ on } \Gamma^N \text{ and } \tau n = 0 \text{ on } \Gamma\}$ and $\Sigma_{\Omega,0} := \{\tau \in H(\text{div}, \Omega; \mathbb{S}_d) : \tau n = 0 \text{ on } \Gamma^N \cup \Gamma\}$ and we seek $(\sigma_\Omega, u_\Omega) \in \Sigma_\Omega \times V_\Omega$ such that (6.16) stands with the following bilinear and linear forms:

$$a_\Omega(\sigma_\Omega, \delta\sigma) := \int_\Omega A^{-1}\sigma_\Omega : \delta\sigma \, dx \quad , \quad b_\Omega(\sigma_\Omega, \delta u) := \int_\Omega (\nabla \cdot \sigma_\Omega) \cdot \delta u \, dx, \quad (6.19)$$

$$F_\Omega(\delta u) := - \int_\Omega f \cdot \delta u \, dx. \quad (6.20)$$

Existence and uniqueness of the solution of the dual mixed variational formulation of the linear elasticity problem follow from Brezzi's theory on mixed methods and we refer to [70, 81] for a detailed presentation of the subject. It is important to highlight that in [38] the authors proved that stability estimates for the dual mixed variational formulation do not deteriorate, be it in the case of nearly incompressible materials or in the incompressible limit, making this approach feasible for the whole range of values of the Lamé constants.

Remark 6.5. A major drawback of the previously introduced dual mixed variational formulation lies in the difficulty of constructing a pair of Finite Element spaces that fulfill the requirements of Brezzi's theory in order to guarantee the stability of the method. Several authors have been dealing with this issue in the last forty years. In [42], Arnold and Winther proposed the first stable pair of Finite Element spaces for the discretization of the linear elasticity problem in two space dimensions. The corresponding three-dimensional case was later discussed in [8, 36]. Owing to the large number of Degrees of Freedom and to the high order of the involved polynomials, the construction of the basis functions described in the aforementioned works and their implementation in existing Finite Element libraries is extremely complex. To the best of our knowledge, no library currently offers fully

supported Arnold-Winther elements and their diffusion in the computational mechanics community is still limited. We refer to section 7.3 for some additional remarks on this family of Finite Element functions.

6.1.3 A dual mixed variational formulation with weakly enforced symmetry of the stress tensor

As stated in the previous subsection, the stress tensor is sought in a subspace of $H(\operatorname{div}, \Omega; \mathbb{S}_d)$. In [70], the authors highlight that the choice of this space is strictly connected with the will of strongly imposing conservation laws. In particular, σ_Ω belonging to the space of square-integrable tensors whose row-wise divergence is square-integrable strongly enforces the conservation of momentum. Moreover, the symmetry of the stress tensor is a simplified way of expressing the conservation of angular momentum for the system under analysis. It is well-known that imposing exactly a conservation law is not trivial. Hence, strongly enforcing a second conservation law by requiring the stress tensor to be symmetric is likely to be difficult.

In order to circumvent this issue and before the work [42] by Arnold and Winther appeared, several alternative formulations have been proposed in the literature to weakly enforce the symmetry of the stress tensor via a Lagrange multiplier. Starting from the pioneering work of Brezzi [81] and Fraeijis de Veubeke [138], several authors have proposed mixed formulations in which the symmetry of the stress tensor is either weakly enforced or dropped (cf. e.g. [30, 39, 258]). One of the simplest solutions was developed by Arnold, Brezzi and Douglas Jr. in [37] via the so-called PEERS element: within this framework, the stress tensor is discretized by means of an augmented cartesian product of the Raviart-Thomas Finite Element space, the displacement field using piecewise constant functions and the Lagrange multiplier via a \mathbb{P}^1 Finite Element function. Stemming from the idea of the PEERS element, several other approaches have been proposed in the literature, e.g. [82, 134, 204, 259, 260, 261]. For a complete discussion on this topic, we refer to [69].

In this subsection, we rely on a more recent mixed Finite Element method to approximate the problem of linear elasticity with weakly imposed symmetry of the stress tensor. In particular, we refer to [40] for the construction of the stable pair of Finite Element spaces in two space dimensions, whereas the corresponding three-dimensional case is treated in [41]. The choice of this new approach by Arnold and co-workers, instead of the widely used PEERS, is mainly due to the simpler discretization arising from the novel method and to the possibility of extending it to the three-dimensional case in a straightforward way.

Let \mathbb{M}_d be the space of $d \times d$ matrices and \mathbb{K}_d be the space of $d \times d$ skew-symmetric matrices. We define the spaces $V_\Omega := L^2(\Omega; \mathbb{R}^d)$, $Q_\Omega := L^2(\Omega; \mathbb{K}_d)$, $\Sigma_\Omega := \{\tau \in H(\operatorname{div}, \Omega; \mathbb{M}_d) : \tau n = g \text{ on } \Gamma^N \text{ and } \tau n = 0 \text{ on } \Gamma\}$ and $\Sigma_{\Omega,0} := \{\tau \in H(\operatorname{div}, \Omega; \mathbb{M}_d) : \tau n = 0 \text{ on } \Gamma^N \cup \Gamma\}$. Moreover, we introduce the space $W_\Omega := V_\Omega \times Q_\Omega$. The extended system obtained from (6.16) by relaxing the symmetry condition on the stress tensor through the introduction of a Lagrange multiplier reads as follows: we seek $(\sigma_\Omega, (u_\Omega, \eta_\Omega)) \in \Sigma_\Omega \times W_\Omega$ such that

$$\begin{aligned} a_\Omega(\sigma_\Omega, \delta\sigma) + b_\Omega(\delta\sigma, (u_\Omega, \eta_\Omega)) &= 0 & \forall \delta\sigma \in \Sigma_{\Omega,0} \\ b_\Omega(\sigma_\Omega, (\delta u, \delta\eta)) &= F_\Omega(\delta u) & \forall (\delta u, \delta\eta) \in W_\Omega \end{aligned} \tag{6.21}$$

where the bilinear and linear forms have the following expressions:

$$a_\Omega(\sigma_\Omega, \delta\sigma) := \int_\Omega A^{-1}\sigma_\Omega : \delta\sigma \, dx \quad , \quad b_\Omega(\sigma_\Omega, (\delta u, \delta\eta)) := \int_\Omega (\nabla \cdot \sigma_\Omega) \cdot \delta u \, dx + \frac{1}{2\mu} \int_\Omega \sigma_\Omega : \delta\eta \, dx, \quad (6.22)$$

$$F_\Omega(\delta u) := - \int_\Omega f \cdot \delta u \, dx. \quad (6.23)$$

Existence and uniqueness of the solution for this variant of the dual mixed variational formulation of the linear elasticity problem with weakly imposed symmetry of the stress tensor follow again from Brezzi's theory (cf. [37]).

Remark 6.6. We highlight that if $(\sigma_\Omega, (u_\Omega, \eta_\Omega))$ is solution of (6.21), then σ_Ω is symmetric and $(\sigma_\Omega, u_\Omega) \in H(\text{div}, \Omega; \mathbb{S}_d) \times L^2(\Omega; \mathbb{R}^d)$ is solution of the original dual mixed formulation of the linear elasticity problem with strongly enforced symmetry of the stress tensor discussed in the previous subsection. Though the infinite-dimensional formulation of the problem featuring weak symmetry is equivalent to the one in which the symmetry of the stress tensor is imposed in a strong way, the former allows for novel discretization techniques in which the approximation σ_Ω^h of the stress tensor σ_Ω is not guaranteed to be symmetric, that is σ_Ω^h solely fulfills the following condition

$$\int_\Omega \sigma_\Omega^h : \delta\eta^h \, dx = 0 \quad \forall \delta\eta^h \in Q_\Omega^h$$

where Q_Ω^h is an appropriate discrete space approximating $L^2(\Omega; \mathbb{K}_d)$.

As stated at the beginning of this subsection, several choices are possible for the discrete spaces Σ_Ω^h , V_Ω^h and Q_Ω^h respectively approximating $H(\text{div}, \Omega; \mathbb{S}_d)$, $L^2(\Omega; \mathbb{R}^d)$ and $L^2(\Omega; \mathbb{K}_d)$. In the rest of this thesis, we consider the approach discussed in [40], in which the stress tensor is approximated by the cartesian product of two pairs of Brezzi-Douglas-Marini Finite Element spaces while the displacement field and the Lagrange multiplier are both discretized using piecewise constant functions.

6.2 Minimization of the compliance under a volume constraint

Let us consider a vector field $\theta \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d)$. We introduce a transformation $X_\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and we define the open subset $\Omega_\theta \subset \mathbb{R}^d$ as $\Omega_\theta = X_\theta(\Omega)$. Moreover, we set that $\Gamma_\theta^N = X_\theta(\Gamma^N)$, $\Gamma_\theta = X_\theta(\Gamma)$ and $\Gamma_\theta^D = X_\theta(\Gamma^D)$. The displacement of an initial point $x \in \Omega$ is governed by the following differential equation:

$$\begin{cases} \frac{dx_\theta}{dt}(t) = \theta(x_\theta(t)) \\ x_\theta(0) = x \end{cases} \quad (6.24)$$

which admits a unique solution $t \mapsto x_\theta(t, x)$ in $C^1(\mathbb{R}; \mathbb{R}^d)$. Owing to (6.24), the initial point $x \in \Omega$ is transported by the field θ to the point $x_\theta = X_\theta(x)$ which belongs to the deformed domain Ω_θ . Moreover, we denote by D_θ the Jacobian matrix of the transformation X_θ and by $I_\theta = \det D_\theta$ its determinant.

Within the framework of shape optimization, a common choice for the transformation X_θ is a perturbation of the identity map, that is

$$X_\theta = \text{Id} + \theta + o(\theta) \quad , \quad \theta \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d). \quad (6.25)$$

Hence, $\Omega_\theta = X_\theta(\Omega) = \{x + \theta(x) : x \in \Omega\}$ and under the assumption of a small perturbation θ , X_θ is a diffeomorphism and belongs to the following space (cf. [17]):

$$\mathcal{X} := \left\{ X_\theta : (X_\theta - \text{Id}) \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d) \text{ and } (X_\theta^{-1} - \text{Id}) \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d) \right\}.$$

By exploiting the notation above, we introduce the set of shapes that may be obtained as result of a deformation of the reference domain Ω :

$$\mathcal{U}_{\text{def}} := \{\Omega_\theta : \exists X_\theta \in \mathcal{X}, \Omega_\theta = X_\theta(\Omega)\}. \quad (6.26)$$

As stated at the beginning of this chapter, we are interested in the optimal design of compliant structures in the framework of linear elasticity. In this section, we introduce the formulation of the aforementioned shape optimization problem as a PDE-constrained optimization problem of a shape-dependent functional. Let us define the compliance on a deformed domain Ω_θ as

$$J(\Omega_\theta) = \int_{\Omega_\theta} A^{-1} \sigma_{\Omega_\theta} : \sigma_{\Omega_\theta} \, dx_\theta. \quad (6.27)$$

Following the framework introduced in section 1.1, the resulting shape optimization problem reads as (1.5) where the objective functional is the compliance (6.27) and we seek a domain $\Omega_\theta \in \mathcal{U}_{\text{ad}}$ which is the set of shapes in (6.26) such that σ_{Ω_θ} is the stress tensor fulfilling the linear elasticity problem (6.12) on Ω_θ .

Remark 6.7. In the pure displacement formulation (cf. subsection 6.1.1), the stress tensor can be expressed in terms of the displacement field through the relationship $\sigma_{\Omega_\theta} = Ae(u_{\Omega_\theta})$. Hence, (6.27) may be rewritten as

$$J(\Omega_\theta) = \int_{\Omega_\theta} Ae(u_{\Omega_\theta}) : e(u_{\Omega_\theta}) \, dx_\theta = \int_{\Omega_\theta} f \cdot u_{\Omega_\theta} \, dx_\theta + \int_{\Gamma_\theta^N} g \cdot u_{\Omega_\theta} \, ds_\theta, \quad (6.28)$$

that is we can equivalently reinterpret the compliance as the work of the external forces applied to the domain Ω_θ .

In real-life problems, the optimal design of compliant structures is usually subject to additional constraints, either imposed by the end-user (e.g. volume/perimeter [17] or stress [126] constraints) or by the manufacturing process (e.g. maximum/minimum thickness [24] or molding direction [23] constraints). Several sophisticated strategies (e.g. quadratic penalty and augmented Lagrangian methods) may be considered to handle the constraints involved in optimization problems and we refer to [212] for a thorough introduction to this subject. Within the field of shape optimization, an algorithm based on a Lagrangian functional featuring an efficient update strategy for the Lagrange multiplier has been proposed in [28]. Several other approaches have known a great success in the literature, e.g. the Method of Moving Asymptotes [265] and the Method of Feasible Directions [271]. In this thesis, we restrict ourselves to the classical volume constraint and we enforce it through a penalty method using a fixed Lagrange multiplier γ . Thus the resulting unconstrained shape optimization problem reads as follows:

$$\min_{\Omega_\theta \in \mathcal{U}_{\text{ad}}} L(\Omega_\theta) \quad , \quad L(\Omega_\theta) := J(\Omega_\theta) + \gamma V(\Omega_\theta) \quad (6.29)$$

where $J(\Omega_\theta)$ is the compliance (6.27), $V(\Omega_\theta) := |\Omega_\theta|$ is the volume of the domain and \mathcal{U}_{ad} is the previously defined set of admissible shapes.

In order to construct a gradient-based strategy to solve (6.29) as discussed in section 2.1, the analytical expression of the shape gradient of $L(\Omega_\theta)$ is required. In particular, owing to the better numerical accuracy that volumetric formulations of the shape gradient have shown with respect to the surface ones for elliptic problems (cf. [166]) and to our interest in *a posteriori* error estimators, in the following sections we compute the volumetric expression of the shape gradient of $L(\Omega_\theta)$. We remark that the volume $V(\Omega_\theta)$ only depends on the geometry of the domain and its shape gradient has the following expression (cf. e.g [118]):

$$\langle dV(\Omega), \theta \rangle = \int_{\Omega} \nabla \cdot \theta \, dx. \quad (6.30)$$

Hence, in sections 6.3 and 6.4 we will only derive the volumetric expression of the shape gradient of the compliance (6.27), respectively for the pure displacement formulation (cf. subsection 6.1.1) and for the mixed formulations (cf. subsections 6.1.2 and 6.1.3) of the linear elasticity problem.

6.3 Volumetric shape gradient of the compliance via the pure displacement formulation

In this section we derive the shape gradient of the compliance using the pure displacement formulation of the linear elasticity problem. Owing to the principle of minimum potential energy for the problem (6.14)-(6.15) on the domain Ω_θ and to (6.28), we may write the compliance as follows:

$$J_1(\Omega_\theta) := - \min_{u_{\Omega_\theta} \in V_{\Omega_\theta}} \int_{\Omega_\theta} Ae(u_{\Omega_\theta}) : e(u_{\Omega_\theta}) \, dx_\theta - 2 \int_{\Omega_\theta} f \cdot u_{\Omega_\theta} \, dx_\theta - 2 \int_{\Gamma_\theta^N} g \cdot u_{\Omega_\theta} \, ds_\theta, \quad (6.31)$$

where $V_{\Omega_\theta} := H_{0,\Gamma_\theta^D}^1(\Omega_\theta; \mathbb{R}^d) = \{v \in H^1(\Omega_\theta; \mathbb{R}^d) : v = 0 \text{ on } \Gamma_\theta^D\}$.

Let $j_1(\theta) := J_1(\Omega_\theta)$. We are interested in computing the shape gradient of the functional $J_1(\Omega)$, that is

$$\langle dJ_1(\Omega), \theta \rangle := \lim_{\theta \searrow 0} \frac{J_1(\Omega_\theta) - J_1(\Omega)}{\theta} = \lim_{\theta \searrow 0} \frac{j_1(\theta) - j_1(0)}{\theta} =: j_1'(0). \quad (6.32)$$

We refer to [118] for a result on the differentiability of a minimum with respect to a parameter. Moreover, we remark that the space V_{Ω_θ} in (6.31) depends on the parameter θ . We use the function space parameterization technique described in [118] to transport the quantities defined on the deformed domain Ω_θ back to the reference domain Ω . Thus, we are able to rewrite (6.31) using solely functions of the space V_Ω which no longer depends on θ and we apply elementary differential calculus techniques to compute the derivative of the objective functional with respect to the parameter θ .

Let us introduce the following transformation to parameterize the functions in $H_{0,\Gamma_\theta^D}^1(\Omega_\theta; \mathbb{R}^d)$ in terms of the elements of $H_{0,\Gamma^D}^1(\Omega; \mathbb{R}^d)$:

$$\mathcal{P}_\theta : H_{0,\Gamma^D}^1(\Omega; \mathbb{R}^d) \rightarrow H_{0,\Gamma_\theta^D}^1(\Omega_\theta; \mathbb{R}^d) \quad , \quad v_{\Omega_\theta} = \mathcal{P}_\theta(v_\Omega) = v_\Omega \circ X_\theta^{-1}. \quad (6.33)$$

Lemma 6.8. *Let $u_\Omega \in H_{0,\Gamma^D}^1(\Omega; \mathbb{R}^d)$. We consider $u_{\Omega_\theta} = \mathcal{P}_\theta(u_\Omega)$ according to the transformation (6.33). It follows that*

$$\frac{1}{2} (\nabla_{x_\theta} u_{\Omega_\theta} + \nabla_{x_\theta} u_{\Omega_\theta}^T) =: e_{x_\theta}(u_{\Omega_\theta}) = \frac{1}{2} (\nabla_x u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla_x u_\Omega^T) \quad (6.34)$$

where ∇_{x_θ} (respectively ∇_x) represents the gradient with respect to the coordinate of the deformed (respectively reference) domain.

Proof. Owing to (6.33), $u_{\Omega_\theta} = u_\Omega \circ X_\theta^{-1}$. Thus,

$$\frac{\partial (u_{\Omega_\theta})_i}{\partial (x_\theta)_j} = \frac{\partial (u_\Omega)_i}{\partial (x)_m} \frac{\partial (X_\theta^{-1})_m}{\partial (x_\theta)_j} = \frac{\partial (u_\Omega)_i}{\partial (x)_m} (D_\theta^{-1})_{mj}.$$

Hence, the result follows directly:

$$(e_{x_\theta}(u_{\Omega_\theta}))_{ij} = \frac{1}{2} \left(\frac{\partial (u_{\Omega_\theta})_i}{\partial (x_\theta)_j} + \frac{\partial (u_{\Omega_\theta})_j}{\partial (x_\theta)_i} \right) = \frac{1}{2} \left(\frac{\partial (u_\Omega)_i}{\partial (x)_m} (D_\theta^{-1})_{mj} + \frac{\partial (u_\Omega)_j}{\partial (x)_m} (D_\theta^{-1})_{mi} \right),$$

that is

$$e_{x_\theta}(u_{\Omega_\theta}) = \frac{1}{2} (\nabla_x u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla_x u_\Omega^T).$$

□

For the sake of readability and except in the case of ambiguity, henceforth we will omit the subscript specifying the spatial coordinate with respect to which the gradient is computed, that is with an abuse of notation we consider $\nabla u_\Omega = \nabla_x u_\Omega$ and $\nabla u_{\Omega_\theta} = \nabla_{x_\theta} u_{\Omega_\theta}$.

Before moving to the derivation of the shape gradient of (6.31), we recall the definition of the cofactor matrix which will later appear in the mapping of the surface $\partial\Omega_\theta$ to $\partial\Omega$.

Definition 6.9. The cofactor matrix $\text{Cof } C$ of a matrix $C \in \mathbb{M}_d$ is the $d \times d$ matrix formed by all its cofactors:

$$(\text{Cof } C)_{ij} := (-1)^{i+j} \det \tilde{C}_{ij}$$

where \tilde{C}_{ij} is the $(d-1) \times (d-1)$ submatrix obtained by deleting the i -th row and the j -th column of the matrix C . Using the Levi-Civita symbol ε_{ijk} and the Einstein summation convention, the cofactor matrix for the three-dimensional case reads as follows:

$$(\text{Cof } C)_{ij} := \frac{1}{2} \varepsilon_{imn} \varepsilon_{jpq} c_{pm} c_{qn} \quad (6.35)$$

where c_{pm} is the (p, m) -th element of the matrix C .

Now, we use the transformation (6.33) and the property (6.34) to map the first term in (6.31) to the reference domain Ω :

$$\begin{aligned} \int_{\Omega_\theta} A e(u_{\Omega_\theta}) : e(u_{\Omega_\theta}) \, dx_\theta &= \int_{\Omega} A e(u_{\Omega_\theta} \circ X_\theta) : e(u_{\Omega_\theta} \circ X_\theta) I_\theta \, dx \\ &= \int_{\Omega} A \left(\frac{1}{2} (\nabla u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla u_\Omega^T) \right) : \left(\frac{1}{2} (\nabla u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla u_\Omega^T) \right) I_\theta \, dx. \end{aligned} \quad (6.36)$$

The remaining terms in (6.31) may be transported to the reference domain as follows:

$$-2 \int_{\Omega_\theta} f \cdot u_{\Omega_\theta} dx_\theta = -2 \int_{\Omega} f \circ X_\theta \cdot (u_{\Omega_\theta} \circ X_\theta) I_\theta dx = -2 \int_{\Omega} f \circ X_\theta \cdot u_\Omega I_\theta dx, \quad (6.37)$$

$$-2 \int_{\Gamma_\theta^N} g \cdot u_{\Omega_\theta} ds_\theta = -2 \int_{\Gamma^N} g \circ X_\theta \cdot (u_{\Omega_\theta} \circ X_\theta) \text{Cof } D_\theta ds = -2 \int_{\Gamma^N} g \circ X_\theta \cdot u_\Omega \text{Cof } D_\theta ds. \quad (6.38)$$

By combining (6.36), (6.37) and (6.38), we obtain the following function $j_1(\theta)$ which solely depends on the reference domain Ω :

$$\begin{aligned} j_1(\theta) = & - \min_{u_\Omega \in V_\Omega} \int_{\Omega} A \left(\frac{1}{2} \left(\nabla u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla u_\Omega^T \right) \right) : \left(\frac{1}{2} \left(\nabla u_\Omega D_\theta^{-1} + D_\theta^{-T} \nabla u_\Omega^T \right) \right) I_\theta dx \\ & - 2 \int_{\Omega} f \circ X_\theta \cdot u_\Omega I_\theta dx - 2 \int_{\Gamma^N} g \circ X_\theta \cdot u_\Omega \text{Cof } D_\theta ds. \end{aligned} \quad (6.39)$$

Owing to (6.25), the Jacobian of the transformations X_θ , X_θ^T and X_θ^{-1} read as

$$D_\theta = \text{Id} + \nabla \theta + o(\nabla \theta), \quad (6.40)$$

$$D_\theta^T = \text{Id} + \nabla \theta^T + o(\nabla \theta), \quad (6.41)$$

$$D_\theta^{-1} = \text{Id} - \nabla \theta + o(\nabla \theta). \quad (6.42)$$

Moreover, we recall that

$$\det(\text{Id} + C) = 1 + \text{tr}(C) + o(C), \quad (6.43)$$

$$\text{Cof}(\text{Id} + C) = \text{Id} + \text{tr}(C) \text{Id} - C + o(C). \quad (6.44)$$

We may now differentiate (6.39) with respect to θ in $\theta = 0$ by exploiting (6.41), (6.42), (6.43) and (6.44). The shape gradient of the compliance using the pure displacement formulation for the linear elasticity problem reads as

$$\begin{aligned} \langle dJ_1(\Omega), \theta \rangle = & \int_{\Omega} A e(u_\Omega) : (\nabla u_\Omega \nabla \theta + \nabla \theta^T \nabla u_\Omega^T) dx - \int_{\Omega} A e(u_\Omega) : e(u_\Omega) (\nabla \cdot \theta) dx \\ & + 2 \int_{\Omega} (\nabla f \theta \cdot u_\Omega + f \cdot u_\Omega (\nabla \cdot \theta)) dx + 2 \int_{\Gamma^N} (\nabla g \theta \cdot u_\Omega + g \cdot u_\Omega (\nabla \cdot \theta - \nabla \theta n \cdot n)) ds. \end{aligned} \quad (6.45)$$

6.4 Volumetric shape gradient of the compliance via the dual mixed formulation

The use of mixed formulations in shape optimization was first proposed by Sigmund and Clausen in [252] and later investigated in a series of works by Bruggi and co-workers [84, 85, 86, 87]. In all the aforementioned contributions the authors use SIMP method to perform shape optimization of linear elastic structures and to the best of our knowledge the results of this section are the first attempt to derive the shape gradient of the compliance starting from a mixed formulation of the linear elasticity equation.

Let us consider the notation introduced in the previous section for the transformation X_θ . Following the same procedure as above, we may rewrite the compliance coupled with the constraint that the stress tensor is solution of the linear elasticity equation in the Hellinger-Reissner dual mixed variational formulation (6.16)-(6.19)-(6.20) on Ω_θ by introducing the following objective functional:

$$J_2(\Omega_\theta) := \inf_{\sigma_{\Omega_\theta} \in \Sigma_{\Omega_\theta}} \sup_{u_{\Omega_\theta} \in V_{\Omega_\theta}} \int_{\Omega_\theta} A^{-1} \sigma_{\Omega_\theta} : \sigma_{\Omega_\theta} \, dx_\theta + \int_{\Omega_\theta} (\nabla \cdot \sigma_{\Omega_\theta} + f) \cdot u_{\Omega_\theta} \, dx_\theta \quad (6.46)$$

where $V_{\Omega_\theta} := L^2(\Omega_\theta; \mathbb{R}^d)$ and $\Sigma_{\Omega_\theta} := \{\tau \in H(\operatorname{div}, \Omega_\theta; \mathbb{S}_d) : \tau n_\theta = g \text{ on } \Gamma_\theta^N \text{ and } \tau n = 0 \text{ on } \Gamma_\theta\}$.

In a similar fashion, starting from the dual mixed variational formulation with weakly enforced symmetry of the stress tensor (6.21)-(6.22)-(6.23), we obtain:

$$J_3(\Omega_\theta) := \inf_{\sigma_{\Omega_\theta} \in \Sigma_{\Omega_\theta}} \sup_{(u_{\Omega_\theta}, \eta_{\Omega_\theta}) \in W_{\Omega_\theta}} \int_{\Omega_\theta} A^{-1} \sigma_{\Omega_\theta} : \sigma_{\Omega_\theta} \, dx_\theta + \int_{\Omega_\theta} (\nabla \cdot \sigma_{\Omega_\theta} + f) \cdot u_{\Omega_\theta} \, dx_\theta + \frac{1}{2\mu} \int_{\Omega_\theta} \sigma_{\Omega_\theta} : \eta_{\Omega_\theta} \, dx_\theta \quad (6.47)$$

where $W_{\Omega_\theta} := V_{\Omega_\theta} \times Q_{\Omega_\theta} := L^2(\Omega_\theta; \mathbb{R}^d) \times L^2(\Omega_\theta; \mathbb{K}_d)$ and $\Sigma_{\Omega_\theta} := \{\tau \in H(\operatorname{div}, \Omega_\theta; \mathbb{M}_d) : \tau n_\theta = g \text{ on } \Gamma_\theta^N \text{ and } \tau n = 0 \text{ on } \Gamma_\theta\}$.

Let $j_i(\theta) := J_i(\Omega_\theta)$ $i = 2, 3$. We are interested in computing the shape gradient of the functionals $J_i(\Omega)$, that is

$$\langle dJ_i(\Omega), \theta \rangle := \lim_{\theta \searrow 0} \frac{J_i(\Omega_\theta) - J_i(\Omega)}{\theta} = \lim_{\theta \searrow 0} \frac{j_i(\theta) - j_i(0)}{\theta} =: j'_i(0). \quad (6.48)$$

We refer to [117] for a general result on the differentiability of a min-max function, whereas in [118, 141] some examples of shape differentiability of min-max functions are provided.

As in section 6.3, we apply the function space parameterization technique to transport the quantities defined on Ω_θ back to Ω . A key aspect in this procedure is the construction of a transformation that preserves the normal traces of the tensors in (6.46) and (6.47). For this purpose, we rely on a special isomorphism known as contravariant Piola transform and we define the following mappings:

$$\mathcal{Q}_\theta : H(\operatorname{div}, \Omega; \mathbb{M}_d) \rightarrow H(\operatorname{div}, \Omega_\theta; \mathbb{M}_d) \quad , \quad \tau_{\Omega_\theta} = \mathcal{Q}_\theta(\tau_\Omega) = \frac{1}{I_\theta} D_\theta \tau_\Omega \circ X_\theta^{-1} D_\theta^T \quad (6.49)$$

$$\mathcal{R}_\theta : L^2(\Omega; \mathbb{R}^d) \rightarrow L^2(\Omega_\theta; \mathbb{R}^d) \quad , \quad v_{\Omega_\theta} = \mathcal{R}_\theta(v_\Omega) = D_\theta^{-T} v_\Omega \circ X_\theta^{-1}. \quad (6.50)$$

We refer to [112, 196] for a discussion on the Piola transform and its role in the mathematical theory of elasticity, to [233, 268] for its application to mixed Finite Element methods for elliptic problems and to [237] for some technical details on its use to efficiently evaluate variational forms in $H(\operatorname{div})$ and $H(\operatorname{curl})$, that is the Sobolev space of square-integrable vectorfields whose rotation curl is square-integrable.

Before moving to the derivation of the shape gradient via the function space parameterization technique, we prove the following property:

Lemma 6.10. *Let $\sigma_\Omega \in H(\operatorname{div}, \Omega; \mathbb{M}_d)$. We consider $\sigma_{\Omega_\theta} = \mathcal{Q}_\theta(\sigma_\Omega)$ according to the transformation (6.49). It follows that*

$$\nabla_{x_\theta} \cdot \sigma_{\Omega_\theta} = \frac{1}{I_\theta} D_\theta \nabla_x \cdot \sigma_\Omega \quad (6.51)$$

where $\nabla_{x_\theta} \cdot$ (respectively $\nabla_x \cdot$) represents the divergence with respect to the coordinate of the deformed (respectively reference) domain.

Proof. First, we recall that for a given invertible matrix $C \in \mathbb{M}_d$, we get that

$$C^{-1} = \frac{1}{\det C} (\text{Cof } C)^T. \quad (6.52)$$

Owing to this property, we may rewrite (6.49) as

$$\sigma_{\Omega_\theta} = D_\theta \sigma_\Omega \circ X_\theta^{-1} (\text{Cof } D_\theta^{-1}). \quad (6.53)$$

We are interested in computing the divergence of (6.53) with respect to the coordinate x_θ of the deformed domain. Within this framework, we observe that being D_θ the Jacobian of the transformation (6.24) such that $\Omega_\theta \ni x_\theta = X_\theta(x)$, $x \in \Omega$, it is independent on the variable x_θ . Let us now prove the following Piola identity:

$$\nabla_{x_\theta} \cdot (\text{Cof } D_\theta^{-1}) = 0. \quad (6.54)$$

Owing to (6.35), the cofactor matrix of the inverse of the Jacobian D_θ has the form

$$(\text{Cof } D_\theta^{-1})_{ij} = \frac{1}{2} \varepsilon_{imn} \varepsilon_{jpq} \frac{\partial (X_\theta^{-1})_m}{\partial (x_\theta)_p} \frac{\partial (X_\theta^{-1})_n}{\partial (x_\theta)_q}.$$

Its divergence reads

$$\begin{aligned} \frac{\partial (\text{Cof } D_\theta^{-1})_{ij}}{\partial (x_\theta)_j} &= \frac{1}{2} \varepsilon_{imn} \varepsilon_{jpq} \left(\frac{\partial^2 (X_\theta^{-1})_m}{\partial (x_\theta)_j \partial (x_\theta)_p} \frac{\partial (X_\theta^{-1})_n}{\partial (x_\theta)_q} + \frac{\partial (X_\theta^{-1})_m}{\partial (x_\theta)_p} \frac{\partial^2 (X_\theta^{-1})_n}{\partial (x_\theta)_j \partial (x_\theta)_q} \right) \\ &= \frac{1}{2} \varepsilon_{imn} \left(\varepsilon_{pqj} \frac{\partial^2 (X_\theta^{-1})_m}{\partial (x_\theta)_p \partial (x_\theta)_j} \frac{\partial (X_\theta^{-1})_n}{\partial (x_\theta)_q} + \varepsilon_{qpj} \frac{\partial (X_\theta^{-1})_m}{\partial (x_\theta)_p} \frac{\partial^2 (X_\theta^{-1})_n}{\partial (x_\theta)_q \partial (x_\theta)_j} \right) \\ &= -\frac{1}{2} \varepsilon_{imn} \varepsilon_{jpq} \left(\frac{\partial^2 (X_\theta^{-1})_m}{\partial (x_\theta)_j \partial (x_\theta)_p} \frac{\partial (X_\theta^{-1})_n}{\partial (x_\theta)_q} + \frac{\partial (X_\theta^{-1})_m}{\partial (x_\theta)_p} \frac{\partial^2 (X_\theta^{-1})_n}{\partial (x_\theta)_j \partial (x_\theta)_q} \right) \\ &= -\frac{\partial (\text{Cof } D_\theta^{-1})_{ij}}{\partial (x_\theta)_j}, \end{aligned}$$

where the third equality follows from the definition of the Levi-Civita symbol. Hence, we can conclude that (6.54) stands. We may now compute the divergence of (6.53):

$$\begin{aligned} \nabla_{x_\theta} \cdot \sigma_{\Omega_\theta} &= \frac{\partial (\sigma_{\Omega_\theta})_{ij}}{\partial (x_\theta)_j} e_i = \frac{\partial}{\partial (x_\theta)_j} \left((D_\theta)_{im} (\sigma_\Omega \circ X_\theta^{-1})_{mq} (\text{Cof } D_\theta^{-1})_{qj} \right) e_i \\ &= (D_\theta)_{im} \frac{\partial (\sigma_\Omega)_{mn}}{\partial (x)_n} \frac{\partial (X_\theta^{-1})_n}{\partial (x_\theta)_j} (\text{Cof } D_\theta^{-1})_{qj} e_i \\ &= (D_\theta)_{im} \frac{\partial (\sigma_\Omega)_{mn}}{\partial (x)_n} (D_\theta^{-1})_{nj} (\text{Cof } D_\theta^{-1})_{qj} e_i \\ &= \frac{1}{\det D_\theta} (D_\theta)_{im} \frac{\partial (\sigma_\Omega)_{mn}}{\partial (x)_n} \delta_{nq} e_i \end{aligned}$$

where the last equality follows from (6.52). Hence, it is straightforward to retrieve the result (6.51):

$$\nabla_{x_\theta} \cdot \sigma_{\Omega_\theta} = \frac{1}{I_\theta} (D_\theta)_{im} \frac{\partial (\sigma_\Omega)_{mq}}{\partial (x)_q} e_i = \frac{1}{I_\theta} D_\theta \nabla_x \cdot \sigma_\Omega$$

□

From now on, if there is no ambiguity we will assume that the differential operators act on the space to which the functions belong and we will omit the subscript associated with the spatial coordinate used to compute the derivatives (e.g. $\nabla \cdot \sigma_{\Omega_\theta} = \nabla_{x_\theta} \cdot \sigma_{\Omega_\theta}$ and $\nabla \cdot \sigma_\Omega = \nabla_x \cdot \sigma_\Omega$).

As stated at the beginning of this section, in order to compute the shape gradients (6.48), we have to express the functionals $J_2(\Omega_\theta)$ and $J_3(\Omega_\theta)$ in terms of the reference domain Ω and of functions defined solely on it. Thus, in the following subsections we use the transformations (6.49) and (6.50) to map (6.46) and (6.47) back to the reference domain and differentiate them with respect to θ .

6.4.1 The case of strongly enforced symmetry of the stress tensor

We consider the Hellinger-Reissner mixed variational formulation of the linear elasticity problem and the corresponding objective functional (6.46). We remark that the symmetry of the stress tensor σ_{Ω_θ} is strongly enforced using the space $\Sigma_{\Omega_\theta} := \{\tau \in H(\text{div}, \Omega_\theta; \mathbb{S}_d) : \tau n_\theta = g \text{ on } \Gamma_\theta^N \text{ and } \tau n = 0 \text{ on } \Gamma_\theta\}$. It is straightforward to observe that the transformation (6.49) holds true for the space of $d \times d$ symmetric matrices \mathbb{S}_d , that is $\mathcal{Q}_\theta : H(\text{div}, \Omega; \mathbb{S}_d) \rightarrow H(\text{div}, \Omega_\theta; \mathbb{S}_d)$. As a matter of fact, being $\tau_\Omega \in H(\text{div}, \Omega; \mathbb{S}_d)$, it follows that

$$(\tau_{\Omega_\theta})^T = \left(\frac{1}{I_\theta} D_\theta \tau_\Omega \circ X_\theta^{-1} D_\theta^T \right)^T = \frac{1}{I_\theta} D_\theta \tau_\Omega \circ X_\theta^{-1} D_\theta^T = \tau_{\Omega_\theta}.$$

We use the definition of the compliance tensor in (6.11) and we map the first term in (6.46) to the reference domain Ω by means of the transformation (6.49):

$$\begin{aligned} \frac{1}{2\mu} \int_{\Omega_\theta} \sigma_{\Omega_\theta} : \sigma_{\Omega_\theta} \, dx_\theta &= \frac{1}{2\mu} \int_{\Omega} (\sigma_{\Omega_\theta} \circ X_\theta) : (\sigma_{\Omega_\theta} \circ X_\theta) I_\theta \, dx \\ &= \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta^2} (D_\theta \sigma_\Omega D_\theta^T) : (D_\theta \sigma_\Omega D_\theta^T) I_\theta \, dx \\ &= \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta} D_\theta^T D_\theta \sigma_\Omega D_\theta^T D_\theta : \sigma_\Omega \, dx, \end{aligned} \tag{6.55}$$

where the last equality follows from the definition of the Frobenius product and the cyclic property of the trace. In a similar fashion, we obtain

$$\begin{aligned} -\frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega_\theta} \text{tr}(\sigma_{\Omega_\theta}) \text{tr}(\sigma_{\Omega_\theta}) \, dx_\theta &= -\frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} \text{tr}(\sigma_{\Omega_\theta} \circ X_\theta) \text{tr}(\sigma_{\Omega_\theta} \circ X_\theta) I_\theta \, dx \\ &= -\frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} \frac{1}{I_\theta^2} \text{tr}(D_\theta \sigma_\Omega D_\theta^T) \text{tr}(D_\theta \sigma_\Omega D_\theta^T) I_\theta \, dx \\ &= -\frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} \frac{1}{I_\theta} \text{tr}(D_\theta^T D_\theta \sigma_\Omega) \text{tr}(D_\theta^T D_\theta \sigma_\Omega) \, dx. \end{aligned} \tag{6.56}$$

We consider now the second term in (6.46). Owing to (6.51) and (6.50) it follows

$$\begin{aligned} \int_{\Omega_\theta} (\nabla \cdot \sigma_{\Omega_\theta}) \cdot u_{\Omega_\theta} \, dx_\theta &= \int_{\Omega} (\nabla \cdot (\sigma_{\Omega_\theta} \circ X_\theta)) \cdot (u_{\Omega_\theta} \circ X_\theta) I_\theta \, dx \\ &= \int_{\Omega} \frac{1}{I_\theta} (D_\theta \nabla \cdot \sigma_\Omega) \cdot (D_\theta^{-T} u_\Omega) I_\theta \, dx = \int_{\Omega} (\nabla \cdot \sigma_\Omega) \cdot u_\Omega \, dx, \end{aligned} \quad (6.57)$$

$$\int_{\Omega_\theta} f \cdot u_{\Omega_\theta} \, dx_\theta = \int_{\Omega} f \circ X_\theta \cdot (u_{\Omega_\theta} \circ X_\theta) I_\theta \, dx = \int_{\Omega} f \circ X_\theta \cdot (D_\theta^{-T} u_\Omega) I_\theta \, dx. \quad (6.58)$$

By combining the above information, we obtain the following min-max function which no longer depends on the space Ω_θ :

$$\begin{aligned} j_2(\theta) &= \inf_{\sigma_\Omega \in \Sigma_\Omega} \sup_{u_\Omega \in V_\Omega} \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta} D_\theta^T D_\theta \sigma_\Omega D_\theta^T D_\theta : \sigma_\Omega \, dx \\ &\quad - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} \frac{1}{I_\theta} \operatorname{tr} (D_\theta^T D_\theta \sigma_\Omega) \operatorname{tr} (D_\theta^T D_\theta \sigma_\Omega) \, dx \\ &\quad + \int_{\Omega} (\nabla \cdot \sigma_\Omega) \cdot u_\Omega \, dx + \int_{\Omega} f \circ X_\theta \cdot (D_\theta^{-T} u_\Omega) I_\theta \, dx. \end{aligned} \quad (6.59)$$

We may now exploit (6.40), (6.41) and (6.43) to differentiate (6.59) with respect to θ and evaluate the resulting quantity in $\theta = 0$. Thus, the shape gradient of the compliance using the Hellinger-Reissner dual mixed variational formulation for the linear elasticity problem reads as

$$\begin{aligned} \langle dJ_2(\Omega), \theta \rangle &= \frac{1}{2\mu} \int_{\Omega} 2N(\theta) \sigma_\Omega : \sigma_\Omega \, dx - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} 2 \operatorname{tr} (N(\theta) \sigma_\Omega) \operatorname{tr} (\sigma_\Omega) \, dx \\ &\quad + \int_{\Omega} (\nabla f \theta \cdot u_\Omega + f \cdot u_\Omega (\nabla \cdot \theta) - f \cdot (\nabla \theta^T u_\Omega)) \, dx \end{aligned} \quad (6.60)$$

where $N(\theta) := \nabla \theta + \nabla \theta^T - \frac{1}{2}(\nabla \cdot \theta) \operatorname{Id}$.

