THÈSE

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

École doctorale des sciences exactes et leurs applications

présentée et soutenue publiquement le 29 avril 2024

par **Giulio Gargantini**

pour l'obtention du

Doctorat de l'Université de Pau et des Pays de l'Adour

(Mathématiques)

Optimisation de forme et incertitudes dans la conception de structures mécaniques

COMPOSITION DU JURY

Rapporteurs :	Grégoire Allaire	École Polytechnique - CMAP
	Christian Rey	Safran Tech
Examinateurs :	Marc Albertelli	Ansys
	Helmut Harbrecht	Universität Basel
	Chiara Nardoni	IRT SystemX
Invité :	Charles Dapogny	Université Grenoble Alpes - LJK
Directeur de thèse :	Marc Dambrine	Université de Pau et des pays de l'Adour - LMAP
Co-directeurs de thèse :	Fabien Caubet	Université de Pau et des pays de l'Adour - LMAP
	Jérôme Maynadier	Safran Helicopter Engines

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

École doctorale des sciences exactes et leurs applications

Optimisation de forme et incertitudes dans la conception de structures mécaniques

THÈSE

présentée et soutenue publiquement le 29 avril 2024

pour l'obtention du

Doctorat de l'Université de Pau et des Pays de l'Adour

(Mathématiques)

par

Giulio Gargantini

Composition du jury

Rapporteurs :	Grégoire Allaire	École Polytechnique - CMAP
	Christian Rey	Safran Tech
Examinateurs :	Marc Albertelli	Ansys
	Helmut Harbrecht	Universität Basel
	Chiara Nardoni	IRT SystemX
Invité :	Charles Dapogny	Université Grenoble Alpes - LJK
Directeur de thèse :	Marc Dambrine	Université de Pau et des pays de l'Adour - LMAP
Co-directeurs de thèse :	Fabien Caubet	Université de Pau et des pays de l'Adour - LMAP
	Jérôme Maynadier	Safran Helicopter Engines

Laboratoire de Mathématiques et de leurs Applications

Résumé

Cette thèse porte sur l'optimisation de la forme de structures élastiques soumises à des chargements mécaniques incertains. On a considéré trois approches différentes pour la prise en compte des incertitudes en optimisation de forme. Dans la première approche, on a étudié un problème d'optimisation avec une contrainte sur le maximum d'une fonctionnelle critique. Ensuite, on a examiné le problème d'optimisation topologique robuste, où la contrainte porte sur l'espérance d'une fonctionnelle d'intérêt. Enfin, on a abordé le problème de l'optimisation topologique fiable, qui vise à contrôler la probabilité d'une variable aléatoire de dépasser un seuil donné. Toutes les contraintes ont été imposées sur des quantités couramment manipulées en mécanique des structures, telles que la compliance et la contrainte de von Mises. On a adopté la méthode de variation des frontières de Hadamard pour dériver l'objectif et les contraintes des problèmes d'optimisation par rapport à la forme, et on a représenté les structures par la méthode des lignes de niveaux. Les résultats théoriques sont supportés par des simulations numériques en deux et trois dimensions.

Mots-clés: Optimisation de forme, élasticité linéaire, quantification d'incertitudes, méthode des lignes de niveau, contrainte de von Mises.

Abstract

The present thesis focuses on the optimization of linear elastic structures subject to uncertain mechanical loads. We studied three different approaches to factor uncertainties in shape optimization. In the first approach we considered an optimization problem with a constraint on the maximum of a critical functional. Then, we studied the case of robust topology optimization, where the constraint holds on the expectation of a functional of interest. Finally, we addressed the problem of reliability-based topology optimization, aiming to control the probability of a stochastic quantity not to exceed a given threshold. All constraints have been imposed on quantities of proven interest in structural mechanics, as the compliance and the von Mises stress. We adopted Hadamard's boundary variation method to differentiate the objective and the constraints of the optimization problems with respect to the shape, and we represented the structures by the level-set method. The theoretical results are supported by numerical simulations in two and three dimensions.

Keywords: Shape optimization, linear elasticity, uncertainty quantification, level-set method, von Mises stress.

Remerciements

Wir sollen heiter Raum um Raum durchschreiten,
An keinem wie an einer Heimat hängen,
Der Weltgeist will nicht fesseln uns und engen,
Er will uns Stuf' um Stufe heben, weiten.

Franchissons donc, sereins, espace après espace ;
n'acceptons en aucun les liens d'une patrie,
pour nous l'esprit du monde n'a ni chaînes, ni murs ;
par degrés il veut nous hausser, nous grandir.

Hermann Hesse, *Le Jeu des perles de verre*
Trad. Jacques Martin

Avant de commencer, je tiens à exprimer en quelques lignes ma gratitude envers tous ceux qui m'ont accompagné le long de ces trois ans de thèse, et sans lesquels ce travail n'aurait pas vu le jour dans sa forme actuelle.

Mes plus grands remerciements vont, en premier abord, à mes encadrants académiques Marc Dambrine et Fabien Caubet, pour m'avoir initié au domaine de l'optimisation de forme, et pour leur soutien constant tout au long de mon parcours. Leur disponibilité pour répondre à mes questions, leur engagement lors de nos discussions dans les moments d'impasse, ainsi que leurs suggestions toujours pertinentes pour enrichir mes recherches, ont été d'une importance capitale pour la réussite de cette thèse, tant d'un point de vue humain que scientifique.

Également, je remercie mon encadrant industriel Jérôme Maynadier pour m'avoir accueilli dans l'équipe Méthodes et Outils de Safran Helicopter Engines et pour m'avoir accompagné dans cette première expérience dans un grand groupe industriel. Son expertise dans les mathématiques et sa connaissance approfondie des méthodes et des besoins de Safran Helicopter Engines ont grandement contribué à orienter ma thèse vers des cas d'application concrètes et pertinentes pour l'industrie.

Je remercie Grégoire Allaire et Christian Rey d'avoir accepté de rapporter ma thèse et pour avoir consacré du temps à la lecture de mon manuscrit. Je suis également reconnaissant vers Marc Albertelli, Chiara Nardoni et Helmut Harbrecht pour avoir accepté de faire partie du jury. Je tiens à exprimer ma gratitude particulière envers ce dernier pour son accueil chaleureux lors de mon séjour à l'université de Bâle et pour ses précieux renseignements sur le traitement des aspects aléatoires et sur la propagation d'incertitudes. Je souhaite également remercier Charles Dapogny pour l'intérêt prêté à mon travail, pour la mise en œuvre des outils numériques utilisés dans cette thèse, pour son assistance dans le déploiement, et pour sa disponibilité à assister à ma soutenance.

Je ne saurais assez remercier l'équipe du service Méthodes et Outils de Safran Helicopter Engines pour leur accueil chaleureux. Je remercie Olivier pour le suivi continu de ma thèse, Loïc et Julien pour leur support avec python et Z-set, et Patricio et Anthony pour leur bienveillance. Un merci tout particulier va à Alexandre et Marianne pour leur travail avec les ressources de calcul de SHE : sans eux l'application de mon travail à des structures industrielles n'aurait pas été possible. Enfin, je souhaite remercier les alternants Louis, Clément Perron, Clément Port et Victor pour les moments agréables partagés et la bonne ambiance au sein du service.

Mes années à Pau n'auraient pas été aussi enrichissantes sans mes chers collègues doctorants (et, pour certains, docteurs) du LMAP. Je tiens donc à exprimer ma gratitude envers Salah et Aymeric pour m'avoir accompagné lors des premières réunions de la communauté d'optimisation de forme, ainsi qu'envers Joyce pour nos conversations avant et après le déménagement de notre bureau poussiéreux. Mes remerciements vont également à Ibtissem et Aimen, ainsi qu'à Nicolas, Sarah, Adrien, Magdalena, Rongquan, Manos, Puneeth, Thibault, Loïc, Sara et Nassim. Un remerciement spécial va à Rodrigo, et par extension à Amanda, pour avoir animé les week-ends avec des sorties, des soirées et des dîners toujours excellents. Je leur souhaite à tous le meilleur dans leurs projets futurs.

Je suis très reconnaissant aussi envers tous les membres de la communauté académique qui qui m'ont apporté leur aide au cours de ce travail, en particulier à Florian Feppon pour avoir mis à disposition les outils développés au cours de sa thèse, et Algiane Froehly pour l'aide fournie avec *mmg*. Je remercie chaleureusement Viacheslav pour la collaboration au sujet du problème élasto-thermique, son hospitalité, et pour les discussions toujours intéressantes.

Je remercie de tout mon cœur Coline de m'avoir accompagné au cours de cette thèse à partir, littéralement, du premier jour. Je suis très heureux d'avoir partagé avec elle les moments les plus joyeux de ces trois ans, et très reconnaissant de sa présence lors des épreuves plus difficiles. Malgré la distance, parfois intercontinentale, entre nous, elle a toujours su m'offrir la chaleur de son sourire, son précieux point de vue de collègue doctorante, et son soutien en toutes circonstances.

Enfin, je remercie ma famille d'avoir toujours cru en moi et de m'avoir toujours supporté. Sans leur encouragement, je n'aurais jamais pu entreprendre ce parcours d'études en France. Si des longues distances nous séparaient géographiquement, ils ont toujours été présents pour célébrer mes succès et m'apporter leur soutien dans les moments difficiles. Pour tout ce qu'ils ont fait pour moi et pour leur amour sans limites, je serai toujours reconnaissant.

Contents

Introduction générale

Généralités en optimisation de structures

Les progrès technologiques de ces dernières années ont entraîné une augmentation de la puissance de calcul, et ont permis le développement de techniques avancées pour la conception de structures complexes. Les techniques d'optimisation de forme et topologique ont été appliquées dans plusieurs branches de l'ingénierie, du génie civil [237, 257], à l'industrie aérospatiale [186, 183, 199], au secteur de l'automobile [227, 204, 175], aux échangeurs de chaleur [264, 259, 238, 243, 4], à la micro-électronique et photonique [146, 109, 163], à la biomécanique [161, 144, 127] et au-delà. Les progrès de la fabrication additive et des fraiseuses à commande numérique ont permis la réalisation de structures complexes, et ont favorisé le développement de méthodes d'optimisation de forme et topologique [260, 243, 15]. Consultez [207] pour une liste exhaustive des différentes applications de l'optimisation de forme dans différents domaines de l'ingénierie. Les approches à l'optimisation de structures sont généralement divisées en trois catégories principales [238, 177] :

- **optimisation paramétrique**, où la structure d'intérêt est représentée par un nombre réduit de variables ;
- **optimisation de forme**, où le paramètre d'optimisation est la frontière de la structure elle-même ;
- **optimisation topologique**, où la variable d'optimisation est la distribution du matériau dans un domaine de référence, sans *a priori* sur la topologie de la structure optimale.

Méthodes de densité et homogénéisation

Les *méthodes de densité* sont une vaste classe de techniques d'optimisation topologique, dont l'origine remonte à la méthode SIMP (*Solid Isotropic Material with Penalization* ou matériau solide isotrope avec pénalisation) conçue par Bendsøe et Sigmund [47]. Cette méthode considère un domaine de calcul D dans lequel est défini une densité $\rho : D \to [0, 1]$. Les propriétés des matériaux en chaque point du domaine sont une interpolation des propriétés de deux matériaux différents (comme une matrice solide et un fluide, ou deux matériaux élastiques) en fonction de ρ. Afin d'assurer que le problème d'optimisation est bien posé dans le cas de l'optimisation de structures élastiques, un matériau *ersatz* suffisamment léger remplace l'espace vide entourant le domaine. Le paramètre d'optimisation est le paramètre de densité ρ lui-même de sorte que, après la convergence de l'algorithme d'optimisation, la distribution optimale des matériaux dans le domaine de calcul est identifiée par les régions où ρ est proche de 0 ou de 1. La méthode SIMP a été largement étudiée et adaptée à l'optimisation des structures élastiques [45], et aux

structures couplées à des problèmes de dynamique des fluides [53]. En outre, cette méthode a été introduite depuis longtemps dans les logiciels commerciaux d'éléments finis tels que ANSYS, Optistruct et ABAQUS [262] et dans des logiciels de dynamique des fluides [195].

La convergence vers une solution avec une distinction nette entre les deux phases dépend de la nature du problème et des paramètres de l'algorithme d'optimisation. Par conséquent, de larges zones grises peuvent exister même à la fin du processus d'optimisation. L'application de contraintes sur l'épaisseur de la structure est rendue particulièrement difficile par l'absence d'une frontière nette. Des contraintes relatives à la topologie du domaine ou à la fabricabilité présentent des difficultés similaires. Comme affirmé dans [167], de tels objectifs nécessitent l'application d'algorithmes de reconnaissance de surface afin d'identifier les différents composants des structures et de quantifier les contraintes. Une fois la frontière clairement identifiée, les contraintes peuvent être appliquées *a posteriori*, en dehors de l'algorithme d'optimisation topologique.

Un exemple particulier d'optimisation de structures est l'optimisation topographique, concernant l'optimisation d'une plaque mince en variant son épaisseur en chaque point. La solution des problèmes d'optimisation topographique approche une méthode de densité où le paramètre ρ représente l'épaisseur de la membrane plutôt que les propriétés du matériau en chaque point. Des problèmes d'optimisation topographique peuvent émerger de l'application de techniques de réduction de modèles à des problèmes plus complexes, comme dans [246, 239].

Une approche différente à l'optimisation de forme consiste en l'optimisation évolutive de structures (*Evolutionary Structural Optimization*, ou ESO) développée par Xie et Steven [256] et étendue à la méthode d'optimisation évolutive bidirectionnelle de structures (*Bidirectional Evolutionary Structural Optimization*, ou BESO) avec la collaboration de Querin en 1998 [203]. La technique BESO considère un maillage fixe englobant l'ensemble du domaine de calcul, et où chaque élément est soit vide, soit rempli par le matériau composant la structure. À chaque étape de l'optimisation, un problème aux éléments finis est résolu dans la région occupée par le matériau, la sensibilité nodale de la fonction objectif par rapport à la forme est calculée, et une partie du matériau est ajoutée ou retirée de la structure en conséquence. Voir [135, 257, 241] pour des applications numériques, et [140] pour des améliorations apportées à la méthode BESO originale, notamment en ce qui concerne la dépendance par rapport au maillage.

Les progrès technologiques concernant les matériaux composites et l'impression 3D ont stimulé l'intérêt pour les structures poreuses, et ont été à l'origine de techniques d'optimisation topologique adaptées à de telles conceptions. À titre d'exemple, nous citons la classe des *méthodes d'homogénéisation*. Introduites entre les années 1970 et 1980 par Tartar et Murat [242, 187], Gibiansky, Lurie et Cherkaev [172, 123, 124], et Kohn and Strang [153], ces méthodes ont montré leur efficacité dans le cadre de l'optimisation de structures depuis la publication de l'article [44] par Bendsøe et Kikuchi en 1988. Au lieu de modéliser directement la structure au niveau microscopique, ce qui peut être extrêmement coûteux du point de vue computationnel, les propriétés mécaniques de la structure sont résumées par un seul tenseur dépendant d'un ou de plusieurs paramètres, et la distribution de ces paramètres dans l'espace est optimisée. Les changements de paramètres dans l'espace représentent des variations de la structure du matériau poreux, comme la direction des fibres d'un matériau anisotrope ou la densité de chaque cellule du lattice. Voir [5, 197, 122] et les références qui y figurent pour plus d'informations sur les techniques d'homogénéisation dans le cadre de l'optimisation topologique. Des courbes paramétrées telles que les *Non-Uniform Rational Basis Splines* (NURBS) peuvent remplacer l'approche de densité pour représenter le champ variable des paramètres dans la structure [185, 48]. Le problème d'homogénéisation inverse, qui consiste à reconstruire une structure poreuse conforme à une distribution donnée des propriétés des matériaux dans un domaine, peut également se traiter à l'aide de techniques d'optimisation de forme [261].

Optimisation de forme par la méthode de Hadamard

Dans la présente thèse, afin de traiter des problèmes d'optimisation de structures, on adopte une approche reposant sur la notion de dérivée par rapport à la forme telle qu'introduite par Jacques Hadamard dans son mémoire [131] et présentée en détail dans [6, 100, 138]. Soit Ω un domaine dans $\mathbb{R}^d$ avec une frontière Lipschitzienne pour $d = 2$ ou 3, $\Phi(\cdot)$ une fonctionnelle à valeurs réelles définie sur un ensemble approprié de domaines admissibles $\mathcal{S}_{\text{adm}}$, et $\boldsymbol{\theta}$ un champ de vecteurs continus et Lipschitziens. La dérivée de forme au sens de Hadamard décrit la sensibilité de $\Phi(\Omega)$ par rapport à de petites perturbations du domaine Ω, et est définie comme la dérivée au sens de Fréchet de la fonction $\boldsymbol{\theta} \mapsto \Phi((\mathbf{I} + \boldsymbol{\theta})\Omega)$, où $\mathbf{I}$ désigne l'opérateur identité. Telle dérivée est utilisée pour calculer un champ de déformation approprié pour optimiser la structure. Plus de détails sur la définition et le calcul des dérivées de forme sont fournis dans la section 1.2.

Dans les premières applications de la méthode de Hadamard à l'optimisation de forme, les structures étaient modifiées en déplaçant les nœuds du maillage dans la direction indiquée par la dérivée de forme [198, 184]. Un problème majeur de cette approche est la dégradation rapide du maillage après seulement quelques étapes de l'algorithme d'optimisation [41]. Parmi les solutions proposées pour ce problème on trouve des opérations de remaillage fréquentes [36], ainsi que l'utilisation de techniques de remaillage adaptées au problème d'optimisation, comme le *Deformable Simplician Complex* [182, 72, 71].

Une autre approche adaptée au contexte de l'optimisation de forme est la *méthode des lignes de niveaux*. Cette technique a été introduite pour la première fois par Osher et Sethian en 1988 [193] pour représenter une interface dans un problème de dynamique des fluides. Cependant, depuis que Haber et Bendsøe ont reconnu son potentiel en 1996 [130], la méthode des lignes de niveaux est devenue l'une des approches les plus populaires pour les applications d'optimisation de forme (voir [247] pour plus de détails et d'exemples). Afin de représenter une structure $\Omega \subset \mathbb{R}^d$, la méthode des lignes de niveaux considère un domaine de calcul $D \subset \mathbb{R}^d$ plus grand que Ω, et associe à la structure une fonction continue $\phi_\Omega : D \mapsto \mathbb{R}$ qui est négative à l'intérieur de la structure, positive à l'extérieur, et égale à zéro sur sa frontière. Le transport de la fonction de niveau par un champ vectoriel donné peut se calculer numériquement à l'aide d'une équation d'advection appropriée sur un maillage couvrant l'ensemble du domaine de calcul D.

Si la fonction ϕ_Ω est définie sur un maillage fixe $\mathcal{T}_D$, certains des éléments du maillage sont traversés par l'isovaleur 0. Une approche simple consiste à modifier les propriétés des matériaux dans chaque élément du maillage en fonction de la valeur locale de la fonction-lignes de niveaux [19]. Cette technique est simple à mettre en œuvre et donne des résultats fiables, comme le montrent les simulations numériques de la section 4.4. Cependant, elle souffre de deux inconvénients importants. Tout d'abord, il est souvent nécessaire de remplacer l'espace extérieur à la structure par un *matériau ersatz* léger, imitant l'espace vide mais qui garantit que les problèmes différentiels définis sur D sont bien posés. L'erreur introduite par le matériau ersatz est étudiée en [84]. Deuxièmement, si la ligne de niveau 0 de ϕ_Ω traverse des éléments du maillage, l'interface entre l'intérieur et l'extérieur de la structure n'est pas clairement définie. Certaines techniques abordant ce problème sont l'approche de l'interface immergée [254, 230], l'approche *XFEM* [107, 249] et *CutFEM* [59, 250], ainsi que l'algorithme *LEVITY* [77]. Une discussion plus détaillée de ces méthodes est présentée dans la section 1.3.2. Pour la plupart des simulations de cette thèse, nous adoptons l'approche proposée par Dapogny et Feppon en [91], qui joint la méthode des lignes de niveaux avec des techniques d'adaptation de maillage de sorte que, à chaque étape de l'optimisation, la structure Ω est explicitement représentée par un sous-maillage de $\mathcal{T}_D$.

Une définition précise des frontières de la structure par la fonction ϕ_Ω permet de traiter

plus facilement les contraintes de fabricabilité et d'accessibilité [167]. Différentes méthodes d'application des contraintes d'épaisseur minimale ou maximale ont été proposées dans la littérature, reposant soit sur l'utilisation de fonctionnelles d'énergie [65], soit sur les propriétés de la fonction *distance signée* [181, 115]. Les contraintes d'accessibilité, en particulier dans le contexte de la fabrication additive, ont également été étudiées dans le cadre d'une approche par lignes de niveaux [9, 7, 8].

Un inconvénient important de la dérivée de forme au sens de Hadamard est qu'elle ne prend pas en compte les changements de topologie de la forme au cours de l'optimisation. Si la structure à optimiser est représentée par un maillage déformable, la modification de la topologie est impossible, car elle nécessiterait des changements dans la matrice de connectivité. Cependant, dans le cadre de la méthode des lignes de niveaux, la nature discrète de l'étape d'advection peut entraîner des modifications topologiques de la forme, comme indiqué dans [112, Section 1.4.2]. Typiquement, une paroi fine peut se briser ou deux parties peuvent fusionner. Comme la dérivée de forme au sens de Hadamard n'est définie que pour de petites déformations qui n'entraînent pas de changements topologiques, on peut observer des discontinuités dans les fonctions objectif ou dans les contraintes à chaque fois qu'un changement de topologie se produit.

Un autre inconvénient de la méthode des lignes de niveaux est le fait que, contrairement à la méthode de la densité, elle ne permet pas l'apparition de nouvelles cavités à l'intérieur du matériau. Ce problème a été résolu par la définition de la *dérivée topologique*, qui mesure la sensibilité d'une fonction de forme par rapport à la création de trous dans la structure [233, 191]. Les dérivés topologiques et de forme ont été utilisées avec la méthode des lignes de niveaux afin de diminuer la dépendance du résultat par rapport à la forme initiale [19, 21, 91].

Généralités sur les incertitudes en optimisation de forme

Les avancements dans les techniques d'optimisation de forme et topologique ont permis de développer des structures très performantes dans plusieurs domaines industriels. Cependant, la performance de la structure dans un cas de référence n'est pas le seul aspect à prendre en compte, car son comportement peut être perturbé par un certain nombre de facteurs différents [10]. Parmi les raisons, on compte la déviation de la structure par rapport à sa conception originale en raison d'imperfections dans le processus de fabrication ou dans la qualité du matériau. Deuxièmement, l'usure de la structure dans le temps peut détériorer son matériau et dégrader ses performances. Enfin, les charges appliquées à la structure et les conditions d'utilisation peuvent fluctuer dans le temps ou être connues avec un certain degré d'incertitude.

La présente thèse aborde le problème récurrent dans le secteur aérospatial de la conception de structures aussi légères que possible, mais conformes aux exigences rigoureuses de robustesse et de fiabilité de l'industrie. En particulier, un moteur à turbine à gaz est caractérisé par de nombreux composants interconnectés entre eux, et soumis à des sollicitations mécaniques intenses et variables. Afin de garantir la tenue du moteur dans une vaste gamme de conditions envisageables au cours du vol, la prise en compte des paramètres incertains lors de la conception des composants de la turbine est fondamentale. C'est dans ce contexte que Safran Helicopter Engines a manifesté son intérêt pour l'optimisation de strucures sous incertitudes, avec l'objectif de développer des techniques d'optimisation de forme adaptées aux circonstances particulières d'un moteur, et de les appliquer à certains composants critiques.

Le problème de la prise en compte des perturbations dans l'optimisation de structures est apparu dans de nombreux domaines de l'ingénierie. A titre d'exemple, on peut citer [196, 220, 156, 139] pour des applications à l'ingénierie aérospatiale, [161] pour une étude dans le contexte

biomédical, et [201, 163] pour une application à la micro-électronique et à la photonique. Le sujet de la quantification des incertitudes et de l'optimisation de la forme d'un domaine soumis à des perturbations géométriques aléatoires dépend strictement de la méthode choisie pour le représenter. Les premières approches reposent sur la paramétrisation de la structure par un nombre fini de variables. Dans [196], une approche de Krigeage est utilisée pour évaluer la réponse d'une aile à de petites perturbations de sa géométrie. Dans [63, 222, 81, 11], le domaine incertain est représenté par la perturbation d'une forme de référence par un champ aléatoire Lipschitzien. Une approche différente basée sur la méthode de la densité est proposée dans [162], où l'incertitude de la forme est modélisée par la présence de paramètres aléatoires dans la régularisation de la fonction de densité. Les techniques d'optimisation topologique en présence d'incertitudes sur la structure se sont révélées utiles dans le contexte des problèmes d'identification de formes et d'obstacles [82, 83].

La présence d'incertitudes dans les propriétés du matériau ou dans les charges appliquées a été étudiée par plusieurs auteurs, et de nombreuses perspectives différentes ont été proposées. Une possibilité consiste à supposer que la distribution de probabilité de la perturbation est connue ou qu'elle peut être déduite de mesures expérimentales. Dans cette catégorie d'approches, on trouve les techniques d' *optimisation topologique robuste* (Robust Topology Optimization - RTO) et d'*optimisation topologique fiable* (Reliability-Based Topology Optimization - RBTO) [121]. L'approche RTO évalue l'espérance et la variance des fonctions d'intérêt. Les méthodes RBTO consistent à estimer la probabilité qu'une fonction donnée dépasse un seuil donné, noté *probabilité de défaillance* [10, 11], et nécessitent une connaissance précise de la distribution des incertitudes. Plus de détails et de références sur les problèmes RTO et RBTO sont présentés dans les chapitres 4 et 5 de cette thèse. Les hypothèses sur la distribution des incertitudes peuvent être inexactes, et une approche possible à cette problématique est présentée dans [93].

Une autre classe de problèmes considère que la perturbation appartient à un ensemble borné, et se préoccupe d'optimiser la structure par rapport au scénario le plus défavorable. La fonction dont la valeur doit être contrôlée peut apparaître soit dans l'objectif, soit dans la contrainte du problème d'optimisation, et les approches aux deux situations sont différentes. Le cas où la fonction incertaine apparaît comme une contrainte est le principal intérêt du chapitre 3. Comme indiqué dans [43], l'optimisation du pire cas est préférable aux problèmes d'optimisation robuste ou fiable lorsque les données sont imprécises ou si leur distribution de probabilité est inconnue, si le respect strict de la contrainte en toutes circonstances est d'importance primordiale, ou si la solution optimale est particulièrement difficile à fabriquer.

Les différentes approches à l'inclusion des incertitudes dans les conditions externes ou dans les propriétés du matériau présentent des inconvénients et des avantages. D'une part, l'optimisation sous contraintes du pire cas ne nécessite aucune hypothèse sur la distribution des incertitudes. D'autre part, elle peut aboutir à une forme trop pessimiste, avec des performances médiocres dans la grande majorité des situations afin de couvrir un ensemble invraisemblable de circonstances. Pour la plupart des applications mécaniques, la forme de la structure et les propriétés des matériaux sont connues avec précision, mais les charges appliquées sont sujettes à une plus grande variabilité. Par conséquent, la présente thèse se concentrera sur le cas de conditions externes incertaines, sans aucune hypothèse préalable sur la taille des perturbations.

Propagation des incertitudes pour l'élasticité linéaire

L'objectif principal de cette thèse est d'analyser les problèmes d'optimisation de forme en élasticité linéaire, lorsque les chargements mécaniques présentent des incertitudes dans la direction

et l'intensité. Plus précisément, on considère une structure mécanique composée d'un matériau élastique homogène et isotrope. Soit $\mathcal{S}_{\text{adm}}$ la classe de toutes les formes admissibles que la structure peut prendre, qu'on suppose être des domaines ouverts Lipschitziens dans $\mathbb{R}^d$. Pour toute $\Omega \in \mathcal{S}_{\text{adm}}$, on suppose que sa frontière $\partial\Omega$ se divise en trois parties avec une mesure strictement positive : Γ_{D}, où la structure est fixée, Γ_{N} où les charges mécaniques sont appliquées, et Γ_0 représentant la surface libre. Le déplacement élastique $\mathbf{u}_\Omega$ induit par les chargements mécaniques peut se calculer à l'aide des équations de l'élasticité linéaire. Dans ce cadre, soit $\mathbf{f}$ la densité de force agissant sur l'ensemble du volume, $\mathbf{g}$ le chargement surfacique appliqué à la portion Γ_{N} de la frontière de la structure, $\mathbf{u}_\Omega$ le champ de déplacement, et $\boldsymbol{\sigma}(\mathbf{u}_\Omega)$ le tenseur des contraintes. Des conditions de Dirichlet homogènes sont appliquées sur Γ_{D}, tandis que Γ_0 et Γ_{N} présentent des conditions de Neumann homogènes et non homogènes respectivement.

L'objectif des problèmes considérés dans cette thèse est de trouver la structure admissible avec un volume minimal, qui satisfait un certain critère sous la forme $H(\mathbf{u}_\Omega, \Omega) \leq \tau$, où τ est un seuil approprié. Le problème d'optimisation peut être formulé ainsi

$$
\left|\begin{array}{l}
\text{Trouver la forme admissible } \Omega \in \mathcal{S}_{\text{adm}} \\
\text{qui minimise le volume } \mathrm{Vol}(\Omega) \\
\text{sous la contrainte } H(\mathbf{u}_\Omega, \Omega) \leq \tau, \\
\text{où le déplacement } \mathbf{u}_\Omega \\
\text{satisfait les équations de l'élasticité linéaire} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\right) & = & \mathbf{f} & \text{dans } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{g} & \text{sur } \Gamma_{\text{N}}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{0} & \text{sur } \Gamma_0, \\
\mathbf{u}_\Omega & = & \mathbf{0} & \text{sur } \Gamma_{\text{D}}.
\end{array}\right.
\end{array}\right.
$$

Afin d'introduire les incertitudes, on considère l'espace probabilisé $(\mathcal{O}, \mathcal{A}, \mathbb{P})$, où l'espace des évènements est désigné par $\mathcal{O}$, $\mathcal{A} \subset 2^{\mathcal{O}}$ est une tribu sur $\mathcal{O}$, et $\mathbb{P}$ une mesure de probabilité. On suppose que le chargement mécanique $\mathbf{g}$ est incertain, de sorte qu'il puisse se modéliser comme une variable aléatoire. Ainsi, les incertitudes se propagent au déplacement $\mathbf{u}_\Omega$ et à la contrainte $H(\mathbf{u}_\Omega(\omega), \Omega)$. Afin de résoudre le problème d'optimisation et d'obtenir une solution déterministe, il est nécessaire de remplacer la contrainte $H(\mathbf{u}_\Omega, \Omega)$ par une quantité déterministe $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \in \mathbb{R}$. En tenant compte des incertitudes, le problème d'optimisation de forme ci-dessus peut se formuler comme suit

$$
\left|\begin{array}{l}
\text{Trouver la forme admissible } \Omega \in \mathcal{S}_{\text{adm}} \\
\text{qui minimise le volume } \mathrm{Vol}(\Omega) \\
\text{sous la contrainte } \mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \leq \tau_{\mathcal{F}}, \\
\text{où le déplacement } \mathbf{u}_\Omega(\omega) \\
\text{satisfait les équations de l'élasticité linéaire} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\right) & = & \mathbf{f} & \text{dans } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{g}(\omega) & \text{sur } \Gamma_{\text{N}}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{0} & \text{sur } \Gamma_0, \\
\mathbf{u}_\Omega(\omega) & = & \mathbf{0} & \text{sur } \Gamma_{\text{D}}, \\
\end{array}\right. \\
\text{pour presque tout évènement } \omega \in \mathcal{O}.
\end{array}\right.
$$

Les différentes approches à l'optimisation de forme sous incertitudes peuvent être distinguées selon le choix de l'opérateur $\mathcal{F}[\cdot]$. Pour l'approche du pire cas, la contrainte déterministe peut s'écrire sous la forme $\mathcal{F}[H(\mathbf{u}_\Omega, \Omega)] = \operatorname{ess\,sup}_{\omega\in\mathcal{O}} H(\mathbf{u}_\Omega(\omega), \Omega) \leq \tau$. Dans le cas des problèmes d'optimisation robuste, l'opérateur $\mathcal{F}[\cdot]$ est une combinaison des opérateurs d'espérance et de variance. Enfin, pour les problèmes d'optimisation fiable, $\mathcal{F}[H(\mathbf{u}_\Omega, \Omega)]$ est la probabilité que $H(\mathbf{u}_\Omega, \Omega)$ dépasse une certaine tolérance, et l'inégalité $\mathcal{F}[H(\mathbf{u}_\Omega, \Omega)] \leq \tau_{\mathcal{F}}$ peut s'interpréter comme une contrainte sur la probabilité de défaillance.

Résumé de la thèse par chapitre

Cette thèse se divise en deux parties principales. Dans la partie I on introduit les notions principales d'optimisation de forme pour des structures élastiques, et on étudie des problèmes d'optimisation dans un cadre déterministe. La partie II présente différentes approches pour prendre en compte les incertitudes sur les chargements mécaniques appliqués à la structure. La partie I est divisée en deux chapitres, tandis que la partie II se compose de trois chapitres, un pour chaque approche.

Chapitre 1: methode de Hadamard pour l'optimisation de forme

Dans ce chapitre introductif, on présente les notions principales relatives à l'optimisation de forme utilisées dans cette thèse.

Dans la section 1.1, on commence par rappeler le cadre général des problèmes d'optimisation sous contraintes. Ensuite, on se concentre sur des problèmes d'optimisation contraints par des Équations aux Dérivées Partielles (EDP), et on présente la notion d'algorithmes d'optimisation basés sur le gradient. Une attention particulière est accordée au calcul et à l'étude de l'état adjoint. La section se termine par quelques rappels et notations sur les espaces de fonctions.

La section 1.2 décrit la méthode de dérivation de fonctionnelles de forme telle qu'introduite par Hadamard en 1908 [131], en suivant les notations de [6, 100, 138]. Après avoir rappelé les notions de dérivées de Fréchet et de Gâteaux, on introduit la notion de dérivée de forme, on énonce le théorème de structure de Hadamard (théorème 1.7), et on montre quelques résultats sur les dérivées des fonctionnelles de volume et de surface. Ensuite, on présente deux méthodes pour calculer la dérivée de forme d'une fonctionnelle de volume qui dépend de la solution d'une EDP. Dans la section 1.2.2, la dérivée de forme est obtenue par le calcul de la dérivée lagrangienne de la solution de l'EDP, et par la déduction de l'équation adjointe. Dans la section 1.2.3, on obtient la même expression par la méthode de dérivation rapide de Céa [61]. Plus de détails sur le calcul de certaines expressions se trouvent dans l'annexe A.

Enfin, dans la section 1.3, on discute des méthodes numériques pour les simulations effectuées dans cette thèse, ainsi que d'alternatives possibles. Tout d'abord, on présente la méthode des lignes de niveaux et la fonction de distance signée pour la représentation des domaines. Ensuite, on introduit la notion de maillage adaptatif et la méthode du matériau ersatz pour résoudre des problèmes aux limites pour les domaines encodés par les fonctions-lignes de niveaux. Dans la section 1.3.4, on présente l'algorithme d'optimisation *null space* tel qu'il a été décrit par Feppon, Allaire et Dapogny en [114]. En conclusion, on résume les différentes étapes de l'algorithme d'optimisation de forme utilisé dans cette thèse.

Chapitre 2: Un problème élasto-thermique déterministe

Le deuxième chapitre de la partie I est consacré à l'optimisation de structures mécaniques élastiques dans le cas déterministe. Dans la section 2.1.1, on commence par introduire les équations de l'élasticité linéaire, les tenseurs des déformations et des contraintes, les propriétés de symétrie du tenseur d'élasticité de quatrième ordre, et on exprime les équations de l'élasticité linéaire pour des matériaux homogènes isotropes. Ensuite, on présente deux fonctions de forme particulièrement intéressantes en mécanique des structures : la compliance (le travail des forces externes) et la contrainte de von Mises.

Dans la section 2.2 on étudie l'influence des effets thermiques pour l'optimisation d'une structure élastique. On commence par présenter les équations constitutives d'un matériau élastique soumis à une dilatation thermique. Ce problème est particulièrement intéressant dans le contexte de la fabrication additive (voir Allaire et Jakabčin [15]), ainsi que pour les applications multiphysiques [112].

On suppose que la réponse mécanique est beaucoup plus rapide que les changements dans le champ de température. Ainsi, on utilise l'équation de diffusion dépendante du temps pour calculer la distribution thermique à chaque instant, tandis que le déplacement est modélisé par les équations quasi-statiques de l'élasticité linéaire. Dans le reste de la section, on formule un problème d'optimisation de forme pour une structure élastique soumise à un champ thermique variable dans le temps, avec l'objectif de minimiser son volume sous une contrainte sur la compliance mécanique. La contrainte est composée de deux termes : le premier consiste en la moyenne temporelle de la compliance pour la durée de la simulation. La seconde composante, qui n'est pas examinée aussi souvent que la première pour des problèmes d'optimisation de forme dépendant du temps, ne prend en compte que l'état final de la structure. La dérivée de forme d'une telle fonctionnelle est calculée dans la section 2.2.3.

Dans la section 2.3, on présente les résultats de quelques simulations numériques. On commence par la minimisation du volume d'un cantilever 3D avec des contraintes sur la compliance et sur la contrainte de von Mises, et on termine par quelques résultats sur le problème thermo-élastique. Les résultats de la section 2.2 ont été obtenu en collaboration avec Viacheslav Karnaev, doctorant à l'université de Bâle. L'étude du problème thermo-élastique fait partie d'un travail en cours sur l'optimisation de forme dans un champ thermique incertain et variable dans le temps, et fera l'objet d'une publication.

Chapitre 3: Optimisation du pire cas

Ce premier chapitre de la partie II présente deux méthodes différentes pour traiter le cas où l'une des contraintes du problème d'optimisation concerne la valeur maximale qu'une fonction donnée peut prendre. Dans les deux cas, on suppose que les charges incertaines peuvent se paramétrer par un nombre fini de variables appartenant à un sous-ensemble borné et convexe $\mathcal{G}$ d'un espace de Banach approprié.

La première méthode, détaillée dans la section 3.2, peut s'appliquer lorsque la fonction $\mathcal{G} \ni \mathbf{g} \mapsto H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$ qui associe les paramètres décrivant les charges incertaines à la valeur de la contrainte est une fonction convexe. La stratégie consiste à approcher l'ensemble $\mathcal{G}$ par un polyèdre convexe $\mathcal{G}_N$ avec un nombre fini de sommets N. Grâce à la convexité de la fonction $\mathbf{g} \mapsto H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$, la valeur maximale de la contrainte est atteinte dans l'un des sommets de $\mathcal{G}_N$. Ainsi, le problème d'optimisation de forme peut se reformuler comme un problème équivalent avec à N contraintes, chacune relative à un différent cas de chargement. Un résultat de convergence pour la solution du problème approximé vers la solution de l'original est fourni dans la section

3.2.3. Cette méthode justifie l'approche des ingénieurs, qui consiste en concevoir des structures capables de résister à un nombre limité de cas de chargements représentatifs.

Dans la section 3.3, on propose une méthode alternative. On introduit la notion de sous-différentiel au sens de Clarke, on l'adapte aux fonctions définies sur des formes, et on fournit une condition suffisante pour son existence sous des conditions de régularité de la fonctionnelle et de l'ensemble $\mathcal{G}$. Le cœur de la méthode proposée consiste en l'identification d'un élément du sous-différentiel à chaque étape de l'algorithme d'optimisation, et le calcul du gradient de forme est fait dans la direction correspondante. On décrit dans les détails la mise en œuvre numérique de l'approche sous-différentielle dans la section 3.3.3.

En conclusion, la section 3.4 présente les résultats de quelques simulations numériques, où on vise à minimiser le volume de deux structures différentes sous des contraintes sur la compliance ou sur la norme L^6 de la contrainte de von Mises. L'approximation polyédrique et la méthode de sous-différentiel ont été utilisées pour enforcer la contrainte. Les résultats numériques valident les deux approches, et illustrent la convergence de la solution de la méthode d'approximation polyédrique vers la solution exacte lorsque $\mathcal{G}_N$ s'approche de $\mathcal{G}$. De plus, les simulations montrent que l'approche basée sur le sous-différentiel est généralement plus rapide que l'approximation polyédrique de l'ensemble $\mathcal{G}$, mais qu'elle peut être perturbée par des oscillations si la structure à optimiser et l'ensemble des chargements admissibles sont charactérisés par de multiples symétries. Un article à ce sujet est en cours de rédaction.

Chapitre 4: Optimisation de forme robuste d'une fonctionnelle polynomiale

Plusieurs approches à l'optimisation de forme concernant l'espérance ou la variance d'une fonctionnelle de forme ont été proposées au cours de ces dernières années [106, 11, 211, 80]. En particulier, Dambrine, Dapogny et Harbrecht ont prouvé en [80] que la valeur de l'espérance d'une fonctionnelle quadratique ne dépend que des deux premiers moments stochastiques des variables aléatoires décrivant les charges incertaines. Par conséquent, le calcul de l'espérance et de sa dérivée par rapport à la forme ne demande pas d'utiliser des méthodes d'échantillonnage, qui peuvent être coûteux du point de vue computationnel.

L'objectif principal de ce chapitre est l'extension de l'approche de [80] aux fonctionnelles m-multilinéaires. Le cadre du problème est décrit dans la section 4.2, tandis que les principaux résultats théoriques sont présentés dans la section 4.3. En particulier, on étend la définition de l'opérateur de corrélation de manière similaire à [226] et, dans le contexte de l'élasticité linéaire, on utilise une approche tensorielle pour montrer que l'espérance d'une fonctionnelle m-multilinéaire ne dépend que des m premiers moments stochastiques des chargements aléatoires.

Dans les deux sections suivantes on montre différentes applications de la méthode, avec quelques résultats numériques. Dans la section 4.4, on aborde le problème de la contrainte de la norme L^∞ de la contrainte de von Mises. Comme la norme L^∞ ne peut pas être différentiée, on peut l'approcher par la norme L^m pour m suffisamment grand, qui peut s'écrire en termes d'une fonctionnelle multilinéaire de m. Comme simulation numérique, on considère la minimisation de la masse d'une structure cylindrique 3D soumise à des forces de cisaillement aléatoires sur sa face supérieure, avec une borne imposée sur la norme L^6 de la contrainte de von Mises. Dans la section 4.5, on exprime la variance d'une fonctionnelle quadratique en termes d'une fonctionnelle multilinéaire de degré 4. Ensuite, on adapte un exemple 2D de [80] pour montrer l'effet d'une contrainte sur l'espérance et la variance de la compliance.

Les résultats principaux de ce chapitre ont été soumis sous la forme d'un article intitulé *Shape Optimization of Polynomial Functionals under Uncertainties on the Right-Hand Side of the State Equation*, coécrit avec Fabien Caubet, Marc Dambrine et Jérôme Maynadier.

Chapitre 5: Optimisation de la probabilité de dépassement d'un seuil

Ce dernier chapitre est consacré aux problèmes d'optimisation fiable pour des fonctions quadratiques. L'intérêt des techniques RBTO se trouve dans la possibilité de concevoir des structures conformes aux contraintes relatives à leur probabilité de défaillance [34, 180, 151, 76, 11].

Soit $\mathbf{g}$ un chargement mécanique aléatoire décrit par un vecteur aléatoire, $\mathbf{u}_{\Omega,\mathbf{g}}$ le déplacement élastique associé, et $(\mathbf{u},\mathbf{v}) \mapsto Q(\mathbf{u},\mathbf{v}) \in \mathbb{R}$ une fonctionnelle quadratique continue du déplacement. La fonctionnelle dépend de la forme de la structure (on peut prendre comme exemples la compliance mécanique ou la norme L^2 de la contrainte de von Mises). Soit τ un seuil, et $\Phi(\cdot)$ la fonctionnelle de forme telle que

$$\Phi(\Omega) = \mathbb{P}\left[Q(\mathbf{u}_{\Omega,\mathbf{g}}, \mathbf{u}_{\Omega,\mathbf{g}}) > \tau\right],$$

où $\mathbf{u}_{\Omega,\mathbf{g}}$ est la solution (incertaine) de l'équation de l'élasticité linéaire. Le calcul de la valeur et de la dérivée de forme de $\Phi(\Omega)$ est détaillée dans les sections 5.2 et 5.3, et repose sur l'intégration de densité de probabilité sur le volume et sur la surface de l'ellipsoïde défini par l'inégalité $Q(\mathbf{u}_{\Omega,\mathbf{g}}, \mathbf{u}_{\Omega,\mathbf{g}}) \leq \tau$.

Dans la section 5.4, on présente une technique différente pour évaluer la probabilité de défaillance $\Phi(\Omega)$ et sa dérivée lorsque la charge externe suit une distribution gaussienne. Cette deuxième méthode repose sur la décomposition en série de la fonction de répartition d'une distribution chi-carré non centrale généralisée, comme montré par Ruben en 1962 [214].

Dans la section 5.5, on présente les détails des algorithmes utilisés pour effectuer les simulations numériques de ce chapitre. Afin de tester la méthode proposée dans la section 5.3, on a considéré des problèmes où les incertitudes sont modélisées par des gaussiennes centrées, et on a appliqué les formules de quadrature de l'annexe C afin d'éviter le calcul numérique d'intégrales en grandes dimensions. Les deux méthodes ont été testées pour l'optimisation de deux structures élastiques, avec l'objectif de minimiser leur volume sous des contraintes sur la probabilité pour la compliance de dépasser un seuil. Les résultats sont présentés dans la section 5.6. Le calcul des coefficients dans l'expression de la dérivée de forme est presque instantané, donc la résolution de ces problèmes de type RBTO prend à peu près le même temps que la résolution d'un problème d'optimisation de forme déterministe.

L'essentiel de ce chapitre a fait l'objet d'un article soumis sous le titre *Shape optimization under constraints on the probability of a quadratic functional to exceed a given threshold*, écrit conjointement avec Marc Dambrine, Helmut Harbrecht et Jérôme Maynadier.

General introduction

Generalities on structural optimization

The technological advancements of the recent years have resulted in an increase in computational power and have allowed for the development of advanced techniques for the design of complex structures. Shape and topology optimization techniques have been applied in several branches of engineering, ranging from civil engineering [237, 257], to the aerospace industry [186, 183, 199], automotive components [227, 204, 175], heat exchangers [264, 259, 238, 243, 4], microelectronics and photonics [146, 109, 163], biomechanics [161, 144, 127] and beyond. The advancements in additive manufacturing and Computer Numerical Control (CNC) milling machines have made possible the realization of elaborate structures and have driven further the development of shape and topology optimization methods [260, 243, 15]. Consult [207] for a comprehensive review of the applications of shape optimization in different fields of engineering.

The numerous approaches to structure optimization are usually divided in three main categories [238, 177]:

- **parametric optimization**, where the structure of interest is represented by a small number of variables;
- **shape optimization**, for which the parameters of the optimization are the boundary of the structure itself;
- **topology optimization**, where the parameter of optimization is the distribution of material within a reference domain, with few preconceptions on the topology of the optimal structure.

Density and homogenization methods

Density methods are a wide class of topology optimization techniques, tracing their origin to the seminal paper on the Solid Isotropic Material with Penalization (SIMP) technique devised by Bendsøe and Sigmund [47]. Such method considers a computational domain D on which is defined a density field $\rho : D \to [0, 1]$. The material properties in each point of the domain are supposed to be an interpolation of the properties of two different materials (like a solid matrix and a fluid, or two elastic materials) depending on ρ. In order to ensure the well-posedness of the optimization problem in the case of the optimization of elastic structures, a suitably light ersatz material replaces the empty space surrounding the domain. The optimization parameter is the density parameter ρ itself so that, after the convergence of the optimization algorithm, the

optimal distribution of the materials in the computational domain is identified by the regions where ρ is close to 0 or to 1. The SIMP method has been widely studied and adapted to the optimization of elastic structures [45], and to structures coupled to fluid problems [53]. Moreover, it has long been introduced in commercial Finite Element software as ANSYS, Optistruct and ABAQUS [262] and Computational Fluid Dynamics software [195].

The convergence towards a solution having a sharp distinction between the two phases depends on the nature of the problem and on the parameters of the optimization algorithm [178]. Therefore, wide grey regions might exists even at the end of the optimization. The enforcement of constraints on the thickness of the structure is made particularly challenging by the absence of a sharp boundary. The same can be said for constraints on the topology of the domain or manufacturability. As stated in the review paper [167], such objectives require the use of suitable surface recognition algorithms in order to identify the different components of the structures and quantify the constraints. Once the boundary has been clearly identified, the constraints can be enforced *a posteriori*, outside the topology optimization algorithm.

A particular example of structure optimization is topography optimization, where the shape to be optimized is a thin plate, and the optimization parameter is its thickness in each point. The solution of topography optimization problems is akin to a density method where the parameter ρ represents the thickness of the membrane instead of the material properties in each point. Topography optimization problems can emerge from the application of model reduction techniques to more complex problems, as in [246, 239].

A different approach to shape optimization consists in the Evolutionary Structural Optimization (ESO) developed by Xie and Steven [256] and extended to the Bi-directional Evolutionary Structural Optimization (BESO) method in collaboration with Querin in 1998 [203]. The BESO technique considers a fixed mesh encompassing the entire computational domain, and where each element is either empty or filled by the material composing the structure. At each step of the optimization, a finite element problem is solved in the region occupied by the material, the nodal sensitivity of the objective with respect to the shape is computed, and some material is added or removed from the structure accordingly. Refer to [135, 257, 241] for some numerical applications, and to [140] for some improvements to the original BESO method addressing the issue of mesh dependency.

The technological advancements surrounding composite materials and 3D printing have fostered the interest in lattice structures, and given rise to topology optimization techniques adapted to such designs. As an example, we cite the class of *homogenization methods*. Introduced between the 1970's and the 1980's by Tartar and Murat [242, 187], Gibiansky, Lurie and Cherkaev [172, 123, 124], and Kohn and Strang [153], this class of methods have proven its efficacy in shape optimization since the publication of the seminal paper by Bendsøe and Kikuchi in 1988 [44]. Instead of directly modeling the the lattice, which can be extremely costly from the computational point of view, the material properties of the structure are summarized by a single tensor depending on one or more parameters, and the distribution of said parameters in space is optimized. The changes of the parameters in space represent variations of the lattice structure, like the direction of the fibers of an anisotropic material, or the density of each lattice cell. Refer to [5, 197, 122] and the references within for further information on the homogenization technique in topology optimization. Parametrized curves as the Non-Uniform Rational Basis Splines (NURBS) can replace the density approach to represent the varying field of parameters in the structure [185, 48]. The inverse homogenization problem, consisting in the reconstruction of a lattice structure compliant with a given distribution of the material properties in a domain, can also be addressed using shape optimization techniques [261].

Shape optimization by Hadamard's method

In the present thesis we adopt a shape optimization approach relying on the notion of shape derivative, as introduced by Jacques Hadamard in his memoir [131] and presented in details in [6, 100, 138]. Let Ω be a domain in $\mathbb{R}^d$ with Lipschitz continuous boundary for $d = 2$ or 3, $\Phi(\cdot)$ a real-valued functional defined on a suitable set of admissible domains $\mathcal{S}_{\text{adm}}$, and $\boldsymbol{\theta}$ a Lipschitz continuous vector field. The shape derivative according to Hadamard describes the sensitivity of $\Phi(\Omega)$ with respect to small perturbations of the domain Ω, and is defined as the Fréchet derivative of the mapping $\boldsymbol{\theta} \mapsto \Phi((\mathbf{I}+\boldsymbol{\theta})\Omega)$, where $\mathbf{I}$ denotes the identity operator. Such derivative is then used to compute a suitable deformation field. More details on the definition and computation of shape derivatives are provided in section 1.2.

In the earliest applications of Hadamard's method in shape optimization, the structures were modified by displacing the nodes of the mesh in the directions indicated by the shape derivative [198, 184]. A major issue of this approach is the rapid degradation of the mesh after few steps of the optimization algorithm [41]. Solutions to this issue were frequent remeshing operations [36] or the use of remeshing techniques adapted to the shape optimization problem as the *Deformable Simplician Complex* [182, 72, 71].

Another approach particularly adapted to the context of shape optimization is the *level-set method*. This technique has been firstly introduced by Osher and Sethian in 1988 [193] to represent an interface in a fluid dynamics problem. However, since Haber and Bendsøe recognized its potential in 1996 [130], the level-set method has become one of the most popular approaches in shape optimization applications (see the review paper [247] for further details and examples). In order to represent a structure $\Omega \subset \mathbb{R}^d$, the level-set method considers a computational domain $D \subset \mathbb{R}^d$ larger than Ω, and associates to the structure a continuous *level-set function* $\phi_\Omega : D \mapsto \mathbb{R}$ that is negative in the interior of the structure, positive outside Ω, and equal to zero on its boundary. The transport of the level-set function by a given vector field can be computed numerically by solving a suitable advection equation [56] on a mesh covering the entire computational domain D.

If the level-set function is defined on a fixed mesh $\mathcal{T}_D$, some of the elements are bound to be crossed by the 0-isoline. One of the earliest approaches consists in altering the material properties in each element of the mesh according to the local value of the level-set function [19]. This approach is simple to implement and provides reliable results, as shown in the numerical simulations of section 4.4. However, it suffers from two important drawbacks. First, it is often necessary to replace the space outside the structure by a light *ersatz material*, mimicking the empty space but ensuring the well-posedness of the differential problems defined on D. The error introduced by the ersatz material is discussed in [84]. Second, if the 0-isoline of ϕ_Ω cuts through some elements of the mesh, the interface between the interior and the exterior of the structure is not sharply defined. Some techniques addressing this issue are the *immersed interface* approach [254, 230], the *XFEM* [107, 249] and *CutFEM* [59, 250] methods, and the *LEVITY* algorithm [77]. A more detailed discussion of these methods is presented in section 1.3.2. For most simulations in this thesis we adopt the approach proposed by Dapogny and Feppon in [91], combining the level-set method with mesh adaptation techniques so that, at each step of the optimization, the structure Ω is explicitly represented by a submesh of $\mathcal{T}_D$.

The clear definition of the structure boundary by the level-set function allows for an easier definition of manufacturability and accessibility constraints [167]. Different methods for the enforcement of minimal or maximal thickness constraints have been proposed in the literature, either relying on the use of suitably-defined energy functionals [65], or on the properties of the *signed-distance* function [181, 115]. Accessibility constraints, particularly in the context of

additive manufacturing, have also been studied under a level-set approach [9, 7, 8].

A notable disadvantage of Hadamard's derivative is the fact that it does not take into account changes in the topology of the shape during the optimization. If the structure to be optimized is represented by a deformable mesh, the modification of the topology is impossible because it would require changes in the connectivity matrix. However, under a level-set approach, the discrete nature of the advection step can result in topological changes of the shape, as pointed out in [112, Section 1.4.2]. Typically, a thin wall might break, or two parts might fuse together. Since the shape derivative according to Hadamard is defined only for small deformations, which do not entail topological changes, discontinuities in the trends of the objective or constraint functionals can be observed whenever a change in topology takes place.

A related drawback of the level-set method is the fact that, contrarily to the density method, it does not allow for the appearance of new holes inside the material. This issue has been addressed by the definition of the *topological derivative*, measuring the sensitivity of a shape functional with respect to the creation of new holes in the structure [233, 191]. Topological derivatives have been used in tandem with Hadamard's shape derivatives and the level-set method in order to relax the dependence from the initial shape [19, 21, 91].

Generalities on uncertainties in shape optimization

Shape and topology optimization techniques have allowed the development of highly performing structures in several industrial domains. However, the performance of the structure in a benchmark case is not the only aspect to be taken into account, since its behavior can be perturbed by a number of different factors [10]. One reason can be a deviation of the structure from its original design because of imperfections in the manufacturing process or in the quality of the material. Secondly, the wear of the structure in time can deteriorate its material and degrade its performance. Finally, the loads applied to the structure and the external conditions can fluctuate in time or be known up to some degree of confidence.

The present thesis addresses the recurrent problem in the aerospace sector of the design of structures that are as light as possible, but comply with the high standards of robustness and reliability of the industry. Specifically, a gas turbine engine is characterized by numerous components interconnected with one another and subject to intense and varying mechanical solicitations. in order to ensure the robustness of the engine in a wide array of flight conditions, it is of primary importance to take the uncertain parameters into account while designing the components of the turbine. It is within this context that Safran Helicopter Engines manifested its interest into structural optimization under uncertainties, with the objective to develop shape optimization techniques adapted for the peculiar circumstances of an engine, and to apply them to some critical components.

The problem of factoring perturbations in structural optimization has appeared in numerous fields of engineering. As examples we mention [196, 220, 156, 139] for application to aerospace engineering, [161] for a study in the biomedical context, and [201, 163] for application to microelectronics and photonics. The issue of uncertainty quantification and shape optimization of a domain subject to random geometric perturbations is strictly dependent from the method chosen to represent it. Early approaches rely on the parametrization of the structure by a finite number of variables. In [196] a Kriging approach is used in order to evaluate the response of a wing to moderate perturbations of its geometry. In [63, 222, 81, 11] the uncertain domain is represented by the perturbation of a reference shape by a Lipschitz continuous random field. A different approach based on the density method is proposed in [162], where the uncertainty

of the shape is modeled by the presence of random parameters in the smoothing of the density function. Topology optimization techniques under uncertainties on the structure have proven to be useful in the context of shape identification problems [82, 83].

The presence of uncertainties in the material properties or in the applied loads has been studied by several different authors, and many different perspectives have been proposed. One possibility is to suppose that the probability distribution of the perturbation is known or it can be inferred from experimental measures. In this class of approaches we find the techniques denoted *Robust Topology Optimization* (RTO) and *Reliability-Based Topology Optimization* (RBTO) [121]. The RTO approach evaluates the expectation and the variance of the uncertain functionals of interest. The RBTO methods consist in estimating the probability of some functional to exceed a given threshold, denoted *failure probability* [10, 11], and require an accurate knowledge of the distribution of the uncertainties. More details and references about RTO and RBTO problems are presented in chapter 4 and chapter 5 of this thesis. It should be remarked that assumptions on the distribution of the uncertainties can be inaccurate, and a possible approach to this issue can be found in [93].

A different class of problems assumes that the perturbation belongs to a bounded set and optimize the structure with respect to the worst-case scenario. The functional whose value ought to be controlled can either appear in the objective function or in the constraint of the optimization problem, and the two situations necessitate different approaches. The case where the uncertain functional appears as a constraint, is the main interest of chapter 3. As reported in [43], the optimization of the worst-case scenario is preferable to robust or reliability-based optimization problems when the data are imprecise or if their probability distribution is unknown, if the strict respect of the constraint in all circumstances is of primary importance, or if the optimal solution is particularly difficult to manufacture.

The different approaches to the inclusion of uncertainties in the external condition or in the material properties have some drawbacks and some advantages. On the one hand, the optimization under constraints on the worst case does not require any hypothesis on the distribution of the uncertainties. On the other hand, it can result in a design that is too pessimistic, with poor performances in the vast majority of situations in order to cover an implausible set of circumstances. For most mechanical applications, the shape of the structure and the material properties are known with precision, but the loads applied to them are subject to a higher variability. Therefore, the present thesis will focus on the case of uncertain external conditions, without any precondition on the size of the perturbations.

Propagation of the uncertainties in linear elasticity

The main aim of the present thesis is to analyze shape optimization problems in linear elasticity, when the mechanical load is subject to uncertainties in the direction and intensity. More precisely, let us consider a mechanical structure composed by an elastic homogeneous isotropic material. We denote $\mathcal{S}_{\text{adm}}$ the class of all admissible shapes that the structure can assume, that we suppose to be open Lipschitz continuous domains in $\mathbb{R}^d$. For any $\Omega \in \mathcal{S}_{\text{adm}}$ we assume that its boundary can be partitioned in three portions with strictly positive measure: Γ_{D}, where the structure is clamped, Γ_{N} where mechanical loads are applied, and Γ_0 denoting the free surface. The elastic displacement $\mathbf{u}_\Omega$ induced by the mechanical solicitation of the structure can be computed using the equations of linear elasticity. Under such framework, we denote $\mathbf{f}$ the force density acting on the entire volume, $\mathbf{g}$ the surface load applied to the portion Γ_{N} of the surface of the structure, $\mathbf{u}_\Omega$ the displacement field, and $\boldsymbol{\sigma}(\mathbf{u}_\Omega)$ the stress tensor. Homogeneous

Dirichlet boundary conditions are applied on Γ_{D}, while Γ_0 and Γ_{N} are subject to homogeneous and non-homogeneous Neumann conditions respectively.

We are interested in finding the admissible structure with minimal volume satisfying a certain criterion in the form $H(\mathbf{u}_\Omega, \Omega) \leq \tau$ for a suitable threshold τ. The optimization problem can be formulated as follows

$$
\left|\begin{array}{l}
\text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\
\text{minimizing the volume } \mathrm{Vol}(\Omega) \\
\text{under the constraint } H(\mathbf{u}_\Omega, \Omega) \leq \tau, \\
\text{where the displacement } \mathbf{u}_\Omega \\
\text{solves the elasticity equation} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\
\mathbf{u}_\Omega & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}.
\end{array}\right.
\end{array}\right.
$$

For the sake of introducing uncertainties, we consider the probability space $(\mathcal{O}, \mathcal{A}, \mathbb{P})$, where the outcome space is denoted by $\mathcal{O}$, $\mathcal{A} \subset 2^{\mathcal{O}}$ is a σ-algebra on $\mathcal{O}$, and $\mathbb{P}$ a probability measure. Let us suppose that the mechanical load $\mathbf{g}$ is uncertain, so that it can be modeled as a random variable. Thus, the uncertainties propagate to the displacement $\mathbf{u}_\Omega$ and to the constraint functional $H(\mathbf{u}_\Omega(\omega), \Omega)$. In order to solve the optimization problem and obtain a solution which is deterministic, it is necessary to replace the constraint functional $H(\mathbf{u}_\Omega, \Omega)$ with a deterministic quantity $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \in \mathbb{R}$. By taking uncertainties into account, the shape optimization problem above can be formulated as follows

$$
\left|\begin{array}{l}
\text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\
\text{minimizing the volume } \mathrm{Vol}(\Omega) \\
\text{under the constraint } \mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \leq \tau_{\mathcal{F}}, \\
\text{where the displacement } \mathbf{u}_\Omega(\omega) \\
\text{solves the elasticity equation} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{g}(\omega) & \text{on } \Gamma_{\mathrm{N}}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\
\mathbf{u}_\Omega(\omega) & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}},
\end{array}\right. \\
\text{for almost any event } \omega \in \mathcal{O}.
\end{array}\right.
$$

Different approaches to shape optimization under uncertainties can be distinguished by different choices of the operator $\mathcal{F}\left[\,\cdot\,\right]$. For the worst-case approach the deterministic constraint can be written as $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] = \operatorname{ess\,sup}_{\omega \in \mathcal{O}} H(\mathbf{u}_\Omega(\omega), \Omega) \leq \tau$. In the case of RTO problems, the operator $\mathcal{F}\left[\,\cdot\,\right]$ is a combination of the expectation and variance operators. Finally, for RBTO problems, $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right]$ is the probability for $H(\mathbf{u}_\Omega, \Omega)$ to exceed some tolerance, and the inequality $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \leq \tau_{\mathcal{F}}$ can be interpreted as a constraint on the failure probability.

Summary of the thesis by chapter

This thesis is divided in two main parts. In part I we introduce the main notions on shape optimization for elastic structures, and we study some optimization problems in a deterministic setting. Part II presents different approaches to take into account uncertainties on the mechanical loads applied to the structure. Part I is divided in two chapters, while part II consists of three chapters, one for each approach.

Chapter 1: Hadamard's method for shape optimization

In this introductory chapter we present the main notions related to shape optimization used in the course of this thesis.

In section 1.1 we start by recalling the general framework of constrained optimization problems. Then, we focus on optimization problems under constraints of Partial Differential Equations, and we discuss the notion of *gradient-based* optimization algorithms. Particular attention is given to the computation and meaning of the adjoint state. The section is closed by some recalls and notations of function spaces.

Section 1.2 describes the differentiation method for shape functionals as introduced by Hadamard in 1908 [131], following the notations of [6, 100, 138]. After recalling the notions of Fréchet and Gâteaux derivatives, we introduce the concept of shape derivative, we enunciate Hadamard's structure theorem (theorem 1.7), and we show some results on the derivatives of volume and surface functionals. Next, we present two methods to compute the shape derivative of a volume functional relying on the solution of a Partial Differential Equation (PDE). In section 1.2.2 the shape derivative is obtained by computing the Lagrangian derivative of the solution of the PDE and deducing the adjoint equation from it. In section 1.2.3 the same expression is obtained through Céa's fast derivation method [61]. Further details on the computation of some expressions are reported in appendix A.

Finally, in section 1.3 we discuss the numerical methods for the simulations performed in this thesis, as well as some alternatives. First, we introduce the level-set method and the signed-distance function for shape representation. Next, we discuss the notion of adaptive meshing and the ersatz material method in order to solve boundary value problems for shape encoded by level-set functions. In section 1.3.4 we discuss the *null space* optimization algorithm as introduced by Feppon, Allaire and Dapogny in [114]. In conclusion, we summarize the different steps of the shape optimization algorithm used in this thesis.

Chapter 2: A deterministic thermo-elastic problem

The second chapter of part I focuses on the optimization of elastic mechanical structures in the deterministic case. In section 2.1.1 we start by introducing the equations of linear elasticity, the strain and stress tensors, the symmetry properties of the fourth-order elasticity tensor, and we express the linear elasticity equation for isotropic homogeneous materials. Next, we present two shape functionals of particular interest in the design of mechanical structures: the compliance (the work of the external forces), and the von Mises stress.

In section 2.2 we discuss the influence of thermal effects for the optimization of an elastic structure. We begin by presenting the constitutive equations of an elastic material subject to thermal dilation. This problem is of particular interest in the context of additive manufacturing, as discussed by Allaire and Jakabčin [15], as well as for multiphysics applications [112].

We suppose the mechanical response to be much faster than changes in the temperature field.

Thus, we consider the time-dependent diffusion equation to compute the thermal distribution at each instant, while the displacement is modeled by the quasi-static system for linear elasticity. In the rest of the section we formulate a shape optimization problem for an elastic structure subject to a time-variant thermal field, with the objective of minimizing its volume under a constraint on its mechanical compliance. The constraint is formulated by weighting two terms. The first one consists in the average in time of the compliance for the duration of the simulation. The second component, not examined as often as the first for time-dependent shape optimization problems, considers only the final state of the structure. The shape derivative of such functional is computed in section 2.2.3.

In section 2.3 we present the results of some numerical simulations, starting from the minimization of the volume of a 3D cantilever under constraints on the compliance and the von Mises stress, and concluding with a few results on the thermo-elastic problem. The results of section 2.2 have been obtained in collaboration with Viacheslav Karnaev, PhD student at the University of Basel. The study of the thermo-eslatic problem is part of an ongoing work on shape optimization under a time-variant and uncertain thermal field, and it will be the subject of future publications.

Chapter 3: Optimization of the worst-case scenario

From this chapter on, we consider various approaches to take uncertainties on the mechanical loads into account while solving problems of shape optimization. This first chapter of part II presents two different methods to address the case where one of the constraints in the optimization problem concerns the maximal possible value that a given shape functional can assume. For both cases we suppose that the uncertain loads can be parametrized by a finite number of variables belonging to a bounded and convex subset $\mathcal{G}$ of a suitable Banach space.

The first method, detailed in section 3.2, requires the convexity of the function $\mathcal{G} \ni \mathbf{g} \mapsto H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$ mapping the parameters describing the uncertain loads to the value of the constraint functional. The strategy consists in approximating the set $\mathcal{G}$ by a convex polyhedron $\mathcal{G}_N$ with a finite number of vertices N. Thanks to the convexity of the mapping $\mathbf{g} \mapsto H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$, the maximal value of the constraint is attained in one of the vertices of $\mathcal{G}_N$. Thus, the shape optimization problem can be reformulated as an equivalent problem subject to N constraints, each one relative to a different loading case. A convergence result for the solution of the approximate problem to the solution of the original is provided in section 3.2.3. This method mirrors and justifies the approach of engineers consisting in designing structures that are robust with respect to a limited number of loading cases deemed representative.

In section 3.3 we propose an alternative method. After introducing the notion of subdifferential in the sense of Clarke, we adapt it to shape functionals, and we provide a sufficient condition for its existence under some conditions on the regularity of the functional and of $\mathcal{G}$. The core of the proposed method consists in the identification of one element of the subdifferential at each step of the optimization algorithm, and the computation of the shape gradient in the corresponding direction. Section 3.3.3 details the numerical implementation of the subdifferential approach.

In conclusion, section 3.4 presents the results of some numerical simulations, optimizing the volume of two different structures under constraints on the compliance or the L^6-norm of the von Mises stress. The polyhedral approximation and the subdifferential methods have both been used to enforce the constraint. The numerical results validate both approaches, and illustrate the convergence of the solution of the polyhedral approximation method towards the exact one when $\mathcal{G}_N$ approaches $\mathcal{G}$. Moreover, the simulations show that the approach based on the

subdifferential is generally faster than the polyhedral approximation of the set $\mathcal{G}$, but it can be subject to oscillations if the structure to be optimized and the set of admissible loads have multiple symmetries. A paper on this subject is currently been drafted.

Chapter 4: Shape optimization of a polynomial functional

In this chapter, we model the uncertain loads applied to the structure as random variables in order to study a Robust Topology Optimization method. Different approaches to the shape optimization of the expectation or the variance of a shape functional have been proposed in the recent years [106, 11, 211, 80]. In particular, Dambrine, Dapogny and Harbrecht proved in [80] that the value of the expectation of a quadratic functional depends only on the first two stochastic moments of the random variables describing the uncertain loads. Therefore, the computations of such expectation and its shape derivative do not require computationally expensive sampling methods.

The main objective of this chapter is to extend the approach of [80] to m-multilinear functionals. The setting of the problem is described in section 4.2, while the main theoretical results are presented in section 4.3. In particular, we extend the definition of the correlation operator similarly to [226] and, focusing on the constext of linear elasticity, we use a tensor approach to show that expectation of an m-multilinear functional depends only on the m first stochastic moments of the random loads.

In the following two sections we show different applications of our method completed by some numerical results. In section 4.4 we address the issue of the constraint of the L^∞-norm of the von Mises stress. Since the L^∞-norm cannot be differentiated, we approximate it by the L^m-norm for m sufficiently large, which can be written in terms of an m-multilinear functional. As numerical simulation we consider the optimization of the mass of a 3D cylindrical structure subject to random shear forces on its upper face, under a constraint on the L^6-norm of the von Mises stress. In section 4.5 we express the variance of a quadratic functional in terms of a multilinear functional of degree 4. Then, we adapt a 2D example from [80] to show the effect of a constraint on the expected value and the variance of the compliance.

The main results of this chapter have been submitted as a journal paper titled *Shape Optimization of Polynomial Functionals under Uncertainties on the Right-Hand Side of the State Equation*, co-written with Fabien Caubet, Marc Dambrine, and Jérôme Maynadier.

Chapter 5: Shape optimization of the probability to exceed a threshold

This final chapter focuses on Reliability-Based Topology Optimization problems for quadratic functionals. The interest of RBTO techniques lies in the possibility to design structures compliant with requirement on their probability of failure [34, 180, 151, 76, 11].

Let $\mathbf{g}$ be a random mechanical load described by a random vector, $\mathbf{u}_{\Omega,\mathbf{g}}$ the associated elastic displacement, and $(\mathbf{u},\mathbf{v}) \mapsto Q(\mathbf{u},\mathbf{v}) \in \mathbb{R}$ be a continuous quadratic functional of the displacement. The functional depends on the shape of the structure (like the mechanical compliance or the L^2-norm of the von Mises stress). Let τ be a threshold value, and let us denote $\Phi(\cdot)$ the shape functional such that

$$\Phi(\Omega) = \mathbb{P}\left[Q(\mathbf{u}_{\Omega,\mathbf{g}}, \mathbf{u}_{\Omega,\mathbf{g}}) > \tau\right],$$

where $\mathbf{u}_{\Omega,\mathbf{g}}$ is the solution of the elasticity equation (which is subject to uncertainties). The method to compute the value and the shape derivative of $\Phi(\Omega)$ is detailed in sections 5.2 and 5.3, and relies on the computation of suitable integrals of the probability density on the volume and on the surface of the ellipsoid defined by the inequality $Q(\mathbf{u}_{\Omega,\mathbf{g}}, \mathbf{u}_{\Omega,\mathbf{g}}) \leq \tau$.

In section 5.4 we present a different technique to evaluate the probability of failure $\Phi(\Omega)$ and its derivative when the external load follows a Gaussian distribution. This second method relies on the series decomposition of the cumulative distribution function of a generalized noncentral chi-squared distribution as proven by Ruben in 1962 [214].

In section 5.5 we detail the algorithms used to perform the numerical simulations for this chapter. In order to test the method proposed in section 5.3 we considered problems subject to centered Gaussian uncertainties, and we apply the quadrature formulas reported in appendix C in order to avoid the numerical computation of integrals in large dimensions. Both methods have been tested for the optimization of two elastic structures, with the objective to minimize their volume under constraints on the probability for the compliance to exceed a threshold. The results are presented in section 5.6. The computation of the coefficients appearing in the expression of the shape derivative is almost instantaneous. Thus, the solution of such RBTO problems takes around the same time as the solution of a deterministic shape optimization problem.

The essential of this chapter has been submitted as a journal paper under the title *Shape optimization under constraints on the probability of a quadratic functional to exceed a given threshold*, jointly written with Marc Dambrine, Helmut Harbrecht, and Jérôme Maynadier.

Part I

Shape optimization in a deterministic setting

Chapter 1

Hadamard's method for shape optimization

Contents

1.1 PDE constrained optimization

1.1.1 Notations and general framework

Following the framework proposed in [176], any optimization problem can be characterized in terms of optimal control theory by a *state problem*, an *objective function* and a *control*. The control x is a variable or a function belonging to a given set of admissible controls $\mathcal{X}_{\text{adm}}$. In the domain of constrained optimization, the state is another variable y belonging to a Banach space of admissible states $\mathcal{Y}_{\text{adm}}$ such that, for a given control $x \in \mathcal{X}_{\text{adm}}$, it solves the *state problem*

$$\mathbf{f}(y, x) = \mathbf{0}, \tag{1.1}$$

where $\mathbf{f} : \mathcal{Y}_{\text{adm}} \times \mathcal{X}_{\text{adm}} \to \mathcal{W}$ is a function taking values in a vector space $\mathcal{W}$. If the state problem is well-posed, the state can be seen as an implicit function of the control as $y = y(x)$. The

objective function is a real-valued function of the control and of the state $J : \mathcal{X}_{\text{adm}} \times \mathcal{Y}_{\text{adm}} \to \mathbb{R}$. If the state problem (1.1) is well-posed, a *reduced cost* $\widehat{J} : \mathcal{X}_{\text{adm}} \to \mathbb{R}$ can be defined as a function of the control as

$$\widehat{J} : x \mapsto \widehat{J}(x) = J(y(x), x). \tag{1.2}$$

Using the notation stated above, an optimization problem can be written as

$$\left|\begin{array}{l} \text{Find the } \textit{control } x \in \mathcal{X}_{\text{adm}} \\ \text{minimizing the } \textit{objective function } J(y(x), x) \\ \text{where the } \textit{state } y(x) \in \mathcal{Y}_{\text{adm}} \text{ is solution} \\ \text{of the } \textit{state problem } \mathbf{f}(y, x) = \mathbf{0}. \end{array}\right. \tag{1.3}$$

We remark that in problem (1.3) the state problem (1.1) acts as a constraint of the optimization problem. Problem (1.3) can be expressed more concisely using the reduced objective function as

$$\left|\begin{array}{l} \text{Find the } \textit{control } x \in \mathcal{X}_{\text{adm}} \\ \text{minimizing the } \textit{reduced objective function } \widehat{J}(x). \end{array}\right. \tag{1.4}$$

For the purposes of PDEs constrained optimization, equation (1.1) represents a boundary-value problem for partial differential equations, and the state y is a function belonging to a suitable Sobolev space and defined on a given domain Ω. Problem (1.4) can be further enriched by imposing some constraints on the control x, which may depend on the solution of one or more state equations.

$$\left|\begin{array}{l} \text{Find the } \textit{control } x \in \mathcal{X}_{\text{adm}} \\ \text{minimizing the } \textit{reduced objective function } \widehat{J}(x) \\ \text{under the } \textit{constraints} \\ \qquad \left\{\begin{array}{rcl} \widehat{\mathbf{G}}(x) & = & \mathbf{0}, \\ \widehat{\mathbf{H}}(x) & \leq & \mathbf{0}, \end{array}\right. \end{array}\right. \tag{1.5}$$

where $\widehat{\mathbf{G}} : \mathcal{X}_{\text{adm}} \to \mathbb{R}^{N_{\mathbf{G}}}$ encodes $N_{\mathbf{G}}$ equality constraints, and $\widehat{\mathbf{H}} \to \mathbb{R}^{N_{\mathbf{H}}}$ encodes $N_{\mathbf{H}}$ inequality constraints. The coverage of the constrained optimization problem (1.5) is discussed in section 1.3.4

Optimal control problems can be distinguished in different categories according to the nature of the control function. If $\mathcal{X}_{\text{adm}}$ consists of functions defined on the domain Ω or on some part of it, problem (1.3) is said to be a *distributed control problem*. Density approaches to topology optimization like the SIMP method [46, 47] can be classified in this category. A second class encompasses the boundary control problems, for which the control x is a function defined on $\partial\Omega$ or on part of it. In the context of this thesis we are interested in a third class of optimization problems, where the control is the shape of the domain Ω itself.

1.1.2 Gradient-based optimization algorithms for constrained optimization

The algorithms for the solution of optimization problems can be divided in two classes. *Gradient-free methods* include various evolutionary methods [256, 54] among which we find the genetic algorithm (see [149] and [6, Chapter 8]). *Gradient-based methods* rely on the computation of the sensitivity of the objective function and the constraint with respect to the control, and require some further regularity on J and $\mathbf{f}$ with respect to the *gradient-free methods*. It should

be remarked that gradient-free and gradient-based methods can be used together in order to accelerate the solution of an optimization problem. See [49] as an example in the context of topology optimization, where a genetic algorithm is used to find a suitable initial condition for a gradient-based method.

In general, all gradient-based algorithm follow a similar structure [126, Section 2.3]. Starting from an initial guess $u^{(0)}$ for the control, the following loop is executed starting from $k = 0$ until a certain convergence criterion is satisfied:

1. identification of the state $y^{(k)} = y\left(x^{(k)}\right)$ corresponding to the control $x^{(k)}$ by solving the state problem $\mathbf{f}(y^{(k)}, x^{(k)}) = \mathbf{0}$;

2. computation of the derivative of the reduced objective function

$$\frac{\mathrm{d}}{\mathrm{d}x}\widehat{J}\left(x^{(k)}\right) = \frac{\mathrm{d}}{\mathrm{d}x}J\left(y\left(x^{(k)}\right), x^{(k)}\right); \tag{1.6}$$

3. computation of an increment $\delta x^{(k)}$ from the results of points 1 and 2;

4. update the control $x^{(k+1)} \leftarrow \delta x^{(k)} + x^{(k)}$ and increment k.

The computation of the derivative of the objective function with respect to the control is a crucial step in any gradient-based optimization algorithm [126, Section 2.3.2]. By applying the chain-rule on equation (1.6) we obtain the following expression for the gradient

$$\frac{\mathrm{d}}{\mathrm{d}x}\widehat{J}\left(x^{(k)}\right) = \left.\frac{\partial J}{\partial x}\right|_{x=x^{(k)}} + \left.\frac{\partial J}{\partial y}\right|_{y=y^{(k)}} \left.\frac{\partial y}{\partial x}\right|_{x=x^{(k)}}. \tag{1.7}$$

The terms $\frac{\partial J}{\partial x}|_{x=x^{(k)}}$ and $\frac{\partial J}{\partial y}|_{y=y^{(k)}}$ can easily be computed from the expression of J. The term $\frac{\partial y}{\partial x}|_{x=x^{(k)}}$ is the *sensitivity* of the state with respect to the control, and its computation is more delicate. One possibility is to estimate it numerically after the discretization of the control. Such a strategy can be very costly if the discretized control is represented by a large number of parameters, since it requires to solve the state problem for all possible directions of perturbation of the control.

A different approach to compute the sensitivity $\frac{\partial y}{\partial x}$ consists in the solution of the following *sensitivity equation*, obtained by the differentiation of the state problem

$$\left(\left.\frac{\partial \mathbf{f}}{\partial y}\right|_{y=y^{(k)}}\right) \left.\frac{\partial y}{\partial x}\right|_{x=x^{(k)}} = -\left.\frac{\partial \mathbf{f}}{\partial x}\right|_{x=x^{(k)}}. \tag{1.8}$$

The numerical solution of equation (1.8) is challenging if the control is discretized by a large number of parameters, since the sensitivity equation should be solved for as many different right-hand sides as the number of degrees of freedom of the control.

In order to avoid such complication, it is possible to consider an adjoint approach, as presented in [169, Section 6.5], [126, Section 2.3.3], and [176, Section 9.3.2].

Definition 1.1 (Adjoint operator)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed vector spaces. We denote $\mathcal{X}^*$ the topological dual of $\mathcal{X}$, which is the set of bounded linear operators defined on $\mathcal{X}$. For any $L \in \mathcal{X}^*$ and $x \in \mathcal{X}$, we denote $\langle L, x\rangle_{\mathcal{X}^*,\mathcal{X}}$ the evaluation of L in the point x.*

Let $A : \mathcal{X} \to \mathcal{Y}$ be a bounded linear operator. The adjoint operator $A^* : \mathcal{Y}^* \to \mathcal{X}^*$ is such that, for any $y^* \in \mathcal{Y}^*$ and $x \in \mathcal{X}$

$$\langle A^* y^*, x \rangle_{\mathcal{X}^*,\mathcal{X}} = \langle y^*, Ax \rangle_{\mathcal{Y}^*,\mathcal{Y}} . \tag{1.9}$$

In the finite-dimensional case, A is a matrix, and its adjoint A^* operator coincides with its conjugate transpose.

Let us suppose that the sets of the admissible controls $\mathcal{X}_{\text{adm}}$ and states $\mathcal{Y}_{\text{adm}}$ are subsets of the reflexive Banach spaces $\mathcal{U}$ and $\mathcal{Y}$ respectively. Moreover, we consider $\mathcal{W}$ to be a Banach space as well. Since $\frac{\partial \mathbf{f}}{\partial y}|_{y=y^{(k)}} : \mathcal{Y} \to \mathcal{W}$ is a bounded linear operator, we can identify is dual by the operator $\left(\frac{\partial \mathbf{f}}{\partial y}|_{y=y^{(k)}}\right)^* : \mathcal{W}^* \to \mathcal{Y}^*$. Let us define the *adjoint state* (or *co-state*) $\lambda^{(k)} \in \mathcal{W}^*$, as the solution of the *adjoint equation*

$$\left(\left.\frac{\partial \mathbf{f}}{\partial y}\right|_{y=y^{(k)}}\right)^* \lambda^{(k)} = \left(\left.\frac{\partial J}{\partial y}\right|_{y=y^{(k)}}\right)^* . \tag{1.10}$$

By injecting the result of the adjoint equation (1.10) and the sensitivity equation (1.8) into equation (1.7) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\widehat{J}\left(x^{(k)}\right) = \left.\frac{\partial J}{\partial x}\right|_{x=x^{(k)}} + {\lambda^{(k)}}^* \left.\frac{\partial \mathbf{f}}{\partial y}\right|_{y=y^{(k)}} \left.\frac{\partial y}{\partial x}\right|_{x=x^{(k)}} = \left.\frac{\partial J}{\partial x}\right|_{x=x^{(k)}} - {\lambda^{(k)}}^* \left.\frac{\partial \mathbf{f}}{\partial x}\right|_{x=x^{(k)}} . \tag{1.11}$$

Since the adjoint state $\lambda^{(k)}$ does not depend on the perturbation $\widehat{\boldsymbol{\theta}}$, it can be solved only once per iteration of the optimization algorithm, after the solution of the state problem.

Crucial assumptions for the application of gradient-based optimization algorithms are the differentiability of the objective J and the state function $\mathbf{f}$ with respect to the control, and the ability to upgrade the control by a suitable increment. In the case of distributed or boundary control problem, the control function x is supposed to belongs to a closed subset of a suitable Banach space. However, for shape optimization problems the set of admissible controls $\mathcal{X}_{\text{adm}}$ consists in a set of domains in $\mathbb{R}^d$, which is not provided with a natural vector structure.

One possible solution consists in the parametrization of the domain by a finite number N of parameters, so that $\mathcal{X}_{\text{adm}}$ is isomorphic to a subset of $\mathbb{R}^N$ (see [189] and other chapters of [139]). Despite its simplicity and widespread use in industrial contexts, such approach bounds the search for the optimal shape to a narrow class of structures and does not allow for radical changes of the geometry.

The approach presented in section 1.2 defines the derivative with respect to shape in terms of perturbations of the domain, allowing for a greater flexibility in the choice of the control and a greater complexity in the resulting optimal shapes. In this section we discussed the differentiation of the objective function of an optimization problem. It is evident that the same considerations can be done for the derivation of a constraint functional.

1.1.3 Recalls, remarks, and notations on function spaces

Let us consider a bounded domain $\Omega \subset \mathbb{R}^d$ with Lipschitz continuous boundary, in dimension $d = 2$ or 3. Before stating the definition of shape derivative proposed by Hadamard, we recall the definition of some functional spaces [216, Chapter 3]. For all $1 \leq p < \infty$, the set $\mathrm{L}^p(\Omega)$ is the space of all real-valued functions f defined on Ω such that $|f|^p$ is integrable on Ω. The space $\mathrm{L}^\infty(\Omega)$ contains all essentially bounded functions defined on Ω.

$$\begin{aligned} \mathrm{L}^p(\Omega) &= \{f : \Omega \to \mathbb{R} \ : \ \textstyle\int_\Omega |f|^p \, \mathrm{d}\mathbf{x} < \infty\} \qquad \text{for } 1 < p < \infty, \\ \mathrm{L}^\infty(\Omega) &= \{f : \Omega \to \mathbb{R} \ : \ \text{ess sup } f < \infty\} \end{aligned}$$

The function spaces $\mathrm{L}^p(\Omega)$ are Banach spaces with respect to the norms

$$\begin{array}{rcll} \|f\|_p & = & (\int_\Omega |f|^p \, \mathrm{d}\mathbf{x})^{1/p} & \text{for } 1 \leq p < \infty, \\ \|f\|_\infty & = & \text{ess sup}\, f. & \end{array} \tag{1.12}$$

If the domain Ω has finite measure, the L^p spaces are nested into one another as proven by the following proposition.

Proposition 1.2. *Let Ω be a domain with finite measure $|\Omega| < \infty$. Let $p, q \in [1, \infty[\cup \{+\infty\}$, where $p < q$, and $f \in L^q(\Omega)$. Then, $f \in \mathrm{L}^p(\Omega)$ as well and*

$$\begin{array}{rcll} \frac{\|f\|_p}{\sqrt[p]{|\Omega|}} & \leq & \frac{\|f\|_q}{\sqrt[q]{|\Omega|}} & \text{if } q < \infty, \\ \frac{\|f\|_p}{\sqrt[p]{|\Omega|}} & \leq & \|f\|_\infty & \text{otherwise.} \end{array} \tag{1.13}$$

Equation (1.13) can be proven using the definition of the L^p norms, Lebesgue's definition of integral, and Hölder's inequality. It can also be remarked that the L^p-norm for $p < \infty$ can be used to approximate the supremum of $|f|$ in Ω.

Proposition 1.3. *Let Ω be a domain with finite measure $|\Omega| < \infty$, and f a real-valued function such that $f \in \mathrm{L}^{p_0}(\Omega) \cap \mathrm{L}^\infty(\Omega)$. Then $f \in L^p(\Omega)$ for all $p > p_0$, and*

$$\lim_{p \to \infty} \|f\|_p = \|f\|_\infty .$$

Proof. At first, from equation (1.13), we can deduce that

$$\limsup_{p \to \infty} \|f\|_p \leq \lim_{p \to \infty} \|f\|_\infty = \|f\|_\infty . \tag{1.14}$$

In order to prove the opposite inequality, we define the family of sets A_ϵ for all $0 < \epsilon < \|f\|_\infty$ as

$$A_\epsilon = \{x \in \Omega \, : |f(x)| > \|f\|_\infty - \epsilon\} .$$

We remark that, by the definition (1.12) of the L^∞-norm, $|A_\epsilon| > 0$. We deduce that

$$\|f\|_p^p = \int_\Omega |f|^p \geq \int_{A_\epsilon} |f|^p \geq (\|f\|_\infty - \epsilon)^p \, |A_\epsilon| ,$$

from which follows

$$\|f\|_p \geq (\|f\|_\infty - \epsilon) \, |A_\epsilon|^{1/p} .$$

By taking the limit for $p \to \infty$, we get

$$\liminf_{p \to \infty} \|f\|_p \geq \lim_{p \to \infty} (\|f\|_\infty - \epsilon) \, |A_\epsilon|^{1/p} = (\|f\|_\infty - \epsilon) .$$

Thus, since such inequality is true for all $0 < \epsilon < \|f\|_\infty$:

$$\liminf_{p \to \infty} \|f\|_p \geq \|f\|_\infty . \tag{1.15}$$

Combining (1.14) and (1.15) we conclude that $\lim_{p \to \infty} \|f\|_p = \|f\|_\infty$. □

A major family of function spaces in the context of partial differential equations is composed by the Sobolev spaces $\mathrm{W}^{k,p}(\Omega)$ for $1 \leq k, p < \infty$ [110, Section 5.2]. For any multi-index $\boldsymbol{\alpha} \in \mathbb{N}^d$, we denote $|\boldsymbol{\alpha}|$ the sum of its components. A real-valued function f defined on $\Omega \subset \mathbb{R}^d$ belongs to $\mathrm{W}^{k,p}(\Omega)$ if, for any multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ such that $|\boldsymbol{\alpha}| \leq k$, the partial derivative $\mathrm{D}^{\boldsymbol{\alpha}} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}, \ldots, \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} f$ exists in the weak sense, and it belongs to $\mathrm{L}^p(\Omega)$.

$$\mathrm{W}^{k,p}(\Omega) = \left\{ f \in \mathrm{L}^k(\Omega) \ : \ \int_\Omega |\mathrm{D}^{\boldsymbol{\alpha}} f|^p \, \mathrm{d}\mathbf{x} < \infty \text{ for all multi-index } \boldsymbol{\alpha} \text{ such that } |\boldsymbol{\alpha}| \leq k \right\}.$$

All Sobolev spaces $\mathrm{W}^{k,p}(\Omega)$ for $1 \leq k, p < \infty$ are Banach spaces with respect to the following norm [110, Section 5.2, Theorem 2]

$$\|f\|_{k,p} = \left(\sum_{|\boldsymbol{\alpha}| \leq k} \|\mathrm{D}^{\boldsymbol{\alpha}} f\|_p^p \right)^{1/p}.$$

The spaces $\mathrm{W}^{k,2}(\Omega)$ are also denoted as $\mathrm{H}^k(\Omega)$ for any $1 \leq k < \infty$. The function space $\mathrm{L}^2(\Omega)$, as well as all spaces $\mathrm{H}^k(\Omega)$ are Hilbert spaces with respect to the following scalar products

$$\begin{aligned} \langle f, g \rangle_{\mathrm{L}^2} &= \int_\Omega fg \, \mathrm{d}\mathbf{x} \qquad f, g \in \mathrm{L}^2(\Omega), \\ \langle f, g \rangle_{\mathrm{H}^k} &= \sum_{|\boldsymbol{\alpha}| \leq k} \int_\Omega (\mathrm{D}^{\boldsymbol{\alpha}} f)\,(\mathrm{D}^{\boldsymbol{\alpha}} g) \, \mathrm{d}\mathbf{x} \qquad f, g \in \mathrm{L}^2(\Omega), k \geq 1. \end{aligned}$$

If $\Gamma \subset \partial\Omega$ is a portion of the boundary of Ω with strictly positive measure, we denote $\mathrm{L}^p_\Gamma(\Omega)$ (or $\mathrm{W}^{k,p}{}_\Gamma(\Omega)$, or $\mathrm{H}^k_\Gamma(\Omega)$) the subset of all functions belonging to $\mathrm{L}^p(\Omega)$ (or $\mathrm{W}^{k,p}(\Omega)$ or $\mathrm{H}^k(\Omega)$ respectively) which are equal to 0 almost everywhere on Γ.

The definition of the function spaces introduced here can be extended to vector-valued functions. We denote $\mathrm{L}^p(\Omega)^d$, $\mathrm{W}^{k,p}(\Omega)^d$, and $\mathrm{H}^k(\Omega)^d$ the equivalent spaces of $\mathrm{L}^p(\Omega)$, $\mathrm{W}^{k,p}(\Omega)$, and $\mathrm{H}^k(\Omega)$ for functions with values in $\mathbb{R}^d$. A function $\mathbf{f} : \Omega \to \mathbb{R}^d$ belongs to $\mathrm{L}^p(\Omega)^d$ if all its components belong to $\mathrm{L}^p(\Omega)$ (likewise for $\mathrm{W}^{k,p}(\Omega)^d$, and $\mathrm{H}^k(\Omega)^d$).

Finally, we denote $\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ the space of Lipschitz continuous vector field in $\mathbb{R}^d$, that is

$$\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d = \left\{ \boldsymbol{\theta} \in \mathrm{L}^\infty\left(\mathbb{R}^d\right)^d \ : \ \nabla\boldsymbol{\theta} \in \mathrm{L}^\infty\left(\mathbb{R}^d\right)^{d\times d} \right\}.$$

The set $\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ is a Banach space with respect to the norm $\|\cdot\|_{1,\infty}$ such that

$$\|\boldsymbol{\theta}\|_{1,\infty} = \|\boldsymbol{\theta}\|_\infty + \|\nabla\boldsymbol{\theta}\|_\infty = \sup_{\mathbf{x}\in\mathbb{R}^d} \|\boldsymbol{\theta}(\mathbf{x})\|_\infty + \sup_{\mathbf{x}\in\mathbb{R}^d} \|\nabla\boldsymbol{\theta}(\mathbf{x})\|_\infty.$$

1.2 Hadamard's boundary variation method

1.2.1 Shape derivative of a real-valued functional

The method of differentiation with respect to the domain presented in this section has been proposed by Jacques Hadamard [131]. All concepts, definition and proof presented here can be found in [138, Chapter 5], [6, Chapter 6], and [100, Chapter 9].

If $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ is a Lipschitz continuous vector field such that $\|\boldsymbol{\theta}\|_{1,\infty} < 1$ and denoting $\mathbf{I}$ the identity operator, the mapping $\mathbf{x} \mapsto (\mathbf{I} + \boldsymbol{\theta}(\mathbf{x}))\mathbf{x}$ is a Lipschitz continuous diffeomorphism, meaning that it is differentiable, invertible, and its inverse is differentiable (see [6, Lemma 6.13]).

Definition 1.4 (Lipschitz domain)**.** *A subset Ω of $\mathbb{R}^d$ is said to be a Lipschitz domain if Ω is an open bounded subset of $\mathbb{R}^d$, with positive measure, and such that its boundary $\partial\Omega$ can be described by a finite number of Lipschitz continuous local maps [218, Definition 1.4].*

Given a Lipschitz continuous vector field $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ with $\|\boldsymbol{\theta}\|_{1,\infty} < 1$ and a Lipschitz domain $\Omega \in \mathcal{S}_{\mathrm{adm}}$, we define the deformed domain $\Omega_{\boldsymbol{\theta}}$ as

$$\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega = \left\{\widetilde{\mathbf{x}} \in \mathbb{R}^d \ : \ \exists\, \mathbf{x} \in \Omega \text{ for which } \widetilde{\mathbf{x}} = (\mathbf{I} + \boldsymbol{\theta}(\mathbf{x}))\mathbf{x}\right\}. \tag{1.16}$$

For the sake of simplicity, we consider a class of admissible shapes $\mathcal{S}_{\mathrm{adm}}$, and a class Θ_{adm} of vector fields such that, for all $\boldsymbol{\theta} \in \Theta_{adm}$, the deformed set $\Omega_{\boldsymbol{\theta}}$ belongs to $\mathcal{S}_{\mathrm{adm}}$. Let $\Phi(\cdot)$ be a shape functional $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$. At first, we recall the notion of shape differentiability, as introduced in [138, Chapter 5] or in [6, Section 6.3].

Definition 1.5 (Fréchet differentiable shape functional)**.** *A real-valued functional $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ is shape differentiable according to Fréchet at Ω if there exists a linear continuous function $A_\Omega : \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \to \mathbb{R}$ such that*

$$\Phi(\Omega_{\boldsymbol{\theta}}) = \Phi(\Omega) + A_\Omega(\boldsymbol{\theta}) + o(\boldsymbol{\theta})$$

for all $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$, where $\lim_{\boldsymbol{\theta}\to 0} \frac{o(\boldsymbol{\theta})}{\|\boldsymbol{\theta}\|_{1,\infty}} = 0$. The linear form A_Ω is called shape derivative of Φ in Ω and is denoted as $\frac{\mathrm{d}}{\mathrm{d}\Omega}\Phi(\Omega)$ or $\mathrm{D}\Phi(\Omega)$.

Fréchet's definition of derivative is not the only possible definition of derivative in the context of shape calculus. We report here the definition of Gâteaux differentiability as presented in [100, Definition 3.3], which extends the notion of directional derivative, but is weaker than the Fréchet definition.

Definition 1.6 (Gâteaux differentiable shape functional)**.** *A real-valued shape functional $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ admits a Gâteaux semiderivative $\mathrm{d}\Phi(\Omega;\boldsymbol{\theta})$ in the direction $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ if the following limit exists and is finite*

$$\mathrm{d}\Phi(\Omega;\boldsymbol{\theta}) = \lim_{t \searrow 0^+} \frac{\Phi((\mathbf{I} + t\boldsymbol{\theta})\Omega) - \Phi(\Omega)}{t}.$$

The shape function Φ is said to be Gâteaux differentiable in Ω if it admits a Gâteaux semiderivative for any direction $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$, and the mapping $\boldsymbol{\theta} \mapsto \mathrm{d}\Phi(\Omega;\boldsymbol{\theta})$ is linear and continuous.

Unless stated otherwise, the notion of differentiability of shape functionals considered in this thesis refers to the Fréchet definition. If the domain Ω is sufficiently regular, we can assume that the value of the derivative $\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta})$ depends only on the normal component of the vector field $\boldsymbol{\theta}$ on the surface $\partial\Omega$ of the domain. Such result derives from the following theorem, proven by Hadamard and stated in [138, Proposition 5.9.1].

Theorem 1.7 (Structure theorem)**.** *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a $\mathcal{C}^1$ domain, and let us denote by $\mathbf{n}(\mathbf{s})$ the outwards normal to the surface $\partial\Omega$ in $\mathbf{s} \in \partial\Omega$. We suppose that $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ is a differentiable functional. If $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ are Lipschitz continuous vector fields such that $\boldsymbol{\theta}_1 \cdot \mathbf{n} = \boldsymbol{\theta}_2 \cdot \mathbf{n}$ on $\partial\Omega$, then $\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}_1) = \mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}_2)$.*

The expression of the shape derivative of a shape functional can be computed directly using definition 1.5. We recall the computation of the shape derivative of a volume functional as presented in [138, Section 5.2] and [6, Section 6.3].

Proposition 1.8 (Shape derivative of a volume function)**.** *Let $f \in \mathrm{W}^{1,1}\left(\mathbb{R}^d\right)$ be a real valued function. Let Ω be an admissible domain such that, for any admissible displacement field $\boldsymbol{\theta} \in \Theta_{adm}$, the perturbed domain $\Omega_{\boldsymbol{\theta}}$ belongs to $\mathcal{S}_{\mathrm{adm}}$. Then, the shape functional $\Phi : \Omega \mapsto \int_\Omega f(\mathbf{x})\mathrm{d}\mathbf{x}$ is shape differentiable in Ω, and its derivative can be written as*

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_\Omega \operatorname{div}\left(\widehat{\boldsymbol{\theta}}(\mathbf{x}) f(\mathbf{x})\right) \mathrm{d}\mathbf{x}. \tag{1.17}$$

If, moreover, Ω has a Lipschitz continuous boundary, the shape derivative of Φ can be also expressed as

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_{\partial\Omega} \left(\mathbf{n}(\mathbf{s}) \cdot \widehat{\boldsymbol{\theta}}(\mathbf{s})\right) f(\mathbf{s}) \,\mathrm{d}\mathbf{s}. \tag{1.18}$$

Proof. At first, we compute the value of the shape functional in a perturbed domain. Let $\boldsymbol{\theta} \in \Theta_{adm} \subset \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ be a vector field such that $\|\boldsymbol{\theta}\|_{1,\infty} < 1$. Assuming that the perturbed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$ belongs to the class $\mathcal{S}_{\mathrm{adm}}$, it is possible to evaluate Φ in $\Omega_{\boldsymbol{\theta}}$. Performing the change of variables $\mathbf{x} = (\mathbf{I} + \boldsymbol{\theta})^{-1}\widetilde{\mathbf{x}}$, the integral on $\Omega_{\boldsymbol{\theta}}$ can be written as an integral on the reference domain Ω

$$\Phi(\Omega_{\boldsymbol{\theta}}) = \int_{\Omega_{\boldsymbol{\theta}}} f(\widetilde{\mathbf{x}}) \,\mathrm{d}\widetilde{\mathbf{x}} = \int_\Omega f \circ (\mathbf{I} + \boldsymbol{\theta})(\mathbf{x}) \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \mathrm{d}\mathbf{x}, \tag{1.19}$$

where $\mathbb{I}$ denotes the identity matrix. Since $\mathbf{x} \mapsto (\mathbf{I} + \boldsymbol{\theta}(\mathbf{x}))\mathbf{x}$ is a diffeomorphism and f is a differentiable function, it is possible to differentiate equation (1.19) around $\boldsymbol{\theta} = 0$. Recalling that the derivative of the Jacobian term is

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}} \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \Big|_{\boldsymbol{\theta}=0} (\widehat{\boldsymbol{\theta}}) = \operatorname{div}\widehat{\boldsymbol{\theta}}, \tag{1.20}$$

we find the expression for the functional Φ as

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\Phi(\Omega_{\boldsymbol{\theta}})\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = \int_\Omega \widehat{\boldsymbol{\theta}}(\mathbf{x}) \cdot \nabla f(\mathbf{x}) \mathrm{d}\mathbf{x} + \int_\Omega f(\mathbf{x}) \operatorname{div}\widehat{\boldsymbol{\theta}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \int_\Omega \operatorname{div}\left(\widehat{\boldsymbol{\theta}}(\mathbf{x}) f(\mathbf{x})\right) \mathrm{d}\mathbf{x}. \tag{1.21}$$

The surface expression (1.18) of the shape derivative can be found integrating by parts the volume expression (1.21). □

The computation of the derivative of shape functionals defined on the boundary of Ω is more delicate. We start by recalling some notions of differential geometry as reported in [138, Sections 5.4.1 to 5.4.3].

Definition 1.9 (Differential operators on a smooth surface)**.** *Let $\Gamma \subset \mathbb{R}^d$ be a smooth manifold without boundary of dimension $(d-1)$, and $g : \Gamma \to \mathbb{R}$ a $\mathcal{C}^1$ function. Let $\widetilde{g} : \mathbb{R}^d \to \mathbb{R}$ a $\mathcal{C}^1$ extension of g to $\mathbb{R}^d$. The tangential gradient of g in $\mathbf{s} \in \Gamma$ is defined as*

$$\nabla_\Gamma g(\mathbf{s}) = \nabla\widetilde{g}(\mathbf{s}) - \left(\mathbf{n} \cdot \nabla\widetilde{g}(\mathbf{s})\right) \mathbf{n},$$

where $\mathbf{n}(\mathbf{s})$ *is the unit vector normal to* Γ *in* $\mathbf{s}$.

The tangential divergence of a $\mathcal{C}^1$ *function* $\mathbf{g} : \Gamma \to \mathbb{R}^n$ *in* $\mathbf{s}$ *is defined as*

$$\left(\operatorname{div}_\Gamma \mathbf{g}\right)(\mathbf{s}) = \left(\operatorname{div} \widetilde{\mathbf{g}}\right)(\mathbf{s}) - \left(\nabla \widetilde{\mathbf{g}}\ \mathbf{n}\right) \cdot \mathbf{n},$$

where $\widetilde{\mathbf{g}}$ *denotes a* $\mathcal{C}^1$ *extension of* $\mathbf{g}$ *to* $\mathbb{R}^n$.

It should be remarked that the definition of the tangential gradient and derivative are independent from the choice of the extension function. In order to state the formula for the change of variables for functionals defined on manifolds [138, Proposition 5.4.3], we introduce the notion of mean curvature of a smooth manifold.

Definition 1.10 (Mean curvature)**.** *Let* Γ *be a smooth manifold without boundary, and* $\widetilde{\mathbf{n}}$ *a* $\mathcal{C}^1$ *unitary extension of the normal vector* $\mathbf{n}$ *to a neighborhood of Gamma. Then, the mean curvature of* Γ *in* $\mathbf{s} \in \Gamma$ *is defined as*

$$H(\mathbf{s}) = \operatorname{div}_\Gamma \mathbf{n}(\mathbf{s}) = \operatorname{div} \widetilde{\mathbf{n}}(\mathbf{s}).$$

Lemma 1.11. *Let* Ω *be a* $\mathcal{C}^1$ *bounded domain in* $\mathbb{R}^d$, $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ *a displacement field, and* $g \in \mathcal{C}^1(\partial\Omega)$ *a real-valued continuous function. Then, the following identity holds*

$$\int_{\partial((\mathbf{I}+\boldsymbol{\theta})\Omega)} g \circ (\mathbf{I}+\boldsymbol{\theta})^{-1} \,\mathrm{d}\tilde{\mathbf{s}} = \int_{\partial\Omega} g \operatorname{Jac}_{\partial\Omega}((\mathbf{I}+\boldsymbol{\theta}))\,\mathrm{d}\mathbf{s}.$$

The tangential Jacobian is defined as

$$\operatorname{Jac}_{\partial\Omega}((\mathbf{I}+\boldsymbol{\theta})) = \left\|(\mathbf{I}+\boldsymbol{\theta})^{-\mathrm{T}}\mathbf{n}\right\| \left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|.$$

Moreover, the mapping $\boldsymbol{\theta} \mapsto \operatorname{Jac}_{\partial\Omega}((\mathbf{I}+\boldsymbol{\theta}))$ *is differentiable in* $\boldsymbol{\theta} = \mathbf{0}$*, and its derivative is*

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}} \operatorname{Jac}_{\partial\Omega}((\mathbf{I}+\boldsymbol{\theta}))(\widehat{\boldsymbol{\theta}}) = \operatorname{div}_\Gamma \widehat{\boldsymbol{\theta}}.$$

The proof of lemma 1.11 can be found in [138, Proposition 5.4.3 and Lemma 5.4.15].

Finally, we can state a proposition on the shape derivative of functionals defined as integrals on the boundary of a domain, reported as [6, Proposition 6.24] and [138, Proposition 5.4.18].

Proposition 1.12 (Shape derivative of a boundary functional)**.** *Let* Ω *be a* $\mathcal{C}^2$ *admissible domain, and* $g \in \mathrm{W}^{2,1}\left(\mathbb{R}^d\right)$ *a real valued function. Then, the shape functional* $\Phi : \Omega \mapsto \displaystyle\int_{\partial\Omega} g(\mathbf{s})\mathrm{d}\mathbf{s}$ *is shape differentiable in* Ω*, and its derivative is expressed as*

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_{\partial\Omega} \left(\nabla g \cdot \widehat{\boldsymbol{\theta}} + g \operatorname{div}_\Gamma \widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{s}. = \int_{\partial\Omega} \left(\mathbf{n}(\mathbf{s}) \cdot \widehat{\boldsymbol{\theta}}(\mathbf{s})\right) \left(\frac{\partial g}{\partial \mathbf{n}} + Hg\right) \mathrm{d}\mathbf{s}. \tag{1.22}$$

In order to be able to compute the shape derivative of more complex shape functional, it is necessary to introduce the concept of derivative of a function defined on a variable domain Ω. Let $u : \mathcal{S}_{\mathrm{adm}} \times \mathbb{R}^d \to \mathbb{R}$ be a function such that $u_\Omega(\mathbf{x})$ is well-defined if $\mathbf{x} \in \Omega$. As stated in [6, Section 6.3.3] and [112, Section 1.2.3], there exist two main definitions of derivative of $u_\Omega(\mathbf{x})$ with respect to the domain. Considering $\boldsymbol{\theta}$ as a transport vector field, the two derivatives are denoted Eulerian and Lagrangian derivatives, by analogy to the derivatives in continuum mechanics.

Definition 1.13 (Eulerian derivative)**.** *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be an admissible domain in $\mathbb{R}^d$ and $\boldsymbol{\theta} \in \Theta_{adm} \subset \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ such that $\Omega_{\boldsymbol{\theta}} = (\mathbf{I}+\boldsymbol{\theta})\Omega \in \mathcal{S}_{\mathrm{adm}}$. Let $\mathbf{x} \in \mathbb{R}^d$ such that $\mathbf{x}$ belongs to both Ω and $\Omega_{\boldsymbol{\theta}}$. The Eulerian derivative of u in $\mathbf{x}$ in the direction $\boldsymbol{\theta}$ is the differential of the mapping $\boldsymbol{\theta} \mapsto u_{\Omega_{\boldsymbol{\theta}}}(\mathbf{x})$ and is denoted*

$$u'_\Omega(\mathbf{x})(\widehat{\boldsymbol{\theta}}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}} u_{\Omega_{\boldsymbol{\theta}}}(\mathbf{x})\Big|_{\boldsymbol{\theta}=\mathbf{0}}(\widehat{\boldsymbol{\theta}}).$$

An important drawback of the Eulerian derivative is that the definition 1.13 makes sense only for $\mathbf{x}$ fixed. Moreover, if $\mathbf{x}$ belongs to the boundary $\partial\Omega$, it might not belong to $\Omega_{\boldsymbol{\theta}}$ for all displacement fields $\boldsymbol{\theta} \in \Theta_{adm}$.

Definition 1.14 (Lagrangian derivative)**.** *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be an admissible domain in $\mathbb{R}^d$. For $\boldsymbol{\theta} \in \Theta_{adm}$ we denote $u_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I}+\boldsymbol{\theta})$ the function defined on Ω obtained by pushing back $u_{\Omega_{\boldsymbol{\theta}}}$ to the original domain Ω. The Lagrangian derivative of u in $\mathbf{x}$ in the direction $\boldsymbol{\theta}$ is the differential of the mapping $\boldsymbol{\theta} \mapsto \mathbf{u}_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I}+\boldsymbol{\theta})(\mathbf{x})$ and is denoted*

$$\dot{u}_\Omega(\mathbf{x})(\widehat{\boldsymbol{\theta}}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}} u_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I}+\boldsymbol{\theta})(\mathbf{x})\Big|_{\boldsymbol{\theta}=\mathbf{0}}(\widehat{\boldsymbol{\theta}}).$$

In contrast to the Eulerian derivative $u'_\Omega(\mathbf{x})$, the Lagrangian derivative $\dot{u}_\Omega(\mathbf{x})$ is well-defined on the domain Ω regardless to the direction $\boldsymbol{\theta}$. Wherever they are both defined, the Lagrangian and Eulerian derivatives are related by the following identity

$$\dot{u}_\Omega(\mathbf{x})(\widehat{\boldsymbol{\theta}}) = u'_\Omega(\mathbf{x})(\widehat{\boldsymbol{\theta}}) + \nabla u_\Omega(\mathbf{x}) \cdot \widehat{\boldsymbol{\theta}}(\mathbf{x}). \tag{1.23}$$

Further analysis on the sensitivity of a shape functional can be done by computing its second derivative. We will not present any result on higher-order shape differentiation, and we suggest the reader to consult [138, Section 5.3] and [86] for further information on the subject.

1.2.2 Differentiation under elliptical PDE constraints

If the evaluation of the functional to be differentiated depends on the solution of a boundary-value problem, the expression of the shape derivative is not as simple as in proposition 1.8. In order to illustrate the procedure of computation of the shape derivative, we consider a simple shape optimization problem in the form of problem (1.3), where the state is the solution of an elliptic partial differential equation. We focus on the case where the state $\mathbf{u}_\Omega$ is a vector field, as for the linear elasticity equation. The analogous computation for the scalar case can be found in [112, Section 1.2.3].

Let $\mathcal{S}_{\mathrm{adm}}$ be a family of admissible domains in $\mathbb{R}^d$ with $d = 2$ or 3. We suppose that the boundary of any domain in $\mathcal{S}_{\mathrm{adm}}$ can be decomposed in three disjoint parts denoted Γ_{D}, Γ_{N}, and Γ_0. For the sake of simplicity, we suppose that all admissible domains share the portion Γ_{D} of their boundary. Let $\mathbf{A}$ be a positive definite matrix, and $\mathbf{f} \in \mathrm{H}^1(\Omega)^d$ and $\mathbf{g} \in \mathrm{L}^2(\Gamma_{\mathrm{N}})^d$. For all $\mathbf{s} \in \partial\Omega$, we denote $\mathbf{n}(\mathbf{s})$ the unitary vector which is orthogonal to the surface $\partial\Omega$ in $\mathbf{s}$. We suppose that the objective function J has the following integral expression:

$$J(\mathbf{u}, \Omega) = \int_\Omega \left(j_0(\mathbf{u}(\mathbf{x})) + j_1(\nabla\mathbf{u}(\mathbf{x}))\right) \mathrm{d}\mathbf{x} \qquad \text{for } \mathbf{u} \in \mathrm{H}^1(\Omega)^d, \tag{1.24}$$

where $j_0 : \mathbb{R} \to \mathbb{R}$ and $j_1 : \mathbb{R}^d \to \mathbb{R}$ are two $\mathcal{C}^1$ functions. Finally, we suppose that there exist two positive constants $C_0, C_1 > 0$ such that $|\partial_i j_0(\mathbf{v})| < C_0 \|\mathbf{v}\|_{\mathbb{R}^d}$ and $|\partial_i \partial_j j_1(\mathbf{V})| < C_1 \|\mathbf{V}\|_{\mathbb{R}^{d\times d}}$ for all $i, j \in \{1, \dots, d\}$ and any $\mathbf{v} \in \mathbb{R}^d$ and $\mathbf{V} \in \mathbb{R}^{d\times d}$.

We aim to solve the following shape optimization problem

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the objective function } J(\mathbf{u}_\Omega, \Omega), \\ \text{where the state } \mathbf{u}_\Omega \in \mathrm{H}^1(\Omega)^d \\ \text{is solution of the state problem} \\ \quad \left\{\begin{array}{rcll} -\mathrm{div}(\mathbf{A}\nabla\mathbf{u}_\Omega) & = & \mathbf{f} & \text{in } \Omega, \\ (\mathbf{A}\nabla\mathbf{u}_\Omega)\,\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ (\mathbf{A}\nabla\mathbf{u}_\Omega)\,\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \end{array}\right. \tag{1.25}$$

The weak formulation of the state problem in (1.25) is the following

$$\left|\begin{array}{l} \text{Find } \mathbf{u}_\Omega \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \text{ such that} \\ \text{for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \\ \displaystyle\int_\Omega \mathbf{A}\nabla\mathbf{w}_\Omega : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \mathbf{g}\cdot\mathbf{v}\,\mathrm{d}\mathbf{s}. \end{array}\right. \tag{1.26}$$

In order to use a gradient-based optimization method to solve problem (1.25), it is necessary to compute the shape derivative of the reduced objective functional $\widehat{J} : \Omega \mapsto J(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$. As first step, we compute the expression of the objective function for a deformed domain. Let Ω be a Lipschitz domain, and $\Theta_{adm} \subset \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ a family of Lipschitz continuous vector fields as

$$\Theta_{adm} = \left\{ \boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \ : \ \boldsymbol{\theta} = \mathbf{0} \text{ on } \Gamma_{\mathrm{D}} \right\}.$$

We consider a deformation field $\boldsymbol{\theta} \in \Theta_{adm}$ and a deformed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$, and we compute the value of $\widehat{J}(\Omega_{\boldsymbol{\theta}})$:

$$\widehat{J}(\Omega_{\boldsymbol{\theta}}) = J(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}, \Omega_{\boldsymbol{\theta}}) = \int_{\Omega_{\boldsymbol{\theta}}} \left(j_0(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}})) + j_1(\nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}))\right) \mathrm{d}\widetilde{\mathbf{x}}, \tag{1.27}$$

where $\nabla_{\widetilde{\mathbf{x}}}$ indicates the gradient operator with respect to the perturbed reference. By applying the change of variables $\mathbf{x} = (\mathbf{I} + \boldsymbol{\theta})^{-1}\widetilde{\mathbf{x}}$, the expression (1.27) can be reformulated as an integral on the reference domain Ω. We recall that, for any function $f \in \mathcal{C}^1\left(\mathbb{R}^d\right)$, the gradient in the perturbed reference of f can be expressed as

$$\nabla_{\widetilde{\mathbf{x}}} f(\widetilde{\mathbf{x}}) = (\mathbb{I} + \nabla\boldsymbol{\theta})^{-\mathrm{T}} \nabla\left(f \circ (\mathbf{I} + \boldsymbol{\theta})\right)(\mathbf{x}). \tag{1.28}$$

By consequence, the gradient of a vector function $\mathbf{f} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$ in the perturbed reference is

$$\nabla_{\widetilde{\mathbf{x}}} \mathbf{f}(\widetilde{\mathbf{x}}) = \nabla\left(\mathbf{f} \circ (\mathbf{I} + \boldsymbol{\theta})\right)(\mathbf{x}) \left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-1}. \tag{1.29}$$

We denote $\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega = \mathbf{u}_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I} + \boldsymbol{\theta})$ the solution of the state problem on the perturbed domain under the change of variables, and we remark that $\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega$ is well-defined on the reference domain Ω. Therefore, we can express (1.27) as

$$\widehat{J}(\Omega_{\boldsymbol{\theta}}) = \int_\Omega \left(j_0(\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega(\mathbf{x})) + j_1(\nabla\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega(\mathbf{x})(\mathbb{I} + \nabla\boldsymbol{\theta})^{-1})\right) \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \mathrm{d}\mathbf{x}. \tag{1.30}$$

The shape derivative of $\widehat{J}(\Omega)$ in the direction $\boldsymbol{\theta}$ is obtained differentiating equation (1.30) with respect to $\boldsymbol{\theta}$. By differentiating equation (1.30) with respect to $\boldsymbol{\theta}$ we obtain

$$\begin{aligned}\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) =& \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\widehat{J}(\Omega_{\boldsymbol{\theta}})\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = \int_\Omega \left(j_0(\mathbf{u}_\Omega(\mathbf{x})) + j_1(\nabla\mathbf{u}_\Omega(\mathbf{x}))\right)\operatorname{div}\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\\ &+\int_\Omega \left(\nabla j_0(\mathbf{u}_\Omega(\mathbf{x}))\cdot\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) + \nabla j_1(\nabla\mathbf{u}_\Omega(\mathbf{x})) : \left(\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) - \nabla\mathbf{u}_\Omega(\mathbf{x})\nabla\widehat{\boldsymbol{\theta}}\right)\right)\mathrm{d}\mathbf{x}.\end{aligned} \tag{1.31}$$

where $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ is the Lagrangian derivative of $\mathbf{u}_{\Omega,\mathbf{g}}$ in the direction $\widehat{\boldsymbol{\theta}}$, as introduced in definition 1.13. The terms $\nabla j_0(\mathbf{u}_\Omega(\mathbf{x})) \in \mathrm{H}^1(\Omega)^d$ and $\nabla j_1(\nabla\mathbf{u}_\Omega(\mathbf{x}))\mathrm{L}^2(\Omega)^{d\times d}$ are the dual operators of the derivatives of the functionals j_0 and j_1 in the respective function spaces. The equation defining the material derivative of the state can be found by differentiating the weak expression of the state problem of (1.25) on a perturbed domain with respect to $\boldsymbol{\theta}$.

$$\left|\begin{array}{l}\text{Find } \mathbf{u}_{\Omega_{\boldsymbol{\theta}}} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega_{\boldsymbol{\theta}}) \text{ such that}\\ \text{for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)\\ \displaystyle\int_{\Omega_{\boldsymbol{\theta}}} \left(\mathbf{A}\nabla_{\widetilde{\mathbf{x}}}\left(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}\right)(\widetilde{\mathbf{x}})\right) : \nabla_{\widetilde{\mathbf{x}}}\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}(\widetilde{\mathbf{x}})\,\mathrm{d}\widetilde{\mathbf{x}}\\ \qquad = \displaystyle\int_{\Omega_{\boldsymbol{\theta}}} \mathbf{f}(\widetilde{\mathbf{x}})\cdot\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}(\widetilde{\mathbf{x}})\,\mathrm{d}\widetilde{\mathbf{x}} + \int_{(\mathbf{I}+\boldsymbol{\theta})\Gamma_\mathrm{N}} \mathbf{g}(\widetilde{\mathbf{s}})\cdot\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}(\widetilde{\mathbf{s}})\,\mathrm{d}\widetilde{\mathbf{s}}.\end{array}\right. \tag{1.32}$$

Since the mapping $\mathbf{x}\mapsto\widetilde{\mathbf{x}} = (\mathbf{I}+\boldsymbol{\theta})\mathbf{x}$ is a diffeomorphism, it is possible to take as test function $\mathbf{v}$ defined on Ω. Expressing the equation in problem (1.32) in terms of the reference domain, we get that for all $\mathbf{v}\in\mathrm{H}^1(\Omega)$

$$\begin{aligned}&\int_\Omega \left((\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}}\mathbf{A}\nabla\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega(\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right) : \nabla\mathbf{v}(\mathbf{x})\,\left|\det(\mathbb{I}+\nabla\boldsymbol{\theta})\right|\,\mathrm{d}\mathbf{x}\\ &\quad = \int_\Omega \mathbf{f}\circ(\mathbf{I}+\boldsymbol{\theta})(\mathbf{x})\cdot\mathbf{v}(\mathbf{x})\,\left|\det(\mathbb{I}+\nabla\boldsymbol{\theta})\right|\,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \mathbf{g}\circ(\mathbf{I}+\boldsymbol{\theta})(\mathbf{s})\cdot\mathbf{v}(\mathbf{s})\mathrm{Jac}_{\partial\Omega}((\mathbf{I}+\boldsymbol{\theta}))\,\mathrm{d}\mathbf{s}.\end{aligned} \tag{1.33}$$

By differentiating (1.33) with respect to $\boldsymbol{\theta}$, we obtain the equation solved by the material derivative of the state

$$\begin{aligned}&\int_\Omega \left(\operatorname{div}(\widehat{\boldsymbol{\theta}})\,\mathbf{A}\nabla\mathbf{u}_\Omega - \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}\right) : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{A}\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x}\\ &\qquad = \int_\Omega \left(\nabla\mathbf{f}\,\widehat{\boldsymbol{\theta}} + \operatorname{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{v}(\mathbf{s})\,\mathrm{d}\mathbf{s}.\end{aligned} \tag{1.34}$$

Thus, the material derivative $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ is solution of the following problem

$$\left|\begin{array}{l}\text{Find } \dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \text{ such that}\\ \text{for all } \mathbf{v}\in\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)\\ \displaystyle\int_\Omega \mathbf{A}\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x} = \int_\Omega \left(-\operatorname{div}(\widehat{\boldsymbol{\theta}})\,\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \mathbf{A}\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}\right) : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x}\\ \qquad\qquad + \displaystyle\int_\Omega \left(\nabla\mathbf{f}\widehat{\boldsymbol{\theta}} + \operatorname{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{v}(\mathbf{s})\,\mathrm{d}\mathbf{s}.\end{array}\right. \tag{1.35}$$

The expression (1.31) for the shape derivative of the reduced functional is not suitable for a gradient-based optimization algorithm, since its evaluation in a direction $\widehat{\boldsymbol{\theta}}$ requires to

retrieve the material derivative of the state by solving problem (1.35). One method to avoid the computation of $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ consists in the computation of a suitable adjoint state (or co-state) $\mathbf{w}_\Omega \in \mathrm{H}^1(\Omega)^d$, solving the following adjoint problem

$$\left\{\begin{array}{rcll} -\mathrm{div}\left(\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right) &=& \nabla j_0(\mathbf{u}_\Omega) - \mathrm{div}\left(\nabla j_1(\nabla\mathbf{u}_\Omega)\right) & \text{in } \Omega, \\ \left(\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right)\mathbf{n} &=& \nabla\left(j_1(\nabla\mathbf{u}_\Omega)\right)\mathbf{n} & \text{on } \Gamma_{\mathrm{N}}\cup\Gamma_0, \\ \mathbf{w}_\Omega &=& 0 & \text{on } \Gamma_{\mathrm{D}}, \end{array}\right. \tag{1.36}$$

whose weak formulation is

$$\left|\begin{array}{l} \text{Find } \mathbf{w}_\Omega \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \text{ such that} \\ \text{for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \\ \displaystyle\int_\Omega \mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x} = \int_\Omega \nabla j_0(\mathbf{u}_\Omega)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_\Omega \nabla j_1(\nabla\mathbf{u}_\Omega) : \nabla\mathbf{v}\,\mathrm{d}\mathbf{x}. \end{array}\right. \tag{1.37}$$

The Lax-Milgram theorem ensures that, if the regularity and growth properties of j_0 and j_1 are satisfied, problem (1.37) is well-posed and the adjoint state $\mathbf{w}_\Omega$ is well-defined.

By considering $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ as test function in problem (1.37), recalling the expression of problem (1.35) and the fact that $\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega : \nabla\mathbf{v} = \mathbf{A}\nabla\mathbf{v} : \nabla\mathbf{w}_\Omega$ by definition of the inner product, we get the following identity

$$\begin{aligned} &\int_\Omega \nabla j_0(\mathbf{u}_\Omega)\cdot\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} + \int_\Omega \nabla j_1(\nabla\mathbf{u}_\Omega) : \nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega : \nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} \\ =&\int_\Omega \mathbf{A}\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{w}_\Omega\,\mathrm{d}\mathbf{x} = \int_\Omega \left(\nabla\mathbf{f}\widehat{\boldsymbol{\theta}} + \mathrm{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{w}_\Omega\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{w}_\Omega\,\mathrm{d}\mathbf{s} \\ &+ \int_\Omega \left(-\mathrm{div}(\widehat{\boldsymbol{\theta}})\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \mathbf{A}\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}\right) : \nabla\mathbf{w}_\Omega\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{1.38}$$

When injecting equation (1.38) into equation (1.31), we get the following expression for the derivative of the reduced functional

$$\begin{aligned} \mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) =& \int_\Omega \left(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega)\right)\mathrm{div}(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} - \int_\Omega \nabla j_1(\nabla\mathbf{u}_\Omega) : \left(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}\right)\,\mathrm{d}\mathbf{x} \\ &+ \int_\Omega \left(\nabla\mathbf{f}\ \widehat{\boldsymbol{\theta}} + \mathrm{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{w}_\Omega\ \mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \left(\widehat{\boldsymbol{\theta}}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s})\right)\left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{w}_\Omega\,\mathrm{d}\mathbf{s} \\ &+ \int_\Omega \left(-\mathrm{div}(\widehat{\boldsymbol{\theta}})\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \mathbf{A}\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}\right) : \nabla\mathbf{w}_\Omega\,\mathrm{d}\mathbf{x} \\ =& \int_\Omega \left(\nabla\mathbf{f}^{\mathrm{T}}\mathbf{w}_\Omega\right)\cdot\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{w}_\Omega\,\mathrm{d}\mathbf{s} \\ &+ \int_\Omega \left(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega) + \mathbf{f}\cdot\mathbf{w}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega : \nabla\mathbf{w}_\Omega\right)\mathrm{div}(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} \\ &+ \int_\Omega \left(-\nabla\mathbf{u}_\Omega{}^{\mathrm{T}}\nabla j_1(\nabla\mathbf{u}_\Omega) + \nabla\mathbf{w}_\Omega{}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla\mathbf{u}_\Omega{}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right) : \nabla\widehat{\boldsymbol{\theta}}. \end{aligned} \tag{1.39}$$

Formula (1.39) shows the volume expression of the shape derivative of $\widehat{J}$. If Ω is a $\mathcal{C}^1$ domain, theorem 1.7 ensures that the derivative in the direction $\widehat{\boldsymbol{\theta}}$ can be expressed as an integral on $\partial\Omega$ depending on the normal component of $\widehat{\boldsymbol{\theta}}$ in each point of the surface. In order to compute such expression, we consider the following lemma.