6.4.2 The case of weakly enforced symmetry of the stress tensor

The dual mixed formulation of the linear elasticity problem discussed in subsection 6.1.3 is characterized by the weak imposition of the symmetry of the stress tensor through a Lagrange multiplier η_{Ω_θ} . Thus, besides the spaces V_{Ω_θ} and Σ_{Ω_θ} , the functional (6.47) associated with the minimization of the compliance using the aforementioned framework introduces the additional space $Q_{\Omega_\theta} := L^2(\Omega_\theta; \mathbb{K}_d)$ of the $d \times d$ skew-symmetric square-integrable tensors. In order to map the space $L^2(\Omega; \mathbb{K}_d)$ to $L^2(\Omega_\theta; \mathbb{K}_d)$, we use the previously introduced transformation (6.49): it is straightforward to observe that given $\eta_\Omega \in L^2(\Omega; \mathbb{K}_d)$, the transported $\eta_{\Omega_\theta} = \mathcal{Q}_\theta(\eta_\Omega)$ is skew-symmetric:

$$(\eta_{\Omega_\theta})^T = \left(\frac{1}{I_\theta} D_\theta \eta_\Omega \circ X_\theta^{-1} D_\theta^T \right)^T = \frac{1}{I_\theta} D_\theta (\eta_\Omega \circ X_\theta^{-1})^T D_\theta^T = -\frac{1}{I_\theta} D_\theta \eta_\Omega \circ X_\theta^{-1} D_\theta^T = -\eta_{\Omega_\theta}.$$

The first two integrals in (6.47) may be treated as in the previous subsection and the manipulations that lead to (6.55), (6.56), (6.57) and (6.58) stand. Let us now map the remaining term in (6.47) back

to the reference domain Ω :

$$\begin{aligned}
 \frac{1}{2\mu} \int_{\Omega_\theta} \sigma_{\Omega_\theta} : \eta_{\Omega_\theta} \, dx_\theta &= \frac{1}{2\mu} \int_{\Omega} (\sigma_{\Omega_\theta} \circ X_\theta) : (\eta_{\Omega_\theta} \circ X_\theta) I_\theta \, dx \\
 &= \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta^2} (D_\theta \sigma_\Omega D_\theta^T) : (D_\theta \eta_\Omega D_\theta^T) I_\theta \, dx \\
 &= \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta} D_\theta^T D_\theta \sigma_\Omega D_\theta^T D_\theta : \eta_\Omega \, dx.
 \end{aligned} \tag{6.61}$$

We combine (6.55), (6.56), (6.57), (6.58) and (6.61) to obtain the min-max function associated with $J_3(\Omega_\theta)$ and defined on a space that does not depend on θ :

$$\begin{aligned}
 j_3(\theta) &= \inf_{\sigma_\Omega \in \Sigma_\Omega} \sup_{(u_\Omega, \eta_\Omega) \in W_\Omega} \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta} D_\theta^T D_\theta \sigma_\Omega D_\theta^T D_\theta : \sigma_\Omega \, dx \\
 &\quad - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} \frac{1}{I_\theta} \operatorname{tr} (D_\theta^T D_\theta \sigma_\Omega) \operatorname{tr} (D_\theta^T D_\theta \sigma_\Omega) \, dx \\
 &\quad + \frac{1}{2\mu} \int_{\Omega} \frac{1}{I_\theta} D_\theta^T D_\theta \sigma_\Omega D_\theta^T D_\theta : \eta_\Omega \, dx \\
 &\quad + \int_{\Omega} (\nabla \cdot \sigma_\Omega) \cdot u_\Omega \, dx + \int_{\Omega} f \circ X_\theta \cdot (D_\theta^{-T} u_\Omega) I_\theta \, dx.
 \end{aligned} \tag{6.62}$$

Let us consider the matrix $N(\theta)$ introduced in the previous subsection. By differentiating (6.62) with respect to θ in $\theta = 0$, we obtain the following expression of the shape gradient of the compliance using the dual mixed variational formulation for the linear elasticity with weakly imposed symmetry of the stress tensor:

$$\begin{aligned}
 \langle dJ_3(\Omega), \theta \rangle &= \frac{1}{2\mu} \int_{\Omega} (N(\theta) \sigma_\Omega : \sigma_\Omega + \sigma_\Omega N(\theta) : \sigma_\Omega) \, dx \\
 &\quad - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \int_{\Omega} 2 \operatorname{tr} (N(\theta) \sigma_\Omega) \operatorname{tr} (\sigma_\Omega) \, dx \\
 &\quad + \frac{1}{2\mu} \int_{\Omega} (N(\theta) \sigma_\Omega : \eta_\Omega + \sigma_\Omega N(\theta) : \eta_\Omega) \, dx \\
 &\quad + \int_{\Omega} (\nabla f \theta \cdot u_\Omega + f \cdot u_\Omega (\nabla \cdot \theta) - f \cdot (\nabla \theta^T u_\Omega)) \, dx.
 \end{aligned} \tag{6.63}$$

We remark that the two expressions of the shape gradient obtained using the dual mixed variational formulations in subsections 6.4.1 and 6.4.2 are equivalent:

Lemma 6.11. *Let us consider a symmetric stress tensor $\sigma_\Omega \in H(\operatorname{div}, \Omega; \mathbb{S}_d)$. Then (6.60) and (6.63) are equal.*

Proof. It is straightforward to observe that the second and the fourth integrals in (6.63) correspond to the last two terms in (6.60). Moreover, owing to the symmetry of $N(\theta)$ and σ_Ω , we get:

$$\begin{aligned}
 \int_{\Omega} (N(\theta) \sigma_\Omega : \sigma_\Omega + \sigma_\Omega N(\theta) : \sigma_\Omega) \, dx &= \int_{\Omega} (\operatorname{tr} (N(\theta) \sigma_\Omega \sigma_\Omega^T) + \operatorname{tr} (N(\theta)^T \sigma_\Omega^T \sigma_\Omega)) \, dx \\
 &= \int_{\Omega} 2 \operatorname{tr} (N(\theta) \sigma_\Omega \sigma_\Omega^T) \, dx = \int_{\Omega} 2 N(\theta) \sigma_\Omega : \sigma_\Omega \, dx.
 \end{aligned}$$

In order to prove the equality $\langle dJ_2(\Omega), \theta \rangle = \langle dJ_3(\Omega), \theta \rangle$, we have to show that the following quantity is equal to zero:

$$\int_{\Omega} (N(\theta)\sigma_{\Omega} : \eta_{\Omega} + \sigma_{\Omega}N(\theta) : \eta_{\Omega}) dx = \int_{\Omega} (\text{tr}(N(\theta)\sigma_{\Omega}\eta_{\Omega}^T) + \text{tr}(N(\theta)^T\sigma_{\Omega}^T\eta_{\Omega})) dx.$$

The result follows directly from the symmetry of the matrix $N(\theta)$, the symmetry of σ_{Ω} and the skew-symmetry of η_{Ω} . \square

6.5 Qualitative assessment of the discretized shape gradients via numerical simulations

In this section, we provide some numerical simulations to present a preliminary comparison of the expressions of the shape gradient of the compliance derived in this chapter using different formulations of the linear elasticity problem.

As mentioned in subsection 6.1.2, a major drawback of the Hellinger-Reissner variational formulation for the linear elasticity equation is the complexity of the stable Arnold-Winther pair of Finite Element spaces associated with this discretization (cf. [42]). Hence, for the scope of this section, we restrict ourselves to the expression of the shape gradient obtained by the pure displacement formulation (cf. sections 6.1.1 and 6.3) and to the one arising from the dual mixed formulation with weakly imposed symmetry of the stress tensor (cf. sections 6.1.3 and 6.4.2).

We consider the optimal design of the classical cantilever beam described in figure 6.1. In particular, we assume a zero body forces configuration, a structure clamped on Γ^D , with a load $g = (0, -1)$ applied on Γ^N and a free boundary Γ .

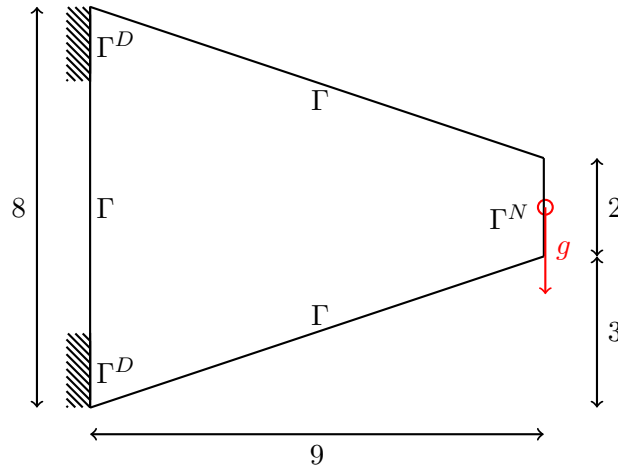


Figure 6.1 – Scheme of a 2D cantilever beam clamped on Γ^D , with a load g applied on the boundary Γ^N and free boundaries Γ .

6.5.1 Experimental analysis of the convergence of the error in the shape gradient

In order to establish an experimental convergence rate for the discretization error associated with the approximation of the pure displacement and the dual mixed formulations of the linear elasticity problem, we consider the cantilever beam described in figure 6.1. In particular, we consider the domain featuring six holes depicted in figure 6.3b. Owing to the fact that the analytical solution of the linear elasticity problem on the aforementioned domain Ω is not known, we solve the linear elasticity problem on an extremely fine mesh and we consider the resulting solution as the exact solution of the problem under analysis. The discretization of the pure displacement formulation of the state problem is performed using $\mathbb{P}^1 \times \mathbb{P}^1$ Finite Element functions to approximate the displacement field. For the dual mixed formulation, we consider the scheme described in subsection 6.1.3 and we approximate the stress tensor using $BDM_1 \times BDM_1$ Finite Elements, the displacement field via $\mathbb{P}^0 \times \mathbb{P}^0$ and the Lagrange multiplier by means of a \mathbb{P}^0 function.

In figure 6.2, we present the convergence history of the discretization error in the shape gradient with respect to the number of Degrees of Freedom for the two discussed Finite Element approximations. In particular, we observe that the strategy based on the dual mixed formulation introduces a lower numerical error than the pure displacement one. Moreover, the blue curve seems slightly steeper than the red one thus the volumetric shape gradient obtained from the dual mixed formulation of the problem may provide better convergence rate than the corresponding expression based on the pure displacement formulation. Nevertheless, this conjecture remains to be proved and a rigorous analysis of the convergence rate by means of *a priori* estimates of the error in the shape gradient is necessary.

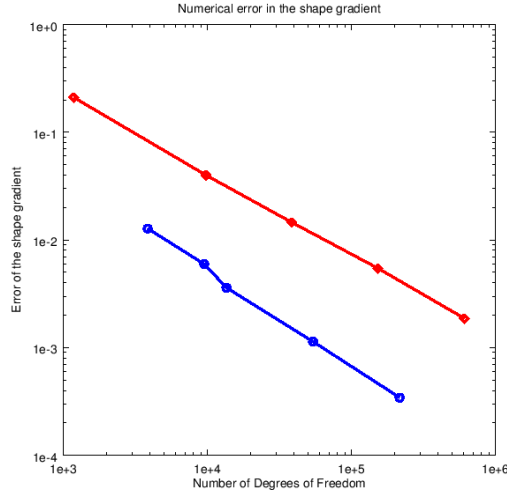


Figure 6.2 – Experimental convergence rate of the error in the shape gradient computed using the pure displacement formulation (red diamond) and the dual mixed formulation (blue circle) with respect to the number of Degrees of Freedom.

6.5.2 Boundary Variation Algorithm using the pure displacement and the dual mixed formulations

In this subsection, we apply the Boundary Variation Algorithm described in subsection 2.1.1 to minimize the compliance of the cantilever in figure 6.1 under a volume constraint. In particular, the volume of the structure under analysis is set to its initial value V_0 and we aim to construct an optimal shape that minimizes the compliance while preserving as much as possible the value V_0 of the volume. As discussed in section 6.2, the volume constraint is handled through a Lagrange multiplier γ . From a theoretical point of view, the value of the Lagrange multiplier should be updated at each iteration in order for the optimal shape to fulfill the volume constraint when the algorithm converges. Nevertheless, enforcing the volume constraint at each iteration would highly increase the complexity of the algorithm and consequently its computational cost. Thus we consider a constant Lagrange multiplier at each iteration of the strategy and starting from the previously computed value γ , we increase it if the current volume V is greater than the target V_0 and we decrease it otherwise.

As mentioned in the introduction and extensively discussed in part II of the thesis, a key aspect of shape optimization procedures is the choice of the criterion to stop the evolution of the optimization strategy. In order to compare the expressions (6.45) and (6.63) of the shape gradient of the compliance, we consider an *a priori* fixed number of iterations for the Boundary Variation Algorithm under analysis. Moreover, the number of connected regions inside the domain is set at the beginning of the procedure and the deformation of the shape is performed via a moving mesh approach. In the rest of this subsection, we present two test cases for the optimal design of the cantilever in figure 6.1, that is a bulky structure (Fig. 6.3a) and a porous one (Fig. 6.3b) featuring 6 internal holes. All the numerical simulations are obtained using FreeFem++ [160].

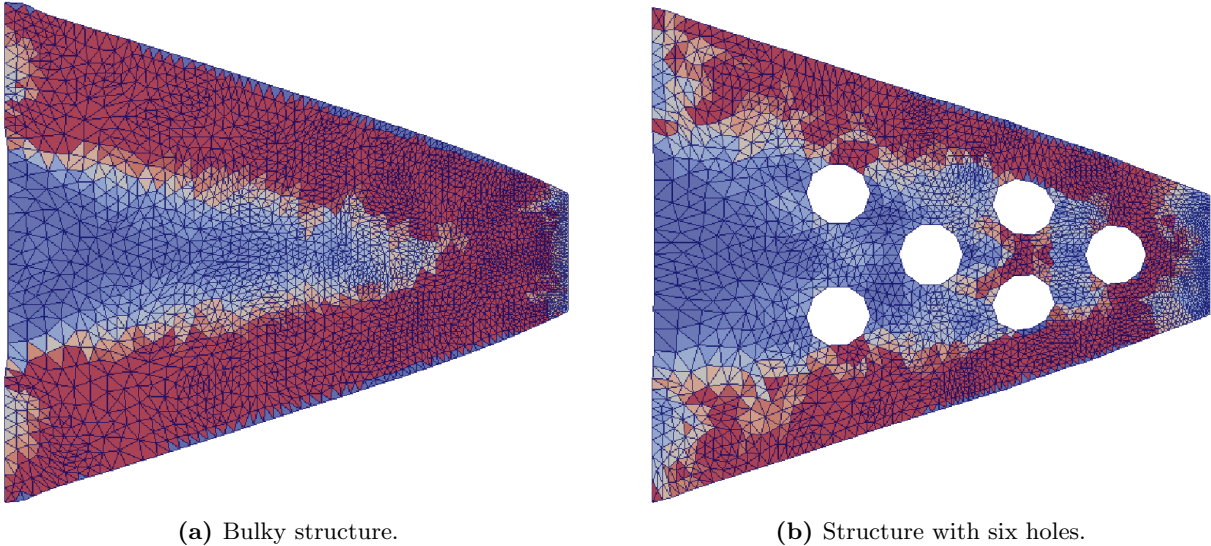


Figure 6.3 – Initial shape and computational mesh for (a) a bulky cantilever and (b) a structure featuring six holes. Density distribution of the elastic energy within the range (a) $(0, 1.5 \cdot 10^{-3})$ and (b) $(0, 3 \cdot 10^{-3})$, the lower values being in blue and the higher ones in red.

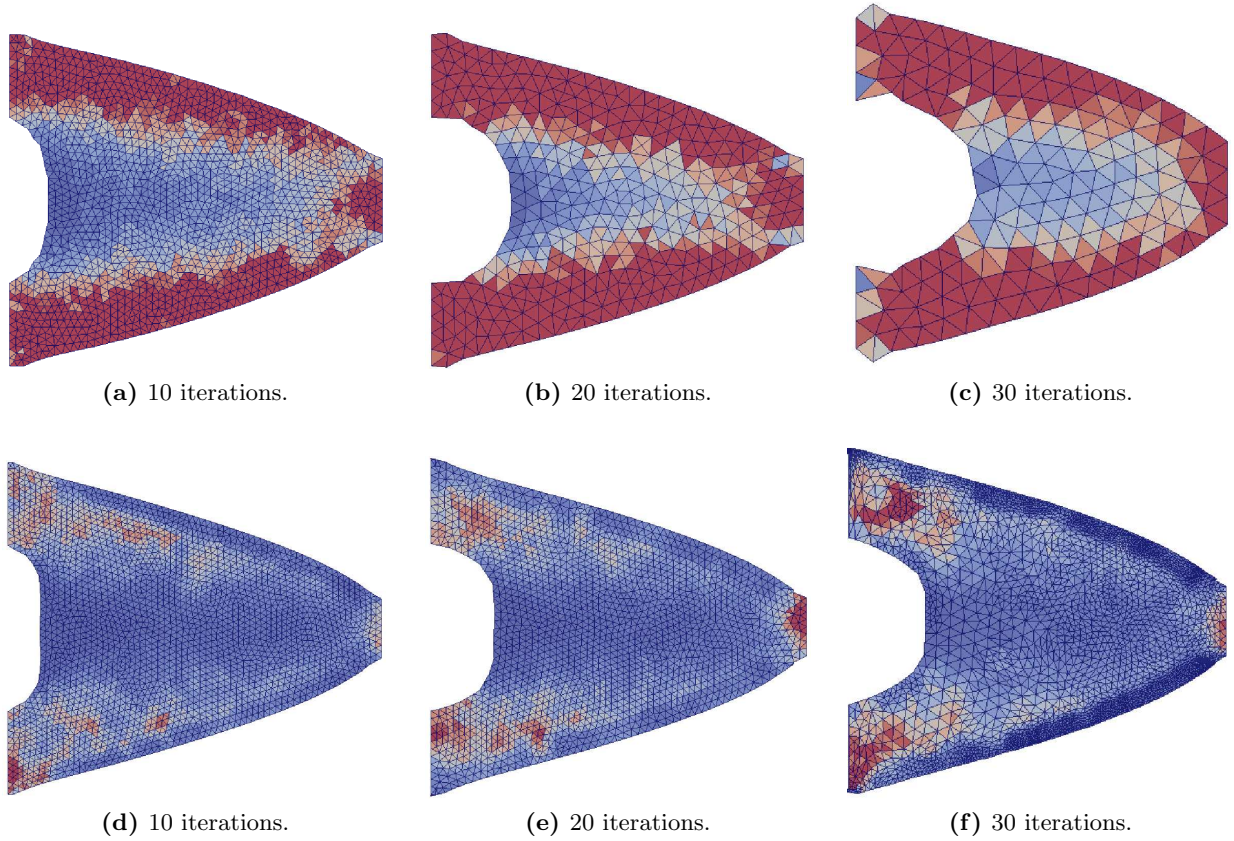
Bulky cantilever beam

Figure 6.4 – Comparison of the BVA after 10, 20 and 30 iterations. At the top: BVA based on the expression of the shape gradient computed using the pure displacement formulation of the linear elasticity problem. At the bottom: BVA using the shape gradient arising from the dual mixed variational formulation. Density distribution of the elastic energy within the range $(0, 1.5 \cdot 10^{-3})$, the lower values being in blue and the higher ones in red.

We consider the initial configuration in figure 6.3a. The volume of the structure under analysis is $V_0 = 45$ and we set the initial value of the Lagrange multiplier to $\gamma_0 = 0.1$. In figure 6.4, we present the shapes obtained using the Boundary Variation Algorithm based on the expressions (6.45) and (6.63) of the shape gradient of the compliance. In particular, we remark that the variant of the BVA which exploits the shape gradient computed via the dual mixed variational formulation of the linear elasticity problem is able to construct configurations in which the total elastic energy is lower than in the corresponding cases obtained starting from the pure displacement formulation of the problem. This remark is confirmed by the comparison plots in figure 6.5 where the BVA based on the dual mixed formulation is depicted by blue curves whereas the red ones represent the results obtained starting from the pure displacement formulation. As a matter of fact, the former approach appears more robust than the latter one: the BVA based on the dual mixed formulation improves both the compliance and the functional $L(\Omega)$ during several iterations, whereas at the beginning of

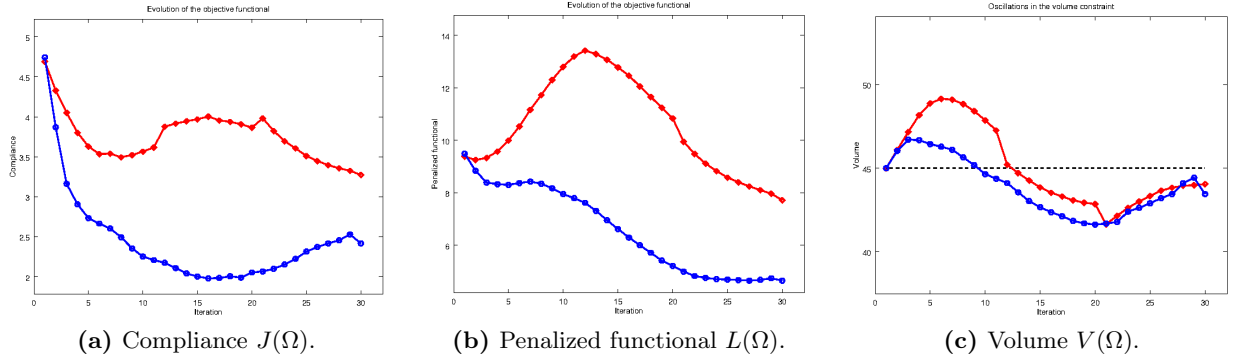


Figure 6.5 – Evolution of the (a) compliance $J(\Omega)$, (b) penalized functional $L(\Omega) = J(\Omega) + \gamma V(\Omega)$ and (c) volume $V(\Omega)$ using the BVA. Results obtained using the pure displacement formulation (red diamond) and the dual mixed one (blue circle). The reference volume V_0 is represented by a black dashed line in (c).

the evolution, the variant exploiting the pure displacement formulation reduces the compliance by enlarging the volume of the structure, thus deteriorating the corresponding value of $L(\Omega)$ (Fig. 6.5b). In a second phase, the BVA based on the pure displacement formulation is able to better control the variation of the volume and the final shapes obtained by the two algorithms have comparable sizes (Fig. 6.5c). Nevertheless, the overall improvement of the compliance is far more limited when using the pure displacement formulation with respect to the one observed starting from the dual mixed formulation (Fig. 6.5a).

Iter. j	$L(\Omega_j)$	$\langle d_h L(\Omega_j), \theta_j^h \rangle$	$L(\Omega_{j+1})$	Iter. j	$L(\Omega_j)$	$\langle d_h L(\Omega_j), \theta_j^h \rangle$	$L(\Omega_{j+1})$
1	9.38	$-5.7 \cdot 10^{-1}$	9.25	1	9.49	-4.78	8.83
5	9.99	$-6.61 \cdot 10^{-2}$	10.52	6	8.37	$-7.28 \cdot 10^{-1}$	8.43
10	12.80	$-1.05 \cdot 10^{-1}$	13.19	10	7.95	$-7.88 \cdot 10^{-1}$	7.79
15	12.78	$-9.4 \cdot 10^{-2}$	12.46	15	6.62	$-5.55 \cdot 10^{-1}$	6.3
25	8.58	$-3.61 \cdot 10^{-2}$	8.4	28	4.68	$-6.55 \cdot 10^{-1}$	4.74
30	7.54	$-9.52 \cdot 10^{-3}$	—	30	4.66	$-6.21 \cdot 10^{-1}$	—

(a) Pure displacement formulation.

(b) Dual mixed formulation.

Table 6.1 – Test case in figure 6.5 using the BVA based on the (a) pure displacement formulation and (b) the dual mixed formulation of the linear elasticity problem. Evolution of the penalized objective functional $L(\Omega)$ with respect to the iteration number. In yellow: the cases in which the discretized direction θ^h fails to be a genuine descent direction for $L(\Omega)$ despite being $\langle d_h L(\Omega), \theta^h \rangle < 0$.

Cantilever beam with six holes

The initial shape for the cantilever beam with six holes is depicted in figure 6.3b and features a reference volume $V_0 = 40.59$ and an initial Lagrange multiplier equal to $\gamma_0 = 0.13$. As for the case of the bulky cantilever, we present snapshots of the shapes obtained at different iterations of the

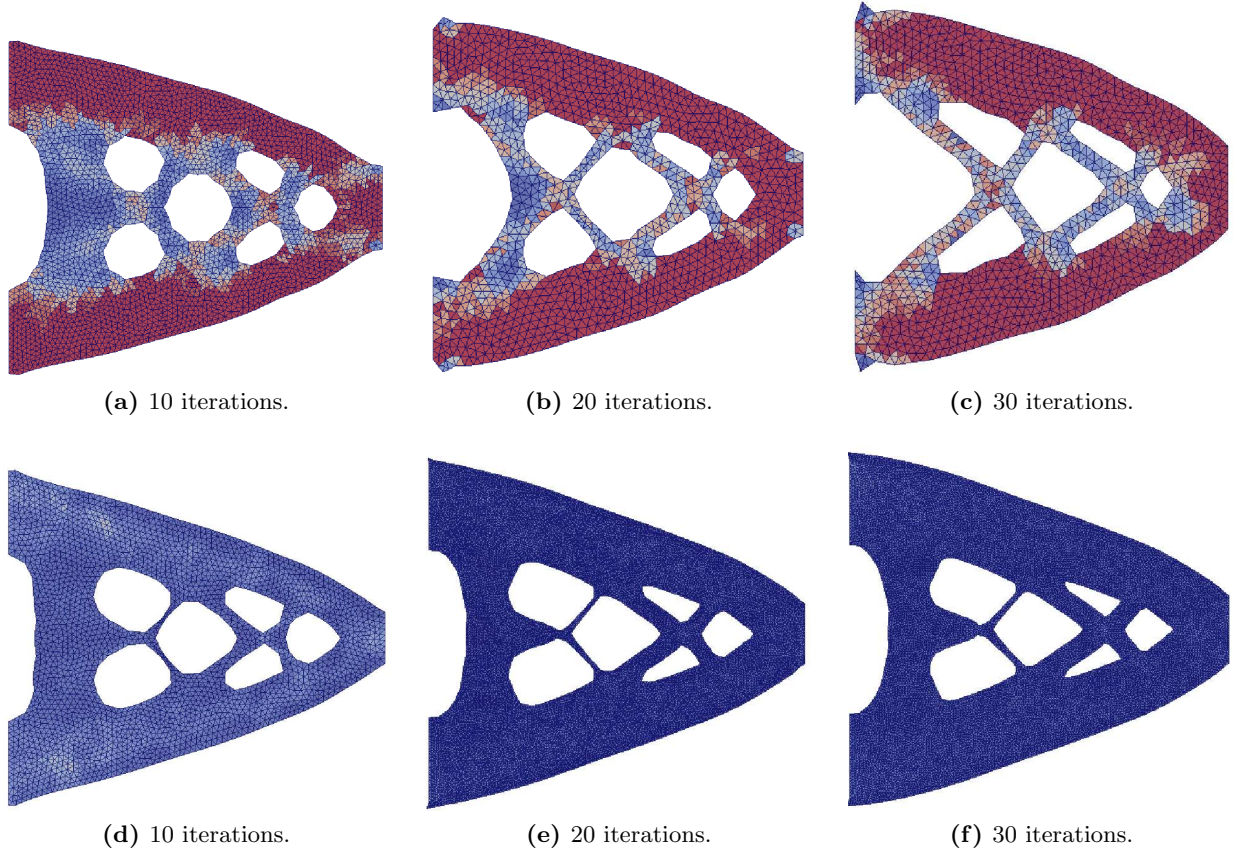


Figure 6.6 – Comparison of the BVA after 10, 20 and 30 iterations. At the top: BVA based on the expression of the shape gradient computed using the pure displacement formulation of the linear elasticity problem. At the bottom: BVA using the shape gradient arising from the dual mixed variational formulation. Density distribution of the elastic energy within the range $(0, 3 \cdot 10^{-3})$, the lower values being in blue and the higher ones in red.

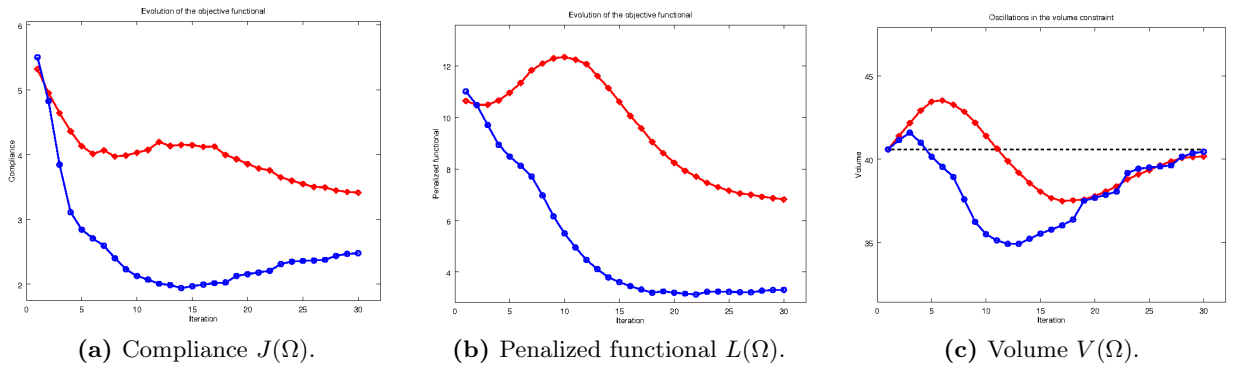


Figure 6.7 – Evolution of the (a) compliance $J(\Omega)$, (b) penalized functional $L(\Omega) = J(\Omega) + \gamma V(\Omega)$ and (c) volume $V(\Omega)$ using the BVA. Results obtained using the pure displacement formulation (red diamond) and the dual mixed one (blue circle). The reference volume V_0 is represented by a black dashed line in (c).

Boundary Variation Algorithm using both the pure displacement and the dual mixed formulation of the linear elasticity problem (Fig. 6.6). Moreover, a qualitative analysis of the evolution of the compliance and of the variation of the volume is discussed starting from figure 6.7. As previously remarked, the Boundary Variation Algorithm based on the dual mixed formulation of the linear elasticity problem leads to configurations with lower elastic energy. Figures 6.7a and 6.7b confirm that the variant of the BVA using the dual mixed formulation generates a sequence of shapes that improve the objective functional for several subsequent iterations. On the contrary, the pure displacement formulation leads to a less robust strategy in which at the beginning of the optimization process, the compliance is reduced by increasing the volume of the structure. Concerning the BVA based on the dual mixed formulation, the comparison of figure 6.6e with figure 6.6f, highlights that only minor modifications of the shape are performed by the algorithm from iteration 20 to iteration 30. As a matter of fact, the evolution of the volume (Fig. 6.7c) shows that after having identified a configuration with low compliance the algorithm tends to correct the shape in order to fulfill the volume constraint which has been violated during the initial iterations. As highlighted by the test case of the bulky cantilever, the Boundary Variation Algorithm based on the dual mixed formulation is able to construct structures with lower compliance than the configurations generated using the pure displacement formulation (Fig. 6.7a). Nevertheless, both the final configuration in figure 6.6c and the one in figure 6.6f, present some issues. On the one hand, the pure displacement solution presents kinks responsible for low compliance near the regions Γ^D where the structure is clamped. On the other hand, the shape obtained by the dual mixed formulation features thin components which may be critical to handle during the manufacturing process. Both these issues may be potentially influenced by the choice of explicitly representing the geometry through the computational mesh and the consequent moving mesh approach to deform the domain. We refer to subsection 6.5.3 for a more detailed comment on this topic.

Iter. j	$L(\Omega_j)$	$\langle d_h L(\Omega_j), \theta_j^h \rangle$	$L(\Omega_{j+1})$	Iter. j	$L(\Omega_j)$	$\langle d_h L(\Omega_j), \theta_j^h \rangle$	$L(\Omega_{j+1})$
1	10.64	-1.00	10.48	1	11.01	-15.32	10.48
3	10.50	$-3.68 \cdot 10^{-1}$	10.66	5	8.49	-2.55	8.12
9	12.30	$-1.74 \cdot 10^{-1}$	12.34	10	5.50	-1.29	4.95
15	10.61	$-1.04 \cdot 10^{-1}$	10.06	18	3.18	-1.36	3.19
25	7.16	$-1.1 \cdot 10^{-2}$	7.05	28	3.32	-2.06	3.36
30	6.82	$-2.75 \cdot 10^{-3}$	—	30	3.37	-2.18	—

(a) Pure displacement formulation.

(b) Dual mixed formulation.

Table 6.2 – Test case in figure 6.7 using the BVA based on the (a) pure displacement formulation and (b) the dual mixed formulation of the linear elasticity problem. Evolution of the penalized objective functional $L(\Omega)$ with respect to the iteration number. In yellow: the cases in which the discretized direction θ^h fails to be a genuine descent direction for $L(\Omega)$ despite being $\langle d_h L(\Omega), \theta^h \rangle < 0$.

Concerning the computational cost of the overall optimization procedures, it is important to remark that the dual mixed formulation features more variables (stress tensor σ_Ω , displacement field u_Ω and Lagrange multiplier η_Ω) than the pure displacement one which - as the name states - solely relies on the displacement field u_Ω . From a practical point of view, this results in a higher number of Degrees of Freedom in the discrete problem and consequently a higher computational cost. Moreover, by comparing the first and the second lines of figures 6.4 and 6.6, we remark that the computations of

the BVA based on the dual mixed formulation were performed on finer meshes than the ones used for the pure displacement one. This turned out to be necessary in order to retrieve an accurate solution of the dual mixed Finite Element problem of linear elasticity, whereas the pure displacement formulation may be easily approximated using Lagrangian Finite Element functions as long as one avoids the nearly incompressible and the incompressible case. Eventually, the linear system obtained by the discretization discussed in subsection 6.1.3 may be extremely ill-posed and the construction of appropriate preconditioners (cf. e.g. [179]) may be necessary. Hence, though the preliminary numerical results suggest that the BVA based on the dual mixed formulation is the best choice when dealing with the minimization of the compliance in linear elasticity, the higher computational cost and the additional numerical difficulties of the overall strategy have to be taken into account to provide a global evaluation of the method. Within this framework, additional investigations have to be performed both from a theoretical point of view (e.g. *a priori* estimate of the error in the shape gradient) and from a computational one, by optimizing and improving the resolution strategy outlined above.

Both figure 6.5 and 6.7 highlight the issue that has motivated the development of the Certified Descent Algorithm in part I of this thesis. It is straightforward to observe that the direction computed using the discretized shape gradient is not always a genuine descent direction for the functional under analysis. This remark is confirmed by tables 6.1 and 6.2 in which we observe that despite being $\langle d_h L(\Omega), \theta^h \rangle < 0$, the functional $L(\Omega)$ may increase when the shape is perturbed accordingly to the field θ^h . Thus, the preliminary simulations of a vectorial shape optimization problem confirm the interest in the development of a certification procedure for the descent direction by accounting for the numerical error introduced by the discretization of the shape gradient. In next chapter, we present the theoretical derivation of the aforementioned *a posteriori* estimator for the error in a Quantity of Interest and we refer to the future work [5] for a more detailed discussion on this topic.

6.5.3 A note on the moving mesh approach

Let us consider the minimization of the compliance under a volume constraint for the cantilever beam initialized in figure 6.8a. We apply the Boundary Variation Algorithm based on the pure displacement formulation and in figure 6.8 we present the snapshots of the configurations computed after 10, 20 and 30 iterations as well as the final shape after 38 iterations. This example summarizes the issues related to the deterioration of the computational mesh mentioned in the previous subsection. In particular, it is straightforward to observe that in the regions near which the structure is clamped the deformation field θ^h is responsible for some spurious oscillations of the existing boundary. Moreover, by intuition, we would expect that the vertical truss which is getting thinner at each iteration is eventually eliminated from the structure. Owing to the considered moving mesh approach, this scenario is impossible and at iteration 38, the optimization algorithm fails to produce a new shape since the thickness of the vertical truss goes below the precision of the mesh generator.