Lemma 1.15. *Let $\Omega \subset \mathbb{R}^d$ be a $\mathcal{C}^1$ domain, $\phi \in \mathcal{C}^1(\Omega)$ and $\mathbf{M} : \Omega \to \mathbb{R}^{d\times d}$ two continuous functions, and $\boldsymbol{\theta} : \mathcal{C}^1(\Omega)^d$ a smooth displacement vector field. For each point $\mathbf{s}$ on $\partial\Omega$ we suppose that $\boldsymbol{\theta}(\mathbf{s})$ can be decomposed as $\boldsymbol{\theta}(\mathbf{s}) = \mathbf{n}(\mathbf{s})\,(\boldsymbol{\theta}\cdot\mathbf{n}) + \boldsymbol{\theta}_\tau$, where $\mathbf{n}(\mathbf{s})$ is the unit vector orthogonal to $\partial\Omega$ in $\mathbf{s}$, and $\boldsymbol{\theta}_\tau$ is the component of $\boldsymbol{\theta}$ tangent to $\partial\Omega$.*

Then, the following identities hold

$$\int_\Omega \phi(\mathbf{x})\ \mathrm{div}\,\boldsymbol{\theta}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} \phi(\mathbf{s})\,(\boldsymbol{\theta}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s}))\ \mathrm{d}\mathbf{s} - \int_\Omega \nabla\phi(\mathbf{x})\cdot\boldsymbol{\theta}(\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{1.40}$$

$$\int_\Omega \mathbf{M}(\mathbf{x}) : \nabla\boldsymbol{\theta}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} \left(\mathbf{n}^{\mathrm{T}}\mathbf{M}\mathbf{n}\right)(\boldsymbol{\theta}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s}))\ \mathrm{d}\mathbf{s} + \int_{\partial\Omega} \boldsymbol{\theta}_\tau^{\mathrm{T}}\mathbf{M}\mathbf{n}\,\mathrm{d}\mathbf{s} - \int_\Omega \mathrm{div}(\mathbf{M}(\mathbf{x}))\cdot\boldsymbol{\theta}(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{1.41}$$

Proof. The proof of the identity (1.41) derives directly from the formula of the divergence of the product of two functions, and from the application of the divergence theorem. Indeed

$$\int_{\partial\Omega} \phi(\mathbf{s})\,(\boldsymbol{\theta}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s}))\ \mathrm{d}\mathbf{s} = \int_\Omega \mathrm{div}\,(\phi\boldsymbol{\theta})\ \mathrm{d}\mathbf{x} = \int_\Omega (\phi(\mathbf{x})\ \mathrm{div}\,\boldsymbol{\theta}(\mathbf{x}) + \nabla\phi(\mathbf{x})\cdot\boldsymbol{\theta}(\mathbf{x}))\ \mathrm{d}\mathbf{x}.$$

In order to prove equation (1.41) we consider the identity $\mathrm{div}(\mathbf{M}\boldsymbol{\theta}) = \mathbf{M} : \nabla\boldsymbol{\theta} + \mathrm{div}(\mathbf{M})\cdot\boldsymbol{\theta}$, and we procede as for (1.40).

$$\int_{\partial\Omega} \mathbf{n}(\mathbf{s})\cdot(\mathbf{M}(\mathbf{s})\boldsymbol{\theta}(\mathbf{s}))\ \mathrm{d}\mathbf{s} = \int_\Omega \mathrm{div}(\mathbf{M}\boldsymbol{\theta})\,\mathrm{d}\mathbf{x} = \int_\Omega (\mathbf{M}(\mathbf{x}) : \nabla\boldsymbol{\theta}(\mathbf{x}) + \mathrm{div}(\mathbf{M})\cdot\boldsymbol{\theta}(\mathbf{x}))\ \mathrm{d}\mathbf{x}. \tag{1.42}$$

Equation (1.41) is obtained by applying the decomposition $\boldsymbol{\theta}(\mathbf{s}) = \mathbf{n}(\mathbf{s})\,(\boldsymbol{\theta}\cdot\mathbf{n}) + \boldsymbol{\theta}_\tau$ to the left-hand side of equation (1.42). □

The surface form of $\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}})$ can be found integrating equation (1.39) by parts. Indeed, using equations (1.40) and (1.41) we find

$$\begin{aligned}
\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) = &\int_{\partial\Omega} \Big(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega) + \mathbf{f}\cdot\mathbf{w}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega : \mathbf{w}_\Omega\Big)\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}\\
&+ \int_\Omega \Big(\nabla\mathbf{f}^{\mathrm{T}}\mathbf{w}_\Omega - \nabla\big(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega) - \mathbf{A}\nabla\mathbf{u}_\Omega : \nabla\mathbf{w}_\Omega\big)\Big)\cdot\widehat{\boldsymbol{\theta}}(\mathbf{x})\,\mathrm{d}\mathbf{x}\\
&+ \int_{\partial\Omega} \mathbf{n}^{\mathrm{T}}\left(-\nabla{\mathbf{u}_\Omega}^{\mathrm{T}}(\nabla\, j_1(\nabla\mathbf{u}_\Omega)) + \nabla{\mathbf{w}_\Omega}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla{\mathbf{u}_\Omega}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right)\mathbf{n}\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}\\
&+ \int_{\partial\Omega} \widehat{\boldsymbol{\theta}}_\tau^{\mathrm{T}}(\mathbf{s})\left(-\nabla{\mathbf{u}_\Omega}^{\mathrm{T}}(\nabla\, j_1(\nabla\mathbf{u}_\Omega)) + \nabla{\mathbf{w}_\Omega}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla{\mathbf{u}_\Omega}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right)\mathbf{n}(\mathbf{s})\,\mathrm{d}\mathbf{s}\\
&- \int_\Omega \mathrm{div}\left(-\nabla{\mathbf{u}_\Omega}^{\mathrm{T}}\nabla j_1(\nabla\mathbf{u}_\Omega) + \nabla{\mathbf{w}_\Omega}^{\mathrm{T}}\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla{\mathbf{u}_\Omega}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right)\cdot\widehat{\boldsymbol{\theta}}(\mathbf{x})\,\mathrm{d}\mathbf{x}\\
&+ \int_{\Gamma_{\mathrm{N}}} \left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\right)\cdot\mathbf{w}_\Omega\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}.
\end{aligned} \tag{1.43}$$

Thanks to theorem 1.7, the second, fourth, and fifth integrals appearing on the right-hand side of (1.43) are equal to zero. Therefore, assuming that $\widehat{\boldsymbol{\theta}} = \mathbf{0}$ on the portion Γ_{D} of the boundary, and considering the boundary conditions of problems (1.25) and (1.36),the shape

derivative of the reduced objective functional can be reduced to

$$\begin{aligned}
\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) = & \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \Big(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega) + \mathbf{f}\cdot\mathbf{w}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega : \nabla\mathbf{w}_\Omega\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \mathbf{n}^\mathrm{T}\Big(\nabla\mathbf{w}_\Omega{}^\mathrm{T}\mathbf{A}\nabla\mathbf{u}_\Omega + \nabla\mathbf{u}_\Omega{}^\mathrm{T}\mathbf{A}^\mathrm{T}\nabla\mathbf{w}_\Omega\Big)\mathbf{n}\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \mathbf{n}^\mathrm{T}\Big(-\nabla\mathbf{u}_\Omega{}^\mathrm{T} - \nabla\, j_1(\nabla\mathbf{u}_\Omega))\Big)\mathbf{n}\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_\mathrm{N}} \Big(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\Big)\cdot\mathbf{w}_\Omega\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
= & \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \Big(\mathbf{f}\cdot\mathbf{w}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega : \mathbf{w}_\Omega - \frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\cdot\frac{\partial j_1}{\partial\mathbf{n}}(\nabla\mathbf{u}_{\Omega,\mathbf{g}})\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \Big(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega)\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} + \int_{\Gamma_\mathrm{N}} \Big(\frac{\partial\mathbf{g}}{\partial\mathbf{n}} + H\mathbf{g}\Big)\cdot\mathbf{w}_\Omega\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \Big(\Big(\frac{\partial\mathbf{w}_\Omega}{\partial\mathbf{n}}\Big)\cdot\Big(\mathbf{A}\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\Big) + \Big(\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\Big)\cdot\Big(\mathbf{A}^\mathrm{T}\Big(\mathbf{w}_\Omega - \frac{\partial j_1}{\partial\mathbf{n}}(\nabla\mathbf{u}_{\Omega,\mathbf{g}})\Big)\mathbf{n}\Big)\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\mathrm{d}\mathbf{s} \\
= & \int_{\Gamma_0\cup\Gamma_\mathrm{N}} \Big(j_0(\mathbf{u}_\Omega) + j_1(\nabla\mathbf{u}_\Omega) + \mathbf{f}\cdot\mathbf{w}_\Omega - \mathbf{A}\nabla\mathbf{u}_\Omega : \nabla\mathbf{w}_\Omega\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s} \\
& + \int_{\Gamma_\mathrm{N}} \Big(\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\cdot\mathbf{w}_\Omega + \frac{\partial\mathbf{w}_\Omega}{\partial\mathbf{n}}\cdot\mathbf{g} + H(\mathbf{g}\cdot\mathbf{w}_\Omega)\Big)\Big(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\Big)\,\mathrm{d}\mathbf{s}.
\end{aligned} \tag{1.44}$$

We remark that, contrarily to equation (1.31), the term $\widehat{\boldsymbol{\theta}}$ appears explicitly in the expressions (1.39) and (1.44) of the shape derivative, since the direction of differentiation does not appear in the definition of the adjoint state in (1.36). Moreover, the expressions (1.39) and (1.44) display clearly the linearity of the function $\mathrm{D}\widehat{J}(\Omega)$, in accordance with definition 1.5.

1.2.3 Céa's fast derivation technique

The procedure detailed in section 1.2.2 allows to rigorously derive the expression of the shape derivative of a shape functional under PDEs constraints. However, the computation of the problem solved by the material derivative and the integrations by parts to obtain a surface expression of the shape derivative can be cumbersome for more complex state problems or objective functionals.

In practice, a faster method developed by Céa in [61] is often used instead. In this section we illustrate Cea's fast derivation method as presented in [6, Section 6.4.3] and [112, Section 1.2.3]. Such technique allows to deduce the equation solved by the adjoint state without having to compute the material derivative of the state, and it provides directly the surface expression of the shape derivative. Nevertheless, this method is purely formal, and does not ensure the existence of the derivative of a shape functional.

Let us consider, once again, the shape optimization problem (1.25). We are interested in computing the shape derivative of $\widehat{J}(\Omega) = J(\mathbf{u}_\Omega, \Omega)$, where $\mathbf{u}_\Omega$ solves the state problem in (1.25) and J is defined as in (1.24).

We define the Lagrangian function $\mathcal{L} : \mathcal{S}_\mathrm{adm} \times \mathrm{H}^1\left(\mathbb{R}^d\right)^d \times \mathrm{H}^1\left(\mathbb{R}^d\right)^d \times \mathrm{H}^1\left(\mathbb{R}^d\right)^d$ such that

$$\begin{aligned}
\mathcal{L}\left(\Omega, \widehat{\mathbf{u}}, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}}\right) = & \int_\Omega \left(j_0(\widehat{\mathbf{u}}(\mathbf{x})) + j_1(\nabla\widehat{\mathbf{u}}(\mathbf{x}))\right)\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{A}\nabla\widehat{\mathbf{u}}(\mathbf{x}) : \nabla\widehat{\mathbf{w}}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
& + \int_\Omega \mathbf{f}(\mathbf{x})\cdot\widehat{\mathbf{w}}(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \mathbf{g}(\mathbf{s})\cdot\widehat{\mathbf{w}}(\mathbf{s})\,\mathrm{d}\mathbf{s} + \int_{\Gamma_\mathrm{D}} \widehat{\boldsymbol{\lambda}}(\mathbf{s})\cdot\widehat{\mathbf{u}}(\mathbf{s})\,\mathrm{d}\mathbf{s}.
\end{aligned} \tag{1.45}$$

We remark that all arguments of the Lagrangian $\mathcal{L}$ are independent from one another, in particular from the shape Ω of the domain. The terms $\widehat{\mathbf{w}}$ and $\widehat{\boldsymbol{\lambda}}$ act as Lagrange multipliers for the PDE constraint and for the Dirichlet boundary condition on Γ_{D}. At first, we remark that the partial derivative $\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{w}}}$ evaluated in $\widehat{\mathbf{u}} = \mathbf{u}_\Omega$ is equal to zero for any choice of $\widehat{\mathbf{w}}$ and $\widehat{\boldsymbol{\lambda}}$. Indeed, by the week formulation of problem (1.26) we have

$$\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{w}}}(\Omega, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}})(\mathbf{v}) = -\int_\Omega \mathbf{A}\nabla \mathbf{u}_\Omega(\mathbf{x}) : \nabla \mathbf{v}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0,$$

and, for the boundary condition on $\mathbf{u}_\Omega$

$$\frac{\partial \mathcal{L}}{\partial \widehat{\boldsymbol{\lambda}}}(\Omega, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}})(\widetilde{\boldsymbol{\lambda}}) = \int_{\Gamma_{\mathrm{D}}} \widehat{\boldsymbol{\lambda}}(\mathbf{s}) \cdot \mathbf{u}_\Omega(\mathbf{s})\,\mathrm{d}\mathbf{s} = 0.$$

Therefore, for any choice of $\widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}} \in \mathrm{H}^1\left(\mathbb{R}^d\right)^d$

$$\widehat{J}(\Omega) = J(\mathbf{u}_\Omega, \Omega) = \mathcal{L}\left(\Omega, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}}\right), \tag{1.46}$$

and

$$\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\mathcal{L}\left(\Omega_{\boldsymbol{\theta}}, \mathbf{u}_{\Omega_{\boldsymbol{\theta}}}, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}}\right)(\widehat{\boldsymbol{\theta}}) = \frac{\partial}{\partial \boldsymbol{\theta}}\mathcal{L}\left(\Omega_{\boldsymbol{\theta}}, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}}\right)(\widehat{\boldsymbol{\theta}}) + \frac{\partial}{\partial \widehat{\mathbf{u}}}\mathcal{L}\left(\Omega, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}}\right)(\mathbf{u}_\Omega{}'(\widehat{\boldsymbol{\theta}})) \tag{1.47}$$

where $\mathbf{u}_\Omega{}'(\widehat{\boldsymbol{\theta}}) \in \mathrm{L}^2(\Omega)^d$ is the Eulerian derivative of the displacement field as introduced in definition 1.13.

The equation for the adjoint state derives from a suitable choice of $\widehat{\mathbf{w}}$ and $\widehat{\boldsymbol{\lambda}}$ that cancels $\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{u}}}$ when evaluated in $\widehat{\mathbf{u}} = \mathbf{u}_\Omega$. Differentiating the Lagrangian and integrating by parts, we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{u}}}(\Omega, \mathbf{u}_\Omega, \widehat{\mathbf{w}}, \widehat{\boldsymbol{\lambda}})(\mathbf{v}) =& \int_\Omega \left(\nabla j_0(\mathbf{u}_\Omega(\mathbf{x})) \cdot \mathbf{v} + \nabla j_1(\nabla \mathbf{u}_\Omega(\mathbf{x})) \cdot \nabla \mathbf{v}\right)\,\mathrm{d}\mathbf{x} \\
&+ \int_{\Gamma_{\mathrm{D}}} \widehat{\boldsymbol{\lambda}}(\mathbf{s}) \cdot \mathbf{v}(\mathbf{s})\,\mathrm{d}\mathbf{s} - \int_\Omega \mathbf{A}\nabla \mathbf{v}(\mathbf{x}) : \nabla \mathbf{w}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
=& \int_\Omega \left(\nabla j_0(\mathbf{u}_\Omega) + \mathrm{div}(\mathbf{A}^{\mathrm{T}}\nabla \widehat{\mathbf{w}}) - \mathrm{div}(\nabla j_1(\nabla \mathbf{u}_\Omega))\right) \cdot \mathbf{v}\,\mathrm{d}\mathbf{x} \\
&+ \int_{\Gamma_{\mathrm{N}} \cup \Gamma_0} \left(\nabla j_1(\nabla \mathbf{u}_\Omega) - \mathbf{A}^{\mathrm{T}}\nabla \widehat{\mathbf{w}}\right)\mathbf{n}(\mathbf{s}) \cdot \mathbf{v}(\mathbf{s})\mathrm{d}\mathbf{s} \\
&+ \int_{\Gamma_{\mathrm{D}}} \left(\left(\nabla j_1(\nabla \mathbf{u}_\Omega) - \mathbf{A}^{\mathrm{T}}\nabla \widehat{\mathbf{w}}\right)\mathbf{n}(\mathbf{s}) + \widehat{\boldsymbol{\lambda}}\right) \cdot \mathbf{v}(\mathbf{s})\mathrm{d}\mathbf{s}.
\end{aligned} \tag{1.48}$$

From equation (1.48) we deduce that the values of $\widehat{\mathbf{w}}$ and $\widehat{\boldsymbol{\lambda}}$ such that the derivative $\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{w}}}$ vanishes are $\mathbf{w}_\Omega$ solution of equation (1.36), and any $\boldsymbol{\lambda}_\Omega \in \mathrm{H}^1\left(\mathbb{R}^d\right)^d$ such that

$$\boldsymbol{\lambda}_\Omega = \left(\mathbf{A}^{\mathrm{T}}\nabla \mathbf{w}_\Omega - \nabla j_1(\nabla \mathbf{u}_\Omega)\right)\mathbf{n} \qquad \text{on } \Gamma_{\mathrm{D}}.$$

By evaluating the Lagrangian in $\widehat{\mathbf{w}} = \mathbf{w}_\Omega$ and $\widehat{\boldsymbol{\lambda}}_\Omega$, the partial derivative $\frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{u}}}$ in equation (1.47) vanishes. Since, for tha sake of simplicity, we assume that all admissible shapes in $\mathcal{S}_{\mathrm{adm}}$ share the portion Γ_{D} of their boundary, we restrict the field of admissible deformation fields to

$$\Theta_{adm} = \left\{\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \ : \ \boldsymbol{\theta} = \mathbf{0} \text{ on } \Gamma_{\mathrm{D}}\right\}.$$

By applying proposition 1.8 and considering only deformations $\widehat{\boldsymbol{\theta}} \in \Theta_{adm}$, the shape derivative of $\widehat{J}$ in Ω becomes

$$\begin{aligned}\mathrm{D}\widehat{J}(\Omega)(\widehat{\boldsymbol{\theta}}) =& \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{L}\left(\Omega_{\boldsymbol{\theta}}, \mathbf{u}_\Omega, \mathbf{w}_\Omega, \boldsymbol{\lambda}_\Omega\right)(\widehat{\boldsymbol{\theta}}) = \int_{\Gamma_0 \cup \Gamma_\mathrm{N}} \left(j_0(\mathbf{u}_\Omega(\mathbf{x})) + j_1(\nabla \mathbf{u}_\Omega(\mathbf{x}))\right)(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{x} \\ &+ \int_{\Gamma_0 \cup \Gamma_\mathrm{N}} \left(\mathbf{f}(\mathbf{x}) \cdot \mathbf{w}_\Omega(\mathbf{x}) - \mathbf{A}\nabla \mathbf{u}_\Omega(\mathbf{x}) : \nabla \mathbf{w}_\Omega(\mathbf{x})\right)(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{x} \\ &+ \int_{\Gamma_\mathrm{N}} \left(\frac{\partial \mathbf{w}}{\partial \mathbf{n}} \cdot \mathbf{g} + \frac{\partial \mathbf{g}}{\partial \mathbf{n}} \cdot \mathbf{w} + H(\mathbf{g} \cdot \mathbf{w})\right)(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{x}.\end{aligned} \tag{1.49}$$

We remark that Céa's fast differentiation method yields the same surface expression for the shape derivative as equation (1.44). However, this method is purely formal and assumes that the state $\mathbf{u}_\Omega$ and the objective functional are regular enough to allow for the definition of the Lagrangian on the appropriate spaces as well as the Eulerian derivative.

1.3 A gradient-based algorithm for shape optimization

1.3.1 Level-set method and signed distance function

In order to solve numerically a shape optimization problem, it is necessary to represent and discretize the structure at each step of the optimization algorithm, and to be able to deform it in accordance to a given deformation field $\boldsymbol{\theta} : \mathbb{R}^d \to \mathbb{R}^d$.

One possible approach consists in representing the structure explicitly by an unstructured mesh and advect its vertices according to the deformation $\boldsymbol{\theta}$ (see [198, Chapter 7] or [184]). Let $\mathcal{T}_\Omega$ be a mesh discretizing the shape Ω. The mesh $\mathcal{T}_{\Omega_{\boldsymbol{\theta}}}$ discretizing $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$ would share its connectivity with $\mathcal{T}_\Omega$, and its vertices $\mathbf{x}'_1, \ldots, \mathbf{x}'_N$ would be obtained by the displacement of the nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$ of $\mathcal{T}_\Omega$ by $\boldsymbol{\theta}$

$$\mathbf{x}'_i = (\mathbb{I} + \boldsymbol{\theta}(\mathbf{x}_i))\mathbf{x}_i \qquad \text{for all nodes } \mathbf{x}_i \text{ of } \mathcal{T}_\Omega.$$

The main benefit of this approach is the direct access to the discretized domain. In particular, the solution of boundary value problems on Ω is made significantly easier since the boundary regions are explicitly defined, and the PDE problem can be solved numerically on $\mathcal{T}_\Omega$ by the Finite Elements Method. Despite its simplicity, this technique is prone to the degradation of the mesh over few advection steps ([91, Section 3.2] and [41]), and it requires frequent remeshing in order to avoid the presence of ill-shaped elements. Moreover, a mesh deformation technique would not be able to deal with topological changes in the shape during the optimization process, since it would require the creation or deletion of some edges or elements around the boundary. Different approaches address these issues, like the partition of the mesh displacement into multiple steps alternated with frequent partial remeshings [36], or the *Deformable Simplicial Complex* (DSC) [182, 72, 71]. It is worth to remark that, contrarily to the simple nodal displacement method, the DSC performs frequent remeshings and allows for topological changes in the structure. Another approach consists in representing the density field by a suitable Non-Uniform Rational Basis Spline (NURBS) as proposed in [185, 49], defining a natural sharp boundary between the interior and the exterior of the structure.

A different technique relies on the introduction of a *level-set* function which provides an implicit representation of the structure. The level-set method has first been introduced in the domain of fluid mechanics [193], but has proven its versatility for shape optimization in several

different contexts (see e.g. [230, 194, 253, 19] for early applications). Refer to [247] for a comprehensive review of the level-set method, and [13, 98, 84] as examples of applications in shape optimization.

We consider the notations adopted in [13, 248, 112, 101].

Definition 1.16 (Level-set function)**.** *A level-set function representing the open domain Ω is any real-valued continuous function ϕ_Ω such that*

$$\begin{cases} \phi_\Omega(\mathbf{x}) < 0 & \text{if} \quad \mathbf{x} \in \Omega, \\ \phi_\Omega(\mathbf{x}) = 0 & \text{if} \quad \mathbf{x} \in \partial\Omega, \\ \phi_\Omega(\mathbf{x}) > 0 & \text{if} \quad \mathbf{x} \in \mathbb{R}^d \setminus \overline{\Omega}. \end{cases}$$

The function ϕ_Ω encodes implicitly the shape Ω in $\mathbb{R}^d$n and the level-set representation allows to account for variable domains. Let $\boldsymbol{\theta}$ be a time-dependent vector field that we suppose to be Lipschitz continuous in space at each instant $t \in [0,T]$ and such that $\|\boldsymbol{\theta}(t)\|_{1,\infty} < 1$. We suppose also that the mapping $t \mapsto \boldsymbol{\theta}(t) \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ is differentiable, and we denote $\mathbf{V}(t,\mathbf{x})$ the Eulerian derivative of the vector field $\boldsymbol{\theta}$ at the time $t \in [0,T]$ in the position $\mathbf{x} \in \mathbb{R}^d$. Let $\Omega_{\boldsymbol{\theta}(t)} = (\mathbf{I} + \boldsymbol{\theta}(t))\,\Omega$ be the domain obtained by applying the deformation field $\boldsymbol{\theta}(t)$ to the original domain Ω, and $\phi_{\Omega_{\boldsymbol{\theta}(t)}}$ the corresponding level-set function. If the velocity field $\mathbf{V}$ is known in $[0,T] \times \mathbb{R}^d$, the function $\phi_{\Omega_{\boldsymbol{\theta}(t)}}$ can be computed by solving the following advection equation

$$\begin{cases} \frac{\partial}{\partial t}\phi_{\Omega_{\boldsymbol{\theta}(t)}} + \mathbf{V}(t,\mathbf{x}) \cdot \nabla \phi_{\Omega_{\boldsymbol{\theta}(t)}} = 0 & \text{in } (0,T] \times \mathbb{R}^d, \\ \phi_{\Omega_{\boldsymbol{\theta}(0)}}(\mathbf{x}) = \phi_\Omega(\mathbf{x}) & \text{in } \mathbb{R}^d. \end{cases} \tag{1.50}$$

If $\Omega_{\boldsymbol{\theta}(t)}$ is of class $\mathcal{C}^1$, it is possible to define a normal vector $\mathbf{n}(t,\mathbf{s})$ for each $t \in [0,T]$ and $\mathbf{s} \in \partial\Omega_{\boldsymbol{\theta}(t)}$. Thus, if there exists a scalar function $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that the advection field can be written as $\mathbf{V}(t,\mathbf{s}) = v(t,\mathbf{s})\mathbf{n}(t,\mathbf{s})$ for all $t \in [0,T]$ and $\mathbf{s} \in \partial\Omega_{\boldsymbol{\theta}(t)}$, the equation (1.50) can be replaced by the Hamilton-Jacobi equation

$$\begin{cases} \frac{\partial}{\partial t}\phi_{\Omega_{\boldsymbol{\theta}(t)}} + v(t,\mathbf{x}) \left\|\nabla \phi_{\Omega_{\boldsymbol{\theta}(t)}}\right\|_{\mathrm{L}^2} = 0 & \text{in } (0,T] \times \mathbb{R}^d, \\ \phi_{\Omega_{\boldsymbol{\theta}(0)}}(\mathbf{x}) = \phi_\Omega(\mathbf{x}) & \text{in } \mathbb{R}^d. \end{cases} \tag{1.51}$$

For numerical applications, the level-set function is considered to be defined on a bounded computational domain $D \subset \mathbb{R}^d$. The library `advection`[1] [56], part of the *ISCD toolbox*[2], is the main tool used in the present thesis for the solution of the advection equation (1.50) on an unstructured mesh. Different numerical approaches to the solution of the advection equation are provided in [150, 62]. As remarked in [56, 101], the problems (1.50) and (1.51) are not well-posed on a bounded domain without a re-entrant boundary condition to account for the values of the level-set function outside D. For further information on the boundary conditions for the Hamilton-Jacobi equation we refer to [78, 40].

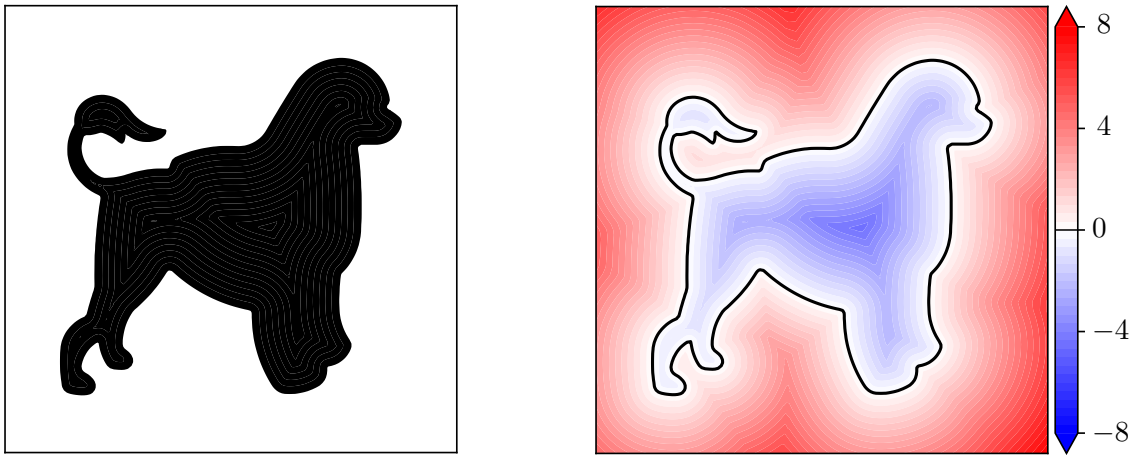
As pointed out in [192, 70], multiple solutions of the advection equation can result in a level-set function that is locally too flat, which can cause problems in numerical applications. Hence, there is a necessity to update the level-set ϕ_Ω every few iterations in order to preserve its quality. Given a domain $\Omega \in \mathbb{R}^d$, there exist infinite level-set functions encoding the shape Ω. A function of particular significance in the context of shape optimization is the signed distance function [228, 18, 90, 113, 94].

[1] `https://github.com/ISCDtoolbox/Advection`
[2] `https://github.com/ISCDtoolbox`

Definition 1.17 (Signed distance function)**.** *For an open domain $\Omega \subset \mathbb{R}^d$, the signed distance function $d_\Omega : \mathbb{R}^d \to \mathbb{R}$ is a real-valued continuous function such that*

$$d_\Omega(\mathbf{x}) = \begin{cases} \inf_{\mathbf{s}\in\partial\Omega} \|\mathbf{x}-\mathbf{s}\|_{\mathbb{R}^d} & \text{if} \quad \mathbf{x} \notin \Omega, \\ -\inf_{\mathbf{s}\in\partial\Omega} \|\mathbf{x}-\mathbf{s}\|_{\mathbb{R}^d} & \text{if} \quad \mathbf{x} \in \Omega. \end{cases} \tag{1.52}$$

At first, we remark that the function d_Ω is a level-set function for Ω according to 1.16 since $d_\Omega(\mathbf{s}) = 0$ if $\mathbf{s} \in \partial\Omega$. Moreover, the signed distance function is 1-Lipschitz continuous [100, 92] and, thanks to Rademacher's theorem [110, Chapter 5], it is differentiable almost everywhere. Let us denote Σ_Ω the skeleton of Ω, that is the set of points in $\mathbb{R}^d$ where the minimization problem $\min_{\mathbf{s}\in\partial\Omega} \|\mathbf{x}-\mathbf{s}\|_{\mathbb{R}^d}$ admits multiple minimizers. For any point $\mathbf{x} \in \mathbb{R}^d \setminus \Sigma_\Omega$, we denote $p_\Omega(\mathbf{x}) \in \partial\Omega$ the projection of $\mathbf{x}$ onto $\partial\Omega$, that is the $\arg\min_{\mathbf{s}\in\partial\Omega} \|\mathbf{x}-\mathbf{s}\|_{\mathbb{R}^d}$. An example of the signed distance function representing a bounded domain in $\mathbb{R}^2$ is reported in fig. 1.1.



(a) Domain Ω included in a square box D.

(b) Representation of the corresponding signed distance function d_Ω. The isoline $d_\Omega(\mathbf{x}) = 0$ is marked in black.

Figure 1.1: Representation of the signed distance function for a domain Ω in a computational domain $D \subset \mathbb{R}^2$. The signed distance function d_Ω has been computed numerically using the library `mshdist`.

For regular domains, it is possible to characterize the signed distance function by the following result [112, Proposition 1.10].

Proposition 1.18. *Let $\Omega \subset \mathbb{R}^d$ be a $\mathcal{C}^1$ domain. Then, for all $\mathbf{x} \in \mathbb{R}^d \setminus \Sigma_\Omega$, the signed distance function d_Ω is differentiable in $\mathbf{x}$, and its gradient is equal to*

$$\nabla d_\Omega(\mathbf{x}) = \mathbf{n}(p_\Omega(\mathbf{x})).$$

Moreover, d_Ω satisfies the eikonal equation

$$\begin{cases} \|\nabla d_\Omega(\mathbf{x})\|_{\mathbb{R}^d} = 1 & \text{for all } \mathbf{x} \in \mathbb{R}^d \setminus \Sigma_\Omega, \\ d_\Omega(\mathbf{s}) = 0 & \text{for all } \mathbf{s} \in \partial\Omega. \end{cases} \tag{1.53}$$

Proposition 1.18 highlights two important properties of the signed distance suncion as level-set for shape optimization. At first, it provides an easy method to compute an extension of the normal vector $\mathbf{n}$ to the entire space $\mathbb{R}^d$. Secondly, since the norm of the gradient of the signed distance is equal to 1 almost everywhere, taking d_Ω as level-set ensures that it is never too flat to cause numerical issues.

It is clear that, if an advection field $\boldsymbol{\theta}$ is applied to a signed distance function d_Ω, the resulting mapping would be a level-set for the perturbed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$, but it would not be a signed distance anymore. Thus, it is necessary to recompute the signed distance function after any perturbation of the domain in order to preserve the properties proven in proposition 1.18.

In the literature, two main families of approaches are proposed for the numerical computation of the signed distance function. The first group, which includes the Fast Marching Method [228, 229, 119] and the Fast Sweeping Method [263, 202, 35] aim to solve directly the stationary eikonal equation (1.53) and rely on the information on the distances between the different elements of the mesh. The second group ,which includes the techniques proposed in [235, 92], relies on the following result on the unsteady eikonal equation (see [92, Theorem 2.5], proven in [32, Section 2.2.2]).

Theorem 1.19. *Let $\overline{\phi_\Omega} : \mathbb{R}^d \to \mathbb{R}$ be a level-set function for an open domain $\Omega \subset \mathbb{R}^d$. We define the unsteady eikonal equation, also called the reinitialization equation [32] or redistanciation equation [92, 112], as the following Hamilton-Jacobi PDE:*

$$\left\{ \begin{array}{rcll} \frac{\partial \varphi}{\partial t}(t,\mathbf{x}) + \operatorname{sign}(\overline{\phi_\Omega}(\mathbf{x}))\left(\|\nabla \varphi(t,\mathbf{x})\|_{\mathbb{R}^d} - 1\right) & = & 0 & \text{for all } t > 0, \mathbf{x} \in \mathbb{R}^d; \\ \varphi(0,\mathbf{x}) & = & \overline{\phi_\Omega}(\mathbf{x}) & \text{for all } \mathbf{x} \in \mathbb{R}^d. \end{array} \right. \tag{1.54}$$

Then, the only viscosity solution of (1.54) is $\varphi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$ defined as

$$\varphi(t,\mathbf{x}) = \left\{ \begin{array}{ll} \operatorname{sign}(\overline{\phi_\Omega}(\mathbf{x})) \inf_{\|\mathbf{y}\| \leq t} \left(\operatorname{sign}(\overline{\phi_\Omega}(\mathbf{x}))\overline{\phi_\Omega}(\mathbf{x}+\mathbf{y}) + t \right) & \text{if} \quad 0 \leq t \leq |d_\Omega(\mathbf{x})|\,, \\ d_\Omega(\mathbf{x}) & \text{if} \quad t > |d_\Omega(\mathbf{x})|\,. \end{array} \right. \tag{1.55}$$

Moreover, the signed distance function is the steady-state solution of equation (1.54).

In this work we used the open-source library `mshdist`[3] developed as part of the *ISCD toolbox* and detailed in [92]. The `mshdist` software relies on the solution of the unsteady eikonal equation (1.54) to compute the signed distance function of domains discretized by body-fitted meshes. Refer to [32, 100, 92] for further information on the eikonal equation for the identification of the signed distance function.

1.3.2 Implicit and explicit shape representation

Let $\mathcal{S}_{\text{adm}}$ be a class of domains in $\mathbb{R}^d$, Ω an open domain in $\mathcal{S}_{\text{adm}}$, and D a sufficiently large subset of $\mathbb{R}^d$ such that all elements of $\mathcal{S}_{\text{adm}}$ are contained in D. We suppose also that all elements of $\mathcal{S}_{\text{adm}}$ share the portion Γ_{D} of their boundaries, and that Γ_{D} is included in ∂D. Let us consider a shape optimization problem in the form (1.25). As we remarked in section 1.3.1, if Ω is explicitly discretized by a mesh $\mathcal{T}_\Omega$, a suitable numerical method can be applied to solve the state and adjoint problems on $\mathcal{T}_\Omega$. Clearly, this consideration is not limited to elliptical problems, but can be extended to any kind of well-posed PDE constraint.

The solution of the state equation is more subtle if the domain Ω is encoded implicitly by a level-set function ϕ_Ω defined on D. Let $\mathcal{T}_D$ be a fixed simplicial mesh covering the computational

[3] `https://github.com/ISCDtoolbox/Mshdist`

box D. The numerical solution of the state problem is complicated by the fact that no mesh covering exclusively Ω is readily available. The *ersatz material* approach is among the early approaches to solve this issue [19], and consists in replacing the state problem defined in Ω with the following relaxation on the entire computational box D.

$$\left\{\begin{array}{rcll} -\mathrm{div}(\mathbf{A}_\varepsilon(\phi_\Omega(\mathbf{x}))\nabla\mathbf{u}_\varepsilon) &=& \mathbf{f}(\mathbf{x})\chi_\Omega(\mathbf{x}) & \text{in } D, \\ (\mathbf{A}_\varepsilon(\phi_\Omega(\mathbf{x}))\nabla\mathbf{u}_\varepsilon)\,\mathbf{n} &=& \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ (\mathbf{A}_\varepsilon(\phi_\Omega(\mathbf{x}))\nabla\mathbf{u}_\varepsilon)\,\mathbf{n} &=& \mathbf{0} & \text{on } \partial D \setminus (\Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}), \\ \mathbf{u} &=& \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \tag{1.56}$$

where χ_Ω is the characteristic function of the domain Ω. The operator $\mathbf{A}_\varepsilon : \mathbb{R} \to \mathbb{R}^{d\times d}$ in (1.56) is defined as

$$\mathbf{A}_\varepsilon = \varepsilon\mathbf{A} + (1-\varepsilon)\mathbf{A}\chi_{\mathbb{R}^+}, \tag{1.57}$$

where $\varepsilon \ll 1$ is a positive small parameter. A relaxed problem analogous to (1.56) can be formulated for the adjont problem (1.36).

Equation (1.56) can be interpreted as an approximation of the state problem, where Ω is surrounded by a light *ersatz material* approximating the void in $D \setminus \Omega$ [84]. The strict positivity of ε is necessary to assure the well-posedness of the relaxed state problem, but it introduces an error in the computation of the state. As proven in [84, Theorem 1.8], the error on the solution of the state problem can be bounded by

$$\|\mathbf{u}_\Omega - \mathbf{u}_\varepsilon\|_{\mathrm{H}^1(\Omega)} \leq C(\Omega, D, \mathbf{f}, \mathbf{g})\,\varepsilon, \tag{1.58}$$

where $C(\Omega, D, \mathbf{f}, \mathbf{g})$ is a constant depending only on the shape of the domain and the computation box, and on the volume and surface loads, and $\mathbf{u}_\Omega$ is the exact solution of the state problem.

In numerical applications, the value of ε is subject to a trade-off. Equation (1.58) encourages to decrease its value, in order to compute a better approximation of $\mathbf{u}_\Omega$. However, a too small value for ε can degrade the conditioning number of the stiffness matrix in the finite element system, and make its numerical solution more difficult to achieve.

As highlighted in [101, Section 2.2], the error between $\mathbf{u}_\varepsilon$ and the exact solution $\mathbf{u}_\Omega$ is not the only difficulty introduced by the ersatz material technique. Indeed, since Ω is included into the computational box D, a finite element problem formulated on $\mathcal{T}_D$ is computationally more expensive than one defined on a smaller mesh covering only Ω because it requires the computation of a fictitious solution far outside Ω. Moreover, if ϕ_Ω is discretized as a $\mathbb{P}^1$ function, the elliptic operator $\mathbf{A}_\varepsilon(\phi_\Omega(\mathbf{x}))$ assumes an intermediate value between $\mathbf{A}$ and $\varepsilon\mathbf{A}$ in the vicinity of the isoline $\phi_\Omega(\mathbf{x}) = 0$, which corresponds to the boundary of Ω. This last issue implies that the error on the numerical solution of (1.56) can be significantly larger than the one on the continuous solution $\mathbf{u}_\varepsilon$. This problem is particularly relevant in the context of shape optimization since the surface expression of the shape derivative of a differentiable functional depends on the value of the state and its adjoint on the free boundary Γ_0. The *immersed interface* approach [254, 230] does not require any modification of the fixed mesh $\mathcal{T}_D$, and deals with the transition between the interior and the exterior of Ω by suitable jump conditions. The *LEVITY* algorithm for shape optimization [77] relies on the ersatz material approximation (1.57) and on the definition of a metric on the computational box D as function of the level-set function and calls for a mesh adaptation step only when a criterion on the convergence of the objective function.

A different strategy addressing all these issues consists in replacing the fixed mesh $\mathcal{T}_D$ by a simplicial mesh $\mathcal{T}_{D,\Omega}$ where Ω is explicitly discretized as a submesh $\mathcal{T}_\Omega$. Multiple approaches

have been proposed in order to extract an explicit discretization $\mathcal{T}_{D,\Omega}$ of Ω from a mesh $\mathcal{T}_D$ and a level-set function ϕ_Ω.

The *XFEM* method [107, 249] consists in cutting the elements of $\mathcal{T}_D$ crossed by ϕ_Ω, obtaining an explicit representation of Ω. An important drawback of this approach is the fact that cutting the elements of $\mathcal{T}_D$ may give rise to thin, low-quality elements in $\mathcal{T}_{D,\Omega}$ around the isoline $\phi_\Omega(\mathbf{x}) = 0$, affecting the numerical solution of the state and adjoint problems. The *CutFEM* method [59, 250] addresses the low-quality issue of the mesh near the boundary resulting from the XFEM method by refining the mesh around the interface between the interior and the exterior of Ω.

In this thesis we considered the approach proposed in [88, 13, 89], which relies on a local remeshing of $\mathcal{T}_D$ around the isoline $\phi_\Omega(\mathbf{x}) = 0$ every time the level-set function is advected, in order to construct a body-fitted mesh for the domain Ω. Let $\mathcal{T}_{D,\Omega}$ be a mesh for D where a domain Ω is explicitly represented as a submesh, and let $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}(D)^d$ a Lipschitz continuous vector field. We suppose that the domain Ω is represented by a level-set function ϕ_Ω discretized on $\mathcal{T}_{D,\Omega}$. We aim to obtain a mesh $\mathcal{T}_{D,\Omega_{\boldsymbol{\theta}}}$ where the perturbed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$ is explicitly discretized. We start by solving the advection equation (1.50) in order to compute a level set function $\widetilde{\phi_{\Omega_{\boldsymbol{\theta}}}}$ for the perturbed domain on the mesh $\mathcal{T}_{D,\Omega_{\boldsymbol{\theta}}}$. Then, a local remeshing is performed around the isoline $\phi_{\Omega_{\boldsymbol{\theta}}}(\mathbf{x}) = 0$, producing a body-fitted mesh $\mathcal{T}_{D,\Omega_{\boldsymbol{\theta}}}$ covering the entire computational box D. An example of the advection and remeshing process is presented in fig. 1.2.

The local remeshing is performed using the implicit domain meshing option of the open-source `mmg` platform[4] [89, 37], in 2 dimensions (`mmg2d_O3`) and 3 dimensions (`mmg3d_O3`). The implicit domain meshing of `mmg` relies on the isovalue discretization algorithm, which takes as input an unstructured simplicial mesh $\mathcal{T}_D$ and a $\mathbb{P}^1$ level-set function ϕ_Ω defined on the nodes of $\mathcal{T}_D$, and returns a simplicial mesh $\mathcal{T}_{D,\Omega}$ where the domain Ω is explicitly represented by a submesh $\mathcal{T}_\Omega$. The execution of the isovalue discretization algorithm can be adjusted by the following options:

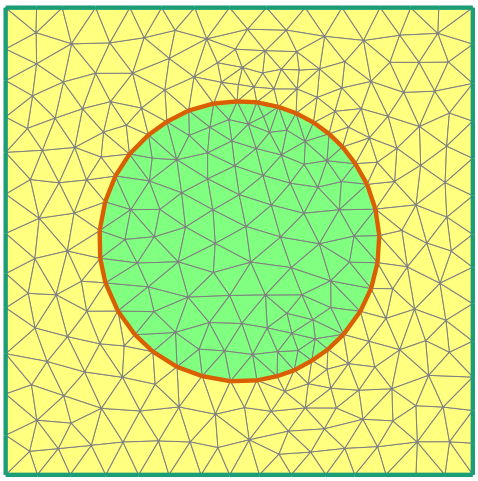
- `hmin` and `hmax` are real values defining the minimal and maximal authorized length of an edge in the mesh;
- `hausd` defines an upper bound on the Hausdorff distance between an edge (in 2D) or a triangle (in 3D), and the ideal continuous boundary $\partial\Omega$, in order to assure a good representation of the domain Ω;
- `hgrad` is a real value, larger than 1 (typically around 1.4), that controls the ratio between adjacent elements of $\mathcal{T}_{D,\Omega}$, in order to avoid flat elements, and assure the overall quality of the mesh.

1.3.3 Shape gradient and the Hilbertian regularization-extension

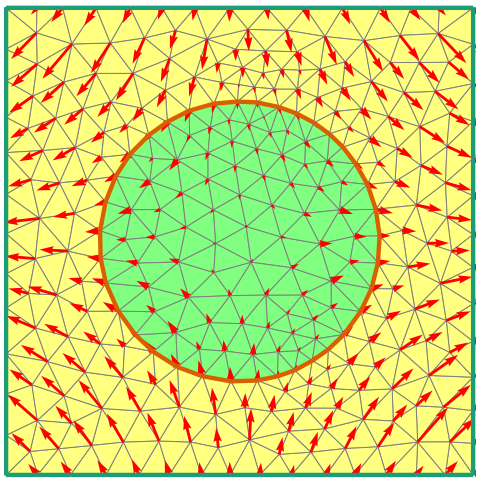
In order to apply a gradient-based algorithm to shape optimization problems, it is worth recalling the definition of gradient in Hilbert spaces in analogy with [114]. Let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ be a Hilbert space, and $J : \mathcal{H} \to \mathbb{R}$ a function that is Fréchet differentiable in $x \in \mathcal{H}$, and its derivative is denoted as $\mathrm{D}J(x)(\cdot)$. The derivative of J in x is a continuous linear functional defined on $\mathcal{H}$. Thus, $\mathrm{D}J(x)$ belongs to the dual space of $\mathcal{H}$

$$\mathrm{D}J(x) \in \mathcal{H}^* = \{L : \mathcal{H} \to \mathbb{R} \ : \ L \text{ linear and bounded}\}.$$

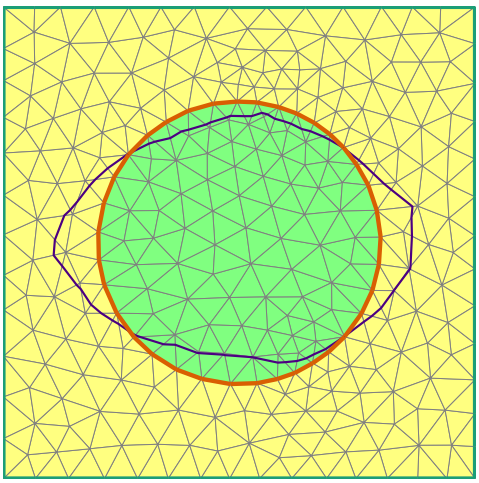
[4] `http://www.mmgtools.org`



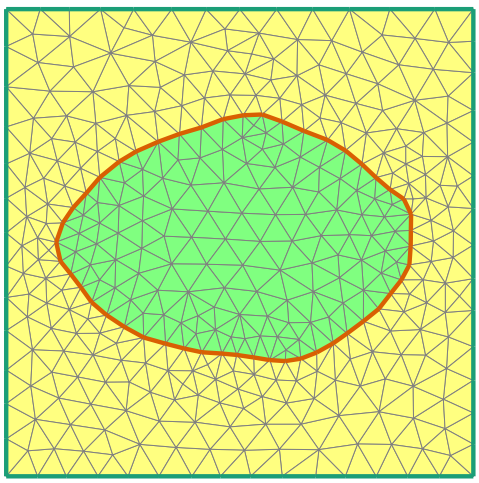
(a) Mesh $\mathcal{T}_{D,\Omega}$ fitting the circle Ω defined by a level-set $\phi_{\Omega_{\boldsymbol{\theta}}}$.



(b) Definition of a displacement field $\boldsymbol{\theta}$ for the level-set $\phi_{\Omega_{\boldsymbol{\theta}}}$



(c) Advection of ϕ_Ω by the field $\boldsymbol{\theta}$ to the new level-set function $\widetilde{\phi_{\Omega_{\boldsymbol{\theta}}}}$



(d) Body-fitting mesh $\mathcal{T}_{D,\Omega_{\boldsymbol{\theta}}}$ for the shape defined by the advected level-set.

Figure 1.2: Level-set advection and mesh adaptation for a domain Ω and a given displacement field $\boldsymbol{\theta}$ using `mshdist`, `advect`, and `mmg`.

An element $y \in \mathcal{H}$ is said to be a direction of descent for J in x if $\mathrm{D}J(x)(y) < 0$.

The gradient of J is an element of the Hilbert space $\mathcal{H}$, uniquely identified by Riesz's identification theorem as the only element $\nabla J(x)$ of $\mathcal{H}$ such that

$$\langle \nabla J(x), y \rangle_{\mathcal{H}} = \mathrm{D}J(x)(y) \qquad \text{for all } y \in \mathcal{H}. \tag{1.59}$$

We suppose that x is not a stationary point, and thus $\mathrm{D}J(x)(\cdot)$ is not identically equal to 0. An evident direction of descent is given by $-\tau \nabla J(x)$ for any $\tau > 0$, since

$$\mathrm{D}J(x)(-\tau \nabla J(x)) = -\tau \langle \nabla J(x), \nabla J(x) \rangle_{\mathcal{H}} = -\tau \|\nabla J(x)\|_{\mathcal{H}}^2 < 0.$$

Similar concepts can be introduced for smooth functions defined on a Riemannian manifold [104, 160]. Let $\mathcal{M}$ be a differential manifold endowed with the Riemannian metric g. We recall

that a Riemannian metric on a differential manifold is a mapping that associates any point $x \in \mathcal{M}$ with an inner product $\langle \cdot, \cdot \rangle_{g(x)}$ defined on the tangent space $T_x\mathcal{M}$ [104, Definition 2.1]. Thus, if J is a function defined on $\mathcal{M}$ which is differentiable in x, its derivative $\mathrm{D}J(x)$ belongs to the dual space $(T_x\mathcal{M})^*$, and its gradient $\nabla J(x)$ is the only element of $T_x\mathcal{M}$ such that

$$\langle \nabla J(x), y \rangle_{g(x)} = \mathrm{D}J(x)(y) \qquad \text{for all } y \in T_x\mathcal{M}. \tag{1.60}$$

We suppose that the derivative of J does not vanish in x. Then, any element of $T_x\mathcal{M}$ in the form $-\tau\nabla J(x)$ is a direction of descent for J in $x \in \mathcal{M}$, if $\tau > 0$.

In the context of shape optimization introduced in section 1.2, the space of definition of a shape functional $\widehat{J}$ is a class of open Lipschitz continuous domains $\mathcal{S}_{\mathrm{adm}}$. We suppose that there exists an open bounded domain $D \subset \mathbb{R}^d$ containing all admissible domains. The class $\mathcal{S}_{\mathrm{adm}}$ is not even a vector space, and thus it cannot be provided with a Hilbertian structure. The idea proposed in [223] and expanded in [224, 114] consists in considering $\mathcal{S}_{\mathrm{adm}}$ as a smooth manifold such that, for any domain $\Omega \in \mathcal{S}_{\mathrm{adm}}$, the tangent space at Ω coincides with $\mathrm{W}^{1,\infty}(D)^d$. Therefore, the shape derivative of $\widehat{J}$ in Ω as introduced in definition 1.5 belongs to the dual of the tangent space of $\mathcal{M}$ in Ω. However, $\mathrm{W}^{1,\infty}(D)^d$ is not a Hilbert space, and the gradient corresponding to a shape derivative cannot be immediately identified as in equation (1.60).

The solution proposed in [114] consists in identifying a suitable Hilbert space $\mathcal{H} \subset \mathrm{W}^{1,\infty}(D)^d$ endowed with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. By limiting the set of admissible deformation fields to $\mathcal{H}$, the restriction of $\mathrm{D}\widehat{J}(\Omega)$ to $\mathcal{H}$ is associated to a unique gradient $\nabla\widehat{J}(\Omega) \in \mathcal{H}$ such that equation (1.59) holds for any $y \in \mathcal{H}$. The choice of the Hilbert space is not unique, as remarked in [58, 98, 114]. One possibility is to identify $\mathcal{H}$ with the Sobolev space $\mathrm{H}^k(D)$ with the standard inner product, with $k > 1 + \frac{d}{2}$. Indeed, thanks to the Sobolev embedding theorem [2, Theorem 4.12, part I], $\mathrm{H}^k(D)$ is embedded into $\mathrm{W}^{1,\infty}(D)^d$.

A more common choice in literature consists in considering the space $\mathrm{H}^1(D)^d$, with the inner product

$$\langle \cdot, \cdot \rangle_\gamma : (\boldsymbol{\theta}, \boldsymbol{\xi}) \mapsto \langle \boldsymbol{\theta}, \boldsymbol{\xi} \rangle_\gamma = \int_D \left(\boldsymbol{\theta} \cdot \boldsymbol{\xi} + \gamma^2 \nabla\boldsymbol{\theta} : \nabla\boldsymbol{\xi} \right) \mathrm{d}\mathbf{x},$$

where the constant γ is a positive regularization parameter (see [19, 11, 224, 15, 113]). In numerical application, γ is usually taken as three times the minimal length of the edges of the mesh covering D [114].

The space $\mathrm{H}^1(D)^d$ is not a subspace of $\mathrm{W}^{1,\infty}(D)^d$. However, if Ω and $\widehat{J}$ are regular enough to satisfy the hypotheses of theorem 1.7, the shape derivative of $\widehat{J}$ can be written as

$$\mathrm{D}\widehat{J}(\Omega) : \widehat{\boldsymbol{\theta}} \mapsto \int_{\partial\Omega} v_{\widehat{J}} (\widehat{\boldsymbol{\theta}} \cdot \mathbf{n}), \tag{1.61}$$

for some $v_{\widehat{J}} \in \mathrm{L}^2(\partial\Omega)$, which is a continuous linear form on $\mathrm{H}^1(\Omega)$ by the trace theorem [218, Theorem 7.85]. Therefore, the shape gradient $\nabla\widehat{J}(\Omega)$ if the only element of $\mathrm{H}^1(D)^d$ such that, for all $\boldsymbol{\theta} \in \mathrm{H}^1(D)^d$,

$$\left\langle \nabla\widehat{J}(\Omega), \boldsymbol{\theta} \right\rangle_\gamma = \int_{\partial\Omega} v_{\widehat{J}} \, (\boldsymbol{\theta} \cdot \mathbf{n}) \, \mathrm{d}\mathbf{s}. \tag{1.62}$$

The use of the shape gradient defined as in (1.62) is convenient to solve a few numerical difficulties. Firstly, the gradient $\nabla\widehat{J}(\Omega)$ is not bounded to $\partial\Omega$, but is defined on D, extending the vector field $v_{\widehat{J}}\mathbf{n}$ to the entire computational box. Secondly, the shape derivative as defined in (1.61) can be affected by regularity issues [33, 19, 98, 101] that may compromise the numerical solution, while the gradient is smooth by definition. One cause of the irregularity in the shape

derivative can be the presence of sharp corners in the domain Ω. Another source of irregularity is linked to the numerical computation of $v_{\widehat{J}}$. Indeed, let us consider $\widehat{J}$ to be a PDE-constrained functional as in (1.24), which depends on a state $\mathbf{u}_\Omega$ solving the state problem (1.26). If $\mathcal{T}_D$ is a mesh covering the computational box D and $\mathbf{u}_\Omega$ and the corresponding adjoint state $\mathbf{w}_\Omega$ are discretized as $\mathbb{P}^1$, then normal velocity $v_{\widehat{J}}$ in the shape derivative holds only a $\mathbb{P}^0$ regularity [101, Subsection 2.2.4]. Moreover, as remarked in [98], the vector velocity field $v_{\widehat{J}}\mathbf{n}$ is even less regular than the normal velocity in presence of two separate boundaries tending to merge. The use of a H^1 shape gradient $\nabla\widehat{J}(\Omega)$ avoids the regularity issues and ensures better stability in the numerical solution [198]. Thanks to these two aforementioned properties, this method to compute a shape gradient is known in literature as the Hilbertian regularization-extension technique [98, 112, 14, 101] or smoothing [58].

A final advantage of the Hilbertian regularization-extension method is the possibility to enforce non-optimizable boundaries in the shape optimization process. Let us consider that all admissible shapes share a portion $\overline{\Gamma}$ of their boundary, or that there exists a subdomain $\overline{\Omega}$ contained in all admissible shapes in $\mathcal{S}_{\mathrm{adm}}$, [91]. In order to enforce the fact that $\overline{\Gamma}$ and $\overline{\Omega}$ are not altered by the advection field, it is sufficient to consider the Hilbert space $\mathcal{H}_{\overline{\Gamma},\overline{\Omega}} \subset \mathcal{H}$ such that

$$\mathcal{H}_{\overline{\Gamma},\overline{\Omega}} = \left\{ \boldsymbol{\theta} \in \mathcal{H} \ : \ \boldsymbol{\theta} = 0 \text{ on } \overline{\Gamma} \cup \overline{\Omega} \right\}.$$

It should be remarked that the computation of the shape gradient depends on the choice of the Hilbert space $\mathcal{H}$ extending $\mathrm{W}^{1,\infty}(D)^d$. However, regardless of the choice of $\mathcal{H}$, the vector field $-\nabla\widehat{J}(\Omega)$ is always a direction of descent for $\widehat{J}$ in Ω.

In conclusion, a gradient-based optimization algorithm for the solution of the shape optimization problem (1.25) can be summarized as in algorithm 1.

Algorithm 1 Procedure for the solution of the shape optimization problem (1.25) by a gradient-based method.

Require: a simplicial mesh $\mathcal{T}_D$ for the computational box D
Require: a level-set ϕ_{Ω^0} representing the initial domain $\Omega^0 \subset D$ on $\mathcal{T}_D$
  Compute a body-fitted mesh $\mathcal{T}_{D,\Omega^0}$
  Compute the signed distance function d_{Ω^0}
  Solve the state equation for $\mathbf{u}_{\Omega^0}$
  Compute the value of the objective functional $J(\mathbf{u}_{\Omega^0}, \Omega^0)$
  while $n < n_{\max}$ **or** a given convergence criterion is satisfied **do**
    Solve the adjoint problem for $\mathbf{w}_{\Omega^n}$
    Compute the shape derivative $\mathrm{D}\widehat{J}(\Omega^n)$
    Compute the shape gradient $\nabla\widehat{J}(\Omega^n)$ by a Hilbertian regularization-extension procedure
    Define a deformation field $\boldsymbol{\theta}^n = -\tau^n \nabla\widehat{J}(\Omega^n)$, where τ^n is a small discretization step
    Advect d_{Ω^n} by the vector field $\boldsymbol{\theta}^n$ in order to obtain a new level-set $\phi_{\Omega^{n+1}}$
    Compute a body-fitted mesh $\mathcal{T}_{D,\Omega^{n+1}}$
    Compute the signed distance function $d_{\Omega^{n+1}}$
    Solve the state equation for $\mathbf{u}_{\Omega^{n+1}}$
    Compute the value of the objective functional $J(\mathbf{u}_{\Omega^{n+1}}, \Omega^{n+1})$
    $n \leftarrow n + 1$
  end while

1.3.4 Null space optimization algorithm

Thus far we considered shape optimization problems where the shape Ω is subject to no explicit constraint, and the state is defined by one or more partial differential equations. In practice it is often necessary to complete the optimization problem (1.25) with one or more constraints.

Multiple approaches have been proposed for the solution of constrained optimization problems in the form of problem (1.5) outside the scope of shape optimization. Among the gradient-based algorithms we mention the augmented Lagrangian [51], penalty methods [117, Section 12.1], interior point algorithms [200], the method of moving asymptotes [240], and techniques based on an Ordinary Differential Equation (ODE) like Yamashita's method [258], and its offshoots [221, 39, 114]. In this thesis we focus on the *null space* optimization approach first introduced by Feppon, Allaire and Dapogny in [114].

Equality constraints

In order to understand the *null space optimization* algorithm, we start by considering a generic optimization problem in the form (1.5), where the control variable x belongs to a subset $\mathcal{X}_{\text{adm}}$ of the Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$.

The main idea behind ODE based methods consists in the definition of a dynamical system on $\mathcal{H}\times[0,+\infty)$, whose solution converges to the solution of the underlying optimization problem for $t\to\infty$. Yamashita's gradient flow technique [258] has been developed for optimization problems featuring only equality constraints. Inequality constraints can be taken into account by slack variables [221], or by suitable adaptation of the dynamical system. In particular, in this section we present the *null space* optimization algorithm as presented by Feppon, Allaire et Dapogny in [114] and its python implementation[5].

At first, let us consider an optimization problem featuring only $N_{\mathbf{G}}$ equality constraints

$$\left|\begin{array}{l}\text{Find } x\in\mathcal{H}\\ \text{minimizing the objective function } J(x)\\ \text{under the constraints}\\ \qquad \mathbf{G}(x)=\mathbf{0}.\end{array}\right. \tag{1.63}$$

The function $\mathbf{G}:\mathcal{H}\to\mathbb{R}^{N_{\mathbf{G}}}$ encodes $N_{\mathbf{G}}$ equality constraints, and we suppose that the objective and the constraint functions to be differentiable in the entire space $\mathcal{H}$. We denote $\mathrm{D}J(x)$ and $\mathrm{D}\mathbf{G}(x)$ the respective sensitivities, and $\nabla J(x)$ and $\mathrm{D}\mathbf{G}(x)^{\mathcal{T}}$ their transpose with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$.

The vector $\mathbf{G}(x)$ satisfies the *Linear Independence Constraint Qualification* (LICQ) condition in $x\in\mathcal{H}$ if $\mathrm{D}\mathbf{G}(x)$ has full rank (or, equivalently, if $\mathrm{D}\mathbf{G}(x)\,\mathrm{D}\mathbf{G}(x)^{\mathcal{T}}$ is a full-rank $N_{\mathbf{G}}\times N_{\mathbf{G}}$ matrix). Let x_0 be a point in the Hilbert space $\mathcal{H}$. The gradient flow ODE associated with the optimization problem (1.63) is

$$\left\{\begin{array}{rcll}\frac{\mathrm{d}}{\mathrm{d}x}t(t) &=& -\alpha_J\boldsymbol{\xi}_J(x)-\alpha_G\boldsymbol{\xi}_G(x), & \text{for } t\in(0,+\infty)\\ x(0) &=& x_0,\end{array}\right. \tag{1.64}$$

where α_J and α_G are positive parameters. The terms $\boldsymbol{\xi}_J(x)$ and $\boldsymbol{\xi}_G(x)$ belong to $\mathcal{H}$ and are defined as the *null space* and *range space* directions respectively. It is not necessary for x_0 to

[5] `https://gitlab.com/florian.feppon/null-space-optimizer`

satisfy the equality constraints $\mathbf{G}(x) = \mathbf{0}$. For any point $x \in \mathcal{H}$ that satisfies the LICQ condition, the null space and range space directions are defined as

$$\begin{aligned} \boldsymbol{\xi}_J(x) &= \left(\mathbb{I} - \mathrm{D}\mathbf{G}^{\mathcal{T}}(\mathrm{D}\mathbf{G}\ \mathrm{D}\mathbf{G}^{\mathcal{T}})^{-1}\mathrm{D}\mathbf{G}\right)\nabla J(x), \\ \boldsymbol{\xi}_G(x) &= \mathrm{D}\mathbf{G}^{\mathcal{T}}(\mathrm{D}\mathbf{G}\ \mathrm{D}\mathbf{G}^{\mathcal{T}})^{-1}\mathbf{G}(x). \end{aligned} \tag{1.65}$$

As proven in [114, Section 2], $-\boldsymbol{\xi}_G(x)$ is a descent direction for the vector of the violations of the equality constraints, and it belongs to the range space of $\mathrm{D}\mathbf{G}(x)$ for all $x \in \mathcal{H}$ respecting the LICQ condition. Moreover, the null space direction $\boldsymbol{\xi}_J(x)$ is the least-square approximation of $\nabla J(x)$ onto $\mathrm{Ker}(\mathrm{D}\mathbf{G}(x))$, and can be seen as

$$\boldsymbol{\xi}_J(x) = \nabla J(x) + \mathrm{D}\mathbf{G}^{\mathcal{T}}\boldsymbol{\lambda}(x).$$

We denoted $\boldsymbol{\lambda}(x) \in \mathbb{R}^{N_G}$ the vector of the Lagrange multipliers for the orthogonality constraint between $\boldsymbol{\xi}_J$ and the range space of $\mathrm{D}\mathbf{G}$, and it solves the optimization problem

$$\boldsymbol{\lambda}(x) = \underset{\boldsymbol{\lambda}\in\mathbb{R}^{N_G}}{\arg\min} \left\| \nabla J(x) + \mathrm{D}\mathbf{G}^{\mathcal{T}}\boldsymbol{\lambda} \right\|_{\mathcal{H}}. \tag{1.66}$$

Problem (1.66) can be solved numerically by dedicated solvers like `cvxopt`[6][24], `OSQP`[7][234], or `IPOPT`[8] [252]. The aforementioned two properties ensure that the vector of violation of the constraints $\mathbf{G}(x)$ tends to $\mathbf{0}$ along the trajectories of (1.64) and that, when $\mathbf{G}(x)$ is small enough, the objective function $J(x)$ decreases without affecting the decrease of $\mathbf{G}(x)$. The proofs are provided in [258] and [114, Annex A].

Inequality constraints

The handling of the inequality constraints in the *null space* optimization algorithm is more complex. We consider the generic optimization problem

$$\left|\begin{array}{l} \text{Find } x \in \mathcal{X}_{\text{adm}} \\ \text{minimizing the objective function } J(x) \\ \text{under the constraints} \\ \qquad \mathbf{G}(x) = \mathbf{0}, \\ \qquad \mathbf{H}(x) \leq \mathbf{0}. \end{array}\right. \tag{1.67}$$

simplifying the notation of problem (1.5). The functions $\mathbf{G}$ and $\mathbf{H}$ encode respectively $N_{\mathbf{G}}$ equality constraints and $N_{\mathbf{H}}$ inequality constraints, and are supposed to be differentiable in the entire space $\mathcal{H}$.

At first, we number the equality constraints with indices from 1 to $N_{\mathbf{G}}$, and the inequality constraints from $N_{\mathbf{G}} + 1$ to $N_{\mathbf{G}} + N_{\mathbf{H}}$. Let $I = \{1, \ldots, N_{\mathbf{G}} + N_{\mathbf{H}}\}$ be the set of indices of all constraints, and $\mathbf{C} = [\mathbf{G}^{\mathrm{T}}, \mathbf{H}^{\mathrm{T}}]^{\mathrm{T}}$ the vector obtained by the concatenation of all equality and inequality constraints. A major complication related to the presence of inequality constraints is the fact that an inequality constraint that is satisfied strictly in any $x \in \mathcal{H}$ should not be considered in the identification of the gradient flow in x, since they are satisfied everywhere in an open neighborhood of x. In order to take into account only the inequality constraints

[6] `https://cvxopt.org/`
[7] `https://osqp.org/docs/release-0.6.3`
[8] `https://coin-or.github.io/Ipopt/`

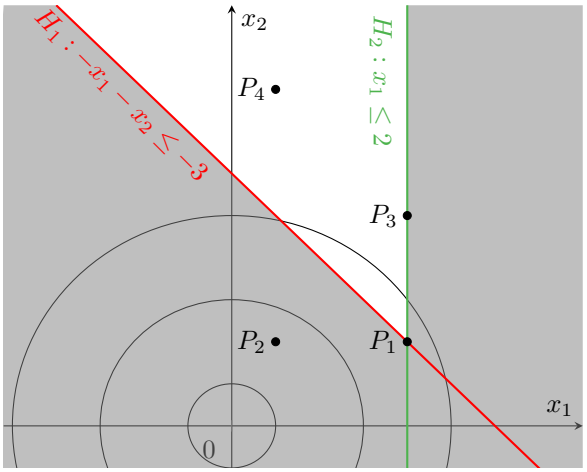


Figure 1.3: Representation of the space of phases $\mathbb{R}^2$ for the problem of minimizing $\|\mathbf{x}\|^2$ under the constraints $H_1 : -x_1 - x_2 \leq 3$ and $H_2 : x_1 \leq 2$. The sets of active constraints for the points $P_1, \dots, P_4$ are $\widetilde{I}(P_1) = \{1, 2\}$, $\widetilde{I}(P_2) = \{1\}$, $\widetilde{I}(P_3) = \{2\}$, $\widetilde{I}(P_4) = \{\}$.

which are violated or saturated, an active-set methodology analogous to [39] is used. We denote $\widetilde{I}(x) \subset I$ the set of all indices corresponding to the active constraints in x, which are all the equality constraints and the violated or saturated inequality constraints.

$$\widetilde{I}(x) = \{i \in I \ : \ i \leq N_{\mathbf{G}} \ \text{ or } \ C_i(x) \geq 0\}.$$

Moreover, we consider $\mathbf{C}_{\widetilde{I}(x)}$ to be the vector of all active constraints in x, $\mathbf{H}_{\widetilde{I}(x)}$ the vector of all active inequality constraints in x, and $\widetilde{N_{\mathbf{H}}}(x)$ the number of active inequality constraints. An example of the definition of active constraints is illustrated in fig. 1.3.