In order to bypass these issues, we refer to the strategy for the implicit description of the geometry mentioned in section 1.2. Currently, we are investigating the approach proposed by Allaire *et al.* in [20] (cf. algorithm 6.1). At each iteration of the optimization strategy, after solving the state problem and computing the descent direction as in the classical Boundary Variation Algorithm [step 1, 2 and 3], we construct a signed distance function to the boundary of the domain [step 4]. Then, we define a level-set function as the solution of an Hamilton-Jacobi equation in which the descent direction

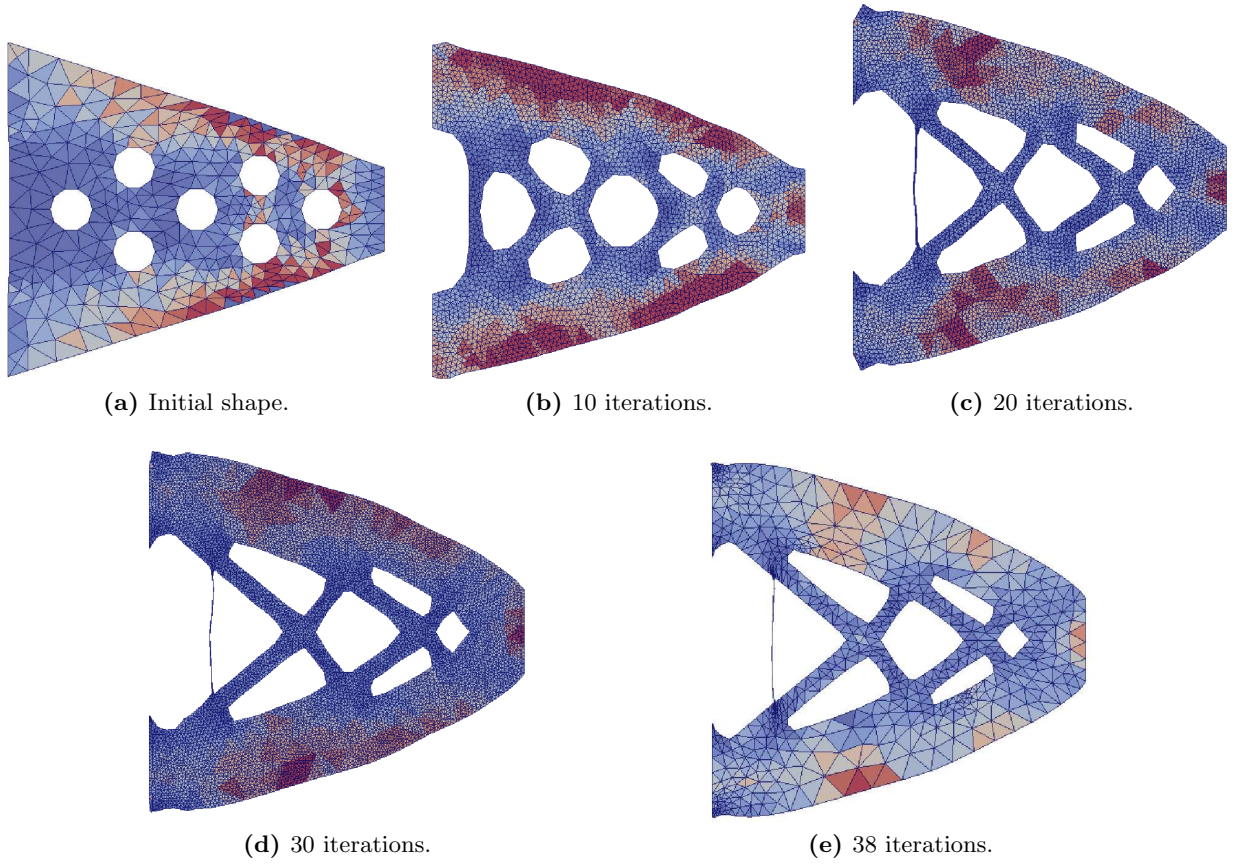


Figure 6.8 – Evolution of the shape of a cantilever with seven holes using the BVA based on the pure displacement formulation. Deterioration of the mesh along the iterations.

computed using the shape gradient of the objective functional acts as transport field and the value of the aforementioned signed distance function as initial condition [step 6]. Eventually, the zero-level curve of the solution is used as the boundary of the newly generated shape [step 7] and a stopping criterion [step 8] verifies whether to end the evolution of the algorithm or to continue it.

Algorithm 6.1 – The discretized BVA with deformation of the domain driven by a level-set function

Given the domain Ω_0 and the box $D \supset \Omega_0$, set $\text{tol} > 0$, $j = 0$ and iterate:

1. Compute the solution $u_{\Omega_j}^h$ of the state equation;
2. Compute the solution $p_{\Omega_j}^h$ of the adjoint equation;
3. Compute a descent direction $\theta_j^h \in X$ solving

$$(\theta_j^h, \delta\theta)_X + \langle d_h J(\Omega_j), \delta\theta \rangle = 0 \quad \forall \delta\theta \in X ;$$

4. Generate the signed distance function to $\partial\Omega_j$ in D ;
 5. Extend the deformation field θ_j^h to the box D ;
 6. Solve the Hamilton-Jacobi equation to propagate the level-set function ϕ_j in D along the transport field θ_j^h using the signed distance function as initial condition;
 7. Retrieve the zero-level curve of ϕ_{j+1} and build the mesh Ω_{j+1} ;
 8. While $|\langle d_h J(\Omega_j), \theta_j^h \rangle| > \text{tol}$, $j = j + 1$ and repeat.
-

Chapter 7

Complementary energy-based estimate of the error in the shape gradient of the compliance

We consider the pure displacement variational formulation of the linear elasticity problem and we discretize it by means of Lagrangian Finite Element functions. Owing to the complementary energy principle, we define a variational problem to compute an accurate approximation of the linearized strain tensor and we derive a fully-computable, constant-free a posteriori estimator of the error in the shape gradient.

7.1 Discretization of the pure displacement formulation of the linear elasticity problem

Let us consider the notation introduced in section 4.1 for the description of the computational mesh \mathcal{T}_h approximating the domain Ω . In this section, we describe a discretization strategy for the linear elasticity problem based on conforming Finite Element functions. In particular, we follow the procedure discussed in chapter 4 for the case of Electrical Impedance Tomography. We discretize the pure displacement formulation (cf. subsection 6.1.1) of the linear elasticity state equation and we construct the adjoint problem to estimate the error in a Quantity of Interest using the expression of the shape gradient of the compliance computed in section 6.3.

First, we introduce the following Lagrangian Finite Element space to approximate the vectorial functions belonging to $H_{0,\Gamma_D}^1(\Omega)$ (cf. (6.13)):

$$V_{\Omega}^{h,\kappa} := \{u^h \in \mathcal{C}^0(\bar{\Omega}; \mathbb{R}^d) : u^h|_T \in [\mathbb{P}^{\kappa}(T)]^d \ \forall T \in \mathcal{T}_h\}$$

where $[\mathbb{P}^{\kappa}(T)]^d$ is the set of d -dimensional vectors such that each component is a polynomial of degree less than or equal to κ on an element T .

7.1.1 The state problem

As in section 6.5, we consider a structure represented by a domain Ω subject to a zero body forces configuration ($f \equiv 0$), clamped on Γ^D , with a load g applied on Γ^N and a free boundary Γ . Let us recall the expression of the bilinear and linear forms associated with the pure displacement formulation of the linear elasticity problem (6.14):

$$a_\Omega(u_\Omega, \delta u) := \int_\Omega A e(u_\Omega) : e(\delta u) \, dx = 2\mu \int_\Omega e(u_\Omega) : e(\delta u) \, dx + \lambda \int_\Omega \text{tr}(e(u_\Omega)) \text{tr}(e(\delta u)) \, dx, \quad (7.1)$$

$$F_\Omega(\delta u) := \int_{\Gamma^N} g \cdot \delta u \, ds. \quad (7.2)$$

Within the framework of conforming Finite Element discretizations, the continuous and discrete bilinear (respectively linear) forms have the same expressions. Hence, the discretization of equation (6.14) may be derived by replacing the analytical solution u_Ω with its approximation u_Ω^h belonging to the space $V_\Omega^{h,\ell}$ of vectorial Lagrangian Finite Element functions of degree ℓ .

In order to solve the problem of minimizing the compliance under a volume constraint using a gradient-based method, we consider the expression $\langle dJ_1(\Omega), \theta \rangle$ of the shape gradient in (6.45). Thus, the approximation of (2.1) for the problem under analysis reads as (2.5) where θ^h is computed using a vectorial Lagrangian Finite Element space and the expression of the shape gradient $\langle dJ_1(\Omega), \theta \rangle$ is replaced by its discretized counterpart. From a practical point of view, the discretized shape gradient $\langle d_h J_1(\Omega), \theta^h \rangle$ is obtained by substituting the analytical solution u_Ω with its Finite Element approximation u_Ω^h in (6.45).

7.1.2 The adjoint problem

In order to minimize the compliance under a volume constraint, we consider the Lagrangian functional (6.29). Gradient-based methods to solve this problem rely on the computation of the shape gradient of $L(\Omega)$, that is

$$\langle dL(\Omega), \theta \rangle = \langle dJ_1(\Omega) + \gamma dV(\Omega), \theta \rangle$$

where γ is a Lagrange multiplier to handle the volume constraint. An overview of the strategies to choose and update γ has been presented in section 6.2 and here we consider the previously discussed approach based on a fixed Lagrange multiplier. We recall that the volume $V(\Omega)$ depends solely on the geometry of the domain, whence the discretization of the state problem does not affect the accuracy of the quantity $\langle dV(\Omega), \theta \rangle$. Thus, only the numerical error in the term $\langle dJ_1(\Omega), \theta \rangle$ has to be estimated and we introduce the influence function r_Ω to evaluate the contribution of the error $u_\Omega - u_\Omega^h$ due to the discretization of the state problem to the error in the approximation of the shape gradient.

In particular, r_Ω is the solution of the following adjoint problem, in which we seek $r_\Omega \in V_\Omega$ such that

$$a_\Omega(\delta r, r_\Omega) = H_\Omega(\delta r) \quad \forall \delta r \in V_\Omega \quad (7.3)$$

where the linear form $H_\Omega(\delta r)$ is the linearization of the shape gradient $\langle dJ_1(\Omega), \theta^h \rangle$ and acts as our Quantity of Interest. Owing to the notation (1.16)-(2.21) introduced in the previous chapters, the shape gradient and its linearization have the following form:

$$\langle dJ_1(\Omega), \theta \rangle = \left\langle \frac{\partial \mathcal{L}}{\partial \varphi}(\Omega, u_\Omega, p_\Omega), \theta^h \right\rangle \quad , \quad H_\Omega(\delta r) := \frac{\partial^2 \mathcal{L}}{\partial \varphi \partial u}(\Omega, u_\Omega^h, p_\Omega^h)[\theta^h, \delta r].$$

In particular, from the expression (6.45) of the shape gradient of the compliance, $H_\Omega(\cdot)$ reads as

$$\begin{aligned} H_\Omega(\delta r) := & \int_\Omega Ae(u_\Omega^h) : \left(\nabla \delta r \nabla \theta^h + (\nabla \theta^h)^T (\nabla \delta r)^T \right) dx \\ & + \int_\Omega A \left(\nabla u_\Omega^h \nabla \theta^h + (\nabla \theta^h)^T (\nabla u_\Omega^h)^T \right) : e(\delta r) dx \\ & - 2 \int_\Omega Ae(u_\Omega^h) : e(\delta r) (\nabla \cdot \theta^h) dx + 2 \int_{\Gamma^N} \left(\nabla g \theta^h \cdot \delta r + g \cdot \delta r (\nabla \cdot \theta^h - \nabla \theta^h n \cdot n) \right) ds. \end{aligned} \quad (7.4)$$

As per the state problems, the discretized solution r_Ω^h is obtained solving problem (7.3) within the space $V_\Omega^{h,m}$ of vectorial Lagrangian Finite Element functions of degree m .

7.2 Estimate of the error in the shape gradient via the complementary energy principle

Following the framework introduced in section 2.3, we construct an *a posteriori* estimator for the error in the shape gradient of the compliance inspired by the complementary energy principle. As previously discussed for the case of Electrical Impedance Tomography, also the problem of minimizing the compliance is self-adjoint. Hence, owing to (2.20) we may write

$$E^h = \langle dJ_1(\Omega) - d_h J_1(\Omega), \theta^h \rangle \simeq H_\Omega(u_\Omega - u_\Omega^h) =: \tilde{E}^h. \quad (7.5)$$

In this chapter, we neglect the linearization error in (7.5) and we construct an estimator for the error \tilde{E}^h via the complementary energy principle. In order to do so, we introduce a pair of dual variables to better approximate the linearized strain tensors $e(u_\Omega)$ and $e(r_\Omega)$. Thus, we rewrite both the state and the adjoint equations as the minimization of some appropriate energy functionals that will be detailed in the following subsections.

We remark that this strategy relies on the possibility of writing the state and adjoint problems through energy functionals solely depending on a primal variable δu (respectively δr) and on a dual one approximating $e(\delta u)$ (respectively $e(\delta r)$). Nevertheless, the linear functional $H_\Omega(\cdot)$ also features a term depending on the gradient of the primal variable which cannot be directly handled using the complementary energy principle. Let us decompose $\nabla \delta r$ into its symmetric part $e(\delta r)$ and its skew-symmetric one $\eta(\delta r) := \frac{1}{2} (\nabla \delta r - \nabla \delta r^T)$. It follows that

$$\nabla \delta r \nabla \theta^h + (\nabla \theta^h)^T (\nabla \delta r)^T = \left(\eta(\delta r) \nabla \theta^h - (\nabla \theta^h)^T \eta(\delta r) \right) + \left(e(\delta r) \nabla \theta^h + (\nabla \theta^h)^T e(\delta r) \right). \quad (7.6)$$

Hence, we may rewrite the first term in (7.4) as follows:

$$\begin{aligned} \int_\Omega Ae(u_\Omega^h) : \left(\nabla \delta r \nabla \theta^h + (\nabla \theta^h)^T (\nabla \delta r)^T \right) dx = & \int_\Omega Ae(u_\Omega^h) : \left(\eta(\delta r) \nabla \theta^h - (\nabla \theta^h)^T \eta(\delta r) \right) dx \\ & + \int_\Omega Ae(u_\Omega^h) : \left(e(\delta r) \nabla \theta^h + (\nabla \theta^h)^T e(\delta r) \right) dx \end{aligned} \quad (7.7)$$

and the linear functional $H_\Omega(\delta r)$ may be decomposed into a term depending solely on $\eta(\delta r)$ and another one on $e(\delta r)$:

$$H_\Omega(\delta r) = H_{\Omega,K}(\delta r) + H_{\Omega,S}(\delta r), \quad (7.8)$$

$$H_{\Omega,K}(\delta r) := \int_{\Omega} Ae(u_\Omega^h) : \left(\eta(\delta r) \nabla \theta^h - \left(\nabla \theta^h \right)^T \eta(\delta r) \right) dx, \quad (7.9)$$

$$\begin{aligned} H_{\Omega,S}(\delta r) := & \int_{\Omega} Ae(u_\Omega^h) : \left(e(\delta r) \nabla \theta^h + \left(\nabla \theta^h \right)^T e(\delta r) \right) dx \\ & + \int_{\Omega} A \left(\nabla u_\Omega^h \nabla \theta^h + \left(\nabla \theta^h \right)^T \left(\nabla u_\Omega^h \right)^T \right) : e(\delta r) dx \\ & - 2 \int_{\Omega} Ae(u_\Omega^h) : e(\delta r) (\nabla \cdot \theta^h) dx \\ & + 2 \int_{\Gamma^N} \left(\nabla g \theta^h \cdot \delta r + g \cdot \delta r \left(\nabla \cdot \theta^h - \nabla \theta^h n \cdot n \right) \right) ds. \end{aligned} \quad (7.10)$$

By some simple manipulations, the first term in (7.10) may be rewritten as follows:

$$\int_{\Omega} Ae(u_\Omega^h) : \left(e(\delta r) \nabla \theta^h + \left(\nabla \theta^h \right)^T e(\delta r) \right) dx = \int_{\Omega} \left(Ae(u_\Omega^h) \left(\nabla \theta^h \right)^T + \nabla \theta^h Ae(u_\Omega^h) \right) : e(\delta r) dx.$$

For the sake of readability, we introduce the following matrices:

$$M_1(\theta^h, u_\Omega^h) := Ae(u_\Omega^h) \left(\nabla \theta^h \right)^T + \nabla \theta^h Ae(u_\Omega^h), \quad (7.11)$$

$$M_2(\theta^h, u_\Omega^h) := \nabla u_\Omega^h \nabla \theta^h + \left(\nabla \theta^h \right)^T \left(\nabla u_\Omega^h \right)^T \quad (7.12)$$

and we rewrite (7.10) as

$$\begin{aligned} H_{\Omega,S}(\delta r) = & \int_{\Omega} M_1(\theta^h, u_\Omega^h) : e(\delta r) dx + \int_{\Omega} M_2(\theta^h, u_\Omega^h) : Ae(\delta r) dx - 2 \int_{\Omega} Ae(u_\Omega^h) : e(\delta r) (\nabla \cdot \theta^h) dx \\ & + 2 \int_{\Gamma^N} \left(\nabla g \theta^h \cdot \delta r + g \cdot \delta r \left(\nabla \cdot \theta^h - \nabla \theta^h n \cdot n \right) \right) ds. \end{aligned} \quad (7.13)$$

Owing to (7.3) and (7.8), we decompose the influence function as $r_\Omega = r_{\Omega,K} + r_{\Omega,S}$, where the contribution $r_{\Omega,K}$ (respectively $r_{\Omega,S}$) is associated with $H_{\Omega,K}(\delta r)$ (respectively $H_{\Omega,S}(\delta r)$) and we seek $r_{\Omega,K}, r_{\Omega,S} \in V_\Omega$ such that

$$a_\Omega(\delta r, r_{\Omega,K}) = H_{\Omega,K}(\delta r) \quad \forall \delta r \in V_\Omega, \quad (7.14)$$

$$a_\Omega(\delta r, r_{\Omega,S}) = H_{\Omega,S}(\delta r) \quad \forall \delta r \in V_\Omega. \quad (7.15)$$

Hence we may now rewrite the error (7.5) in the Quantity of Interest as follows:

$$\begin{aligned} |E^h| & \simeq |\tilde{E}^h| \leq |H_{\Omega,K}(u_\Omega - u_\Omega^h)| + |H_{\Omega,S}(u_\Omega - u_\Omega^h)| \\ & = |a_\Omega(u_\Omega - u_\Omega^h, r_{\Omega,K})| + |a_\Omega(u_\Omega - u_\Omega^h, r_{\Omega,S})| \\ & = |a_\Omega(u_\Omega - u_\Omega^h, r_{\Omega,K} - r_{\Omega,K}^h)| + |a_\Omega(u_\Omega - u_\Omega^h, r_{\Omega,S} - r_{\Omega,S}^h)| \\ & \leq \|u_\Omega - u_\Omega^h\|_\Omega \left(\|r_{\Omega,K} - r_{\Omega,K}^h\|_\Omega + \|r_{\Omega,S} - r_{\Omega,S}^h\|_\Omega \right). \end{aligned} \quad (7.16)$$

In (7.16), the first and last inequalities respectively follow from the triangle and the Cauchy-Schwartz inequality whereas the third line stems from Galerkin orthogonality.

In order to obtain a computable upper bound for the error in the shape gradient, we seek an estimate of the energy-norm of the error for the state solution $u_\Omega - u_\Omega^h$ and for the two components $r_{\Omega,K} - r_{\Omega,K}^h$ and $r_{\Omega,S} - r_{\Omega,S}^h$ of the adjoint solution appearing in (7.16). It is straightforward to observe that the state error and the latter contribution of the adjoint error solely depend on the variables δu (respectively δr) and $e(\delta u)$ (respectively $e(\delta r)$) making the complementary energy strategy a feasible approach to estimate them.

On the contrary, an analogous approach cannot be applied to derive an upper bound of the quantity $\|r_{\Omega,K} - r_{\Omega,K}^h\|_\Omega$ owing to the presence of the skew-symmetric term $\eta(\delta r)$ in (7.9). Moreover, we remark that (7.16) is not fully computable since the exact solution of (7.14) is unknown. A possible workaround for this issue relies on solving problem (7.14) using higher-order Finite Element functions as discussed in chapter 5 for the goal-oriented estimator based on equilibrated fluxes. Let $V_\Omega^{h,n}$ be the space of vectorial Lagrangian Finite Element functions of degree $n > m$ and $I_n^m : V_\Omega^{h,n} \rightarrow V_\Omega^{h,m}$ be the projection operator from the space of high-order functions to the low-order one. Within this framework, we may replace the solution $r_{\Omega,K}$ with its high-order numerical counterpart belonging to $V_\Omega^{h,n}$ whereas $r_{\Omega,K}^h$ is substituted by the projection of the higher-order approximation onto the space of low-order Finite Element functions used to compute the solution of problem (7.15):

$$\|r_{\Omega,K} - r_{\Omega,K}^h\|_\Omega \simeq \|r_{\Omega,K} - I_n^m r_{\Omega,K}^h\|_\Omega. \quad (7.17)$$

In the following subsections, we provide the details of the application of the complementary energy principle to the derivation of the *a posteriori* estimator of the error in the energy-norm for both the state solution u_Ω and the contribution $r_{\Omega,S}$ to the adjoint solution r_Ω . Moreover, we exploit the higher-order approximation of problem (7.14) to construct the fully computable estimator (7.16) of the error in the shape gradient for the compliance.

7.2.1 Energy-norm error estimates for the state equation

In order to estimate the error $e_\Omega := u_\Omega - u_\Omega^h$ due to the approximation of the state problem, we seek the solution of the residual equation (2.18). Following the procedure discussed in subsection 2.3.1, we introduce the potential energy functional $E_\Omega(w)$:

$$E_\Omega(w) := \frac{1}{2} \int_\Omega A e(w) : e(w) \, dx + \int_\Omega A e(u_\Omega^h) : e(w) \, dx - F_\Omega(w). \quad (7.18)$$

The solution of the residual equation for the state problem is equivalent to the minimization of (7.18):

$$E_\Omega(e_\Omega) = \min_{w \in V_\Omega} E_\Omega(w). \quad (7.19)$$

By introducing an additional variable $\rho = e(w)$ and a dual variable $\psi \in H(\text{div}, \Omega; \mathbb{S}_d)$, we construct the Lagrangian functional $L_\Omega : V_\Omega \times L^2(\Omega; \mathbb{S}_d) \times H(\text{div}, \Omega; \mathbb{S}_d) \rightarrow \mathbb{R}$ which reads as

$$\begin{aligned} L_\Omega(w, \rho, \psi) = & \frac{1}{2} \int_\Omega A \rho : \rho \, dx + \int_\Omega A e(u_\Omega^h) : \rho \, dx - \int_\Omega (\psi : \rho + (\nabla \cdot \psi) \cdot w) \, dx \\ & - \int_{\Gamma^N} g \cdot w \, ds + \int_{\partial\Omega \setminus \Gamma^D} (\psi n) \cdot w \, ds. \end{aligned} \quad (7.20)$$

From the system of first-order optimality conditions of (7.20), we get

$$\begin{cases} \rho = A^{-1}\psi - e(u_\Omega^h) & \text{in } \Omega \\ \nabla \cdot \psi = 0 & \text{in } \Omega \\ \psi n = g & \text{on } \Gamma^N \\ \psi n = 0 & \text{on } \Gamma \end{cases} \quad (7.21)$$

Let us define the space $\Psi_\Omega := \{\tau \in H(\text{div}, \Omega; \mathbb{S}_d) : \nabla \cdot \tau = 0 \text{ in } \Omega\}$. We consider the dual complementary energy functional $C_\Omega(\psi)$ for the linear elasticity state problem and we maximize it over the set of all the tensorfields $\psi \in \Psi_\Omega$ fulfilling the boundary conditions in (7.21):

$$\max_{\substack{\psi \in \Psi_\Omega \\ \psi n = g \text{ on } \Gamma^N \\ \psi n = 0 \text{ on } \Gamma}} C_\Omega(\psi) \quad , \quad C_\Omega(\psi) := -\frac{1}{2} \int_\Omega A\rho : \rho \, dx \quad (7.22)$$

where ρ has the form stated in (7.21).

In order to compute the aforementioned dual variable, we define a new variational problem in which we seek $\psi_\Omega \in \Psi_\Omega$ s.t. $\psi_\Omega n = g$ on Γ^N and $\psi_\Omega n = 0$ on Γ which verifies

$$\int_\Omega A^{-1}\psi_\Omega : \delta\psi \, dx = \int_\Omega e(u_\Omega^h) : \delta\psi \, dx \quad \forall \delta\psi \in \Psi_\Omega \text{ s.t. } \delta\psi n = 0 \text{ on } \Gamma^N \cup \Gamma. \quad (7.23)$$

Let ψ_Ω^h be the approximation of the dual variable ψ_Ω via an appropriate Finite Element space. By combining the definition of the energy-norm induced by the bilinear form $\|\cdot\|_\Omega^2 = a_\Omega(\cdot, \cdot)$ with the information in (7.21) and (7.23), we get the following estimator for the error in the discretization of the linear elasticity state problem:

$$\|u_\Omega - u_\Omega^h\|_\Omega^2 \leq \int_\Omega A \left(A^{-1}\psi_\Omega^h - e(u_\Omega^h) \right) : \left(A^{-1}\psi_\Omega^h - e(u_\Omega^h) \right) \, dx. \quad (7.24)$$

7.2.2 Energy-norm error estimates for the adjoint equation

Let $\epsilon_\Omega := r_{\Omega,S} - r_{\Omega,S}^h$. As in the previous subsection, we define the optimization problem associated with the residual equation obtained from problem (7.15):

$$E_\Omega(\epsilon_\Omega) = \min_{w \in V_\Omega} E_\Omega(w) \quad , \quad E_\Omega(w) := \frac{1}{2} \int_\Omega Ae(w) : e(w) \, dx + \int_\Omega Ae(r_{\Omega,S}^h) : e(w) \, dx - H_{\Omega,S}(w). \quad (7.25)$$

Let us introduce a variable $\rho = e(w)$ and a dual variable $\xi \in H(\text{div}, \Omega; \mathbb{S}_d)$. We construct the following Lagrangian functional $L_\Omega : V_\Omega \times L^2(\Omega; \mathbb{S}_d) \times H(\text{div}, \Omega; \mathbb{S}_d) \rightarrow \mathbb{R}$:

$$\begin{aligned} L_\Omega(w, \rho, \xi) = & \frac{1}{2} \int_\Omega A\rho : \rho \, dx + \int_\Omega Ae(r_{\Omega,S}^h) : \rho \, dx - \int_\Omega M_1(\theta^h, u_\Omega^h) : \rho \, dx \\ & - \int_\Omega M_2(\theta^h, u_\Omega^h) : A\rho \, dx + 2 \int_\Omega Ae(u_\Omega^h) : \rho(\nabla \cdot \theta^h) \, dx - \int_\Omega (\xi : \rho + (\nabla \cdot \xi) \cdot w) \, dx \\ & - 2 \int_{\Gamma^N} \left(\nabla g \theta^h \cdot w + g \cdot w \left(\nabla \cdot \theta^h - \nabla \theta^h n \cdot n \right) \right) ds + \int_{\partial\Omega \setminus \Gamma^D} (\xi n) \cdot w \, ds. \end{aligned} \quad (7.26)$$

The first-order optimality conditions of (7.26) read as:

$$\begin{cases} \rho = A^{-1}\xi - e(r_{\Omega,S}^h) + A^{-1}M_1(\theta^h, u_{\Omega}^h) + M_2(\theta^h, u_{\Omega}^h) - 2e(u_{\Omega}^h)(\nabla \cdot \theta^h) & \text{in } \Omega \\ \nabla \cdot \xi = 0 & \text{in } \Omega \\ \xi n = 2 \left(\nabla g \theta^h + g \left(\nabla \cdot \theta^h - \nabla \theta^h n \cdot n \right) \right) =: q & \text{on } \Gamma^N \\ \xi n = 0 & \text{on } \Gamma \end{cases} \quad (7.27)$$

and by plugging (7.27) into (7.26) we may derive the corresponding dual optimization problem:

$$\max_{\substack{\xi \in \Psi_{\Omega} \\ \xi n = q \text{ on } \Gamma^N \\ \xi n = 0 \text{ on } \Gamma}} C_{\Omega}(\xi) \quad , \quad C_{\Omega}(\xi) := -\frac{1}{2} \int_{\Omega} A\rho : \rho \, dx. \quad (7.28)$$

The complementary energy $C_{\Omega}(\xi)$ is defined using the expression of ρ in (7.27) and is maximized among all tensorfields $\xi \in \Psi_{\Omega}$ that satisfy the boundary conditions in (7.27).

The dual variable ξ_{Ω} is determined as the solution of an additional variational problem. In particular, we seek $\xi_{\Omega} \in \Psi_{\Omega}$ s.t. $\xi_{\Omega} n = q$ on Γ^N and $\xi_{\Omega} n = 0$ on Γ which verifies $\forall \delta \xi \in \Psi_{\Omega}$ such that $\delta \xi n = 0$ on $\Gamma^N \cup \Gamma$:

$$\begin{aligned} \int_{\Omega} A^{-1}\xi_{\Omega} : \delta \xi \, dx &= \int_{\Omega} e(r_{\Omega,S}^h) : \delta \xi \, dx - \int_{\Omega} A^{-1}M_1(\theta^h, u_{\Omega}^h) : \delta \xi \, dx \\ &\quad - \int_{\Omega} M_2(\theta^h, u_{\Omega}^h) : \delta \xi \, dx + 2 \int_{\Omega} e(u_{\Omega}^h)(\nabla \cdot \theta^h) : \delta \xi \, dx. \end{aligned} \quad (7.29)$$

Let ξ_{Ω}^h be a Finite Element approximation of the dual variable ξ_{Ω} . The resulting upper bound of the error of the adjoint problem (7.15) in the energy-norm has the following form:

$$\begin{aligned} \|r_{\Omega,S} - r_{\Omega,S}^h\|_{\Omega}^2 &\leq \int_{\Omega} A \left(A^{-1}\xi - e(r_{\Omega,S}^h) + A^{-1}M_1(\theta^h, u_{\Omega}^h) + M_2(\theta^h, u_{\Omega}^h) - 2e(u_{\Omega}^h)(\nabla \cdot \theta^h) \right) : \\ &\quad \left(A^{-1}\xi - e(r_{\Omega,S}^h) + A^{-1}M_1(\theta^h, u_{\Omega}^h) + M_2(\theta^h, u_{\Omega}^h) - 2e(u_{\Omega}^h)(\nabla \cdot \theta^h) \right) \, dx. \end{aligned} \quad (7.30)$$

Eventually, by combining the upper bounds (7.24) and (7.30) computed using the complementary energy principle with the estimate (7.17) obtained from the higher-order discretization of problem (7.14) and plugging them into (7.16), we are able to evaluate the error in the shape gradient and retrieve a fully computable upper bound \overline{E} of the quantity (7.5).

7.3 A note on the Finite Element approximation of the complementary energy problem

The solutions of the dual complementary energy problems (7.23) and (7.29) are sought in a subspace of the space of $d \times d$ square-integrable symmetric tensors with vanishing divergence in Ω and appropriate traction conditions on $\partial\Omega$. From a practical point of view, the construction of a discrete Finite Element space fulfilling these conditions is not trivial and has been an open problem in the scientific community

for several decades. A family of Finite Element spaces suitable to handle this problem is the Arnold-Winther one introduced in [36, 42] to solve the Hellinger-Reissner dual mixed formulation of the linear elasticity problem. Despite this class of Finite Element functions is the most straightforward way to handle the aforementioned problem and some recent works [97, 100] have shown their efficiency for mixed formulations of the linear elasticity problem, the Arnold-Winther Finite Element spaces are currently far from being a widely spread standard in the community. To the best of our knowledge, among the most common Finite Element libraries, only the FEniCS Project (cf. [29], <http://www.fenics.org>) provides partial support for the Arnold-Winther functions.

Alternatively, following the idea discussed in subsection 6.1.3 for the mixed formulation with weakly enforced symmetry of the stress tensor, the relaxation of the symmetry of the target tensor via the introduction of a Lagrange multiplier may be exploited. The main advantage of this approach is the possibility of using more common Finite Element functions to discretize problems (7.23) and (7.29). Despite this approach seems to simplify the approximation of the aforementioned problems allowing the use of well-known Finite Element functions - inspired e.g. by the Brezzi-Douglas-Marini and the PEERS elements (cf. subsection 6.1.3 for a detailed list of options) -, the weak enforcement of the symmetry of the tensor introduces some critical issues when dealing with the complementary energy formulation discussed above. First, a rigorous analysis of the well-posedness of the newly set discrete problem is necessary. Secondly, the tensors ψ_Ω, ξ_Ω are sought in the space $H(\text{div}, \Omega; \mathbb{M}_d)$ such that the divergence of their symmetric parts vanishes in Ω . Eventually, appropriate traction conditions on the symmetric parts of ψ_Ω and ξ_Ω have to be applied on $\partial\Omega$. Though the boundary conditions may probably be handled with some technical complications via a weak imposition, e.g. as discussed in [73, 142, 143], we are not aware of any method to enforce the condition requiring the symmetric part of a given tensor to be divergence-free.

Owing to the previous remarks, we consider the family of Arnold-Winther Finite Element spaces the best approach to approximate the dual problems arising from the complementary energy formulation in section 7.2 and to derive the *a posteriori* error estimators (7.24) and (7.30). Hence, for the numerical results obtained by the application of the Certified Descent Algorithm based on the complementary energy principle to the minimization of the compliance of an elastic structure under a volume constraint we refer to the future work [5].

Summary of the results and prospective developments

In this part, we studied a vectorial shape optimization problem arising from the field of structural mechanics. In particular, we analyzed the well-known problem of designing compliant structures under a volume constraint and we proposed two novel volumetric expressions of the shape gradient of the compliance starting from the pure displacement and the dual mixed formulations of the governing linear elasticity equation. The interest of using volumetric formulations of the shape gradient was first suggested in [64] and later rigorously investigated in [166] for the case of elliptic problems. To the best of our knowledge, the results in chapter 6 are the first attempt to derive volumetric expressions of the shape gradient for a shape-dependent functional within the framework of linear elasticity. A preliminary comparison of the aforementioned expressions by means of numerical simulations showed extremely promising results, especially using the dual mixed variational formulation of the linear elasticity equation. As a matter of fact, the global optimization strategy based on this approach seems more robust than the one obtained from the pure displacement formulation. Moreover, it provides final configurations which appear *visually* better than the ones arising from the pure displacement formulation. Nevertheless, a rigorous and detailed analysis both from an analytical and a numerical point of view is necessary to validate the aforementioned statement.

Eventually, in chapter 7 we presented the derivation of the *a posteriori* estimator of the error due to the discretization of the shape gradient of the compliance. Starting from the pure displacement formulation of the linear elasticity equation, we developed a strategy based on the complementary energy principle. First, we introduced an additional variational problem to accurately evaluate the error in the computation of the strain tensor. Then, we derived a certified and fully computable upper bound of the error in our Quantity of Interest. This estimate represents the starting point for the application of the previously discussed Certified Descent Algorithm to the optimal design of elastic structures which will be treated in the future work [5].

Several aspects related to the discretization of shape optimization problems within the field of linear elasticity still remain to be treated. First, the implementation of the Certified Descent Algorithm using the *a posteriori* error estimate presented in chapter 7. Secondly, a reduction of the computational cost of the certification step is necessary to make more complex problems treatable using the CDA. Within this framework, an extremely interesting line of investigation focuses on the construction of goal-oriented error estimators which rely solely on local quantities like the equilibrated fluxes approach discussed in part II for the scalar problem. The aforementioned framework for the construction of the equilibrated fluxes may be extended to the pure displacement formulation of the linear elasticity equation using the Arnold-Winther Finite Element functions introduced in [42] (respectively

[36]) for the two-dimensional (respectively three-dimensional) case. An alternative approach based on flux-free estimators (cf. [223, 224]) may also be investigated. Concerning the analytical derivation of the volumetric shape gradient of the compliance, a rigorous proof of the equivalence of the expressions obtained using the pure displacement and the dual mixed variational formulations of the linear elasticity problem is required. Moreover, following the analysis performed for the elliptic case in [166], *a priori* estimates of the error in the shape gradient may be derived. This analysis seems particularly interesting since it may provide additional information on the convergence of the shape gradient using different discretization techniques, thus possibly fostering one formulation over the other to achieve better accuracy in the approximation of the shape gradient. Stemming from the work in chapter 6, alternative mixed formulations of the linear elasticity problem may be investigated, e.g. the Tangential-Displacement Normal-Normal-Stress (TD-NNS) recently proposed by Pechstein and Schöberl in [226]. Eventually, we may develop *a posteriori* estimators for the error in the shape gradient using the expressions obtained from the mixed variational formulation of the linear elasticity problem.

Conclusion

In this thesis, we studied the coupling of gradient-based shape optimization methods with *a posteriori* estimators of the error due to the Finite Element approximation of the shape gradient. The main contribution of this work is the development of the Certified Descent Algorithm (CDA) for shape optimization (cf. chapter 2). This method exploits the information of the upper bound of the error in the shape gradient to implement a certification procedure that is able to identify a genuine descent direction for the objective functional at each iteration. Moreover, owing to the quantitative information encapsulated in the estimator in the Quantity of Interest, a novel stopping criterion based on the norm of the shape gradient is defined.

We analyzed two applications which may be formulated as optimization problems of functionals which depend on the shape: a scalar inverse identification problem issue of the field of Electrical Impedance Tomography (cf. part II) and a vectorial forward problem of optimal design of elastic structures (cf. part III).

In part II, we used the identification problem issue of the Electrical Impedance Tomography (EIT) as a proof of concept to establish some properties of the algorithm: the CDA generates a minimizing sequence of shapes, that is, the objective functional is monotonically decreasing along the iteration of the procedure; moreover, the overall strategy automatically ends when the stopping criterion based on the norm of the shape gradient is fulfilled. Concerning the solution of the EIT problem, we discussed two strategies to compute the goal-oriented estimator: on the one hand, we constructed an *a posteriori* estimator of the error in the shape gradient inspired the complementary energy principle (cf. chapter 4); on the other hand, we investigated a technique based on the local reconstruction of equilibrated fluxes (cf. chapter 5) in order to reduce the overall computational cost of the certification step.

In part III, we tackled the problem of designing a compliant structure governed by the linear elasticity equation. First, we analyzed the pure displacement and the dual mixed formulations of the linear elasticity problem and we derived the corresponding volumetric expressions of the shape gradient of the compliance (cf. chapter 6). A preliminary qualitative comparison of the aforementioned expressions was performed by means of numerical simulations. Eventually, in chapter 7 we constructed an *a posteriori* estimator of the error in the shape gradient via the complementary energy principle. This estimate represents the starting point for the application of the CDA to the optimal design of a compliant elastic structures under a volume constraint which will be the subject of the future work [5].

Besides the developments already suggested in the concluding remarks of part II and III, some topics may be of interest for both the inverse identification problem of Electrical Impedance Tomography and the optimal design of elastic structures. First, an accurate investigation of the contribution of the approximation of the geometry to the error in the shape gradient may clarify some issues observed in

the numerical simulations when moving to higher-order Finite Element schemes. In particular, more precise descriptions of the shape via NURBS (cf. [244]) or FreeForm Deformation maps (cf. [51]) may be explored as well as hybrid approaches as the NURBS-Enhanced Finite Element Method (cf. [247]). Other interesting lines of investigation are represented by the reference domain approach proposed by Kiniger and Vexler in [177] and later investigated also by Fumagalli *et al.* in [139] and Hiptmair *et al.* in [165] and by the coupling of shape optimization with Reduced Order Models (cf. e.g. [187]) for which an additional verification step would be required in order to certify the reliability of the reduced model with respect to the high-fidelity one. An implicit representation of the geometry may also be considered and interesting results are expected by the application of the certification procedure to the level-set formulation of PDE-constrained optimization problems of shape-dependent functionals (cf. [25]). Moreover, the certification step may be extended to second order methods [10, 18, 188, 236] and to problems in the field of topology optimization, e.g to drive the insertion of a new inclusion (respectively hole) in a shape identification (respectively shape optimization) problem through the analysis of the error due to the discretization of the topological gradient. Within this latter framework, it is interesting to recall the remarks by Bonnet and Guzina in [71] and Sigmund and Maute in [253] on the numerical use of topological gradients: as a matter of fact, despite the topological derivatives are computed for infinitesimal holes, in practical applications holes of finite size are introduced. Hence, the criteria currently used in the literature to drive the insertion of new holes are solely of heuristic nature and may be improved by a validation step that accounts for the numerical error due to the Finite Element discretization of the problem. Eventually, the analysis of more realistic problems in linear elasticity may generate interesting synergies with the rapidly-expanding community of engineers and industries involved in the field of additive manufacturing which may greatly benefit from the automatization of shape optimization algorithms.

Conclusion (en français)

Dans cette thèse, on a étudié le couplage entre les méthodes d'optimisation basées sur le gradient et les estimations a posteriori de l'erreur due à l'approximation de la dérivée de forme par la méthode des éléments finis. La contribution principale de ce travail est l'Algorithme de Descente Certifiée (CDA) pour l'optimisation de forme (cf. chapitre 2). Cette méthode exploite l'information de la borne supérieure de l'erreur dans la dérivée de forme pour mettre en œuvre une procédure de certification qui identifie une véritable direction de descente pour la fonctionnelle coût à chaque itération. De plus, grâce à l'information quantitative encapsulée dans l'estimation d'erreur sur notre Quantité d'Intérêt, un nouveau critère d'arrêt - basé sur la norme de la dérivée de forme - est défini.

On a considéré deux applications qui peuvent être formulées comme des problèmes d'optimisation d'une fonctionnelle qui dépend de la forme : un problème inverse scalaire d'identification, issu de la tomographie par impédance électrique (cf. partie II) et un problème direct vectoriel de conception optimale de structures élastiques (cf. partie III).

Dans la deuxième partie, on a utilisé le problème de la tomographie par impédance électrique pour valider le CDA et identifier quelques propriétés de la méthode étudiée : le CDA produit une suite minimisante de formes, c'est-à-dire la fonctionnelle coût décroît monotoniquement pendant l'évolution de l'algorithme ; de plus, la stratégie termine automatiquement quand le critère d'arrêt basé sur la norme de la dérivée de forme est atteint.

On a considéré deux stratégies pour calculer l'erreur due à l'approximation numérique de la dérivée de forme. Premièrement, à partir du principe de l'énergie complémentaire (cf. chapitre 4) on a introduit un problème variationnel additionnel pour chaque variable d'état et adjointe et on a récupéré une approximation précise des flux numériques pour construire une bonne estimation de l'erreur en norme de l'énergie. L'inconvénient principal de cette approche découle de la nécessité de résoudre des problèmes locaux globaux pour dériver une borne supérieure explicitement calculable de l'erreur dans la dérivée de forme. En effet, bien que le CDA puisse utiliser des maillages grossiers pour identifier une véritable direction de descente pour la fonctionnelle coût pendant les itérations initiales, le coût global de la procédure de certification tend à être non négligeable. En particulier, cette version de l'Algorithme de Descente Certifiée induit un important coût de calcul additionnel par rapport à l'Algorithme de Variation de la Frontière appliqué sur des maillages fins.

Afin d'éviter de résoudre des problèmes variationnels globaux pour construire l'estimation d'erreur, on a étudié une stratégie qui utilise uniquement des quantités locales. En particulier, dans le chapitre 5 on a proposé une version du CDA qui exploite une estimation dans notre Quantité d'Intérêt au moyen de la reconstruction des flux équilibrés. Cette approche a été développée dans un cadre générique, valide pour des discrétisations par éléments finis conformes et par Galerkin discontinu. D'une part, le nombre de degrés de liberté pour une discrétisation avec éléments finis conformes étant petit, la

procédure pour reconstruire les flux équilibrés demande la résolution d'un sous-problème local défini sur un patch d'éléments pour chaque sommet du maillage. D'autre part, bien que l'approximation Galerkin discontinu des problèmes utilise un nombre de degrés de liberté par élément de maillage plus élevé, le calcul de l'estimation d'erreur basée sur les flux équilibrés est extrêmement simple grâce à une procédure de post-traitement de la solution qui utilise seulement des quantités locales. Les deux stratégies étant valides, on a observé que les bornes obtenues par la méthode de Galerkin discontinu sont plus précises et plus efficaces du point de vue du temps de calcul.

Une validation des deux approches au CDA pour le problème de la tomographie par impédance électrique a été présentée à l'aide des simulations numériques sur plusieurs cas test. Les résultats ont mis en évidence l'importance de prendre en compte l'erreur due à la discrétisation de la dérivée de forme afin d'automatiser l'algorithme d'optimisation de forme.

Plusieurs développements sont possibles dans le domaine de la tomographie par impédance électrique. Premièrement, on devrait essayer d'améliorer la reconstruction de la forme finale. Cet objectif pourrait être atteint en couplant l'information électrique avec une onde acoustique [32] ou une onde élastique [33] et en résolvant le problème d'optimisation de forme correspondant. D'autres améliorations peuvent être obtenues par la construction d'une fonctionnelle coût alternative qui exploite l'information associée à plusieurs mesures prises sur le bord. En ce qui concerne la procédure de certification de la direction de descente à l'aide des estimations a posteriori, une piste intéressante est représentée par la comparaison de l'approche basée sur les flux équilibrés avec une stratégie sans flux [223, 225]. De plus, prendre en compte l'anisotropie du maillage [135, 136] pourrait réduire le nombre global de degrés de liberté en générant un maillage plus adapté au problème physique étudié.

Dans la troisième et dernière partie, on a étudié un problème vectoriel d'optimisation de forme en mécanique. Plus précisément, on a analysé le problème de la conception optimale d'une structure compliant sous une contrainte de volume. On a proposé deux nouvelles expressions volumiques pour la dérivée de forme de la compliance à partir des formulations variationnelles primale et duale mixte de l'équation de l'élasticité linéaire. L'intérêt pour les formulations volumiques de la dérivée de forme a été suggéré pour la première fois dans [64] et ensuite a été rigoureusement étudié dans [166] pour le cas des problèmes elliptiques. À ma connaissance, les résultats dans le chapitre 6 représentent le premier essai de calcul des expressions volumiques de la dérivée de forme d'une fonctionnelle qui dépend de la forme dans le cadre de l'élasticité linéaire. Une comparaison préliminaire des expressions susdites au moyen des simulations numériques a montré des résultats extrêmement prometteurs, en particulier en utilisant la formulation duale mixte du problème de l'élasticité linéaire. Cette approche semble assurer une plus grande robustesse de la stratégie globale d'optimisation et des meilleures conceptions finales mais une analyse plus détaillée et rigoureuse est nécessaire.

Enfin, dans le chapitre 7, on a présenté le calcul d'une estimation a posteriori de l'erreur de discrétisation pour la dérivée de forme de la compliance au moyen du principe de l'énergie complémentaire. À partir de la formulation primale en déplacements du problème de l'élasticité linéaire, on a introduit un problème variationnel additionnel pour évaluer précisément l'erreur dans le calcul du tenseur des déformations. Ensuite, on a construit une borne supérieure certifiée et explicitement calculable de l'erreur effectuée sur le calcul de notre Quantité d'Intérêt. Cette estimation représente le point de départ pour l'application de l'Algorithme de Descente Certifiée, préalablement présenté, à la conception optimale de structures élastiques qui fera l'objet du prochain article [5].

Plusieurs aspects qui concernent la discrétisation des problèmes d'optimisation de forme dans le domaine de l'élasticité linéaire restent à traiter. Premièrement, la mise en œuvre de l'Algorithme

de Descente Certifiée en utilisant l'estimation d'erreur présentée dans le chapitre 7. Ensuite, une réduction du coût de calcul est nécessaire pour rendre faisable l'application du CDA à des problèmes plus complexes. Une piste de recherche extrêmement intéressante est la construction d'estimations d'erreur qui exploitent uniquement des quantités locales comme les flux équilibrés utilisés pour le problème scalaire dans la partie II. La construction des flux équilibrés mentionnée ci-dessus peut être appliquée aussi à la formulation primale en déplacement de l'équation de l'élasticité linéaire en utilisant les fonctions éléments finis d'Arnold-Winther introduites dans [42] (respectivement [36]) pour le cas bidimensionnel (respectivement tridimensionnel). Une approche alternative basée sur les estimations sans flux [223, 224] pourrait également être envisagée. En ce qui concerne la dérivation analytique de la dérivée de forme volumique de la compliance, une preuve rigoureuse de l'équivalence des expressions obtenues utilisant les formulations primale et duale mixte du problème de l'élasticité linéaire est nécessaire. De plus, on pourrait calculer des estimations a priori de l'erreur pour la dérivée de forme comme Hiptmair *et al.* dans [166] pour le cas elliptique. Cette analyse semble particulièrement intéressante puisqu'elle fournirait des informations supplémentaires sur la convergence de la dérivée de forme calculée utilisant différentes techniques de discrétisation. Ainsi, des informations additionnelles seraient acquises permettant de choisir une formulation plutôt qu'une autre afin d'obtenir une meilleure précision dans l'approximation de la dérivée de forme. À partir du travail dans le chapitre 6, des formulations mixtes alternatives pour le problème de l'élasticité linéaire pourraient être analysées, par exemple la formulation TD-NNS récemment proposée par Pechstein et Schöberl [226]. Finalement, des estimations a posteriori de l'erreur de la dérivée de forme calculée utilisant les expressions obtenues à partir des formulations mixtes du problème de l'élasticité linéaire pourraient être développées.

Sans compter les développements suggérés ci-dessus, d'autres améliorations dans l'analyse des deux problèmes traités dans les parties II et III sont possibles. Premièrement, une analyse détaillée de l'influence de l'approximation de la géométrie sur l'erreur de la dérivée de forme permettrait de clarifier quelques questions induites par les simulations numériques quand on considère des schémas éléments finis d'ordre élevé. En particulier, des techniques pour décrire plus précisément la forme au moyen de courbes NURBS [244] ou des déformations de forme libre FFD [51] mériteraient d'être explorées aussi bien que des approches hybrides comme la méthode NEFEM des éléments finis améliorée par les courbes NURBS [247]. Des autres pistes de recherche intéressantes sont représentées par l'approche basée sur un domaine de référence proposée par Kiniger et Vexler [177] et étudiée ensuite aussi par Fumagalli *et al.* [139] et Hiptmair *et al.* [165] et le couplage des méthodes d'optimisation de forme avec des modèles d'ordre réduit (cf. [187]) pour lesquels une étape additionnelle de vérification serait nécessaire afin de certifier la fiabilité du modèle réduit par rapport au modèle complet. Des résultats intéressants sont attendus aussi d'une représentation implicite de la géométrie et de l'application de la procédure de certification à une formulation par lignes de niveau (cf. [25]) des problèmes d'optimisation de forme. De plus, il serait légitime d'appliquer l'étape de certification à des méthodes du second ordre [10, 18, 188, 236] et à des problèmes d'optimisation topologique : des estimations a posteriori de l'erreur due à la discrétisation de la dérivée topologique dirigeraient l'insertion d'une nouvelle inclusion (respectivement trou) dans un problème d'identification (respectivement d'optimisation) de forme. À ce titre, il est intéressant de rappeler les remarques de Bonnet et Guzina [71] et Sigmund et Maute [253] sur l'utilisation des dérivées topologiques approximées dans des algorithmes numériques : en fait, bien que les dérivées topologiques soient calculées pour des trous infinitésimaux, dans la pratique des trous de taille finie sont introduits. Par conséquent, les critères actuellement utilisés dans la littérature pour diriger l'insertion de nouveaux trous sont exclusivement heuristiques et pourraient être améliorés par

Conclusion

une étape de validation qui prend en compte l'erreur numérique due à la discrétisation du problème par la méthode des éléments finis. Finalement, l'analyse de problèmes réalistes plus complexes en élasticité linéaire pourrait mettre en œuvre des synergies intéressantes avec la communauté des ingénieurs et des industries actifs dans le domaine de la fabrication additive qui peut obtenir des bénéfices considérables du processus d'automatisation des algorithmes d'optimisation de forme.

Appendix A

Identification of optimal grasping points for an Unmanned Aerial Vehicle

This appendix is devoted to the presentation of the preliminary results obtained in a joint work with J.-A. Escareno and co-workers on the use of optimization techniques for the identification of the optimal grasping points in Unmanned Aerial Systems. This research belongs to the line of investigation focusing on the development of Micro Unmanned Aerial Vehicles at the Institut Polytechnique des Sciences Avancées.

This project tackles the practical engineering problem of conceiving an optimization algorithm that a Micro Air Vehicle (MAV) can execute in real-time to identify the optimal points to grasp an object starting from the images acquired by an on-board camera. Working on this subject has been extremely important to understand the constraints arising in real-world engineering applications and to develop an efficient algorithm starting from a mathematical model. This approach to a real-life problem may be usefully applied to the shape optimization problems described in the core of this thesis in order to develop novel variants of the Certified Descent Algorithm able to handle practical engineering constraints and thus viable for industrial applications.

Optimal Grasping Points Identification for a Rotational Four-fingered Aerogripper

J. Vazquez¹, M. Giacomini^{2,3}, J. Escareno³, E. Rubio¹ and H. Sossa¹

Abstract—In the present paper, we deal with the grasping aspect regarding reactive object retrieving during fast aerial pick-and-place maneuvers. The paper addresses the problem of the optimal grabbing of an object by means of a 5 degrees of freedom (DoF) gripper. Also, is introduced the optimization framework for the identification of the optimal contact points between the end-effector and the object. In particular, we define an objective functional to evaluate the optimality of the contact points and we propose a preliminary strategy to identify the optimal contact points. Simulations and experimental are presented to support the actual proposal.

I. INTRODUCTION

In the last years the Unmanned Aerial Vehicles (UAVs) were used for a wide variety of applications, either civilian and military. The research and development of MAVs have different facets according to the application scope and its main feature is the constant evolution of such domain. However, the enhancement of classical UAVs-based applications requires aerial robots endowed with novel capabilities. Among these novel applications it is possible to cite aerial interactivity (aerial manipulation/grasping), swarm-based operations (multi-uav cargo carrying, natural disaster assessment, convoy escort/surveillance,...), natural science (zoological research, environmental observation,...).

Moving towards the operational limits of current UAV configurations unveils new challenges involving different scientific disciplines. For instance, maintaining the balance between mechanical enhancements and flight stability (translational and rotational motion). In this paper we are interested on aerial interactivity, specifically on aerial grasping which is an active research trend due to its application potential. Various contributions can be found on the literature. [1] presents the planar model, attitude control analysis and outdoors experimental validation of a middle-size helicopter equipped with a compliant gripper capable of robust grasping and transporting objects of different shapes and dimensions. In [2], a classical quadrotor featuring a home-customized

1DOF gripper performs an aerial grasping based IR camera evaluate. In [3], the experiments are extended to outdoor, using a GPS system and a Kalman filter to improve the precision in the position system. But both contributions are limited to the use of a 1-DOF gripper, which reduces the precision of manipulation. [4] presents a quadrotor equipped with a four-fingered gripper which enables to perform aerial grasping and perching. The gripper is directly attached to the vehicle, this fact restricts the grasping workspace, i.e. the vehicle's center of mass (CoM) must be aligned to the object to be grasped (target). From the mechanical point of view, the gripper is significantly complex featuring 16 DOF, 4 joints per finger. In [5], the authors present a classical quadrotor equipped with a monocular camera. The proposed control strategy enables performing aggressive grasping maneuvers via an Image Based Visual Servoing (IBVS). It is claimed that unlike most IBVS approaches, the dynamics is obtained directly in the image to deal with a second order system. An alternative UAV configuration equipped with a hook intended to deliver/retrieving cargo using a vision-based strategy is presented in [6].

In this paper is proposed a grasping strategy adapted for aerial pick-and-place operations using a quadrotor equipped with a 3DOF delta parallel robot. The advantages of parallel against serial robots such as more rigid structure, high orientation accuracy, enhanced stability and repetitiveness, functioning, control on the limits of velocities. These manipulation features render the delta robot as the ideal platform for aerial pick-and-place tasks. The actual proposal is based on a novel 5 DoF four-fingered rotational gripper (see Fig. 1), which, in terms of mechanics, is simple and robust. The rotational capability of the proposed end-effector decouples the motion of the rotorcraft. The current paper focuses on the strategy to improve the grasping effectiveness. The latter represents the cornerstone to achieve the final goal of the ongoing project, i.e. performing high-reactive aerial pick-and-place operations. In this paper we focus our efforts to increase Aerogripper's grasping effectiveness by introducing a grasping criteria to detect optimal contact points. Simulations results are presented to illustrate the objective of the proposed approach. Lastly, embedded image processing was used to validate and evaluate experimentally the optimal grasping-points detector.

The paper is organized as follows. Section II the grasping problem is discussed. In section III the proposed grasp-

*This work was supported by Polytechnic Institute of Advanced Sciences and CONACYT Mexico

¹J. Vazquez, E. Rubio and H. Sossa are with Computer Research Center of National Polytechnic Institute, Mexico City, MEXICO. erubio/hsossa@cic.ipn.mx

²M. Giacomini is with the DEFI team at INRIA Saclay Île-de-France, CMAP Ecole Polytechnique, Palaiseau, France. matteo.giacomini@polytechnique.edu

³Polytechnic Institute of Advanced Sciences, Ivry-sur-Seine, FRANCE. email: juan-antonio.escareno@ipsa.fr

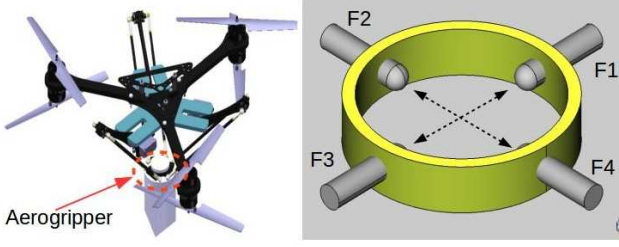


Fig. 1: Aerial grasping system (AGS): [left] AGS with cargo [right] Aerogripper scheme

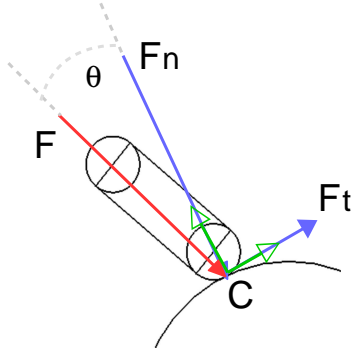


Fig. 2: Contact point: contact force components. Where the contact frame is denoted by $\{C\}$ with axes $\{\hat{n}, \hat{t}\}$

ing strategy is described mathematically. Illustration and validation of the approach is provided in sections IV and V respectively. Finally, section VI presents the concluding remarks as well as the future work of the project.

II. THE GRASPING PROBLEM

A. Contacts Points

During grasping or manipulation operations, reliable contacts between the fingers and the object depend on different aspects, such as: friction, vector force direction and magnitude. In the current paper it is assumed the following.

- A1. Non-sliding contacts are considered, i.e.

$$\|F_t\| \leq \mu \|F_n\| \quad (1)$$

where μ stands for the dry friction coefficient

- A2. The static and dynamic friction coefficients are equal.

The finger force $F \in \mathbb{R}^2$, two-dimensional for this case, applied on the surface is features two components the normal and tangential forces (see Fig. 2). The actual grasping strategy aims to reduce as much as possible the tangential component, i.e. the applied force F is orthogonal to the contact surface.

B. Aerogripper Description

Guaranteeing the grasping of nonregular-form objects often requires dexterous anthropomorphic-like grippers which

implies having mechanically complex end-effectors. The Aerogripper concept is meant to enhance aerial grasping with a simpler and robust design. The gripper features 5 degrees of freedom (DoF) (Aerogripper) having four linear actuators (fingers) orthogonally arranged. Rotational motion of the gripper is also possible. Each finger is actuated separately to adapt to object geometry ((see Fig. 3)). In addition, beyond the grasping capability, the gripper design allows different configurations to manipulate the object with the grasping plane.

III. OPTIMAL GRASPING POINTS DETECTION

As previously stated, a reliable contact point between the gripper and the object is considered as optimal if the grabbing force is orthogonal to the object's surface. The strategy is based in finding the optimal grasping points of the objects while respecting the orthogonal position of the fingers.

Let $\varphi = \varphi(\tau)$ be the boundary of the object parametrized via the curvilinear abscissa τ and $t(\tau) = \varphi'(\tau)$ the tangential direction to the boundary itself. The optimality condition for a contact point τ_k on $\varphi(\tau)$ reads as the following orthogonality condition

$$t(\tau_k) \cdot p_k = 0 \quad (2)$$

where p_k is the direction of the force associated with the k -th grabbing point of the Aerogripper. Since the scalar product is zero between two orthogonal vectors, the bigger is the dot product $t(\tau_k) \cdot p_k$, the less optimal the corresponding contact point τ_k is.

In order to establish the optimality for the current grasping configuration, we have to evaluate the concurrent optimality for all the m contact points. For this purpose, let us introduce an objective functional which provides the discrepancy of each point τ_k from the optimal configuration:

$$J(\tau_1, \dots, \tau_m) = \sum_{k=1}^m |t(\tau_k) \cdot p_k|. \quad (3)$$

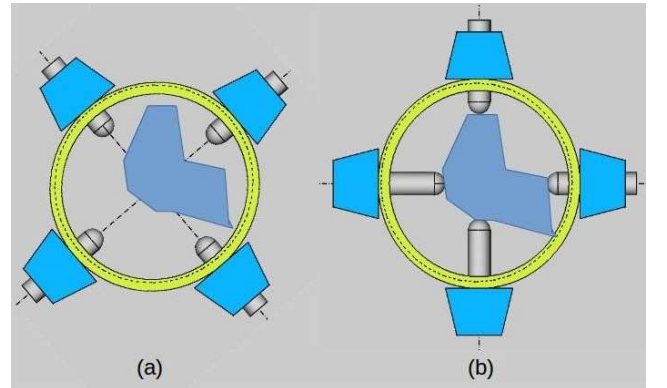


Fig. 3: Aerogripper operation: (a) Object within the Aerogripper's workspace (b) Object grasping after contact points identification and gripper rotation

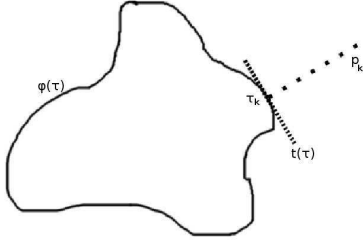


Fig. 4: Orthogonality property in a given contact point τ_k .

A. Identification of the contact points

In order to fully-compute the objective functional (Eq. 3) for a given grabbing configuration, first we have to identify all the contact points τ_k 's, for $k = 1, \dots, m$. The contact points τ_k 's stand at the intersection of the boundary $\varphi(\tau)$ and the vectors representing the directions of the grasping forces p_k 's (see Fig. 4).

Remark 3.1: The number of the contact points is fixed *a priori* and depends on the configuration design of the grasping mechanism.

First, we consider a simplified model of the rotational gripper with four grasping points (see Fig. 5 and Fig. 3). We introduce a parametrization using an angle θ such that the four grasping points are located in the the following configuration

$$\theta_k = \theta^* + (k-1)\frac{\pi}{2}, \quad k = 1, \dots, 4$$

where θ^* represents the angle of the first grasping point. In figure 5 we present a configuration in which $\theta^* = 0$, that is the first grabbing point is located along the horizontal diameter of the circle.

Remark 3.2: The angle θ^* is not known *a priori* and will act as the optimization variable in the algorithm to identify the optimal grabbing configuration in section III-C.

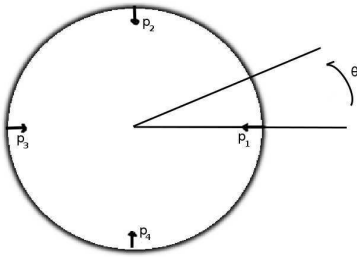


Fig. 5: Aerogripper scheme: scheme of the circular four contact points. The grabbing points are located in the configurations θ_k , $k = 1, \dots, 4$ and the associated grabbing forces are p_k , $k = 1, \dots, 4$.

By exploiting this parametrization, the directions of the grasping forces p_k 's associated with proposed configuration θ_k 's are given by

$$p_k = [-\cos \theta_k, -\sin \theta_k]^T, \quad k = 1, \dots, 4.$$

Hence, to identify the contact points τ_k 's, for $k = 1, \dots, 4$ it is sufficient to solve the problem

$$\varphi(\tau_k) = p_k, \quad k = 1, \dots, 4. \quad (4)$$

B. Computation of the tangential directions

Let us assume that the boundary $\varphi(\tau) \in \mathcal{C}^1(\mathbb{R})$. Under this assumption, we are allowed to compute the first-order derivative in all the points of the boundary, thus retrieving the information on the tangential vectors needed to evaluate the functional (Eq. 3).

Since the analytical expression of the boundary is not known, from a practical point of view we compute the numerical approximation of the first-order derivative using a Finite Difference scheme (Fig. 6). Let $\delta\tau$ be a small step size along the curvilinear abscissa τ . The first-order derivative of $\varphi(\tau)$ in a given point τ_i may be approximated using the following centered Finite Difference formula:

$$t(\tau_i) := \varphi'(\tau_i) \simeq \frac{\varphi(\tau_i + \delta\tau) - \varphi(\tau_i - \delta\tau)}{2\delta\tau}. \quad (5)$$

Having introduced all the components required for the evaluation of the optimality of a grabbing configuration, we may now sketch a preliminary procedure to identify the optimal contact points for the Aerogripper.

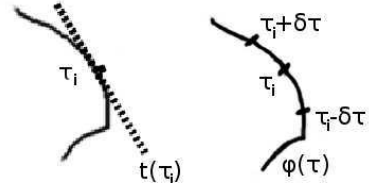


Fig. 6: Approximation of the tangential direction $t(\tau)$ by means of the centered Finite Difference scheme (5).

C. A simple procedure for the optimization of the contact points

In this section we introduce a simple procedure for the identification of the optimal contact points between an object and the Aerogripper fingers. The actual grasping strategy relies on the comparison of the objective functional (Eq. 3) computed for different grasping configurations, each of which is represented by means of a different value of the angle $\theta^* \in [0, 2\pi)$. By exploiting the symmetrical structure of the circular mechanical gripper, the optimization problem might be reduced to the analysis of the values within the interval $[0, \frac{\pi}{2})$ for the angle θ^* since the other configurations will be associated with the remaining three grasping points via rigid rotations.

a) Step 1. Identification of the boundary: Let us now assume that the analytical parameterization of the boundary of the object is not known. We consider a cartesian grid of points - representing the pixels of an image - and the only available information is the list of grid points belonging to the boundary of the object.

Before running the optimization procedure, we need to reconstruct the boundary using a numerical interpolation tool, i.e. a B-Spline interpolation algorithm.

b) *Step 2. Optimization loop:* Let us consider $\theta^* = 0$ as the initial configuration of the grabbing mechanism. The *greedy* optimization procedure relies on the idea of introducing a uniform sampling of the interval $[0, \frac{\pi}{2})$ and evaluating the objective functional (Eq. 3) for all these entries. Among the computed values, the minimal one is picked and the associated θ^* is chosen as the angle for the optimal grabbing configuration.

For every value θ^* in $[0, \frac{\pi}{2})$:

- 1) Identify the four contact points by solving equation (Eq. 4) for $k = 1, \dots, 4$, i.e. by means of a Newton method.
- 2) Approximate the tangential direction in the four points computed in the previous step using the Finite Difference scheme (5).
- 3) Evaluate the objective functional (3) associated with the current contact points.
- 4) Store the optimal values. During the first iteration, store the value of the objective functional computed in the current step. In the following iterations, check whether the newly computed value is smaller than the previous one. If so, flag the new angle as the optimal one and store the value of the objective functional, otherwise keep the previous data.
- 5) Update the angle.

Hence, the optimal grabbing configuration is represented by the setting for which the contact points are

$$\begin{aligned} \theta_1 &= \theta^* & , & \quad \theta_2 = \theta^* + \frac{\pi}{2} \\ \theta_3 &= \theta^* + \pi & , & \quad \theta_4 = \theta^* + \frac{3}{2}\pi \end{aligned}$$

where the angle θ^* arises from the procedure in step 2.

IV. SIMULATION STUDY

In this section is presented the numerical study in Matlab in order to provide a in-depth idea about the proposed grasping discussed in the previous section. For the sake of illustration, let us present two examples:

A. Square geometry

It is straightforward to guess the optimal contact points for this geometric shape. Thus, applying our algorithm we obtain four points where the finger axis and the surface are orthogonal, i.e. the dot product is zero (see Fig. 7). The points are located at 45, 135, 225 and 315 degrees as is depicted by (Fig. 8). Thus, these points represent the commanded grasping point. It is important to say that the dot product is almost zero since the orthogonal grippers configuration and the shape are perfectly orthogonal which is not the case for non-regular geometries.

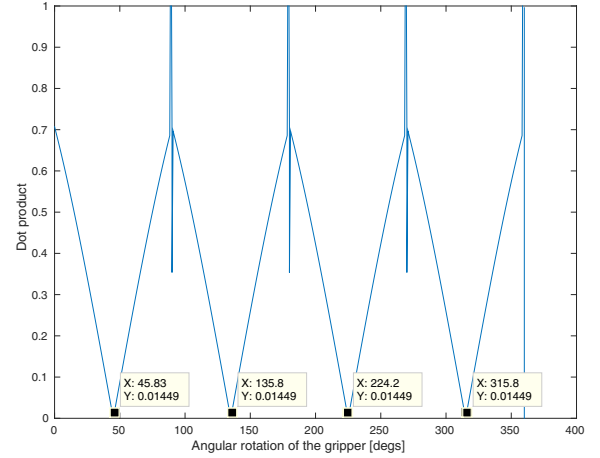


Fig. 7: Gripper angular displacement based on minimal dot-product contact points

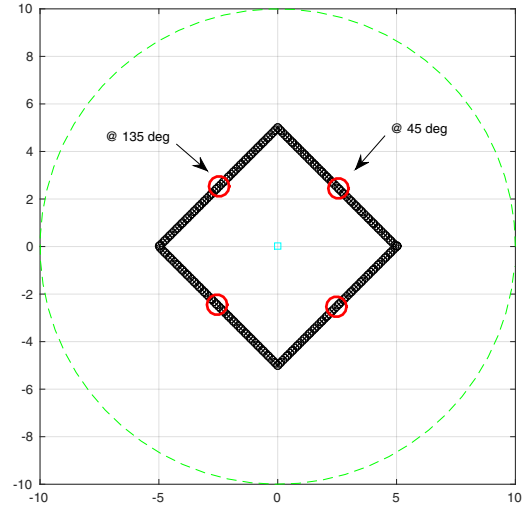


Fig. 8: Optimal grasping points

B. Non-regular geometry

In this case the shape is not regular and the the optimal contact points are obtained through the proposed algorithm. It is worth to recall that the dot-product values selected are the minimal allowed for the orthogonal gripper configuration.

V. EXPERIMENTAL RESULTS

A. UDOO Microprocesor

The microprocessor used to evaluate the proposed grasping strategy is UDO0 which is an embedded computer developed by Aidilab and SECO USA inc. This board features the equivalent of 4 raspberry pi board plus one arduino, the processor for realizing this purpose is the ARM Cortex-A9 CPU Dual/Quad core 1 GHz. An interesting feature is that UDO0 also has an embedded Arduino DUE (working 80 Mhz). The microprocessor features an embedded camera

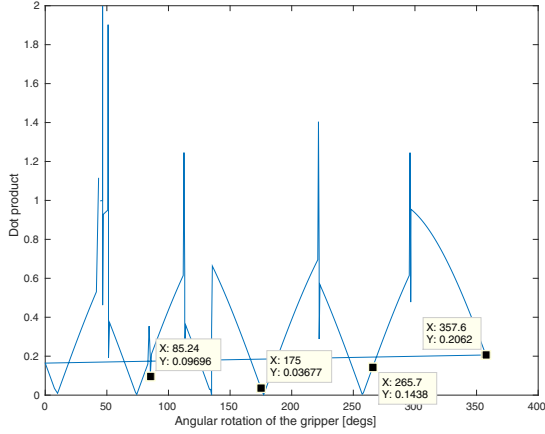


Fig. 9: Gripper angular displacement based on minimal dot-product contact points

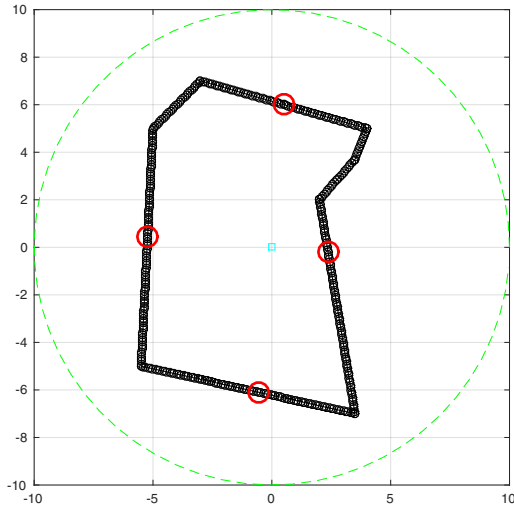


Fig. 10: Optimal grasping points

MIPI of 5 Megapixels which acquires images in QVGA resolution (320x480) 120 FPS, such specification only corresponds for recording purposes.

B. Vision-based Optimal Grasping Points Detection

The current proposal lies within an undergoing project to realise fast aerial pick-and-place operations using an aerial hexarotor configuration having a parallel Delta robot. In general, it is expected that the vehicle recognize the object to be grasped and defines the corresponding trajectory. The current proposal deals with a preliminary section of the vision-based identification process. The actual proposal seeks to maximize the possibilities of a successful grasping during an aerial pick-and-place task.

The vision processing is performed by the microprocessor UDOO within an embedded Linux environment capable of executing C++ routines, OpenCV libraries in this case.

Since the grasping strategy requires the contour vector (x_c, y_c) of the image, the actual study uses the Canny edges

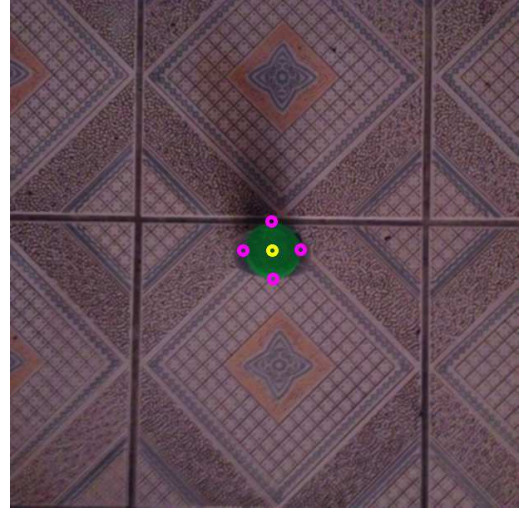


Fig. 11: Identification of the four optimal points for a circular shape

detector (CED) to extract the object's contour. The algorithm is executed while optimizing three main criteria.

- Noise robustness
- Accurate edges extraction
- Response unicity

Classically, the CED algorithm consist of four well-known steps:

1) *Filtering.*

A low-pass filter is applied to the image to reduce noise (image smoothing). The smoothing process require a Gaussian Mask $\mathcal{M} \in \mathbb{R}^{5 \times 5}$

2) *Gradient*

Since the edges of the images are distributed in different directions, it then possible to compute the gradient in the horizontal (G_x) and vertical (G_y) directions. From the latter allows to determine the magnitude and angle of the total edge gradient.

3) *Thresholding*

This process uses window of two thresholds values to get rid of edge's pixels caused by image parasitic effects (noise and color variation).

4) *Edge thinning*

The less representative pixels along the gradient direction are eliminated.

Figures 11-13 show the performance of the approach implementation. They demonstrate that four points are successfully obtained for three regular geometries, i.e. simulation results match with thos obtained from the experiments.

VI. CONCLUDING REMARKS AND FUTURE WORKS

In this paper we have presented preliminary results towards high-reactive grasping pick-and-place maneuvers. It was presented a novel concept of end-effector for aerial robots. Moreover, it is proposed an adapted-to-gripper grasping strategy based on the optimal identification of contact



Fig. 12: Identification of the four optimal points for a squared shape



Fig. 13: Identification of the four optimal points for an elliptical shape

points. In order to support our proposal, the simulation stage has validated the point-detection strategy. In this regard, the experimental stage has demonstrated the effectiveness of the proposed approach to find the optimal grasping profile for real objects. The future work consist in evaluate the grasping strategy mounted on the complete aerial grasping system (rotorcraft + parallel robot + aerogripper) which is under development.