In this case, the vector $\mathbf{C}_{\widetilde{I}(x)}$ is said to satisfy the LICQ condition in x if $\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}(x)$ has full rank $N_{\mathbf{G}} + \widetilde{N_{\mathbf{H}}}(x)$, which is equivalent to require that $\left(\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}(x)\ \mathrm{D}\mathbf{C}_{\widetilde{I}(x)}(x)^{\mathcal{T}}\right)$ is an invertible square matrix.

As for the equality constrained problem, we aim to define an ODE in the form

$$\left\{ \begin{array}{rcll} \frac{\mathrm{d}}{\mathrm{d}x}t(t) & = & -\alpha_J \boldsymbol{\xi}_J(x) - \alpha_C \boldsymbol{\xi}_C(x), & \text{for } t \in (0, +\infty) \\ x(0) & = & x_0, & \end{array} \right. \tag{1.68}$$

whose solution converges towards the optimal solution of problem (1.67). We look for the definition of a null-space and a range-space direction analogous to (1.65). Let us consider $x \in \mathcal{H}$ that satisfies the LICQ condition. The range-space direction $\boldsymbol{\xi}_C$ is defined similarly to $\boldsymbol{\xi}_G$ in (1.65) as

$$\boldsymbol{\xi}_C(x) = \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widetilde{I}(x)}(\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}\ \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widetilde{I}(x)})^{-1}\mathbf{C}_{\widetilde{I}(x)}(x). \tag{1.69}$$

Once again, $-\boldsymbol{\xi}_C(x)$ is a descent direction for the violation of active constraints and it belongs to the range space of $\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}(x)$.

The computation of the null space direction $\boldsymbol{\xi}_J$ is not obtained simply by replacing $\mathrm{D}\mathbf{G}$ with $\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}$ in equation (1.65), since the projection of $-\nabla J$ onto $\mathrm{Ker}(\mathrm{D}\mathbf{C}_{\widetilde{I}(x)}(x))$ may not be the optimal direction for the solution of problem (1.67), as shown by fig. 1.4. Instead, we consider

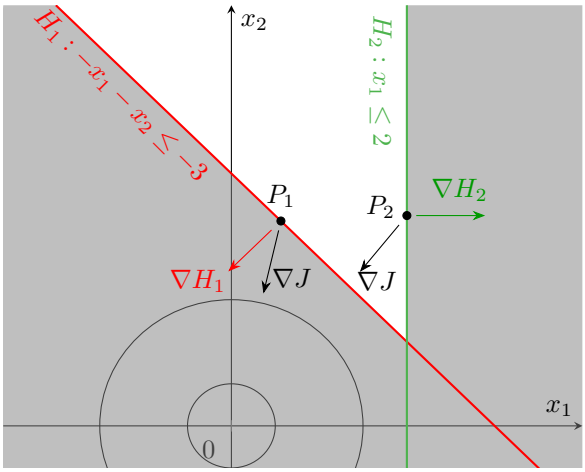


Figure 1.4: We consider once again the problem of minimizing $\|\mathbf{x}\|^2$ under the constraints $H_1 : -x_1 - x_2 \leq 3$ and $H_2 : x_1 \leq 2$. Constraints H_1 and H_2 are active in P_1 and P_2 respectively. The minimization of the objective and the satisfaction of the constraint are antagonistic in P_1, so $\widetilde{I}(P_1) = \widehat{I}(P_1) = \{1\}$. On the contrary, the improvement of the objective in P_2 does not violate H_2. Thus $\widetilde{I}(P_2) = \{2\}$, while $\widehat{I}(P_2) = \{\}$.

$-\boldsymbol{\xi}_J$ to be positively proportional to the solution $\boldsymbol{\xi}^\star$ of the following constrained optimization problem

$$\left|\begin{array}{l} \text{Find } \boldsymbol{\xi}^\star \in \mathcal{H} \\ \text{minimizing } \mathrm{D}J(x)\boldsymbol{\xi} \\ \text{under the constraints} \\ \qquad \mathrm{D}\mathbf{G}(x)\boldsymbol{\xi} = \mathbf{0}, \\ \qquad \mathrm{D}\mathbf{H}(x)\boldsymbol{\xi} \leq \mathbf{0}, \\ \qquad \|\boldsymbol{\xi}\|_{\mathcal{H}} \leq 0. \end{array}\right. \tag{1.70}$$

Instead of solving directly problem (1.70), it is suitable to consider its dual problem

$$\left|\begin{array}{l} \text{Find } \boldsymbol{\lambda}(x) \in \mathbb{R}^{N_{\mathbf{G}}} \text{ and } \boldsymbol{\mu}(x) \in \mathbb{R}^{\widetilde{N_{\mathbf{H}}}(x)}, \\ \text{minimizing } \left\| \nabla J(x) + \mathrm{D}\mathbf{G}^{\mathcal{T}}\boldsymbol{\lambda} + \mathrm{D}\mathbf{H}^{\mathcal{T}}_{\widehat{I}(x)}\boldsymbol{\mu} \right\|_{\mathcal{H}} \\ \text{under the constraints } \boldsymbol{\mu} \geq \mathbf{0}. \end{array}\right. \tag{1.71}$$

If $x \in \mathcal{H}$ satisfies the LICQ condition, problem (1.71) admits a unique solution [114, Proposition 3.2]. Let us consider the set of indices $\widehat{I}(x) \subset \widetilde{I}(x)$ defined as

$$\widehat{I}(x) = \left\{ i \in \widetilde{I}(x) \ : \ i \leq N_{\mathbf{G}} \ \text{ or } \ \boldsymbol{\mu}_{i-N_{\mathbf{G}}}(x) > 0 \right\}, \tag{1.72}$$

representing the subset of active constraints whose gradients are not aligned with $\nabla J(x)$.

Then, if x satisfies the LICQ condition, the null-space direction $\boldsymbol{\xi}_J(x)$ is defined as the projection of $\nabla J(x)$ onto the kernel of $\mathrm{D}\mathbf{C}_{\widehat{I}(x)}(x)$, and can be expressed as

$$\begin{aligned} \boldsymbol{\xi}_J(x) &= \left(\mathbb{I} - \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widehat{I}(x)} (\mathrm{D}\mathbf{C}_{\widehat{I}(x)} \ \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widehat{I}(x)})^{-1} \mathrm{D}\mathbf{C}_{\widehat{I}(x)}(x) \right) \nabla J(x) \\ &= \nabla J(x) + \mathrm{D}\mathbf{G}(x)^{\mathcal{T}} \ \boldsymbol{\lambda}(x) + \mathrm{D}\mathbf{H}_{\widehat{I}(x)}(x)^{\mathcal{T}} \ \boldsymbol{\mu}(x). \end{aligned} \tag{1.73}$$

Moreover, the solution of problem (1.70) is such that $\boldsymbol{\xi}^\star(x) = -\boldsymbol{\xi}_J(x)/\left\|\boldsymbol{\xi}_J(x)\right\|_{\mathcal{H}}$.

Discretization of the ODE

The *null space* optimization algorithm consists in a discretization of the ODE (1.68). Given a time step Δt sufficiently small, the discretized dynamical system evolves according to the following law

$$\left\{\begin{array}{rcll} \Delta x^n & = & (\alpha_J^n\ \boldsymbol{\xi}_J(x^n) + \alpha_C^n\ \boldsymbol{\xi}_C(x^n)) & \text{for } n \in \{1,\dots,N_{\max}\} \\ x^{n+1} & = & x^n - \Delta t\, \Delta x^n, & \text{for } n \in \{1,\dots,N_{\max}\} \\ x^0 & = & x_0. & \end{array}\right. \tag{1.74}$$

The coefficients α_J and α_C are tuned throughout the execution of the optimization algorithm as follows, in order to assure the normalization of the descent directions and avoid oscillations after a large number of iterations (see [91])

$$\alpha_J^n = \left\{\begin{array}{ll} \frac{A_J}{\|\boldsymbol{\xi}_J(x^n)\|_{\mathcal{H}}} & \text{if } n \le \overline{n}, \\ \frac{A_J}{\max\left(\|\boldsymbol{\xi}_J(x^n)\|_{\mathcal{H}}, \left\|\boldsymbol{\xi}_J(x^{\overline{n}})\right\|_{\mathcal{H}}\right)} & \text{if } n > \overline{n}; \end{array}\right.$$

$$\alpha_C^n = \min\left(\frac{A_C}{\|\boldsymbol{\xi}_C(x^n)\|_{\mathcal{H}}}, \overline{\alpha_C}\right).$$

The terms A_J, A_C, $\overline{n}$, and $\overline{\alpha_C}$ are fixed positive parameters.

An issue related to inequality constraints consists in the fact that the mapping of the active set $x \mapsto \widetilde{I}(x)$ is discontinuous on the boundary of the admissible set of problem (1.67). Such discontinuity can cause wide oscillations in the discretization of the ODE (1.68) [103]. The authors of [114] propose to introduce a vector of tolerances $\boldsymbol{\epsilon}(x) \in \mathbb{R}^{\widetilde{N_{\mathbf{H}}}(x)}$ defined as

$$\epsilon_i = \|\nabla H_i(x)\|_{\mathcal{H}}\, \mathtt{h} \qquad \text{for all } i \in I, \tag{1.75}$$

where $\mathtt{h}$ is a positive parameter. In order to avoid oscillations near the border of the admissible region, the set of active constraints $\widetilde{I}(x)$ is replaced by the set of active constraints up to ϵ_i

$$\widetilde{I}_{\boldsymbol{\epsilon}}(x) = \{i \in I\ :\ i \le N_{\mathbf{G}}\ \text{ or }\ C_i(x) \ge -\epsilon_i\}. \tag{1.76}$$

Similarly, if $\widetilde{I}_{\boldsymbol{\epsilon}}(x)$ replaces $\widetilde{I}(x)$ in the solution of problem (1.71) and denoting $\boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(x)$ and $\boldsymbol{\mu}_{\boldsymbol{\epsilon}}(x)$ its solution, it is possible to define the subset $\widehat{I}_{\boldsymbol{\epsilon}}(x) \subset \widetilde{I}_{\boldsymbol{\epsilon}}(x)$ as the set of all active constraints up to $\boldsymbol{\epsilon}$ which are either equality constraints, or whose gradient is not aligned with $\nabla J(x)$

$$\widehat{I}_{\boldsymbol{\epsilon}}(x) = \{i \in I\ :\ i \le N_{\mathbf{G}}\ \text{ or }\ \mu_{\boldsymbol{\epsilon},i-N_{\mathbf{G}}}(x) > \mathtt{tolLag}\}, \tag{1.77}$$

with $0 < \mathtt{tolLag} \ll 1$. Thus, for the discretization of the solution of the ODE (1.68), the null-space and range-space directions become

$$\boldsymbol{\xi}_{J,\boldsymbol{\epsilon}}(x) = \left(\mathbb{I} - \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widehat{I}_{\boldsymbol{\epsilon}}(x)}(\mathrm{D}\mathbf{C}_{\widehat{I}_{\boldsymbol{\epsilon}}(x)}\ \mathrm{D}\mathbf{C}^{\mathcal{T}}_{\widehat{I}_{\boldsymbol{\epsilon}}(x)})^{-1}\mathrm{D}\mathbf{C}_{\widehat{I}_{\boldsymbol{\epsilon}}(x)}\right)\nabla J(x), \tag{1.78}$$

$$\boldsymbol{\xi}_{C,\boldsymbol{\epsilon}}(x) = \mathrm{D}\mathbf{C}^{\mathcal{T}}_{I^\star_{\boldsymbol{\epsilon}}(x)}(\mathrm{D}\mathbf{C}_{I^\star_{\boldsymbol{\epsilon}}(x)}\ \mathrm{D}\mathbf{C}^{\mathcal{T}}_{I^\star_{\boldsymbol{\epsilon}}(x)})^{-1}\mathbf{C}_{I^\star_{\boldsymbol{\epsilon}}(x)}(x), \tag{1.79}$$

where $I^\star_{\boldsymbol{\epsilon}}(x) = \widehat{I}_{\boldsymbol{\epsilon}}(x) \cup \widetilde{I}(x)$. In order to ease the computations, the expression (1.78) is replaced by an equivalent formula

$$\boldsymbol{\xi}_{J,\boldsymbol{\epsilon}}(x) = \nabla J(x) + \mathrm{D}\mathbf{G}^{\mathcal{T}}\boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(x) + \mathrm{D}\mathbf{H}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x)}(x)\,\boldsymbol{\mu}_{\boldsymbol{\epsilon}}(x), \tag{1.80}$$

where the Lagrange multipliers solve the following problem

$$\begin{cases} \text{Find } \boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(x) \in \mathbb{R}^{N_{\mathbf{G}}} \text{ and } \boldsymbol{\mu}_{\boldsymbol{\epsilon}}(x) \in \mathbb{R}^{\mathrm{Card}\left(\mathbf{H}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x)}(x)\right)}, \\ \text{minimizing } \left\| \nabla J(x) + \mathrm{D}\mathbf{G}^{\mathcal{T}} \boldsymbol{\lambda}_{\boldsymbol{\epsilon}} + \mathrm{D}\mathbf{H}^{\mathcal{T}}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x)} \boldsymbol{\mu}_{\boldsymbol{\epsilon}} \right\|_{\mathcal{H}} \\ \text{under the constraints } \boldsymbol{\mu}_{\boldsymbol{\epsilon}} \geq \mathbf{0}. \end{cases} \tag{1.81}$$

The size of the time step is adapted during the optimization algorithm through the use of a *merit function* $\mathfrak{M}_n : \mathcal{H} \to \mathbb{R}$ defined at each iteration n as

$$\mathfrak{M}_n(x) = \alpha_J^n \left(J(x) + \mathbf{G}(x)\, \boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(x^n) + \mathbf{H}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)}(x)\, \boldsymbol{\mu}_{\boldsymbol{\epsilon}}(x^n) \right) + \frac{\alpha_C^n}{2} \left(\mathbf{C}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)}(x)^{\mathrm{T}} \mathbf{S}(x^n) \mathbf{C}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)}(x) \right). \tag{1.82}$$

The square positive definite matrix $\mathbf{S}(x^n)$ is defined as

$$\mathbf{S}(x^n) = \left(\mathrm{D}\mathbf{H}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)}(x^n)\ \mathrm{D}\mathbf{H}_{\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)}(x^n)^{\mathcal{T}} \right)^{-1}.$$

The merit function $\mathfrak{M}_n$ is defined in such a way that its gradient computed in x^n is equal to the increment Δx^n of equation (1.74). At each iteration of the algorithm, the merit function computed in x^{n+1} is compared with its value in x^n. If $\mathfrak{M}_n(x^{n+1}) < \mathfrak{M}_n(x^n)$, the algorithm proceeds to the next iteration. Otherwise, x^{n+1} is discarded and recomputed with a halved time step Δt. This procedure is repeated until an iterate for which the merit function decreases is found, or a maximal number of halvings $k_{\max}$ is attained.

The *null space* optimization algorithm for the solution of problem (1.67) can be summarized as in algorithm 2, and requires the input of the following parameters:

- Δt (denoted `dt` in the python implementation) virtual time-step of the discretized ODE;
- A_J and A_C scaling the null-space and range-space direction respectively in the definition of the increment (in the implementation, the parameters `alphaJ` $= A_J/\Delta t$ and `alphaC` $= A_C/\Delta t$ are provided instead);
- $\overline{n}$ (`itnormalisation`) the iteration number after which the normalization of the null-space step $\boldsymbol{\xi}_J$ is no further updated;
- $\overline{\alpha_C}$ (equal to 0.9 in the implementation) minimal normalization term for the range-space direction;
- the distance `h` used in the definition of the tolerance vector $\boldsymbol{\epsilon}$ (in the implementation is given the parameter `J` $=$ `h`$/\Delta t$);
- $n_{\max}$ (`maxit`) maximal number of steps in the optimization algorithm;
- $k_{\max}$ (`maxtrials`) maximal number of halvings in the search for a suitable increment step.

1.3.5 Summary of the algorithm for constrained shape optimization

Having stated in the previous sections the definitions and the main results about shape derivatives, we can now consider a more general shape optimization problem written similarly to

Algorithm 2 *null space* optimization algorithm for the solution of a constrained optimization problem in the form (1.67).

Require: a starting point $x^0 \in \mathcal{H}$
  Compute $J(x^0)$, $\mathbf{G}(x^0)$, and $\mathbf{H}(x^0)$
  while $n < n_{\max}$ **do**
    Compute the vector of tolerances $\boldsymbol{\epsilon}$ (1.75)
    Identify the sets $\widetilde{I}(x^n)$ and $\widetilde{I}_{\boldsymbol{\epsilon}}(x^n)$ of active constraints, and active constraints up to $\boldsymbol{\epsilon}$
    Compute the derivatives of the objective$\mathrm{D}J(x^n)$ and of the constraints $\mathrm{D}\mathbf{G}(x^n)$, $\mathrm{D}\mathbf{H}(x^n)$
    Compute the gradients $\nabla J(x^n)$, $\mathrm{D}\mathbf{G}(x^n)^{\mathcal{T}}$, and $\mathrm{D}\mathbf{H}(x^n)^{\mathcal{T}}$
    Compute the range-space direction $\boldsymbol{\xi}_{C,\boldsymbol{\epsilon}}(x^n)$ by equation (1.79)
    Solve the problem (1.81) to find the Lagrange multipliers $\boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(x^n)$ and $\boldsymbol{\mu}_{\boldsymbol{\epsilon}}(x^n)$
    Identify the set $\widehat{I}_{\boldsymbol{\epsilon}}(x^n)$ of active constraints up to $\boldsymbol{\epsilon}$ non aligned with ∇J
    Compute the null-space direction $\boldsymbol{\xi}_{J,\boldsymbol{\epsilon}}(x^n)$ by equation (1.80)
    Compute the increment $\Delta x^n = (\alpha_J^n\ \boldsymbol{\xi}_J(x^n) + \alpha_C^n\ \boldsymbol{\xi}_C(x^n))$
    Compute the merit function and evaluate $\mathfrak{M}_n(x^n)$
    $k \leftarrow 0$
    while $k \leq k_{\max}$ **do**
      Increment $x^{n+1} = x^n + \frac{\Delta t}{2^k}\,\Delta x$
      Evaluate $\mathfrak{M}_n(x^{n+1})$
      if $\mathfrak{M}_n(x^{n+1}) < \mathfrak{M}_n(x^n)$ **then**
        break
      else
        Reject x^{n+1}
        $k \leftarrow k + 1$
      end if
    end while
    Accept x^{n+1}
    Compute $J(x^{n+1})$, $\mathbf{G}(x^{n+1})$, and $\mathbf{H}(x^{n+1})$
    $n \leftarrow n + 1$
  end while

problem (1.3) as

$$\left|\begin{array}{l}\text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}}\\ \text{minimizing the objective function } J(\mathbf{u}_\Omega, \Omega)\\ \text{under the costraints}\\ \quad \left\{\begin{array}{rcl}\mathbf{G}(\mathbf{u}_\Omega, \Omega) & = & \mathbf{0},\\ \mathbf{H}(\mathbf{u}_\Omega, \Omega) & \leq & \mathbf{0},\end{array}\right.\\ \text{where the state } \mathbf{u}_\Omega \in \mathcal{Y}_{\text{adm}} \text{ is solution}\\ \quad \text{of the state problem } \mathbf{f}(y, \Omega) = \mathbf{0}.\end{array}\right. \tag{1.83}$$

We consider $\mathcal{S}_{\text{adm}}$ to be a class of Lipschitz continuous open and bounded domains in $\mathbb{R}^d$. The functions $\mathbf{G}$ and $\mathbf{H}$ encode respectively $N_{\mathbf{G}}$ equality constraint and $N_{\mathbf{H}}$ inequality constraint, and each one may depend on a state defined by a partial differential equation. For the sake of simplicity, we consider a single state function $\mathbf{u}_\Omega$ for the objective and the constraints.

If the state problem $\mathbf{f}(y, \Omega) = \mathbf{0}$ is well-posed for all $\Omega \in \mathcal{S}_{\text{adm}}$, problem (1.83) can be written

in terms of the reduced functionals

$$\left|\begin{array}{l}\text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the objective function } \widehat{J}(\Omega) \\ \text{under the costraints} \\ \quad \left\{\begin{array}{rcl} \widehat{\mathbf{G}}(\Omega) & = & \mathbf{0}, \\ \widehat{\mathbf{H}}(\Omega) & \leq & \mathbf{0}. \end{array}\right.\end{array}\right. \tag{1.84}$$

We suppose also that the state function $\mathbf{f}$, the objective J, and the constraints are regular enough to compute the shape derivatives $\mathrm{D}J$, $\mathrm{D}\mathbf{G}$ and $\mathrm{D}\mathbf{H}$ as continuous linear functionals on the space of Lipschitz continuous vector fields $\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$.

The *null space* optimization algorithm presented in section 1.3.4 can be adapted to shape and topology optimization by considering the set of admissible shapes as a Riemaniann manifold, whose metric is the scalar product associated to the regularization-extension procedure (see [114, Sections 5] for further information). This approach has proven its efficacy in level-set based shape optimization [112, 188, 91]. Other techniques that have been used in the context of structural optimization include the method of moving asymptotes for density-based approaches to topology optimization [45, 1], and the augmented Lagrangian [19, 91, 261] for level-set based optimization.

As remarked in [112, pg. 57], a shape optimization method based on Hadamard's derivative with respect to the domain should not allow for changes in the topology of the structure. Strictly speaking, the shape derivatives introduced in definition 1.5 describe the sensitivity of shape functionals with respect to small deformation which do not entail topological changes. However, the discretized nature of the algorithm makes them possible, in particular the piercing and disappearance of thin structures and the merging of two or more holes next to each other. Thus, sudden variations and even increases of the value of the functional of interest should be expected whenever changes in the topology occur during the execution of a shape optimization algorithm.

A shape optimization procedure can be developed by taking into account the definition of shape derivative of section 1.2, the discussion around body-fitting meshes of section 1.3.2, the Hilbertian regularization-extension procedure of section 1.3.3, and the *null space* optimization algorithm, outlined in section 1.3.4. The structure of the shape optimization procedure considered in this thesis follows the structure outlined in algorithm 3, similarly to the *sotuto*[9] platform devised by Dapogny and Feppon [91]. A similar solution is provided by the *OpenPisco*[10] software [188] developed by Safran Tech and IRT SystemX[11].

[9] `https://github.com/dapogny/sotuto`
[10] `https://openpisco.irt-systemx.fr/`
[11] `https://www.irt-systemx.fr/`

Algorithm 3 Procedure for shape optimization based on the *null space* optimization algorithm, using a level-set method and body-fitted meshes.

Require: a simplicial mesh $\mathcal{T}_D$ for the computational box D
Require: a level-set ϕ_{Ω^0} representing the initial domain $\Omega^0 \subset D$ on $\mathcal{T}_D$
  Compute a body-fitted mesh $\mathcal{T}_{D,\Omega^0}$
  Compute the signed distance function d_{Ω^0}
  Solve the state equation for $\mathbf{u}_{\Omega^0}$
  Compute the value of $J(\mathbf{u}_{\Omega^0}, \Omega^0)$, $\mathbf{G}(\mathbf{u}_{\Omega^0}, \Omega^0)$, $\mathbf{H}(\mathbf{u}_{\Omega^0}, \Omega^0)$
  while $n < n_{\max}$ **do**
    Compute the vector of tolerances $\boldsymbol{\epsilon}$ (1.75)
    Identify the sets of active constraints $\widetilde{I}(\Omega^n)$ as in (1.72), and $\widetilde{I}_{\boldsymbol{\epsilon}}(\Omega^n)$ as in (1.76)
    Compute the adjoint states for the objective or the constraints, where necessary
    Compute the shape derivatives $\mathrm{D}\widehat{J}(\Omega^n)$, $\mathrm{D}\widehat{\mathbf{G}}(\Omega^n)$, $\mathrm{D}\widehat{\mathbf{H}}(\Omega^n)$
    Compute the gradients $\nabla\widehat{J}(\Omega^n)$, $\mathrm{D}\widehat{\mathbf{G}}(\Omega^n)^{\mathcal{T}}$, $\mathrm{D}\widehat{\mathbf{H}}(\Omega^n)^{\mathcal{T}}$ by regularization-extension (1.62)
    Compute the range-space direction $\boldsymbol{\theta}_C^n$ by (1.79)
    Solve the problem (1.81) to find the Lagrange multipliers $\boldsymbol{\lambda}_{\boldsymbol{\epsilon}}(\Omega^n)$ and $\boldsymbol{\mu}_{\boldsymbol{\epsilon}}(\Omega^n)$
    Identify the set $\widehat{I}_{\boldsymbol{\epsilon}}(\Omega^n)$ defined in (1.77)
    Compute the null-space direction $\boldsymbol{\theta}_J^n$ by (1.80)
    Compute the displacement field $\boldsymbol{\theta}^n = (\alpha_J^n\ \boldsymbol{\theta}_J^n + \alpha_C^n\ \boldsymbol{\theta}_C^n)$
    Compute the merit function (1.82) and evaluate $\mathfrak{M}_n(\Omega^n)$
    $k \leftarrow 0$
    while $k \leq k_{\max}$ **do**
      Transport d_{Ω^n} by $\frac{1}{2^k}\boldsymbol{\theta}^n$ to get a new level-set $\phi_{\Omega^{n+1}}$ for Ω^{n+1} using `advect`
      Compute a body-fitted mesh $\mathcal{T}_{D,\Omega^{n+1}}$ using `mmg`
      Compute the signed distance function $d_{\Omega^{n+1}}$ using `mshdist`
      Evaluate $\mathfrak{M}_n(\Omega^{n+1})$
      if $\mathfrak{M}_n(\Omega^{n+1}) < \mathfrak{M}_n(\Omega^n)$ **then**
        break
      else
        Reject Ω^{n+1}
        $k \leftarrow k + 1$
      end if
    end while
    Accept Ω^{n+1}, $\mathcal{T}_{D,\Omega^{n+1}}$, and $d_{\Omega^{n+1}}$
    Solve the state equation for $\mathbf{u}_{\Omega^{n+1}}$
    Compute the value of $J(\mathbf{u}_{\Omega^{n+1}}, \Omega^{n+1})$, $\mathbf{G}(\mathbf{u}_{\Omega^{n+1}}, \Omega^{n+1})$, and $\mathbf{H}(\mathbf{u}_{\Omega^{n+1}}, \Omega^{n+1})$
    $n \leftarrow n + 1$
  end while

Chapter 2

A deterministic thermo-elastic problem

All formulas in section 2.2 have been obtained in collaboration with Viacheslav Karnaev[12], PhD student at the university of Basel. The results on deterministic thermo-elastic coupling are part of an ongoing work about shape optimization under uncertainties in a multiphysics context.

Contents

2.1 Linear elasticity framework

2.1.1 Introduction to linear elasticity

The subject of the optimization of static structure modeled by the equations of linear elasticity has been addressed since the early applications of the shape optimization algorithms (see [46, 47, 213, 210] for examples of density methods, and [17, 23, 205] for boundary variation methods). In the recent years, shape optimization of mechanical structures has been studied for more complex cases. Notable examples include the coupling with heat and fluid equations in a multiphysics system [113, 112], the application to connected structures [205, 204], and the consideration of

[12] viacheslav.karnaev@unibas.ch

contact phenomena [101, 143, 3]. Plastic deformations have been taken into account in their temporal dimension in [102, 101].

The problems studied in part II of this thesis consider elastic structure whose behavior is modeled by Hooke's linear elasticity equation, for which the displacement field in a mechanically charged structure is linearly proportional to the applied force. In particular, we focus our attention on the case of linear homogeneous isotropic materials. For further information on the theory of linear elasticity we refer to [232]. The impact of thermal coupling and time-dependent problems will be discussed in section 2.2.

Let us consider a structure represented by a Lipschitz continuous bounded domain $\Omega \subset \mathbb{R}^d$. We consider its surface $\partial\Omega$ to be divided in three disjoint parts with strictly positive measure: Γ_{D}, Γ_{N}, and Γ_0. We suppose that the structure is clamped in Γ_{D}, and a force $\mathbf{g}$ is applied on Γ_{N}. Moreover, the structure is subject to a volume force $\mathbf{f}$. The forces $\mathbf{f}$ and $\mathbf{g}$ induce a displacement of the structure represented by a vector field $\mathbf{u}_\Omega : \Omega \to \mathbb{R}^d$. In a steady problem, the conservation of mechanical momentum yields the following system of equations

$$\begin{cases} -\mathrm{div}\left(\boldsymbol{\sigma}\right) = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\ \mathbf{n} = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{\sigma}\ \mathbf{n} = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}, \end{cases} \tag{2.1}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor.

In the context of linear elasticity, the displacement is supposed to be small, and the stress tensor is taken, according to Hooke's law, as a linear function of the linearized strain tensor $\boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega\right)$ defined as

$$\boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega\right) = \frac{\nabla \mathbf{u}_\Omega + \nabla \mathbf{u}_\Omega{}^{\mathrm{T}}}{2}.$$

The linear relation between the stress and strain tensors is the constitutive relation of the material

$$\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right) = \mathbb{C}\boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega\right). \tag{2.2}$$

The expression of the fourth order elasticity tensor $\mathbb{C}$ depends on the material of the structure. However, in order to ensure that the operator $\nabla \mathbf{v} \mapsto \mathbb{C}\boldsymbol{\varepsilon}\left(\mathbf{v}\right)$ is uniformly elliptic and the elasticity problem is well posed, the elasticity tensor $\mathbb{C}$ must fulfill the following three conditions [116, 157]:

1. $\mathbb{C}_{ijkl} = \mathbb{C}_{klij}$, ensuring that $\mathbb{C}$ is a symmetric operator;
2. $\mathbb{C}_{ijkl} = \mathbb{C}_{jikl}$ and $\mathbb{C}_{ijkl} = \mathbb{C}_{ijlk}$, so that $\mathbb{C}$ is an endomorphism in the space of symmetric matrices of order d;
3. there exist two positive constants κ and K such that, for any $d \times d$ symmetric matrix $\mathbf{M}$, $\kappa\ (\mathbf{M} : \mathbf{M}) < (\mathbb{C}\mathbf{M} : \mathbf{M}) < K(\mathbf{M} : \mathbf{M})$, meaning that $\mathbb{C}$ is a strictly positive and continuous operator.

Under the aforementioned hypotheses, equation (2.1) becomes the equation of linear elasticity

$$\begin{cases} -\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\right) = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\ \mathbf{n} = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\ \mathbf{n} = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{cases} \tag{2.3}$$

Problem (2.3) is written in weak form as

$$\begin{cases} \text{Find } \mathbf{u}_\Omega \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \text{ such that} \\ \text{for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \\ \displaystyle\int_\Omega \boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) = \int_\Omega \mathbf{f}\cdot\mathbf{v} + \int_{\Gamma_\mathrm{N}} \mathbf{g}\cdot\mathbf{v}. \end{cases} \tag{2.4}$$

The symmetries required for the fourth order elasticity tensor $\mathbb{C}$ limit drastically the number of independent components (6 for two-dimensional problems, and 21 for three-dimensional problems). For homogeneous isotropic materials, the independent components of the elasticity tensor are reduced to two. In this case $\mathbb{C}$ is written as function of the Lamé parameters λ and μ as

$$\boldsymbol{\sigma}(\mathbf{u}_\Omega) = \mathbb{C}\boldsymbol{\varepsilon}(\mathbf{u}_\Omega) = 2\mu\boldsymbol{\varepsilon}(\mathbf{u}_\Omega) + \lambda(\operatorname{div}\mathbf{u}_\Omega)\mathbb{I}. \tag{2.5}$$

In three dimensions, the Lamé parameters can be expressed in terms of Young's modulus E and Poisson's ratio ν as

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

In order to ensure the positivity and continuity of the elasticity tensor, the Lamé parameters must satisfy the following two inequalities

$$\mu > 0, \qquad 2\mu + d\lambda > 0.$$

We remark that the parameter λ can technically be negative. However, it is positive for most materials. The parameter μ is also known as the shear modulus of the material, and it measures the ratio between the shear stress and the shear strain. Both Lamé parameters have the dimension of a pressure.

The well-posedness of problem (2.4) is a classical result (see [116] for a generic elasticity tensor and [218, Section 9.3] for the homogeneous and isotropic case), and relies on Lax-Milgram's lemma and Korn's inequality (see [157, 73] and [218, Lemma 9.5]).

2.1.2 The mechanical compliance

A shape functional that appears very often in studies on shape and topology optimization of elastic structures is the mechanical compliance, which is defined as the work of all forces acting on the structure [6, Section 5.1]:

$$\widehat{\mathcal{C}}(\Omega) = \mathcal{C}(\Omega, \mathbf{u}_\Omega) = \int_\Omega \mathbf{f}\cdot\mathbf{u}_\Omega \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \mathbf{g}\cdot\mathbf{u}_\Omega \,\mathrm{d}\mathbf{s}. \tag{2.6}$$

The expression of the compliance coincides with the left-hand side of the variational formulation (2.4), where the displacement $\mathbf{u}_\Omega \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)$ is taken as test function. Thus, the compliance can be written as twice the total density of the elastic energy in the structure

$$\mathcal{C}(\Omega, \mathbf{u}_\Omega) = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) \,\mathrm{d}\mathbf{x}, \tag{2.7}$$

which is a quadratic function of the displacement $\mathbf{u}_\Omega$. Expression (2.7) highlights the interpretation of the compliance as the integral over the volume of the structure of the elastic energy induced by the forces $\mathbf{f}$ and $\mathbf{g}$. As remarked in [6, 101], a large value of the compliance implies a high transfer of energy to the domain, and thus a significant deformation. Thus, the compliance

can be interpreted as a measure of the flexibility of the structure with respect to the applied loads $\mathbf{f}$ and $\mathbf{g}$.

An important property of the compliance as a differentiable shape functional is the fact that it does not require the computation of an adjoint state for the formulation of its shape derivative. Indeed, let $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ be a Lipschitz continuous vector field such that $\|\boldsymbol{\theta}\|_{1,\infty} < 1$ and $\boldsymbol{\theta} = \mathbf{0}$ on Γ_{D}. The mechanical compliance on the deformed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I} + \boldsymbol{\theta})\Omega$ is

$$\begin{aligned}
&\mathcal{C}\left(\Omega_{\boldsymbol{\theta}}, \mathbf{u}_{\Omega_{\boldsymbol{\theta}}}\right) = \int_{\Omega_{\boldsymbol{\theta}}} \boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}) : \boldsymbol{\varepsilon}_{\widetilde{\mathbf{x}}}(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}})\, \mathrm{d}\widetilde{\mathbf{x}} \\
&= \int_{\Omega_{\boldsymbol{\theta}}} \left(\mu\left(\nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}} : \nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}} + \nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}} : \nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}{}^{\mathrm{T}}\right) + \lambda\left(\mathrm{div}_{\widetilde{\mathbf{x}}}\, \mathbf{u}_{\Omega_{\boldsymbol{\theta}}}\right)^2\right) \mathrm{d}\widetilde{\mathbf{x}} \\
&= \int_{\Omega} \Big(\mu\left(\nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}}\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-1} : \nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}}\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-1} + \nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}}\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-1} : \left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-\mathrm{T}}\left(\nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right) \\
&\quad + \lambda\left(\nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}} : \left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-\mathrm{T}}\right)^2\Big) \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \mathrm{d}\mathbf{x},
\end{aligned} \tag{2.8}$$

where $\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}$ is the solution of the elasticity equation (2.3) on the deformed domain, $\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}} = \mathbf{u}_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I} + \boldsymbol{\theta})$ its push-back on the reference domain, and the index "$\widetilde{\mathbf{x}}$" under the differential operators indicates that they are computed with respect to the perturbed coordinates. By differentiating equation (2.8) with respect to $\boldsymbol{\theta}$ and using the symmetries of the elasticity tensor we obtain

$$\begin{aligned}
&\mathrm{D}\widehat{\mathcal{C}}(\Omega)(\widehat{\boldsymbol{\theta}}) = 2\int_{\Omega} \mu\left(\nabla\dot{\mathbf{u}}_{\Omega}(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{u}_{\Omega} + \nabla\dot{\mathbf{u}}_{\Omega}(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{u}_{\Omega}{}^{\mathrm{T}} - \nabla\mathbf{u}_{\Omega}\nabla\widehat{\boldsymbol{\theta}} : \nabla\mathbf{u}_{\Omega} - \nabla\mathbf{u}_{\Omega}\nabla\widehat{\boldsymbol{\theta}} : \nabla\mathbf{u}_{\Omega}{}^{\mathrm{T}}\right) \mathrm{d}\mathbf{x} \\
&+ 2\int_{\Omega} \lambda\left((\mathrm{div}\,\dot{\mathbf{u}}_{\Omega}(\widehat{\boldsymbol{\theta}}))(\mathrm{div}\,\mathbf{u}_{\Omega}) - \left(\nabla\mathbf{u}_{\Omega} : \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\right)(\mathrm{div}\,\mathbf{u}_{\Omega})\right) \mathrm{d}\mathbf{x} + \int_{\Omega} (\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega}\right))(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\, \mathrm{d}\mathbf{x} \\
&= 2\int_{\Omega} \boldsymbol{\sigma}\left(\dot{\mathbf{u}}_{\Omega}(\widehat{\boldsymbol{\theta}})\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega}\right) \mathrm{d}\mathbf{x} - 2\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega}\right) : \left(\nabla\mathbf{u}_{\Omega}\nabla\widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{x} + \int_{\Omega} (\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega}\right))(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\, \mathrm{d}\mathbf{x}.
\end{aligned} \tag{2.9}$$

The equation solved by the Lagrangian derivative is found by differentiating the variational form of problem (2.4). For any $\mathbf{v} \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)$ defined on the reference domain, the weak form of the state equation is

$$\int_{\Omega_{\boldsymbol{\theta}}} \boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}) : \boldsymbol{\varepsilon}_{\widetilde{\mathbf{x}}}(\mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1})\, \mathrm{d}\widetilde{\mathbf{x}} = \int_{\Omega_{\boldsymbol{\theta}}} \mathbf{f}(\widetilde{\mathbf{x}}) \cdot \mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\, \mathrm{d}\widetilde{\mathbf{x}} + \int_{(\mathbf{I}+\boldsymbol{\theta})\Gamma_{\mathrm{N}}} \mathbf{g}(\widetilde{\mathbf{x}}) \cdot \mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\, \mathrm{d}\widetilde{\mathbf{s}}$$

which, transported on Ω, yields

$$\begin{aligned}
&\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{v}\right) : \left(\nabla\widetilde{\mathbf{u}}_{\Omega}^{\boldsymbol{\theta}}\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)^{-1}\right) \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \mathrm{d}\mathbf{x} \\
&= \int_{\Omega} \mathbf{f} \circ (\mathbf{I} + \boldsymbol{\theta}) \cdot \mathbf{v} \left|\det\left(\mathbb{I} + \nabla\boldsymbol{\theta}\right)\right| \mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \mathbf{g} \circ (\mathbf{I} + \boldsymbol{\theta}) \cdot \mathbf{v} \mathrm{Jac}_{\Gamma_{\mathrm{N}}}(\mathbf{I} + \boldsymbol{\theta})\, \mathrm{d}\mathbf{s}.
\end{aligned} \tag{2.10}$$

By differentiating both sides of equation (2.10), we obtain the following expression

$$\begin{aligned}
&\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{v}\right) : \left(\nabla\dot{\mathbf{u}}_{\Omega}(\widehat{\boldsymbol{\theta}}) - \nabla\mathbf{u}_{\Omega}\nabla\widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{x} + \int_{\Omega} (\boldsymbol{\sigma}\left(\mathbf{v}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega}\right))(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\, \mathrm{d}\mathbf{x} \\
&= \int_{\Omega} \left(\nabla\mathbf{f}\widehat{\boldsymbol{\theta}} \cdot \mathbf{v} + (\mathbf{f} \cdot \mathbf{v})(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\right) \mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \left(\nabla\mathbf{g}\widehat{\boldsymbol{\theta}} \cdot \mathbf{v} + (\mathbf{g} \cdot \mathbf{v})(\mathrm{div}_{\Gamma_{\mathrm{N}}}\,\widehat{\boldsymbol{\theta}})\right) \mathrm{d}\mathbf{x}.
\end{aligned}$$

Since the equation above holds for any test function $\mathbf{v} \in \mathrm{H}^1(\Gamma_\mathrm{D})\,\Omega$ and thanks to the symmetry properties of the elasticity tensor, the following variational formulation for $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ holds

$$\left|\begin{array}{l}\text{Find } \dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \text{ such that}\\ \text{for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)\\ \displaystyle\int_\Omega \boldsymbol{\sigma}\left(\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\right) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\mathbf{x} = \int_\Omega \boldsymbol{\sigma}(\mathbf{v}) : \left(\nabla \mathbf{u}_\Omega \nabla\widehat{\boldsymbol{\theta}}\right)\,\mathrm{d}\mathbf{x} - \int_\Omega (\boldsymbol{\sigma}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega))(\operatorname{div}\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\\ \qquad\qquad + \displaystyle\int_\Omega \left(\nabla\mathbf{f}\widehat{\boldsymbol{\theta}} + \operatorname{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \left(\nabla\mathbf{g}\widehat{\boldsymbol{\theta}} + \operatorname{div}_{\Gamma_\mathrm{N}}(\widehat{\boldsymbol{\theta}})\mathbf{g}\right)\cdot\mathbf{v}\,\mathrm{d}\mathbf{s}.\end{array}\right. \tag{2.11}$$

By choosing $\mathbf{u}_\Omega$ as test function, we obtain the term with the Lagrangian derivative of the displacement of equation (2.9) on the right-hand side of the equation in (2.11). Thus, it is possible to formulate the shape derivative of the mechanical compliance without the need of an adjoint state

$$\begin{aligned}\mathrm{D}\widehat{\mathcal{C}}(\Omega)(\widehat{\boldsymbol{\theta}}) &= -\int_\Omega (\boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega))(\operatorname{div}\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\\ +2\int_\Omega \left(\nabla\mathbf{f}\widehat{\boldsymbol{\theta}} + \operatorname{div}(\widehat{\boldsymbol{\theta}})\mathbf{f}\right)\cdot\mathbf{u}_\Omega\,\mathrm{d}\mathbf{x} &+ 2\int_{\Gamma_\mathrm{N}} \left(\nabla\mathbf{g}\widehat{\boldsymbol{\theta}} + \operatorname{div}_{\Gamma_\mathrm{N}}(\widehat{\boldsymbol{\theta}})\mathbf{g}\right)\cdot\mathbf{u}_\Omega\,\mathrm{d}\mathbf{s}.\end{aligned} \tag{2.12}$$

If Ω is a $\mathcal{C}^1$ domain, the hypotheses of theorem 1.7 hold, and the shape derivative of the compliance can be written as

$$\begin{aligned}\mathrm{D}\widehat{\mathcal{C}}(\Omega)(\widehat{\boldsymbol{\theta}}) = &\int_{\Gamma_\mathrm{N}\cup\Gamma_0} (\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\left(-\boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) + 2\mathbf{f}\cdot\mathbf{u}_\Omega\right)\,\mathrm{d}\mathbf{s}\\ &+ 2\int_{\Gamma_\mathrm{N}} (\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\left(\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\cdot\mathbf{g} + \frac{\partial\mathbf{g}}{\partial\mathbf{n}}\cdot\mathbf{u}_\Omega + H(\mathbf{g}\cdot\mathbf{u}_\Omega)\right)\,\mathrm{d}\mathbf{s}.\end{aligned} \tag{2.13}$$

2.1.3 The von Mises criterion

Let us consider a structure covering a domain Ω, composed of a linear elastic, homogeneous and isotropic material, characterized by the Lamé parameters λ and μ. A material can lose its elastic properties and assume a plastic behavior if it is subject to a sufficiently intense stress. Notable aspects of plastic deformation are the non-linearity of the relation between stress and strain, and the fact that the structure does not return to its initial configuration once the mechanical load is removed. These phenomena can be explained by permanent modifications in the crystalline structure of the material. Therefore, plastic deformations are generally avoided in many industrial applications, since they can entail a degradation of the material and of its elastic properties.

A classic behavior of an elasto-plastic material subject to uniaxial tensile loading is represented in the diagram of fig. 2.1. If a sample is progressively stretched, the strain-stress relation remains linear until the stress reaches the yield point σ_Y. After this threshold, plasticity behavior appears and the relation between the stress and the strain becomes non-linear (we refer to [145] for a detailed description of the non-linear phenomena). If the sample is stretched even further, it may break once the strain reaches the fracture point in the diagram. Otherwise, if the sample is released past the yield point, it does not return to the initial configuration, but it maintains a residual strain.

In order to describe the limit of the elastic behavior of materials for two or three dimensional deformations, we introduce the decomposition of the stress tensor as the sum of a hydrostatic

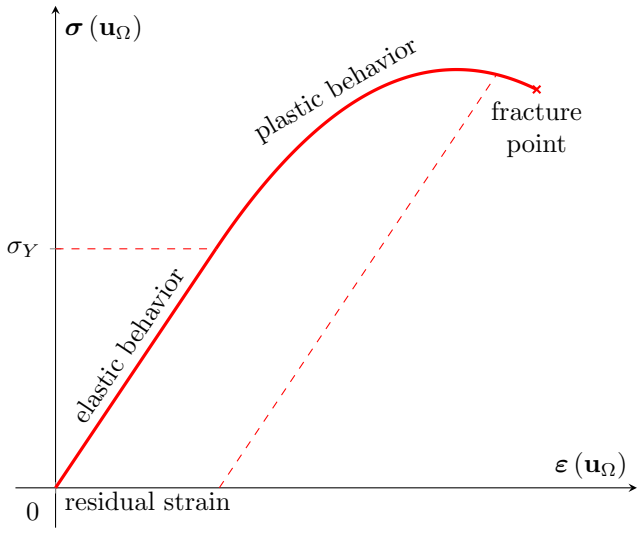


Figure 2.1: Example of a stress-strain curve for an elasto-plastic material subject to uniaxial strain.

and deviatoric component

$$\boldsymbol{\sigma}(\mathbf{u}_\Omega) = \boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) + \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega), \tag{2.14}$$

where

$$\boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) = \mathbb{I}\,\frac{1}{d}\,\mathrm{tr}\left(\boldsymbol{\sigma}(\mathbf{u}_\Omega)\right), \qquad \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) = \boldsymbol{\sigma}(\mathbf{u}_\Omega) - \boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega). \tag{2.15}$$

The hydrostatic component identifies the mean normal stress exerted in each point, while the deviatoric part is related to distortion [31]. The same decomposition can be applied to the strain tensor $\boldsymbol{\varepsilon}(\mathbf{u}_\Omega) = \boldsymbol{\varepsilon}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) + \boldsymbol{\varepsilon}_{\mathrm{D}}(\mathbf{u}_\Omega)$:

$$\boldsymbol{\varepsilon}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) = \mathbb{I}\,\frac{1}{d}\left(\mathrm{div}\,\mathbf{u}_\Omega\right), \qquad \boldsymbol{\varepsilon}_{\mathrm{D}}(\mathbf{u}_\Omega) = \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) - \boldsymbol{\varepsilon}_{\mathrm{Hyd}}(\mathbf{u}_\Omega).$$

For linear homogeneous isotropic materials, the tensors $\boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega)$ and $\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)$ can be expressed in terms of the Lamé parameters as

$$\boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) = \left(\frac{2\mu}{d} + \lambda\right)(\mathrm{div}\,\mathbf{u}_\Omega)\,\mathbb{I}, \qquad \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) = 2\mu\boldsymbol{\varepsilon}(\mathbf{u}_\Omega) - \frac{2\mu}{d}(\mathrm{div}\,\mathbf{u}_\Omega)\,\mathbb{I} = 2\mu\,\boldsymbol{\varepsilon}_{\mathrm{D}}(\mathbf{u}_\Omega). \tag{2.16}$$

Thanks to the orthogonality between the identity tensor and the space of traceless tensors, the elastic energy density U in each point of a loaded elastic structure can be written as the sum of the dilatation and distortion energy densities U_{Hyd} and U_{D} as

$$U(\mathbf{u}_\Omega) = \frac{1}{2}\,\boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) = U_{\mathrm{Hyd}}(\mathbf{u}_\Omega) + U_{\mathrm{D}}(\mathbf{u}_\Omega) \tag{2.17}$$

in each point of Ω, where

$$U_{\mathrm{Hyd}}(\mathbf{u}_\Omega) = \frac{1}{2}\,\boldsymbol{\sigma}_{\mathrm{Hyd}}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}_{\mathrm{Hyd}}(\mathbf{u}_\Omega), \qquad U_{\mathrm{D}}(\mathbf{u}_\Omega) = \frac{1}{2}\,\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}_{\mathrm{D}}(\mathbf{u}_\Omega).$$

Historically, the von Mises stress has been developed in the context of plasticity criteria [145, Section 1.4.4]. The intuition of the dependency of plastic deformation from the deviatoric part

of the stress tensor emerged in England in the nineteenth century. The idea was then adopted as an hypothesis for subsequent plasticity criteria, like Tresca's and von Mises', and confirmed by experiments. It should be remarked that such assumption is adapted only for ductile materials as aluminum or steel. Indeed, structures made up of brittle materials, like concrete, can undergo permanent deformations and fissuring following isotropic compression or dilatation. For these cases, the hydrostatic and deviatoric components of the stress tensor must be taken into account by the material failure criterion.

Von Mises' plasticity criterion has been formulated for the first time by Richard von Mises in 1913 [251], but its physical interpretation in terms of the deformation energy has been provided only in 1924 by Heinrich Hencky [137]. Under this interpretation, plastic deformation occurs when the distortion energy density in some point of Ω exceeds the distortion energy density of a sample of the same material subject to uniaxial traction stress at the yield point. For a sample of linear isotropic material subject to uniaxial traction of module σ_Y, the stress tensor $\boldsymbol{\sigma}_Y$ and its deviatoric component $\boldsymbol{\sigma}_{Y,\mathrm{D}}$ are

$$\boldsymbol{\sigma}_Y = \begin{bmatrix} \sigma_Y & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{\sigma}_{Y,\mathrm{D}} = \begin{bmatrix} \frac{2\sigma_Y}{3} & 0 & 0 \\ 0 & -\frac{\sigma_Y}{3} & 0 \\ 0 & 0 & -\frac{\sigma_Y}{3} \end{bmatrix}.$$

Therefore, thanks to the proportionality between the deviatoric strain and stress, the distortion energy density related to the uniaxial traction is

$$U_{Y,\mathrm{D}} = \frac{1}{2}\,\boldsymbol{\sigma}_{Y,\mathrm{D}} : \left(\frac{1}{2\mu}\boldsymbol{\sigma}_{Y,\mathrm{D}}\right) = \frac{1}{2}\,\frac{6\,\sigma_Y^2}{9}\,\frac{1}{2\,\mu} = \frac{\sigma_Y^2}{6\,\mu} \tag{2.18}$$

According to Hencky's interpretation of the von Mises criterion, a loaded structure preserves its elastic behavior if, in each point, the distortion energy density is smaller than $U_{Y,\mathrm{D}}$ as computed in equation (2.18).

$$\frac{\sigma_Y^2}{6\,\mu} = U_{Y,\mathrm{D}} \geq U_{\mathrm{D}}(\mathbf{u}_\Omega) = \frac{1}{2}\,\boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\varepsilon}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) = \frac{1}{4\,\mu}\,\boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right).$$

Thus, the criterion can be written as

$$\sqrt{\frac{3}{2}}\left(\boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right)\right)^{1/2} \leq \sigma_Y. \tag{2.19}$$

The left-hand side of the inequality (2.19) is defined as the von Mises stress

$$s_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) = \sqrt{\frac{3}{2}}\Big(\boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{u}_\Omega\right)\Big)^{1/2}. \tag{2.20}$$

We refer to [145, Section 4.5.6] for further details on the interpretation of the criterion and its comparison with other plasticity criteria.

The interpretation in terms of a plasticity criterion suggests that, in order to avoid plastic deformations and deterioration of the material, it is important to control the value of the von Mises stress in the entire domain while designing the structure. Requiring the von Mises stress to remain below a given threshold in a domain is equivalent to control its supremum, or its L^∞-norm, which is defined like in (1.12) as

$$\left\|s_{\mathrm{D}}\left(\mathbf{u}_\Omega\right)\right\|_\infty = \operatorname*{ess\,sup}_{\mathbf{x}\in\Omega} \left|s_{\mathrm{D}}\left(\mathbf{u}_\Omega\right)(\mathbf{x})\right|. \tag{2.21}$$

However, the quantity defined in (2.21) is not differentiable with respect to the domain, and thus cannot be considered as a constraint functional in a shape optimization problem solved by gradient-based algorithms.

An alternative to the supremum of the von Mises stress is given by its L^p-norm for p sufficiently large, as proven by proposition 1.3. This approximation has the advantage of being differentiable for all $p < \infty$ if the domain and the loads are sufficiently regular. The order p of the norm ought to be chosen carefully. On the one hand, if p is too small, substantial concentrations of stress could be overlooked, especially for wide domains. On the other hand, the computation of $\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_p$ can be numerically challenging for large value of p and requires highly refined meshes to capture the stress peaks inside the structure.

By the expression (2.20), the L^p-norm of the von Mises stress is

$$\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_p = \sqrt{\frac{3}{2}}\left(\int_\Omega \Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)^{p/2} \mathrm{d}\mathbf{x}\right)^{1/p} \tag{2.22}$$

Let us consider an elastic structure $\Omega \subset \mathbb{R}^d$ which has a $\mathcal{C}^1$ boundary and whose displacement $\mathbf{u}_\Omega$ solves problem (2.3). Then, we can apply the results of section 1.2.2 and obtain the following expression for the shape derivative of $\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_p$

$$\begin{aligned}\mathrm{D}\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_p(\widehat{\boldsymbol{\theta}}) = \left(\frac{3}{2}\right)^{1/2} &\frac{1}{p\,\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_p^{p-1}}\left(\int_{\Gamma_0\cup\Gamma_{\mathrm{N}}} \Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)^{p/2}\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}\right.\\ &+\int_{\Gamma_0\cup\Gamma_{\mathrm{N}}} \left(\mathbf{f}\cdot\mathbf{w}_\Omega - \boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{w}_\Omega)\right)\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}\\ &\left.+\int_{\Gamma_{\mathrm{N}}}\left(\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\cdot\mathbf{w}_\Omega + \frac{\partial\mathbf{w}_\Omega}{\partial\mathbf{n}}\cdot\mathbf{g} + H(\mathbf{g}\cdot\mathbf{w}_\Omega)\right)\left(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}\right)\mathrm{d}\mathbf{s}\right)\end{aligned} \tag{2.23}$$

The adjoint state $\mathbf{w}_\Omega$ satisfies the following boundary values problem

$$\left\{\begin{array}{rcll} -\mathrm{div}\left(\boldsymbol{\sigma}(\mathbf{w}_\Omega)\right) = & -\mathrm{div}\left(\frac{p}{2}\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)^{\frac{p}{2}-1}\nabla\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)\right) & \text{in } \Omega,\\ \left(\mathbf{A}^{\mathrm{T}}\nabla\mathbf{w}_\Omega\right)\mathbf{n} = & \frac{p}{2}\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)^{\frac{p}{2}-1}\nabla\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)\mathbf{n} \quad \text{on} & \Gamma_{\mathrm{N}}\cup\Gamma_0,\\ \mathbf{w}_\Omega = & 0 & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \tag{2.24}$$

The variational formulation of (2.24)

$$\left|\begin{array}{l}\text{Find } \mathbf{w}_\Omega \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \text{ such that for all } \mathbf{v} \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)\\ \displaystyle\int_\Omega \boldsymbol{\sigma}(\mathbf{w}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\mathbf{x} = \int_\Omega \frac{p}{2}\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega)\Big)^{\frac{p}{2}-1}\Big(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\Omega) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{v})\Big)\,\mathrm{d}\mathbf{x}.\end{array}\right. \tag{2.25}$$

2.2 Formulation of the thermo-elastic problem

2.2.1 Time dependent thermo-elasticity

An aspect that can introduce some perturbations in the elastic behavior of a mechanical structure consists in taking into account the influence of the temperature on the material properties . In particular, unless the temperature reaches large enough values to provoke the melting of the material or alter its elasto-plastic properties, thermal dilation is an important phenomenon

to take into account. Thermal effects has been taken into account in shape optimization problems in the literature (see, for example, [255, 15, 112]). In particular, applications of topology optimization to additive manufacturing require to pay attention to the different effects of high temperature gradients on the structure as well as their temporal dimension (refer to [15] for more details on thermal effects in 3D printing).

In this chapter we focus on the thermal dilation and the mechanical stress caused by the presence of a time-variant temperature field surrounding an elastic structure. Let us consider a bounded domain $D \in \mathbb{R}^d$, where the temperature field is defined, containing an open domain Ω, representing an elastic structure. We consider the evolution of the temperature inside Ω in the time interval $[0, t_f]$. We suppose that the temperature field in $D \setminus \Omega$ is given as a function of space and time, and that the material of the structure interacts with the temperature of its surroundings by Fourier's law of thermal conduction. The variation of temperature in Ω is modeled as the solution of the heat equation, where Fourier-Robin boundary conditions are imposed on the interface between D and Ω. Let ρ, k, and β be positive real parameters representing the product of the mass density and the thermal capacity of the material composing the structure, its thermal conductivity, and the heat transfer coefficient respectively. We consider also $Q \in \mathrm{L}^2(\Omega)$ to be a heat source internal to the structure. Finally, we assume that the temperature in the external medium is known, we denote its expression by the function $T_{\text{ext}} : (D \setminus \Omega) \times [0, t_f] \to \mathbb{R}$, and we suppose that

$$T_{\text{ext}} \in \mathrm{L}^2\left([0,t_f]; \mathrm{L}^2(\partial\Omega)\right) = \left\{ T : (D \setminus \Omega) \times [0,t_f] \to \mathbb{R} \ : \ \int_0^t \left(\int_{\Gamma_{\mathrm{N}} \cup \Gamma_0} |T|^2 \, \mathrm{d}\mathbf{x} \right) \mathrm{d}t < +\infty \right\}.$$

Under the assumptions above, the differential equation describing the evolution of the temperature field T_Ω in Ω for the time interval $[0, t_f]$ is

$$\left\{ \begin{array}{rcll} \rho \frac{\partial T_\Omega}{\partial t} - \operatorname{div}(k \nabla T_\Omega) & = & Q & \text{in } (0,t_f] \times \Omega, \\ k \nabla T_\Omega \cdot \mathbf{n} + \beta (T_\Omega - T_{\text{ext}}) & = & 0 & \text{on } (0,t_f] \times \partial\Omega, \\ T_\Omega(0, \mathbf{x}) & = & T_0 & \text{in } \Omega, \end{array} \right. \tag{2.26}$$

where $T_0 \in \mathrm{L}^2(\Omega)$ is the initial temperature distribution. The parabolic boundary value problem (2.26) is well-posed, as proven in [218, Theorem 2.2], and its solution belongs to the function space $\mathrm{L}^2([0,t_f]; \mathrm{H}^1(\Omega))$.

We consider that the temperature field affects the mechanics of the structure through thermal dilation. Let $\alpha > 0$ be the thermal dilation coefficient, and $T_{\text{ref}} \in \mathbb{R}$ a reference temperature. The phenomenon of thermal dilation for linear elastic systems can be taken into account by adding a term depending on the temperature to the Cauchy stress tensor. For linear, homogeneous, isotropic materials, the modified stress tensor follows the Duhamel-Neumann law [168] and assumes the following expression

$$\boldsymbol{\sigma}(\mathbf{u}_\Omega, T_\Omega) = \mathbb{C} \nabla \mathbf{u}_\Omega + \boldsymbol{\sigma}_{\text{Th}}(T_\Omega),$$

where $\mathbb{C}$ is the elasticity tensor introduced in (2.5), and $\boldsymbol{\sigma}_{\text{Th}}(\mathbf{u}_\Omega)$ the component dependent on the temperature, which is defined as

$$\boldsymbol{\sigma}_{\text{Th}}(\mathbf{u}_\Omega) = -\alpha (T_\Omega - T_{\text{ref}}) \mathbb{I}. \tag{2.27}$$

We consider that the surface $\partial\Omega$ of the structure can be divided in three disjoint parts: Γ_{D}, Γ_{N}, and Γ_0. Let $\mathbf{f} \in \mathrm{L}^2([0,t_f]; \mathrm{H}^{-1}(\Omega))$ and $\mathbf{g} \in \mathrm{L}^2\left([0,t_f]; \mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})\right)$ be the mechanical solicitations applied in the volume of the structure and on the surface Γ_{N} respectively. We consider

that the elastic displacements caused by the external mechanical loads propagate faster than the heat, so that at each instant the structure is supposed at its mechanical equilibrium. Thus, for all $t \in (0, t_f]$, the displacement $\mathbf{u}_\Omega(t, \mathbf{x})$ can be computed solving the following boundary values problem

$$\left\{\begin{array}{rcll} -\text{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega, T_\Omega\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega, T_\Omega\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega, T_\Omega\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega(\mathbf{s}, t) & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}, \end{array}\right. \tag{2.28}$$

where T_Ω is the solution of (2.26).

The temperature appears in the expression of the elasticity problem, but we ignore the heat generated by internal friction or other mechanical phenomena. Thus, the thermal and mechanical problems are said to be weakly coupled. In order to compute the state of the system in the time interval $[0, t_f]$, it is possible to solve first the time-dependent heat equation, and then solve for the displacement.

2.2.2 The optimization problem

Having stated the heat and elasticity equations for an elastic structure subject to a time-dependent temperature field, we can state a shape optimization problem. The shape functional we are interested in is the sum of two parts. The first one, similarly to the problem considered in [15], consists in the average over time of a differentiable functional, while the second takes into account the state of the structure at the final time t_f. In particular, we focus on a functional $\mathcal{C}_{\mathrm{T}}$ that evaluates the compliance of an elastic structure in the time interval $(0, t_f)$

$$\begin{aligned} \mathcal{C}_{\mathrm{T}}\left(\mathbf{u}, T; \Omega\right) = & \frac{\gamma}{t_f}\int_0^{t_f}\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_\Omega(t,\mathbf{x}), T_\Omega(t,\mathbf{x})\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega(t,\mathbf{x})\right) \mathrm{d}\mathbf{x}\,\mathrm{d}t \\ & + (1-\gamma)\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_\Omega(t_f,\mathbf{x}), T_\Omega(t_f,\mathbf{x})\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega(t_f,\mathbf{x})\right) \mathrm{d}\mathbf{x}, \end{aligned} \tag{2.29}$$

for any $\mathbf{u} \in \mathrm{L}^2\left([0,t_f], \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)^d\right)$ and any $T \in \mathrm{L}^2\left([0,t_f], \mathrm{H}^1(\Omega)\right)$. The parameter $\gamma \in [0,1]$ weights the two components of the functional $\mathcal{C}_{\mathrm{T}}$. As in section 1.1.1, we denote the reduced functional by $\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega) = \mathcal{C}_{\mathrm{T}}\left(\mathbf{u}_\Omega, T_\Omega; \Omega\right)$, where T_Ω and $\mathbf{u}_\Omega$ solve problems (2.26) and (2.28) respectively. Similarly to the optimization problems studied in part II, we consider an optimization problem where the objective is the minimization of the volume of a structure, under a constraint on the functional $\mathcal{C}_{\mathrm{T}}$.

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\ \text{minimizing the volume } \mathrm{Vol}(\Omega) \\ \text{under the constraint } \mathcal{C}_{\mathrm{T}}\left(\mathbf{u}_\Omega, T_\Omega; \Omega\right) \leq \tau, \text{ (see equation (2.29))}, \\ \text{where the temperature } T_\Omega \in \mathrm{L}^2\left([0,t_f], \mathrm{H}^1(\Omega)\right) \\ \text{solves the heat equation (2.26)}, \\ \text{and the displacement } \mathbf{u}_\Omega \in \mathrm{L}^2\left([0,t_f], \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)^d\right) \\ \text{solves the thermo-elasticity equation (2.28)}. \end{array}\right. \tag{2.30}$$

2.2.3 Computation of the shape derivative

In order to solve problem (2.30) using a gradient-based optimization method, it is necessary to compute the shape derivative $\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)$ of the functional $\widehat{\mathcal{C}_{\mathrm{T}}}$. We proceed as in section 2.1.2. First, we compute the sensitivity of $\widehat{\mathcal{C}_{\mathrm{T}}}$ to the shape Ω, then, we compute the Lagrangian derivatives of the temperature and the displacement. Finally, since the functional $\mathcal{C}_{\mathrm{T}}$ is not self-adjoint with respect to the state $(\mathbf{u}_\Omega, T_\Omega)$, we introduce two adjoint states R_Ω and $\mathbf{w}_\Omega$ aiming to erase the Lagrangian derivatives in the expression of $\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)$.