REFERENCES

- [1] P. E. Pounds, D. Bersak, and A. Dollar, "Grasping from the air: Hovering capture and load stability," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 2491–2498.
- [2] G. J. Ghadiok, V. and W. Ren, "Autonomous indoor aerial gripping using a quadrotor," in *In 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012.
- [3] —, "On the design and development of attitude stabilization, vision-based navigation, and aerial gripping for a low-cost quadrotor."
- [4] J. W. Yeol and C.-H. Lin, "Development of multi-tentacle micro air vehicle," in *Unmanned Aircraft Systems (ICUAS), 2014 International Conference on*, May 2014, pp. 815–820.
- [5] J. Thomas, G. Loianno, K. Sreenath, and V. Kumar, "Toward Image Based Visual Servoing for Aerial Grasping and Perching," in *ICRA*. IEEE, 2014.
- [6] N. R. Kuntz and P. Y. Oh, "Towards autonomous cargo deployment and retrieval by an unmanned aerial vehicle using visual servoing," in *ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers, 2008, pp. 841–849.
- [7] M. Bernard, K. Kondak, and G. Hommel, "Load transportation system based on autonomous small size helicopters," *Aeronautical Journal*, vol. 114, no. 1153, pp. 191–198, 2010.
- [8] D. Mellinger, M. Shomin, and V. Kumar, "Control of quadrotors for robust perching and landing," in *Proceedings of the International Powered Lift Conference*, 2010, pp. 205–225.

Appendix B

Modeling and simulation of a bioreactor landfill

In this appendix, we attach the work developed within the framework of a collaboration with G. Dollé, O. Duran, N. Feyeux, E. Frénod and C. Prud'homme on the mathematical modeling and the numerical simulation of a bioreactor landfill. This research started during the 2015 edition of the CEMRACS (*Centre d'Été de Mathématique de Recherche Avancée en Calcul Scientifique*) at CIRM (*Centre International de Rencontres Mathématiques*) of Luminy, near Marseille.

CEMRACS 2015 was devoted to the subject of *Coupling Multi-Physics Models involving Fluids* and this work is issue of the project *Simulation of a Virtual BioReactor using Feel++* (SiViBiR++). The goal of the project is the development of a multiphysics model of the physical and chemical phenomena taking place inside a bioreactor landfill and leading to the production of methane. The main challenge is represented by the complexity of the industrial system which is characterized by lack of data and no *a priori* known optimal configuration. As the work presented in appendix A, the interest of this project resides in the development of a practical approach to a real-life industrial problem and in the correct handling of the constraints it introduces. This approach may turn out to be extremely useful when dealing with shape optimization problems in an industrial context.

MATHEMATICAL MODELING AND NUMERICAL SIMULATION OF A BIOREACTOR LANDFILL USING FEEL++ *

GUILLAUME DOLLÉ¹, OMAR DURAN², NELSON FEYEUX³, EMMANUEL FRÉNOT⁴,
MATTEO GIACOMINI^{5,6} AND CHRISTOPHE PRUD'HOMME¹

Abstract. In this paper, we propose a mathematical model to describe the functioning of a bioreactor landfill, that is a waste management facility in which biodegradable waste is used to generate methane. The simulation of a bioreactor landfill is a very complex multiphysics problem in which bacteria catalyze a chemical reaction that starting from organic carbon leads to the production of methane, carbon dioxide and water. The resulting model features a heat equation coupled with a non-linear reaction equation describing the chemical phenomena under analysis and several advection and advection-diffusion equations modeling multiphase flows inside a porous environment representing the biodegradable waste. A framework for the approximation of the model is implemented using FEEL++, a C++ open-source library to solve Partial Differential Equations. Some heuristic considerations on the quantitative values of the parameters in the model are discussed and preliminary numerical simulations are presented.

1. INTRODUCTION

Waste management and energy generation are two key issues in nowadays societies. A major research field arising in recent years focuses on combining the two aforementioned topics by developing new techniques to handle waste and to use it to produce energy. A very active field of investigation focuses on bioreactor landfills which are facilities for the treatment of biodegradable waste. The waste is accumulated in a humid environment and its degradation is catalyzed by bacteria. The main process taking place in a bioreactor landfill is the methane generation starting from the consumption of organic carbon due to waste decomposition. Several by-products appear during this reaction, including carbon dioxide and leachate, that is a liquid suspension containing particles of the waste material through which water flows.

* This work has been supported by the LMBA Université de Bretagne-Sud, the project PEPS Amies VirtualBioReactor and the private funding of See-d and Entreprise Charier.

The project is hosted on the facilities at CEMOSIS whose support is kindly acknowledged.

M. Giacomini is member of the DeFI team at Inria Saclay Île-de-France.

¹ Université de Strasbourg, IRMA UMR 7501, 7 rue René Descartes, 67084 Strasbourg, France. e-mail: dolle@math.unistra.fr; prudhomme@unistra.fr

² State University of Campinas, SP, Brazil. e-mail: omar@dep.fem.unicamp.br

³ MOISE team, INRIA Grenoble Rhône-Alpes. Université de Grenoble, Laboratoire Jean Kuntzmann, UMR 5224, Grenoble, France. e-mail: nelson.feyeux@imag.fr

⁴ Université de Bretagne-Sud, UMR 6205, LMBA, F-56000 Vannes, France. e-mail: emmanuel.frenod@univ-ubs.fr

⁵ CMAP, Inria, Ecole polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France.

e-mail: matteo.giacomini@polytechnique.edu

⁶ DRI Institut Polytechnique des Sciences Avancées, 63 Boulevard de Brandebourg, 94200 Ivry-sur-Seine, France.

Several works in the literature have focused on the study of bioreactor landfills but to the best of our knowledge none of them tackles the global multiphysics problem. On the one hand, [20, 21, 25] present mathematical approaches to the problem but the authors deal with a single aspect of the phenomenon under analysis focusing either on microbiota activity and leachates recirculation or on gas dynamic. On the other hand, this topic has been of great interest in the engineering community [3, 16, 17] and several studies using both numerical and experimental approaches are available in the literature. We refer the interested reader to the review paper [1] on this subject.

In this work, we tackle the problem of providing a mathematical model for the full multiphysics problem of methane generation inside a bioreactor landfill. Main goal is the development of a reliable model to simulate the long-time behavior of these facilities in order to be able to perform forecasts and process optimization [19]. This paper represents a preliminary study of the problem starting from the physics of the phenomena under analysis and provides a first set of equations to describe the methane generation inside a bioreactor landfill. In a more general framework, we aim to develop a model sufficiently accurate to be applied to an industrial context limiting at the same time the required computational cost. Thus, a key aspect of this work focused on the identification of the most important features of the functioning of a bioreactor landfill in order to derive the simplest model possible to provide an accurate description of the aforementioned methanogenic phenomenon. The proposed model has been implemented using FEEL++ and the resulting tool to numerically simulate the dynamic of a bioreactor landfill has been named SiViBiR++ which stands for *Simulation of a Virtual BioReactor using FEEL++*.

The rest of the paper is organized as follows. After a brief description of the physical and chemical phenomena taking place inside these waste management facilities (Section 2), in section 3 we present the fully coupled mathematical model of a bioreactor landfill. Section 4 provides details on the numerical strategy used to discretize the discussed model. Eventually, in section 5 preliminary numerical tests are presented and section 6 summarizes the results and highlights some future perspectives. In appendix A, we provide a table with the known and unknown parameters featuring our model.

2. WHAT IS A BIOREACTOR LANDFILL?

As previously stated, a bioreactor landfill is a facility for the treatment of biodegradable waste which is used to generate methane, electricity and hot water. Immediately after being deposited inside a bioreactor, organic waste begins to experience degradation through chemical reactions. During the first phase, degradation takes place via aerobic metabolic pathways, that is a series of concatenated biochemical reactions which occur within a cell in presence of oxygen and may be accelerated by the action of some enzymes. Thus bacteria begin to grow and metabolize the biodegradable material and complex organic structures are converted to simpler soluble molecules by hydrolysis.

The aerobic degradation is usually short because of the high demand of oxygen which may not be fulfilled in bioreactor landfills. Moreover, as more material is added to the landfill, the layers of waste tend to be compacted and the upper strata begin to block the flow of oxygen towards the lower parts of the bioreactor. Within this context, the dominant reactions inside the facility become anaerobic. Once the oxygen is exhausted, the bacteria begin to break the previously generated molecules down to organic acids which are readily soluble in water and the chemical reactions involved in the metabolism provide energy for the growth of population of microbiota.

After the first year of life of the facility, the anaerobic conditions and the soluble organic acids create an environment where the methanogenic bacteria can proliferate [28]. These bacteria become the major actors inside the landfill by using the end products from the first stage of degradation to drive the methane fermentation and convert them into methane and carbon dioxide. Eventually, the chemical reactions responsible for the generation of these gases gradually decrease until the material inside the landfill is inert (approximately after 40 years).

In this work, we consider the second phase of the degradation process, that is the methane fermentation during the anaerobic stage starting after the first year of life of the bioreactor.

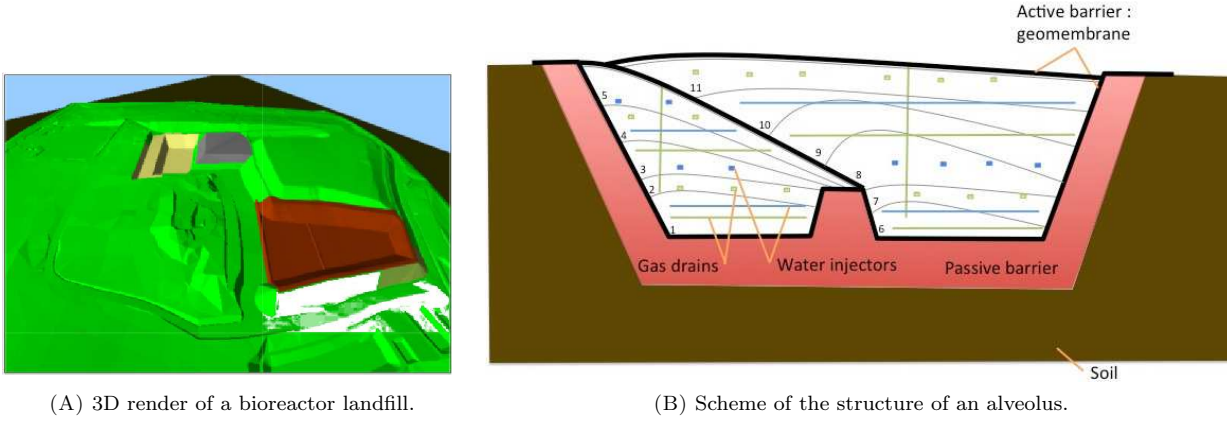


FIGURE 1. Structure of a bioreactor landfill and its composing alveoli. Image courtesy of Entreprise Charier <http://www.charier.fr>.

2.1. Structure of a bioreactor landfill

A bioreactor landfill counts several unit structures - named alveoli - as shown in the 3D rendering of a facility in Drouès, France (Fig. 1A). We focus on a single alveolus and we model it as a homogeneous porous medium in which the bulk material represents the solid waste whereas the void parts among the organic material are filled by a mixture of gases - mainly methane, carbon dioxide, oxygen, nitrogen and water vapor - and leachates, that is a liquid suspension based on water. For the rest of this paper, we will refer to our domain of interest by using indifferently the term alveolus and bioreactor, though the latter one is not rigorous from a modeling point of view.

Each alveolus is filled with several layers of biodegradable waste and the structure is equipped with a network of horizontal water injectors and production pipes respectively to allow the recirculation of leachates and to extract the gases generated by the chemical reactions. Moreover, each alveolus is isolated from the surrounding ground in order to prevent pollutant leaks and is covered by means of an active geomembrane. Figure 1B provides a schematic of an alveolus in which the horizontal pipes are organized in order to subdivide the structure in a cartesian-like way. Technical details on the construction and management of a bioreactor landfill are available in [8, 14].

2.2. Physical and chemical phenomena

Let us define the porosity ϕ as the fraction of void space inside the bulk material:

$$\phi = \frac{\text{Pore Volume}}{\text{Total Volume}}. \quad (2.1)$$

For biodegradable waste, we consider $\phi = 0.3$ as by experimental measurements in [26], whereas it is known in the literature that for generic waste the value drops to 0.1. Within this porous environment, the following phenomena take place:

- chemical reaction for the methane fermentation;
- heat transfer driven by the chemical reaction;
- transport phenomena of the gases;
- transport and diffusion phenomena of the leachates;

Here, we briefly provide some details about the chemical reaction for the methane generation, whereas we refer to section 3 for the description of the remaining phenomena and the derivation of the full mathematical model for the coupled system. As previously mentioned, at the beginning of the anaerobic stage the bacteria break the previously generated molecules down to organic acids, like the propionic acid CH_3CH_2COOH . This acid acts as a reacting term in the following reaction:



The microbiota activity drives the generation of methane (CH_4) and is responsible for the production of other by-products, mainly water (H_2O) and carbon dioxide (CO_2). As per equation (2.2), for each consumed mole of propionic acid - equivalently referred to as organic carbon with an abuse of notation -, two moles of methane are generated and three moles of water and one of carbon dioxide are produced as well.

Remark 2.1. We remark that in order for reaction (2.2) to take place, water has to be added to the propionic acid. This means that the bacteria can properly catalyze the reaction only if certain conditions on the temperature and the humidity of the waste are fulfilled. This paper presents a first attempt to provide a mathematical model of a bioreactor landfill, thus both the temperature and the quantity of water inside the facility will act as unknowns in the model (cf. section 3). In a more general framework, the proposed model will be used to perform long-time forecasts of the methane generation process and the temperature and water quantity will have the key role of control variables of the system.

3. MATHEMATICAL MODEL OF A BIOREACTOR LANDFILL

In this section, we describe the equations modeling the phenomena taking place inside an alveolus. As stated in the introduction, the final goal of the SiViBiR++ project is to control and optimize the functioning of a real bioreactor landfill, hence a simple model to account for the phenomena under analysis is sought. Within this framework, in this article we propose a first mathematical model to describe the coupled physical and chemical phenomena involved in the methanogenic fermentation. In the following sections, we will provide a detailed description of the chemical reaction catalyzed by the methanogenic bacteria, the evolution of the temperature inside the alveolus and the transport phenomena driven by the dynamic of a mixture of gases and by the liquid water. An extremely important aspect of the proposed model is the interaction among the variables at play and consequently the coupling among the corresponding equations. In order to reduce the complexity of the model and to keep the corresponding implementation in FEEL++ as simple as possible, some physical phenomena have been neglected. In the following subsections, we will detail the simplifying assumptions that allow to neglect some specific phenomena without degrading the reliability of the resulting model, by highlighting their limited impact on the global behavior of the overall system.

Let $\Omega \subset \mathbb{R}^3$ be an alveolus inside the landfill under analysis. We split the boundary $\partial\Omega$ of the computational domain into three non-void and non-overlapping regions Γ_t , Γ_b and Γ_l , representing respectively the membrane covering the top surface of the bioreactor, the base of the alveolus and the ground surrounding the lateral surface of the structure.

3.1. Consumption of the organic carbon

As previously stated, the functioning of a bioreactor landfill relies on the consumption of biodegradable waste by means of bacteria. From the chemical reaction in (2.2), we may derive a relationship between the concentration of bacteria b and the concentration of the consumable organic material which we denote by C^{org} . The activity of the bacteria takes place if some environmental conditions are fulfilled, namely the waste humidity and the bioreactor temperature. Let w_{\max} and T_{opt} be respectively the maximal quantity of water and the optimal temperature that allow the microbiota to catalyze the chemical reaction (2.2). We introduce the

following functions to model the metabolism of the microbiota:

$$\Psi_1(w) = w \max\left(0, 1 - \frac{w}{w_{\max}}\right), \quad \Psi_2(T) = \max\left(0, 1 - \frac{|T - T_{\text{opt}}|}{A_T}\right) \quad (3.1)$$

where A_T is the amplitude of the variation of the temperature tolerated by the bacteria. On the one hand, Ψ_1 models the fact that the bacterial activity is proportional to the quantity of liquid water - namely leachates - inside the bioreactor and it is prevented when the alveolus is flooded. On the other hand, according to Ψ_2 the microbiota metabolism is maximum when the current temperature equals T_{opt} and it stops when it exceeds the interval of admissible temperatures $[T_{\text{opt}} - A_T; T_{\text{opt}} + A_T]$.

Since the activity of the bacteria mainly consists in consuming the organic waste to perform reaction (2.2), it is straightforward to deduce a proportionality relationship between b and C^{org} . By combining the information in (3.1) with this relationship, we may derive the following law to describe the evolution of the concentration of bacteria inside the bioreactor:

$$\partial_t b \propto b C^{\text{org}} \Psi_1(w) \Psi_2(T) \quad (3.2)$$

and consequently, we get a proportionality relationship for the consumption of the biodegradable material C^{org} :

$$\partial_t C^{\text{org}} \propto -\partial_t b. \quad (3.3)$$

Let a_b and c_b be two proportionality constants associated respectively with (3.2) and (3.3). By integrating (3.3) in time and introducing the proportionality constant c_b , we get that the concentration of bacteria reads as

$$b(x, t) = b_0 + c_b(C_0^{\text{org}} - C^{\text{org}}(x, t)) \quad (3.4)$$

where $b_0 := b(\cdot, 0)$ and $C_0^{\text{org}} := C^{\text{org}}(\cdot, 0)$ are the initial concentrations respectively of bacteria and organic material inside the alveolus. Thus, by plugging (3.4) into (3.2) we get the following equation for the consumption of organic carbon between the instant $t = 0$ and the final time S^{fin} :

$$\begin{cases} (1 - \phi) \partial_t C^{\text{org}}(x, t) = -a_b b(x, t) C^{\text{org}}(x, t) \Psi_1(w(x, t)) \Psi_2(T(x, t)) , & \text{in } \Omega \times (0, S^{\text{fin}}] \\ C^{\text{org}}(\cdot, 0) = C_0^{\text{org}} , & \text{in } \bar{\Omega} \end{cases} \quad (3.5)$$

We remark that the organic material filling the bioreactor is only present in the bulk part of the porous medium and this is modeled by the factor $1 - \phi$ which features the information about the porosity of the environment. Moreover, we highlight that in equation (3.5) a non-linear reaction term appears and in section 4 we will discuss a strategy to deal with this non-linearity when moving to the Finite Element discretization.

For the sake of readability, from now on we will omit the dependency on the space and time variables in the notation for both the organic carbon and the concentration of bacteria.

3.2. Evolution of the temperature

The equation describing the evolution of the temperature T inside the bioreactor is the classical heat equation with a source term proportional to the consumption of bacteria. We consider the external temperature to be fixed by imposing Dirichlet boundary conditions on $\partial\Omega$.

Remark 3.1. Since we are interested in the long-time evolution of the system ($S^{\text{fin}} = 40$ years), the unit time interval is sufficiently large to allow daily variations of the temperature to be neglected. Moreover, we assume that the external temperature remains constant during the whole life of the bioreactor. From a physical point of view, this assumption is not realistic but we conjecture that only small fluctuations would arise by the relaxation of this hypothesis. A future improvement of the model may focus on the integration of dynamic boundary conditions in order to model seasonal changes of the external temperature.

The resulting equation for the temperature reads as follows:

$$\begin{cases} \partial_t T(x, t) - k_T \Delta T(x, t) = -c_T \partial_t C^{\text{org}}(x, t) , & \text{in } \Omega \times (0, S^{\text{fin}}] \\ T(x, \cdot) = T_m , & \text{on } \Gamma_t \times (0, S^{\text{fin}}] \\ T(x, \cdot) = T_g , & \text{on } \Gamma_b \cup \Gamma_l \times (0, S^{\text{fin}}] \\ T(\cdot, 0) = T_0 , & \text{in } \bar{\Omega} \end{cases} \quad (3.6)$$

where k_T is the thermal conductivity of the biodegradable waste and c_T is a scaling factor that accounts for the heat transfer due to the chemical reaction catalyzed by the bacteria. The values T_m , T_g and T_0 respectively represent the external temperature on the membrane Γ_t , the external temperature of the ground $\Gamma_b \cup \Gamma_l$ and the initial temperature inside the bioreactor.

For the sake of readability, from now on we will omit the dependency on the space and time variables in the notation of the temperature.

3.3. Velocity field of the gas

In order to model the velocity field of the gas inside the bioreactor, we have to introduce some assumptions on the physics of the problem. First of all, we assume the gas to be incompressible. This hypothesis stands if a very slow evolution of the mixture of gases takes place and this is the case for a bioreactor landfill in which the methane fermentation gradually decreases along the 40 years lifetime of the facility. Additionally, the decompression generated by the extraction of the gases through the pipes is negligible due to the weak gradient of pressure applied to the production system. Furthermore, we assume low Reynolds and low Mach numbers for the problem under analysis: this reduces to having a laminar slow flow which, as previously stated, is indeed the dynamic taking place inside an alveolus. Eventually, we neglect the effect due to the gravity on the dynamic of the mixture of gases: owing to the small height of the alveolus (approximately 90 m), the variation of the pressure in the vertical direction due to the gravity is limited and in our model we simplify the evolution of the gas by neglecting the hydrostatic component of the pressure.

Under the previous assumptions, the behavior of the gas mixture inside a bioreactor landfill may be described by a mass balance equation coupled with a Darcy's law

$$\begin{cases} \nabla \cdot u = 0 , & \text{in } \Omega \\ u = -\nabla p , & \text{in } \Omega \end{cases} \quad (3.7)$$

where $p := \frac{D}{\phi \mu_{\text{gas}}} P$, D is the permeability of the porous medium, ϕ its porosity and μ_{gas} the gas viscosity whereas P is the pressure inside the bioreactor. In (3.7), the incompressibility assumption has been expressed by stating that the gas flow is isochoric, that is the velocity is divergence-free. This equation is widely used in the literature to model porous media (cf. e.g. [10, 18]) and provides a coherent description of the phenomenon under analysis in the bioreactor landfill. As a matter of fact, it is reasonable to assume that the density of the gas mixture is nearly constant inside the domain, owing to the weak gradient of pressure applied to extract the gas via the production system and to the slow rate of methane generation via the fermentation process, that lasts approximately 40 years.

To fully describe the velocity field, the effect of the production system that extracts the gases from the bioreactor has to be accounted for. We model the production system as a set of N_g cylinders Θ_g^i 's thus the effect of the gas extraction on each pipe results in a condition on the outgoing flow. Let $J_{\text{out}} > 0$ be the mass flow rate exiting from the alveolus through each production pipe. The system of equations (3.7) is coupled with the following conditions on the outgoing normal flow on each drain used to extract the gas:

$$\int_{(\partial \Theta_g^i)^n} (C^{\text{dx}} + M + O + N + h) u \cdot n \, d\sigma = J_{\text{out}} \quad \forall i = 1, \dots, N_g. \quad (3.8)$$

In (3.8), n is the outward normal vector to the surface, $(\partial\Theta_g^i)^n$ is the part of the boundary of the cylinder Θ_g^i which belongs to the lateral surface of the alveolus itself and the term $(C^{dx} + M + O + N + h)$ represents the total concentration of the gas mixture starring carbon dioxide, methane, oxygen, nitrogen and water vapor. Since the cross sectional area of the pipes belonging to the production system is negligible with respect to the size of the overall alveolus, we model these drains as 1D lines embedded in the 3D domain. Owing to this, in the following subsection we present a procedure to integrate the information (3.8) into a source term named F^{out} in order to simplify the problem that describes the dynamic of the velocity field inside a bioreactor landfill.

Remark 3.2. According to conditions (3.8), the velocity u depends on the concentrations of the gases inside the bioreactor, thus is a function of both space and time. Nevertheless, the velocity field at each time step is independent from the previous ones and is only influenced by the distribution of gases inside the alveolus. For this reason, we neglect the dependency on the time variable and we consider u being only a function of space.

3.3.1. The source term F^{out}

As previously stated, each pipe Θ_g^i is modeled as a cylinder of radius R and length L . Hence, the cross sectional area $(\partial\Theta_g^i)^n$ and the lateral surface $(\partial\Theta_g^i)^l$ respectively measure πR^2 and $2\pi RL$. We assume the gas inside the cylinder to instantaneously exit the alveolus through its boundary $(\partial\Theta_g^i)^n$, that is the outgoing flow (3.8) is equal to the flow entering the drain through its lateral surface. Thus we may neglect the gas dynamic inside the pipe and (3.8) may be rewritten as

$$\int_{(\partial\Theta_g^i)^n} (C^{dx} + M + O + N + h) u \cdot n \, d\sigma = \int_{(\partial\Theta_g^i)^l} (C^{dx} + M + O + N + h) u \cdot n \, d\sigma = J_{out} \quad \forall i = 1, \dots, N_g. \quad (3.9)$$

Moreover, under the hypothesis that the quantity of gas flowing from the bioreactor to the inside of the cylinder Θ_g^i is uniform over its lateral surface, that is the same gas mixture surrounds the drain in all the points along its dominant size, we get

$$\left(\int_{(\partial\Theta_g^i)^l} (C^{dx} + M + O + N + h) \, d\sigma \right) u \cdot n = J_{out} \quad \text{on } (\partial\Theta_g^i)^l \quad \forall i = 1, \dots, N_g. \quad (3.10)$$

We remark that gas densities may be considered uniform along the perimeter of the cylinder only if the latter is small enough, that is the aforementioned assumption is likely to be true if the radius of the pipe is small in comparison with the size of the alveolus. Within this framework, (3.10) reduces to

$$u \cdot n = \frac{J_{out}}{2\pi R \int_{\mathfrak{L}^i} (C^{dx} + M + O + N + h) \, dl} \quad \text{on } (\partial\Theta_g^i)^l \quad \forall i = 1, \dots, N_g \quad (3.11)$$

where \mathfrak{L}^i is the centerline associated with the cylinder Θ_g^i . By coupling (3.7) with (3.11), we get the following PDE to model the velocity field:

$$\begin{cases} -\Delta p = 0, & \text{in } \Omega \\ \nabla p \cdot n = -\frac{J_{out}}{2\pi R \int_{\mathfrak{L}^i} (C^{dx} + M + O + N + h) \, dl}, & \text{on } (\partial\Theta_g^i)^l \end{cases} \quad \forall i = 1, \dots, N_g \quad (3.12)$$

Let us consider the variational formulation of problem (3.12): we seek $p \in H^1(\Omega)$ such that

$$\int_{\Omega} \nabla p \cdot \nabla \delta p \, dx = \sum_{i=1}^{N_g} -\frac{J_{out}}{2\pi R \int_{\mathfrak{L}^i} (C^{dx} + M + O + N + h) \, dl} \int_{(\partial\Theta_g^i)^l} \delta p \, d\sigma \quad \forall \delta p \in C_0^1(\Omega). \quad (3.13)$$

We may introduce the term F^{out} as the limit when R tends to zero of the right-hand side of (3.13):

$$F^{\text{out}} := \sum_{i=1}^{N_g} - \frac{J_{\text{out}}}{\int_{\mathfrak{L}^i} (C^{\text{dx}} + M + O + N + h) dl} \delta_{\mathfrak{L}^i} \quad (3.14)$$

where $\delta_{\mathfrak{L}^i}$ is a Dirac mass concentrated along the centerline \mathfrak{L}^i of the pipe Θ_g^i .

Hence, the system of equations describing the evolution of the velocity inside the alveolus may be written as

$$\begin{cases} \nabla \cdot u = F^{\text{out}} , & \text{in } \Omega \\ u = -\nabla p , & \text{in } \Omega \\ u \cdot n = 0 , & \text{on } \partial\Omega \end{cases} \quad (3.15)$$

where the right-hand side of the mass balance equation may be either (3.14) or a mollification of it.

3.4. Transport phenomena for the gas components

Inside a bioreactor landfill the pressure field is comparable to the external atmospheric pressure. This low-pressure does not provide the physical conditions for gases to liquefy. Hence, the gases are not present in liquid phase and solely the dynamic of the gas phases has to be accounted for. Within this framework, in section 3.5 we consider the case of water for which phase transitions driven by heat transfer phenomena are possible, whereas in the current section we focus on the remaining gases (i.e. oxygen, nitrogen, methane and carbon dioxide) which solely exist in gas phase.

Let u be the velocity of the gas mixture inside the alveolus. We consider a generic gas whose concentration inside the bioreactor is named G . The evolution of G fulfills the classical pure advection equation:

$$\begin{cases} \phi \partial_t G(x, t) + u \cdot \nabla G(x, t) = F^G(x, t) , & \text{in } \Omega \times (0, S^{\text{fin}}] \\ G(\cdot, 0) = G_0 , & \text{in } \overline{\Omega} \end{cases} \quad (3.16)$$

where ϕ is again the porosity of the waste. The source term $F^G(x, t)$ depends on the gas and will be detailed in the following subsections.

3.4.1. The case of oxygen and nitrogen

We recall that the oxygen concentration is named O , whereas the nitrogen one is N . Neither of these components appears in reaction (2.2) thus the associated source terms are $F^O(x, t) = F^N(x, t) = 0$. The resulting equations (3.17) are closed by the initial conditions $O(\cdot, 0) = O_0$ and $N(\cdot, 0) = N_0$.

$$\begin{aligned} \phi \partial_t O + u \cdot \nabla O &= 0 \\ \phi \partial_t N + u \cdot \nabla N &= 0 \end{aligned} \quad (3.17)$$

Both the oxygen and the nitrogen are extracted by the production system thus their overall concentration may be negligible with respect to the quantity of carbon dioxide and methane inside the alveolus. Hence, for the rest of this paper we will neglect equations (3.17) by considering $O(x, t) \simeq O_0 \simeq 0$ and $N(x, t) \simeq N_0 \simeq 0$.

3.4.2. The case of methane and carbon dioxide

As previously stated, (2.2) describes the methanogenic fermentation that starting from the propionic acid drives the production of methane, having carbon dioxide as by-product. Equation (3.16) stands for both the methane M and the carbon dioxide C^{dx} . For these components, the source terms have to account for the production of gas starting from the transformation of biodegradable waste. Thus, the source terms are

proportional to the consumption of the quantity C^{org} through some constants c_M and c_C specific to the chemical reaction and the component:

$$F^j(x, t) = -c_j \partial_t C^{\text{org}} \quad , \quad j = M, C$$

In a similar fashion as before, the resulting equations read as

$$\begin{aligned} \phi \partial_t M + u \cdot \nabla M &= -c_M \partial_t C^{\text{org}} \\ \phi \partial_t C^{\text{dx}} + u \cdot \nabla C^{\text{dx}} &= -c_C \partial_t C^{\text{org}} \end{aligned} \quad (3.18)$$

and they are coupled with appropriate initial conditions $M(\cdot, 0) = M_0$ and $C^{\text{dx}}(\cdot, 0) = C_0^{\text{dx}}$.

3.5. Dynamic of water vapor and liquid water

Inside a bioreactor landfill, water exists both in vapor and liquid phase. Let h be the concentration of water vapor and w the one of liquid water. The variation of temperature responsible for phase transitions inside the alveolus is limited, whence we do not consider a two-phase flow for the water but we describe separately the dynamics of the gas and liquid phases of the fluid. On the one hand, the water vapor inside the bioreactor landfill evolves as the gases presented in section 3.4: it is produced by the chemical reaction (2.2), it is transported by the velocity field u and is extracted via the pipes of the production system; as previously stated, no effect of the gravity is accounted for. This results in a pure advection equation for h . On the other hand, the dynamic of the liquid water may be schematized as follows: it flows in through the injector system at different levels of the alveolus, is transported by a vertical field u_w due to the effect of gravity and is spread within the porous medium. The resulting governing equation for w is an advection-diffusion equation. Eventually, the phases h and w are coupled by a source term that accounts for phase transitions.

Owing to the different nature of the phenomena under analysis and to the limited rate of heat transfer inside a bioreactor landfill, in the rest of this section we will describe separately the equations associated with the dynamics of the water vapor and the liquid water, highlighting their coupling due to the phase transition phenomena.

3.5.1. Phase transitions

Two main phenomena are responsible for the production of water vapor inside a bioreactor landfill. On the one hand, vapor is a product of the chemical reaction (2.2) catalyzed by the microbiota during the methanogenic fermentation process. On the other hand, heat transfer causes part of the water vapor to condensate and part of the liquid water to evaporate.

Let us define the vapor pressure of water P^{vp} inside the alveolus as the pressure at which water vapor is in thermodynamic equilibrium with its condensed state. Above this critical pressure, water vapor condenses, that is it turns to the liquid phase. This pressure is proportional to the temperature T and may be approximated by the following Rankine law:

$$P^{\text{vp}}(T) = P_0 \exp \left(s_0 - \frac{s_1}{T} \right) \quad (3.19)$$

where P_0 is a reference pressure, s_0 and s_1 are two constants known by experimental results and T is the temperature measured in Kelvin. If we restrict to a range of moderate temperatures, we can approximate the exponential in (3.19) by means of a linear law. Let H_0 and H_1 be two known constants, we get

$$P^{\text{vp}}(T) \simeq H_0 + H_1 T. \quad (3.20)$$

Let P^h be the partial pressure of the water vapor inside the gas mixture. We can compute P^h multiplying the total pressure p by a scaling factor representing the ratio of water vapor inside the gas mixture:

$$P^h = \frac{h}{C^{\text{dx}} + M + O + N + h} p. \quad (3.21)$$

The phase transition process features two different phenomena. On the one hand, when the pressure P^h is higher than the vapor pressure of water P^{vp} the vapor condensates. By exploiting (3.21) and (3.20), the condition $P^h > P^{\text{vp}}(T)$ may be rewritten as

$$h - H(T) > 0 \quad , \quad H(T) := (C^{\text{dx}} + M + O + N + h) \frac{H_0 + H_1 T}{p}.$$

We assume the phase transition to be instantaneous, thus the condensation of water vapor may be expressed through the following function

$$F^{\text{cond}} := c_{h \rightarrow w} \max(h - H(T), 0) \quad (3.22)$$

where $c_{h \rightarrow w}$ is a scaling factor. As per (3.22), the production of vapor from liquid water is 0 as soon as the concentration of vapor is larger than the threshold $H(T)$, that is the air is saturated.

In a similar fashion, we may model the evaporation of liquid water. When P^h is below the vapor pressure of water P^{vp} - that is $h - H(T) < 0$ - part of the liquid water generates vapor. The evaporation rate is proportional to the difference $P^{\text{vp}}(T) - P^h$ and to the quantity of liquid water w available inside the alveolus. Hence, the evaporation of liquid water is modeled by the following expression

$$F^{\text{evap}} := c_{w \rightarrow h} \max(H(T) - h, 0)w. \quad (3.23)$$

Remark 3.3. Since the quantity of water vapor inside a bioreactor landfill is negligible, we assume that the evaporation process does not significantly affect the dynamic of the overall system. Hence, in the rest of this paper, we will neglect this phenomenon by modeling only the condensation (3.22).

3.5.2. The case of water vapor

The dynamic of the water vapor may be modeled using (3.16) as for the other gases. In this case, the source term has to account for both the production of water vapor due to the chemical reaction (2.2) and its decrease as a consequence of the condensation phenomenon:

$$F^h(x, t) := -c_h \partial_t C^{\text{org}} - F^{\text{cond}}$$

where c_h is a scaling factor describing the relationship between the consumption of organic carbon and the generation of water vapor. The resulting advection equation reads as

$$\phi \partial_t h + u \cdot \nabla h = -c_h \partial_t C^{\text{org}} - F^{\text{cond}} \quad (3.24)$$

and it is coupled with the initial condition $h(\cdot, 0) = h_0$.

3.5.3. The case of liquid water

The liquid water inside the bioreactor is modeled by an advection-diffusion equation in which the drift term is due to the gravity, that is the transport phenomenon is mainly directed in the vertical direction and is associated with the liquid flowing downward inside the alveolus.

$$\begin{cases} \phi \partial_t w(x, t) + u_w \cdot \nabla w(x, t) - k_w \Delta w(x, t) = F^{\text{cond}}(x, t) , & \text{in } \Omega \times (0, S^{\text{fin}}] \\ k_w \nabla w(x, t) \cdot n = 0 \text{ and } u_w \cdot n = 0 , & \text{on } \Gamma_t \cup \Gamma_l \times (0, S^{\text{fin}}] \\ k_w \nabla w(x, t) \cdot n = 0 , & \text{on } \Gamma_b \times (0, S^{\text{fin}}] \\ w(\cdot, 0) = w_0 , & \text{in } \overline{\Omega} \end{cases} \quad (3.25)$$

where $u_w := (0, 0, -\|u_w\|)^T$ is the vertical velocity of the water and k_w is its diffusion coefficient. The right-hand side of the first equation accounts for the water production by condensation as described in section 3.5.1. On the one hand, the free-slip boundary conditions on the lateral and top surfaces allow water to slide but

prevent its exit, that is the top and lateral membranes are waterproof. On the other hand, the homogeneous Neumann boundary condition on the bottom of the domain describes the ability of the water to flow through this membrane. These conditions are consistent with the impermeability of the geomembranes and with the recirculation of leachates which are extracted when they accumulate in the bottom part of the alveolus and are reinjected in the upper layers of the waste management facility.

Remark 3.4. It is well-known in the literature that the evolution of an incompressible fluid inside a given domain is described by the Navier-Stokes equation. In (3.25), we consider a simplified version of the aforementioned equation by linearizing the inertial term. As previously stated, the dynamic of the fluids inside the bioreactor landfill is extremely slow and we may assume a low Reynolds number regime for the water as well. Under this assumption, the transfer of kinetic energy in the turbulent cascade due to the non-linear term of the Navier-Stokes equation may be neglected. Moreover, by means of a linearization of the inertial term $(w \cdot \nabla)w$, the transport effect is preserved and the resulting parabolic advection-diffusion problem (3.25) may be interpreted as an unsteady version of the classical Oseen equation [9].

Remark 3.5. Equation (3.25) may be furtherly interpreted as a special advection equation modeling the transport phenomenon within a porous medium. As a matter of fact, the diffusion term $-k_w \Delta w$ accounts for the inhomogeneity of the environment in which the water flows and describes the fact that the liquid spreads in different directions while flowing downwards due to the encounter of blocking solid material along its path. The distribution of the liquid into different directions is random and is mainly related to the nature of the surrounding environment thus we consider an isotropic diffusion tensor k_w . The aforementioned equation is widely used (cf. e.g. [27]) to model flows in porous media and is strongly connected with the description of the porous environment via the Darcy's law introduced in section 3.3.

Within the framework of our problem, the diffusion term is extremely important since it models the spread of water and leachates inside the bioreactor landfill and the consequent humidification of the whole alveolus and not solely of the areas neighboring the injection pipes.

Eventually, problem (3.25) is closed by a set of conditions that describe the injection of liquid water and leachates through N_w pipes Θ_w^i 's. As previously done for the production system, we model each injector as a cylinder of radius R and length L and we denote by $(\partial\Theta_w^i)^n$ and $(\partial\Theta_w^i)^l$ respectively the part of the boundary of the cylinder which belongs to the boundary of the bioreactor and its lateral surface. The aforementioned inlet condition reads as

$$\int_{(\partial\Theta_w^i)^n} k_w \nabla w \cdot n \, d\sigma = -J_{\text{in}} \quad \forall i = 1, \dots, N_w$$

where $J_{\text{in}} > 0$ is the mass flow rate entering the alveolus through each injector. As for the production system in section 3.3.1, we may now integrate this condition into a source term for equation (3.25). Under the assumption that the flow is instantaneously distributed along the whole cylinder in a uniform way, we get

$$\int_{(\partial\Theta_w^i)^n} k_w \nabla w \cdot n \, d\sigma = - \int_{(\partial\Theta_w^i)^l} k_w \nabla w \cdot n \, d\sigma = -J_{\text{in}} \quad \forall i = 1, \dots, N_w.$$

Consequently the condition on each injector reads as

$$k_w \nabla w \cdot n = \frac{J_{\text{in}}}{2\pi RL} \quad \text{on } (\partial\Theta_w^i)^l \quad \forall i = 1, \dots, N_w$$

and we obtain the following source term F^{in}

$$F^{\text{in}} := \sum_{i=1}^{N_w} \frac{J_{\text{in}}}{L} \delta_{\Sigma^i} \quad (3.26)$$

where $\delta_{\mathfrak{L}^i}$ is a Dirac mass concentrated along the centerline \mathfrak{L}^i of the pipe Θ_w^i . Hence, the resulting dynamic of the liquid water inside an alveolus is modeled by the following PDE:

$$\begin{cases} \phi \partial_t w + u_w \cdot \nabla w - k_w \Delta w = F^w, & \text{in } \Omega \times (0, S^{\text{fin}}] \\ k_w \nabla w \cdot n = 0 \text{ and } u_w \cdot n = 0, & \text{on } \Gamma_t \cup \Gamma_l \times (0, S^{\text{fin}}] \\ k_w \nabla w \cdot n = 0, & \text{on } \Gamma_b \times (0, S^{\text{fin}}] \end{cases} \quad (3.27)$$

with $F^w := F^{\text{cond}} + F^{\text{in}}$ and the initial condition $w(\cdot, 0) = w_0$. By analyzing the right-hand side of equation (3.27), we remark that neglecting the effect of evaporation in the phase transition allows to decouple the dynamics of liquid water and water vapor. Moreover, as previously stated for equation (3.15), F^{in} may be chosen either according to definition (3.26) or by means of an appropriate mollification.

4. NUMERICAL APPROXIMATION OF THE COUPLED SYSTEM

This section is devoted to the description of the numerical strategies used to discretize the fully coupled model of the bioreactor landfill introduced in section 3. We highlight that one of the main difficulties of the presented model is the coupling of all the equations and the multiphysics nature of the problem under analysis. Here we propose a first attempt to discretize the full model by introducing an explicit coupling of the equations, that is by considering the source term in each equation as function of the variables at the previous iteration.

4.1. Geometrical model of an alveolus

As previously stated, a bioreactor landfill is composed by several alveoli. Each alveolus may be modeled as an independent structure obtained starting from a cubic reference domain (Fig. 2A) to which pure shear transformations are applied (Fig. 2B-2C). For example, the pure lateral shear in figure 2B allows to model

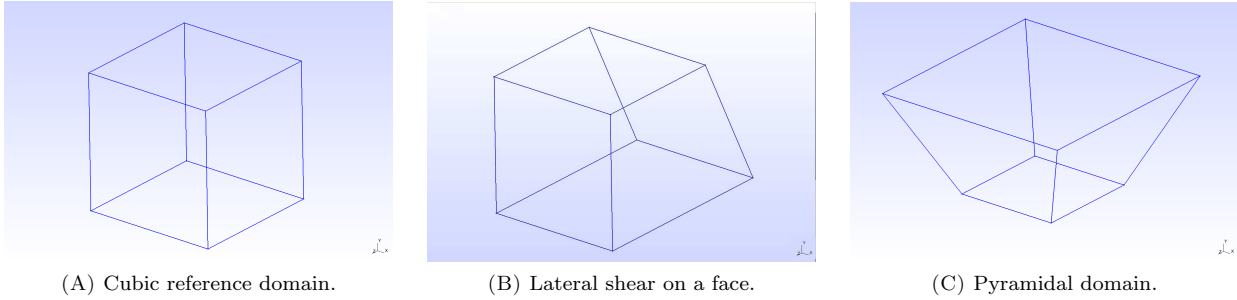


FIGURE 2. Reference domain for an alveolus and admissible transformations.

the left-hand side alveolus in figure 1B whereas the right-hand side one may be geometrically approximated by means of the pyramid in figure 2C.

In order to model the network of the water injectors and the one of the drains extracting the gas, the geometrical domains in figure 2 are equipped with a cartesian distribution of horizontal lines, the 1D model being justified by the assumption in section 3.

4.2. Finite Element approximations of the organic carbon and heat equations

Both equations (3.5) and (3.6) are discretized using Lagrangian Finite Element functions. In particular, the time derivative is approximated by means of an implicit Euler scheme, whereas the basis functions for the spatial discretization are the classical \mathbb{P}_k Finite Element functions of degree k .

Let $t = t_n$. We consider the following quantities at time t_n as known variables: $C_n^{\text{org}} := C^{\text{org}}(x, t_n)$, $T_n :=$

$T(x, t_n)$ and $w_n := w(x, t_n)$. The consumption of organic carbon is described by equation (3.5) coupled with equation (3.4) for the dynamic of the bacteria. At each time step, we seek $C_{n+1}^{\text{org}} \in H^1(\Omega)$ such that

$$\int_{\Omega} (1 - \phi) \frac{C_{n+1}^{\text{org}} - C_n^{\text{org}}}{\Delta t} \delta C \, dx = - \int_{\Omega} a_b C_{n+1}^{\text{org}} [b_0 + c_b(C_0^{\text{org}} - C_n^{\text{org}})] \Psi_1(w_n) \Psi_2(T_n) \delta C \, dx \quad \forall \delta C \in H^1(\Omega)$$

We remark that in the previous equation the non-linear reaction term has been handled in a semi-implicit way by substituting $(C_{n+1}^{\text{org}})^2$ by $C_{n+1}^{\text{org}} C_n^{\text{org}}$ in the right-hand side. Hence, the bilinear and linear forms associated with the variational formulation at $t = t_n$ respectively read as

$$a_{C^{\text{org}}}(C_{n+1}^{\text{org}}, \delta C) = \int_{\Omega} A_C C_{n+1}^{\text{org}} \delta C \, dx \quad , \quad l_{C^{\text{org}}}(\delta C) = \int_{\Omega} (1 - \phi) C_n^{\text{org}} \delta C \, dx \quad (4.1)$$

where $A_C = (1 - \phi) + \Delta t a_b [b_0 + c_b(C_0^{\text{org}} - C_n^{\text{org}})] \Psi_1(w_n) \Psi_2(T_n)$ and $a_{C^{\text{org}}}(C_{n+1}^{\text{org}}, \delta C) = l_{C^{\text{org}}}(\delta C) \, \forall \delta C \in H^1(\Omega)$.

In a similar fashion, we derive the variational formulation of the heat equation and at $t = t_n$ we seek $T_{n+1} \in H^1(\Omega)$ such that $T_{n+1}|_{\Gamma_t} = T_m$, $T_{n+1}|_{\Gamma_b \cup \Gamma_1} = T_g$ and $a_T(T_{n+1}, \delta T) = l_T(\delta T) \, \forall \delta T \in H_0^1(\Omega)$, where

$$a_T(T_{n+1}, \delta T) = \int_{\Omega} T_{n+1} \delta T \, dx + \Delta t \int_{\Omega} k_T \nabla T_{n+1} \cdot \nabla \delta T \, dx, \quad (4.2)$$

$$l_T(\delta T) = \int_{\Omega} T_n \delta T \, dx - \int_{\Omega} c_T (C_{n+1}^{\text{org}} - C_n^{\text{org}}) \delta T \, dx. \quad (4.3)$$

Remark 4.1. In (4.2) and (4.3), we evaluate the time derivative $\partial_t C^{\text{org}}$ at $t = t_n$, that is we consider the current value and not the previous one as stated at the beginning of this section. This is feasible because the solution of problem (4.1) precedes the one of the heat equation thus the value C_{n+1}^{org} is known when solving (4.2)-(4.3).

Remark 4.2. In (4.3) we assume that the same time discretization is used for both the organic carbon and the temperature. If this is not the case, the second term in the linear form $l_T(\cdot)$ would feature a scaling factor $\frac{\Delta t_T}{\Delta t_C}$, the numerator being the time scale associated with the temperature and the denominator the one for the organic carbon. For the rest of this paper, we will assume that all the unknowns are approximated using the same time discretization.

By substituting C_n^{org} and T_n with their Finite Element counterparts $C_{h,n}^{\text{org}}$ and $T_{h,n}$ in (4.1), (4.2) and (4.3) we obtain the corresponding discretized equations for the organic carbon and the temperature.

4.3. Stabilized dual-mixed formulation of the velocity field

A good approximation of the velocity field is a key point for a satisfactory simulation of all the transport phenomena. In order for problem (3.15) to be well-posed, the following compatibility condition has to be fulfilled

$$\int_{\Omega} F^{\text{out}} = 0. \quad (4.4)$$

Nevertheless, (4.4) does not stand for the problem under analysis thus we consider a slightly modified version of problem (3.15) by introducing a small perturbation parameter $\lambda = \mathcal{O}(\ell_K)$, ℓ_K being the diameter of the element K of the triangulation \mathcal{T}_h :

$$\begin{cases} \nabla \cdot u + \lambda p = F^{\text{out}} , & \text{in } \Omega \\ u = -\nabla p , & \text{in } \Omega \\ u \cdot n = 0 , & \text{on } \partial\Omega \end{cases} \quad (4.5)$$

Hence the resulting problem (4.5) is well-posed even if (4.4) does not stand.

It is well-known in the literature [12] that classical discretizations of problem (4.5) by means of Lagrangian Finite Element functions lead to poor approximations of the velocity field. A widely accepted workaround relies on the derivation of mixed formulations in which a simultaneous approximation of pressure and velocity fields is performed by using different Finite Element spaces [24].

4.3.1. Dual-mixed formulation

Let $H(\text{div}) = \{v \in [L^2(\Omega)]^3 \text{ s.t. } \nabla \cdot v \in L^2(\Omega)\}$ and $H_0(\text{div}) = \{v \in H(\text{div}) \text{ s.t. } v \cdot n = 0 \text{ on } \partial\Omega\}$. The dual-mixed formulation of problem (4.5) is obtained by seeking $(u, p) \in H_0(\text{div}) \times L^2(\Omega)$ such that

$$\begin{cases} \int_{\Omega} \nabla \cdot u \delta p \, dx + \int_{\Omega} \lambda p \delta p \, dx = \int_{\Omega} F^{\text{out}} \delta p \, dx \\ \int_{\Omega} u \cdot \delta u \, dx - \int_{\Omega} p \nabla \cdot \delta u \, dx = 0 \end{cases}, \quad \forall (\delta u, \delta p) \in H_0(\text{div}) \times L^2(\Omega).$$

Hence, the bilinear and linear forms associated with the variational formulation of the problem respectively read as

$$a_{\text{vel}}(\{u, p\}, \{\delta u, \delta p\}) = \int_{\Omega} u \cdot \delta u \, dx - \int_{\Omega} p \nabla \cdot \delta u \, dx - \int_{\Omega} \nabla \cdot u \delta p \, dx - \int_{\Omega} \lambda p \delta p \, dx \quad (4.6)$$

$$l_{\text{vel}}(\{\delta u, \delta p\}) = - \int_{\Omega} F^{\text{out}} \delta p \, dx \quad (4.7)$$

To overcome the constraint due to the LBB compatibility condition that the Finite Element spaces have to fulfill [5], several stabilization approaches have been proposed in the literature over the years and in this work we consider a strategy inspired by the Galerkin Least-Squares method and known as CGLS [11].

4.3.2. Galerkin Least-Squares stabilization

The GLS formulation relies on adding one or more quantities to the bilinear form of the problem under analysis in order for the resulting bilinear form to be strongly consistent and stable. Let L be the abstract operator for the Boundary Value Problem $L\varphi = g$. We introduce the solution φ_h of the corresponding problem discretized via the Finite Element Method. The GLS stabilization term reads as

$$\mathcal{L}^{\text{GLS}}(\varphi_h, g; \psi_h) = d \int_{\Omega} (L\varphi_h - g) L\psi_h \, dx.$$

GLS formulation of Darcy's law

Following the aforementioned framework, we have

$$L_1(\{u, p\}) = g_1 \quad \text{with} \quad L_1(\{u, p\}) := u + \nabla p, \quad g_1 := 0.$$

Thus the GLS term associated with Darcy's law reads as

$$\mathcal{L}_1^{\text{GLS}}(\{u_h, p_h\}, g_1; \{\delta u_h, \delta p_h\}) = d_1 \int_{\Omega} (u_h + \nabla p_h) \cdot (\delta u_h + \nabla \delta p_h) \, dx. \quad (4.8)$$

GLS formulation of the mass balance equation

The equation describing the mass equilibrium may be rewritten as

$$L_2(\{u, p\}) = g_2 \quad \text{with} \quad L_2(\{u, p\}) := \nabla \cdot u + \lambda p, \quad g_2 := F^{\text{out}}.$$

Consequently, the Least-Squares stabilization term has the following form

$$\mathcal{L}_2^{\text{GLS}}(\{u_h, p_h\}, g_2; \{\delta u_h, \delta p_h\}) = \underbrace{d_2 \int_{\Omega} (\nabla \cdot u_h + \lambda p_h)(\nabla \cdot \delta u_h + \lambda \delta p_h) \, dx}_{\mathcal{L}_{2a}^{\text{GLS}}(\{u_h, p_h\}, g_2; \{\delta u_h, \delta p_h\})} - \underbrace{d_2 \int_{\Omega} F^{\text{out}}(\nabla \cdot \delta u_h + \lambda \delta p_h) \, dx}_{\mathcal{L}_{2l}^{\text{GLS}}(\{u_h, p_h\}, g_2; \{\delta u_h, \delta p_h\})}. \quad (4.9)$$

GLS formulation of the curl of Darcy's law

Let us consider the rotational component of Darcy's law. Since p is a scalar field, $\nabla \times (\nabla p) = 0$ and we get

$$L_3(\{u, p\}) = g_3 \quad \text{with} \quad L_3(\{u, p\}) := \nabla \times u, \quad g_3 := 0.$$

Thus, the GLS term associated with the curl of Darcy's law reads as

$$\mathcal{L}_3^{\text{GLS}}(\{u_h, p_h\}, g_3; \{\delta u_h, \delta p_h\}) = d_3 \int_{\Omega} (\nabla \times u_h)(\nabla \times \delta u_h) \, dx. \quad (4.10)$$

The stabilized CGLS dual-mixed formulation

The stabilized CGLS formulation arises by combining the previous terms. In particular, we consider the bilinear form (4.6), we subtract the Least-Squares stabilization (4.8) for Darcy's law and we add the corresponding GLS terms (4.9) and (4.10) for the mass balance equation and the curl of Darcy's law itself. In a similar fashion, we assemble the linear form for the stabilized problem, starting from (4.7). The resulting CGLS formulation of problem (4.5) has the following form:

$$\begin{aligned} a_{\text{CGLS}}(\{u_h, p_h\}, \{\delta u_h, \delta p_h\}) &= a_{\text{vel}}(\{u_h, p_h\}, \{\delta u_h, \delta p_h\}) - \mathcal{L}_1^{\text{GLS}}(\{u_h, p_h\}, g_1; \{\delta u_h, \delta p_h\}) \\ &\quad + \mathcal{L}_{2a}^{\text{GLS}}(\{u_h, p_h\}, g_2; \{\delta u_h, \delta p_h\}) + \mathcal{L}_3^{\text{GLS}}(\{u_h, p_h\}, g_3; \{\delta u_h, \delta p_h\}) \end{aligned} \quad (4.11)$$

$$l_{\text{CGLS}}(\{\delta u_h, \delta p_h\}) = l_{\text{vel}}(\{\delta u_h, \delta p_h\}) + \mathcal{L}_{2l}^{\text{GLS}}(\{u_h, p_h\}, g_2; \{\delta u_h, \delta p_h\}) \quad (4.12)$$

To accurately approximate problem (4.11)-(4.12), we consider the product space $RT_0 \times \mathbb{P}_0$, that is we use lowest-order Raviart-Thomas Finite Element for the velocity and piecewise constant functions for the pressure.

4.4. Streamline Upwind Petrov Galerkin for the dynamics of gases and liquid water

The numerical approximation of pure advection and advection-diffusion transient problems has to be carefully handled in order to retrieve accurate solutions. It is well-known in the literature [4] that classical Finite Element Method suffers from poor accuracy when dealing with steady-state advection and advection-diffusion problems and requires the introduction of numerical stabilization to construct a strongly consistent scheme. When moving to transient advection and advection-diffusion problems, time-space elements are the most natural setting to develop stabilized methods [15].

Let L_{ad} be the abstract operator to model an advection - respectively advection-diffusion - phenomenon. The resulting transient Boundary Value Problem may be written as

$$\phi \partial_t \varphi + L_{\text{ad}} \varphi = g_{\text{ad}}. \quad (4.13)$$

We consider the variational formulation of (4.13) by introducing the corresponding abstract bilinear form $B_{\text{ad}}(\varphi, \psi)$ which will be detailed in next subsections. Let φ_h be the solution of the discretized PDE via the Finite Element Method. The SUPG stabilization term for the transient problem reads as

$$\mathcal{L}^{\text{SUPG}}(\varphi_h, g_{\text{ad}}; \psi_h) = d \int_{\Omega} (\phi \partial_t \varphi_h + L_{\text{ad}} \varphi_h - g_{\text{ad}}) L_{\text{ad}}^{\text{SS}} \psi_h \, dx$$

where $L_{\text{ad}}^{\text{SS}}$ is the skew-symmetric part of the advection - respectively advection-diffusion - operator and d is a stabilization parameter constant in space and time. Let \mathcal{T}_h be the computational mesh that approximates the domain Ω and ℓ_K the diameter of each element $K \in \mathcal{T}_h$. We choose:

$$d := \frac{1}{2\|u\|_2} \max_{K \in \mathcal{T}_h} \ell_K. \quad (4.14)$$

Let $V_h := \{\psi_h \in \mathcal{C}(\bar{\Omega}) \text{ s.t. } \psi_h|_K \in \mathbb{P}_k(K) \ \forall K \in \mathcal{T}_h\}$ be the space of Lagrangian Finite Element of degree k , that is the piecewise polynomial functions of degree k on each element K of the mesh \mathcal{T}_h . The stabilized SUPG formulation of the advection - respectively advection-diffusion - problem (4.13) reads as follows: for all $t \in (0, S^{\text{fin}}]$ we seek $\varphi_h(t) \in V_h$ such that

$$\int_{\Omega} \phi \partial_t \varphi_h \psi_h \, dx + B_{\text{ad}}(\varphi_h, \psi_h) + \mathcal{L}^{\text{SUPG}}(\varphi_h, g_{\text{ad}}; \psi_h) = \int_{\Omega} g_{\text{ad}} \psi_h \, dx \quad \forall \psi_h \in V_h. \quad (4.15)$$

Concerning time discretization, it is well-known in the literature that implicit schemes tend to increase the overall computational cost associated with the solution of a PDE. Nevertheless, precise approximations of advection and advection-diffusion equations via explicit schemes usually require high-order methods and are subject to stability conditions that may be responsible of making computation unfeasible. On the contrary, good stability and convergence properties of implicit strategies make them an extremely viable option when dealing with complex - possibly coupled - phenomena and with equations featuring noisy parameters. In particular, owing to the coupling of several PDE's, the solution of the advection and advection-diffusion equations presented in our model turned out to be extremely sensitive to the choice of the involved parameters. Being the tuning of the unknown coefficients of the equations one of the main goal of the SiViBiR++ project, a numerical scheme unconditionally stable and robust to the choice of the discretization parameters is sought. Within this framework, we consider an implicit Euler scheme for the time discretization and we stick to low-order Lagrangian Finite Element functions for the space discretization. The numerical scheme arising from the solution of equation (4.15) by means of the aforementioned approximation is known to be stable and to converge quasi-optimally [6].

In the following subsections, we provide some details on the bilinear and linear forms involved in the discretization of the advection and advection-diffusion equations as well as on the formulation of the associated stabilization terms.

4.4.1. The case of gases

Let us consider a generic gas G whose Finite Element counterpart is named G_h . Within the previously introduced framework, we get

$$L_G G := u \cdot \nabla G, \quad g_G := F^G, \quad B_G(G, \delta G) = \int_{\Omega} (u \cdot \nabla G) \delta G \, dx.$$

Hence, the SUPG stabilization term reads as

$$\mathcal{L}_G^{\text{SUPG}}(G_h, g_G; \delta G_h) = d \int_{\Omega} (\phi \partial_t G_h + u \cdot \nabla G_h - F^G)(u \cdot \nabla \delta G_h) \, dx.$$

By introducing an implicit Euler scheme to approximate the time derivative in (4.15), we obtain the fully discretized advection problem for the gas G : at $t = t_n$ we seek $G_{h,n+1} \in V_h$ such that $a_G(G_{h,n+1}, \delta G_h) = l_G(\delta G_h) \ \forall \delta G_h \in V_h$, where

$$a_G(G_{h,n+1}, \delta G_h) = \int_{\Omega} \phi G_{h,n+1} (\delta G_h + d \, u \cdot \nabla \delta G_h) \, dx + \Delta t \int_{\Omega} u \cdot \nabla G_{h,n+1} (\delta G_h + d \, u \cdot \nabla \delta G_h) \, dx, \quad (4.16)$$

$$l_G(\delta G_h) = \int_{\Omega} \phi G_{h,n} (\delta G_h + d \, u \cdot \nabla \delta G_h) \, dx + \Delta t \int_{\Omega} F^G (\delta G_h + d \, u \cdot \nabla \delta G_h) \, dx. \quad (4.17)$$

Remark 4.3. As the authors highlight in [4], from a practical point of view the implementation of (4.16)-(4.17) may not be straightforward due to the non-symmetric mass matrix resulting from the discretization of the first term in (4.16).

By considering $F^O = F^N = 0$, we get the linear forms associated with the dynamic of oxygen and nitrogen:

$$\begin{aligned} l_O(\delta O_h) &= \int_{\Omega} \phi_{O,h,n}(\delta O_h + d \, u \cdot \nabla \delta O_h) \, dx, \\ l_N(\delta N_h) &= \int_{\Omega} \phi_{N,h,n}(\delta N_h + d \, u \cdot \nabla \delta N_h) \, dx. \end{aligned} \quad (4.18)$$

In a similar way, when $F^j = -c_j \partial_t C^{\text{org}}$ $j = M, C$, we obtain the linear forms for methane and carbon dioxide:

$$\begin{aligned} l_M(\delta M_h) &= \int_{\Omega} \phi_{M,h,n}(\delta M_h + d \, u \cdot \nabla \delta M_h) \, dx - \int_{\Omega} c_M(C_{h,n+1}^{\text{org}} - C_{h,n}^{\text{org}})(\delta M_h + d \, u \cdot \nabla \delta M_h) \, dx, \\ l_{C^{\text{dx}}}(\delta C_h) &= \int_{\Omega} \phi_{C,h,n}^{\text{dx}}(\delta C_h + d \, u \cdot \nabla \delta C_h) \, dx - \int_{\Omega} c_C(C_{h,n+1}^{\text{org}} - C_{h,n}^{\text{org}})(\delta C_h + d \, u \cdot \nabla \delta C_h) \, dx. \end{aligned} \quad (4.19)$$

Eventually, the dynamic of the water vapor is obtained when considering $F^h = -c_h \partial_t C^{\text{org}} - F^{\text{cond}}$:

$$l_h(\delta h_h) = \int_{\Omega} \phi_{h,n}(\delta h_h + d \, u \cdot \nabla \delta h_h) \, dx - \int_{\Omega} c_h(C_{h,n+1}^{\text{org}} - C_{h,n}^{\text{org}})(\delta h_h + d \, u \cdot \nabla \delta h_h) \, dx - \Delta t \int_{\Omega} F^{\text{cond}}(\delta h_h + d \, u \cdot \nabla \delta h_h) \, dx. \quad (4.20)$$

4.4.2. The case of liquid water

The dynamic of the liquid water being described by an advection-diffusion equation, the SUPG framework may be written in the following form:

$$L_w w := u_w \cdot \nabla w - k_w \Delta w, \quad g_w := F^{\text{cond}} + F^{\text{in}}, \quad B_w(w, \delta w) = \int_{\Omega} \left((u_w \cdot \nabla w) \delta w + k_w \nabla w \cdot \nabla \delta w \right) dx.$$

Hence the stabilization term reads as

$$\mathcal{L}_w^{\text{SUPG}}(w_h, g_w; \delta w_h) = d_w \int_{\Omega} (\phi \partial_t w_h + u_w \cdot \nabla w_h - k_w \Delta w_h - F^{\text{cond}} - F^{\text{in}})(u_w \cdot \nabla \delta w_h) \, dx$$

where d_w is obtained by substituting u_w in (4.14). We obtain the fully discretized problem in which at each $t = t_n$ we seek $w_{h,n+1} \in V_h$ such that

$$\begin{aligned} a_w(w_{h,n+1}, \delta w_h) &= l_w(\delta w_h) \quad \forall \delta w_h \in V_h \\ a_w(w_{h,n+1}, \delta w_h) &= \int_{\Omega} \phi w_{h,n+1}(\delta w_h + d_w \, u_w \cdot \nabla \delta w_h) \, dx + \Delta t \int_{\Omega} u_w \cdot \nabla w_{h,n+1}(\delta w_h + d_w \, u_w \cdot \nabla \delta w_h) \, dx \\ &\quad + \Delta t \int_{\Omega} k_w \nabla w_{h,n+1} \cdot \nabla \delta w_h \, dx - \Delta t \int_{\Omega} k_w \Delta w_{h,n+1}(d_w \, u_w \cdot \nabla \delta w_h) \, dx, \end{aligned} \quad (4.21)$$

$$l_w(\delta w_h) = \int_{\Omega} \phi w_{h,n}(\delta w_h + d_w \, u_w \cdot \nabla \delta w_h) \, dx + \Delta t \int_{\Omega} (F^{\text{cond}} + F^{\text{in}})(\delta w_h + d_w \, u_w \cdot \nabla \delta w_h) \, dx. \quad (4.22)$$

5. NUMERICAL RESULTS

In this section we present some preliminary numerical simulations to test the proposed model. The SiViBiR++ project implements the discussed numerical methods for the approximation of the phenomena inside a bioreactor landfill. It is based on a C++ library named FEEL++ which provides a framework to solve PDE's using the Finite Element Method [23].

5.1. Feel++

FEEL++ stands for *Finite Element Embedded Language in C++* and is a C++ library for the solution of Partial Differential Equations using generalized Galerkin methods. It provides a framework for the implementation of advanced numerical methods to solve complex systems of PDE's. The main advantage of FEEL++ for the applied mathematicians and engineers community relies on its design based on the Domain Specific Embedded Language (DSEL) approach [22]. This strategy allows to decouple the difficulties encountered by the scientific community when dealing with libraries for scientific computing. As a matter of fact, DSEL provides a high-level language to handle mathematical methods without losing abstraction. At the same time, due to the continuing evolution of the state-of-the-art techniques in computer science (e.g. new standards in programming languages, parallel architectures, etc.) the choice of the proper tools in scientific computing may prove very difficult. This is even more critical for scientists who are not specialists in computer science and have to reach a compromise between user-friendly interfaces and high performances.

FEEL++ proposes a solution to hide these difficulties behind a user-friendly language featuring a syntax that mimics the mathematical formulation by using a much more common low-level language, namely C++. Moreover, FEEL++ integrates the latest C++ standard - currently C++14 - and provides seamless parallel tools to handle mathematical operations such as projection, integration or interpolation through C++ keywords. FEEL++ is regularly tested on High Performance Computing facilities such as the PRACE research infrastructures (e.g. Tier-0 supercomputer CURIE, Supermuc, etc.) via multidisciplinary projects mainly gathered in the FEEL++ consortium.

The embedded language provided by the FEEL++ framework represents a powerful engineering tool to rapidly develop and deploy production of ready-to-use softwares as well as prototypes. This results in the possibility to treat physical and engineering applications featuring complex coupled systems from early-stage exploratory analysis till the most advanced investigations on cutting-edge optimization topics. Within this framework, the use of FEEL++ for the simulation of the dynamic inside a bioreactor landfill seemed promising considering the complexity of the problem under analysis featuring multiphysics phenomena at different space and time scales.

A key aspect in the use of FEEL++ for industrial applications like the one presented in this paper is the possibility to operate on parallel infrastructures without directly managing the MPI communications. Here, we briefly recall the main steps for the parallel simulation of the dynamic inside a bioreactor landfill highlighting the tools involved in FEEL++ and in the external libraries linked to it:

- we start by constructing a computational mesh using GMSH [13];
- a mesh partition is generated using Chaco or Metis and additional information about ghost cells with communication between neighbors is provided [13];
- FEEL++ generates the required parallel data structures and create a table with global and local views for the Degrees of Freedom;
- FEEL++ assembles the system of PDE's starting from the variational formulations and the chosen Finite Element spaces;
- the algebraic problem is solved using the efficient solvers and preconditioners provided by PETSc [2].

A detailed description of the high performance framework within FEEL++ is available in [7].

5.2. Geometric data

We consider a reverse truncated pyramid domain as in figure 2C. The base of the domain measures 90 m \times 90 m and its height is 90 m. All the lateral walls feature a slope of $\pi/6$. The alveolus counts 20 extraction drains

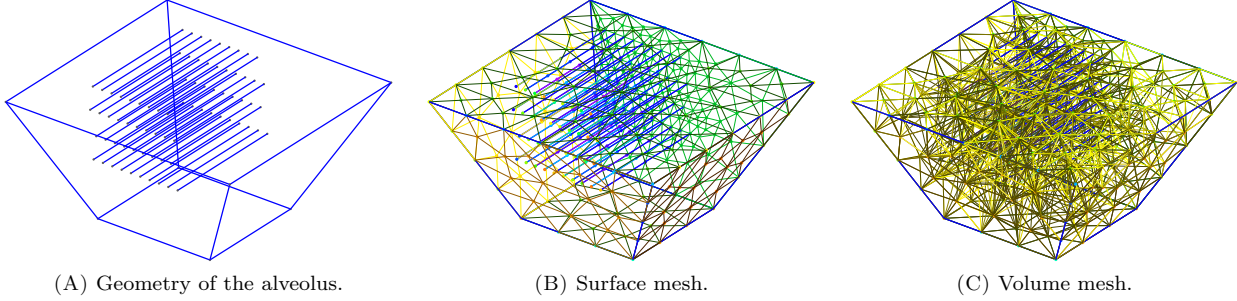


FIGURE 3. Geometry of the alveolus and computational mesh generated using GMSH. Each layer of 1D lines alternatively represents a set of water injectors or a group of extraction drains for the gas.

organized on 2 levels and 20 injection pipes, distributed on 2 levels as well. All the pipes are 25 m long and are modeled as 1D lines since their diameters are of order 10^{-1} m. A simplified scheme of the alveolus under analysis is provided in figure 3A and the corresponding computational domain is obtained by constructing a triangulation of mesh size 5 m (Fig. 3C).

5.3. Heuristic evaluation of the unknown constants in the model

The model presented in this paper features a large set of unknown variables (i.e. diffusion coefficients, scaling factors, ...) whose role is crucial to obtain realistic simulations. In this section, we propose a first set of values for these parameters that have been heuristically deduced by means of some qualitative and numerical considerations. A major improvement of the model is expected by a more rigorous tuning of these parameters which will be investigated in future works.

Porous medium

The physical and chemical properties of the bioreactor considered as a porous environment have been derived by experimental results in the literature. In particular, we consider a porosity $\Phi = 0.3$ and a permeability $D = 10^{-11} \text{ m}^2$.

Bacteria and organic carbon

We consider both the concentration of bacteria b and of organic carbon C^{org} as non-dimensional quantities in order to estimate their evolution. Thus we set $b_0 = 1$ and $C_0^{\text{org}} = 1$ and we derive the values $a_b = 10^{-5} \text{ m}^6 \text{ kg}^{-2} \text{ d}^{-1}$ and $c_b = 1$ respectively for the rate of consumption of the organic carbon and for the rate of reproduction of bacteria. Within this framework and under the optimal conditions of reaction prescribed by (3.1), C^{org} decreases to 2% of its initial value during the lifetime of the bioreactor whereas bacteria concentration b remains bounded ($b < 2b_0$).

Temperature

As per experimental data, the optimal temperature for the methanogenic fermentation to take place is $T_{\text{opt}} = 35^\circ\text{C} = 308 \text{ K}$ with an admissible variation of temperature of $\pm A_T = 20^\circ\text{C} = 20 \text{ K}$ to guarantee the survival of bacteria. The factor c_T represents the heat produced per unit of consumed organic carbon and per unit of time and is estimated to $c_T = 10^2 \text{ K m}^2 \text{ d}^{-1}$. The thermal conductivity of the waste inside the alveolus is fixed at $k_T = 9 \times 10^{-2} \text{ m}^2 \text{ d}^{-1}$. In order to impose realistic boundary conditions for the heat equation, we consider different values for the temperature of the lateral surface of the alveolus $T_g = 5^\circ\text{C} = 278 \text{ K}$ and the one of the top geomembrane $T_m = 20^\circ\text{C} = 293 \text{ K}$.

Water

The production of methane takes place only when less than 10% of water is present inside the bioreactor. Since the alveolus is completely flooded when $w = 1000 \text{ kg m}^{-3}$, we get that $w_{\max} = 100 \text{ kg m}^{-3}$. The vertical velocity drift due to gravity is estimated from Darcy's law to the value $\|u_w\| = 2.1 \text{ m d}^{-1}$ and the diffusion coefficient is set to $k_w = 8.6 \times 10^{-2} \text{ m}^2 \text{ d}^{-1}$.

Phase transition

In order to model the phase transitions, we have to consider the critical values of the pressure associated with evaporation and condensation. The Rankine law for the vapor pressure of water is approximated using the following values: $P_0 = 133.322 \text{ Pa}$, $s_0 = 20.386$ and $s_1 = 5132 \text{ K}$ and its linearization arises when considering $H_0 = -9.56 \times 10^4 \text{ Pa}$ and $H_1 = 337.89 \text{ Pa K}^{-1}$ for the range of temperatures $[288 \text{ K}; 328 \text{ K}]$. Moreover, we set the value $c_{h \rightarrow w}$ that represents the speed for the condensation of vapor to liquid water: $c_{h \rightarrow w} = 10^{-1} \text{ d}^{-1}$.

Gases

We consider a gas mixture made of methane, carbon dioxide and water vapor. Its dynamical viscosity is set to $\mu_{\text{gas}} = 1.3 \text{ Pa d}^{-1}$. Other parameters involved represent the rate of production of a specific gas (methane, carbon dioxide and water vapor) per unit of consumed organic carbon and per unit of time: $c_M = 1.8 \times 10^7 \text{ kg m}^{-3}$; $c_C = 2.6 \times 10^7 \text{ kg m}^{-3}$; $c_h = 2.5 \times 10^6 \text{ kg m}^{-3}$.

Remark 5.1. A key aspect in the modeling of a bioreactor landfill is the possibility to adapt the incoming flow of water and leachates J_{in} and the outgoing flow of biogas J_{out} . These values are user-defined parameters which are kept constant to $258 \text{ m}^3 \text{ d}^{-1}$ for the simulations in this paper but should act as control variables in the framework of the forecast and optimization procedures described in the introduction.

5.4. A preliminary test case

In this section we present some preliminary numerical results obtained by using the SiViBiR++ module developed in FEEL++ to solve the equations presented in section 3 using the numerical schemes discussed in section 4. In particular, we remark that in all the following simulations we neglect the effects due to the gas and fluid dynamics inside the bioreactor landfill. Though this choice limits the ability of the discussed results to correctly describe the complete physical behavior of the system, this simplification is a necessary starting point for the validation of the mathematical model in section 3. As a matter of fact, the equations describing the gas and fluid dynamics feature several unknown parameters whose tuning - independent and coupled with one another - has to be accurately performed before linking them to the problems modeling the consumption of organic carbon and the evolution of the temperature.

Thus, here we restrict our numerical simulations to two main phenomena occurring inside the bioreactor landfill: first, we describe the consumption of organic carbon under some fixed optimal conditions of humidity and temperature; then we introduce the evolution in time of the temperature and we discuss the behavior of the coupled system given by equations (3.5)-(3.6).

The test cases are studied in the computational domain introduced in section 5.2: in particular, we consider the triangulation of mesh size 5 m in figure 3C and we set the unit measure for the time evolution to $\Delta t = 365 \text{ d}$. The final time for the simulation is $S^{\text{fin}} = 40$ years. The parameters inside the equations are set according to the values in section 5.3 but a thorough investigation of these quantities has to be performed to verify their accuracy. The computations have been executed using up to 32 processors and below we present some simulations for the aforementioned preliminary test cases.

Evolution of the organic carbon under optimal hydration and temperature conditions

First of all, we consider the case of a single uncoupled equation, that is the evolution of the organic carbon in a scenario in which the concentration of water and the temperature are fixed. Starting from equation (3.5), we assume fixed optimal conditions for the humidity and the temperature inside the bioreactor landfill. We set

a fixed amount of water $w = \frac{w_{\max}}{2}$ inside the alveolus and a constant temperature $T = T_{\text{opt}}$. Thus, from (3.1) we get

$$\Psi_1(w) \equiv \frac{w_{\max}}{4}, \quad \Psi_2(T) \equiv 1$$

and equation (3.5) reduces to

$$\begin{cases} (1 - \phi) \partial_t C^{\text{org}}(x, t) = -a_b \frac{w_{\max}}{4} [b_0 + c_b(C_0^{\text{org}} - C^{\text{org}}(x, t))] C^{\text{org}}(x, t), & \text{in } \Omega \times (0, S^{\text{fin}}] \\ C^{\text{org}}(\cdot, 0) = C_0^{\text{org}}, & \text{in } \bar{\Omega} \end{cases} \quad (5.1)$$

It is straightforward to observe that equation (5.1) only features one unknown variable - namely the organic carbon - since the concentration of bacteria $b(x, t)$ is an affine function of the concentration of organic carbon itself (cf. equation (3.4)).

As stated in section 4.2, the key aspect in the solution of equation (5.1) is the handling of the non-linear reaction term. In order to numerically treat this term as described, at each time step we need the value C_n^{org} at the previous iteration to compute the semi-implicit quantity $C_{n+1}^{\text{org}} C_n^{\text{org}}$. To provide a suitable value of C_n^{org} during the first iteration, we solve a linearization of equation (5.1) and we use the corresponding solution to evaluate $C_{n+1}^{\text{org}} C_n^{\text{org}}$.