We refer to appendix A for the list of formulas used to compute the shape derivative of $\mathcal{C}_{\mathrm{T}}$. An analogous result can be obtained using the Céa fast derivation technique described in section 1.2.3, as proposed in [15].

Differentiation of the thermo-elastic compliance

We start by considering a bounded $\mathcal{C}^1$-domain Ω and a Lipschitz continuous vector field $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ such that $\|\boldsymbol{\theta}\|_{1,\infty} < 1$. We denote $\Omega_{\boldsymbol{\theta}} = (\mathbf{I}+\boldsymbol{\theta})\Omega$ the perturbed domain by the vector field $\boldsymbol{\theta}$, and $T_{\Omega_{\boldsymbol{\theta}}}$ and $\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}$ the temperature and displacement respectively obtained by solving equations (2.26) and (2.28) on $\Omega_{\boldsymbol{\theta}}$. By supposing that the portion of the boundary where the structure is clamped is not optimizable, we can consider that $\boldsymbol{\theta} = 0$ on Γ_{D}.

$$
\begin{aligned}
&\mathcal{C}_{\mathrm{T}}\left(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}, T_{\Omega_{\boldsymbol{\theta}}}; \Omega_{\boldsymbol{\theta}}\right) = \frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega_{\boldsymbol{\theta}}} \boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}\left(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}t), T_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}})\right) : \boldsymbol{\varepsilon}_{\widetilde{\mathbf{x}}}\left(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}})\right) \mathrm{d}\widetilde{\mathbf{x}}\,\mathrm{d}t \\
&\quad + (1-\gamma)\int_{\Omega_{\boldsymbol{\theta}}} \boldsymbol{\sigma}_{\widetilde{\mathbf{x}}}\left(\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f), T_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f)\right) : \boldsymbol{\varepsilon}_{\widetilde{\mathbf{x}}}\left(\mathbf{u}_{\Omega}(\widetilde{\mathbf{x}}, t_f)\right) \mathrm{d}\widetilde{\mathbf{x}} \\
=&\frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega_{\boldsymbol{\theta}}} \left(\mathbb{C}\nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}}) : \nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}}) - \alpha(T_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}}) - T_{\mathrm{ref}})\mathrm{div}_{\widetilde{\mathbf{x}}}\,\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(t,\widetilde{\mathbf{x}})\right)\,\mathrm{d}\widetilde{\mathbf{x}}\,\mathrm{d}t \\
&\quad + (1-\gamma)\int_{\Omega_{\boldsymbol{\theta}}} \left(\mathbb{C}\nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f) : \nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f) - \alpha(T_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f) - T_{\mathrm{ref}})\mathrm{div}_{\widetilde{\mathbf{x}}}\,\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}(\widetilde{\mathbf{x}}, t_f)\right)\,\mathrm{d}\widetilde{\mathbf{x}}.
\end{aligned}
$$

Using the results of section A.2, we can express $\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega_{\boldsymbol{\theta}})$ in terms of integrals on the reference domain.

$$
\begin{aligned}
\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega_{\boldsymbol{\theta}}) = &\frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega} \left(\mathbb{C}\left(\nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right) : \left(\nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right)\right)\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\mathrm{d}\mathbf{x}\mathrm{d}t \\
&- \frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega} \left(\alpha(\widetilde{T}_\Omega^{\boldsymbol{\theta}}(t,\mathbf{x}) - T_{\mathrm{ref}})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}} : \nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right)\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&+ (1-\gamma)\int_{\Omega} \left(\mathbb{C}\left(\nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t_f,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right) : \left(\nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t_f,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right)\right)\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x} \\
&- (1-\gamma)\int_{\Omega} \left(\alpha(\widetilde{T}_\Omega^{\boldsymbol{\theta}}(t_f,\mathbf{x}) - T_{\mathrm{ref}})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}} : \nabla\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}}(t_f,\mathbf{x})(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right)\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x},
\end{aligned}
\tag{2.31}
$$

where $\mathbb{C}$ is the fourth order Hooke tensor, and $\widetilde{\mathbf{u}}_\Omega^{\boldsymbol{\theta}} = \mathbf{u}_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I}+\boldsymbol{\theta})$ and $\widetilde{T}_\Omega^{\boldsymbol{\theta}} = T_{\Omega_{\boldsymbol{\theta}}} \circ (\mathbf{I}+\boldsymbol{\theta})$ are the results of pushing $\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}$ and $T_{\Omega_{\boldsymbol{\theta}}}$ on the reference domain. Then, we differentiate equation (2.31) with respect to $\boldsymbol{\theta}$ in $\mathbf{0}$ using the expressions in section A.3. Denoting $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$ and $\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})$ the Lagrangian derivatives of the displacement and the temperature we get

$$
\begin{aligned}
\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}}) = &2\frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega} \left(\mathbb{C}\nabla\mathbf{u}_\Omega(t,\mathbf{x}) : \left(\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) - \nabla\mathbf{u}_\Omega(t,\mathbf{x})\nabla\widehat{\boldsymbol{\theta}}\right)\right)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&- \frac{\gamma}{t_f}\int_0^{t_f}\int_{\Omega} \left(\alpha\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\mathrm{div}\,\mathbf{u}_\Omega + \alpha(T_\Omega - T_{\mathrm{ref}})\left(\mathrm{div}\,\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) - \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} : \nabla\mathbf{u}_\Omega\right)\right)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\gamma}{t_f}\int_0^{t_f}\int_\Omega (\mathbb{C}\nabla \mathbf{u}_\Omega(t,\mathbf{x}):\mathbf{u}_\Omega-\alpha(T_\Omega-T_{\text{ref}})\text{div}\,\mathbf{u}_\Omega)\,\text{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
&+2(1-\gamma)\int_\Omega\left(\mathbb{C}\nabla\mathbf{u}_\Omega(t_f,\mathbf{x}):\left(\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})-\nabla\mathbf{u}_\Omega(t_f,\mathbf{x})\nabla\widehat{\boldsymbol{\theta}}\right)\right)\,\mathrm{d}\mathbf{x}\\
&-(1-\gamma)\int_\Omega\left(\alpha\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\text{div}\,\mathbf{u}_\Omega+\alpha(T_\Omega-T_{\text{ref}})\left(\text{div}\,\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})-\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}:\nabla\mathbf{u}_\Omega\right)\right)\,\mathrm{d}\mathbf{x}\\
&+(1-\gamma)\int_\Omega(\mathbb{C}\nabla\mathbf{u}_\Omega(t_f,\mathbf{x}):\mathbf{u}_\Omega-\alpha(T_\Omega-T_{\text{ref}})\text{div}\,\mathbf{u}_\Omega)\,\text{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}.
\end{aligned}
\tag{2.32}
$$

Computation of the Lagrangian derivatives

In order to remove the terms of equation (2.32) containing the Lagrangian derivatives of the temperature and the displacement, we compute their expression by differentiating the state equations (2.26) and (2.28) with respect to the domain.

We start by multiplying the time-dependent heat equation (2.26) by a generic test function $S\in\mathcal{V}(0,t_f;\Omega)$, where

$$
\mathcal{V}(0,t_f;\Omega)=\mathcal{C}^1\left([0,t_f];\mathrm{L}^2(\Omega)\right)\cap\mathrm{L}^2\left((0,t_f);\mathrm{H}^1(\Omega)\right).
$$

After integrating over Ω, on the time interval $[0,t_f]$ and by parts, we obtain that T_Ω satisfies the following identity for all $S\in\mathcal{V}(0,t_f;\Omega)$

$$
\begin{aligned}
\int_0^{t_f}\int_\Omega\rho\frac{\partial T_\Omega}{\partial t}S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_\Omega k\,\nabla T_\Omega\cdot\nabla S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_{\partial\Omega}\beta T_\Omega S\,\mathrm{d}\mathbf{s}\,\mathrm{d}t\\
=\int_0^{t_f}\int_\Omega S\,Q\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_{\partial\Omega}\beta T_{\text{ext}}S\,\mathrm{d}\mathbf{s}\,\mathrm{d}t.
\end{aligned}
\tag{2.33}
$$

Let $T_{\Omega_{\boldsymbol{\theta}}}$ be the solution of problem (2.26) on the perturbed domain $\Omega_{\boldsymbol{\theta}}=(\mathbf{I}+\boldsymbol{\theta})\Omega$. Thus, considering a test function S still defined on the reference domain, (2.33) becomes

$$
\begin{aligned}
&\int_0^{t_f}\int_{\Omega_{\boldsymbol{\theta}}}\rho\frac{\partial T_{\Omega_{\boldsymbol{\theta}}}}{\partial t}S\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\,\mathrm{d}\widetilde{\mathbf{x}}\,\mathrm{d}t+\int_0^{t_f}\int_{\Omega_{\boldsymbol{\theta}}}k\nabla_{\widetilde{\mathbf{x}}}T_{\Omega_{\boldsymbol{\theta}}}\cdot\nabla_{\widetilde{\mathbf{x}}}\left(S\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\right)\,\mathrm{d}\widetilde{\mathbf{x}}\,\mathrm{d}t\\
&+\int_0^{t_f}\int_{(\mathbf{I}+\boldsymbol{\theta})\partial\Omega}\beta T_{\Omega_{\boldsymbol{\theta}}}\,S\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\,\mathrm{d}\widetilde{\mathbf{s}}\,\mathrm{d}t=\int_0^{t_f}\int_{\Omega_{\boldsymbol{\theta}}}Q\,S\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\,\mathrm{d}\widetilde{\mathbf{x}}\,\mathrm{d}t\\
&+\int_0^{t_f}\int_{(\mathbf{I}+\boldsymbol{\theta})\partial\Omega}\beta T_{\text{ext}}\,S\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\,\mathrm{d}\widetilde{\mathbf{s}}\,\mathrm{d}t.
\end{aligned}
\tag{2.34}
$$

By writing equation (2.34) with respect to the reference domain we get

$$
\begin{aligned}
&\int_0^{t_f}\int_\Omega\rho\frac{\partial\widetilde{T}_\Omega^{\boldsymbol{\theta}}}{\partial t}S\,\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_{\partial\Omega}\beta\widetilde{T}_\Omega^{\boldsymbol{\theta}}\,S\,\text{Jac}_{\Gamma_{\mathrm{N}}}(\mathbf{I}+\boldsymbol{\theta})\,\mathrm{d}\mathbf{s}\,\mathrm{d}t\\
&+\int_0^{t_f}\int_\Omega k\left((\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}}\nabla\widetilde{T}_\Omega^{\boldsymbol{\theta}}\right)\cdot\left((\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}}\nabla S\right)\,\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
&=\int_0^{t_f}\int_\Omega S\,Q\circ(\mathbf{I}+\boldsymbol{\theta})\,\left|\det\left(\mathbb{I}+\nabla\boldsymbol{\theta}\right)\right|\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_{\partial\Omega}\beta\,S\,T_{\text{ext}}\circ(\mathbf{I}+\boldsymbol{\theta})\,\text{Jac}_{\Gamma_{\mathrm{N}}}(\mathbf{I}+\boldsymbol{\theta})\,\mathrm{d}\mathbf{s}\,\mathrm{d}t.
\end{aligned}
\tag{2.35}
$$

The expression for the Lagrangian derivative $\dot{T}_\Omega$ is obtained by differentiating equation (2.35)

in $\boldsymbol{\theta} = 0$:

$$\begin{aligned}
&\int_0^{t_f}\int_\Omega \rho S\left(\frac{\partial \dot{T}_\Omega(\widehat{\boldsymbol{\theta}})}{\partial t} + \frac{\partial T_\Omega}{\partial t}(\operatorname{div}\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_\Omega k\nabla\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\cdot\nabla S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
&-\int_0^{t_f}\int_\Omega k\left(\nabla\widehat{\boldsymbol{\theta}} + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} - \operatorname{div}\widehat{\boldsymbol{\theta}}\right)\nabla T_\Omega\cdot\nabla S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_{\partial\Omega}\beta S\left(\dot{T}_\Omega + \operatorname{div}_{\partial\Omega}\widehat{\boldsymbol{\theta}}\right)\mathrm{d}\mathbf{s}\,\mathrm{d}t\\
&=\int_0^{t_f}\int_\Omega S\operatorname{div}(\widehat{\boldsymbol{\theta}}\,Q)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_{\partial\Omega}\beta S\left(\operatorname{div}_{\partial\Omega}(\widehat{\boldsymbol{\theta}}\,T_{\mathrm{ext}}) + \frac{\partial T_{\mathrm{ext}}}{\partial\mathbf{n}}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\,\mathrm{d}t
\end{aligned} \tag{2.36}$$

for any $\widehat{\boldsymbol{\theta}} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$. By rearranging the terms of equation (2.36), we obtain the following identity, satisfied for all $S \in \mathcal{V}(0,t_f;\Omega)$

$$\begin{aligned}
&\int_0^{t_f}\int_\Omega \rho\frac{\partial\dot{T}_\Omega}{\partial t}S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_\Omega k\,\nabla\dot{T}_\Omega\cdot\nabla S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_{\partial\Omega}\beta\dot{T}_\Omega S\,\mathrm{d}\mathbf{s}\,\mathrm{d}t\\
&=-\int_0^{t_f}\int_\Omega \rho\frac{\partial T_\Omega}{\partial t}\,S(\operatorname{div}\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_\Omega k\left(\nabla\widehat{\boldsymbol{\theta}} + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\right)\nabla T_\Omega\cdot\nabla S\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
&-\int_0^{t_f}\int_\Omega k\nabla T_\Omega\cdot\nabla S\,(\operatorname{div}\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t - \int_0^{t_f}\int_{\partial\Omega}\beta T_\Omega S\,(\operatorname{div}_{\partial\Omega}\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{s}\,\mathrm{d}t\\
&+\int_0^{t_f}\int_\Omega S\operatorname{div}(\widehat{\boldsymbol{\theta}}\,Q)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t + \int_0^{t_f}\int_{\partial\Omega}\beta S\left(\operatorname{div}_{\partial\Omega}(\widehat{\boldsymbol{\theta}}\,T_{\mathrm{ext}}) + \frac{\partial T_{\mathrm{ext}}}{\partial\mathbf{n}}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\,\mathrm{d}t.
\end{aligned} \tag{2.37}$$

Moreover, thanks to the initial condition $T_\Omega(0,\mathbf{x}) = T_{\mathrm{ref}}$, we have that $\dot{T}_\Omega(0,\mathbf{x})$ 0 on Ω.

In order to identify the equation solved bu $\dot{\mathbf{u}}_\Omega$ we proceed in a similar way. Since the elasticity equation is supposed to be quasi-static, the following equations hold for any $t \in [0,t_f]$. We start by considering a generic test function $\mathbf{v} \in \mathcal{W}(0,t_f;\Omega)$, where

$$\mathcal{W}(0,t_f;\Omega) = \mathcal{C}^1\left([0,t_f];\mathrm{L}^2(\Omega)^d\right)\cap\mathrm{L}^2\left((0,t_f);\mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)^d\right).$$

We multiply the first equation of problem (2.28) by $\mathbf{v}$ and we integrate over Ω and by parts, and we obtain the following identity, which hold for all $t \in [0,t_f]$

$$\int_\Omega \mathbb{C}\nabla\mathbf{u}_\Omega(\mathbf{x},t) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\mathbf{x} = \int_\Omega \alpha(T_\Omega - T_{\mathrm{ref}})\,\operatorname{div}\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}}\mathbf{g}\cdot\mathbf{v}\,\mathrm{d}\mathbf{s}. \tag{2.38}$$

Denoting $\mathbf{u}_{\Omega_{\boldsymbol{\theta}}}$ the displacement solving (2.26) on the perturbed domain $\Omega_{\boldsymbol{\theta}} = (\mathbf{I}+\boldsymbol{\theta})\Omega$ and considering a test function $\mathbf{v}$ defined on Ω, (2.38) is written as

$$\begin{aligned}
&\int_{\Omega_{\boldsymbol{\theta}}}\mathbb{C}\nabla_{\widetilde{\mathbf{x}}}\mathbf{u}_{\Omega_{\boldsymbol{\theta}}} : \nabla_{\widetilde{\mathbf{x}}}\left(\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\right)\mathrm{d}\widetilde{\mathbf{x}} = \int_{\Omega_{\boldsymbol{\theta}}}\alpha(T_{\Omega_{\boldsymbol{\theta}}} - T_{\mathrm{ref}})\,\operatorname{div}_{\widetilde{\mathbf{x}}}(\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1})\,\mathrm{d}\widetilde{\mathbf{x}}\\
&+\int_{\Omega_{\boldsymbol{\theta}}}\mathbf{f}\cdot\left(\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\right)\mathrm{d}\widetilde{\mathbf{x}} + \int_{(\mathbf{I}+\boldsymbol{\theta})\Gamma_{\mathrm{N}}}\mathbf{g}\cdot\left(\mathbf{v}\circ(\mathbf{I}+\boldsymbol{\theta})^{-1}\right)\mathrm{d}\widetilde{\mathbf{s}}.
\end{aligned} \tag{2.39}$$

Using the formulas of section A.2 we express equation (2.39) in terms of integrals on the reference domain

$$\begin{aligned}
&\int_\Omega \mathbb{C}\left(\nabla\widetilde{\mathbf{u}}^{\boldsymbol{\theta}}_\Omega(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right) : \left(\nabla\mathbf{v}(\mathbb{I}+\nabla\boldsymbol{\theta})^{-1}\right)\,|\det(\mathbb{I}+\nabla\boldsymbol{\theta})|\,\mathrm{d}\mathbf{x}\\
&=\int_\Omega \alpha\,(\widetilde{T}^{\boldsymbol{\theta}}_\Omega - T_{\mathrm{ref}})\left((\mathbb{I}+\nabla\boldsymbol{\theta})^{-\mathrm{T}} : \nabla\mathbf{v}\right)\,|\det(\mathbb{I}+\nabla\boldsymbol{\theta})|\,\mathrm{d}\mathbf{x} + \int_\Omega(\mathbf{f}\circ(\mathbf{I}+\boldsymbol{\theta}))\cdot\mathbf{v}\,|\det(\mathbb{I}+\nabla\boldsymbol{\theta})|\,\mathrm{d}\mathbf{x}\\
&+\int_{\Gamma_{\mathrm{N}}}(\mathbf{g}\circ(\mathbf{I}+\boldsymbol{\theta}))\,\operatorname{Jac}_{\Gamma_{\mathrm{N}}}(\mathbf{I}+\boldsymbol{\theta})\,\mathrm{d}\mathbf{s}.
\end{aligned} \tag{2.40}$$

Then, we differentiate (2.40) in $\boldsymbol{\theta} = \mathbf{0}$ using the expressions formulated in section A.3. Thus, for all $\widehat{\boldsymbol{\theta}} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ such that $\boldsymbol{\theta} = \mathbf{0}$ on Γ_{D} we have

$$\begin{aligned}
&\int_\Omega \left(\mathbb{C}\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{v} - \mathbb{C}(\nabla\mathbf{u}\nabla\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{v} - \mathbb{C}\nabla\mathbf{u} : (\nabla\mathbf{v}\nabla\widehat{\boldsymbol{\theta}}) + (\mathbb{C}\nabla\mathbf{u} : \nabla\mathbf{v})\,(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\right)\,\mathrm{d}\mathbf{x} \\
=&\int_\Omega \alpha\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\mathrm{div}\,\mathbf{v}\,\mathrm{d}\mathbf{x} - \int_\Omega \alpha(T_\Omega - T_{\mathrm{ref}})\left(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} : \nabla\mathbf{v}\right)\,\mathrm{d}\mathbf{x} + \int_\Omega \alpha(T_\Omega - T_{\mathrm{ref}})(\mathrm{div}\,\mathbf{v})(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} \\
&+ \int_\Omega \mathbf{v}\cdot\mathrm{div}\left(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\right)\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}} \mathbf{v}\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right)\,\mathrm{d}\mathbf{s} + \int_{\Gamma_{\mathrm{N}}}\left(\mathbf{v}\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}.
\end{aligned}$$

The equation above can be reformulated as follows, highlighting the terms containing the Lagrangian derivative $\dot{\mathbf{u}}_\Omega$

$$\begin{aligned}
\int_\Omega \mathbb{C}\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\mathbf{x} =& \int_\Omega \left(\mathbb{C}(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}) : \nabla\mathbf{v} + \mathbb{C}\nabla\mathbf{u}_\Omega : (\nabla\mathbf{v}\nabla\widehat{\boldsymbol{\theta}}) - (\mathbb{C}\nabla\mathbf{u}_\Omega : \nabla\mathbf{v})\,(\mathrm{div}\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{x} \\
&+ \int_\Omega \alpha\left(\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\mathrm{div}\,\mathbf{v} - (T_\Omega - T_{\mathrm{ref}})\left(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} : \nabla\mathbf{v}\right) + (T_\Omega - T_{\mathrm{ref}})(\mathrm{div}\,\mathbf{v})(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\right)\,\mathrm{d}\mathbf{x} \\
&+ \int_\Omega \mathbf{v}\cdot\mathrm{div}\left(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\right)\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}}\left(\mathbf{v}\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right) + \left(\mathbf{v}\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\,\mathrm{d}\mathbf{s}.
\end{aligned} \tag{2.41}$$

Introduction of the adjoint states

In order to cancel the Lagrangian derivatives in the expression (2.32), we introduce the adjoint states R_Ω and $\mathbf{w}_\Omega$, related to the temperature and the elastic displacement respectively.

We start by considering the adjoint state $\mathbf{w}_\Omega \in \mathcal{V}(0, t_f, \Omega)$ solving the following quasi-static differential equation for all $t \in [0, t_f]$

$$\left\{\begin{array}{rcll}
-\mathrm{div}(\mathbb{C}\nabla\mathbf{w}_\Omega) &=& -\mathrm{div}(\mathbb{C}\nabla\mathbf{u}_\Omega) + \frac{\alpha}{2}\nabla T_\Omega & \text{in } \Omega, \\
(\mathbb{C}\nabla\mathbf{w}_\Omega)\mathbf{n} &=& \mathbf{g} + (T_\Omega - T_{\mathrm{ref}})\mathbf{n} & \text{on } \Gamma_{\mathrm{N}}, \\
(\mathbb{C}\nabla\mathbf{w}_\Omega)\mathbf{n} &=& (T_\Omega - T_{\mathrm{ref}})\mathbf{n} & \text{on } \Gamma_0, \\
\mathbf{w}_\Omega &=& \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}.
\end{array}\right. \tag{2.42}$$

In order to inject the adjoint state in the expression (2.32) of the shape derivative of $\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)$, we multiply the first equation of (2.42) by $\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) \in \mathcal{W}(0, t_f; \Omega)$, and we integrate over Ω, obtaining

$$\int_\Omega \mathbb{C}\nabla\mathbf{w}_\Omega : \nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbb{C}\nabla\mathbf{u}_\Omega : \nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x} - \int_\Omega \frac{\alpha}{2}\nabla T_\Omega\cdot\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}. \tag{2.43}$$

We remark that all terms in the right-hand side of equation (2.43) appear in the expression (2.32) of the shape derivative of the compliance. Moreover, the left-hand side of equation (2.43) coincides with the left-hand side of (2.41) for $\mathbf{v} = \mathbf{w}_\Omega$ because, thanks to the definition and properties of the Hooke elasticity tensor,

$$\mathbb{C}\nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}}) : \boldsymbol{\varepsilon}(\mathbf{v}) = \mathbb{C}\nabla\mathbf{v} : \nabla\dot{\mathbf{u}}_\Omega(\widehat{\boldsymbol{\theta}})$$

Since $\mathbf{w}_\Omega \in \mathcal{W}(0,t_f,\Omega)$, the identity (2.41) holds, and the shape derivative of $\widehat{\mathcal{C}_{\mathrm{T}}}$ becomes

$$
\begin{aligned}
\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}}) = \frac{\gamma}{t_f}\int_0^{t_f}\Bigg(\int_\Omega\Big(\alpha\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{div}(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)-\alpha(T_\Omega-T_{\mathrm{ref}})\Big(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}:\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\Big)\Big)\,\mathrm{d}\mathbf{x}\\
-\int_\Omega(\mathbb{C}\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega):\nabla\mathbf{u}_\Omega-\alpha(T_\Omega-T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega-\mathbf{u}_\Omega))\,\mathrm{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\\
+\int_\Omega\Big(\mathbb{C}(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}):\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)+\mathbb{C}\nabla\mathbf{u}_\Omega:(\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\nabla\widehat{\boldsymbol{\theta}})(\mathrm{div}\widehat{\boldsymbol{\theta}})\Big)\,\mathrm{d}\mathbf{x}\\
+2\int_\Omega\mathbf{w}_\Omega\cdot\mathrm{div}\Big(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\Big)\,\mathrm{d}\mathbf{x}+2\int_{\Gamma_{\mathrm{N}}}\Big(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\Big(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\Big)+\Big(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\Big)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\Big)\,\mathrm{d}\mathbf{s}\Bigg)\mathrm{d}t\\
+(1-\gamma)\Bigg(\int_\Omega\Big(\alpha\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{div}(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)-\alpha(T_\Omega-T_{\mathrm{ref}})\Big(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}:\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\Big)\Big)\,\mathrm{d}\mathbf{x}\\
-\int_\Omega(\mathbb{C}\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega):\nabla\mathbf{u}_\Omega-\alpha(T_\Omega-T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega-\mathbf{u}_\Omega))\,\mathrm{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\\
+\int_\Omega\Big(\mathbb{C}(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}):\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)+\mathbb{C}\nabla\mathbf{u}_\Omega:(\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\nabla\widehat{\boldsymbol{\theta}})(\mathrm{div}\widehat{\boldsymbol{\theta}})\Big)\,\mathrm{d}\mathbf{x}\\
+2\int_\Omega\mathbf{w}_\Omega\cdot\mathrm{div}\Big(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\Big)\,\mathrm{d}\mathbf{x}+2\int_{\Gamma_{\mathrm{N}}}\Big(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\Big(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\Big)+\Big(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\Big)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\Big)\,\mathrm{d}\mathbf{s}\Bigg)\Bigg|_{t=t_f}.
\end{aligned}
\tag{2.44}
$$

For the purpose of canceling the dependence of the expression (2.44) from $\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})$, we introduce the adjoint state $R_\Omega \in \mathcal{V}(0,t_f;\Omega)$ solving the following differential equation backwards in time

$$
\left\{\begin{array}{rcll}
-\rho\frac{\partial R_\Omega}{\partial t}-\mathrm{div}\,(k\nabla R_\Omega) &=& \gamma\,\alpha\,\mathrm{div}\,(2\mathbf{w}_\Omega-\mathbf{u}_\Omega) & \text{in } [0,t_f)\times\Omega,\\
k\,\nabla R_\Omega\cdot\mathbf{n}+\beta R_\Omega &=& 0 & \text{on } [0,t_f)\times\partial\Omega,\\
R_\Omega(t_f,\mathbf{x}) &=& \frac{(1-\gamma)}{\rho}\,\alpha\,\mathrm{div}\,(2\mathbf{w}_\Omega(t_f,\mathbf{x})-\mathbf{u}_\Omega(t_f,\mathbf{x})) & \text{in } \Omega.
\end{array}\right.
\tag{2.45}
$$

We multiply the first equation of (2.45) by $\dot{T}_\Omega(\widehat{\boldsymbol{\theta}}) \in \mathcal{V}(0,t_f;\Omega)$, and we integrate over Ω and $[0,t_f]$, as well as by parts in space and time, and we obtain

$$
-\int_0^{t_f}\int_\Omega\Big(\rho\frac{\partial R_\Omega}{\partial t}\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})+\mathrm{div}\,(k\nabla R_\Omega)\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\Big)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t=\gamma\int_0^{t_f}\int_\Omega\alpha\,\mathrm{div}\,(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\,\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t;
$$

$$
\begin{aligned}
-\int_\Omega\rho\Big(R_\Omega\,\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\Big)\Big|_{t=t_f}\mathrm{d}\mathbf{x}+\int_0^{t_f}\int_\Omega\rho\frac{\partial\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})}{\partial t}R_\Omega\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_\Omega k\,\nabla R_\Omega\cdot\nabla\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
-\int_0^{t_f}\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}k\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\frac{\partial R_\Omega}{\partial\mathbf{n}}\,\mathrm{d}\mathbf{x}\,\mathrm{d}t=\gamma\int_0^{t_f}\int_\Omega\alpha\,\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\mathrm{div}\,(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\int_0^{t_f}\int_\Omega\rho\frac{\partial\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})}{\partial t}R_\Omega\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_\Omega k\,\nabla R_\Omega\cdot\nabla\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^{t_f}\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\beta\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\,R_\Omega\,\mathrm{d}\mathbf{x}\,\mathrm{d}t\\
&=\gamma\int_0^{t_f}\int_\Omega\alpha\,\dot{T}_\Omega(\widehat{\boldsymbol{\theta}})\mathrm{div}\,(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\,\mathrm{d}\mathbf{x}\,\mathrm{d}t,+(1-\gamma)\int_\Omega\alpha\,(\mathrm{div}(2\mathbf{w}_\Omega(t_f,\mathbf{x})-\mathbf{u}_\Omega(t_f,\mathbf{x})))\Big|_{t=t_f}\mathrm{d}\mathbf{x}.
\end{aligned}
\tag{2.46}
$$

By combining equations (2.37) and (2.46), and injecting the result in the expression (2.44), the shape derivative of the compliance can be written as

$$
\begin{aligned}
\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}}) = \frac{\gamma}{t_f}\int_0^{t_f}\Bigg(&-\int_\Omega\left(\alpha(T_\Omega - T_{\mathrm{ref}})\left(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} : \nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\right)\right)\mathrm{d}\mathbf{x}\\
&-\int_\Omega\left(\mathbb{C}\nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega) : \nabla\mathbf{u}_\Omega - \alpha(T_\Omega - T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\right)\mathrm{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\\
&+\int_\Omega\left(\mathbb{C}(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}) : \nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega) + \mathbb{C}\nabla\mathbf{u}_\Omega : (\nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\nabla\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{x}\\
&+2\int_\Omega\mathbf{w}_\Omega\cdot\mathrm{div}\left(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\right)\mathrm{d}\mathbf{x} + 2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right) + \left(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\mathrm{d}t\\
+(1-\gamma)\Bigg(&-\int_\Omega\left(\alpha(T_\Omega - T_{\mathrm{ref}})\left(\nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}} : \nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\right)\right)\mathrm{d}\mathbf{x}\\
&-\int_\Omega\left(\mathbb{C}\nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega) : \nabla\mathbf{u}_\Omega - \alpha(T_\Omega - T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\right)\mathrm{div}\,\widehat{\boldsymbol{\theta}}\,\mathrm{d}\mathbf{x}\\
&+\int_\Omega\left(\mathbb{C}(\nabla\mathbf{u}_\Omega\nabla\widehat{\boldsymbol{\theta}}) : \nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega) + \mathbb{C}\nabla\mathbf{u}_\Omega : (\nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\nabla\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{x}\\
&+2\int_\Omega\mathbf{w}_\Omega\cdot\mathrm{div}\left(\mathbf{f}\otimes\widehat{\boldsymbol{\theta}}\right)\mathrm{d}\mathbf{x} + 2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right) + \left(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\Bigg|_{t=t_f}\\
+\int_0^{t_f}\Bigg(&-\int_\Omega\left(\rho\frac{\partial T_\Omega}{\partial t}R_\Omega(\mathrm{div}\,\widehat{\boldsymbol{\theta}}) + k\left(\nabla\widehat{\boldsymbol{\theta}} + \nabla\widehat{\boldsymbol{\theta}}^{\mathrm{T}}\right)\nabla T_\Omega\cdot\nabla R_\Omega - k\nabla T_\Omega\cdot\nabla R_\Omega\,(\mathrm{div}\,\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{x}\\
&+\int_\Omega R_\Omega\,\mathrm{div}\,(\widehat{\boldsymbol{\theta}}\,Q)\,\mathrm{d}\mathbf{x} + \int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\beta R_\Omega\left(\mathrm{div}_{\partial\Omega}(\widehat{\boldsymbol{\theta}}\,T_{\mathrm{ext}}) + \frac{\partial T_{\mathrm{ext}}}{\partial\mathbf{n}}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}) - T_\Omega\,(\mathrm{div}_{\partial\Omega}\,\widehat{\boldsymbol{\theta}})\right)\mathrm{d}\mathbf{s}\Bigg)\mathrm{d}t.
\end{aligned}
\tag{2.47}
$$

Looking at the equations (2.42) and (2.45), we remark that the adjoint states are weakly coupled with one another, but the dependence is reversed with respect to the relation between the temperature and the displacement fields. Indeed, in section 2.2.1 we supposed that the equation for $\mathbf{u}_\Omega$ relies on the solution of T_Ω, while the expression of the adjoint state for the temperature R_Ω depends on the adjoint state for the displacement $\mathbf{w}_\Omega$.

The expression (2.47) of $\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}})$ is the volume form of the shape derivative of the compliance. However, theorem 1.7 ensures that, if $\Omega \subset \mathbb{R}^d$ is a $\mathcal{C}^1$ domain, $\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}})$ can be expressed as an integral over $\partial\Omega$ depending only on the component of $\widehat{\boldsymbol{\theta}}$ normal to $\partial\Omega$. We suppose that the part Γ_{D} of the boundary is non-optimizable, therefore we restrict the Lipschitz continuous vector field $\widehat{\boldsymbol{\theta}} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ to be such that $\widehat{\boldsymbol{\theta}} = \mathbf{0}$ on Γ_{D}. With the help of the formulas of section A.4 and considering the boundary conditions of problems (2.26) and (2.28) we obtain

$$
\begin{aligned}
\mathrm{D}\widehat{\mathcal{C}_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}}) = \frac{\gamma}{t_f}\int_0^{t_f}\Bigg(&-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\alpha(T_\Omega - T_{\mathrm{ref}})\mathbf{n}\cdot\left(2\frac{\partial\mathbf{w}_\Omega}{\partial\mathbf{n}} - \frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\mathbb{C}\nabla(2\mathbf{w}_\Omega - \mathbf{u}_\Omega) : \nabla\mathbf{u}_\Omega - \alpha(T_\Omega - T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega - \mathbf{u}_\Omega)\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\cdot\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} + 2\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\mathbf{f}\cdot\mathbf{w}_\Omega + (T_\Omega - T_{\mathrm{ref}})\mathbf{n}\cdot\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}
\end{aligned}
$$

$$
\begin{aligned}
&+2\int_{\Gamma_N}\left(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_N}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right)+\left(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\mathrm{d}t\\
&+(1-\gamma)\Bigg(-\int_{\Gamma_N\cup\Gamma_0}\left(\alpha(T_\Omega-T_{\mathrm{ref}})\mathbf{n}\cdot\left(2\frac{\partial\mathbf{w}_\Omega}{\partial\mathbf{n}}-\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&-\int_{\Gamma_N\cup\Gamma_0}\Big(\mathbb{C}\nabla(2\mathbf{w}_\Omega-\mathbf{u}_\Omega):\nabla\mathbf{u}_\Omega-\alpha(T_\Omega-T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}_\Omega-\mathbf{u}_\Omega)\Big)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_N}\left(\mathbf{g}\cdot\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}+2\int_{\Gamma_N\cup\Gamma_0}\left(\mathbf{f}\cdot\mathbf{w}_\Omega+(T_\Omega-T_{\mathrm{ref}})\mathbf{n}\cdot\frac{\partial\mathbf{u}_\Omega}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_N}\left(\mathbf{w}_\Omega\cdot\mathrm{div}_{\Gamma_N}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right)+\left(\mathbf{w}_\Omega\cdot\frac{\partial\mathbf{g}}{\partial\mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\Bigg|_{t=t_f}\\
&+\int_0^{t_f}\Bigg(\int_{\Gamma_N\cup\Gamma_0}\left(k(\nabla T_\Omega\cdot\nabla R_\Omega)-\rho\frac{\partial T_\Omega}{\partial t}R_\Omega-2\frac{\beta^2}{k}R_\Omega(T_\Omega-T_{\mathrm{ref}})+R_\Omega\,Q\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+\int_{\Gamma_N\cup\Gamma_0}\beta R_\Omega\left(\mathrm{div}_{\partial\Omega}(\widehat{\boldsymbol{\theta}}\,T_{\mathrm{ext}})+\frac{\partial T_{\mathrm{ext}}}{\partial\mathbf{n}}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})-T_\Omega\left(\mathrm{div}_{\partial\Omega}\,\widehat{\boldsymbol{\theta}}\right)\right)\mathrm{d}\mathbf{s}\Bigg)\mathrm{d}t. \qquad (2.48)
\end{aligned}
$$

2.2.4 Discretization of the state and adjoint equations

In order to solve numerically the optimization problem (2.30) using the techniques presented in section 1.3, it is necessary to discretize and solve the state problems (2.26) and (2.28) for the temperature and the displacement, as well as the adjoint problems (2.42) and (2.45), at each step of the optimization algorithm. Let $D\subset\mathbb{R}^d$ be a computational domain such that any admissible domain in $\mathcal{S}_{\mathrm{adm}}$ is contained in D, and let $\Omega\subset D$ be an admissible domain. We consider a mesh $\mathcal{T}_D$ covering the entire domain D, and such that Ω is explicitly discretized by a submesh $\mathcal{T}_\Omega$. Moreover, we discretize the time interval $[0,t_f]$ by dividing it in N_t sub-intervals, each with length $\Delta t=\frac{t_f}{N_t}$. We introduce the finite-dimensional spaces $\mathcal{V}^h(\Omega)$ and $\mathcal{W}^h(\Omega)$ discretizing $\mathcal{V}(\Omega)$ and $\mathcal{W}(\Omega)$ respectively by $\mathbb{P}^1$ Lagrangian finite elements:

$$
\begin{aligned}
\mathcal{V}^h(\Omega)&=\left\{S\in\mathrm{H}^1(\Omega)\ :\ S\Big|_K\in\mathbb{P}^1(K)\text{ for all element }K\in\mathcal{T}_D\right\};\\
\mathcal{W}^h(\Omega)&=\left\{\mathbf{v}\in\mathrm{H}^1_{\Gamma_D}(\Omega)^d\ :\ \mathbf{v}\Big|_K\in\mathbb{P}^1(K)^d\text{ for all element }K\in\mathcal{T}_D\right\}.
\end{aligned}
$$

As first step, we aim to solve the time-dependent thermal equation. We start by writing the discretization in space of the variational form of equation (2.26).

$$
\left|\begin{array}{l}
\text{Find } T^h\in\mathcal{C}^1\left([0,t_f];\mathcal{V}^h(\Omega)\right)\cap\mathrm{L}^2\left((0,t_f);\mathcal{V}^h(\Omega)\right)\\
\text{ such that for all } S^h\in\mathcal{V}^h(\Omega)\\
\dfrac{\mathrm{d}}{\mathrm{d}t}\displaystyle\int_\Omega\rho\,T^h\,S^h\,\mathrm{d}\mathbf{x}+\int_\Omega k\nabla T^h\cdot\nabla S^h\,\mathrm{d}\mathbf{x}+\int_{\Gamma_N\cup\Gamma_0}\beta\,T^h\,S^h\,\mathrm{d}\mathbf{s}=\int_\Omega S^h\,Q\,\mathrm{d}\mathbf{x}+\int_{\Gamma_N\cup\Gamma_0}T_{\mathrm{ref}}\,S^h\,\mathrm{d}\mathbf{s},\\
\text{with the initial condition } T^h(0,\cdot)=T_0.
\end{array}\right. \qquad (2.49)
$$

In order to solve problem (2.49) in time we denote $T^h_i\in\mathcal{V}^h(\Omega)$ the discretization of $T_\Omega(i\Delta t,\cdot)$ for all $i\in\{0,\dots,N_t\}$. We consider as initial condition $T^h_0=T_0$, and we compute the subsequent

temperature fields by an implicit Euler method

$$\left|\begin{array}{l}\text{For any } i \in \{1,\dots,N_t\}, \text{ find } T_i^h \in \mathcal{V}^h(\Omega)\\ \text{such that for all } S^h \in \mathcal{V}^h(\Omega)\\ \displaystyle\int_\Omega \rho \left(\frac{T_i^h - T_{i-1}^h}{\Delta t}\right) S^h \,\mathrm{d}\mathbf{x} + \int_\Omega k \nabla T_i^h \cdot \nabla S^h \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}\cup\Gamma_0} \beta\, T_i^h\, S^h \,\mathrm{d}\mathbf{s}\\ \qquad\qquad = \displaystyle\int_\Omega S^h\, Q \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}\cup\Gamma_0} T_\mathrm{ref}\, S^h \,\mathrm{d}\mathbf{s},\\ \text{knowing the initial condition } T_0^h = T_0.\end{array}\right. \tag{2.50}$$

Since the state equation (2.28) for the displacement $\mathbf{u}_\Omega$ is quasi-static, we can compute $\mathbf{u}_i^h \in \mathcal{W}^h(\Omega)$ approximating $\mathbf{u}_\Omega(i\Delta t, \cdot)$ for all $i \in \{0,\dots,N_t\}$ by solving the following problem by a finite-element method

$$\left|\begin{array}{l}\text{For any } i \in \{0,\dots,N_t\},\\ \text{find } \mathbf{u}_i^h \in \mathcal{W}^h(\Omega) \text{ such that for all } \mathbf{v}^h \in \mathcal{W}^h(\Omega)\\ \displaystyle\int_\Omega \mathbb{C}\nabla \mathbf{u}_i^h : \nabla \mathbf{v}^h \,\mathrm{d}\mathbf{x} = \int_\Omega (T_i^h - T_\mathrm{ref})(\operatorname{div} \mathbf{v}^h)\,\mathrm{d}\mathbf{x} + \int_\Omega \mathbf{f}\cdot\mathbf{v}^h \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}} \mathbf{g}\cdot\mathbf{v}^h \,\mathrm{d}\mathbf{s}.\end{array}\right. \tag{2.51}$$

Having discretized the state in space and time, we can express the discrete thermal compliance (2.29) by replacing each state variable with its approximation, and the time integral with a Riemann sum

$$\begin{aligned}\widehat{\mathcal{C}_\mathrm{T}^h}(\Omega) =& \mathcal{C}_\mathrm{T}^h\left((\mathbf{u}_0^h,\dots,\mathbf{u}_{N_t}^h),(T_0^h,\dots,T_{N_t}^h);\Omega\right) = \frac{\gamma}{t_f}\sum_{i=1}^{N_t}\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_i^h, T_i^h\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_i^h\right)\mathrm{d}\mathbf{x}\\ &+ (1-\gamma)\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_{N_t}^h, T_{N_t}^h\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{N_t}^h\right)\mathrm{d}\mathbf{x}.\end{aligned} \tag{2.52}$$

The computation of a discretized shape derivative of the compliance requires the discretization of the adjoint states is done similarly to the forward equations. Since equation equation (2.42) for $\mathbf{w}_\Omega$ is quasi-static and it depends on both $\mathbf{u}_\Omega$ and T_Ω, we can discretize it as follows

$$\left|\begin{array}{l}\text{For any } i \in \{0,\dots,N_t\},\\ \text{find } \mathbf{w}_i^h \in \mathcal{W}^h(\Omega) \text{ such that for all } \mathbf{v}^h \in \mathcal{W}^h(\Omega)\\ \displaystyle\int_\Omega \mathbb{C}\nabla \mathbf{w}_i^h : \nabla \mathbf{v}^h \,\mathrm{d}\mathbf{x} = \int_\Omega \mathbb{C}\nabla \mathbf{u}_i^h : \nabla \mathbf{v}^h \,\mathrm{d}\mathbf{x} + \int_\Omega \frac{\alpha}{2}\nabla T_i^h \cdot \mathbf{v}^h \,\mathrm{d}\mathbf{x}.\end{array}\right. \tag{2.53}$$

For the adjoint state R_Ω we discretize in space the variational form of problem (2.45), and we consider an implicit Euler scheme backwards in time for the integration in time.

$$\left|\begin{array}{l}\text{For any } i \in \{0,\dots,(N_t-1)\},\\ \text{find } R_i^h \in \mathcal{V}^h(\Omega) \text{ such that for all } S^h \in \mathcal{V}^h(\Omega)\\ \displaystyle -\int_\Omega \rho \left(\frac{R_i^h - R_{i+1}^h}{\Delta t}\right) S^h \,\mathrm{d}\mathbf{x} + \int_\Omega k\nabla R_i^h \cdot \nabla S^h \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}\cup\Gamma_0} \beta\, R_i^h\, S^h \,\mathrm{d}\mathbf{s}\\ \qquad\qquad = \displaystyle\int_\Omega \gamma\,\alpha \operatorname{div}\left(2\mathbf{w}_i^h - \mathbf{u}_i^h\right) S^h \,\mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{N}\cup\Gamma_0} T_\mathrm{ref}\, S^h \,\mathrm{d}\mathbf{s},\\ \text{knowing the final condition } R_{N_t}^h = \dfrac{(1-\gamma)}{\rho}\,\alpha \operatorname{div}\left(2\mathbf{w}_{N_t}^h - \mathbf{u}_{N_t}^h\right).\end{array}\right. \tag{2.54}$$

The final condition on the adjoint state is fixed. Thus, we start by computing $R^h_{N_t-1}$, and we proceed backwards towards R^h_0.

Having computed the discrete solutions of problems (2.50), (2.51), (2.53), and (2.54) for all timestep $i = 0, \dots, N_t$, we can finally express the discretization of the shape derivative of the functional $\widehat{\mathcal{C}_{\mathrm{T}}}$ in Ω as

$$
\begin{aligned}
\mathrm{D}\widehat{\mathcal{C}^h_{\mathrm{T}}}(\Omega)(\widehat{\boldsymbol{\theta}}) =& \frac{\gamma}{t_f}\sum_{i=1}^{N_t}\Bigg(-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\alpha(T^h_i - T_{\mathrm{ref}})\mathbf{n}\cdot\left(2\frac{\partial \mathbf{w}^h_i}{\partial \mathbf{n}} - \frac{\partial \mathbf{u}^h_i}{\partial \mathbf{n}}\right)\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\Big(\mathbb{C}\nabla(2\mathbf{w}^h_i - \mathbf{u}^h_i) : \nabla\mathbf{u}^h_i - \alpha(T^h_i - T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}^h_i - \mathbf{u}^h_i)\Big)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\cdot\frac{\partial \mathbf{u}^h_i}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} + 2\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\mathbf{f}\cdot\mathbf{w}^h_i + (T^h_i - T_{\mathrm{ref}})\mathbf{n}\cdot\frac{\partial \mathbf{u}_\Omega}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{w}^h_i\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right) + \left(\mathbf{w}_\Omega\cdot\frac{\partial \mathbf{g}}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\\
&+(1-\gamma)\Bigg(-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\alpha(T^h_i - T_{\mathrm{ref}})\mathbf{n}\cdot\left(2\frac{\partial \mathbf{w}^h_i}{\partial \mathbf{n}} - \frac{\partial \mathbf{u}^h_i}{\partial \mathbf{n}}\right)\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&-\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\Big(\mathbb{C}\nabla(2\mathbf{w}^h_i - \mathbf{u}^h_i) : \nabla\mathbf{u}_\Omega - \alpha(T^h_i - T_{\mathrm{ref}})\mathrm{div}(2\mathbf{w}^h_i - \mathbf{u}^h_i)\Big)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\cdot\frac{\partial \mathbf{u}^h_{N_t}}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} + 2\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(\mathbf{f}\cdot\mathbf{w}_\Omega + (T^h_{N_t} - T_{\mathrm{ref}})\mathbf{n}\cdot\frac{\partial \mathbf{u}^h_{N_t}}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+2\int_{\Gamma_{\mathrm{N}}}\left(\mathbf{w}^h_{N_t}\cdot\mathrm{div}_{\Gamma_{\mathrm{N}}}\left(\mathbf{g}\otimes\widehat{\boldsymbol{\theta}}\right) + \left(\mathbf{w}^h_{N_t}\cdot\frac{\partial \mathbf{g}}{\partial \mathbf{n}}\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\right)\mathrm{d}\mathbf{s}\Bigg)\Bigg|_{t=t_f}\\
&+\sum_{i=1}^{N_t}\Bigg(\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\left(k(\nabla T^h_i\cdot\nabla R^h_i) - \rho\frac{T^h_i - T^h_{i-1}}{t}R^h_i - 2\frac{\beta^2}{k}R^h_i(T^h_i - T_{\mathrm{ref}}) + R^h_i\,Q\right)(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\\
&+\int_{\Gamma_{\mathrm{N}}\cup\Gamma_0}\beta R^h_i\left(\mathrm{div}_{\partial\Omega}(\widehat{\boldsymbol{\theta}}\,T_{\mathrm{ext}}) + \frac{\partial T_{\mathrm{ext}}}{\partial \mathbf{n}}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n}) - T^h_i\left(\mathrm{div}_{\partial\Omega}\,\widehat{\boldsymbol{\theta}}\right)\right)\mathrm{d}\mathbf{s}\Bigg).
\end{aligned}
\tag{2.55}
$$

2.3 Numerical simulations

2.3.1 Optimization of elastic structures

As a first example, we consider the optimization of the 3D cantilever structure shown in fig. 2.2. The cantilever is clamped on the four portions Γ_{D} of its boundary, and a traction load $\mathbf{g}$ oriented as $-\mathbf{e}_z$ is applied on Γ_{N}. We suppose that the structure is composed by a linear elastic, homogeneous and isotropic material, characterized by the Young modulus E and the Poisson ratio ν.

We considered three different optimization problems. For all three, the objective is to minimize the volume of the cantilever. However, different constraints are imposed to each case. In the first one we impose that the mechanical compliance should not exceed a threshold τ_C. For the two remaining problems, the constraint is replaced by imposing an upper bound on the L^2-norm and the L^6-norm of the von Mises stress respectively.

The optimization problems can thus be formulated as follows

$$\left|\begin{array}{l}\text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the volume } \text{Vol}(\Omega) \\ \text{under the constraint } H(\mathbf{u}_\Omega, \Omega) \leq \tau, \\ \text{where the displacement } \mathbf{u}_\Omega \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \\ \text{solves the elasticity equation} \\ \quad \left\{\begin{array}{rcll} -\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_\mathrm{N}, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega & = & \mathbf{0} & \text{on } \Gamma_\mathrm{D}. \end{array}\right. \end{array}\right. \tag{2.56}$$

The three different constraint functionals taken in consideration here are:

- $H(\mathbf{u}_\Omega, \Omega) = \mathcal{C}\left(\mathbf{u}_\Omega, \Omega\right) = \displaystyle\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega\right) \mathrm{d}\mathbf{x}$ for the mechanical compliance;
- $H(\mathbf{u}_\Omega, \Omega) = \left\| s_\mathrm{D}\left(\mathbf{u}_\Omega\right)\right\|_{\mathrm{L}^2(\Omega)} = \sqrt{\dfrac{3}{2}}\left(\displaystyle\int_\Omega \boldsymbol{\sigma}_\mathrm{D}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\sigma}_\mathrm{D}\left(\mathbf{u}_\Omega\right) \mathrm{d}\mathbf{x}\right)^{1/2}$ for the L^2-norm of the von Mises stress;
- $H(\mathbf{u}_\Omega, \Omega) = \left\| s_\mathrm{D}\left(\mathbf{u}_\Omega\right)\right\|_{\mathrm{L}^6(\Omega)} = \sqrt{\dfrac{3}{2}}\left(\displaystyle\int_\Omega \left(\boldsymbol{\sigma}_\mathrm{D}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\sigma}_\mathrm{D}\left(\mathbf{u}_\Omega\right)\right)^3 \mathrm{d}\mathbf{x}\right)^{1/6}$ for the L^6-norm of the von Mises stress.

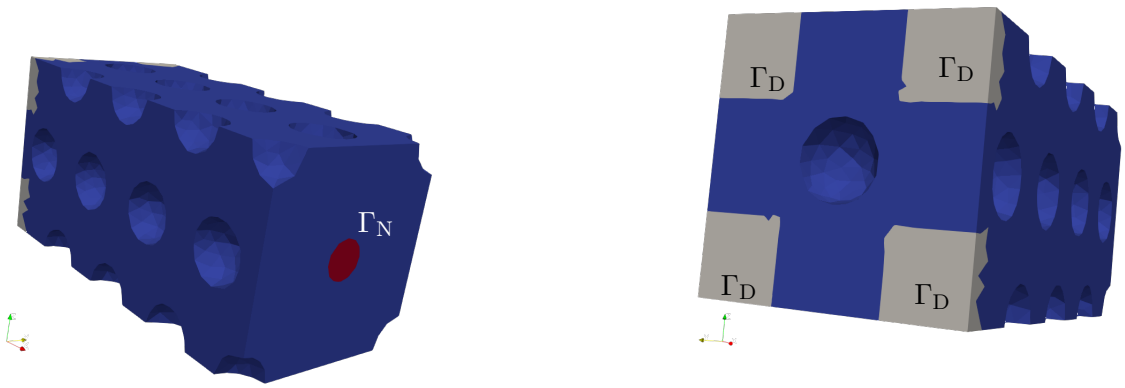


Figure 2.2: Initial and final condition for the optimization problems (2.56) of the 3D cantilever.

All numerical results of this section and the next rely on the *sotuto* platform developed by Dapogny and Feppon in [91] and summarized in section 1.3.5 The simulations have been performed on a Virtualbox virtual machine Linux with 1GB of dedicated memory, on a Dell PC equipped with a 2.80 GHz Intel i7 processor. The numerical values for the geometry, the material properties, and the discretization parameters are presented in table 2.1. The numerical results are compiled in table 2.2.

In fig. 2.3 we show the optimal shape of the cantilever under a constraint on the mechanical compliance, and in fig. 2.4 we report the trends of the objective and the constraint functions

Geometry of the structure		
cross section length	ℓ_s	1.0 cm
longitudinal length	ℓ_x	2.0 cm
sidelength of Γ_{D}		0.3 cm
radius of Γ_{N}		0.1 cm
Elastic coefficients		
young's modulus	E	200 MPa
poisson's ratio	ν	0.3
Mechanical loads		
vertical load	$\lvert\mathbf{g}\rvert$	10 kPa
Mesh size parameters		
minimal mesh size	`hmin`	0.025 cm
maximal mesh size	`hmin`	0.10 cm
Thresholds for the inequality constraints		
threshold on the compliance	τ_C	5×10^{-2} kPa cm^3
threshold on the norm of the von Mises stress	τ_{D}	3 MPa

Table 2.1: Numerical data about the geometry and the mechanics of the cantilever of fig. 2.2.

throughout the optimization process. The optimized cantilever has a hollow structure that is reinforced in the vertical direction, in order to resist the vertical traction. It is worth remarking that the upper and lower halves of the structure are identical. This symmetry is due to the invariance of the compliance with respect to a change of sign of the displacement.

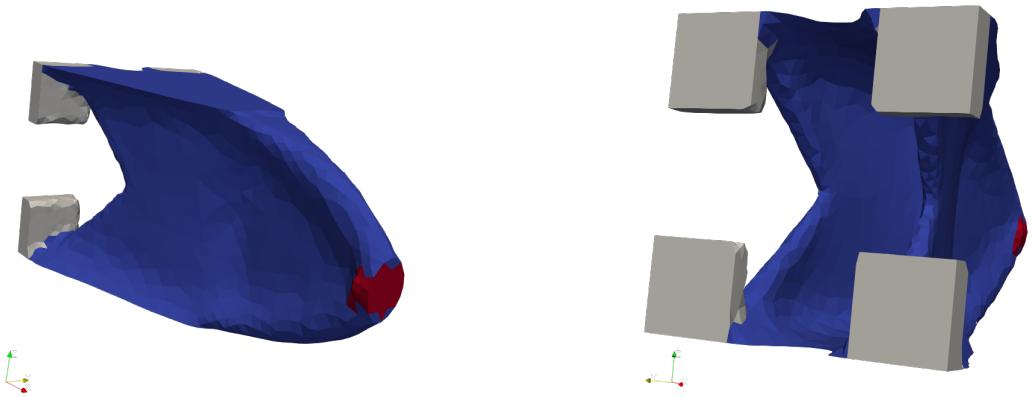
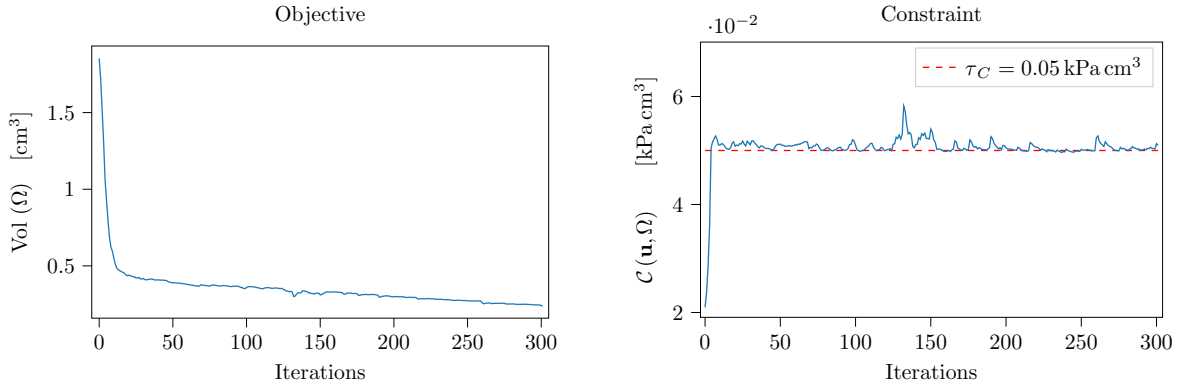


Figure 2.3: Optimal shape of the 3D cantilever under a constraint on the mechanical compliance.

The results of the optimization under constraint on the L^p-norm of the von Mises stress are reported in fig. 2.5 (for $p = 2$) and fig. 2.6 (for $p = 6$). The evolution of the objective and the constraint is shown in fig. 2.7 for both cases.

As for the case of the constraint on the compliance, the optimized shapes are reinforced on the vertical direction and present a symmetry between the upper and lower halves. However, the structures are not hollow and show a pronounced branched structure. In graph of fig. 2.7b



(a) Evolution of the objective function.

(b) Evolution of the constraint.

Figure 2.4: Convergence of the objective and the constraints for the 3D cantilever under constraints on the mechanical compliance.

we can observe a few spikes in the trend of the constraint on the L^6-norm of the von Mises stress. This phenomenon shows that the enforcement of the constraint on the L^6-norm of the von Mises stress is more difficult than for the constraint on the L^2-norm. Indeed, the L^6-norm is more sensible to the concentration of the von Mises stress than the L^2-norm. Thus, numerical instabilities can arise due to an inadequate mesh refinement in the critical regions for some steps of the optimization.

The numerical results of table 2.2 compare the value of the compliance and the L^2 and L^6 - norms of the von Mises stress for the three optimal structures. At first, we can remark that the constraints enforced in each problem are saturated and satisfied. By comparing the norms of the von Mises stress for the three cases it appears that the L^6-norm is always larger than the L^2-norm. Despite their similar shape, the structure optimized for the L^2-norm does not satisfy the constraint on the L^6-norm of the von Mises stress, while the optimal structure for the L^6-norm satisfies both. This difference can be explained by a stress concentration in the first structure that has been avoided in the second one. As discussed in section 2.1.3, the avoidance of high concentrations of the von Mises stress is a problem of primary importance in structural mechanics. The different results yielded by the cantilevers optimized for the L^2 and the L^6-norm highlight the need to consider norms of higher degree in order to obtain robust structures with respect to the von Mises criterion.

2.3.2 A 2D thermo-elastic problem

In this section we present the results of the numerical solution of a 2D thermo-elastic problem.

Let us consider a square computational domain D and an elastic structure $\Omega \subset D$. We divide the boundary of Ω in three disjoint regions: Γ_D consists in two support region on the lower face of D, Γ_N is the upper surface, and Γ_0 the remaining portion of $\partial\Omega$. A uniform mechanical load is applied on the upper surface of the structure (see section 2.3.2), while a time-dependent thermal field T_{ext} swipes the space $D \setminus \Omega$. We suppose that T_{ext} models an ascending heatwave with the following expression

$$T_{\text{ext}}(t, \mathbf{x}) = T_0 + \max\left(0.0, -T_{\max} \sin\left(\pi\left(y - 2\frac{t}{t_f}\right)\right)\right) \qquad \begin{array}{rcl} t & \in & (0, t_f], \\ \mathbf{x} = (x, y) & \in & D \setminus \Omega. \end{array} \tag{2.57}$$

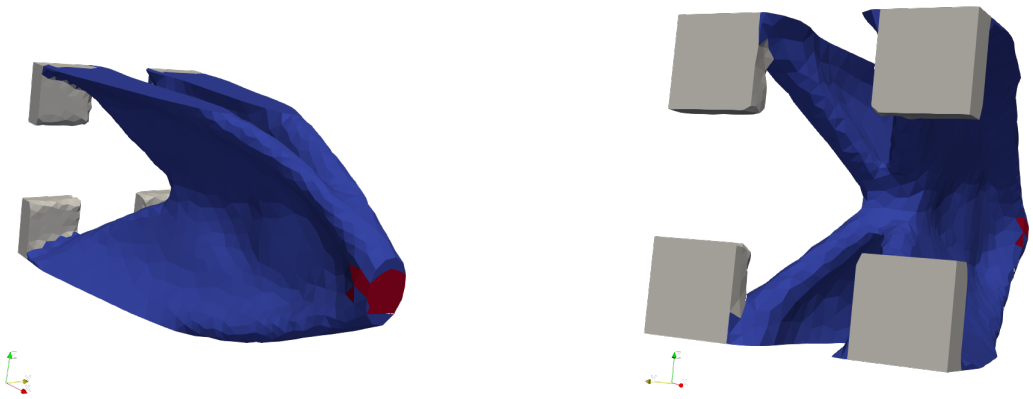


Figure 2.5: Optimal shape of the 3D cantilever under a constraint on the L^2-norm of the von Mises stress.

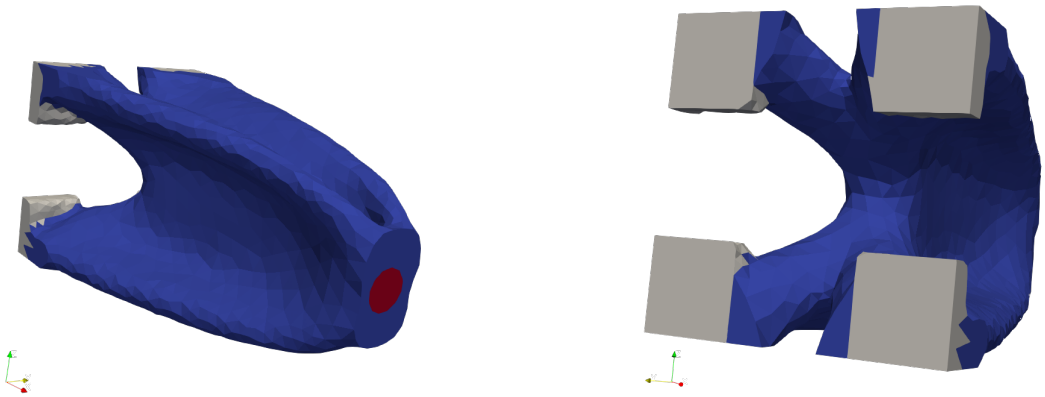
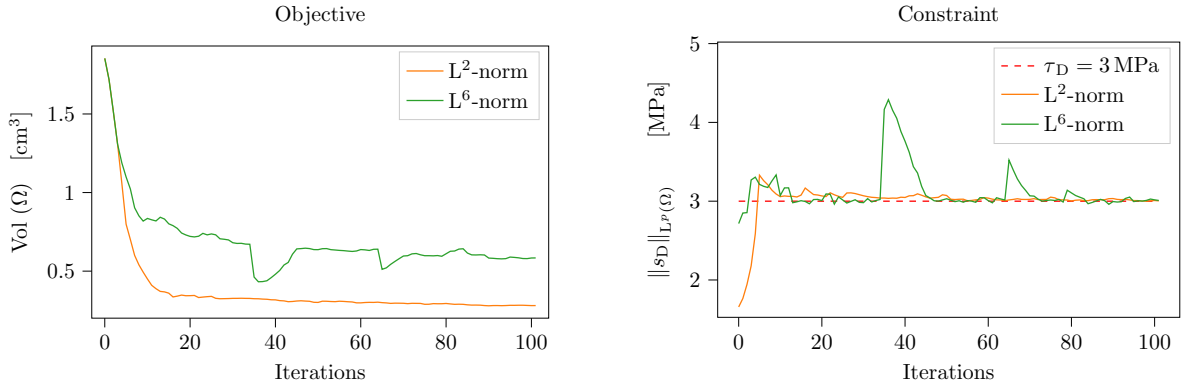


Figure 2.6: Optimal shape of the 3D cantilever under a constraint on the L^6-norm of the von Mises stress.



(a) Evolution of the objective function.

(b) Evolution of the constraint.

Figure 2.7: Convergence of the objective and the constraints for the 3D cantilever under constraints on the L^2-norm and the L^6-norm of the von Mises stress.

Elastic problem		**Constraint functional**		
		Compliance	L^2-norm of s_D	L^6-norm of s_D
Execution				
Number of iterations		300	100	100
Execution time	[min]	143.7	47.6	40.9
Objective				
Final volume	[cm^3]	0.2402	0.2493	0.5840
Constraints				
$\mathcal{C}(\mathbf{u}, \Omega)$	[kPa cm^3]	**5.0936e − 2**	6.6118×10^{-2}	3.5763×10^{-2}
$\|s_D(\mathbf{u}_\Omega)\|_{L^6(\Omega)}$	[MPa]	2.645	**3.003**	2.203
$\|s_D(\mathbf{u}_\Omega)\|_{L^2(\Omega)}$	[MPa]	4.906	5.571	**2.989**

Table 2.2: Numerical results for the cantilever problem (2.56)

The heat is transmitted from the exterior of Ω to the structure through the free surface Γ_0, and than it diffuses in the interior of Ω.

The elastic properties of the material are described by the Lamé parameters λ and μ, the density ρ_M, the heat capacity C, the thermal expansion coefficient α, the thermal conductivity k, and the heat transfer coefficient β between the structure Ω and the exterior. The term ρ appearing in the heat equation (2.26) is the product between the mass density ρ_M and the heat capacity of the material C. The equations describing the heat diffusion in Ω and the elastic displacement are presented in section 2.2 as equation (2.26) and equation (2.26). For the sake of simplicity, we neglect the volume forces term ($\mathbf{f} = \mathbf{0}$) and the thermal generation inside Ω ($Q = 0$). A vertical pressure of module $|\mathbf{g}|$ is applied on the upper surface of the structure.

We aim to find the shape Ω of minimal volume for which the functional $\mathcal{C}_T(\mathbf{u}, T; \Omega)$ introduced in equation (2.29) is kept below a threshold τ. The shape optimization problem can thus be summarized as in problem (2.30). We suppose that small regions around Γ_N and Γ_D are non-optimizable. The numerical values of the parameters cited above are presented in table 2.3, while the procedure for the numerical optimization is detailed in section 2.2.4.

Geometry of the structure		
Length	ℓ_x	1.0 cm
Heigth	ℓ_y	1.0 cm
Sidelength of each support Γ_{D}		0.1 cm
Thickness of the non-optimizable regions		0.05 m
Duration of the simulation	t_f	5.0 s
Discretization parameters		
Minimal mesh size	`hmin`	0.01
Maximal mesh size	`hmax`	0.03
Time step length	Δt	0.1 s
Number of time steps	N_t	50
Elastic coefficients		
Young modulus	E	200 GPa cm
Poisson ratio	ν	0.3
Thermo-mechanical properties		
Mass density	ρ	8×10^{-3} kg cm^{-2}
Heat capacity	C	450 J kg^{-1} K^{-1}
Thermal conductivity	k	15×10^{-2} W m^{-1} K^{-1}
Heat transfer coefficient	β	10 W m^{-1} K^{-1}
Thermal expansion coefficient	α	0.05 GPa K^{-1}
Mechanical and thermal loads		
Vertical mechanical load	$\lvert\mathbf{g}\rvert$	10 GPa cm
Maximal external temperature	$T_{\max}$	100 °C
Parameters on the constraint functional		
Weight parameter	γ	0.5
Threshold on the compliance	τ	3.0×10^{-3} MPa cm^2

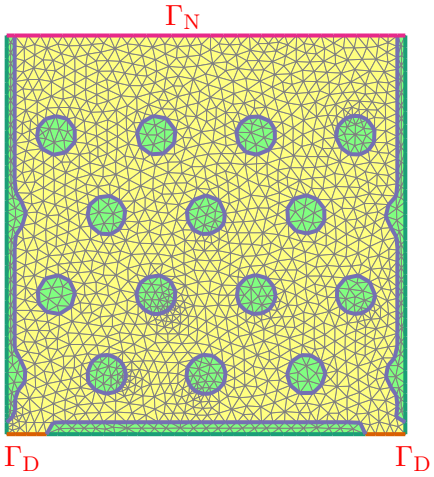
Table 2.3: Numerical data concerning the geometry, the mechanics, and the material properties of the thermo-elastic problem (2.30).

We compare two cases. In the first one we consider the thermal coupling according to the equations of section 2.2. In the first one, all dependence on time and temperature is disregarded. For this case, the functional $\mathcal{C}_{\mathrm{T}}(\mathbf{u}, T; \Omega)$ is equivalent to the simple mechanical compliance. The results of the two simulations are presented in fig. 2.8a and fig. 2.8b respectively. Figure 2.10 compare the evolution of the compliance through the optimization for both cases, and compare it with the imposed threshold. In fig. 2.10a the compliance $\mathcal{C}_{\mathrm{T}}(\mathbf{u}, T; \Omega)$ is decomposed in two parts $\mathcal{C}_{\mathrm{T}}^{\mathrm{Mech}}$ and $\mathcal{C}_{\mathrm{T}}^{\mathrm{Th}}$, in order to highlight the contributions of the two parts of the thermo-elastic stress tensor $\boldsymbol{\sigma}(\mathbf{u}_\Omega, T_\Omega)$

$$
\begin{aligned}
\mathcal{C}_{\mathrm{T}}^{\mathrm{Mech}} &= \tfrac{\gamma}{t_f}\int_0^{t_f}\int_\Omega \mathbb{C}\nabla\mathbf{u}_\Omega(t,\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega(t,\mathbf{x}))\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&\quad +(1-\gamma)\int_\Omega \mathbb{C}\nabla\mathbf{u}_\Omega(t_f,\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega(t_f,\mathbf{x}))\,\mathrm{d}\mathbf{x}, \\
\mathcal{C}_{\mathrm{T}}^{\mathrm{Th}} &= -\alpha\tfrac{\gamma}{t_f}\int_0^{t_f}\int_\Omega (T_\Omega(t,\mathbf{x}) - T_{\mathrm{ref}})\,\mathrm{div}\,\mathbf{u}_\Omega(t,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \\
&\quad -\alpha(1-\gamma)\int_\Omega (T_\Omega(t_f,\mathbf{x}) - T_{\mathrm{ref}})\,\mathrm{div}\,\mathbf{u}_\Omega(t_f,\mathbf{x})\,\mathrm{d}\mathbf{x},
\end{aligned}
$$

The numerical results of the two optimization problems are compared in table 2.4. In fig. 2.9

we show the decrease of the volume (the objective function) through the optimization process. Figure 2.10 illustrates the conformity to the constraint for all three examples. In fig. 2.11 is shown the evolution of the temperature inside the structure optimized for the thermo-elastic problem.



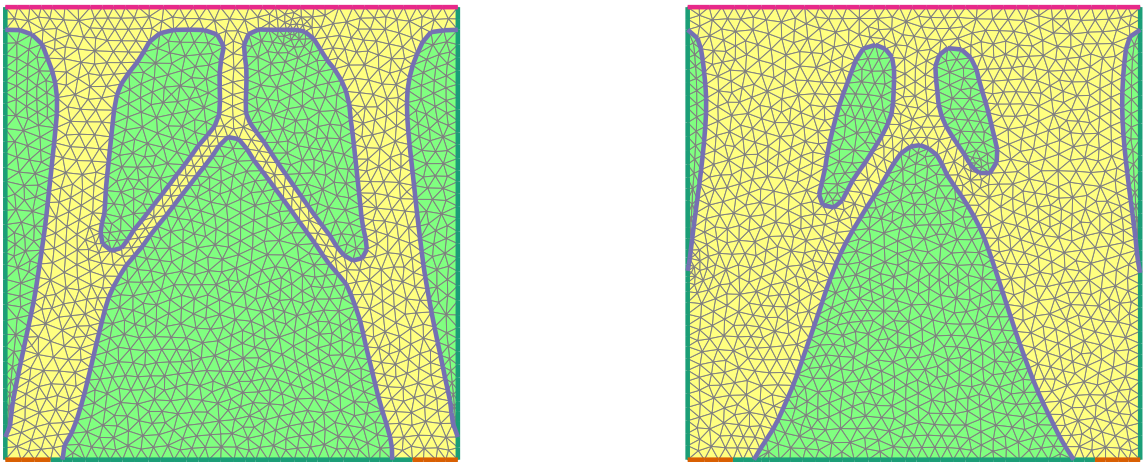
		Thermo-elasticity	**Elasticity**
Execution			
Number of iterations		100	100
Execution time	[min]	18.07	1.30
Results			
Final volume	$\mathrm{Vol}\,(\Omega_{\mathrm{opt}})$ [cm^2]	0.6423	0.3825
Compliance	$\mathcal{C}_{\mathrm{T}}\,(\mathbf{u}, T; \Omega)$ [MPa cm^2]	$3.004\,51 \times 10^{-3}$	$3.010\,18 \times 10^{-3}$

Table 2.4: Numerical results for the thermo-elastic problem (2.30)

A notable difference between the mechanical compliance for purely elastic problems and problems where the thermal effects are taken into account is the fact that the second one can be negative. If the thermal expansion is predominant over the deformation due to mechanical solicitation, it is possible that the pressure applied to the structure exerts a negative work. Such situation can be observed for the initial condition of the thermo-elastic problem in fig. 2.10a.

Figure 2.10a shows also the antagonistic role played by the components $\mathcal{C}_{\mathrm{T}}^{\mathrm{Mech}}$ and $\mathcal{C}_{\mathrm{T}}^{\mathrm{Th}}$ of the constraint functional. Being constantly negative, the thermal component balances the mechanical component, which is always positive, lowering the value of the constraint.The effects of this feature can be observed also by comparing the optimal structure sensible to the thermal variations (fig. 2.8a) with the optimal shape for simple linear elasticity (fig. 2.8b). Indeed, the lowering of the compliance by the thermal expansion effect allows, in this case, for a thinner and lighter structure.

The extremely short duration of the optimization in the purely elastic case is expected. The mechanical response of the structure is supposed to be instantaneous, and the displacement is computed by the quasi-static elasticity equations. Since the dependence from the temperature is neglected for that example, and since the mechanical load is supposed as constant, there is no need to compute the evolution of the displacement field in time.



(a) Thermo-elastic problem.

(b) Purely elastic problem.

Figure 2.8: Optimal shapes for the optimization problem (2.30).

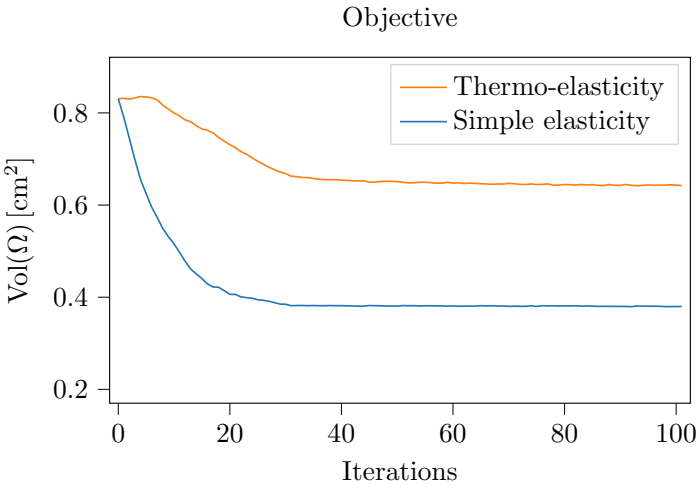
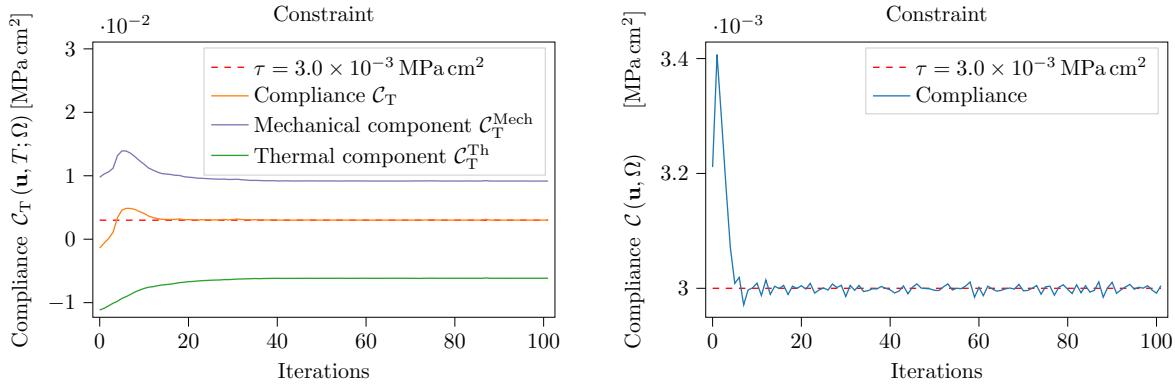


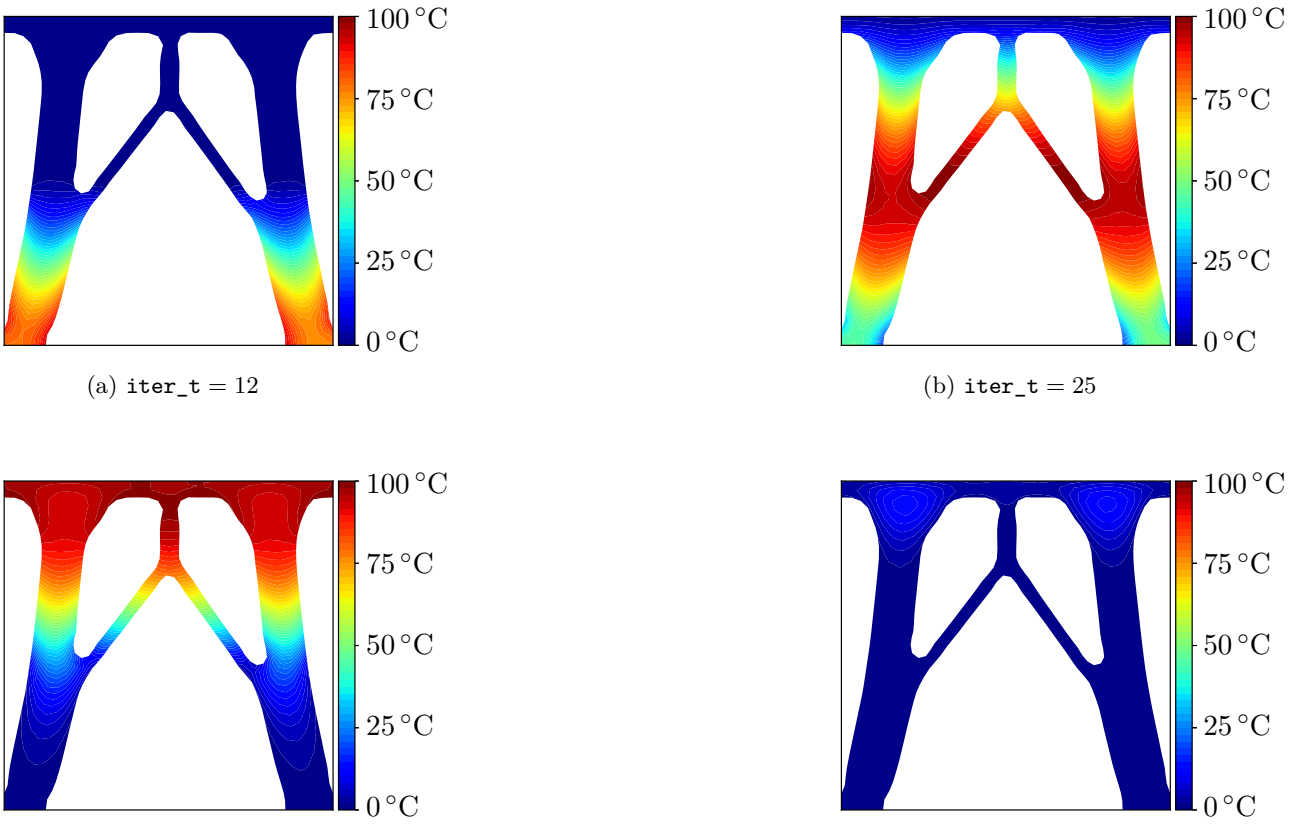
Figure 2.9: Evolution of the volume of the structures for the two cases of the thermo-elastic optimization problem.



(a) Compliance for the thermo-elastic problem, with the mechanical and thermal components.

(b) Compliance for the elastic problem.

Figure 2.10: Convergence of the constraints for the 2D thermo-elastic optimization problem (2.30).



(a) `iter_t` = 12

(b) `iter_t` = 25

(c) `iter_t` = 36

(d) `iter_t` = 50

Figure 2.11: Temperature of the optimal structure for the thermo-elastic problem subject to vertical compression at the instants `iter_t` = 12, 25, 36 and 50.

Part II

Shape optimization under uncertainties

Chapter 3

Optimization of the worst-case scenario

Contents

3.1 Introduction

3.1.1 State of the art and contents of the chapter

An important class of optimization problem under uncertainties focuses on the evaluation of the worst-case scenario of a given functional, which can either appear as the objective function or as the constraint of the optimization problem. The two situations require different approaches. As reported in [43], the optimization of the worst-case scenario is preferable to robust or reliability-based optimization problems when the data are imprecise or if they are uncertain with an unknown probability distribution, or if the strict respect of the constraint in all circumstances is of primary importance. Different techniques to solve shape optimization problems with worst-case functionals have been proposed. In [68, 99, 161, 22] the objective of the optimization problem consists in minimizing the maximal possible value of a given functional.

In [128, 10, 22] are considered also problems where the uncertain functional acts as a constraint, and numerical examples are provided for density and level-set methods. The authors of [10] consider smooth functionals subject to small perturbations, and propose a method to compute the shape derivative of their supremum using linearization techniques. For further information about worst-case problems, outside the domain of structure optimization, we refer to [43, 42].

In this chapter we focus on the optimization of linear elastic structures subject to uncertain mechanical loads. In section 3.1.2 we present an optimization problem under constraints on the worst possible value of a continuous functional of the displacement. In the next two sections we present two different methods to solve shape optimization problems under constraints on the worst-case scenario for a given functional.

The first, discussed in 3.2, requires the convexity of the set of admissible loads and of the function mapping the uncertain parameters with the corresponding value of the constraint function. This first method consists in the replacement of the constraint on the worst-case by a finite number of deterministic constraints, obtaining an optimization problem that approximates the initial one. Theorem 3.12 provides a result on the convergence of the solution of the approximated problem towards the exact solution when the number of constraints increases.

In 3.3 a second method is discussed. After recalling the notion of generalized directional derivative and of subdifferential in the sense of Clarke, we extend them to the context of shape functionals. The optimization method, that is discussed in 3.3.3, consists in the identification of one element of the subdifferential at each step of the shape optimization algorithm, and in the definition of a direction of descent for the constraint function with respect to it.

Finally, in section 3.4 we present some numerical results. We compare the two methods in two different situations. In fig. 3.9b we optimize the shape of a 3D cantilever subject to a traction and compression load on its extremity. In section 3.4.2 we aim to optimize a disc-like structure showing a rotational symmetry. The two methods are compared for both examples, highlighting the advantages of the technique based on the subdifferential in terms of the duration of the computations, as well as the challenges provided by a problem with a symmetric geometry.

3.1.2 Problem framework

Let us consider the following shape optimization problem, where the objective consists in the minimization of the volume of the structure, the state solves the linear elasticity equation, and the mechanical loads applied are uncertain. Let us suppose that $\mathbf{g}$ is a random variable that belongs almost surely to a bounded set $\mathcal{G} \subset \mathrm{L}^2(\Gamma_\mathrm{N})^d$. Therefore, the optimization problem can be written as follows, hiding the dependence from the event space.

$$
\left|\begin{array}{l}
\text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\
\text{minimizing the volume } \mathrm{Vol}(\Omega) \\
\text{under the constraint } \sup_{\mathbf{g}\in\mathcal{G}} H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega) \leq \tau, \\
\text{where the displacement } \mathbf{u}_{\Omega,\mathbf{g}} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \\
\text{solves the elasticity equation} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_\mathrm{N}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\
\mathbf{u}_{\Omega,\mathbf{g}} & = & \mathbf{0} & \text{on } \Gamma_\mathrm{D}.
\end{array}\right.
\end{array}\right. \tag{3.1}
$$

We suppose that the constraint functional $H(\cdot,\cdot)$ can be written in integral form as

$$H(\mathbf{u},\Omega)=\int_{\Omega}\left(j_0(\mathbf{u}(\mathbf{x}))+j_1(\nabla\mathbf{u}(\mathbf{x}))\right)\,\mathrm{d}\mathbf{x}\qquad\text{for }\mathbf{u}\in\mathrm{H}^1\left(\Omega\right)^d,$$

with j_0 and j_1 continuous functions. If the mapping $\mathbf{g}\mapsto H(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)$ is continuous and the set $\mathcal{G}$ is compact, $H(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)$ reaches its supremum for some $\mathbf{g}\in\mathcal{G}$, and the constraint can be replaced by $\max_{\mathbf{g}\in\mathcal{G}}H(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)\leq\tau_{\mathcal{G}}$.

As remarked in [10], two different points of view can be adopted, according to whether the functional affected by the perturbation acts as the objective of an optimization problem or as a constraint. If the objective of the optimization is the minimization of the maximal possible level of a functional, the problem can be formulated as a *min-max* problem. This interpretation can also be applied to the case of constrained optimization problems where an upper bound on the maximum of a functional is imposed. However, this constraint can also be interpreted as imposing an upper bound on the constraining functional for all possible configuration of the uncertain parameters.

3.2 A convexity-based approach for the worst-case

3.2.1 Definitions of convex sets and functions

Before describing the first approach, we recall a few definitions about convex functions and sets.

Definition 3.1 (Convex function)**.** *Let $\mathcal{X}$ be a vector space. A real valued function $f:\mathcal{X}\to\mathbb{R}$ is convex if, for all $\mathbf{x}_1,\mathbf{x}_2\in\mathcal{X}$ such that $\mathbf{x}_1\neq\mathbf{x}_2$, the following inequality holds*

$$f(t\mathbf{x}_1+(1-t)\,\mathbf{x}_2)\leq tf(\mathbf{x}_1)+(1-t)f(\mathbf{x}_2),\qquad\text{for all } t\in(0,1).\tag{3.2}$$

If the inequality (3.2) *is strictly satisfied, the function f is said to be strictly convex.*

In other terms, a function f is (strictly) convex if, for any pair of points $(\mathbf{x}_1,\mathbf{x}_2)$, the graph of f is (strictly) below the segment connecting $\mathbf{x}_1$ to $\mathbf{x}_2$.

Definition 3.2 (Convex set)**.** *A subset $\mathcal{S}$ of the vector space $\mathcal{X}$ is a convex set if, for all $\mathbf{x}_1,\mathbf{x}_2\in\mathcal{S}$ and for all $0<t<1$, the point $y_{\mathbf{x}_1,\mathbf{x}_2}(t)=t\mathbf{x}_1+(1-t)\mathbf{x}_2$ belongs to $\mathcal{S}$. The set $\mathcal{S}$ is said to be strictly convex if $y_{\mathbf{x}_1,\mathbf{x}_2}(t)$ belongs to the interior of $\mathcal{S}$. The minimal convex set containing a set $\mathcal{T}\subset\mathcal{X}$ is the convex hull of $\mathcal{T}$, denoted* $\mathrm{hull}(\mathcal{T})$.

Definition 3.3 (Polyhedral set)**.** *A subset $\mathcal{P}$ of the vector space $\mathcal{X}$ is a polyhedral convex set if it can be expressed as the convex hull of a set containing a finite number of points in $\mathcal{X}$.*

Having stated the definitions of convex functions and sets, we can introduced a classical result concerning the maximization of convex functionals on convex sets. Such result descends directly from [209, Theorem 32.3] and its corollaries.

Proposition 3.4. *Let $f:\mathcal{X}\to\mathbb{R}$ be a convex and bounded function defined on the vector space $\mathcal{X}$, and let $\mathcal{S}\subset\mathcal{X}$ be a compact convex set. Then, f attains* $\sup_{\mathbf{x}\in\mathcal{S}}f(\mathbf{x})$ *in at least a point $\bar{\mathbf{x}}\in\mathcal{S}$, and $\bar{\mathbf{x}}$ belongs to the border $\partial\mathcal{S}$ of the convex set $\mathcal{S}$. Moreover, if $\mathcal{S}$ is a convex, closed, and bounded polyhedral set, $\bar{\mathbf{x}}$ can be found among the vertices of $\mathcal{S}$.*

(a) Convex hull of a set $\mathcal{T}$.

(b) Convex polyhedral set.

Figure 3.1: Examples of the convex hull of a set and a convex polyhedral set in $\mathbb{R}^2$.

3.2.2 The Hausdorff distance and the continuity of the constraint

In order to define the concept of convergence for the domains, it is necessary to introduce a topology on the set of admissible shapes $\mathcal{S}_{\text{adm}}$ and among the sets of possible mechanical loads. At first, we introduce the notion of Hausdorff distance between subsets of metric spaces as in [138, Definition 2.2.7].

Definition 3.5 (Hausdorff distance)**.** *Let $\mathcal{M}$ be a metric space provided with the distance $d_{\mathcal{M}}$, and let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two subsets of $\mathcal{M}$. The distance of a point $\mathbf{x} \in \mathcal{M}$ to a set $\mathcal{A}_1 \subset \mathcal{M}$ is defined as*

$$d_{\mathcal{M}}(\mathbf{x}, \mathcal{A}_1) = \inf_{\mathbf{y} \in \mathcal{A}_1} d_{\mathcal{M}}(\mathbf{x}, \mathbf{y}).$$

The Hausdorff distance between the sets $\mathcal{A}_1$ and $\mathcal{A}_2$ is

$$d_{\text{H}}(\mathcal{A}_1, \mathcal{A}_2) = \max \left\{ \sup_{\mathbf{x} \in \mathcal{A}_1} d_{\mathcal{M}}(\mathbf{x}, \mathcal{A}_2), \sup_{\mathbf{y} \in \mathcal{A}_2} d_{\mathcal{M}}(\mathbf{y}, \mathcal{A}_1) \right\}.$$

The fact that d_{H} is a distance between subsets of $\mathcal{M}$ is proven in [29, Proposition 1.2].

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two subsets of the metric space $(\mathcal{M}, \|\cdot\|)$ such that $d_{\text{H}}(\mathcal{A}_1, \mathcal{A}_2) = 0$. As remarked in [28], we can deduce that $\mathcal{A}_1 = \mathcal{A}_2$ only assuming that both $\mathcal{A}_1$ and $\mathcal{A}_2$ are closed, meaning that d_{H} is a metric for the class of compact subsets of $\mathcal{M}$. However said implication does not hold true for open sets.

In order to define a metric and the notion of convergence for open subsets of $\mathcal{M}$, we limit our study to uniformly bounded open sets. Let $\mathcal{B} \subset \mathcal{M}$ be a fixed compact subset of $\mathcal{M}$. We can introduce a metric on the class of the open subsets of $\mathcal{B}$ as proposed in [138, Definition 2.2.8 and Remark 2.2.10].

Definition 3.6 (Metric among open spaces and Hausdorff convergence)**.** *We consider the following function m_{H} defined on the class of subsets of $\mathcal{B}$*

$$m_{\text{H}}{}^{\mathcal{B}}(\mathcal{A}_1, \mathcal{A}_2) = d_{\text{H}}(\mathcal{B} \setminus \mathcal{A}_1, \mathcal{B} \setminus \mathcal{A}_2).$$

The function m_{H} defines a metric structure on the class of open subsets of $\mathcal{B}$. Moreover, if $\widehat{\mathcal{B}}$ is another compact subset of $\mathcal{M}$, and $\mathcal{A}_1$ and $\mathcal{A}_2$ are open sets contained in both $\mathcal{B}$ and $\widehat{\mathcal{B}}$, we have the identity

$$m_{\text{H}}{}^{\mathcal{B}}(\mathcal{A}_1, \mathcal{A}_2) = m_{\text{H}}{}^{\widehat{\mathcal{B}}}(\mathcal{A}_1, \mathcal{A}_2).$$

For the sake of simplicity, if the compact subset of the metric space including all subspaces of interest is fixed, we denote the Hausdorff metric introduced in (3.6) *as* $m_{\mathrm{H}}(\cdot,\cdot)$. *Finally, if* $\{\mathcal{A}_n\}_{n=1}^{\infty}$ *and* $\mathcal{A}$ *are open subsets of* $\mathcal{B}$, *we say that the sequence* $\{\mathcal{A}_n\}_{n=1}^{\infty}$ *converges in the sense of Hausdorff towards* $\mathcal{A}$ *if*

$$\lim_{n\to\infty} m_{\mathrm{H}}(\mathcal{A}_n,\mathcal{A}) = 0$$

and we denote such convergence as $\mathcal{A}_n \xrightarrow{\mathrm{H}} \mathcal{A}$.

We recall that the Hausdorff metric m_{H} does not take into account the topology of the open sets, and it is possible to design two open sets arbitrarily close to each other that are different from a topological point of view.

The shape functional $\Omega \mapsto \sup_{\mathbf{g}\in\mathcal{G}} H(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)$ is not continuous if $\mathcal{S}_{\mathrm{adm}}$ is a generic class of open domains in $\mathbb{R}^d$ since it relies on the computation of $\mathbf{u}_\Omega$. Therefore, it is necessary to restrict the class of admissible domains in order to ensure the continuity of the constraint functional.

Definition 3.7 (ε-cone condition)**.** *Let* ε *be a positive parameter. For any* $\mathbf{x}\in\mathbb{R}^d$ *and any unit vector* $\boldsymbol{\xi}\in\mathbb{R}^d$, *we denote* $B_\varepsilon(\mathbf{x})$ *the ball of radius* ε *centered in* $\mathbf{x}$, *and* $C_\varepsilon(\mathbf{x},\boldsymbol{\xi})$ *the open cone of vertex* $\mathbf{x}$ *(without its vertex), of direction* $\boldsymbol{\xi}$ *defined as*

$$C_\varepsilon(\mathbf{x},\boldsymbol{\xi}) = \left\{\mathbf{y}\in\mathbb{R}^d \ : \ \langle \mathbf{y}-\mathbf{x},\boldsymbol{\xi}\rangle_{\mathbb{R}^d} \geq \cos(\varepsilon)\,|\mathbf{y}-\mathbf{x}| \quad \textit{and} \quad 0<|\mathbf{y}-\mathbf{x}|<\varepsilon\right\}.$$

An open set $\mathcal{A}$ *is said to satisfy the* ε*-cone condition if, for all* $\mathbf{x}$ *on the boundary of* $\mathcal{A}$, *there exists a unit vector* $\boldsymbol{\xi}$ *such that for any* $\mathbf{y}\in B_\varepsilon(\mathbf{x})$, *the cone* $C_\varepsilon(\mathbf{y},\boldsymbol{\xi})$ *is contained in* $\mathcal{A}$. *We denote* $\mathcal{D}_\varepsilon(\mathbb{R}^d)$ *the set of all open domains in* $\mathbb{R}^d$ *fulfilling the* ε*-cone condition. For any compact set* $\mathcal{B}\subset\mathbb{R}^d$, *we denote* $\mathcal{D}_\varepsilon(\mathcal{B})$ *the subset of all open domains in* $\mathcal{D}_\varepsilon(\mathbb{R}^d)$ *contained into* $\mathcal{B}$

$$\mathcal{D}_\varepsilon(\mathcal{B}) = \{\mathcal{A}\subset\mathcal{B} \ : \ \mathcal{A} \textit{ open, and satisfying the } \varepsilon\textit{-cone condition}\}. \tag{3.3}$$

This definition can be found in [138, Definition 2.4.1]. Moreover, according to [138, Theorem 2.4.7], a domain Ω satisfies the ε-cone condition for some $\varepsilon>0$ if and only if it has a uniformly Lipschitz continuous boundary.

The continuity of the reduced constraint functional $\widehat{H}$ relies on the following result, proven in [67] and reported in [138, Theorem 3.2.13].

Theorem 3.8. *We consider* $\varepsilon>0$, *and* $\mathcal{B}\subset\mathbb{R}^d$ *to be a compact set. Let* $\{\Omega_n\}_{n=1}^{\infty}$ *and* Ω *be open domains in* $\mathcal{S}_{\mathrm{adm}}\subset\mathcal{D}_\varepsilon(\mathcal{B})$, *and let* $\Omega_n \xrightarrow{\mathrm{H}} \Omega$. *Then, the sequence* $\{\mathbf{u}_{\Omega_n}\}_{n=1}^{\infty}$ *converges towards* $\mathbf{u}_\Omega$, *where* $\mathbf{u}_\Omega$ *is the solution of the elasticity problem on* Ω *and* $\mathbf{u}_{\Omega_n}$ *is the solution on* Ω_n *for all* $n\geq 0$.

The result of theorem 3.8 combined with the structure of the constraint stated in (3.1.2) ensures that the reduced shape functional $\Omega\mapsto\sup_{\mathbf{g}\in\mathcal{G}} H(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)$ is continuous on a class of uniformly bounded open sets satisfying the ε-cone condition for some positive ε.

3.2.3 Theoretical results

In industrial applications, the principal technique to approximate the solution of problem (3.1) consists in the identification of a number N of loading conditions and consider them as separate constraints of the shape optimization problem. From now on, we consider that all admissible domains satisfy the ε-cone condition, that they are uniformly bounded by a compact set $\mathcal{B}$, and that $\mathcal{S}_{\mathrm{adm}}$ is a closed subset of $\mathcal{D}_\varepsilon(\mathcal{B})$. We suppose that the set of admissible loads is bounded,

convex and finite-dimensional, and the mapping $\mathbf{g} \mapsto H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$ is a convex function. The objective of this section is to justify the approach of the engineers and provide some results on the convergence of the solution when the number of loading conditions N increases.

We state all the results of this section for the following generic optimization problem, where the dependence from the state is kept implicit

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the volume } \operatorname{Vol}(\Omega) \\ \text{under the constraint } \sup_{\mathbf{g}\in\mathcal{G}} h\left(\mathbf{g}, \Omega\right) \leq \tau\,. \end{array}\right. \tag{3.4}$$

We suppose that $\mathcal{G}$ is contained in a finite-dimensional normed vector space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, and that $h : \mathcal{G} \times \mathcal{S}_{\text{adm}} \to \mathbb{R}$ is a real-valued function. We assume also that $\mathcal{S}_{\text{adm}}$ consists in a class of open subsets of $\mathbb{R}^d$ which are uniformly bounded and satisfy the ε-cone condition for some $\varepsilon > 0$. The mapping $\Omega \mapsto h\left(\mathbf{g}, \Omega\right)$ can be seen as an instance of a family of shape functions depending on the parameter $\mathcal{G}$. Moreover, we suppose that said functions satisfy the following conditions:

(i) the mapping $\mathbf{g} \mapsto h\left(\mathbf{g}, \Omega\right)$ is convex and bounded for all admissible $\Omega \in \mathcal{S}_{\text{adm}}$;

(ii) for all choice of the parameter $\mathbf{g} \in \mathcal{G}$, the mapping $\Omega \mapsto h\left(\mathbf{g}, \Omega\right)$ is shape differentiable.

A first result on the solution of problem (3.1) is provided by the following proposition, which applies to the case where the set of admissible loads is polyhedral.

Proposition 3.9. *Let $\mathcal{G}$ be a compact convex polyhedral subset of the Banach space $\mathcal{Y}$ with N vertices $\mathbf{g}_1, \ldots, \mathbf{g}_N$, and $h : \mathcal{G} \times \mathcal{S}_{\text{adm}} \to \mathbb{R}$ a real-valued function satisfying the assumptions (i) - (ii) stated above. Then, the solution of problem* (3.4) *is equivalent to the solution of the following constrained optimization problem:*

$$\left|\begin{array}{l} \textit{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \textit{minimizing the volume } \operatorname{Vol}(\Omega) \\ \textit{under the } N \textit{ inequality constraints:} \\ \qquad h\left(\mathbf{g}_1, \Omega\right) \leq \tau, \\ \qquad\qquad \ldots \\ \qquad h\left(\mathbf{g}_N, \Omega\right) \leq \tau\,. \end{array}\right. \tag{3.5}$$

Proposition 3.9 follows directly from the application of proposition 3.4 to the inequality constraint. Moreover, the formulation of proposition 3.4 as an optimization problem with multiple constraints makes it conforming with the *nullspace* optimization algorithm introduced in [114, 112] and presented in section 1.3.4.

Having proven a result on the solution of problem (3.4) for convex polyhedra, we aim to extend it to more general compact convex sets. Let $\mathcal{G}$ be a compact and convex subset of the Banach space $\mathcal{Y}$, and $\{\mathcal{G}_n\}_{n=1}^{\infty}$ a sequence of convex compact polyhedral subsets of $\mathcal{Y}$ converging towards $\mathcal{G}$ with respect to the Hausdorff distance. The next step is the evaluation of the convergence of the minimizers of a sequence of problems in the form (3.4). A first important remark concerns the relation of the admissible sets in two different problems, when the corresponding sets of parameters are nested one into the other.

Lemma 3.10. *Let us consider two subsets $\mathcal{G}_1$, $\mathcal{G}_2$ of the Banach space $\mathcal{Y}$ such that $\mathcal{G}_1 \subset \mathcal{G}_2$. We denote E_1, E_2 the subsets of $\mathcal{S}_{\text{adm}} \subset \mathcal{D}_\varepsilon(\mathcal{B})$ where the inequality constraint of problem (3.4) is satisfied for the sets of parameters $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively:*

$$E_i = \left\{ \Omega \in \mathcal{S}_{\text{adm}} \ : \ \sup_{\mathbf{g} \in \mathcal{G}_i} h\left(\mathbf{g}, \Omega\right) \leq \tau \right\}, \qquad \text{for } i \in \{1, 2\}.$$

Then, $E_2 \subset E_1$.

Proof. Let us consider $\Omega \in E_2$. Since all $\mathbf{g}_1 \in \mathcal{G}_1$ belongs also to $\mathcal{G}_2$, we have that $h\left(\mathbf{g}_1, \Omega\right) \leq \tau$ for all $\mathbf{g}_1 \in \mathcal{G}_1$. Thus $\Omega \in E_1$. □

Thanks to lemma 3.10, we can prove the following result about the convergence of the solutions of a sequence of problems in the form (3.4).

Proposition 3.11. *We consider $\{\mathcal{G}_n\}_{n=1}^{\infty}$ to be an increasing sequence of compact subsets of $\mathcal{Y}$ where $\mathcal{G}_i \subset \mathcal{G}_j$ if $i < j$ and such that $\mathcal{G} = \overline{\bigcup_{i=1}^{\infty} \mathcal{G}_i}$ is compact as well. Let $h : \mathcal{G} \times \mathcal{S}_{\text{adm}} \to \mathbb{R}$ be a function that satisfies the assumptions (i) and (ii), $\mathcal{S}_{\text{adm}} \subset \mathcal{D}_\varepsilon(\mathcal{B})$ closed, and $\tau \in \mathbb{R}$ be a given threshold. As in lemma 3.10, we denote E_i the subset of admissible domains $\mathcal{S}_{\text{adm}}$ such that, if $\Omega \in E_i$, then $h\left(\mathbf{g}, \Omega\right) \leq \tau$ for all $g \in \mathcal{G}_i$. Finally, we denote E the set of admissible domains such defined as*

$$E = \left\{ \Omega \in \mathcal{S}_{\text{adm}} \ : \ \sup_{\mathbf{g} \in \mathcal{G}} h\left(\mathbf{g}, \Omega\right) \leq \tau \right\},$$

and we suppose that neither any set E_i, nor the set E is empty. Then, the sequence $\{E_i\}_{i=1}^{\infty}$ is decreasing, in the sense that $E_i \supseteq E_j$ if $i < j$, $E = \bigcap_{i=1}^{\infty} E_i$, and E as well as all E_i are closed subsets of $\mathcal{S}_{\text{adm}}$ with respect to its metric m_{H}.

Proof. The fact that $\{E_i\}_{i=1}^{\infty}$ is a decreasing sequence follows directly from lemma 3.10. Next, we prove the identity $E = \bigcap_{i=1}^{\infty} E_i$. The inclusion $E \subset \bigcap_{i=1}^{\infty} E_i$ is, again, a direct consequence of lemma 3.10 since, for all $i > 0$, we suppose that $\mathcal{G} \supseteq \mathcal{G}_i$. In order to prove the converse inclusion we suppose that $\Omega \in E_i$ for all $i > 0$. Since we defined the set $\mathcal{G}$ as $\mathcal{G} = \overline{\bigcup_{i=1}^{\infty} \mathcal{G}_i}$, for all $\mathbf{g} \in \mathcal{G}$, there exists a sequence $\{\mathbf{g}_i\}_{i=1}^{\infty}$ such that $\mathbf{g}_i \in \mathcal{G}_i$ for all $i > 0$, and $\mathbf{g}_i \to \mathbf{g}$. By hypothesis, $\mathbf{g} \mapsto h\left(\mathbf{g}, \Omega\right)$ is convex on the finite-dimensional space $\mathcal{Y}$ and it is bounded. Thus, such mapping is also continuous (see [209, Corollary 10.1.1]). By the definition of the sets $\{E_i\}_{i=1}^{\infty}$ and the sequence $\{\mathbf{g}_i\}_{i=1}^{\infty}$, and by the continuity of $\mathbf{g} \mapsto h\left(\mathbf{g}, \Omega\right)$, we deduce that $h\left(\mathbf{g}, \Omega\right) \leq \tau$, and we conclude that $\Omega \in E$.

Let $i \in \mathbb{N}$. In order to prove that E_i is a closed set, we consider the function $\Phi_i : \mathcal{S}_{\text{adm}} \to \mathbb{R}^+$, mapping $\Omega \mapsto \sup_{\mathbf{g} \in \mathcal{G}_i} h\left(\mathbf{g}, \Omega\right)$. Such function is well-defined and continuous on $\mathcal{S}_{\text{adm}}$, since the set of parameters $\mathcal{G}_i$ is compact. Thus, we deduce that $E_i = \Phi_i^{-1}([0, M])$ is a closed subset of $\mathcal{S}_{\text{adm}}$. Since $E = \bigcap_{i=1}^{\infty} E_i$, we conclude that E is closed with respect to the Hausdorff metric in $\mathcal{S}_{\text{adm}}$ as well. □

Now, we can state the main result of this section, which is about the convergence of the solution of a sequence of shape optimization problems in the form (3.9) with an increasingly accurate approximation of the set $\mathcal{G}$.

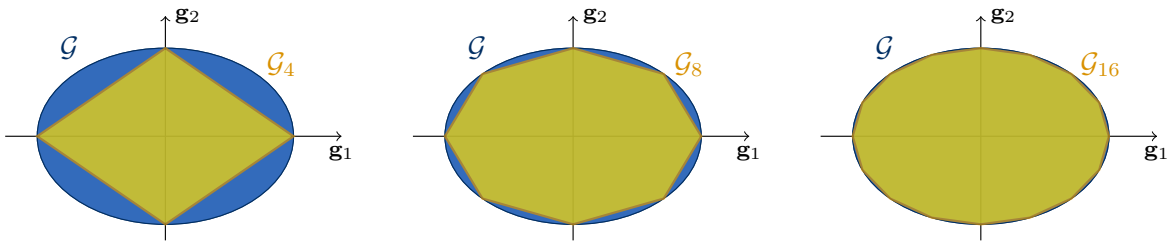
Theorem 3.12. *We consider a compact set $\mathcal{B} \subset \mathbb{R}^d$, and a family of open domains $\mathcal{S}_{\text{adm}}$, uniformly bounded by $\mathcal{B}$ and closed in $\mathcal{D}_\varepsilon(\mathcal{B})$. Let $h : \mathcal{G} \times \mathcal{S}_{\text{adm}} \to \mathbb{R}$ be a function fulfilling assumptions (i) and (ii), $\tau \in \mathbb{R}$ a given threshold, and $\{\mathcal{G}_n\}_{n=1}^{\infty}$ a sequence of compact subsets of $\mathcal{Y}$ satisfying the hypotheses of proposition 3.11. Let $\{\Omega_i\}_{i=1}^{\infty}$ be a sequence of domains such*

that $\Omega_i \in \arg\min_{\Omega \in E_i} \mathrm{Vol}(\Omega)$ *for all* $i \in \mathbb{N}$. *Then,* $\{\Omega_i\}_{i=1}^{\infty}$ *admits a converging subsequence with respect to the Hausdorff metric, and any* Ω_∞ *in the limit class is a solution of problem (3.4).*

Proof. Let us consider a sequence $\{\Omega_i\}_{i=1}^{\infty}$ such that $\Omega_i \in E_i$ for all $n > 0$. All sets E_i are closed subset of $\mathcal{S}_{\mathrm{adm}} \subset \mathcal{D}_\varepsilon(\mathcal{B})$ embedded into one another, and $\mathcal{D}_\varepsilon(\mathcal{B})$ is sequentially compact with respect to the Hausdorff metric [138, Theorem 2.4.10]. Thus, $\{\Omega_i\}_{i=1}^{\infty}$ admits a subsequence converging towards $\Omega_\infty \in \mathcal{D}_\varepsilon(\mathcal{B})$, and $\Omega_\infty \in E_i$ for all $i \in \mathbb{N}$. Thanks to proposition 3.11 we deduce that $\Omega_\infty \in E$.

Finally, in order to prove that $\Omega_\infty \in \arg\min_{\Omega \in E} \mathrm{Vol}(\Omega)$, we reason by contradiction. Let $\varepsilon > 0$ and $\Omega_\varepsilon \in E$ such that $\mathrm{Vol}(\Omega_\infty) = \mathrm{Vol}(\Omega_\varepsilon) + \varepsilon$. Since $\mathrm{Vol}(\cdot)$ is a continuous function with respect to the metric m_{H}, there exists $N_\varepsilon > 0$ such that, for all $n > N_\varepsilon$, $\mathrm{Vol}(\Omega_n) > \mathrm{Vol}(\Omega_\varepsilon) + \varepsilon/2$. This result is in contradiction with the assumption $\Omega_n \in \arg\min_{\Omega \in E_n} \mathrm{Vol}(\Omega)$, since $\Omega_\varepsilon \in E \subset E_n$. Therefore, we conclude that $\Omega_\infty \in \arg\min_{\Omega \in E} \mathrm{Vol}(\Omega)$. □

One technique to solve problem (3.4) for a constraint functional $h(\cdot, \cdot)$ satisfying conditions (i) and (ii) consists in solving an approximate problem where the set of admissible parameters $\mathcal{G}$ is replaced by a convex polyhedral set $\mathcal{G}_N$ with N vertices. Theorem 3.12 suggests that, by increasing the accuracy of the approximation of $\mathcal{G}$ by $\mathcal{G}_N$, the solution of the approximate problem converges towards the solution of the original one. The approximate problems can then be solved as simple constrained optimization problem using proposition 3.9.



(a) $N = 4$ (b) $N = 8$ (c) $N = 16$

Figure 3.2: Approximation of the set of admissible parameters by convex polyhedra with an increasing number of vertices N.

This approach suffers from two notable drawbacks. The first one is the fact that none of the polyhedral sets considered by proposition 3.9 is a conservative approximation of the original set of admissible parameters $\mathbf{g}$. Therefore, for any possible approximation $\mathcal{G}_N$ of $\mathcal{G}$, denoting Ω_N the solution of the corresponding optimization problem, there exists a parameter $\tilde{\mathbf{g}} \in (\mathcal{G} \setminus \mathcal{G}_N) \neq \emptyset$ such that $h(\mathbf{g}, \Omega_N) > \tau$. A possible solution to this issue consists in considering a sequence of polyhedral sets converging towards $\mathcal{G}^*$ strictly containing $\mathcal{G}$. However, the convergence of the solutions of the approximated problems towards the solution of the original would be lost.

A second important issue from the numerical point of view concerns the number of points that are necessary to accurately approximate the finite-dimensional set $\mathcal{G}$. Indeed, as shown in [38], given a convex set $\mathcal{G} \subset \mathbb{R}^n$ of class $\mathcal{C}^2$ and a tolerance $\epsilon > 0$, the minimal number of vertices $N_{\mathcal{G},\epsilon}$ such that the Hausdorff distance between their convex hull and $\mathcal{G}$ is bounded by

$$N_{\mathcal{G},\epsilon} \geq \left(\frac{c(\mathcal{G})}{\epsilon} \right)^{\frac{n-1}{2}}, \tag{3.6}$$

where $c(\mathcal{G})$ is a constant depending on the shape of the convex set. Equation (3.6) proves that the number of vertices that are necessary to approximate a given convex set for a given precision increases exponentially with respect to the dimension of the space of parameters. Since any vertex in the approximating polyhedron corresponds to a constraint in problem (3.4), an exponentially increasing number of constraints has to be evaluated for the solution of the optimization problem, indicating that this approach suffers from the curse of dimensionality.

3.3 An approach based on subdifferentials

3.3.1 Introduction to Clarke's subdifferential

A different approach to the solution of problem (3.1) by a gradient-based method consists in differentiating directly the constraint function $\Omega \mapsto \sup_{\mathbf{g}\in\mathcal{G}} H(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$. The question of the derivative with respect to the domain of non-differentiable shape functionals has been considered in literature from different points of view. The authors of [170, 3] are interested in the optimization with respect to non-smooth functionals In [16, 85, 60], the quantity of interest consists in the first eigenvalue of different functionals, which can be expressed as minima of suitable Rayleigh quotients. In particular, the approach proposed in [60] consists in the computation of a semiderivative in the sense of Danskin [87] by applying a result from Delfour and Zolésio [100, Theorem 2.1, Chapter 10] on the sensitivity of a minimum with respect to a parameter. In this section we present an approach inspired by the methods of [129], which relies on the notion of subdifferentiability of nonsmooth functions as introduced by Clarke in [74].

At first, we recall the definitions of generalized directional derivative, strict differentiability, and subdifferential in the sense of Clarke, as found in [74, Section 2.1].

Definition 3.13 (Generalized directional derivative)**.** *Let $\mathcal{X}$ be a Banach space, $x, v \in \mathcal{X}$, and $f : \mathcal{X} \to \mathbb{R}$ a function which is Lipschitz continuous in a neighborhood of x. The generalized directional derivative of f in x in the direction v is defined as*

$$f^\circ(x; v) = \lim_{\substack{y \to x \\ t \searrow 0^+}} \frac{f(y + tv) - f(y)}{t}.$$

Definition 3.14 (Strictly differentiable function)**.** *Let us denote $\mathcal{X}^*$ the topological dual of the Banach space $\mathcal{X}$. A real-valued function f defined on $\mathcal{X}$ is strictly differentiable in $x \in \mathcal{X}$ if it admits a generalized directional derivative $f^\circ(x; v)$ for all $v \in \mathcal{X}$, and there exists a linear functional $\mathrm{D}_x^s \in \mathcal{X}^*$ such that*

$$f^\circ(x; v) = \mathrm{D}_x^s(v)$$

for all $v \in \mathcal{X}$.

Next, we present the notion of subdifferential in the sense of Clarke, introduced in [74] as "generalized gradient" and extending the results of Danskin [87].

Definition 3.15 (subdifferential in the sense of Clarke)**.** *Let $f : \mathcal{X} \to \mathbb{R}$ be a function defined on the Banach space $\mathcal{X}$, which is Lipschitz continuous in a neighborhood of $x \in \mathcal{X}$. The subdifferential in the sense of Clarke of f in x is the subset of $\mathcal{X}^*$ defined as*

$$\partial f(x) = \{L \in \mathcal{X}^* \ : \ f^\circ(x; v) \geq L(v) \text{ for all } v \in \mathcal{X}\}.$$

We state now a result presented as [74, Corollary 2 of Theorem 2.8.2] and referred in [129, Section 4.2].

Proposition 3.16. *Let $\mathcal{X}$ be a Banach space, $\mathcal{T}$ a metrizable, sequentially compact topological space, $x \in \mathcal{X}$, and $\{f(\cdot,t)\}_{t\in\mathcal{T}}$ a family of functions such that*

(1) $y \mapsto f(y,t)$ is Lipschitz continuous for all y in a neighborhood U_x of x and for all $t \in \mathcal{T}$;

(2) for any $y \in U_x$ the mapping $t \mapsto f(y,t)$ is upper-semicontinuous;

(3) the set $\{f(x,t) \ : \ t \in \mathcal{T}\} \subset \mathbb{R}$ is bounded;

(4) $f(\cdot,t)$ is strictly differentiable in U_x, and the strict derivative $\mathrm{D}_y^s f(y,t)$ is continuous in $U_x \times \mathcal{T}$.

We denote $F : \mathcal{X} \to \mathbb{R}$ the function mapping $y \mapsto F(y) = \sup_{t\in\mathcal{T}} f(y,t)$, which is defined, finite, and Lipschitz continuous in U_x. Moreover, we denote $\mathcal{T}(y) = \{t \in \mathcal{T} \ : \ f(y,t) = F(y)\}$ the subset of the set of the parameters $\mathcal{T}$ where the maximum in the definition of $F(\cdot)$ is attained, and we remark that $\mathcal{T}(y)$ is not empty for all $y \in U_x$.

Then, the subdifferential in the sense of Clarke of F exists for all $y \in U_x$ and is given by:

$$\partial F\,(y) = \left\{ \int_{\mathcal{T}} \mathrm{D}_y^s f(y,t)\,\mathrm{d}\mu(t) \ : \ \mu \in P[\mathcal{T}(y)] \right\}, \tag{3.7}$$

where the notation $P[\mathcal{S}]$ denotes the collection of probability Radon measures on a measurable subset $\mathcal{S}$ of $\mathcal{T}$.

Remark 3.17. *In [74, Section 2.2] is recalled that any continuously differentiable function in $x \in \mathcal{X}$ is also strictly differentiable in x. Thus, in proposition 3.16 we can suppose that $f(\cdot,t)$ is differentiable in U_x and the derivative $\frac{\partial f}{\partial y}(y,t)$ is continuous on $U_x \times \mathcal{T}$, and the proposition still holds true.*

Let $f : \mathcal{X} \times \mathcal{T} \to \mathbb{R}$ satisfy the hypotheses of proposition 3.16, and $F : \mathcal{X} \to \mathbb{R}$ be the mapping $x \mapsto \max_{t\in\mathcal{T}} f(x,t)$. Finally, let us consider $y \in \mathcal{X}$, and $\bar{t} \in \mathcal{T}(y)$ a parameter for which $f(x,t) = F(x)$. Since $P[\mathcal{T}(y)]$ contains the Dirac measure concentrated in $\bar{t}$, the derivative $\frac{\partial f}{\partial y}(y,\bar{t}) \in \mathcal{X}^*$ belongs to the subdifferential $\partial F\,(y)$. Moreover, if F is differentiable in $y \in \mathcal{X}$, the subdifferential reduces to a singleton.

3.3.2 Application to shape optimization problems

Similarly to what has been done in section 3.2.3, we consider a generic optimization problem without an explicit expression of the state like problem (3.4). For the sake of simplicity, we denote $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ the shape functional defined as

$$\Phi(\Omega) = \sup_{\mathbf{g}\in\mathcal{G}} h\,(\mathbf{g},\Omega). \tag{3.8}$$

Unfortunately, we cannot apply directly proposition 3.16 to differentiate Φ, since the space $\mathcal{S}_{\mathrm{adm}}$ provided with the Hausdorff metric m_{H} defined in (3.6) is not a Banach space. Such issue can be bypassed thanks to the definition of the shape derivative according to Hadamard introduced in section 1.2.1. Indeed, for a given admissible domain $\Omega \in \mathcal{S}_{\mathrm{adm}}$, the deformation field $\boldsymbol{\theta}$ at the core of Hadamard's moving boundaries approach belongs to the Banach space $\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$. With this in mind, we can extend the concepts of subdifferential to shape functionals.

Definition 3.18 (Subdifferential of a shape functional)**.** *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a domain in $\mathbb{R}^d$, and $J : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ a shape functional such that the mapping $\boldsymbol{\theta} \mapsto J(\Omega_{\boldsymbol{\theta}})$ admits a Gâteaux derivative $\mathrm{d}J(\Omega;\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$. Then, the subdifferential of J in Ω is defined as*

$$\partial J(\Omega) = \left\{ L \in \left(\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \right)^* \ : \ \mathrm{d}J(\Omega;\boldsymbol{\theta}) \geq L(\boldsymbol{\theta}) \text{ for all } \boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \right\}.$$

We can now state a result for shape functionals analogous to proposition 3.16.

Proposition 3.19. *Let $\mathcal{S}_{\mathrm{adm}}$ be a family of uniformly bounded open domains in $\mathbb{R}^d$ endowed with the topology induced by the Hausdorff metric m_{H}, and $\mathcal{G}$ a compact subset of a Banach space $(\mathcal{Y}, \|\cdot\|)$. Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be an admissible domain, and $h : \mathcal{G} \times \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ a shape functional such that:*

(i) $\widetilde{\Omega} \mapsto h\left(\mathbf{g}, \widetilde{\Omega}\right)$ is Lipschitz continuous in a neighborhood U_Ω of Ω for all $\mathbf{g} \in \mathcal{G}$;

(ii) $\mathbf{g} \mapsto h\left(\mathbf{g}, \widetilde{\Omega}\right)$ is convex and bounded for all $\tilde{\Omega} \in U_\Omega$;

(iii) $\widetilde{\Omega} \mapsto h\left(\mathbf{g}, \widetilde{\Omega}\right)$ is Fréchet differentiable in U_Ω for any choice of the parameter $\mathbf{g} \in \mathcal{G}$;

(iv) The Fréchet derivative $\frac{\partial h}{\partial \Omega}(\mathbf{g}, \tilde{\Omega}) \in \left(\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \right)^*$ is continuous in $\mathcal{G} \times U_\Omega$.

We denote $\Phi : \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ the shape functional

$$\Omega \mapsto \Phi(\Omega) = \max_{\mathbf{g} \in \mathcal{G}} h(\mathbf{g}, \Omega),$$

where the maximum is attained thanks to the convexity of $h(\cdot, \mathbf{g})$ and compactness of $\mathcal{G}$. Then, the functional $\Phi(\cdot)$ admits a subdifferential $\partial\Phi(\Omega)$ in Ω, and its expression is given by

$$\partial\Phi(\Omega) = \left\{ \int_{\mathcal{G}} \frac{\partial h}{\partial \Omega}(\mathbf{g}, \Omega)\, \mathrm{d}\mu(\mathbf{g}) \ : \ \mu \in P[\mathcal{G}(\Omega)] \right\} \subset \left(\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \right)^*.$$

Proof. We consider the class $\Theta_\Omega \subset \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ of admissible deformations defined as

$$\Theta_\Omega = \left\{ \boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \ : \ \Omega_{\boldsymbol{\theta}} \in U_\Omega \right\}.$$

We introduce the function $f_\Omega : \mathcal{G} \times \Theta_\Omega \to \mathbb{R}$ mapping $(\mathbf{g}, \boldsymbol{\theta}) \mapsto f_\Omega(\mathbf{g}, \boldsymbol{\theta}) = h(\mathbf{g}, \Omega_{\boldsymbol{\theta}})$. In order to prove proposition 3.19, we verify that f_Ω satisfies all the hypotheses of proposition 3.16. At first, we observe that the set $\mathcal{G}$ is compatible with the hypotheses of proposition 3.16, since it is a compact subset of the Banach space $\mathcal{Y}$ with respect to the Hausdorff metric m_{H}. The set Θ_Ω is a neighborhood of the origin in the Banach space $\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$.

The conditions (1) - (3) of proposition 3.16 are satisfied by f_Ω thanks to assumptions (i) and (ii). In particular, the continuity of $f_\Omega(\cdot, \boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Theta_\Omega$ is ensured by the convexity of $\mathbf{g} \mapsto h\left(\mathbf{g}, \widetilde{\Omega}\right)$ for all $\widetilde{\Omega} \in U_\Omega$. The existence and continuity of the strict derivative of $\boldsymbol{\theta} \mapsto f_\Omega(\mathbf{g}, \boldsymbol{\theta})$ follow from assumptions (iii) and (iv) and from remark 3.17. Therefore, the function

$\boldsymbol{\theta} \mapsto F_\Omega(\boldsymbol{\theta}) = \max_{\mathbf{g}\in\mathcal{G}} f_\Omega(\mathbf{g}, \boldsymbol{\theta})$ is well defined, and it admits a subdifferential $\partial F_\Omega(\boldsymbol{\theta})$ with the following expression

$$\partial F_\Omega(\boldsymbol{\theta}) = \left\{ \int_{\mathcal{G}} \frac{\partial f_\Omega}{\partial \boldsymbol{\theta}}(\mathbf{g}, \boldsymbol{\theta}) \, \mathrm{d}\mu(\mathbf{g}) \ : \ \mu \in P[\mathcal{G}(\Omega)] \right\} \subset \left(\mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d \right)^* .$$

Thanks to the definitions of the functionals Φ and F_Ω we have that, for all $\boldsymbol{\theta} \in \Theta_\Omega$, $\Phi(\Omega_{\boldsymbol{\theta}}) = F_\Omega(\boldsymbol{\theta})$ and, in particular, $\Phi(\Omega) = F_\Omega(\mathbf{0})$. Thus, we conclude that $\partial\Phi(\Omega) = \partial F_\Omega(\mathbf{0})$. □

Proposition 3.19 provides a method to compute elements of the subdifferential $\partial\Phi(\Omega)$ by computing a value of the parameter $\mathbf{g} \in \mathcal{G}$ where $h(\Omega, \mathbf{g})$ attains its maximum. Differently from the case studied in section 3.2, the convexity of the set $\mathcal{G}$ of external loads is not required. It should also be remarked that any element of the subdifferential defines a direction of descent for the functional Φ, albeit not necessarily optimal if $\partial\Phi(\Omega)$ is not a singleton.

3.3.3 Algorithmic implementation

Let us consider the same notations of proposition 3.19: $\mathcal{G}$ is a compact subset of a Banach space $(\mathcal{Y}, \|\cdot\|)$, $h : \mathcal{G} \times \mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ a function satisfying the conditions (i)-(iv), and $\Phi(\cdot)$ the shape functional mapping $\Omega \mapsto \max_{\mathbf{g}\in\mathcal{G}} h(\mathbf{g}, \Omega)$ as in (3.8). In section 3.3.2 we provided the theoretical framework for the computation of the subdifferential of $\Phi(\cdot)$.

Here we provide a procedure to compute one element to the subdifferential $\partial\Phi(\Omega)$ in the context of the optimization algorithm presented in section 1.3. The procedure can be divided in two steps. First, we identify a parameter $\overline{\mathbf{g}} \in \arg\max_{\mathbf{g}\in\mathcal{G}} h(\mathbf{g}, \Omega) \subset \mathcal{G}$ for which the maximum of $h(\cdot, \Omega)$ is attained. The maximum is attained in at least one point, since $\mathcal{G}$ is compact, and $h(\cdot, \Omega)$ is convex and bounded (see proposition 3.4). Next, the shape derivative of the term $h(\overline{\mathbf{g}}, \Omega)$ is computed using the classical methods of section 1.2. Proposition 3.19 ensures that the shape derivative of $h(\overline{\mathbf{g}}, \Omega)$ belongs to the subdifferential $\partial\Phi(\Omega)$.

Different possible methods can be considered to identify the parameter $\overline{\mathbf{g}}$ depending on the nature of the set $\mathcal{G}$ and the function $h(\cdot, \Omega)$. If the mapping $\mathbf{g} \mapsto h(\mathbf{g}, \Omega)$ is differentiable with respect to $\mathbf{g}$ and $\mathcal{G}$ is a subset of a Hilbert space, a simple gradient-descent method can be implemented to identify $\overline{\mathbf{g}}$. If further hypotheses apply on the constraint functional or on the set of admissible parameters, *ad hoc* methods can be used. An example for the case where $\mathbf{g} \mapsto h(\mathbf{g}, \Omega)$ is a quadratic function and $\mathcal{G}$ an ellipsoid is provided in section 3.4.1.

3.4 Numerical results

3.4.1 3D Cantilever

As first numerical application, we consider the optimization of a 3D cantilever structure under a constraint on the mechanical compliance. The initial condition of the structure is presented in fig. 3.3: the structure is clamped on the four corners marked as Γ_{D}, and a mechanical load $\mathbf{g}$ is applied on the region Γ_{N} on the opposite side.

We suppose that the load $\mathbf{g}$ applied to Γ_{N} consists of two components: one of traction-compression (oriented along the x axis), and a vertical one (along the z axis):

$$\mathbf{g} = X\mathbf{e}_x + Z\mathbf{e}_z. \tag{3.9}$$

We suppose also that the parameters X and Z belong to the intervals $[-\overline{g}_x, \overline{g}_x]$ and $[-\overline{g}_z, \overline{g}_z]$ respectively. Moreover, we suppose that they are bounded by the following inequality:

$$\frac{X^2}{\overline{g}_x^2} + \frac{Z^2}{\overline{g}_z^2} \leq 1. \tag{3.10}$$

The inequality (3.10) states that the set of admissible mechanical loads can be parametrized by an ellipse in $\mathbb{R}^2$ with semi-axes equal to $\overline{g}_x$ and $\overline{g}_z$. To sum up, the optimization problem can be expressed as

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the volume } \mathrm{Vol}(\Omega) \\ \text{under the constraint } \sup_{\mathbf{g}\in\mathcal{G}} \mathcal{C}\left(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega\right) \leq \tau, \\ \text{where the displacement } \mathbf{u}_{\Omega,\mathbf{g}} \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)^d \\ \text{solves the elasticity equation} \\ \quad \left\{\begin{array}{rcll} -\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_{\Omega,\mathbf{g}} & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \end{array}\right. \tag{3.11}$$

The numerical parameters considered for this problem are reported in table 3.1. The simulations of this section and of section 3.4.2 have been performed on a Virtualbox virtual machine Linux with 1GB of dedicated memory, installed on a Dell PC equipped with a 2.80 GHz Intel i7 processor.