The initial concentration of organic carbon inside the bioreactor is set to 1 and in figure 4 we observe several snapshots of the quantity of organic carbon inside the alveolus at time $t = 1$ year, $t = 10$ years, $t = 20$ years and $t = 40$ years. At the end of the life of the facility, the amount of organic carbon inside the alveolus is $C^{\text{org}} = 2.0 \times 10^{-3}$. In figure 5, we plot the evolution of the overall quantity of organic carbon with respect to

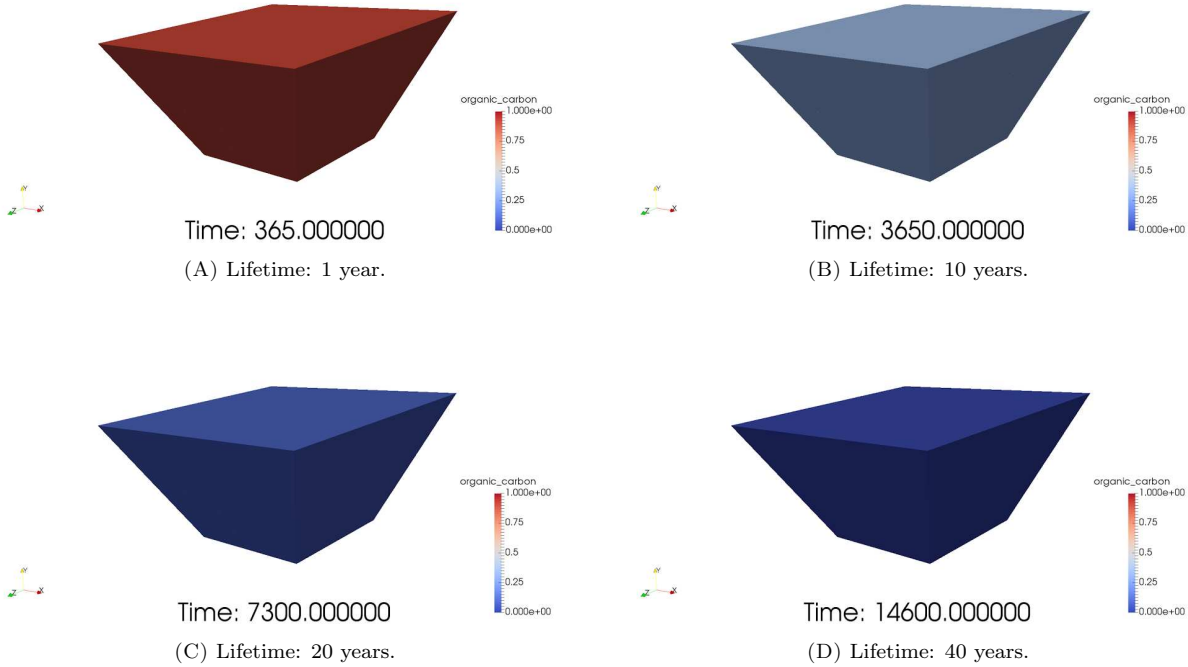


FIGURE 4. Evolution of the organic carbon inside the alveolus at $t = 1$ year, $t = 10$ years, $t = 20$ years and $t = 40$ years.

time. In particular, as expected we observe that C^{org} decreases in time as the methanogenic fermentation takes place.

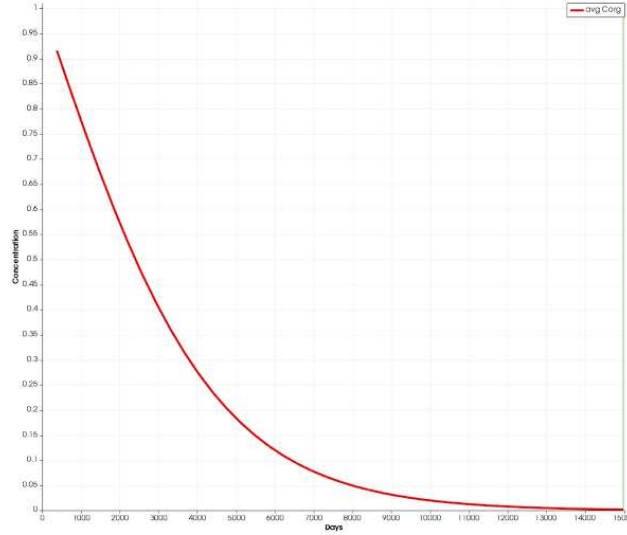


FIGURE 5. Evolution of the organic carbon between $t = 0$ and $t = 40$ years under optimal hydration and temperature conditions.

Evolution of the temperature as a function of organic carbon

We introduce a novel variable depending both on space and time to model the temperature inside the bioreactor. Figure 6 presents the snapshots of the value of the temperature in a section of the alveolus under analysis at time $t = 1$ year, $t = 10$ years, $t = 20$ years and $t = 40$ years. These result from the solution of equation (3.6) for a given trend of the organic carbon. Let us consider the evolution of the organic carbon obtained from the previous test case. The corresponding profile is given by

$$C^{\text{org}}(x, t) = e^{-\alpha t} \quad , \quad \alpha = 10^{-3}.$$

We observe that the consumption of organic carbon by means of the chemical reaction (2.2) is responsible for the generation of heat in the middle of the domain. As expected by the physics of the problem, the heat tends to diffuse towards the external boundaries where the temperature is lower. After a first phase which lasts approximately 10 years in which the methanogenic process produces heat and the temperature rises, the consumption of organic carbon slows down and the temperature as well starts to decrease until the end-of-life of the facility (Fig. 7A-7B).

Evolution of the coupled system of organic carbon and temperature under optimal hydration condition

Starting from the previously discussed cases, we now proceed to the coupling of the organic carbon with the temperature. We keep the optimal hydration condition as in the previous simulations - that is $w = \frac{w_{\text{max}}}{2}$ - and we consider the solution of the coupled equations (3.5)-(3.6).

From the numerical point of view, this scenario introduces several difficulties, mainly due to the fact that the two equations are now dependent on one another. As mentioned in section 4, the coupling is handled explicitly, that is, first we solve the problem featuring the organic carbon with fixed temperature then we approximate the heat equation using the information arising from the previously computed C^{org} .

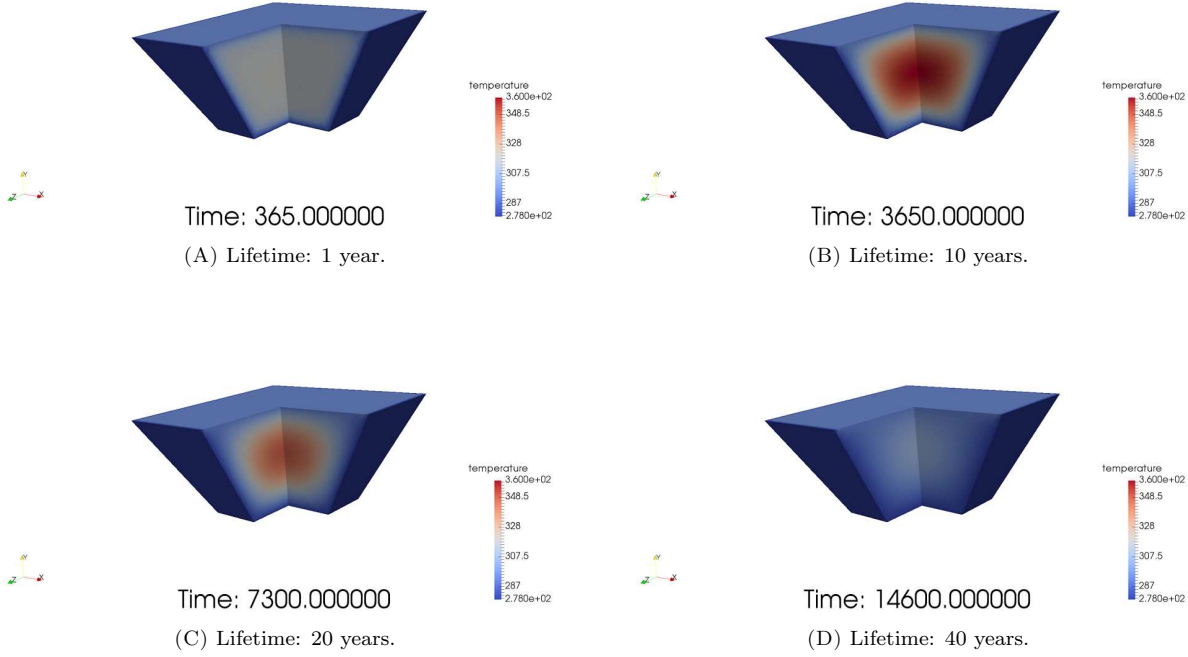


FIGURE 6. Evolution of the temperature inside the alveolus at $t = 1$ year, $t = 10$ years, $t = 20$ years and $t = 40$ years.

Within this framework, at time $t = t_n$ the conditions of humidity and temperature for the organic carbon equation read as

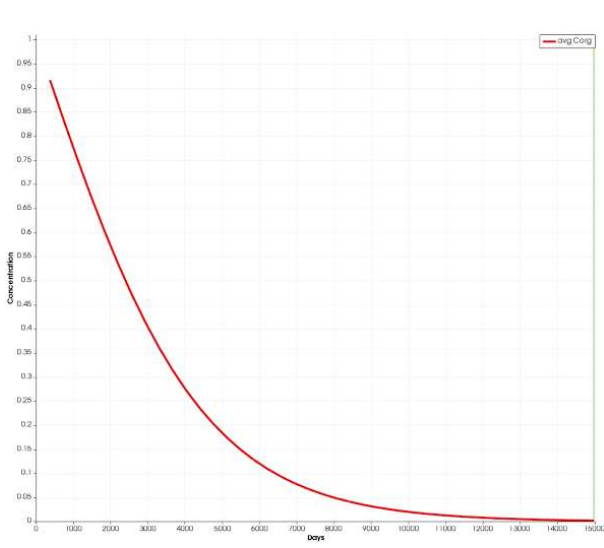
$$\Psi_1(w) \equiv \frac{w_{\max}}{4}, \quad \Psi_2(T) = \max\left(0, 1 - \frac{|T_n - T_{\text{opt}}|}{A_T}\right)$$

where T_n is the temperature at the previous iteration.

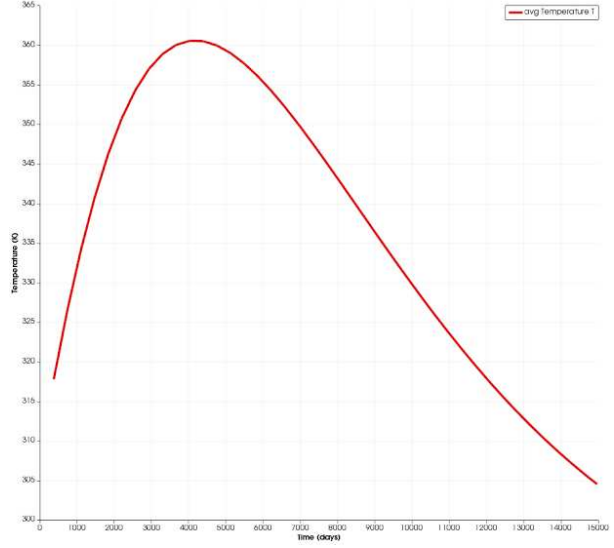
As in the previous case, we consider an initial concentration of organic carbon equal to 1 and we observe it decreasing in figure 8A due to the bacterial activity. We verify that the quantity of organic carbon inside the bioreactor landfill decays towards zero during the lifetime of the facility. At the same time, the temperature increases as a result of the methanogenic process catalyzed by the microbiota (Fig. 8B). Nevertheless, when the temperature goes beyond the tolerated variation A_T , the second condition in (3.1) is no more fulfilled and the chemical reaction is prevented. We may observe this behavior in figures 8A-8B between $t = 3$ years and $t = 20$ years. Once the temperature is inside the admissible range $[T_{\text{opt}} - A_T; T_{\text{opt}} + A_T]$ again (starting approximately from $t = 20$ years), the reaction (2.2) is allowed, the organic carbon is consumed and influences the temperature which slightly increases again before eventually decreasing towards the end-of-life of the bioreactor. Eventually, in figure 9 we report some snapshots of the solutions of the coupled system (3.5)-(3.6).

6. CONCLUSION

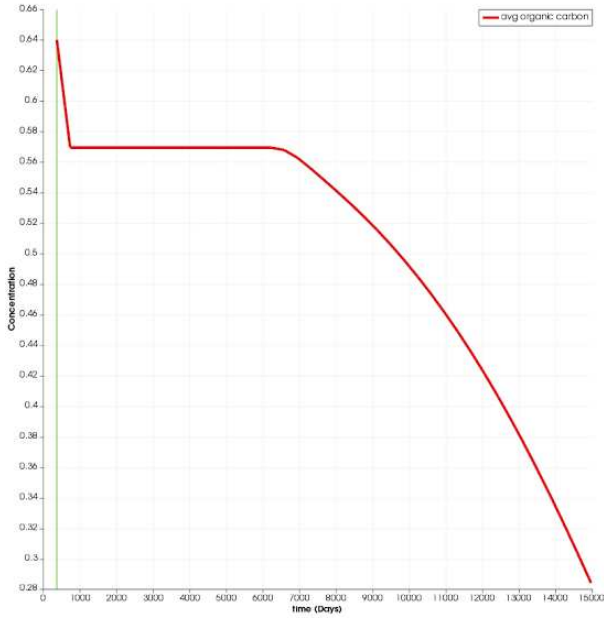
In this work, we proposed a first attempt to mathematically model the physical and chemical phenomena taking place inside a bioreactor landfill. A set of 7 coupled equations has been derived and a Finite Element discretization has been introduced using FEEL++. A key aspect of the discussed model is the tuning of the



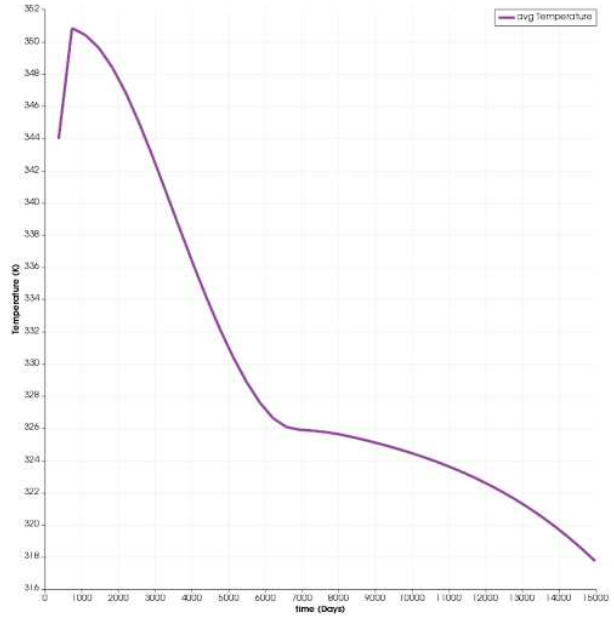
(A) Evolution of the organic carbon.



(B) Evolution of the temperature.

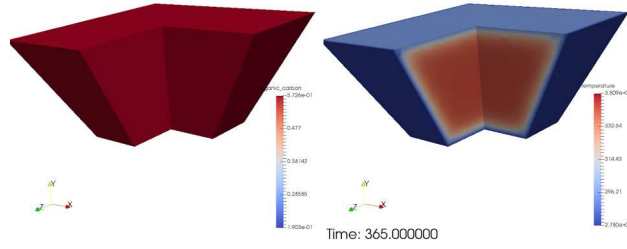
FIGURE 7. Evolution of the quantity of organic carbon and temperature inside the alveolus between $t = 0$ and $t = 40$ years.

(A) Evolution of the organic carbon.

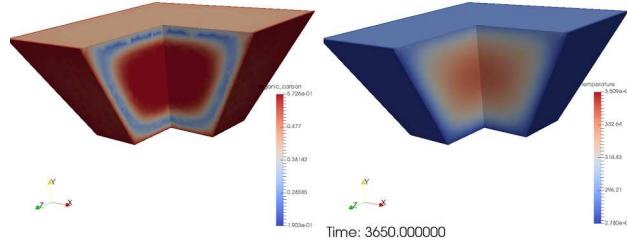


(B) Evolution of the temperature.

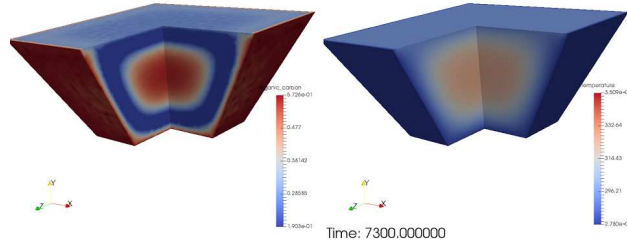
FIGURE 8. Evolution of the quantity of organic carbon and the temperature inside the alveolus between $t = 0$ and $t = 40$ years.



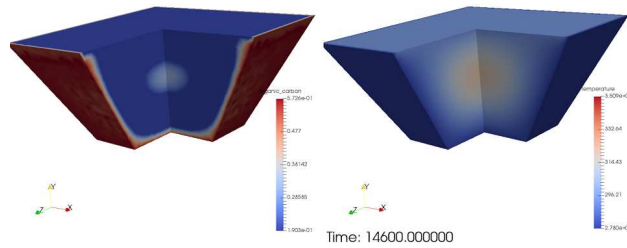
(A) Lifetime: 1 year.



(B) Lifetime: 10 years.



(C) Lifetime: 20 years.



(D) Lifetime: 40 years.

FIGURE 9. Coupled evolution of the organic carbon (left) and the temperature (right) inside the alveolus at $t = 1$ year, $t = 10$ years, $t = 20$ years and $t = 40$ years.

coefficients appearing in the equations. On the one hand, part of these unknowns represents physical quantities whose values may be derived from experimental studies. On the other hand, some parameters are scalar factors that have to be estimated by means of heuristic approaches. A rigorous tuning of these quantities represents a major line of investigation to finalize the implementation of the model in the SiViBiR++ module and its validation with real data.

The present work represents a starting point for the development of mathematically-sound investigations on bioreactor landfills. From a modeling point of view, some assumptions may be relaxed, for example by adding a term to account for the death rate of bacteria or the cooling effect due to water injection inside the bioreactor. The final goal of SiViBiR++ project is the simulation of long-time behavior of the bioreactor in order to perform forecasts on the methane production and optimize the control strategy. The associated inverse problems and PDE-constrained optimization problems are likely to be numerically intractable due to their complexity and their dimension thus the study of reduced order models may be necessary to decrease the overall computational cost.

A. SUMMARY OF THE UNKNOWN PARAMETERS

In the following table, we summarize the values of some unknown parameters which were deduced during the present work. We highlight that all these quantities have been estimated via heuristic approaches and a rigorous verification/validation procedure remains necessary before their application to real-world problems.

Parameter	Description	Value	Unit
Φ	Porosity of the medium	0.3	
D	Permeability	10^{-11}	m^2
b_0	Initial concentration of bacteria	1.0	
C_0^{org}	Initial concentration of organic carbon	1.0	
a_b	Rate of consumption of organic carbon	10^{-5}	$\text{m}^6\text{kg}^{-2}\text{d}^{-1}$
c_b	Rate of creation of bacteria	1.0	
T_{opt}	Optimal temperature for the reaction	308	K
A_T	Tolerated variation of temperature	20	K
c_T	Rate of production of heat by the chemical reaction	10^2	K
k_T	Thermal conductivity	9×10^{-2}	m^2d^{-1}
T_g	Temperature of the soil	278	K
T_m	Temperature of the geomembrane	293	K
T_0	Initial temperature	293	K
w_{max}	Maximal admissible quantity of water	100	kg m^{-3}
$\ u_w\ $	Velocity of the water	2.1	m d^{-1}
k_w	Diffusion coefficient of the water	8.6×10^{-2}	m^2d^{-1}
w_0	Initial quantity of water	50	kg m^{-3}
H_0	Constant for the vapor pressure	-9.56×10^4	Pa
H_1	Constant for the vapor pressure	337.89	Pa K^{-1}
$c_{h \rightarrow w}$	Condensation rate	10^{-1}	d^{-1}
μ_{gas}	Dynamical viscosity of the gas	1.3	Pa d^{-1}
c_M	Rate of production of methane	1.8×10^7	kg m^{-3}

Parameter	Description	Value	Unit
c_C	Rate of production of carbon dioxide	2.6×10^7	kg m^{-3}
c_h	Rate of production of water vapor	2.5×10^6	kg m^{-3}
M_0	Initial concentration of methane	1.0	
C_0^{dx}	Initial concentration of carbon dioxide	1.0	
h_0	Initial concentration of water vapor	1.0	
J_{out}	Outgoing flow of biogas	258	m^3d^{-1}
J_{in}	Incoming flow of water and leachates	258	m^3d^{-1}

TABLE 1. Summary of the parameters involved in the 7-equations model.

ACKNOWLEDGMENTS

The authors are grateful to Alexandre Ancel (Université de Strasbourg) and the FEEL++ community for the technical support. The authors wish to thank the CEMRACS 2015 and its organizers.

REFERENCES

- [1] F. Agostini, C. Sundberg, and R. Navia. Is biodegradable waste a porous environment? A review. *Waste Manage. Res.*, 30(10):1001–1015, 2012.
- [2] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. Gropp, D. Karpeyev, D. Kaushik, M. Knepley, L. Curfman McInnes, K. Rupp, B. Smith, S. Zampini, and H. Zhang. PETSc Users Manual. Technical report, Argonne National Laboratory, 2015.
- [3] F. Bezzo, S. Macchietto, and C. C. Pantelides. General hybrid multizonal/CFD approach for bioreactor modeling. *AIChE Journal*, 49(8):2133–2148, 2003.
- [4] P. B. Bochev, M. D. Gunzburger, and J. N. Shadid. Stability of the SUPG finite element method for transient advection-diffusion problems. *Comput. Method. Appl. M.*, 193(23-26):2301–2323, 2004.
- [5] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.
- [6] E. Burman. Consistent SUPG-method for transient transport problems: Stability and convergence. *Comput. Method. Appl. M.*, 199(17-20):1114–1123, 2010.
- [7] V. Chabanne. *Vers la simulation des écoulements sanguins*. Ph.D. thesis, Université Joseph Fourier, Grenoble, 2013.
- [8] T. Chassagnac. Etat des connaissances techniques et recommandations de mise en oeuvre pour une gestion des installations de stockage de déchets non dangereux en mode bioréacteur. Technical report, ADEME - Agence Nationale de l'Environnement et de la Maîtrise de l'Energie. FNADE - Fédération Nationale des Activités de la Dépollution et de l'Environnement, 2007.
- [9] A. J. Chorin and J. E. Marsden. *A mathematical introduction to fluid mechanics*, volume 4 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 1993.
- [10] D. Córdoba, F. Gancedo, and R. Orive. Analytical behavior of two-dimensional incompressible flow in porous media. *J. Math. Phys.*, 48(6):065206, 2007.
- [11] M. R. Correa and A. F. D. Loula. Unconditionally stable mixed finite element methods for Darcy flow. *Comput. Method. Appl. M.*, 197(17-18):1525–1540, 2008.
- [12] L. J. Durlofsky. Accuracy of mixed and control volume finite element approximations to Darcy velocity and related quantities. *Water Resour. Res.*, 30(4):965–973, 1994.
- [13] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and postprocessing facilities, 2009.
- [14] C. Johnson et al. Characterization, Design, Construction and Monitoring of Bioreactor Landfills. Technical report, ITRC - Interstate Technology & Regulatory Council, 2006.
- [15] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Comput. Method. Appl. M.*, 45(1):285–312, 1984.

- [16] J. Kindlein, D. Dinkler, and H. Ahrens. Numerical modelling of multiphase flow and transport processes in landfills. *Waste Manage. Res.*, 24(4):376–387, 2006.
- [17] T. Kling and J. Korkealaakso. Multiphase modeling and inversion methods for controlling landfill bioreactor. In *Proceedings of TOUGH Symposium*. Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 2006.
- [18] M. Lebeau and J.-M. Konrad. Natural convection of compressible and incompressible gases in undeformable porous media under cold climate conditions. *Comput. Geotech.*, 36(3):435–445, 2009.
- [19] C. N. Liu, R. H. Chen, and K. S. Chen. Unsaturated consolidation theory for the prediction of long-term municipal solid waste landfill settlement. *Waste Manage. Res.*, 24(1):80–91, 2006.
- [20] S. Martín, E. Marañón, and H. Sastre. Mathematical modelling of landfill gas migration in MSW sanitary landfills. *Waste Manage. Res.*, 19(5):425–435, 2001.
- [21] P. McCreanor. *Landfill leachate recirculation systems: Mathematical modelling and validation*. PhD thesis, University of Central Florida, USA, 1998.
- [22] C. Prud’homme. A Domain Specific Embedded Language in C++ for automatic differentiation, projection, integration and variational formulations. *Sci. Program.*, 14(2):81–110, 2006.
- [23] C. Prud’homme, V. Chabannes, V. Doyeux, M. Ismail, A. Samake, and G. Pena. Feel++: A computational framework for Galerkin methods and advanced numerical methods. *ESAIM: ProcS*, 38:429–455, 2012.
- [24] P. A. Raviart and J. M. Thomas. A mixed finite element method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of Finite Element Methods*, volume 606 of *Lect. Notes Math*, pages 292–315. Springer-Verlag, 1977.
- [25] D. R. Reinhart and B. A. Al-Yousfi. The impact of leachate recirculation on municipal solid waste landfill operating characteristics. *Waste Manage. Res.*, 14(4):337–346, 1996.
- [26] H. S. Sidhu, M. I. Nelson, and X. D. Chen. A simple spatial model for self-heating compost piles. In W. Read and A. J. Roberts, editors, *Proceedings of the 13th Biennial Computational Techniques and Applications Conference, CTAC-2006*, volume 48 of *ANZIAM J.*, pages C135–C150, 2007.
- [27] K. Vafai. *Handbook of porous media*. Crc Press, 2015.
- [28] J. J. Walsh and R. N. Kinman. Leachate and gas production under controlled moisture conditions. In *Municipal Solid Waste: Land Disposal, Proceedings of the 5th Annual Research Symposium*, pages 41–57. EPA-600/9-79-023a, 1979.

Published and preprint articles

- [1] G. Dollé, O. Duran, N. Feyeux, E. Frénod, M. Giacomini, and C. Prud'homme. Mathematical modeling and numerical simulation of a bioreactor landfill using Feel++. *ESAIM: Proceedings and Surveys*, 2016. To appear.
- [2] M. Giacomini. An equilibrated fluxes approach to the Certified Descent Algorithm for shape optimization using conforming Finite Element and Discontinuous Galerkin discretizations. Submitted, 2016.
- [3] M. Giacomini, O. Pantz, and K. Trabelsi. An a posteriori error estimator for shape optimization: application to EIT. *J. Phys.: Conf. Ser.*, 657(1):012004, 2015.
- [4] M. Giacomini, O. Pantz, and K. Trabelsi. Certified Descent Algorithm for shape optimization driven by fully-computable a posteriori error estimators. *ESAIM: Contr. Op. Ca. Va.*, 2016. To appear.
- [5] M. Giacomini, O. Pantz, and K. Trabelsi. Certified Descent Algorithm for structural shape optimization. In preparation, 2016.
- [6] M. Giacomini, O. Pantz, and K. Trabelsi. Volumetric expressions for the shape gradient of the compliance in structural shape optimization. To be submitted, 2016.
- [7] J. Vazquez, M. Giacomini, J. Escareno, E. Rubio, and H. Sossa. Optimal grasping points identification for a rotational four-fingered aerogripper. In *2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, pages 272–277, Nov 2015.

Bibliography

- [8] S. Adams and B. Cockburn. A mixed finite element method for elasticity in three dimensions. *J. Sci. Comput.*, 25(3):515–521, 2005.
- [9] L. Afraites, M. Dambrine, and D. Kateb. Shape methods for the transmission problem with a single measurement. *Numer. Func. Anal. Opt.*, 28(5-6):519–551, 2007.
- [10] L. Afraites, M. Dambrine, and D. Kateb. On second order shape optimization methods for electrical impedance tomography. *SIAM J. Control Optim.*, 47(3):1556–1590, 2008.
- [11] M. Ainsworth. A Posteriori Error Estimation for Discontinuous Galerkin Finite Element Approximation. *SIAM J. Numer. Anal.*, 45(4):1777–1798, 2007.
- [12] M. Ainsworth and J. Oden. A posteriori error estimators for second order elliptic systems part 2. An optimal order process for calculating self-equilibrating fluxes. *Comput. Math. Appl.*, 26(9):75 – 87, 1993.
- [13] M. Ainsworth and J. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. A Wiley-Interscience publication. Wiley, 2000.
- [14] M. Ainsworth and R. Rankin. Guaranteed computable bounds on quantities of interest in finite element computations. *Int. J. Numer. Meth. Eng.*, 89(13):1605–1634, 2012.
- [15] F. Alauzet, B. Mohammadi, and O. Pironneau. Mesh adaptivity and optimal shape design for aerospace. In G. Buttazzo and A. Frediani, editors, *Variational Analysis and Aerospace Engineering: Mathematical Challenges for Aerospace Design*, Springer Optimization and Its Applications, pages 323–337. Springer US, 2012.
- [16] G. Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002.
- [17] G. Allaire. *Conception optimale de structures*. Springer, 2006.
- [18] G. Allaire, E. Cancès, and J.-L. Vié. Second-order shape derivatives along normal trajectories, governed by Hamilton-Jacobi equations. *Struct. Multidiscip. O.*, 54(5):1245–1266, 2016.
- [19] G. Allaire, C. Dapogny, and P. Frey. A mesh evolution algorithm based on the level set method for geometry and topology optimization. *Struct. Multidiscip. O.*, 48(4):711–715, Oct. 2013.

- [20] G. Allaire, C. Dapogny, and P. Frey. Shape optimization with a level set based mesh evolution method. *Comput. Method. Appl. M.*, 282:22 – 53, 2014.
- [21] G. Allaire, F. de Gournay, F. Jouve, and A.-M. Toader. Structural optimization using topological and shape sensitivity via a level set method. *Control Cybern.*, 34(1):59–80, 2005.
- [22] G. Allaire and G. A. Francfort. A numerical algorithm for topology and shape optimization. In *Topology design of structures (Sesimbra, 1992)*, volume 227 of *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, pages 239–248. Kluwer Acad. Publ., Dordrecht, 1993.
- [23] G. Allaire, F. Jouve, and G. Michailidis. Molding direction constraints in structural optimization via a level-set method. working paper or preprint, Dec. 2015.
- [24] G. Allaire, F. Jouve, and G. Michailidis. Thickness control in structural optimization via a level set method. *Struct. Multidiscip. O.*, 53(6):1349–1382, 2016.
- [25] G. Allaire, F. Jouve, and A.-M. Toader. A level-set method for shape optimization. *C. R. Acad. Sci. I-Math.*, 334(12):1125 – 1130, 2002.
- [26] G. Allaire, F. Jouve, and A.-M. Toader. Structural optimization using sensitivity analysis and a level-set method. *J. Comput. Phys.*, 194(1):363 – 393, 2004.
- [27] G. Allaire and R. V. Kohn. Optimal design for minimum weight and compliance in plane stress using extremal microstructures. *Eur. J. Mech. A-Solid.*, 12(6):839–878, 1993.
- [28] G. Allaire and O. Pantz. Structural optimization with FreeFem++. *Struct. Multidiscip. O.*, 32(3):173–181, 2006.
- [29] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. Rognes, and G. Wells. The FEniCS Project Version 1.5. *Archive of Numerical Software*, 3(100), 2015.
- [30] M. Amara and J. M. Thomas. Equilibrium finite elements for the linear elastic problem. *Numer. Math.*, 33(4):367–383, 1979.
- [31] L. Ambrosio and G. Buttazzo. An optimal design problem with perimeter penalization. *Calc. Var. Partial Dif.*, 1(1):55–69, 1993.
- [32] H. Ammari, E. Bossy, J. Garnier, and L. Seppecher. Acousto-electromagnetic tomography. *SIAM J. Appl. Math.*, 72(5):1592–1617, 2012.
- [33] H. Ammari, E. Bretin, J. Garnier, H. Kang, H. Lee, and A. Wahab. *Mathematical Methods in Elasticity Imaging*. Princeton University Press, Princeton, NJ, USA, 2015.
- [34] S. Amstutz. Connections between topological sensitivity analysis and material interpolation schemes in topology optimization. *Struct. Multidiscip. O.*, 43(6):755–765, 2011.
- [35] D. N. Arnold. Mixed finite element methods for elliptic problems. *Comput. Method. Appl. M.*, 82(1-3):281–300, 1990. Reliability in computational mechanics (Austin, TX, 1989).

-
- [36] D. N. Arnold, G. Awanou, and R. Winther. Finite elements for symmetric tensors in three dimensions. *Math. Comput.*, 77(263):1229–1251, 2008.
 - [37] D. N. Arnold, F. Brezzi, and J. Douglas, Jr. PEERS: a new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1(2):347–367, 1984.
 - [38] D. N. Arnold, J. Douglas, and C. P. Gupta. A family of higher order mixed finite element methods for plane elasticity. *Numer. Math.*, 45(1):1–22, 1984.
 - [39] D. N. Arnold and R. S. Falk. A new mixed formulation for elasticity. *Numer. Math.*, 53(1-2):13–30, 1988.
 - [40] D. N. Arnold, R. S. Falk, and R. Winther. Differential complexes and stability of finite element methods. II. The elasticity complex. In *Compatible spatial discretizations*, volume 142 of *IMA Vol. Math. Appl.*, pages 47–67. Springer, New York, 2006.
 - [41] D. N. Arnold, R. S. Falk, and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Comput.*, 76(260):1699–1723, 2007.
 - [42] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92(3):401–419, 2002.
 - [43] J. P. Aubin and H. G. Burchard. Some aspects of the method of the hypercircle applied to elliptic variational problems. In B. Hubbard, editor, *Numerical Solution of Partial Differential Equations-II*, pages 1 – 67. Academic Press, 1971.
 - [44] H. Azegami, S. Kaizu, M. Shimoda, and E. Katamine. Irregularity of shape optimization problems and an improvement technique. In S. Hernandez and C. Brebbia, editors, *Computer Aided Optimization Design of Structures V*, pages 309–326. Computational Mechanics Publications, 1997.
 - [45] I. Babuška and A. Miller. A feedback finite element method with a posteriori error estimation: Part i. the finite element method and some basic properties of the a posteriori error estimator. *Comput. Method. Appl. M.*, 61(1):1 – 40, 1987.
 - [46] I. Babuška and W. C. Rheinboldt. A-posteriori error estimates for the finite element method. *Int. J. Numer. Meth. Eng.*, 12(10):1597–1615, 1978.
 - [47] I. Babuška and W. C. Rheinboldt. Analysis of optimal finite-element meshes in R^1 . *Math. Comput.*, 33(146):435–463, 1979.
 - [48] I. Babuška and W. C. Rheinboldt. A posteriori error analysis of finite element solutions for one-dimensional problems. *SIAM J. Numer. Anal.*, 18(3):565–589, 1981.
 - [49] I. Babuška and T. Strouboulis. *The Finite Element Method and Its Reliability*. Numerical mathematics and scientific computation. Clarendon Press, 2001.
 - [50] I. Babuška, T. Strouboulis, and S. Gangaraj. Guaranteed computable bounds for the exact error in the finite element solution part I: One-dimensional model problem. *Comput. Method. Appl. M.*, 176(14):51 – 79, 1999.

- [51] F. Ballarin, A. Manzoni, G. Rozza, and S. Salsa. Shape optimization by free-form deformation: existence results and numerical solution for Stokes flows. *J. Sci. Comput.*, 60(3):537–563, 2014.
- [52] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.
- [53] N. V. Banichuk. *Introduction to optimization of structures*. Springer-Verlag, New York, 1990. Translated from the Russian by Vadim Komkov.
- [54] N. V. Banichuk, F.-J. Barthold, A. Falk, and E. Stein. Mesh refinement for shape optimization. *Struct. Optimization*, 9:46–51, 1995.
- [55] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comput.*, 44(170):283–301, 1985.
- [56] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwend.*, 22(4):751–756, 2003.
- [57] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.
- [58] Z. Belhachmi and H. Meftahi. Shape sensitivity analysis for an interface problem via minimax differentiability. *Appl. Math. Comput.*, 219(12):6828 – 6842, 2013.
- [59] M. P. Bendsøe. Optimal shape design as a material distribution problem. *Struct. Optimization*, 1(4):193–202, 1989.
- [60] M. P. Bendsøe. *Optimization of Structural Topology, Shape, and Material*. Springer, 1995.
- [61] M. P. Bendsøe and N. Kikuchi. Generating optimal topologies in structural design using a homogenization method. *Comput. Method. Appl. M.*, 71(2):197–224, 1988.
- [62] M. P. Bendsøe and O. Sigmund. Material interpolation schemes in topology optimization. *Arch. Appl. Mech.*, 69(9):635–654, 1999.
- [63] M. P. Bendsøe and O. Sigmund. *Topology optimization*. Springer-Verlag, Berlin, 2003. Theory, methods and applications.
- [64] M. Berggren. A unified discrete–continuous sensitivity analysis method for shape optimization. In W. Fitzgibbon, Y. Kuznetsov, P. Neittaanmäki, J. Périaux, and O. Pironneau, editors, *Applied and Numerical Partial Differential Equations: Scientific Computing in Simulation, Optimization and Control in a Multidisciplinary Context*, pages 25–39. Springer Netherlands, Dordrecht, 2010.
- [65] J. T. Betts. *Practical methods for optimal control using nonlinear programming*, volume 3 of *Advances in Design and Control*. SIAM, Philadelphia, PA, 2001.
- [66] J. T. Betts and S. L. Campbell. Discretize then optimize. In *Mathematics for industry: challenges and frontiers*, pages 140–157. SIAM, Philadelphia, PA, 2005.

-
- [67] S. Bhavikatti and C. Ramakrishnan. Optimum shape design of shoulder fillets in tension bars and t-heads. *Int. J. Mech. Sci.*, 21(1):29 – 39, 1979.
- [68] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.
- [69] D. Boffi, F. Brezzi, and M. Fortin. Reduced symmetry elements in linear elasticity. *Commun. Pure Appl. Anal.*, 8(1):95–121, 2009.
- [70] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.
- [71] M. Bonnet and B. B. Guzina. Sounding of finite solid bodies by way of topological derivative. *Int. J. Numer. Meth. Eng.*, 61(13):2344–2373, 2004.
- [72] L. Borcea. Electrical impedance tomography. *Inverse Probl.*, 18(6):R99, 2002.
- [73] L. Boulaajine, S. Nicaise, L. Paquet, and Rafilipojaona. Dual mixed finite element methods for the elasticity problem with Lagrange multipliers. *J. Comput. Appl. Math.*, 221(1):234–260, 2008.
- [74] B. Bourdin. Filters in topology optimization. *Int. J. Numer. Meth. Eng.*, 50(9):2143–2158, 2001.
- [75] B. Bourdin and A. Chambolle. Design-dependent loads in topology optimization. *ESAIM: Contr. Op. Ca. Va.*, 9:19–48, 8 2003.
- [76] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2001.
- [77] D. Braess, T. Fraunholz, and R. H. W. Hoppe. An equilibrated a posteriori error estimator for the interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 52(4):2121–2136, 2014.
- [78] D. Braess and J. Schöberl. Equilibrated residual error estimator for edge elements. *Math. Comput.*, 77(262):651–672, 2008.
- [79] V. Braibant and C. Fleury. Shape optimal design using b-splines. *Comput. Method. Appl. M.*, 44(3):247 – 267, 1984.
- [80] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.
- [81] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. *ESAIM: Math. Model. Num.*, 8(R2):129–151, 1974.
- [82] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Recent results on mixed finite element methods for second order elliptic problems. In *Vistas in applied mathematics*, Transl. Ser. Math. Engrg., pages 25–43. Optimization Software, New York, 1986.