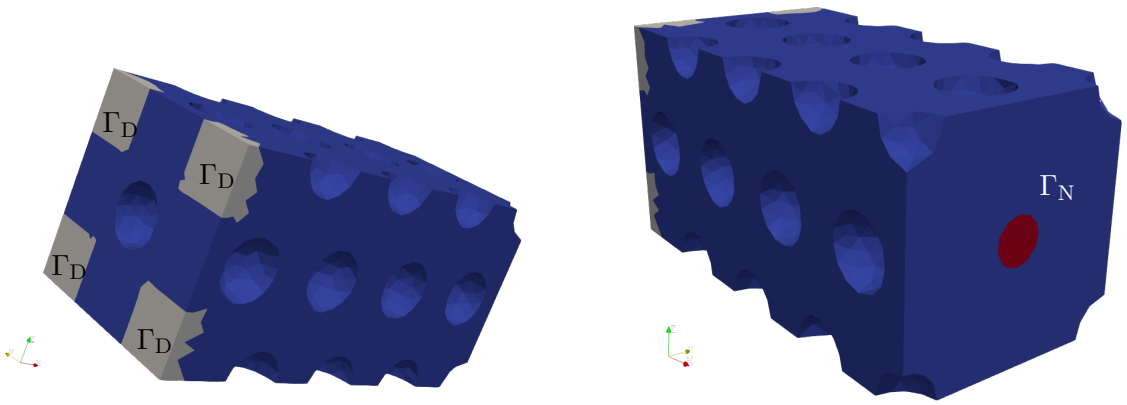


Figure 3.3: Structure of the 3D cantilever structure. The region Γ_{N} where the uncertain mechanical load is applied is marked in red, while the clamping region Γ_{D} is highlighted in grey.

In order to solve the optimization problem (3.11) we consider both the polyhedral approximation approach of 3.2, and the method based on the subdifferential as in section 3.3. Both methods can be applied since the set $\mathcal{G}$ is convex, the mapping $\mathbf{g} \mapsto \mathcal{C}\left(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega\right)$ is a convex function, and the compliance operator satisfies the conditions (i)-(iv) of proposition 3.19.

Geometry of the structure		
cross section length	ℓ_s	1.0 cm
longitudinal length	ℓ_x	2.0 cm
sidelength of Γ_{D}		0.3 cm
radius of Γ_{N}		0.1 cm
Elastic coefficients		
Young's modulus	E	200 MPa
Poisson's ratio	ν	0.3
Mechanical loads		
compression load	$\overline{g}_x$	25 kPa
vertical load	$\overline{g}_z$	10 kPa
Mesh size parameters		
minimal mesh size	`hmin`	0.025 cm
maximal mesh size	`hmax`	0.10 cm
Thresholds for the inequality constraints		
threshold on the compliance	τ	2.5×10^{-2} kPa cm^3
bound on the probability of failure	$\bar{p}$	1.0%

Table 3.1: Numerical data concerning the geometry and the mechanics of the cantilever structure of fig. 3.3.

For the polyhedral approach, we approximated the ellipse $\mathcal{G}$ by polygons with 4, 8, and 16 vertices denoted $\mathcal{G}_4$, $\mathcal{G}_8$, and $\mathcal{G}_{16}$ respectively. The polygons $\mathcal{G}_4$, $\mathcal{G}_8$, and $\mathcal{G}_{16}$ are defined as convex hulls of N points as follows

$$\begin{aligned}
\mathcal{G}_4 &= \operatorname{hull}\left\{\left(\overline{g}_x \sin\left(\tfrac{2n\pi}{4}\right)\mathbf{e}_x + \overline{g}_z \cos\left(\tfrac{2n\pi}{4}\right)\mathbf{e}_y\right) \;:\; n \in \{0,\dots,3\}\right\} \subset \mathbb{R}^2;\\
\mathcal{G}_8 &= \operatorname{hull}\left\{\left(\overline{g}_x \sin\left(\tfrac{2n\pi}{8}\right)\mathbf{e}_x + \overline{g}_z \cos\left(\tfrac{2n\pi}{8}\right)\mathbf{e}_y\right) \;:\; n \in \{0,\dots,7\}\right\} \subset \mathbb{R}^2;\\
\mathcal{G}_{16} &= \operatorname{hull}\left\{\left(\overline{g}_x \sin\left(\tfrac{2n\pi}{16}\right)\mathbf{e}_x + \overline{g}_z \cos\left(\tfrac{2n\pi}{16}\right)\mathbf{e}_y\right) \;:\; n \in \{0,\dots,15\}\right\} \subset \mathbb{R}^2.
\end{aligned}$$

It should be remarked that, since the compliance is invariant with respect to a change of sign in the applied load, it is necessary to consider only half of the vertices of $\mathcal{G}_4$, $\mathcal{G}_8$, and $\mathcal{G}_{16}$ to define the constraints of the approximated optimization problem. The structures optimized for the three cases are denoted as Ω_4, Ω_8, and Ω_{16} respectively, and are represented in fig. 3.4, fig. 3.5, and fig. 3.6.

In the subdifferential approach, it is necessary to identify the parameter $\overline{\mathbf{g}}$ maximizing $\mathcal{C}(\cdot,\Omega)$ at each step of the optimization. We can assume, by proposition 3.4, that $\overline{\mathbf{g}}$ belongs to the boundary of $\mathcal{G}$. Therefore, there exists an angle $\alpha \in [0, 2\pi)$ such that

$$\overline{\mathbf{g}} = \sin\alpha\, g_x \mathbf{e}_x + \cos\alpha\, g_z \mathbf{e}_z.$$

Thanks to the symmetry of the compliance operator, we can restrict the search for α to the interval $[-\frac{\pi}{2}, \frac{\pi}{2})$. The angle α yielding the maximal compliance for a given shape Ω can be identified by interpreting the compliance as a quadratic functional. Indeed, there exists a matrix $\mathbf{M}_\Omega \in \mathbb{R}^{2\times 2}$ such that, for all $\hat{\alpha} \in [0, 2\pi)$,

$$\mathcal{C}(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega) = (\sin\hat{\alpha},\, \cos\hat{\alpha})\mathbf{M}_\Omega \begin{pmatrix} \sin\hat{\alpha} \\ \cos\hat{\alpha} \end{pmatrix},$$

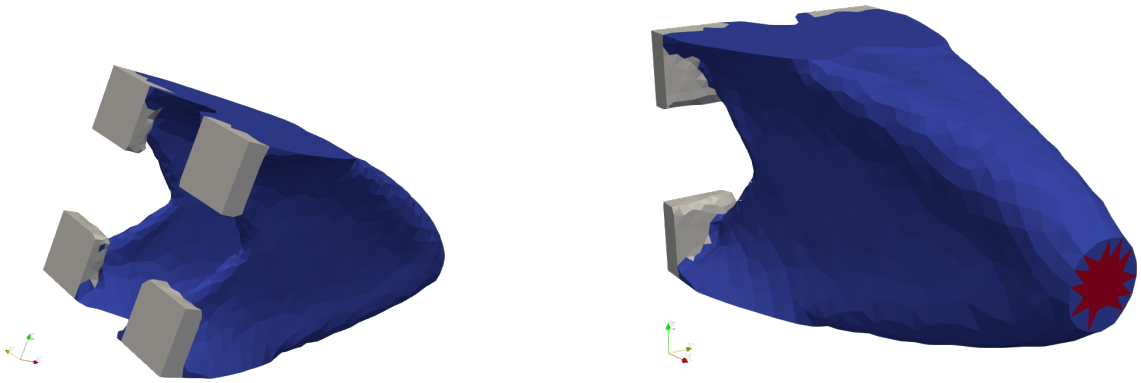


Figure 3.4: Optimal shape Ω_4 resulting from the polyhedron approach with $N = 4$ vertices.

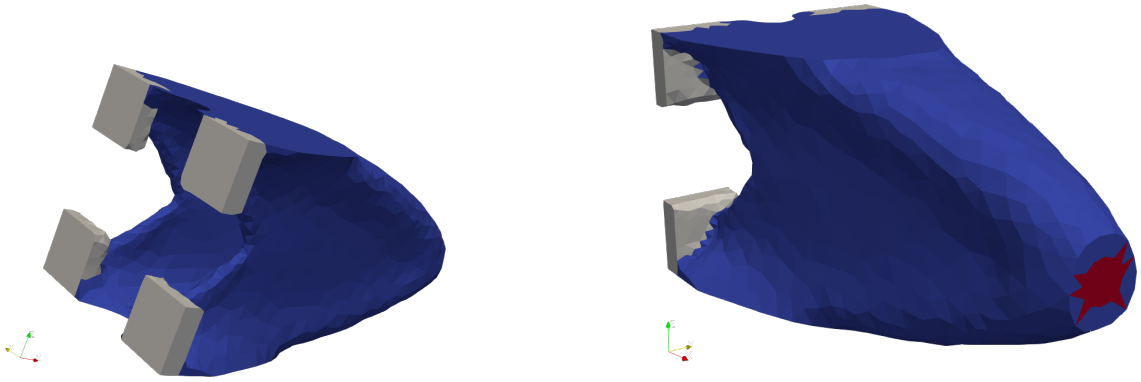


Figure 3.5: Optimal shape Ω_8 resulting from the polyhedron approach with $N = 8$ vertices.

where the load associated to $\mathbf{u}_{\Omega,\mathbf{g}}$ is $\mathbf{g}(\hat{\alpha}) = \sin\hat{\alpha} g_x \mathbf{e}_x + \cos\hat{\alpha} g_z \mathbf{e}_z$. The entries of the matrix $\mathbf{M}_\Omega$ are

$$[\mathbf{M}_\Omega]_{11} = m_{11} = \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}_x}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega,\mathbf{g}_x}\right) \mathrm{d}\mathbf{x},$$
$$[\mathbf{M}_\Omega]_{22} = m_{22} = \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}_z}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega,\mathbf{g}_z}\right) \mathrm{d}\mathbf{x},$$
$$[\mathbf{M}_\Omega]_{12} = [\mathbf{M}_\Omega]_{21} = m_{12} = \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}_x}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_{\Omega,\mathbf{g}_z}\right) \mathrm{d}\mathbf{x}.$$

The angle α for which the maximum of the compliance is attained depends on the eigenvector related to the maximal eigenvalue of $\mathbf{M}_\Omega$. In particular, α can be computed explicitly by the following expression

$$\alpha = \begin{cases} \frac{\pi}{4} - \frac{\beta}{2} & \text{if } m_{12} \geq 0, \\ \frac{3\pi}{4} + \frac{\beta}{2} & \text{if } m_{12} < 0, \end{cases} \qquad \text{where } \beta = \arcsin\left(\frac{m_{22} - m_{11}}{2\sqrt{\left(\frac{m_{22}-m_{11}}{2}\right)^2 + m_{12}^2}}\right). \tag{3.12}$$

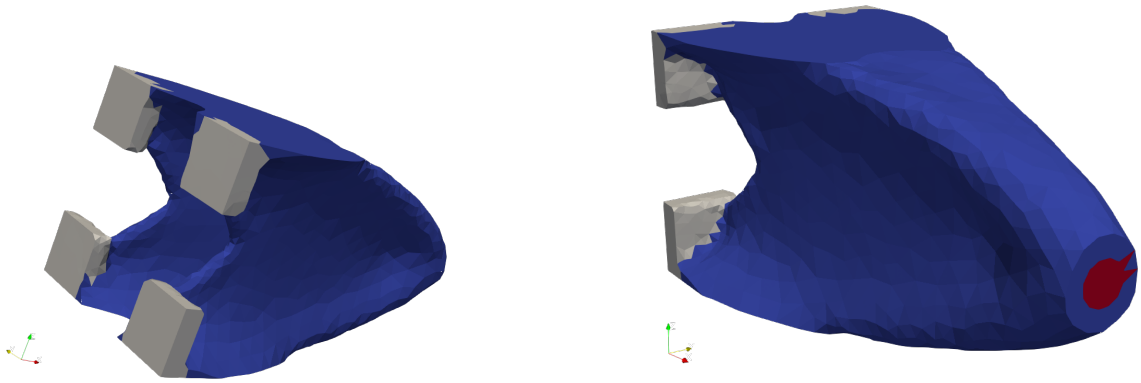


Figure 3.6: Optimal shape Ω_{16} resulting from the polyhedron approach with $N = 16$ vertices.

The optimal shape resulting from the optimization based on the subdifferential is reported in fig. 3.7, and we denote it as Ω_S. In the graph of fig. 3.8 is reported the evolution of the angle α along the iterations. We remark that α oscillates around 0, underlying the fact that vertical loads which are orthogonal to the main axis of the cantilever are responsible for the largest values of the compliance.

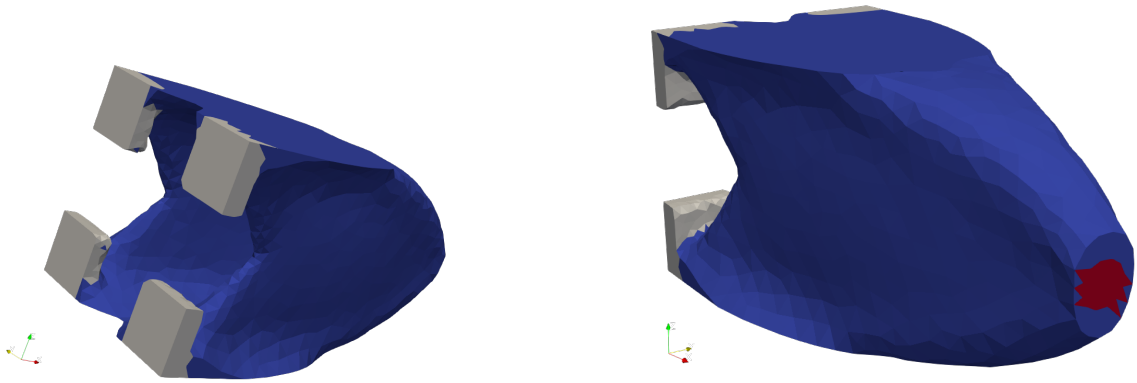


Figure 3.7: Optimal shape Ω_S resulting from the subdifferential approach.

In table 3.2 we reported the numerical results of the optimization of the cantilever using the method of polyhedral approximation with three increasing degrees of precision, as well as the results of the subdifferential technique. The graph showing the progressive decrease of the volume of the structure is presented in fig. 3.9a, while fig. 3.9 follows the evolution of the constraint in each numerical example.

A first remark concerns the slow rate of convergence of the four examples, as shown in fig. 3.9a. This issue seems to be proper to the 3D cantilever structure, as pointed out also in [112, Section 6.2.1]. Next, we can observe in fig. 3.9b that in all four cases the constraint on

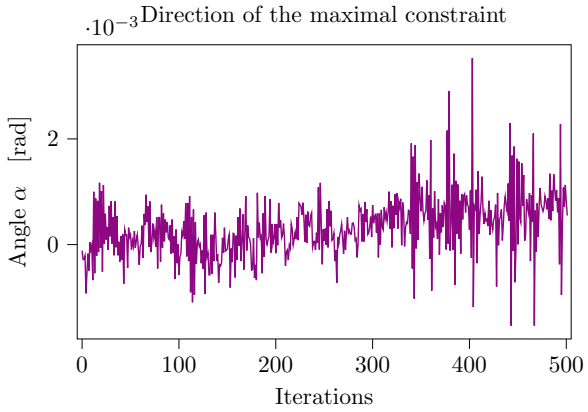
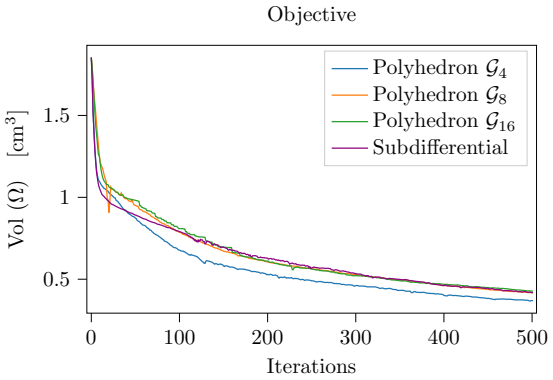


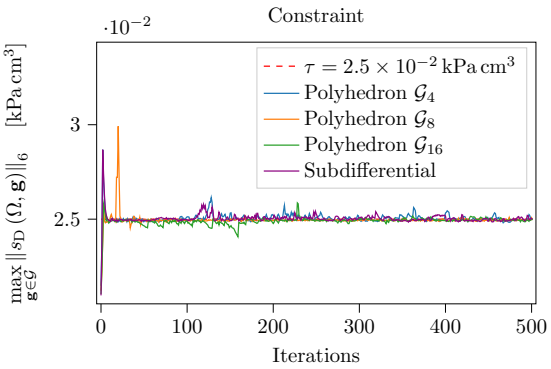
Figure 3.8: Evolution of the direction of the maximal constraint (in terms of the angle α) during the optimization process.

		Polyhedron			Subdifferential
		$N = 4$	$N = 8$	$N = 16$	
Optimization					
Duration	[min]	180	187	208	180
Number of iterations		500	500	500	500
Results					
Final volume Vol(Ω)	[cm^3]	0.369	0.418	0.428	0.419
Maximal constraint	[kPa cm^3]	2.500×10^{-2}	2.495×10^{-2}	2.496×10^{-2}	2.499×10^{-2}
$d_{\mathrm{H}}(\Omega_S, \Omega_N)$	[cm]	0.1561	0.1361	0.2130	–

Table 3.2: Numerical results for the optimization of the volume of the cantilever under constraints on the mechanical compliance, obtained using the Polyhedron method (with an increasing number of vertices), and the Subdifferential method.



(a) Evolution of the objective function.



(b) Evolution of the constraint for three instances of the polyhedron method, and the Subdifferential method.

Figure 3.9: Convergence of the objective (volume) and the constraint (compliance) for the cantilever.

the maximum of the compliance is satisfied. By comparing the duration of the four simulations we can state that the method based on the subdifferential is efficient and reliable to solve problem (3.11) since it yields a similar result as the three other simulations while requiring fewer computations of the shape derivative.

Finally, we can see that the four simulations yield similar results, as a consequence of the preeminence of the vertical load to the optimization of the structure. The fact that the Hausdorff distances between Ω_N and Ω_S is of the order of the mesh size for each $N \in \{4, 8, 16\}$ supports the conclusion that all four simulations have reached a result close to the exact solution of problem (3.11).

3.4.2 Disc

In this section we consider the optimization of a cylinder-like structure. Once again we aim to minimize its volume, but we replace the constraint on the compliance on a constraint on the L^6-norm of the von Mises stress. The initial condition is presented in fig. 3.10: the structure is fixed on a region Γ_D on its side, while shear loads are applied tangentially to a ring-like surface Γ_N on the top of the cylinder. The optimization problem to be solved is the following:

$$\left|\begin{array}{l}
\text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\
\text{minimizing the volume } \mathrm{Vol}(\Omega) \\
\text{under the constraint } \sup\limits_{\mathbf{g}\in\mathcal{G}} \left\| s_\mathrm{D}\left(\mathbf{u}_\Omega\right)\right\|_6 \leq \tau, \\
\text{where the displacement } \mathbf{u}_{\Omega,\mathbf{g}} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}\left(\Omega\right)^d \\
\text{solves the elasticity equation} \\
\quad \left\{\begin{array}{rcll}
-\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{g} & \text{on } \Gamma_\mathrm{N}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_{\Omega,\mathbf{g}}\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\
\mathbf{u}_{\Omega,\mathbf{g}} & = & \mathbf{0} & \text{on } \Gamma_\mathrm{D}.
\end{array}\right.
\end{array}\right. \tag{3.13}$$

Similarly to the model considered in the previous section, we suppose that the load $\mathbf{g}$ can be written as sum of two terms, aligned with the axes x and y

$$\mathbf{g} = X\mathbf{e}_x + Y\mathbf{e}_y.$$

We suppose that the intensity of the applied force is bounded by $\overline{g}$, so that the set of admissible loads $\mathcal{G}$ can be parametrized by a circle in $\mathbb{R}^2$ with radius $\overline{g}$. The geometric and material properties of the structure, the mesh size, the maximal value of the applied force and the threshold τ on the L^6-norm of the von Mises stress are reported in table 3.3.

Similarly to the previous section, we consider three different approximations for the polyhedral approach, where $\mathcal{G}$ is replaced by inscribed regular polygons with $N = 4$, 8, and 16 vertices denoted $\mathcal{G}_4$, $\mathcal{G}_8$, and $\mathcal{G}_{16}$ respectively. These polygons can thus be defined as convex hulls of N points as follows

$$\begin{array}{rll}
\mathcal{G}_4 = & \mathrm{hull}\left\{\overline{g}\left(\sin\left(\frac{2n\pi}{4}\right)\mathbf{e}_x + \cos\left(\frac{2n\pi}{4}\right)\mathbf{e}_y\right) \ : \ n \in \{0,\dots,3\}\right\} & \subset \mathbb{R}^2; \\
\mathcal{G}_8 = & \mathrm{hull}\left\{\overline{g}\left(\sin\left(\frac{2n\pi}{8}\right)\mathbf{e}_x + \cos\left(\frac{2n\pi}{8}\right)\mathbf{e}_y\right) \ : \ n \in \{0,\dots,7\}\right\} & \subset \mathbb{R}^2; \\
\mathcal{G}_{16} = & \mathrm{hull}\left\{\overline{g}\left(\sin\left(\frac{2n\pi}{16}\right)\mathbf{e}_x + \cos\left(\frac{2n\pi}{16}\right)\mathbf{e}_y\right) \ : \ n \in \{0,\dots,15\}\right\} & \subset \mathbb{R}^2.
\end{array}$$

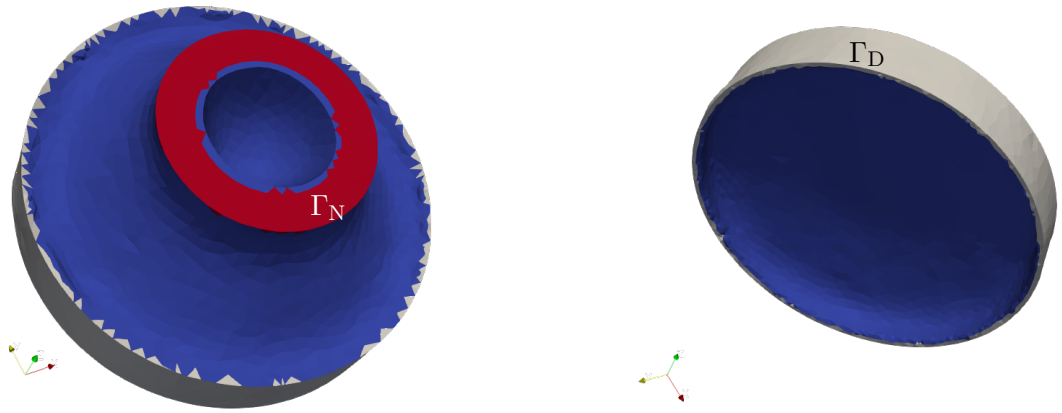


Figure 3.10: Structure of the 3D disc structure. The region Γ_N where the random load is applied is marked in red, while the clamping region Γ_D is highlighted in grey.

Geometry of the structure		
height of the domain		12.0 cm
maximal radius of the romain		12.0 cm
Region Γ_N		
inner radius of Γ_N		4.0 cm
outer radius of Γ_N		6.0 cm
Region Γ_D		
thickness of Γ_D		2.0 cm
Mesh size parameters		
minimal mesh size	`hmin`	0.75 cm
maximal mesh size	`hmax`	1.25 cm
Elastic coefficients		
Young's modulus	E	200 MPa
Poisson's ratio	ν	0.3
Mechanical loads		
maximal load in any direction	$\overline{g}$	10 kPa
threshold on $\|s_D\|_6$	τ	5.0 kPa

Table 3.3: Numerical data concerning the geometry and the mechanics of the disc structure of fig. 3.10.

Thanks to the symmetry of the constraint with respect to a change of sign in the applied force (and thus in the displacement $\mathbf{u}_{\Omega,\mathbf{g}}$), only $N/2$ constraints need to be evaluated at each step of the solution of problem (3.13). The structures resulting from applying the polyhedral approximation method are shown in fig. 3.11, fig. 3.12, and fig. 3.13 and denoted as Ω_4, Ω_8, and Ω_{16} respectively.

Contrarily to the example discussed in section 3.4.1, the increasing refinement in the approximation of $\mathcal{G}$ results in structures that differ significantly from one another. Indeed, we can see how Ω_4, represented infig. 3.11, is optimized to resist the forces applied in the directions of the four edges of $\mathcal{G}_4$. The structure Ω_8 of fig. 3.12 is similar, but its four branches are wider, responding to forces oriented in the direction bisecting the main axes. Finally, Ω_{16}, shown in fig. 3.13, is characterized by a rotational symmetry, thus resisting to forces applied in 16 different directions.

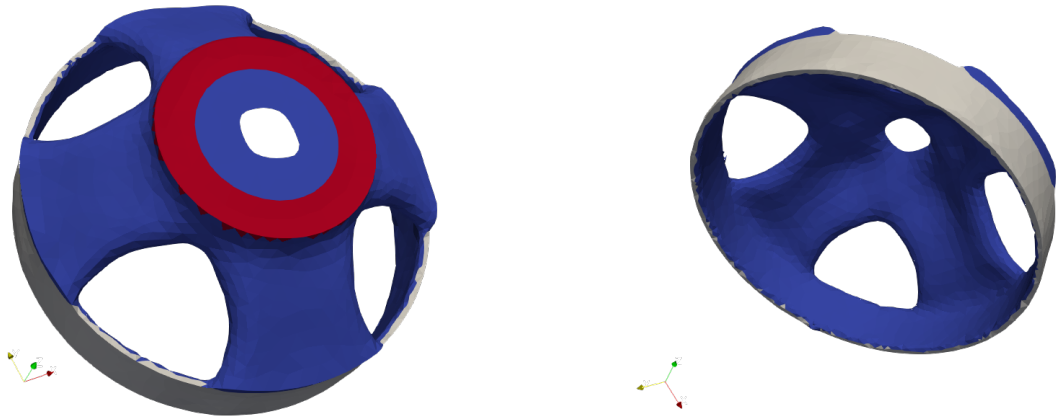


Figure 3.11: Optimal shape Ω_4 resulting from the polyhedron approach with $N = 4$ vertices.

The similarities between the two cases allow to use a similar parametrization of the set $\mathcal{G}$. Let us denote $\overline{\mathbf{g}}$ the element of $\mathcal{G}$ maximizing the constraint functional. Thanks to the convexity of the mapping $\mathbf{g} \mapsto \|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_6$ proposition 3.4 applies, and we deduce that $\overline{\mathbf{g}}$ belongs to the boundary of $\mathcal{G}$. Thus, there exist $\alpha \in [0, 2\pi)$ such that

$$\overline{\mathbf{g}} = \overline{g}\left(\sin\alpha\,\mathbf{e}_x + \cos\alpha\,\mathbf{e}_y\right).$$

The constraint functional considered in problem (3.13) is the L^6-norm of the von Mises stress, which is not a quadratic function. Therefore, the method used in section 3.4.1 to identify the load maximizing the constraint functional cannot be applied. Instead, we identify the value of α maximizing the constraint function by applying the Newton method to the function $\alpha \mapsto \left(\|s_{\mathrm{D}}(\mathbf{u}_\Omega)\|_6\right)^6$. It should be remarked that such a function can be expressed analytically in terms of the displacement fields generated by the application of the loads $\overline{g}\mathbf{e}_x$ and $\overline{g}\mathbf{e}_y$. Thus its evaluation is extremely fast and does not require the solution of an expensive boundary value problem. Once again, thanks to the symmetry of the constraint under a change of sign of the mechanical load, the search of the critical direction α can be limited to the interval $[-\frac{\pi}{2}, \frac{\pi}{2})$. The shape Ω_S resulting from the application of the subdifferential approach is reported in fig. 3.14.

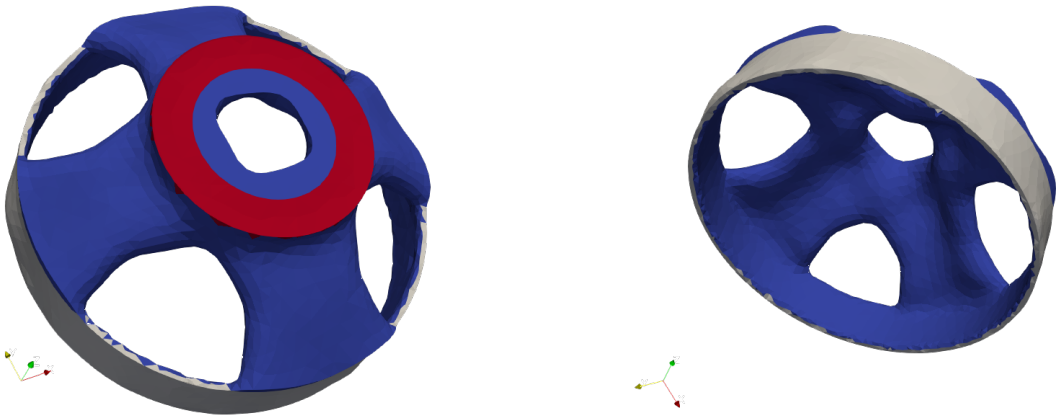


Figure 3.12: Optimal shape Ω_8 resulting from the polyhedron approach with $N = 8$ vertices.

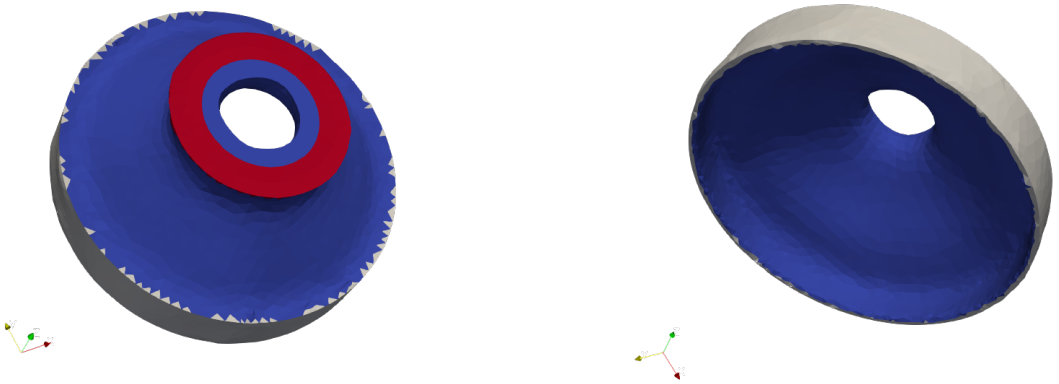


Figure 3.13: Optimal shape Ω_{16} resulting from the polyhedron approach with $N = 16$ vertices.

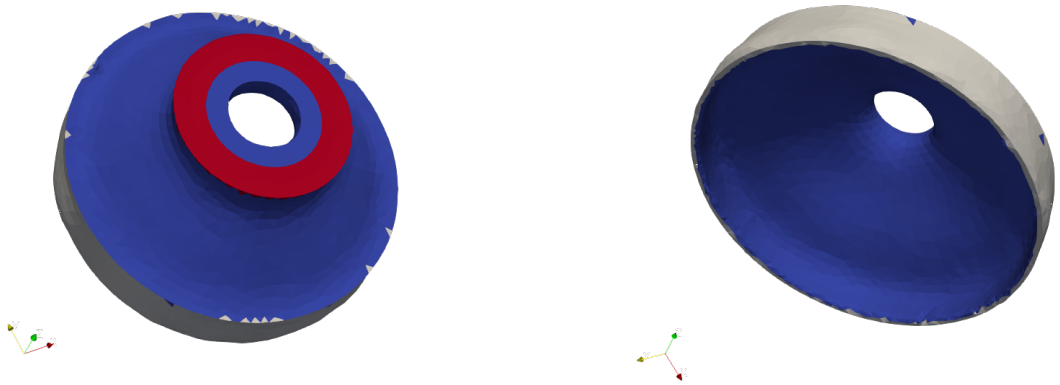


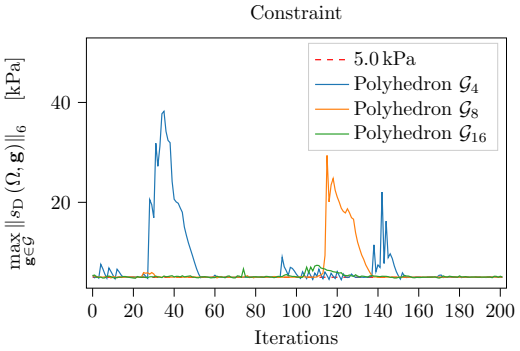
Figure 3.14: Optimal shape Ω_S resulting from the Subdifferential approach.

The numerical results of the optimization performed using the polyhedral approximation and the subdifferential method are reported in table 3.4. In fig. 3.15 we plotted the value of the maximal constraint throughout the optimization: firstly for the three instances of the polyhedral approximation algorithm (fig. 3.15a), next comparing them with the evolution of the constraints for the subdifferential approach (fig. 3.15b). The trend of the objective function for all four simulations is represented in fig. 3.16a. In fig. 3.16b we report the evolution of the angle α parametrizing the direction of the load maximizing the L^6-norm of the von Mises stress at each step.

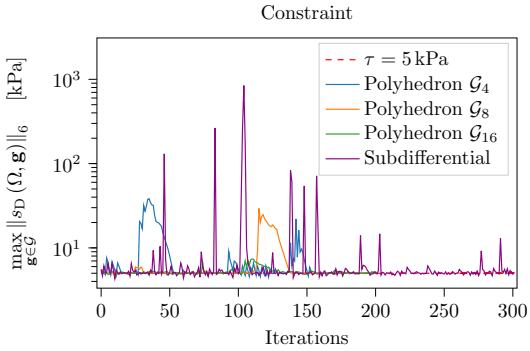
The method of subdifferential yields an optimal structure Ω_S with rotational symmetry similar to the most precise polyedral approximation Ω_{16}, as shown by fig. 3.14. If we assume that Ω_S is representative of the exact solution of problem (3.13), the comparison of the illustrations of the optimal shapes validates the convergence result of theorem 3.12. Indeed, the similarity between Ω_N and Ω_S increases when $\mathcal{G}_N$ approximates better the original set $\mathcal{G}$. This statement is corroborated by the numerical computation of the Hausdorff distances between Ω_S and the shapes resulting from the polyhedral approximation, as shown in table 3.4.

By looking at the graph of fig. 3.15b we remark that the constraint on the maximum of the L^6-norm of the von Mises stress is overall satisfied by the method of subdifferential, but more significant perturbations can be observed. A more difficult convergence compared to the polyhedron method can be remarked in fig. 3.16a, where a slower decrease in the objective function is evident, to the point that a larger number of iterations has been necessary in order to reach a stable configuration (200 for the polyhedron method and 300 for the subdifferential). Both issues can be justified by the rotational symmetry of the optimization problem. As the graph in 3.16b shows, the critical direction α varies widely at each step of the optimization algorithm, even from one iteration to the next.

By comparing the duration of the simulations as presented in table 3.4, we remark that the subdifferential approach is overall faster than the polyhedral approximation, since it requires fewer evaluations of the constraint functional. Therefore, the shorter duration of each step compensates the smaller contribution of each iteration to the decrease of the objective function.

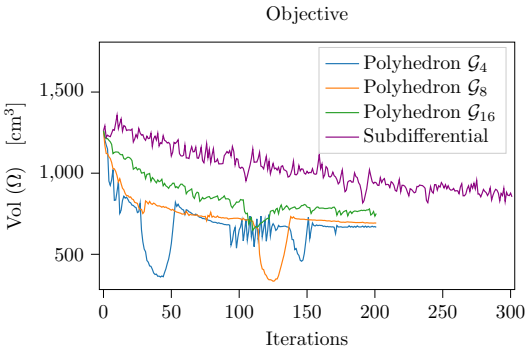


(a) Evolution of the constraint for three instances of the polyhedron method (4, 8, and 16 vertices respectively).

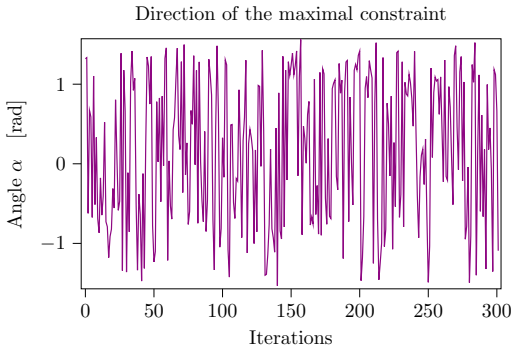


(b) Evolution of the constraint for three instances of the Polyhedron method, and the Subdifferential method.

Figure 3.15: Convergence of the constraint (L^6-norm of the von Mises stress) for the disc structure.



(a) Evolution of the objective function.



(b) Evolution of the direction of the maximal constraint (in terms of the angle α) during the optimization process.

Figure 3.16: Convergence of the objective (volume) for the Polyhedron and Subdifferential methods, and direction of the largest constraint under the Subdifferential method for the disc structure.

		Polyhedron			Subdifferential
		$N = 4$	$N = 8$	$N = 16$	
Execution of the optimization					
Duration of the optimization	[min]	189	296	499	159
Number of iterations		200	200	200	300
Results					
Final volume Vol(Ω)	[cm^3]	666.27	692.18	751.46	874.07
Maximal constraint	[kPa]	5.015	5.053	5.164	5.183
$d_{\mathrm{H}}(\Omega_S, \Omega_N)$	[cm]	3.001	3.037	1.347	–

Table 3.4: Numerical results for the optimization of the volume of a disc structure under constraints on the L^6-norm of the von Mises stress, obtained using the Polyhedron method (with an increasing number of vertices), and the Subdifferential method.

3.5 Conclusions and perspectives

In this chapter we have compared two different methods to address shape optimization problems under constraints on the worst-case scenario of a given functional. The first method can be applied only to convex functions of the displacement, and relies on the approximation of the set $\mathcal{G}$ of admissible loads by polyhedra. This method corresponds to the design of a structure which complies with the constraint in a finite number of representative load cases. The second method is based on the computation of an element of the subdifferential of the constraint by identifying the critical element of $\mathcal{G}$ maximizing the constraint, and differentiating in the relative direction.

The numerical simulations of section 3.4 support the efficacy of both methods in the case of an admissible set of loads parametrized by an ellipse in $\mathbb{R}^2$. We observed that in both cases the method based on the subdifferential is faster than the polyhedral approximation since it requires fewer evaluations of the constraint function. However, we remark that, if the admissible load maximizing the constraint is not unique, the convergence of method based on the subdifferential is degraded, and smaller and more numerous optimization steps are required to converge.

One way to improve the method of the subdifferential consists in taking into account multiple elements of the subdifferential of the constraint functional. Such variant of the algorithm would require to identify if multiple mechanical loads maximize the constraint for a given shape. A possible direction of development could be the adaptation of the proximal algorithm to shape optimization, since it already relies on the subdifferential in the sense of Clarke. Refer to [27, 190] and the references therein for further information on the proximal algorithm in non-smooth optimization.

Chapter 4

Shape optimization of a polynomial functional

A large portion of this chapter has been submitted to publication as a journal paper under the title *Shape Optimization of Polynomial Functionals under Uncertainties on the Right-Hand Side of the State Equation*, co-written by Fabien Caubet, Marc Dambrine, and Jérôme Maynadier.

Contents

4.1 Introduction

The design of mechanical structures satisfying several constraints of different natures is a difficult problem for engineers, and shape optimization techniques offer an automated approach to devise original designs which satisfies the given constraints. However, as discussed in chapter 3, it is unrealistic to consider that all information on the problem is perfectly known for industrial applications. In the previous chapter, from section 3.1.2 forward, we focused on optimization problem with constraints on the worst-case scenario for an uncertain functional. Here, we

consider a different approach to topology optimization under uncertainties, where the boundary conditions are modeled by random variables, and the constraint is imposed on the expected value or the variance of a given functional of the state.

This approach, known in the literature as *Robust Topology Optimization* (RTO) [121] has been studied by numerous different authors, providing multiple perspectives on the problem. In [64] the mean and the variance of a generic objective functional are estimated by using a dimension reduction method followed by a Gauss-type quadrature sampling, while the shape sensitivities are computed by using the analytic derivatives of the random moments. In [11] the authors address the issue of small uncertainties on the material properties, on the external loads, and on the geometry of the structure by linearizing the perturbation around their mean value. In [26, 245, 79, 212] the authors consider RTO problems for structures with uncertainties on the applied loads and on the shape itself, where the structure is represented using density methods. In particular, in [245] the propagation of the uncertainties is studied using a polynomial chaos method, while in [211, 212] the authors represent the uncertainties by a Karhunen-Loève expansion, and model their propagation by a Sparse Grid Collocation method. The issue of the curse of dimensionality is addressed in [97], where the uncertainties are taken into account using a Stochastic Gradient Descent algorithm. In [76, 30] the structures are represented by a level-set technique, and the optimization under uncertainties is performed using techniques of stochastic programming. The authors of [106] study the minimization of the mean and the variance of the mechanical compliance of an elastic structure, considering an exact expression of the random moments and their sensitivities with respect to the shape. A similar approach is adopted in [80], where the authors provide a method to compute analytically the expected value of a generic quadratic functional in terms of the first and second moments of the random variables modeling the uncertainties.

The present chapter adapts and extends the approach of [80] to the case of polynomial functionals of the right-hand side of the state equation. We consider the shape optimization problem as an instance of a PDE-constrained optimization problem. We suppose the right-hand side of the partial differential equation to be subject to uncertainties, without any assumption on their amplitude. The uncertainties are modeled as random variables, by using suitable Bochner spaces. Let us consider a functional of the shape that can be expressed as a polynomial function of degree m of the solution of the state equation. Similarly to the procedure detailed in [80], we introduce a deterministic correlation tensor of order m, which depends only on the first m random moments. Consequently, it is possible to compute exactly the expected value of the functional of interest, as well as its shape derivative.

The main contribution of this chapter is theorem 4.17, that provides the analytic expression of its shape derivative in terms of the first m moments of the random variables modeling the uncertainties, without any further assumption on their distributions. In the case of a finite dimensional valued uncertainty, tensor representation of the uncertainties is not needed and we present in proposition 4.18 the corresponding result. Notably, no sampling method requiring a large number of simulations is used in the method presented here.

An application of the proposed procedure is related to the utilization of the L^m-norm of a function as a smooth approximation of its L^∞-norm (i.e. its supremum) in a given domain. Indeed, by considering the L^m-norm of the stress in the domain as functional of interest, we are able to derive shapes where, on average, stress is less concentrated than in the ones obtained by controlling the expectation of the mechanical compliance.

This chapter is organized as follows. In section 4.2 we recall some definitions and results about Bochner spaces and tensor products among Hilbert spaces. Section 4.3 states the main results of the present chapter: it introduces the correlation operator for multilinear functionals

and its applications in the context of shape optimization, with a particular focus on the context of linear elasticity. Section 4.4 and section 4.5 provide two examples of numerical applications. In section 4.4 a tridimensional structure is subject to an uncertain load, and its mass is minimized under a constraint on the L^6-norm of the von Mises stress in the domain. Such example adresses the common concern in structural mechanics in avoiding the concentration of stress in a small region of the structure. Section 4.5 treats a bidimensional example, and shows how taking into account the variance of the compliance in a shape optimization problem can be crucial when the random variable modeling the mechanical loads are heavily correlated. Conclusions are drawn in section 4.6.

4.2 Mathematical setting

4.2.1 Tensor product in Hilbert spaces

In [225, 80] tensor products between Hilbert spaces have been used to work with the stochastic moment of order 2 of random quantities. In [226], the same technique has been extended to treat the stochastic moments of order $m > 2$. Here, we recall the main definitions and results about the tensor product in Hilbert spaces as presented in [206, 217, 225, 226, 80].

Definition 4.1 (Tensor product in vector spaces)**.** *For a positive integer $m \geq 2$, let us consider the vector spaces $\mathcal{X}_1, \ldots, \mathcal{X}_m$. We denote $\hat{\mathfrak{P}}_m(\mathcal{X}_1, \ldots, \mathcal{X}_m)$ the space of all m-multilinear forms on $\prod_{i=1}^m \mathcal{X}_i$. For $(\mathcal{X}_1, \ldots, \mathcal{X}_m) \in \prod_{i=1}^m \mathcal{X}_i$, the tensor product $x_1 \otimes \ldots \otimes x_m$, also written as $\bigotimes_{i=1}^m x_i$, is a real valued linear application defined on $\hat{\mathfrak{P}}_m(\mathcal{X}_1, \ldots, \mathcal{X}_m)$ such that, for all $P_m \in \hat{\mathfrak{P}}_m(\mathcal{X}_1, \ldots, \mathcal{X}_m)$,*

$$\left(\bigotimes_{i=1}^m x_i\right)(P_m) = P_m(\mathcal{X}_1, \ldots, \mathcal{X}_m).$$

If all spaces $\mathcal{X}_1, \ldots, \mathcal{X}_m$ are Hilbert spaces, it is possible to define a product space with a Hilbertian structure.

Definition 4.2 (Tensor product between Hilbert spaces)**.** *Let $\left(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i}\right)$ be m Hilbert spaces, with $m \geq 2$. We define the set $\mathcal{V}$ as*

$$\mathcal{V} = \operatorname{span}\left\{\bigotimes_{i=1}^m x_i \quad \text{such that} \quad x_i \in \mathcal{H}_i \quad \forall i = 1 \ldots m\right\}.$$

Let $\langle\cdot,\cdot\rangle_\otimes : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ be a bilinear operation such that

$$\left\langle \bigotimes_{i=1}^m x_i, \bigotimes_{i=1}^m y_i \right\rangle_\otimes = \prod_{i=1}^m \langle x_i, y_i\rangle_{\mathcal{H}_i}. \tag{4.1}$$

for any choice of $x_i, y_i \in \mathcal{H}_i$. We denote $\|\cdot\|_\otimes$ the norm induced by the tensor product $\langle\cdot,\cdot\rangle_\otimes$. The operation introduced in (4.1) is an inner product in $\mathcal{V}$ [206, Section II.4]. The tensor product of the Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_m$ is the completion of $\mathcal{V}$ with respect to the inner product $\langle\cdot,\cdot\rangle_\otimes$, and is denoted $\bigotimes_{i=1}^m \mathcal{H}_i$. If all the m Hilbert spaces coincide with a single Hilbert space $\mathcal{H}$, we denote their tensor product as $\mathcal{H}^{\otimes m}$.

Definition 4.3 (Operator norm)**.** *For any real valued linear operator P defined on a normed vector space $\mathcal{X}$, we denote its operator norm as*

$$\|P\|_{\mathrm{OP}} = \sup_{\|x\|_{\mathcal{X}}=1} |P(x)|.$$

For any m-multilinear functional $P_m : \mathcal{X}_1, \ldots, \mathcal{X}_m \to \mathbb{R}$ its operator norm is defined as

$$\|P_m\|_{\mathrm{OP}} = \sup_{\|x_i\|_{\mathcal{X}_i}=1 \,\forall i} |P_m(\mathcal{X}_1, \ldots, \mathcal{X}_m)|.$$

As stated in [217, section 1.2], a primary purpose of the tensor product is the multilinear mappings into linear ones.

Proposition 4.4 (Linearization of bounded multilinear functionals)**.** *Let us consider a real-valued, bounded, multilinear functional $P : \prod_{i=1}^m \mathcal{H}_i \to \mathbb{R}$ defined on the separable Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_m$. Then, there exists a unique linear functional $\widehat{P}_m : \widehat{\bigotimes}_{i=1}^m \mathcal{H}_i \to \mathbb{R}$ such that $\widehat{P}_m$ is continuous, and for all $(x_1, \ldots, x_m) \in \prod_{i=1}^m \mathcal{H}_i$, $\widehat{P}_m \left(\bigotimes_{i=1}^m x_i\right) = P(x_1, \ldots, x_m)$.*

The existence of the functional $\widehat{P}_m$ for any m-multilinear continuous mapping P is often referred as the *universal property of the tensor product* [57, Chapter 9] and is proven in [148, Theorem 2.6.4].

4.2.2 Modeling of the uncertainties

In order to model the uncertainties, we use the formalism of Bochner spaces, which extends the theory of integration to Banach-valued functions [141, Chapter 1]. Let $(\mathcal{O}, \mathcal{A}, \mu)$ be a measure space, characterized by the σ-algebra $\mathcal{A}$ and the measure μ. We recall the definition of measurable and integrable functions in the context of Bochner spaces.

Definition 4.5 (μ-simple and strongly μ-measurable functions)**.** *A function $g : \mathcal{O} \to \mathcal{X}$ is said to be μ-simple if it can be written in the form*

$$\sum_{i=1}^{N} \chi_{A_i} \mathbf{x}_i,$$

where N is a finite positive integer, $\mathbf{x}_i \in \mathcal{X}$, $A_i \in \mathcal{A}$, and $\mu(A_i) < \infty$ for all $i \in \{1, \ldots, N\}$, and χ_A is the characteristic function for the set A. A function $f : \mathcal{O} \to \mathcal{X}$ is said to be strongly μ-measurable if there exists a sequence $\{g_i\}_{i=1}^{\infty}$ of μ-simple functions converging to f μ-almost everywhere.

Definition 4.6 (Bochner integral)**.** *The Bochner integral of a simple function $g : \mathcal{O} \to \mathcal{X}$ such that $g = \sum_{i=1}^N \chi_{A_i} \mathbf{x}_i$ with respect to the measure μ is defined by*

$$\int_{\mathcal{O}} g \, \mathrm{d}\mu = \sum_{i=1}^{N} \mu(A_i) \, \mathbf{x}_i \quad \in \mathcal{X}.$$

A strongly μ-measurable function f is Bochner integrable with respect to the measure μ if there exists a sequence $\{g_i\}_{i=1}^{\infty}$ of μ-simple functions $g_i : \mathcal{O} \to \mathcal{X}$ such that

$$\lim_{i \to \infty} \int_{\mathcal{O}} \|f - g_i\|_{\mathcal{X}} \, \mathrm{d}\mu = 0,$$

where the (real) integral is intended at the sense of Lebesgue. The Bochner integral of such a Bochner integrable function function is defined as

$$\int_{\mathcal{O}} f \,\mathrm{d}\mu = \lim_{i\to\infty} \int_{\mathcal{O}} g_i \,\mathrm{d}\mu \quad \in \mathcal{X}.$$

Moreover, the value of $\int_{\mathcal{O}} f \,\mathrm{d}\mu$ *is independent from the choice of the sequence* $\{g_i\}_{i=1}^{\infty}$.

Once defined the integration for Banach-valued functions, we can introduce the Bochner spaces as the equivalent of the usual L^p spaces for real-valued functions.

Definition 4.7 (Bochner spaces and equivalence)**.** *A Bochner integrable function* $f : \mathcal{O} \to \mathcal{X}$ *belongs to the space* $\mathcal{L}^p(\mathcal{O}, \mu; \mathcal{X})$ *for* $i \leq p < \infty$ *if and only if* $\int_{\mathcal{O}} \|f\|_{\mathcal{X}}^p \,\mathrm{d}\mu < \infty$. *A Bochner integrable function* $f : \mathcal{O} \to \mathcal{X}$ *belongs to the space* $\mathcal{L}^\infty(\mathcal{O}, \mu; \mathcal{X})$ *if and only if there exist a real positive number* $r < \infty$ *such that* $\mu(\{\Omega \in \mathcal{O} : \|f\|_{\mathcal{X}} \geq r\}) = 0$. *Two strongly* μ*-measurable function* f *and* g *are said to be equivalent if the subset of* $\mathcal{O}$ *where* f *is different from* g *has measure* 0. *The equivalence relation is denoted as* $f \sim g$. *The Bochner space* $\mathrm{L}^p(\mathcal{O}, \mu; \mathcal{X})$ *for* $1 \leq p \leq \infty$ *is defined as the quotient of* $\mathcal{L}^p(\mathcal{O}, \mu; \mathcal{X})$ *with respect to the equivalence relation "*$\sim$*". Bochner spaces are also Banach spaces with respect to the following norms:*

$$\|f\|_p = \left(\int_{\mathcal{O}} \|f\|_{\mathcal{X}}^p \,\mathrm{d}\mathbb{P}\right)^{1/p} \qquad \text{for} \quad 1 \leq p < \infty;$$
$$\|f\|_\infty = \inf\{r \geq 0 : \mu(\{\Omega \in \mathcal{O} : \|f\|_{\mathcal{X}} \geq r\}) = 0\}.$$

Having stated the main definition about generic Bochner spaces, let us focus on the case where a probability measure $\mathbb{P}$ replaces the generic measure μ. At first, we can remark the following embedding of Bochner spaces.

Proposition 4.8 (Embeddings in Bochner spaces)**.** *Let* ℓ *and* m *be real numbers such that* $1 \leq \ell < m < \infty$. *Then, the following inclusion is true:*

$$\mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{X}) \subset \mathrm{L}^\ell(\mathcal{O}, \mathbb{P}; \mathcal{X}).$$

In particular, if $f \in \mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{X})$, *then* f *belongs also to* $\mathrm{L}^1(\mathcal{O}, \mathbb{P}; \mathcal{X})$.

Proof. The proof relies simply on Hölder's inequality [218, Equation (1.9)]. Let us denote $p = \frac{m}{l}$ and q its conjugate such that $\frac{1}{p} + \frac{1}{q} = 1$. Then, we have

$$\int_{\mathcal{O}} \|f\|_{\mathcal{X}}^l \,\mathrm{d}\mathbb{P} = \int_{\mathcal{O}} \|f\|_{\mathcal{X}}^l \, 1 \,\mathrm{d}\mathbb{P} = \left(\int_{\mathcal{O}} \|f\|_{\mathcal{X}}^{l\frac{m}{l}} \,\mathrm{d}\mathbb{P}\right)^{1/p} \left(\int_{\mathcal{O}} 1\right)^{1/q} = \|f\|_{\mathrm{L}^m(\mathcal{O},\mathbb{P};\mathcal{X})}^l < \infty.$$

□

We recall the definition of the expectation operator in Bochner spaces and a classical result about the commutation of the expectation and a closed linear operator.

Definition 4.9 (Expectation)**.** *The expectation operator* $\mathbb{E}[\,\cdot\,] : \mathrm{L}^1(\mathcal{O}, \mathbb{P}; \mathcal{X}) \to \mathcal{X}$ *is the bounded linear operator such that, for all* $f \in \mathrm{L}^1(\mathcal{O}, \mathbb{P}; \mathcal{X})$,

$$\mathbb{E}[f] = \int_{\mathcal{O}} f \,\mathrm{d}\mathbb{P} \quad \in \mathcal{X}.$$

The following result, proven by [141, Proposition 1.2.3 and Equation (1.2)] allows to swap the expectation and a continuous linear operator.

Proposition 4.10. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, $f : \mathcal{O} \to \mathcal{X}$ be a Bochner-integrable function, and $T : \mathcal{X} \to \mathcal{Y}$ be a continuous linear operator. Then, $Tf : \mathcal{O} \to \mathcal{Y}$ is a Bochner-integrable function, and*

$$\mathbb{E}[T\,f] = T\,\mathbb{E}[f].$$

A more general version of proposition 4.10 is known in literature as Hille's theorem [141, Theorem 1.2.4], which does not require T to be a continuous operator, but only a closed one on a subspace of $\mathcal{X}$.

Theorem 4.11 (Hille)**.** *Let $f : \mathcal{O} \to \mathcal{X}$ be a Bochner-integrable function valued in the Banach space $\mathcal{X}$, and let T be a closed linear operator whose domain $D(T)$ is a subspace of $\mathcal{X}$ and has values in another Banach space Y. We suppose that f takes its values in $D(T)$ almost everywhere and the almost everywhere defined function $Tf : \mathcal{O} \to Y$ is Bochner-integrable. Then, f is Bochner-integrable as a $D(T)$-valued function (i.e. the equivalence class of f belongs to $\mathrm{L}^1(\mathcal{O}, \mathbb{P}; D(T))$), $\mathbb{E}[f] \in D(T)$, and $\mathbb{E}[T\,f] = T\,\mathbb{E}[f]$.*

Hille's theorem is reported in [141, Theorem 1.2.4]. We can remark that Hille's theorem is valid if T is a continuous operator, since all continuous operator is closed (see [141, page 15, definition of closed operator]). However, [141, Proposition 1.2.3 and Equation (1.2)] points out that, for continuous operators, it is not necessary to prove Hille's theorem to get the same properties, since they descend directly from the definition of Bochner integral.

4.3 Main results

4.3.1 Correlation operator and multilinear functionals

We begin our study by introducing the correlation operator for multilinear functionals defined on Hilbert spaces under uncertainties. The random component of the problem is treated by considering functions defined on suitable Bochner spaces. The correlation operator has been studied in the context of shape optimization under uncertainties in [80], limitedly to bilinear functionals defined on Hilbert spaces, and in [225, 226] for bilinear and multilinear functionals. Let us consider the measure space $(\mathcal{O}, \mathcal{A}, \mathbb{P})$, where $\mathcal{O}$ is the event space, $\mathcal{A} \subset 2^{\mathcal{O}}$ is a σ-algebra on $\mathcal{O}$, and $\mathbb{P}$ is a probability measure, and let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space. First of all, we can state a result about the Bochner-integrability of the tensor product.

Proposition 4.12. *Let $\mathcal{H}_1, \ldots, \mathcal{H}_m$ be Hilbert spaces, each endowed with the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}_i}$ for $i = 1, \ldots, m$. Let us consider $x_1, \ldots, x_m$, each belonging to the Bochner space $\mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{H}_i)$. Finally, we define the mapping $\omega \mapsto \bigotimes_{i=1}^m x_i(\omega)$ from the event space $\mathcal{O}$ to the Hilbert space $\widehat{\bigotimes}_{i=1}^m \mathcal{H}_i$. Then, such a function belongs to the Bochner space $\mathrm{L}^1\left(\mathcal{O}, \mathbb{P}; \widehat{\bigotimes}_{i=1}^m \mathcal{H}_i\right)$.*

Proof. We consider the Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_m$ as well as their tensor product $\widehat{\bigotimes}_{i=1}^m \mathcal{H}_i$ as Banach spaces with respect to the norms induced by their respective inner products. In order to prove that $\bigotimes_{i=1}^m x_i \in \mathrm{L}^1\left(\mathcal{O}, \mathbb{P}; \widehat{\bigotimes}_{i=1}^m \mathcal{H}_i\right)$, we estimate its norm, and we use Hölder's inequality

extended to multiple terms

$$\int_{\mathcal{O}} \left\| \bigotimes_{i=1}^{m} x_i(\omega) \right\|_{\widehat{\bigotimes}_{i=1}^{m} \mathcal{H}_i} \mathrm{d}\mathbb{P}(\omega) = \int_{\mathcal{O}} \left(\prod_{i=1}^{m} \|x_i(\omega)\|_{\mathcal{H}_i} \right) \mathrm{d}\mathbb{P}(\omega)$$
$$\leq \prod_{i=1}^{m} \left(\int_{\mathcal{O}} \|x_i(\omega)\|_{\mathcal{H}_i}^{m} \, \mathrm{d}\mathbb{P}(\omega) \right)^{\frac{1}{m}} = \prod_{i=1}^{m} \|x_i\|_{\mathrm{L}^m(\mathcal{O},\mathbb{P};\mathcal{H}_i)} < \infty.$$

□

Next, the correlation operator is introduced. As it is remarked in [80], the literature is not consistent in the definition of the correlation between random variables. In the present thesis, we adopt the following definition.

Definition 4.13 (Correlation operator on Bochner spaces)**.** *Let* $\left(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i}\right)$ *be Hilbert spaces for* $i = 1, \ldots, m$*. Let us consider the linear operator defined on* $\prod_{i=1}^{m} \mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{H}_i)$*, mapping* $(x_1, \ldots, x_m)$ *to* $\bigotimes_{i=1}^{m} x_i(\omega)$*. Proposition 4.12 ensures that the function* $\omega \mapsto \bigotimes_{i=1}^{m} x_i(\omega)$ *is Bochner-integrable. The correlation between the* m *functions* $x_1, \ldots, x_m$ *is defined as*

$$\mathrm{Cor}(x_1, \ldots, x_m) = \mathbb{E}\left[\bigotimes_{i=1}^{m} x_i(\omega) \right] \in \widehat{\bigotimes_{i=1}^{m}} \mathcal{H}_i,$$

and the correlation operator $\mathrm{Cor} : \prod_{i=1}^{m} \mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{H}_i) \to \widehat{\bigotimes}_{i=1}^{m} \mathcal{H}_i$ *is a bounded linear operator associating* m *random vectors with their correlation. If all arguments of the correlation operator are the same, we denote* $\mathrm{Cor}_m(x) = \mathrm{Cor}(x, \ldots, x)$*.*

In [225, 226] the term $\mathrm{Cor}_m(x)$ is denoted as the stochastic moment of order m of x.

Finally, we state a proposition that allows the expression of the expected value of a multilinear expression in terms of a correlation tensor.

Proposition 4.14. *Let* $(\mathcal{O}, \mathcal{A}, \mathbb{P})$ *be a probability space,* $\mathcal{H}_1, \ldots, \mathcal{H}_m$ *Hilbert spaces provided with the norms* $\|\cdot\|_{\mathcal{H}_i}$ *for* $i = 1 \ldots m$*, and* $P : \prod_{i=1}^{m} \mathcal{H}_i \to \mathbb{R}$ *a bounded multilinear operator. Then, there exists a unique bounded, real-valued, linear operator* $\widehat{P}_m$ *defined on* $\widehat{\bigotimes}_{i=1}^{m} \mathcal{H}_i$ *such that the following three statements hold true for all* $(x_1, \ldots, x_m) \in \prod_{i=1}^{m} \mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathcal{H}_i)$*:*

1. $P(x_1, \ldots, x_m) \in \mathrm{L}^1(\mathcal{O}, \mathbb{P})$*,*
2. $P(x_1(\omega), \ldots, x_m(\omega)) = \widehat{P}_m\left(\bigotimes_{i=1}^{m} x_i(\omega)\right)$*, for almost all* $\omega \in \mathcal{O}$*,*
3. $\mathbb{E}[P(x_1, \ldots, x_m)] = \widehat{P}_m(\mathrm{Cor}(x_1, \ldots, x_m))$*.*

Proof. The first point comes directly from the continuity of the operator P and the application of Hölder's inequality. The second can be deduced from the universal property of the tensor product (see [57, Chapter 9], [148, Theorem 2.6.4], and proposition 4.4).

In order to prove the third statement, we show that the hypotheses of [141, Proposition 1.2.3] (reported in section 4.2 as proposition 4.10) are verified. The function $\omega \mapsto P(x_1(\omega), \ldots, x_m(\omega))$ is Bochner-integrable thanks to the first point of this proposition. The operator $\widehat{P}_m$ is continuous, as proven by the second point of this proposition. Therefore, we can apply [141, Proposition 1.2.3] and conclude

$$\mathbb{E}[P(x_1, \ldots, x_m)] = \mathbb{E}\left[\widehat{P}_m\left(\bigotimes_{i=1}^{m} x_i \right) \right] = \widehat{P}_m\left(\bigotimes_{i=1}^{m} x_i \right) = \widehat{P}_m(\mathrm{Cor}(x_1, \ldots, x_m)).$$

□

4.3.2 Uncertain loads in linear elasticity

From now on, we focus on Robust Topology Optimization problems in the context of linear elasticity.

$$\left|\begin{array}{l}\text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{minimizing the volume } \mathrm{Vol}(\Omega) \\ \text{under the constraint } \mathbb{E}\left[H(\mathbf{u}_\Omega(\cdot),\Omega)\right] \leq \tau, \\ \text{where the displacement } \mathbf{u}_\Omega(\omega) \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \\ \text{solves the elasticity equation} \\ \quad \left\{\begin{array}{rcll} -\mathrm{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{g}(\omega) & \text{on } \Gamma_\mathrm{N}, \\ \boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega(\omega) & = & \mathbf{0} & \text{on } \Gamma_\mathrm{D}, \end{array}\right. \\ \text{for almost all event } \omega \in \mathcal{O}. \end{array}\right. \tag{4.2}$$

We remark that, since $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O},\mathbb{P};\mathrm{L}^2(\Gamma_\mathrm{N})^d\right)$ is a random process, $\mathbf{u}_\Omega$ is a random process in the Bochner space $\mathrm{L}^m(\mathcal{O},\mathbb{P};\mathcal{X}_\Omega)$ thanks to the usual elliptic *a priori* estimates. If we consider the mechanical load applied on Γ_N to be a random variable $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O},\mathbb{P};\mathrm{L}^2(\Gamma_\mathrm{N})^d\right)$, we can conclude that $\mathbf{u}_\Omega(\omega) \in \mathrm{H}^1(\Omega)^d$ for almost all event ω, and $\mathbf{u}_\Omega \in \mathrm{L}^m\left(\mathcal{O},\mathbb{P};\mathrm{H}^1(\Omega)^d\right)$. Denoting $\mathbb{C}$ the fourth order elasticity tensor, and $\gamma : \mathrm{H}^1(\Omega)^d \to \mathrm{L}^2(\Gamma_\mathrm{N})^d$ the operator mapping $\mathbf{v} \mapsto \mathbf{v}|_{\Gamma_\mathrm{N}}$, the problem defining the state equation can be written in variational form

$$\left|\begin{array}{l}\text{Find } \mathbf{u}_\Omega(\omega) \in \mathcal{V} = \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \text{ such that for all } \mathbf{v} \in \mathcal{V}: \\ \qquad \langle \mathbb{C}\nabla\mathbf{u}_\Omega(\omega), \nabla\mathbf{v}\rangle_{\mathrm{L}^2(\Omega)^d} = \langle \gamma(\mathbf{v}), \mathbf{g}(\omega)\rangle_{\mathrm{L}^2(\Gamma_\mathrm{N})^d}, \\ \text{where} \\ \qquad \langle \mathbb{C}\nabla\mathbf{u}_\Omega(\omega), \nabla\mathbf{v}\rangle_{\mathrm{L}^2(\Omega)^d} = \displaystyle\int_\Omega \left(\mathbb{C}\nabla\mathbf{u}_\Omega(\omega) : \nabla\mathbf{v}\right) \mathrm{d}\mathbf{x}, \\ \qquad \langle \gamma(\mathbf{v}), \mathbf{g}(\omega)\rangle_{\mathrm{L}^2(\Gamma_\mathrm{N})^d} = \displaystyle\int_{\Gamma_\mathrm{N}} \mathbf{g}(\omega)\cdot\mathbf{v}\,\mathrm{d}\mathbf{s}. \end{array}\right. \tag{4.3}$$

For simplicity, we suppose that all admissible shapes in $\mathcal{S}_{\text{adm}}$ share the portions Γ_N and Γ_D, constraining the displacements fields $\boldsymbol{\theta} \in \Theta_{adm} \subset \mathrm{W}^{1,\infty}\left(\mathbb{R}^d,\mathbb{R}^d\right)$ to be equal to $\mathbf{0}$ on these surfaces. Moreover, we focus our study on functionals P^Ω with the following structure

$$P^\Omega(\mathbf{v}_1,\ldots,\mathbf{v}_m) = \int_\Omega q_0(\mathbf{v}_1(\mathbf{x}),\ldots,\mathbf{v}_m(\mathbf{x}))\,\mathrm{d}\mathbf{x} + \int_\Omega q_1(\nabla\mathbf{v}_1(\mathbf{x}),\ldots,\nabla\mathbf{v}_m(\mathbf{x}))\,\mathrm{d}\mathbf{x}, \tag{4.4}$$

where $q_0 : \prod_{i=1}^m \mathbb{R}^d \to \mathbb{R}$ and $q_1 : \prod_{i=1}^m \mathbb{R}^{d\times d} \to \mathbb{R}$ are multilinear and continuous.