- [83] R. A. Brockman. Geometric sensitivity analysis with isoparametric finite elements. *Commun. Appl. Numer. M.*, 3(6):495–499, 1987.
- [84] M. Bruggi. Topology optimization with mixed finite elements on regular grids. *Comput. Method. Appl. M.*, 305:133 – 153, 2016.
- [85] M. Bruggi and C. Cinquini. An alternative truly-mixed formulation to solve pressure load problems in topology optimization. *Comput. Method. Appl. M.*, 198(17–20):1500 – 1512, 2009.
- [86] M. Bruggi and P. Venini. Topology optimization of incompressible media using mixed finite elements. *Comput. Method. Appl. M.*, 196(33–34):3151 – 3164, 2007.
- [87] M. Bruggi and P. Venini. A mixed FEM approach to stressconstrained topology optimization. *Int. J. Numer. Meth. Eng.*, 73(12):1693–1714, 3 2008.
- [88] D. Bucur and G. Buttazzo. *Variational methods in shape optimization problems*. Progress in Nonlinear Differential Equations and their Applications, 65. Birkhäuser Boston, Inc., Boston, MA, 2005.
- [89] M. Burger, B. Hackl, and W. Ring. Incorporating topological derivatives into level set methods. *J. Comput. Phys.*, 194(1):344 – 362, 2004.
- [90] M. Burger and S. J. Osher. A survey on level set methods for inverse problems and optimal design. *Eur. J. Appl. Math.*, null:263–301, 4 2005.
- [91] M. Burger and R. Stainko. Phasefield relaxation of topology optimization with local stress constraints. *SIAM J. Control Optim.*, 45(4):1447–1466, 2006.
- [92] E. Burman and P. Zunino. A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 44(4):1612–1638, 2006.
- [93] J. Cahn and J. Hilliard. Free Energy of a Nonuniform System I - Interfacial Free Energy. *J. Chem. Phys.*, 28:258–267, 1958.
- [94] A. Calderón. On an inverse boundary value problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro 1980)*, pages 65–73. Soc. Brasil. Mat., 1980.
- [95] P. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J. Theor. Biol.*, 26(1):61 – 81, 1970.
- [96] A. Carpio and M.-L. Rapún. Hybrid topological derivative and gradient-based methods for electrical impedance tomography. *Inverse Probl.*, 28(9):095010, 22, 2012.
- [97] C. Carstensen, M. Eigel, and J. Gedicke. Computational competition of symmetric mixed FEM in linear elasticity. *Comput. Method. Appl. M.*, 200(4144):2903 – 2915, 2011.
- [98] C. Carstensen and S. A. Funken. Fully reliable localized error control in the fem. *SIAM J. Sci. Comput.*, 21(4):1465–1484, 1999.

-
- [99] C. Carstensen and S. A. Funken. Constants in Clément-interpolation error and residual based a posteriori estimates in finite element methods. *East-West J. Numer. Math.*, 8(3):153–175, 2000.
 - [100] C. Carstensen, D. Günther, J. Reininghaus, and J. Thiele. The Arnold-Winther mixed FEM in linear elasticity. part I: Implementation and numerical verification. *Comput. Method. Appl. M.*, 197(3340):3014 – 3023, 2008.
 - [101] C. Carstensen and C. Merdon. Effective postprocessing for equilibration a posteriori error estimators. *Numer. Math.*, 123(3):425–459, 2013.
 - [102] J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.
 - [103] J. Céa. Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût. *ESAIM: Math. Model. Num.*, 20(3):371–402, 1986.
 - [104] V. J. Challis and J. K. Guest. Level set topology optimization of fluids in stokes flow. *Int. J. Numer. Meth. Eng.*, 79(10):1284–1308, 2009.
 - [105] A. Chambolle. A density result in two-dimensional linearized elasticity, and applications. *Arch. Ration. Mech. An.*, 167(3):211–233, 2003.
 - [106] D. Chenais. On the existence of a solution in a domain identification problem. *J. Math. Anal. Appl.*, 52(2):189–219, 1975.
 - [107] M. Cheney, D. Isaacson, and J. Newell. Electrical impedance tomography. *SIAM Rev.*, 41:85–101, 1999.
 - [108] K.-S. Cheng, D. Isaacson, J. C. Newell, and D. G. Gisser. Electrode models for electric current computed tomography. *IEEE T. Bio-med. Eng.*, 36(9):918–924, sep 1989.
 - [109] H.-W. Choi and M. Paraschivoiu. Adaptive computations of a posteriori finite element output bounds: a comparison of the hybrid-flux approach and the flux-free approach. *Comput. Method. Appl. M.*, 193(3638):4001 – 4033, 2004.
 - [110] P. W. Christensen and A. Klarbring. *An introduction to structural optimization*, volume 153 of *Solid Mechanics and its Applications*. Springer, New York, 2009.
 - [111] E. Chung, T. Chan, and X.-C. Tai. Electrical impedance tomography using level set representation and total variational regularization. *J. Comput. Phys.*, 205(1):357–372, 2005.
 - [112] P. G. Ciarlet. *Mathematical elasticity. Vol. I*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.
 - [113] P. Clément. Approximation by finite element functions using local regularization. *ESAIM: Math. Model. Num.*, 9(R2):77–84, 1975.
 - [114] S. Cochez-Dhondt and S. Nicaise. Equilibrated error estimators for discontinuous Galerkin methods. *Numer. Meth. Part. D. E.*, 24(5):1236–1252, 2008.

- [115] R. Codina and J. Baiges. Weak imposition of essential boundary conditions in the finite element approximation of elliptic problems with non-matching meshes. *Int. J. Numer. Meth. Eng.*, 104(7):624–654, 2015.
- [116] T. Colding and W. Minicozzi, II. *Minimal surfaces*, volume 4 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York, 1999.
- [117] R. Correa and A. Seeger. Directional derivative of a minimax function. *Nonlinear Anal.-Theor.*, 9(1):13–22, 1985.
- [118] M. Delfour and J.-P. Zolésio. *Shapes and geometries: analysis, differential calculus, and optimization*. SIAM, Philadelphia, USA, 2001.
- [119] J.-A. Désidéri. Cooperation and competition in multidisciplinary optimization: application to the aero-structural aircraft wing shape optimization. *Comput. Optim. Appl.*, 52(1):29–68, 2012.
- [120] P. Destuynder and B. Métivet. Explicit error bounds in a conforming finite element method. *Math. Comput.*, 68(228):1379–1396, 1999.
- [121] D. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012.
- [122] A. R. Díaz and M. P. Bendsøe. Shape optimization of structures for multiple loading conditions using a homogenization method. *Struct. Optimization*, 4(1):17–22, 1992.
- [123] G. Dogǎn, P. Morin, R. H. Nochetto, and M. Verani. Discrete gradient flows for shape optimization and applications. *Comput. Method. Appl. M.*, 196(3740):3898 – 3914, 2007. Special Issue Honoring the 80th Birthday of Professor Ivo Babuška.
- [124] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
- [125] J. Drury. *Ultrasonic Flaw Detection for Technicians*. Silverwing Ltd., 2004.
- [126] P. Duysinx, L. Van Miegroet, E. Lemaire, O. Brûls, and M. Bruyneel. Topology and generalized shape optimization: Why stress constraints are so important? *Int. J. Simul. Multidisci. Des. Optim.*, 2(4):253–258, 2008.
- [127] K. Eppler and H. Harbrecht. A regularized Newton method in electrical impedance tomography using shape Hessian information. *Control Cybern.*, 34(1):203–225, 2005.
- [128] K. Eppler and H. Harbrecht. Shape optimization for 3D electrical impedance tomography. In *Free and moving boundaries*, volume 252 of *Lect. Notes Pure Appl.*, pages 165–183. Chapman & Hall/CRC, Boca Raton, FL, 2007.
- [129] A. Ern, S. Nicaise, and M. Vohralík. An accurate $\mathbf{H}(\text{div})$ flux reconstruction for discontinuous Galerkin approximations of elliptic problems. *C. R. Acad. Sci. I-Math.*, 345(12):709–712, 2007.

-
- [130] A. Ern, A. Stephansen, and M. Vohralík. Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection-diffusion-reaction problems. *J. Comput. Appl. Math.*, 234(1):114–130, 2010.
- [131] A. Ern, A. Stephansen, and P. Zunino. A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.*, 29(2):235–256, 2009.
- [132] A. Ern and M. Vohralík. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous galerkin, and mixed discretizations. *SIAM J. Numer. Anal.*, 53(2):1058–1081, 2015.
- [133] H. A. Eschenauer, V. V. Kobelev, and A. Schumacher. Bubble method for topology and shape optimization of structures. *Struct. Optimization*, 8(1):42–51, 1994.
- [134] M. Farhloul and M. Fortin. Dual hybrid methods for the elasticity and the Stokes problems: a unified approach. *Numer. Math.*, 76(4):419–440, 1997.
- [135] L. Formaggia, S. Micheletti, and S. Perotto. Anisotropic mesh adaption in computational fluid dynamics: application to the advection-diffusion-reaction and the Stokes problems. *Appl. Numer. Math.*, 51(4):511–533, 2004.
- [136] L. Formaggia and S. Perotto. Anisotropic error estimates for elliptic problems. *Numer. Math.*, 94(1):67–92, 2003.
- [137] B. Fraeijs de Veubeke. Displacement and equilibrium models in the finite element method. In O. C. Zienkiewicz and G. S. Holister, editors, *Stress Analysis*, pages 145–197. John Wiley & Sons, Ltd., 1965.
- [138] B. Fraeijs de Veubeke. Stress function approach. In *Proceedings of the World Congress on Finite Element Methods in Structural Mechanics*, volume 1. Dorset, 1975.
- [139] I. Fumagalli, N. Parolini, and M. Verani. Shape optimization for Stokes flows: a finite element convergence analysis. *ESAIM: Math. Model. Num.*, 49(4):921–951, 2015.
- [140] A. L. Gain and G. H. Paulino. A critical comparative assessment of differential equation-driven methods for structural topology optimization. *Struct. Multidiscip. O.*, 48(4):685–710, 2013.
- [141] Z. Gao, Y. Ma, and H. Zhuang. Optimal shape design for Stokes flow via minimax differentiability. *Math. Comput. Model.*, 48(3-4):429–446, 2008.
- [142] G. N. Gatica. Analysis of a new augmented mixed finite element method for linear elasticity allowing \mathbb{RT}_0 - \mathbb{P}_1 - \mathbb{P}_0 approximations. *ESAIM: Math. Model. Num.*, 40(1):1–28, 2006.
- [143] G. N. Gatica, A. Márquez, and S. Meddahi. A new dual-mixed finite element method for the plane linear elasticity problem with pure traction boundary conditions. *Comput. Method. Appl. M.*, 197(9-12):1115–1130, 2008.

- [144] L. Gibiansky and A. Cherkaev. Design of composite plates of extremal rigidity. Preprint, Ioffe Physicotechnical Institute, 1984. In russian.
- [145] M. Giles and E. Süli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numer.*, 11:145–236, 2002.
- [146] R. Glowinski. Numerical simulation for some applied problems originating from continuum mechanics. In *Trends and applications of pure mathematics to mechanics (Palaiseau, 1983)*, volume 195 of *Lect. Notes Phys.*, pages 96–145. Springer, Berlin, 1984.
- [147] J. Goodman, R. V. Kohn, and L. Reyna. Numerical study of a relaxed variational problem from optimal design. *Comput. Method. Appl. M.*, 57(1):107–127, 1986.
- [148] P. Gould. *Introduction to Linear Elasticity*. Introduction to Linear Elasticity. Springer, 1993.
- [149] T. Grätsch and K.-J. Bathe. Goal-oriented error estimation in the analysis of fluid flows with structural interactions. *Comput. Method. Appl. M.*, 195:5673–5684, 2006.
- [150] J. Guedes and N. Kikuchi. Preprocessing and postprocessing for materials based on the homogenization method with adaptive finite element methods. *Comput. Method. Appl. M.*, 83(2):143 – 198, 1990.
- [151] J. K. Guest, J. H. Prévost, and T. Belytschko. Achieving minimum length scale in topology optimization using nodal design variables and projection functions. *Int. J. Numer. Meth. Eng.*, 61(2):238–254, 2004.
- [152] A. Habbal, J. Petersson, and M. Thellner. Multidisciplinary topology optimization solved as a Nash game. *Int. J. Numer. Meth. Eng.*, 61(7):949–963, 2004.
- [153] J. Hadamard. Mémoire sur le problème d’analyse relatif à l’équilibre des plaques élastiques encastrées. *B. Soc. Math. Fr.*, 1907.
- [154] M. Hanke, B. Harrach, and N. Hyvönen. Justification of point electrode models in electrical impedance tomography. *Math. Mod. Meth. Appl. S.*, 21(6):1395–1413, 2011.
- [155] R. Hanke, T. Fuchs, and N. Uhlmann. X-ray based methods for non-destructive testing and material characterization. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 591(1):14 – 18, 2008. Radiation Imaging Detectors 2007 - Proceedings of the 9th International Workshop on Radiation Imaging Detectors.
- [156] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, III. *J. Comput. Phys.*, 71(2):231 – 303, 1987.
- [157] R. Hartmann. Adjoint consistency analysis of discontinuous Galerkin discretizations. *SIAM J. Numer. Anal.*, 45(6):2671–2696 (electronic), 2007.
- [158] J. Haslinger and I. Hlaváček. Convergence of a finite element method based on the dual variational formulation. *Apl. Mat.*, 21(1):43–65, 1976.

-
- [159] F. Hecht. BAMG: Bidimensional Anisotropic Mesh Generator. Research report, INRIA, 2006.
- [160] F. Hecht. New development in **FreeFem++**. *J. Numer. Math.*, 20(3-4):251–265, 2012.
- [161] W. Helfrich. Elastic properties of lipid bilayers: theory and possible experiments. *Z. Naturforsch.*, 28(11):693–703, 1973.
- [162] A. Henrot and M. Pierre. *Variation et optimisation de formes: Une analyse géométrique*. Springer, 2005.
- [163] M. Hintermüller and A. Laurain. Electrical impedance tomography: from topology to shape. *Control Cybern.*, 37(4):913–933, 2008.
- [164] M. Hintermüller, A. Laurain, and A. A. Novotny. Second-order topological expansion for electrical impedance tomography. *Adv. Comput. Math.*, 36(2):235–265, 2012.
- [165] R. Hiptmair and A. Paganini. Shape optimization by pursuing diffeomorphisms. *Comput. Methods Appl. Math.*, 15(3):291–305, 2015.
- [166] R. Hiptmair, A. Paganini, and S. Sargheini. Comparison of approximate shape gradients. *BIT*, 55(2):459–485, 2015.
- [167] D. Holder. *Electrical Impedance Tomography: Methods, History and Applications*. Series in Medical Physics and Biomedical Engineering. CRC Press, 2004.
- [168] C. O. Horgan. Korn’s inequalities and their applications in continuum mechanics. *SIAM Rev.*, 37(4):491–511, 1995.
- [169] N. Hyvönen. Approximating idealized boundary data of electric impedance tomography by electrode measurements. *Math. Mod. Meth. Appl. S.*, 19(7):1185–1202, 2009.
- [170] B. Jin and P. Maass. An analysis of electrical impedance tomography with applications to Tikhonov regularization. *ESAIM: Contr. Op. Ca. Va.*, 18(4):1027–1048, 2012.
- [171] C. Johnson and P. Hansbo. Adaptive finite element methods in computational mechanics. *Comput. Method. Appl. M.*, 101(1):143 – 181, 1992.
- [172] D. W. Kelly, J. P. De S. R. Gago, O. C. Zienkiewicz, and I. Babuška. A posteriori error analysis and adaptive processes in the finite element method: Part I-error analysis. *Int. J. Numer. Meth. Eng.*, 19(11):1593–1619, 1983.
- [173] G. K. W. Kenway and J. R. R. A. Martins. Multipoint high-fidelity aerostructural optimization of a transport aircraft configuration. *J. Aircraft*, 51(1):144–160, jan 2014.
- [174] N. Kikuchi, K. Chung, T. Torigaki, and J. Taylor. Adaptive Finite Element Methods for shape optimization of linearly elastic structures. *Comput. Method. Appl. M.*, 57:67–89, 1986.
- [175] K. Y. Kim. A posteriori error analysis for locally conservative mixed methods. *Math. Comput.*, 76(257):43–66, 2007.

- [176] R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *P. Natl. Acad. Sci. U.S.A.*, 95(15):8431–8435, 1998.
- [177] B. Kiniger and B. Vexler. A priori error estimates for finite element discretizations of a shape optimization problem. *ESAIM: Math. Model. Num.*, 47(6):1733–1763, 2013.
- [178] U. Kirsch. *Optimum structural design: concepts, methods, and applications*. McGraw-Hill, 1981.
- [179] A. Klawonn and G. Starke. A preconditioner for the equations of linear elasticity discretized by the peers element. *Numer. Linear Algebr.*, 11(5-6):493–510, 2004.
- [180] R. Kohn and M. Vogelius. Relaxation of a variational method for impedance computed tomography. *Commun. Pur. Appl. Math.*, 40(6):745–777, 1987.
- [181] R. V. Kohn and G. Strang. Optimal design and relaxation of variational problems. I. *Commun. Pur. Appl. Math.*, 39(1):113–137, 1986.
- [182] R. V. Kohn and G. Strang. Optimal design and relaxation of variational problems. II. *Commun. Pur. Appl. Math.*, 39(2):139–182, 1986.
- [183] R. V. Kohn and G. Strang. Optimal design and relaxation of variational problems. III. *Commun. Pur. Appl. Math.*, 39(3):353–377, 1986.
- [184] S. Kreissl and K. Maute. Levelset based fluid topology optimization using the extended finite element method. *Struct. Multidiscip. O.*, 46(3):311–326, 2012.
- [185] P. Ladevèze and L. Chamoin. *The Constitutive Relation Error Method: A General Verification Tool*, pages 59–94. Springer International Publishing, Cham, 2016.
- [186] P. Ladevèze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20(3):485–509, 1983.
- [187] T. Lassila and G. Rozza. Parametric free-form shape design with PDE models and reduced basis method. *Comput. Method. Appl. M.*, 199(2324):1583 – 1592, 2010.
- [188] M. Laumen. Newton’s method for a class of optimal shape design problems. *SIAM J. Optimiz.*, 10(2):503–533, 2000.
- [189] A. Laurain and K. Sturm. Distributed shape derivative via averaged adjoint method and applications. *ESAIM: Math. Model. Num.*, 50(4):1241–1267, 2016.
- [190] T. Lewiński, M. Zhou, and G. Rozvany. Extended exact solutions for least-weight truss layouts - Part I: Cantilever with a horizontal axis of symmetry. *Int. J. Mech. Sci.*, 36(5):375 – 398, 1994.
- [191] R. Luce and B. I. Wohlmuth. A local a posteriori error estimator based on equilibrated fluxes. *SIAM J. Numer. Anal.*, 42(4):1394–1414, 2004.
- [192] K. A. Lurie and A. V. Cherkasov. *Topics in the Mathematical Modelling of Composite Materials*, chapter Effective Characteristics of Composite Materials and the Optimal Design of Structural Elements, pages 175–271. Birkhäuser Boston, Boston, MA, 1997.

-
- [193] K. A. Lurie, A. V. Cherkaev, and A. V. Fedorov. Regularization of optimal design problems for bars and plates, part 1. *J. Optimiz. Theory App.*, 37(4):499–522, 1982.
- [194] K. A. Lurie, A. V. Cherkaev, and A. V. Fedorov. Regularization of optimal design problems for bars and plates, part 2. *J. Optimiz. Theory App.*, 37(4):523–543, 1982.
- [195] L. Machiels, Y. Maday, and A. T. Patera. A flux-free nodal Neumann subproblem approach to output bounds for partial differential equations. *C. R. Acad. Sci. I-Math.*, 330(3):249 – 254, 2000.
- [196] J. E. Marsden and T. J. R. Hughes. *Mathematical foundations of elasticity*. Dover Publications, Inc., New York, 1994. Corrected reprint of the 1983 original.
- [197] J. R. R. A. Martins and A. B. Lambe. Multidisciplinary design optimization: A survey of architectures. *AIAA J.*, 51(9):2049–2075, sep 2013.
- [198] A. Michell. LVIII. The limits of economy of material in frame-structures. *Philosophical Magazine Series 6*, 8(47):589–597, 1904.
- [199] S. G. Mikhlin. *Variational methods in mathematical physics*. Translated by T. Boddington; editorial introduction by L. I. G. Chambers. A Pergamon Press Book. The Macmillan Co., New York, 1964.
- [200] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Numerical mathematics and scientific computation. Clarendon, Oxford, 2001. Autre tirage : 2005.
- [201] P. Morin, R. H. Nochetto, M. Pauletti, and M. Verani. Adaptive finite element method for shape optimization. *ESAIM: Contr. Op. Ca. Va.*, 18:1122–1149, 2012.
- [202] P. Morin, R. H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.
- [203] P. Morin, R. H. Nochetto, and K. G. Siebert. Local problems on stars: A posteriori error estimators, convergence, and performance. *Math. Comput.*, 72(243):1067–1097, 2003.
- [204] M. E. Morley. A family of mixed finite elements for linear elasticity. *Numer. Math.*, 55(6):633–666, 1989.
- [205] I. Mozolevski and S. Prudhomme. Goal-oriented error estimation based on equilibrated-flux reconstruction for finite element approximations of elliptic problems. *Comput. Method. Appl. M.*, 288:127–145, 2015.
- [206] D. J. Munk, G. A. Vio, and G. P. Steven. Topology and shape optimization methods using evolutionary algorithms: a review. *Struct. Multidiscip. O.*, 52(3):613–631, 2015.
- [207] F. Murat and J. Simon. Sur le contrôle par un domaine géométrique. Internal Report 76 015, Laboratoire d’Analyse Numérique de l’Université de Paris 6, 1976.
- [208] F. Murat and L. Tartar. Calcul des Variations et homogénéisation. *Coll. Dir. Etudes et Recherches EDF*, 57:319–369, 1985.

- [209] F. Murat and L. Tartar. Optimality conditions and homogenization. In *Nonlinear variational problems (Isola d'Elba, 1983)*, volume 127 of *Res. Notes in Math.*, pages 1–8. Pitman, Boston, MA, 1985.
- [210] S. Nicaise, K. Witowski, and B. Wohlmuth. An a posteriori error estimator for the lam equation based on equilibrated fluxes. *IMA J. Numer. Anal.*, 28(2):331–353, 2008.
- [211] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 36(1):9–15, 1971.
- [212] J. Nocedal and S. Wright. *Numerical optimization*. Springer-Verlag New York, USA, 1999.
- [213] R. H. Nochetto, K. G. Siebert, and A. Veiser. Theory of adaptive finite element methods: An introduction. In R. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 409–542. Springer Berlin Heidelberg, 2009.
- [214] R. H. Nochetto, A. Veiser, and M. Verani. A safeguarded dual weighted residual method. *IMA J. Numer. Anal.*, 29(1):126–140, 2009.
- [215] J. Oden and S. Prudhomme. Goal-oriented error estimation and adaptivity for the Finite Element Method. *Comput. Math. Appl.*, 41(5-6):735 – 756, 2001.
- [216] S. Ohnibus, E. Stein, and E. Walhorn. Local error estimates of FEM for displacements and stresses in linear elasticity by solving local Neumann problems. *Int. J. Numer. Meth. Eng.*, 52(7):727–746, 2001.
- [217] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(1):12–49, 1988.
- [218] S. Osher and C.-W. Shu. High-order essentially nonoscillatory schemes for hamilton-jacobi equations. *SIAM J. Numer. Anal.*, 28(4):907–922, 1991.
- [219] O. Pantz. Sensibilité de l'équation de la chaleur aux sauts de conductivité. *C. R. Acad. Sci. I-Math.*, (341):333–337, 2005.
- [220] O. Pantz and K. Trabelsi. Simultaneous shape, topology, and homogenized properties optimization. *Struct. Multidiscip. O.*, 34(4):361–365, 2007.
- [221] O. Pantz and K. Trabelsi. A post-treatment of the homogenization method for shape optimization. *SIAM J. Control Optim.*, 47(3):1380–1398, 2008.
- [222] M. Paraschivoiu, J. Peraire, and A. T. Patera. A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations. *Comput. Method. Appl. M.*, 150(14):289 – 312, 1997. Symposium on Advances in Computational Mechanics.
- [223] N. Parés, J. Bonet, A. Huerta, and J. Peraire. The computation of bounds for linear-functional outputs of weak solutions to the two-dimensional elasticity equations. *Comput. Method. Appl. M.*, 195(46):406 – 429, 2006. Adaptive Modeling and Simulation.

-
- [224] N. Parés, P. Díez, and A. Huerta. Subdomain-based flux-free a posteriori error estimators. *Comput. Method. Appl. M.*, 195(4-6):297–323, 2006.
 - [225] N. Parés, P. Díez, and A. Huerta. Exact bounds for linear outputs of the convection-diffusion-reaction equation using flux-free error estimates. In M. Griebel and M. Schweitzer, editors, *Meshfree Methods for Partial Differential Equations IV*, volume 65 of *Lecture Notes in Computational Science and Engineering*, pages 215–230. Springer Berlin Heidelberg, 2008.
 - [226] A. Pechstein and J. Schöberl. Tangential-displacement and normal-normal-stress continuous mixed finite elements for elasticity. *Math. Mod. Meth. Appl. S.*, 21(8):1761–1782, 2011.
 - [227] J. Petersson. A finite element analysis of optimal variable thickness sheets. *SIAM J. Numer. Anal.*, 36(6):1759–1778, 1999.
 - [228] J. Petersson and O. Sigmund. Slope constrained topology optimization. *Int. J. Numer. Meth. Eng.*, 41(8):1417–1434, 1998.
 - [229] O. Pironneau. *Optimal Shape Design for Elliptic Systems*. Springer-Verlag, 1984.
 - [230] S. Prudhomme, F. Nobile, L. Chamoin, and J. T. Oden. Analysis of a subdomain-based error estimator for finite element approximations of elliptic problems. *Numer. Meth. Part. D. E.*, 20(2):165–192, 2004.
 - [231] S. Prudhomme and J. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Comput. Method. Appl. M.*, 176(14):313 – 331, 1999.
 - [232] S. Prudhomme, J. Oden, T. Westermann, J. Bass, and M. Botkin. Practical methods for a posteriori error estimation in engineering applications. *Int. J. Numer. Meth. Eng.*, 56(8):1193–1224, 2003.
 - [233] P.-A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. In *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*, pages 292–315. Lecture Notes in Math., Vol. 606. Springer, Berlin, 1977.
 - [234] E. Reissner. On a variational theorem in elasticity. *J. Math. Phys. Camb.*, 29(1-4):90–95, 4 1950.
 - [235] S. Repin. A posteriori error estimation for variational problems with uniformly convex functionals. *Math. Comput.*, 69(230):481–500, 2000.
 - [236] J. Roche. Adaptive Newton-like method for shape optimization. *Control Cybern.*, 34(1):363–377, 2005.
 - [237] M. E. Rognes, R. C. Kirby, and A. Logg. Efficient assembly of $H(\text{div})$ and $H(\text{curl})$ conforming finite elements. *SIAM J. Sci. Comput.*, 31(6):4130–4151, 2009/10.
 - [238] G. I. N. Rozvany. *Structural design via optimality criteria*, volume 8 of *Mechanics of Elastic and Inelastic Solids*. Kluwer Academic Publishers Group, Dordrecht, 1989. The Prager approach to structural optimization.

- [239] G. I. N. Rozvany. A critical review of established methods of structural topology optimization. *Struct. Multidiscip. O.*, 37(3):217–237, 2009.
- [240] M. Rüter, T. Gerasimov, and E. Stein. Goal-oriented explicit residual-type error estimates in XFEM. *Comput. Mech.*, 52(2):361–376, 2013.
- [241] A. M. Sauer-Budge, J. Bonet, A. Huerta, and J. Peraire. Computing bounds for linear functionals of exact weak solutions to Poisson’s equation. *SIAM J. Numer. Anal.*, 42(4):1610–1630, 2004.
- [242] M. Save, W. Prager, and G. Sacchi. *Structural optimization. Vol. 1*, volume 34 of *Mathematical Concepts and Methods in Science and Engineering*. Plenum Press, New York, 1985. Optimality criteria.
- [243] A. Schleupen, K. Maute, and E. Ramm. Adaptive FE-procedures in shape optimization. *Struct. Multidiscip. O.*, 19:282–302, 2000.
- [244] U. Schramm and W. D. Pilkey. The coupling of geometric descriptions and finite elements using NURBS - A study in shape optimization. *Finite Elem. Anal. Des.*, 15(1):11 – 34, 1993.
- [245] L. Selvakumaran, Q. Long, S. Prudhomme, and G. Lubineau. On the detectability of transverse cracks in laminated composites using electrical potential change measurements. *Compos. Struct.*, 121:237 – 246, 2015.
- [246] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *P. Natl. Acad. Sci. U.S.A.*, 93(4):1591–1595, 1996.
- [247] R. Sevilla, S. Fernández-Méndez, and A. Huerta. NURBS-enhanced finite element method (NEFEM). *Int. J. Numer. Meth. Eng.*, 76(1):56–83, 2008.
- [248] K. Shahbazi. An explicit expression for the penalty parameter of the interior penalty method. *J. Comput. Phys.*, 205(2):401 – 407, 2005.
- [249] O. Sigmund. On the design of compliant mechanisms using topology optimization. *Mech. Struct. Mach.*, 25(4):493–524, 1997.
- [250] O. Sigmund. Morphology-based black and white filters for topology optimization. *Struct. Multidiscip. O.*, 33(4):401–424, 2007.
- [251] O. Sigmund. On the usefulness of non-gradient approaches in topology optimization. *Struct. Multidiscip. O.*, 43(5):589–596, 2011.
- [252] O. Sigmund and P. M. Clausen. Topology optimization using a mixed formulation: An alternative way to solve pressure load problems. *Comput. Method. Appl. M.*, 196(13–16):1874 – 1889, 2007.
- [253] O. Sigmund and K. Maute. Topology optimization approaches. *Struct. Multidiscip. O.*, 48(6):1031–1055, 2013.
- [254] J. Simon. Differentiation with respect to the domain in boundary value problems. *Numer. Func. Anal. Opt.*, 2(7-8):649–687, 1980.

-
- [255] J. Sokołowski and A. Zochowski. On the topological derivative in shape optimization. *SIAM J. Control Optim.*, 37(4):1251–1272, 1999.
- [256] J. Sokołowski and J. P. Zolésio. *Introduction to shape optimization: shape sensitivity analysis*. Springer-Verlag, 1992.
- [257] E. Somersalo, M. Cheney, and D. Isaacson. Existence and Uniqueness for Electrode Models for Electric Current Computed Tomography. *SIAM J. Appl. Math.*, 52(4):1023–1040, aug 1992.
- [258] E. Stein and R. Rolfes. Mechanical conditions for stability and optimal convergence of mixed finite elements for linear plane elasticity. *Comput. Method. Appl. M.*, 84(1):77–95, 1990.
- [259] R. Stenberg. On the construction of optimal mixed finite element methods for the linear elasticity problem. *Numer. Math.*, 48(4):447–462, 1986.
- [260] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53(5):513–538, 1988.
- [261] R. Stenberg. Two low-order mixed methods for the elasticity problem. In *The mathematics of finite elements and applications, VI (Uxbridge, 1987)*, pages 271–280. Academic Press, London, 1988.
- [262] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.
- [263] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.
- [264] M. Stolpe and K. Svanberg. An alternative interpolation scheme for minimum compliance topology optimization. *Struct. Multidiscip. O.*, 22(2):116–124, 2001.
- [265] K. Svanberg. The method of moving asymptotes a new method for structural optimization. *Int. J. Numer. Meth. Eng.*, 24(2):359–373, 1987.
- [266] V. Šverák. On optimal shape design. *J. Math. Pure. Appl. (9)*, 72(6):537–551, 1993.
- [267] J. Sylvester and G. Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Ann. Math.*, 125:153–169, 1987.
- [268] J.-M. Thomas. Méthode des éléments finis hybrides d’aux pour les problèmes elliptiques du second ordre. *ESAIM: Math. Model. Num.*, 10(R3):51–79, 1976.
- [269] N. P. van Dijk, K. Maute, M. Langelaar, and F. van Keulen. Level-set methods for structural topology optimization: a review. *Struct. Multidiscip. O.*, 48(3):437–472, 2013.
- [270] L. Van Mieghem and P. Duysinx. Stress concentration minimization of 2d fillets using X-FEM and level set description. *Struct. Multidiscip. O.*, 33(4):425–438, 2007.
- [271] G. N. Vanderplaats and F. Moses. Structural optimization by methods of feasible directions. *Comput. Struct.*, 3(4):739 – 755, 1973.

- [272] T. Vejchodský. Complementary error bounds for elliptic systems and applications. *Appl. Math. Comput.*, 219(13):7194 – 7205, 2013.
- [273] R. Verfürth. A posteriori error estimators for the Stokes equations. *Numer. Math.*, 55(3):309–325, 1989.
- [274] R. Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods*. Oxford University Press, 2013.
- [275] M. Vohralík. Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients. *J. Sci. Comput.*, pages 1–40, 2011.
- [276] M. Wallin, M. Ristinmaa, and H. Askfelt. Optimal topologies derived from a phase-field method. *Struct. Multidiscip. O.*, 45(2):171–183, 2012.
- [277] M. Y. Wang and X. Wang. PDE-driven level sets, shape sensitivity and curvature flow for structural topology optimization. *CMES - Comp. Model. Eng.*, 6(4):373–395, 2004.
- [278] M. Y. Wang, X. Wang, and D. Guo. A level set method for structural topology optimization. *Comput. Method. Appl. M.*, 192(12):227 – 246, 2003.
- [279] M. Y. Wang and S. Zhou. Phase field: a variational method for structural topology optimization. *CMES - Comp. Model. Eng.*, 6(6):547–566, 2004.
- [280] A. Wexler, B. Fry, and M. R. Neuman. Impedance-computed tomography algorithm and system. *Appl. Optics*, 24(23):3985–3992, 1985.
- [281] T. Yamada, K. Izui, S. Nishiwaki, and A. Takezawa. A topology optimization method based on the level set method incorporating a fictitious interface energy. *Comput. Method. Appl. M.*, 199(4548):2876 – 2891, 2010.
- [282] S. Yamasaki, T. Nomura, A. Kawamoto, K. Sato, and S. Nishiwaki. A level set-based topology optimization method targeting metallic waveguide design problems. *Int. J. Numer. Meth. Eng.*, 87(9):844–868, 2011.
- [283] L. Younes. *Shapes and diffeomorphisms*, volume 171 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2010.

Titre : Estimations d'erreur a posteriori quantitatives pour l'approximation des problèmes d'optimisation de forme par la méthode des éléments finis

Mots clefs : Optimisation de forme, estimations d'erreur a posteriori, Algorithme de Descente Certifiée, dérivée de forme volumique, Tomographie d'Impédance Electrique, minimisation de la compliance

Résumé : Cette thèse est consacrée à la construction d'une procédure de certification de la direction de descente dans des algorithmes de gradient en optimisation de forme grâce à des estimations a posteriori de l'erreur introduite par l'approximation de la dérivée de forme par la méthode des éléments finis. En construisant une borne supérieure certifiée et explicitement calculable de cette erreur, l'Algorithme de Descente Certifiée pour l'optimisation de forme identifie une véritable direction de descente à chaque itération et permet d'établir un critère d'arrêt fiable basé sur la norme de la dérivée de forme. Deux applications sont abordées dans la thèse. Premièrement, on considère le problème scalaire d'identification de forme en tomographie d'impédance électrique et on étudie différentes estimations d'erreur. Une première approche est basée sur le principe de l'énergie complémentaire et nécessite

la résolution de problèmes globaux additionnels. Afin de réduire le coût de calcul de la procédure de certification, une estimation qui dépend seulement de quantités locales est dérivée par la reconstruction des flux équilibrés. Après avoir validé les estimations de l'erreur pour un cas 2D, des résultats numériques sont présentés. Une deuxième application est centrée sur le problème vectoriel de la conception optimale des structures élastiques. Dans ce contexte, on calcule l'expression volumique de la dérivée de forme de la compliance à partir de la formulation primale en déplacements et de celle duale mixte de l'équation de l'élasticité linéaire. Quelques résultats préliminaires pour la minimisation de la compliance sous une contrainte de volume en 2D sont obtenus à l'aide de l'Algorithme de Variation de Frontière et une estimation de l'erreur de la dérivée de forme est calculée.

Title : Quantitative a posteriori error estimators in Finite Element-based shape optimization

Keywords : Shape optimization, a posteriori error estimators, Certified Descent Algorithm, volumetric shape gradient, Electrical Impedance Tomography, compliance minimization

Abstract : This Ph.D. thesis is devoted to the construction of a certification procedure for the descent direction in gradient-based shape optimization methods using a posteriori estimators of the error due to the Finite Element approximation of the shape gradient. By deriving a fully computable certified upper bound of this error, the Certified Descent Algorithm for shape optimization is able to identify a genuine descent direction at each iteration and features a reliable stopping criterion based on the norm of the shape gradient. Two applications are tackled in the thesis. First, we consider the scalar inverse problem of Electrical Impedance Tomography and we investigate several a posteriori estimators. A first procedure is inspired by the complementary energy principle and requires the solu-

tion of additional global problems. To reduce the computational cost of the certification step, an estimator which depends solely on local quantities is derived via an equilibrated fluxes approach. The estimators are validated for a 2D case and some numerical tests are presented. A second application focuses on the vectorial problem of optimal design of elastic structures. Within this framework, we derive the volumetric expression of the shape gradient of the compliance using the pure displacement and the dual mixed formulations of the linear elasticity equation. Some preliminary numerical tests are performed to minimize the compliance under a volume constraint in 2D using the Boundary Variation Algorithm and an a posteriori error estimator for the shape gradient is derived.