Without any further assumption on the domain, problem (4.2) can be affected by regularity issues [165, 133, 50]. Indeed, the Lax-Milgram theorem ensures that, for almost any $\omega \in \mathcal{O}$, $\mathbf{u}_\Omega(\omega) \in \mathrm{H}^1(\Omega)^d$. However, for $P^\Omega(\mathbf{u}_\Omega(\omega),\ldots,\mathbf{u}_\Omega(\omega))$ to be well-defined we must require that $\mathbf{u}_\Omega(\omega) \in \mathrm{W}^{1,m}(\Omega)$. In order to verify such condition we consider that, for all admissible domain $\Omega \in \mathcal{S}_{\text{adm}}$, the portions of the boundary where Dirichlet and Neumann conditions are applied are fully separated, meaning that

$$\overline{\Gamma_\mathrm{D}} \cap \overline{(\Gamma_\mathrm{N} \cup \Gamma_0)} = \emptyset \tag{4.5}$$

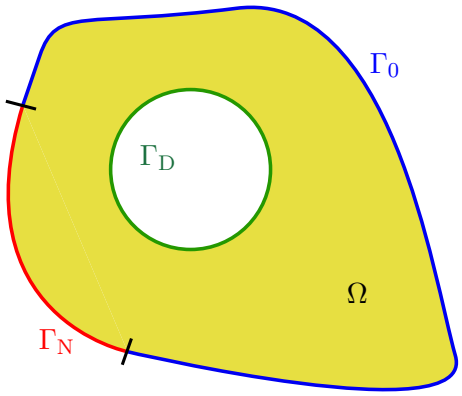


Figure 4.1: Example of a domain Ω where the condition (4.5) is satisfied, since the Neumann and Dirichlet portiond of the boundary are fully disjoint.

as showed in fig. 4.1.

We recall the following result on the regularity of the solution of boundary value problems.

Proposition 4.15. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain of class $\mathcal{C}^{k+2}$, with k integer. We suppose that its boundary can be divided in three parts Γ_D, Γ_N, and Γ_0 mutually disjoint, with strictly positive measure, satisfying the condition (4.5). Let us consider $\mathbf{g} \in \mathrm{H}^{k+\frac{1}{2}}(\Gamma_N)^d$. Then, the solution $\mathbf{u}_\Omega$ of problem (4.3) belongs to the Sobolev space $\mathrm{H}^{k+2}_{\Gamma_D}(\Omega)$, and there exists $C > 0$ such that the following estimate holds*

$$\|\mathbf{u}_\Omega\|_{\mathrm{H}^{k+2}(\Omega)} \leq C \|\mathbf{g}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N)}. \tag{4.6}$$

Proof. At first, we remark that $\widetilde{\mathbf{g}}$, extension of $\mathbf{g}$ to $\Gamma_N \cup \Gamma_0$ such that $\widetilde{\mathbf{g}}|_{\Gamma_0} = 0$ belongs to $\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)^d$. Moreover, $\|\mathbf{g}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N)} = \|\widetilde{\mathbf{g}}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)}$.

Under the hypothesis (4.5), the elliptic regularity estimates apply on the entire domain Ω. In particular, there exists some constant $\widetilde{C} > 0$ such that (see [218, Theorem 8.29])

$$\|\mathbf{u}_\Omega\|_{\mathrm{H}^{k+2}(\Omega)} \leq \widetilde{C} \left(\|\mathbf{u}_\Omega\|_{\mathrm{L}^2(\Omega)^d} + \|\widetilde{\mathbf{g}}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)} \right). \tag{4.7}$$

By the coercivity of the bilinear form and the Poincaré and trace inequalities, there exists three positive constants α, C_1, C_2 such that

$$\begin{aligned} \|\mathbf{u}_\Omega\|^2_{\mathrm{H}^1(\Omega)} &\leq \frac{1}{\alpha} \langle \mathbb{C} \nabla \mathbf{u}_\Omega, \nabla \mathbf{u}_\Omega \rangle_{\mathrm{L}^2(\Omega)^d} = \langle \gamma(\mathbf{u}_\Omega), \mathbf{g} \rangle_{\mathrm{L}^2(\Gamma_N)^d} \\ &\leq C_1 \|\mathbf{u}_\Omega\|_{\mathrm{L}^2(\Gamma_N)^d} \|\widetilde{\mathbf{g}}\|_{\mathrm{L}^2(\Gamma_N \cup \Gamma_0)} \leq C_2 \|\mathbf{u}_\Omega\|_{\mathrm{H}^1(\Omega)} \|\widetilde{\mathbf{g}}\|_{\mathrm{L}^2(\Gamma_N \cup \Gamma_0)}. \end{aligned}$$

Injecting this result into (4.7) we conclude that

$$\begin{aligned} \|\mathbf{u}_\Omega\|_{\mathrm{H}^{k+2}(\Omega)} &\leq \widetilde{C} \left(\|\mathbf{u}_\Omega\|_{\mathrm{H}^1(\Omega)} + \|\widetilde{\mathbf{g}}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)} \right) \\ &\leq \widetilde{C} \left(C_2 \|\widetilde{\mathbf{g}}\|_{\mathrm{L}^2(\Gamma_N \cup \Gamma_0)} + \|\widetilde{\mathbf{g}}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)} \right) \leq C \|\widetilde{\mathbf{g}}\|_{\mathrm{H}^{k+\frac{1}{2}}(\Gamma_N \cup \Gamma_0)} \end{aligned}$$

for some positive constant C. □

The following result details the computation of the shape derivative of $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$ for a deterministic problem. We remark that, even if we only need $\mathbf{u}_\Omega$ to be in $\mathrm{W}^{1,m}_{\Gamma_\mathrm{D}}(\Omega)^d$ to define $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$, we require the higher regularity in $\mathrm{W}^{1,2m-2}_{\Gamma_\mathrm{D}}(\Omega)^d$ to compute its derivative.

Proposition 4.16. *Let k be a positive integer such that $k+1 > d/2$. Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a bounded domain with a $\mathcal{C}^{k+2}$ boundary, $\mathbf{g} \in \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})$, and $\mathbf{u}_\Omega$ solution of problem (4.3). We suppose that $P^{\Omega} : \prod_{i=1}^m \mathrm{W}^{1,m}_{\Gamma_\mathrm{D}}(\Omega)^d \to \mathbb{R}$ has the structure presented in (4.4). Then, $\mathbf{u}_\Omega \in \mathrm{W}^{1,2m-2}_{\Gamma_\mathrm{D}}(\Omega)$, and the quantity $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$ is differentiable with respect to Ω. The adjoint state $\mathbf{w}_\Omega$ solving the following boundary-value problem is well-defined in $\mathrm{H}^1(\Omega)$*

$$\begin{cases} -\operatorname{div} \mathbb{C}\nabla \mathbf{w}_\Omega &= \displaystyle\sum_{i=1}^{m} \partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega) - \operatorname{div} \partial_i q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega) & \text{in } \Omega, \\ \Big(\mathbb{C}\nabla \mathbf{w}_\Omega\Big)\mathbf{n} &= \displaystyle\sum_{i=1}^{m} \partial_i q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega)\mathbf{n} & \text{on } \Gamma_\mathrm{N} \cup \Gamma_0, \\ \mathbf{w}_\Omega &= \mathbf{0} & \text{on } \Gamma_\mathrm{D}, \end{cases} \tag{4.8}$$

where $\partial_i q_0$ and $\partial_i q_1$ denote the derivative of q_0 and q_1 with respect to the i-th component of their respective arguments. Finally, the shape derivative of $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$ can be expressed as follows

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\Omega} P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)(\boldsymbol{\theta}) &= \int_{\Gamma_0} \Big(q_0(\mathbf{u}_\Omega(\mathbf{s}), \dots, \mathbf{u}_\Omega(\mathbf{s}))\Big)(\boldsymbol{\theta} \cdot \mathbf{n})\, \mathrm{d}\mathbf{s} \\ + \int_{\Gamma_0} \Big(q_1(\nabla \mathbf{u}_\Omega(\mathbf{s}), \dots, \nabla \mathbf{u}_\Omega(\mathbf{s}))\Big)(\boldsymbol{\theta} \cdot \mathbf{n})\, \mathrm{d}\mathbf{s} &- \int_{\Gamma_0} \Big(\mathbb{C}\nabla \mathbf{u}_\Omega(\mathbf{s}) : \nabla \mathbf{w}_\Omega(\mathbf{s})\Big)(\boldsymbol{\theta} \cdot \mathbf{n})\, \mathrm{d}\mathbf{s}. \end{aligned} \tag{4.9}$$

Proof. At first, we prove the regularity of $\mathbf{u}_\Omega$ to ensure that $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$ is well-defined. The displacement $\mathbf{u}_\Omega$ solves the elliptic boundary-value problem (4.3). Given the regularity $\mathcal{C}^{k+2}$ of the domain and the fact that $\mathbf{g} \in \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})$, we can apply proposition 4.15 and prove that $\mathbf{u}_\Omega \in \mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)$. The domain Ω is bounded and of class $\mathcal{C}^{k+2}$, thus it complies with the cone condition defined as in [2, Definition 4.6]. Thanks to the Sobolev embedding theorem [2, Theorem 4.12, part I], the space $\mathrm{H}^{k+2}(\Omega)$ is compactly embedded into $\mathrm{W}^{1,m}(\Omega)$ for any $\widetilde{p} > 2$. Thus, we can conclude that $\mathbf{u}_\Omega \in \mathrm{W}^{1,2m-2}_{\Gamma_\mathrm{D}}(\Omega) \subset \mathrm{W}^{1,m}_{\Gamma_\mathrm{D}}(\Omega)$.

The shape derivative and the adjoint problem can be computed by the fast derivation method developed by Céa as presented in section 1.2.3. We introduce a Lagrangian function $\mathcal{L} : \mathcal{S}_{\mathrm{adm}} \times \mathrm{H}^{k+2}\left(\mathbb{R}^d\right)^d \times \mathrm{H}^1\left(\mathbb{R}^d\right)^d \to \mathbb{R}$ such that

$$\begin{aligned} \mathcal{L}(\Omega; \hat{\mathbf{u}}; \hat{\mathbf{w}}) &= \int_\Omega q_0(\hat{\mathbf{u}}, \dots, \hat{\mathbf{u}})\, \mathrm{d}\mathbf{x} + \int_\Omega q_1(\nabla \hat{\mathbf{u}}, \dots, \nabla \hat{\mathbf{u}})\, \mathrm{d}\mathbf{x} - \int_\Omega \mathbb{C}\nabla \hat{\mathbf{u}} : \nabla \hat{\mathbf{w}}\, \mathrm{d}\mathbf{x} \\ &+ \int_{\Gamma_\mathrm{N}} \mathbf{g} \cdot \gamma(\hat{\mathbf{w}})\, \mathrm{d}\mathbf{s} + \int_{\Gamma_\mathrm{D}} \left(\gamma(\hat{\mathbf{w}}) \cdot (\mathbb{C}\nabla \hat{\mathbf{u}})\mathbf{n} + \gamma(\hat{\mathbf{u}}) \cdot (\mathbb{C}\nabla \hat{\mathbf{w}} - \sum_{j=1}^{N} \partial_i q_1(\nabla \hat{\mathbf{u}}, \dots, \nabla \hat{\mathbf{u}}))\mathbf{n} \right) \mathrm{d}\mathbf{s}. \end{aligned}$$

All arguments of the Lagrangian are independent, since $\hat{\mathbf{u}}$ and $\hat{\mathbf{w}}$ are defined on the whole space $\mathbb{R}^d$, and not only on Ω. The term defined as an integral on the portion Γ_D of the boundary enforces the Dirichlet boundary condition, similarly to the proof of [19, Theorem 7]. The partial derivative $\frac{\partial \mathcal{L}}{\partial \tilde{w}}(\Omega, \hat{\mathbf{u}}, \hat{\mathbf{w}})$ vanishes when evaluated in $\hat{\mathbf{u}} = \mathbf{u}_\Omega$ thanks to the definition of the state

equation (4.3). Thus, for any $\hat{\mathbf{w}} \in \mathrm{H}^1\left(\mathbb{R}^d\right)^d$ we have that

$$\mathcal{L}\left(\Omega; \mathbf{u}_\Omega; \hat{\mathbf{w}}\right) = \int_\Omega q_0(\mathbf{u}_\Omega(\mathbf{x}), \dots, \mathbf{u}_\Omega(\mathbf{x}))\, \mathrm{d}\mathbf{x} + \int_\Omega q_1(\nabla\mathbf{u}_\Omega(\mathbf{x}), \dots, \nabla\mathbf{u}_\Omega(\mathbf{x}))\, \mathrm{d}\mathbf{x} = P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega). \tag{4.10}$$

Let us focus on the adjoint state $\mathbf{w}_\Omega$ solution of equation (4.8). The weak form of equation (4.8) can be written as follows

$$\left|\begin{array}{l} \text{Find } \mathbf{w}_\Omega \in \mathcal{V} = \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \text{ such that for all } \mathbf{v} \in \mathcal{V}: \\ \langle \mathbb{C}\nabla\mathbf{w}_\Omega, \nabla\mathbf{v}\rangle_{\mathrm{L}^2(\Omega)^d} = \displaystyle\sum_{i=1}^m \int_\Omega \partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\mathbf{v}\, \mathrm{d}\mathbf{x} + \sum_{i=1}^m \int_\Omega \partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\nabla\mathbf{v}\, \mathrm{d}\mathbf{x}. \end{array}\right. \tag{4.11}$$

The well-posedness of problem (4.11) is proven by the Lax-Milgram theorem. The bilinear form is the classical elasticity operator, which is continuous and coercive. The left-hand side is continuous by the definition of the operator P^Ω and by the regularity of $\mathbf{u}_\Omega$. Indeed, since q_0 and q_1 are m-multilinear continuous operators (see (4.4)), there exist two positive constants C_0 and C_1 such that

$$\begin{aligned} |q_0(\mathbf{y}_1, \dots, \mathbf{y}_m)| &\leq C_0 \prod_{j=1}^m \|\mathbf{y}_j\|_{\mathbb{R}^d} && \text{for any} \quad \mathbf{y}_1, \dots, \mathbf{y}_m \in \mathbb{R}^d, \\ |q_1(\mathbf{Y}_1, \dots, \mathbf{Y}_m)| &\leq C_1 \prod_{j=1}^m \|\mathbf{Y}_j\|_{\mathbb{R}^{d\times d}} && \text{for any} \quad \mathbf{Y}_1, \dots, \mathbf{Y}_m \in \mathbb{R}^{d\times d}. \end{aligned}$$

Thus

$$\begin{aligned} &\left| \sum_{i=1}^m \int_\Omega \partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\mathbf{v}\, \mathrm{d}\mathbf{x} + \sum_{i=1}^m \int_\Omega \partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\nabla\mathbf{v}\, \mathrm{d}\mathbf{x} \right| \\ &\quad \leq \sum_{i=1}^m \int_\Omega |\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\mathbf{v}|\, \mathrm{d}\mathbf{x} + \sum_{i=1}^m \int_\Omega |\partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\nabla\mathbf{v}|\, \mathrm{d}\mathbf{x} \\ &\quad \leq m\left(C_0 \int_\Omega \|\mathbf{u}_\Omega(\mathbf{x})\|^{m-1}_{\mathbb{R}^d} \|\mathbf{v}(\mathbf{x})\|_{\mathbb{R}^d}\, \mathrm{d}\mathbf{x} + C_1 \int_\Omega \|\nabla\mathbf{u}_\Omega(\mathbf{x})\|^{m-1}_{\mathbb{R}^{d\times d}} \|\nabla\mathbf{v}(\mathbf{x})\|_{\mathbb{R}^{d\times d}}\, \mathrm{d}\mathbf{x} \right) \\ &\quad \leq m\, C_0 \|\mathbf{v}\|_{\mathrm{L}^2(\Omega)^d} \|\mathbf{u}_\Omega\|^{m-1}_{\mathrm{L}^{2m-2}(\Omega)^{m-1}} + m\, C_1 \|\nabla\mathbf{v}\|_{\mathrm{L}^2(\Omega)^d} \|\nabla\mathbf{u}_\Omega\|^{m-1}_{\mathrm{L}^{2m-2}(\Omega)^{m-1}} \\ &\quad \leq \left(m(C_0 + C_1) \|\mathbf{u}_\Omega\|^{m-1}_{\mathrm{W}^{1,2m-2}(\Omega)} \right) \|\mathbf{v}\|_{\mathrm{H}^1(\Omega)}. \end{aligned}$$

Having proven that the adjoint problem (4.8) is well-posed, we remark that the derivative $\frac{\partial\mathcal{L}}{\partial\hat{\mathbf{u}}}(\Omega, \mathbf{u}_\Omega, \hat{\mathbf{w}})$ vanishes when evaluated in $\hat{\mathbf{w}} = \mathbf{w}_\Omega$. Indeed

$$\begin{aligned} &\frac{\partial\mathcal{L}}{\partial\hat{\mathbf{u}}}(\Omega, \mathbf{u}_\Omega, \mathbf{w}_\Omega)(\mathbf{v}) = -\int_\Omega (\mathbb{C}\nabla\mathbf{w}_\Omega : \nabla\mathbf{v})\, \mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{D}} \gamma(\mathbf{v})(\mathbb{C}\nabla\mathbf{w}_\Omega\mathbf{n})\, \mathrm{d}\mathbf{s} \\ &- \sum_{i=1}^m \left(\int_{\Gamma_\mathrm{D}} \gamma(\mathbf{v}) \cdot \left(\partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\right)\mathbf{n}\, \mathrm{d}\mathbf{s} \right. \\ &\left. + \int_\Omega (\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)(\mathbf{v}) + \partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)(\nabla\mathbf{v}))\, \mathrm{d}\mathbf{x} \right) \\ &= \int_\Omega \mathbf{v} \cdot \left(\operatorname{div}\mathbb{C}\nabla\mathbf{w}_\Omega\right) \mathrm{d}\mathbf{x} + \sum_{i=1}^m \int_\Omega \mathbf{v} \cdot \left(\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega) - \operatorname{div}\partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\right) \mathrm{d}\mathbf{x} \\ &- \int_{\Gamma_\mathrm{N}\cup\Gamma_0} \gamma(\mathbf{v}) \cdot \left(\mathbb{C}\nabla\mathbf{w}_\Omega - \sum_{i=1}^m \partial_i q_1(\nabla\mathbf{u}_\Omega, \dots, \nabla\mathbf{u}_\Omega)\right)\mathbf{n}\, \mathrm{d}\mathbf{s} + \int_{\Gamma_\mathrm{D}} \gamma(\mathbf{v}) \cdot \left(\mathbb{C}\nabla\mathbf{w}_\Omega\mathbf{n}\right) \mathrm{d}\mathbf{s} = 0. \end{aligned}$$

We conclude computing the expression (4.9) of the shape derivative of $P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)$. Recalling (4.10) and the results of [138, Section 5] on shape differentiation we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\Omega} P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)(\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}\Omega} \mathcal{L}\left(\Omega; \mathbf{u}_\Omega; \mathbf{w}_\Omega\right)(\boldsymbol{\theta}) \\
&= \frac{\partial}{\partial \Omega} \mathcal{L}\left(\Omega; \mathbf{u}_\Omega; \mathbf{w}_\Omega\right)(\boldsymbol{\theta}) + \frac{\partial}{\partial \hat{\mathbf{u}}} \mathcal{L}\left(\Omega; \mathbf{u}_\Omega; \mathbf{w}_\Omega\right) \frac{\mathrm{d}}{\mathrm{d}\Omega} \mathbf{u}_\Omega(\boldsymbol{\theta}) + \frac{\partial}{\partial \hat{\mathbf{w}}} \mathcal{L}\left(\Omega; \mathbf{u}_\Omega; \mathbf{w}_\Omega\right) \frac{\mathrm{d}}{\mathrm{d}\Omega} \mathbf{w}_\Omega(\boldsymbol{\theta}) \\
&= \int_{\Gamma_0} \Big(q_0(\mathbf{u}_\Omega(\mathbf{s}), \dots, \mathbf{u}_\Omega(\mathbf{s})) + q_1(\nabla \mathbf{u}_\Omega(\mathbf{s}), \dots, \nabla \mathbf{u}_\Omega(\mathbf{s})) - \mathbb{C} \nabla \mathbf{u}_\Omega(\mathbf{s}) : \nabla \mathbf{w}_\Omega(\mathbf{s}) \Big) (\boldsymbol{\theta} \cdot \mathbf{n}) \, \mathrm{d}\mathbf{s}.
\end{aligned}$$

□

We have proven the well-posedness of the variational problems and the expression of the shape derivative in the deterministic case. Thus, if we consider the applied load $\mathbf{g}$ to be a random variable belonging to the Bochner space $\mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})\right)$, the results of proposition 4.16 apply for almost any event $\omega \in \mathcal{O}$. Thanks to the framework adopted in [226, Theorem 3.2], we introduce the following tensorialized bounded linear operators, defined on Hilbert spaces

$$\begin{array}{rcll}
\widehat{\mathbb{C}}_m & : & \left(\mathrm{L}^2(\Omega)^d\right)^{\otimes m} \to \left(\mathrm{L}^2(\Omega)^d\right)^{\otimes m} & \text{s.t.} \quad \widehat{\mathbb{C}}_m \bigotimes_{i=1}^m \mathbf{V}_i \mapsto \bigotimes_{i=1}^m (\mathbb{C}\mathbf{V}_i); \\
\widehat{\nabla}_m & : & \left(\mathrm{H}^1(\Omega)^d\right)^{\otimes m} \to \left(\mathrm{L}^2(\Omega)^d\right)^{\otimes m} & \text{s.t.} \quad \widehat{\nabla}_m \bigotimes_{i=1}^m \mathbf{v}_i \mapsto \bigotimes_{i=1}^m (\nabla \mathbf{v}_i); \\
\widehat{\gamma}_m & : & \left(\mathrm{H}^1(\Omega)^d\right)^{\otimes m} \to \left(\mathrm{H}^{-1/2}(\Gamma_\mathrm{N})^d\right)^{\otimes m} & \text{s.t.} \quad \widehat{\gamma}_m \bigotimes_{i=1}^m \mathbf{v}_i \mapsto \bigotimes_{i=1}^m \mathbf{v}_i|_{\Gamma_\mathrm{N}}.
\end{array}$$

Using the tensor notation we can state the following result concerning the computation of the objective of problem (4.2) and its shape derivative.

Theorem 4.17. *Let $\mathcal{S}_{\mathrm{adm}}$ be a class of regular enough admissible shapes sharing the portions Γ_D and Γ_N of their boundaries, so that any $\Omega \in \mathcal{S}_{\mathrm{adm}}$ is of class $\mathcal{C}^{k+2}$ with $k+1 > \frac{dm}{2}$ integer. Moreover, set $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})\right)$. If $\mathbf{u}_\Omega$ solves problem (4.3) almost surely, then*

(i) $\mathbf{u}_\Omega$ *belongs to* $\mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$.

(ii) $\mathrm{Cor}_m(\mathbf{u}_\Omega) \in \left(\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d\right)^{\otimes m}$ *is solution of the following problem*

$$\left| \begin{array}{l}
\text{Find } \mathrm{Cor}_m(\mathbf{u}_\Omega) \in \mathcal{V} = \left(\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d\right)^{\otimes m} \\
\text{such that, for all } V \in \mathcal{V}: \\
\left\langle \widehat{\mathbb{C}}_m \widehat{\nabla}_m \mathrm{Cor}_m(\mathbf{u}_\Omega), \widehat{\nabla}_m V \right\rangle_{\left(\mathrm{L}^2(\Omega)^d\right)^{\otimes m}} = \langle \widehat{\gamma}_m(V), \mathrm{Cor}_m(\mathbf{g}) \rangle_{\left(\mathrm{L}^2(\Gamma_\mathrm{N})^d\right)^{\otimes m}}.
\end{array} \right. \tag{4.12}$$

Moreover, $\mathrm{Cor}_m(\mathbf{u}_\Omega)$ *belongs to* $\bigotimes_{i=1}^m \mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d$.

(iii) Let $\widehat{P^{\Omega}}_m$ *be the tensorization of the operator* P^{Ω} *on* $\left(\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)\right)^{\otimes m}$. *We denote* $\widehat{q}_{0m} : (\mathbb{R}^d)^{\otimes m} \to \mathbb{R}$ *and* $\widehat{q}_{1m} : (\mathbb{R}^{d\times d})^{\otimes m} \to \mathbb{R}$ *the tensorizations of* q_0 *and* q_1 *respectively. Then* $\mathbb{E}\left[P^{\Omega}(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right] = \widehat{P^{\Omega}}_m(\mathrm{Cor}_m(\mathbf{u}_\Omega))$, *with*

$$\widehat{P^{\Omega}}_m(\mathrm{Cor}_m(\mathbf{u}_\Omega)) = \int_\Omega \left(\widehat{q}_{0m} + \widehat{q}_{1m} \widehat{\nabla}_m\right) \Big(\mathrm{Cor}_m(\mathbf{u}_\Omega)\Big)(\mathbf{x}, \dots, \mathbf{x}) \mathrm{d}\mathbf{x}. \tag{4.13}$$

(iv) The shape derivative of $\mathbb{E}\left[P^{\Omega}(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right]$ *can be written as*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathbb{E}\left[P^{\Omega}(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right](\boldsymbol{\theta}) =& \int_{\Gamma_0}\left(\widehat{q_0}_m+\widehat{q_1}_m\widehat{\nabla}_m\right)\Big(\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)\Big)(\mathbf{s},\dots,\mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\mathrm{d}\mathbf{s}\\ &-\int_{\Gamma_0}\langle\mathbb{C}\nabla,\nabla\rangle\Big(\mathrm{Cor}\left(\mathbf{u}_\Omega,\mathbf{w}_\Omega\right)\Big)(\mathbf{s},\mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\mathrm{d}\mathbf{s},\end{aligned}\tag{4.14}$$

where the mapping $\langle\mathbb{C}\nabla,\nabla\rangle : \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)\otimes\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)\to\mathrm{L}^1(\Omega)$ *is induced from the bilinear form* $(\hat{\mathbf{u}},\hat{\mathbf{w}})\mapsto\mathbb{C}\nabla\hat{\mathbf{u}}:\nabla\hat{\mathbf{w}}$*. The term* $\mathrm{Cor}\left(\mathbf{u}_\Omega,\mathbf{w}_\Omega\right)$ *solves the adjoint problem*

$$\left|\begin{array}{l}\text{Find } \mathrm{Cor}\left(\mathbf{u}_\Omega,\mathbf{w}_\Omega\right)\in\mathcal{W}=\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d\otimes\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d\\ \text{such that, for all } W\in\mathcal{W}:\\ \Big\langle\Big(\mathbb{C}\nabla\otimes\mathbb{C}\nabla\Big)\mathrm{Cor}\left(\mathbf{u}_\Omega,\mathbf{w}_\Omega\right),\Big(\nabla\otimes\nabla\Big)W\Big\rangle_{\mathrm{L}^2(\Omega)^d\otimes\mathrm{L}^2(\Omega)^d}\\ \qquad=\displaystyle\sum_{i=1}^m\left\langle(\gamma\otimes\mathbb{I})\,W,\mathrm{Cor}\left(\mathbf{g},\partial_i q_0(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right)\right\rangle_{\mathrm{L}^2(\Gamma_\mathrm{N})^d\otimes\mathrm{L}^2(\Omega)^d}\\ \qquad+\displaystyle\sum_{i=1}^m\left\langle(\gamma\otimes\nabla)\,W,\mathrm{Cor}\left(\mathbf{g},\partial_i q_1(\nabla\mathbf{u}_\Omega,\dots,\nabla\mathbf{u}_\Omega)\right)\right\rangle_{\mathrm{L}^2(\Gamma_\mathrm{N})^d\otimes\mathrm{L}^2(\Omega)^d}.\end{array}\right.\tag{4.15}$$

Proof. In order to prove (i), we recall that the estimate (4.6) on the norm of $\mathbf{u}_\Omega$ holds for almost any $\omega\in\mathcal{O}$. Thus, the solution of (4.3) belongs to the space $\mathrm{H}^{k+2}(\Omega)$ almost surely. Moreover, [226, Theorem 2.1] ensures that, since $\mathbf{g}\in\mathrm{L}^2\left(\mathcal{O},\mathbb{P};\mathrm{H}^{k+1/2}(\Gamma_\mathrm{N})\right)$, the random solution $\mathbf{u}_\Omega$ of problem (4.3) is unique and belongs to $\mathrm{L}^2\left(\mathcal{O},\mathbb{P};\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$. Therefore, by the uniqueness of $\mathbf{u}_\Omega$ and the elliptical estimates, we can state that $\mathbf{u}_\Omega\in\mathrm{L}^m\left(\mathcal{O},\mathbb{P};\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$.

Point (ii) can be proven by [226, Theorem 3.2], which ensures the well-posedness of problem (4.12) and the uniqueness of the solution in $\mathcal{V}$. In order to prove the regularity of $\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)$ we use point (i) and proposition 4.12 to show that, since $\mathbf{u}_\Omega$ belongs to the Bochner space $\mathrm{L}^m\left(\mathcal{O},\mathbb{P};\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$, then

$$\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)=\mathbb{E}\left[\left(\mathbf{u}_\Omega\right)^{\otimes m}\right]\in\left(\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)^{\otimes m}.$$

The existence of the linear continuous operator $\widehat{P^\Omega}_m:\left(\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)^{\otimes m}\to\mathbb{R}$ in point (iii) is a direct application of proposition 4.14. Indeed, since $\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)\subset\mathrm{W}^{1,m}_{\Gamma_\mathrm{D}}(\Omega)$ by the Sobolev embedding theorem, the restriction of the m-multilinear functional P^Ω to $\prod_{i=1}^m\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)$ is well-defined. Hence, proposition 4.14 ensures the existence and uniqueness of the linear operator $\widehat{P^\Omega}_m$ such that $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right]=\widehat{P^\Omega}_m\left(\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)\right)$.

The expression (4.13) for $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right]$ derives from the linearity of the expectation operator

$$\begin{aligned}&\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right]=\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega,\dots,\mathbf{u}_\Omega)\right]\\ &=\int_\Omega\mathbb{E}\left[q_0(\mathbf{u}_\Omega(\mathbf{x}),\dots,\mathbf{u}_\Omega(\mathbf{x}))\right]\mathrm{d}\mathbf{x}+\int_\Omega\mathbb{E}\left[q_1(\nabla\mathbf{u}_\Omega(\mathbf{x}),\dots,\nabla\mathbf{u}_\Omega(\mathbf{x}))\right]\mathrm{d}\mathbf{x}\\ &=\int_\Omega\left(\widehat{q_0}_m\Big(\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)(\mathbf{x},\dots,\mathbf{x})\Big)+\widehat{q_1}_m\Big(\widehat{\nabla}_m\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)(\mathbf{x},\dots,\mathbf{x})\Big)\right)\mathrm{d}\mathbf{x}\\ &=\int_\Omega\left(\widehat{q_0}_m+\widehat{q_1}_m\widehat{\nabla}_m\right)\Big(\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)\Big)(\mathbf{x},\dots,\mathbf{x})\,\mathrm{d}\mathbf{x}=\widehat{P^\Omega}_m\left(\mathrm{Cor}_m\left(\mathbf{u}_\Omega\right)\right).\end{aligned}$$

The variational formulation (4.15) for $\mathrm{Cor}\,(\mathbf{u}_\Omega, \mathbf{w}_\Omega)$ in point (iv) can be deduced from the application of [226, Theorem 3.2] to $\mathbf{u}_\Omega \otimes \mathbf{w}_\Omega$, knowing that, for almost any $\omega \in \mathcal{O}$, $\mathbf{u}_\Omega(\omega)$ solves problem (4.3) and $\mathbf{w}_\Omega(\omega)$ solves problem (4.11). We remark that, by point (i), $\mathbf{u}_\Omega$ belongs to $\mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$ and, by proposition 4.16, also to $\mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{W}^{1,2m-2}_{\Gamma_\mathrm{D}}(\Omega)^d\right)$. Thus

$$\sum_{i=1}^{m} \left(\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega) + \partial_i q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega)\right) \in \mathrm{L}^{1-\frac{1}{m}}\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2(\Omega)^d\right).$$

Since $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})\right) \subset \mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2(\Gamma_\mathrm{N})^d\right)$ and by proposition 4.12 we have

$$\mathbf{g} \otimes \sum_{i=1}^{m} \left(\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega) + \partial_i q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega)\right) \in \mathrm{L}^1\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2(\Gamma_\mathrm{N})^d \otimes \mathrm{L}^2(\Omega)^d\right).$$

Therefore

$$\mathrm{Cor}\left(\mathbf{g}, \sum_{i=1}^{m} \left(\partial_i q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega) + \partial_i q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega)\right)\right) \in \mathrm{L}^2(\Gamma_\mathrm{N})^d \otimes \mathrm{L}^2(\Omega)^d$$

and the right-hand side of the variational formulation in (4.15) is continuous, assuring the well-posedness of problem (4.15) by the Lax-Milgram theorem.

In order to retrieve the expression (4.14) of the shape derivative of $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right]$, we consider the derivative of the objective function for a fixed event $\omega \in \mathcal{O}$ found in the expression (4.9) of proposition 4.16. For each event $\omega \in \mathcal{O}$ we introduce the adjoint state $\mathbf{w}_\Omega(\omega)$ as the solution of problem (4.11). The expression of the derivative of $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right]$ can be found computing the expectation of proposition 4.16 and applying the tensorized operators introduced earlier

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right](\boldsymbol{\theta}) &= \mathbb{E}\left[\int_{\Gamma_0}\Big(q_0(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\right] \\
&\quad + \mathbb{E}\left[\int_{\Gamma_0}\Big(q_1(\nabla \mathbf{u}_\Omega, \dots, \nabla \mathbf{u}_\Omega)\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} - \int_{\Gamma_0}\Big(\mathbb{C}\nabla \mathbf{u}_\Omega : \nabla \mathbf{w}_\Omega\Big)(\mathbf{s}, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\right] \\
&= \mathbb{E}\left[\int_{\Gamma_0}\Big(\widehat{q_0}_m(\mathbf{u}_\Omega)^{\otimes m}\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} + \int_{\Gamma_0}\Big(\widehat{q_1}_m\widehat{\nabla}_m(\mathbf{u}_\Omega)^{\otimes m}\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\right] \\
&\quad - \mathbb{E}\left[\int_{\Gamma_0}\Big(\langle \mathbb{C}\nabla, \nabla\rangle\,(\mathbf{u}_\Omega \otimes \mathbf{w}_\Omega)\Big)(\mathbf{s}, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}\right] \\
&= \int_{\Gamma_0}\widehat{q_0}_m\Big(\mathrm{Cor}_m\,(\mathbf{u}_\Omega)\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\mathrm{d}\mathbf{s} + \int_{\Gamma_0}\widehat{q_1}_m\widehat{\nabla}_m\Big(\mathrm{Cor}_m\,(\mathbf{u}_\Omega)\Big)(\mathbf{s}, \dots, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s} \\
&\quad - \int_{\Gamma_0}\langle \mathbb{C}\nabla, \nabla\rangle\left(\mathrm{Cor}\,(\mathbf{u}_\Omega, \mathbf{w}_\Omega)\right)(\mathbf{s}, \mathbf{s})(\boldsymbol{\theta}\cdot\mathbf{n})\,\mathrm{d}\mathbf{s}
\end{aligned}$$

□

4.3.3 Shape derivatives under finite-rank noise

For this section, we consider k to be a positive integer such that $k + 1 \geq d/2$.

The numerical computation of the correlation tensor by the numerical solution problem (4.12) is not viable, given the high dimension of the tensor space. Therefore, we limit our study to

finite-rank perturbations. We consider that $\mathbf{g}$ can be written in terms of a finite number of random variables $X_1, \dots, X_m \in \mathrm{L}^m(\mathcal{O}, \mathcal{A}, \mathbb{P})$ and loads $\mathbf{g}_1, \dots, \mathbf{g}_N \in \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})$ as

$$\mathbf{g}(\omega) = \sum_{j=1}^{N} X_j(\omega)\, \mathbf{g}_j. \tag{4.16}$$

We introduce the following notations:

- $\mathfrak{A}_{(1,m),N} = \{1, \dots, N\}^m$ is the set of all m-tuples whose elements are integers between 1 and N;
- $\mathfrak{A}^{i,j}_{(1,m),N} = \left\{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N} \text{ such that } k_i = j\right\} \subset \mathfrak{A}_{(1,m),N}$ is the subset of all m-tuples in $\mathfrak{A}_{(1,m),N}$ whose i-th element is equal to j;
- we denote $C^j_{\vec{\mathbf{k}}}$ the number of times the integer j appears in the m-tuple $\vec{\mathbf{k}}$
- finally, we denote $\alpha(\vec{\mathbf{k}})$ the following quantity:

$$\alpha(\vec{\mathbf{k}}) = \alpha(k_1, \dots, k_m) = \prod_{j=1}^{N} \left(\mathbb{E}\left[X_j^{C^j_{\vec{\mathbf{k}}}} \right] \right).$$

Proposition 4.18. *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a $\mathcal{C}^{k+2}$ domain, and P^Ω be an m-multilinear continuous functional following the structure (4.4). Moreover, let $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{k+\frac{1}{2}}(\Gamma_\mathrm{N})\right)$ be a random mechanical load such that it can be decomposed as in (4.16), where the N real random variables $X_i \in \mathrm{L}^m(\mathcal{O}, \mathcal{A}, \mathbb{P})$ are mutually independent, and let $\mathbf{u}_j \in \mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)$ be the solution of the elasticity equation under the load $\mathbf{g}_j$ for $j \in \{1, \dots, N\}$. Then, $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right]$ can be written as*

$$\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right] = \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \left(\alpha(\vec{\mathbf{k}}) \int_\Omega \left(q_0(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_m}) + q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_m}) \right) \mathrm{d}\mathbf{x} \right) \tag{4.17}$$

Furthermore, we can write its the shape derivative in Ω as follows

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\Omega} \mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega)\right](\boldsymbol{\theta}) &= -\sum_{j=1}^{N} \int_{\Gamma_0} (\boldsymbol{\theta} \cdot \mathbf{n})\, (\mathbb{C}\nabla \mathbf{u}_j : \nabla \mathbf{w}_j)\, \mathrm{d}\mathbf{s} \\ + \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \alpha(\vec{\mathbf{k}}) \left(\int_{\Gamma_0} (\boldsymbol{\theta} \cdot \mathbf{n})\, \left(q_0(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_m}) + q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_m}) \right) \mathrm{d}\mathbf{s} \right) \end{aligned} \tag{4.18}$$

where the N states $\mathbf{u}_1, \dots, \mathbf{u}_N$ solve the state equation for $\mathbf{g}_1, \dots, \mathbf{g}_N$ respectively and belong to $\mathrm{H}^{k+2}_{\Gamma_\mathrm{D}}(\Omega)$, while the N adjoint states $\mathbf{w}_1, \dots, \mathbf{w}_N$ belong to $\mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)$ and solve the following adjoint problems

$$\begin{cases} -\mathrm{div}\,(\mathbb{C}\nabla \mathbf{w}_j) = \displaystyle\sum_{i=1}^{m} \sum_{\vec{\mathbf{k}} \in \mathfrak{A}^{i,j}_{(1,m),N}} \alpha(\vec{\mathbf{k}}) \Big(\partial_i q_0(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_m}) - \mathrm{div}\, \partial_i q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_m}) \Big) \text{ in } \Omega, \\ (\mathbb{C}\nabla \mathbf{w}_j)\, \mathbf{n} = \displaystyle\sum_{i=1}^{m} \sum_{\vec{\mathbf{k}} \in \mathfrak{A}^{i,j}_{(1,m),N}} \alpha(\vec{\mathbf{k}})\, (\partial_i q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_m}))^T\, \mathbf{n} \quad \text{on } \Gamma_0 \cup \Gamma_\mathrm{N}, \\ \mathbf{w}_j = \mathbf{0} \quad \text{on } \Gamma_\mathrm{D}. \end{cases} \tag{4.19}$$

Proof. At first we remark that, by the linearity of the elasticity equation, the decomposition (4.16) can be extended to the displacement $\mathbf{u}_\Omega$ as

$$\mathbf{u}_\Omega(\omega) = \sum_{j=1}^{N} X_j(\omega)\mathbf{u}_j$$

The expression (4.18) derives directly from the linearity of the expected value and the m-linearity of P^Ω. Indeed, for almost all event $\omega \in \mathcal{O}$,

$$P^\Omega\left(\mathbf{u}_\Omega(\omega), \dots, \mathbf{u}_\Omega(\omega)\right) = \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \left(\Big(\prod_{j=1}^{N} X_j^{C_{\vec{\mathbf{k}}}^j} \Big) P^\Omega(\mathbf{u}_{k_1}(\omega), \dots, \mathbf{u}_{k_m}(\omega)) \right).$$

Therefore

$$\mathbb{E}\left[P^\Omega\left(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega\right)\right] = \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \alpha(\vec{\mathbf{k}}) P^\Omega(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_m}).$$

In order to compute the shape derivative of $\mathbb{E}\left[P^\Omega\left(\mathbf{u}_\Omega, \dots, \mathbf{u}_\Omega\right)\right]$ we use once again Céa's fast derivative method [61, 6] as done for proposition 4.16. We introduce the following Lagrangian function $\mathcal{L} : \mathcal{S}_{\text{adm}} \times (\mathrm{H}^1\left(\mathbb{R}^d\right)^d)^N \times (\mathrm{H}^1\left(\mathbb{R}^d\right)^d)^N \to \mathbb{R}$ associated to problem (4.2) where the state equation is seen as a PDE constraint

$$\begin{aligned}
\mathcal{L}\left(\Omega; \hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_N; \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N\right) = &- \sum_{j=1}^{N} \Big\{ \int_\Omega \left(\mathbb{C}\nabla\hat{\mathbf{u}}_j : \nabla\hat{\mathbf{w}}_j\right) \mathrm{d}\mathbf{x} - \int_{\Gamma_\mathrm{N}} \mathbf{g}_j \cdot \hat{\mathbf{w}}_j \, \mathrm{d}\mathbf{s} \\
&- \int_{\Gamma_\mathrm{D}} \left(\hat{\mathbf{w}}_j \cdot \left(\mathbb{C}\nabla\hat{\mathbf{u}}_j\right)\mathbf{n} + \hat{\mathbf{u}}_j \cdot \left(\mathbb{C}\nabla\hat{\mathbf{w}}_j - \sum_{\ell}^{N} \partial_\ell q_1(\nabla\hat{\mathbf{u}}_{k_1}, \dots, \nabla\hat{\mathbf{u}}_{k_m}) \right)\mathbf{n} \right) \mathrm{d}\mathbf{s} \Big\} \\
&+ \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \Big\{ \alpha(\vec{\mathbf{k}}) \Big(\int_\Omega q_0(\hat{\mathbf{u}}_{k_1}, \dots, \hat{\mathbf{u}}_{k_m}) \, \mathrm{d}\mathbf{x} + \int_\Omega q_1(\nabla\hat{\mathbf{u}}_{k_1}, \dots, \nabla\hat{\mathbf{u}}_{k_m}) \, \mathrm{d}\mathbf{x} \Big) \Big\}.
\end{aligned} \tag{4.20}$$

The variables $\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N$ act as Lagrange multipliers for the PDE constraints of the terms $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_N$. In order to assure that all arguments of the Lagrangian are independent, the terms $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_N$ and $\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N$ are defined on the whole space $\mathbb{R}^d$, and not only on Ω.

By construction, the terms $\mathbf{u}_1, \dots, \mathbf{u}_N$ solving the equation $\frac{\partial \mathcal{L}}{\partial \hat{\mathbf{w}}_j} = 0$, are also solutions of the state equation for the right-hand side $\mathbf{g} = \mathbf{g}_j$. Thus, we can express the functional $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right]$ in terms of the Lagrangian:

$$\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right] = \mathcal{L}\left(\Omega; \mathbf{u}_1, \dots, \mathbf{u}_N; \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N\right), \tag{4.21}$$

for all $\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N \in \mathrm{H}^1\left(\mathbb{R}^d\right)$.

The expression for the shape derivative of the functional of interest is found differentiating equation (4.21) with respect to the domain Ω

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right](\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}\Omega}\mathcal{L}\left(\Omega; \mathbf{u}_1, \dots, \mathbf{u}_N; \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N\right)(\boldsymbol{\theta}) \\
=&\frac{\partial \mathcal{L}}{\partial \Omega}\left(\Omega, \mathbf{u}_1, \dots, \mathbf{u}_N, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N\right)(\boldsymbol{\theta}) + \sum_{j=1}^{N} \frac{\partial \mathcal{L}}{\partial \hat{\mathbf{u}}_i}\left(\Omega, \mathbf{u}_1, \dots, \mathbf{u}_N, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_N\right)(\mathbf{u}_j').
\end{aligned} \tag{4.22}$$

The term $\mathbf{u}_j'$ denotes the Eulerian derivative of $\mathbf{u}_j$, which is the derivative of the mapping $t \mapsto \mathbf{u}_j(\Omega_{t\boldsymbol{\theta}})$ in $t = 0$, where $\mathbf{u}_j(\Omega_{t\boldsymbol{\theta}})$ is the unique solution of the state equation for the right-hand side $\mathbf{g}_j$ and on the deformed domain $\Omega_{t\boldsymbol{\theta}}$.

Next, we remark that, by choosing $\mathbf{w}_1, \dots, \mathbf{w}_N$ solving the adjoint problem (4.19), the quantity $\frac{\partial \mathcal{L}}{\partial \hat{\mathbf{u}}_j}$ vanishes for $j = 1 \dots N$. Indeed, for any $\mathbf{v} \in \mathrm{H}^1\left(\mathbb{R}^d\right)$ we have

$$
\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial \hat{\mathbf{u}}_j}(\Omega, \mathbf{u}_1, \dots, \mathbf{u}_N, \mathbf{w}_1, \dots, \mathbf{w}_N)(\mathbf{v}) \\
&= \sum_{i=1}^{m} \sum_{\vec{\mathbf{k}} \in \mathfrak{A}^{i,j}_{(1,m),N}} \int_\Omega \alpha(\vec{\mathbf{k}}) \left(\partial_i q_0(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_N})(\mathbf{v}) + \partial_i q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_N}) : (\nabla \mathbf{v})\right) \mathrm{d}\mathbf{x} \\
&\quad - \int_\Omega \left(\mathbb{C} \nabla \mathbf{w}_j : \nabla \mathbf{v}\right) \mathrm{d}\mathbf{x} + \int_{\Gamma_\mathrm{D}} \left(\mathbf{w}_j \cdot (\mathbb{C} \nabla \mathbf{v} \mathbf{n}) + \mathbf{v} \cdot (\mathbb{C} \nabla \mathbf{w}_j \mathbf{n})\right) \mathrm{d}\mathbf{s} \\
&= \int_\Omega \left(\sum_{i=1}^{m} \sum_{\vec{\mathbf{k}} \in \mathfrak{A}^{i,j}_{(1,m),N}} \alpha(\vec{\mathbf{k}}) \left(\partial_i q_0(\mathbf{u}_1, \dots, \mathbf{u}_N) - \left(\operatorname{div} \partial_i q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_N})\right)\right) + \operatorname{div} \mathbf{w}_j \right) \cdot \mathbf{v} \, \mathrm{d}\mathbf{x} \\
&\quad + \int_{\Gamma_\mathrm{N}} \left(-\mathbb{C} \nabla \mathbf{w}_j \, \mathbf{n} + \left(\partial_i q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_N})\right)^T \mathbf{n}\right) \cdot \mathbf{v} \, \mathrm{d}\mathbf{s} + \int_{\Gamma_\mathrm{D}} \mathbf{w}_j \left(\mathbb{C} \nabla \mathbf{v} \, \mathbf{n}\right) \mathrm{d}\mathbf{s} = 0.
\end{aligned}
$$

By taking $\mathbf{w}_1, \dots, \mathbf{w}_N$ as solutions of problem (4.19) for $j = 1 \dots N$, we can further simplify the expression (4.22) for the shape derivative of $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right]$ and obtain

$$
\frac{\mathrm{d}}{\mathrm{d}\Omega} \mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right](\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}\left(\Omega, \mathbf{u}_1, \dots, \mathbf{u}_N, \mathbf{w}_1, \dots, \mathbf{w}_N\right)(\boldsymbol{\theta}). \tag{4.23}
$$

For simplicity, we consider the portions Γ_N and Γ_D of the boundary to be non-optimizable, which is equivalent to narrow the set of admissible displacement fields $\boldsymbol{\theta}$ to the set Θ_{adm} defined as

$$
\Theta_{adm} = \left\{ \boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d, \mathbb{R}^d\right) \, : \, \boldsymbol{\theta} = 0 \text{ on } \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N} \right\}.
$$

Thanks to the restriction of the admissible displacement fields to Θ_{adm} and to [138, Theorem 5.2.2], we conclude that the shape derivative of $\mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right]$ can be expressed as

$$
\frac{\mathrm{d}}{\mathrm{d}\Omega} \mathbb{E}\left[P^\Omega(\mathbf{u}_\Omega, \cdots, \mathbf{u}_\Omega)\right](\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}\left(\Omega, \mathbf{u}_1, \dots, \mathbf{u}_N, \mathbf{w}_1, \dots, \mathbf{w}_N\right)(\boldsymbol{\theta}) = \int_{\Gamma_0} (\boldsymbol{\theta} \cdot \mathbf{n}) A(\mathbf{s}) \, \mathrm{d}\mathbf{s}.
$$

with

$$
A = -\sum_{i=1}^{m} \left(\mathbb{C} \nabla \mathbf{u}_i : \nabla \mathbf{w}_i\right) + \sum_{\vec{\mathbf{k}} \in \mathfrak{A}_{(1,m),N}} \alpha(\vec{\mathbf{k}}) \left(q_0(\mathbf{u}_{k_1}, \dots, \mathbf{u}_{k_m}) + q_1(\nabla \mathbf{u}_{k_1}, \dots, \nabla \mathbf{u}_{k_m})\right).
$$

□

It is worth remarking that the method presented in this section requires the computation of only N adjoint states. Moreover, the PDEs defining the states $\mathbf{u}_1, \dots, \mathbf{u}_N$ and the adjoint states $\mathbf{w}_1, \dots, \mathbf{w}_N$ all share the structure of their left-hand side. This property can be useful for the numerical simulations since, by inverting once the matrix representing the discretization of the bilinear form $(\mathbf{u}, \mathbf{v}) \mapsto \int_\Omega \left(\mathbb{C} \nabla \mathbf{u} : \nabla \mathbf{v}\right) \mathrm{d}\mathbf{x}$, we can solve the $2N$ boundary value problems faster.

The authors of [80] proved that the expected value of a quadratic functional of the displacement and its shape derivative depend only from the first two stochastic moments of the random

variables modeling the uncertain boundary conditions. Theorem 4.17 and proposition 4.18 extend such result to m-multilinear functionals, showing that their expectation can be expressed as a function of the first m stochastic moments of the uncertainties.

Let us denote $\mathfrak{N}\left(P^{\Omega}, N\right)$ the minimal number of terms to be computed in (4.17) and (4.18) to express the expected value of the functional and its derivative. In the most general case, $\mathfrak{N}\left(P^{\Omega}, N\right) = N^m$, since we have to compute all the terms in the form $P^{\Omega}\left(\mathbf{g}_{k_1}, \dots, \mathbf{g}_{k_m}\right)$, as well as their shape derivatives. However, this number can be reduced if the multilinear functional P^{Ω} shows some symmetries among its arguments. Indeed, if P^{Ω} is completely symmetric, we have $\mathfrak{N}\left(P^{\Omega}, N\right) = \binom{N+m-1}{m}$.

4.4 Optimization under constraints on the von Mises stress

4.4.1 Estimate of the expected value of the von Mises stress

An application of polynomial functionals in shape optimization is related to the approximation of the L^{∞}-norm of a given quantity in a structure by the L^m-norm, for m sufficiently large. A significant concern in structural mechanics is the design of structures where the stress is as evenly distributed as possible, preventing stress concentrations that could compromise the integrity of the component. This requirement suggests the use of functionals with order $m > 2$ in order to better penalize stress concentrations than quadratic functionals.

As a showcase, we study the optimization of a 3D linear elastic structure (thus $d = 3$) with respect to its volume and the L^m-norm of the von Mises stress, for $m \geq 2$ even integer. We suppose that the optimization problem is framed as problem (4.2), and that the random external load $\mathbf{g} \in \mathrm{L}^m\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2\left(\Omega\right)^d\right)$ can be decomposed as in equation (4.16), with $k \geq 1$ integer.

We are interested in estimating the expected value of the following quantity $\mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$, representing the L^m-norm of the von Mises stress in a structure Ω subject to a random load $\mathbf{g}$. The functional $\mathcal{H}_m(\cdot,\cdot)$ is defined, for m even, as

$$\Omega \mapsto \mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega) = H_m(\mathbf{u}_{\Omega,\mathbf{g}}, \dots, \mathbf{u}_{\Omega,\mathbf{g}}), \tag{4.24}$$

where $H_m : \left[\mathrm{W}^{1,m}\left(\Omega\right)^3\right]^d \to \mathbb{R}$ is such that

$$H_m(\mathbf{v}_1, \dots, \mathbf{v}_m) = \int_{\Omega} \left(\prod_{i=1}^{m/2} \left(\left(\boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{v}_{2i-1}\right) : \boldsymbol{\sigma}_{\mathrm{D}}\left(\mathbf{v}_{2i}\right) \right) \right) \right) \mathrm{d}\mathbf{x}. \tag{4.25}$$

We remark that, because of the concavity of the mapping $x \mapsto \sqrt[m]{x}$, the following bound on the expectation of the L^m-norm of the von Mises stress holds

$$\mathbb{E}\left[\left\| s_{\mathrm{D}}\left(\mathbf{u}\right) \right\|_{\mathrm{L}^m(\Omega)} \right] \leq \sqrt{\frac{3}{2}} \left(\mathbb{E}\left[H_m(\mathbf{u}, \dots, \mathbf{u}) \right] \right)^{\frac{1}{m}}. \tag{4.26}$$

The functional H_m respects the structure defined in equation (4.4). Therefore, we can apply proposition 4.18 to compute the shape derivative of the functional $\Omega \mapsto \mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$. The expression of the functional can be further simplified by considering the symmetries between the arguments of $\mathcal{H}_m$.

Definition 4.19. *We establish the following notation.*

- *We denote*
$$\mathfrak{B}_{\ell,N} = \left\{ \vec{\boldsymbol{\rho}} \in \mathbb{N}^{N\times N} \ : \ 0 \leq \rho_{ij} \leq \ell \ \text{ and } \ \sum_{i,j=1}^{N} \rho_{ij} = \ell \right\}$$
the set of all $N \times N$ integer matrices whose entries are positive and their sum is equal to ℓ. The cardinality of said set can be computed as $\operatorname{Card}(\mathfrak{B}_{\ell,N}) = \binom{N^2+\ell-1}{\ell}$.

- *For ℓ and N positive integers and $\vec{\boldsymbol{\rho}} \in \mathfrak{B}_{\ell,N}$, we define the following multinomial coefficient*
$$\binom{\ell}{\vec{\boldsymbol{\rho}}} = \frac{\ell!}{\prod_{i,j=1}^{N}(\rho_{i,j}!)}.$$

- *For N real random variables $X_1, \ldots, X_m \in \mathrm{L}^m(\mathcal{O}, \mathbb{P}; \mathbb{R})$ and $\vec{\boldsymbol{\rho}} \in \mathfrak{B}_{\frac{m}{2},N}$, we denote:*
$$K(\vec{\boldsymbol{\rho}}) = \binom{\frac{m}{2}}{\vec{\boldsymbol{\rho}}} \prod_{j=1}^{N} \mathbb{E}\left[X_j^{\sum_{k=1}^{N}(\rho_{kj}+\rho_{jk})} \right].$$

Having introduced the necessary notation to take the symmetries among the arguments into account, we can write the expectation of the functional $\mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)$ as

$$\mathbb{E}\left[\mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)\right] = \sum_{\vec{\boldsymbol{\rho}} \in \mathfrak{B}_{\frac{m}{2},N}} \left\{ K(\vec{\boldsymbol{\rho}}) \int_{\Omega} \prod_{j,k=1}^{N} \left(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_j) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_k)\right)^{\rho_{jk}} \mathrm{d}\mathbf{x} \right\}, \tag{4.27}$$

where each $\mathbf{u}_j$ solves the state equation problem (4.3) with the loadings $\mathbf{g}_j$ for $j \in \{1, \ldots, N\}$. Since the functional F respects the structure defined in equation (4.4), we can apply proposition 4.18 and find the following expression for the shape derivative of $\mathbb{E}\left[\mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)\right]$

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\Omega}\mathbb{E}\left[\mathcal{H}_m(\mathbf{u}_{\Omega,\mathbf{g}}, \Omega)\right](\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}\Omega}\mathbb{E}\left[F(\mathbf{u}, \ldots, \mathbf{u})\right] \\ &= \int_{\Gamma_0} (\boldsymbol{\theta}\cdot\mathbf{n}) \left(-\sum_{j=1}^{N} (\mathbb{C}\nabla\mathbf{u}_j : \nabla\mathbf{w}_j) + \sum_{\vec{\boldsymbol{\rho}} \in \mathfrak{B}_{\frac{m}{2},N}} \left\{ K(\vec{\boldsymbol{\rho}}) \prod_{j,k=1}^{N} \left(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_j) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_k)\right)^{\rho_{jk}} \right\} \right) \mathrm{d}\mathbf{s}. \end{aligned} \tag{4.28}$$

The adjoint states $\mathbf{w}_1, \ldots, \mathbf{w}_N$ solve the following adjoint equations

$$\left\{ \begin{array}{rcll} -\operatorname{div}\mathbb{C}\nabla\mathbf{w}_j &=& -2\mu \operatorname{div}\left(\displaystyle\sum_{k=1}^{N} L_{jk}\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_k) \right) & \text{in } \Omega \\ (\mathbb{C}\nabla\mathbf{w}_j)\mathbf{n} &=& 2\mu \left(\displaystyle\sum_{k=1}^{N} L_{jk}\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_k) \right)\mathbf{n} & \text{on } \Gamma_0 \cup \Gamma_{\mathrm{N}} \\ \mathbf{w}_j &=& \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}, \end{array} \right. \tag{4.29}$$

where the terms $L_{jk} \in \mathrm{L}^{m-1}(\Omega)$ are defined as

$$L_{jk} = 2 \sum_{\vec{\boldsymbol{\rho}} \in \mathfrak{B}_{\frac{m}{2},N}} \left(K(\vec{\boldsymbol{\rho}})\rho_{jk} \left(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_j) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_k)\right)^{\rho_{jk}-1} \prod_{\ell\neq k} \left(\boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_j) : \boldsymbol{\sigma}_{\mathrm{D}}(\mathbf{u}_\ell)\right)^{\rho_{j\ell}} \right).$$

We can notice also that, thanks to the symmetries of the von Mises functional H_m defined as in equation (4.25), it is not necessary to compute all the Card $\left(\mathfrak{B}_{\frac{m}{2},N}\right)$ terms of the sums in the formulas (4.27) and (4.28). Instead, the computation of $\mathfrak{N}(H_m, N) = \binom{\frac{N(N+1)+m}{2} - 1}{\frac{m}{2}}$ terms is sufficient, provided that they are counted with their respective multiplicity.

4.4.2 Numerical application

As showcase we study the following shape optimization problem

$$\left|\begin{array}{l} \text{Find } \Omega \in \mathcal{S}_{\text{adm}} \text{ minimizing } \operatorname{Vol}(\Omega), \\ \text{where, for all } \omega \in \mathcal{O}, \text{ the state } \mathbf{u}_\Omega \in \mathrm{H}^1(\Omega)^3 \text{ solves:} \\ \qquad \left\{\begin{array}{rcll} -\operatorname{div} \boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega)) & = & \mathbf{0} & \text{in } \Omega, \\ \boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\,\mathbf{n} & = & \mathbf{g}(\omega) & \text{on } \Gamma_{\mathrm{N}}, \\ \boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\,\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\ \mathbf{u}_\Omega(\omega) & = & \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \\ \text{with the constraint:} \quad \mathbb{E}\left[\mathcal{H}_6(\Omega, \mathbf{u}_{\Omega,\mathbf{g}})\right] \leq M_0^6, \end{array}\right. \tag{4.30}$$

where $M_0 > 0$ is a given upper bound for the constraint functional $\mathbb{E}\left[\mathcal{H}_6(\Omega, \mathbf{u}_{\Omega,\mathbf{g}})\right]$.

The structure to be optimized is a cylinder-like shape with axis $z = 0$, reported in fig. 4.2. Dirichlet boundary conditions are imposed on a thin stripe on the lateral surface, while the random load $\mathbf{g}$ is applied on a ring-shaped section on the upper surface of the structure. We consider the mechanical load $\mathbf{g} \in \mathrm{L}^6\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2(\Gamma_{\mathrm{N}})^d\right)$ to have the following structure

$$\mathbf{g}(\omega) = \mathbf{g}_1\, X_1(\omega) + \mathbf{g}_2\, X_2(\omega) \qquad \text{for almost all event } \omega \in \mathcal{O}.$$

The loads $\mathbf{g}_1$ and $\mathbf{g}_2$ are set as constant vectors on Γ_{N}, parallel to the axes x and y respectively, thus tangent to the surface. Moreover, we consider the random variables X_1 and X_2 to follow centered Gaussian distributions with variance σ_1 and σ_2 respectively.

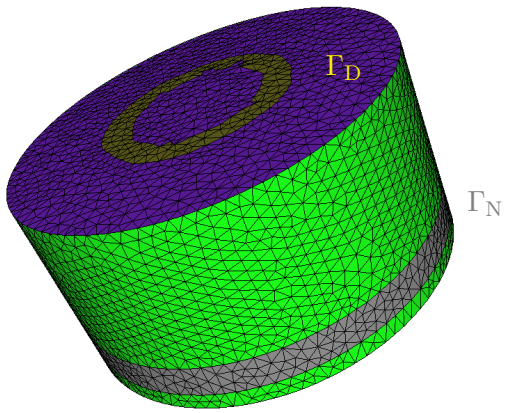
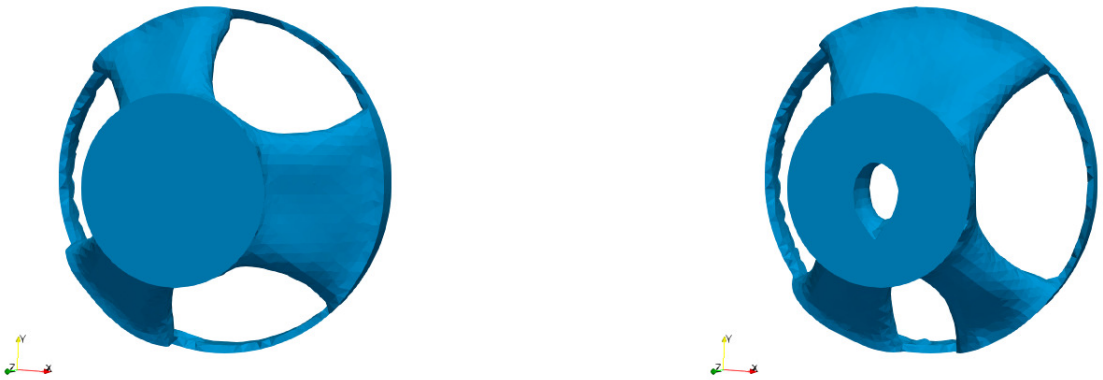


Figure 4.2: Representation of the structure to be optimized. The surface Γ_{D} is the thin grey stripe on the lateral surface, while Γ_{N} is the ring-shaped portion of the upper surface marked in yellow.

Unfortunately, the hypothesis (4.5) about the separation of the Dirichlet and Neumann boundaries cannot be verified in most practical situations. Indeed, the regularity of the displacement $\mathbf{u}_\Omega$ is limited by the possible appearance of a finite number of singularities around the junctions of the two portions of the boundary where natural or essential conditions are imposed [50]. In this section we will not focus on the study of the compatibility conditions to avoid the emergence of singularities, but we present the results of some simulations where no difficulty related to the regularity of the solution has been observed.

From the numerical point of view, we represent the structure by using a level-set function on a fixed mesh $\mathcal{T}$ covering a fixed domain D containing every admissible shape in $\mathcal{S}_{\text{adm}}$. The linear elasticity equations (4.3) and the adjoint problems (4.19) are defined on the entire domain $D = \Omega \cup \Omega^C$, by using an *ersatz material* approximation in Ω^C to assure the well-posedness of the problems (see [19, 84]). The elasticity and adjoint equations are solved by using the *FreeFem++* environment [136].

The numerical results of two different simulations are discussed: in the first case we consider the random variables X and Y to have an identical distribution (isotropic distribution of the external mechanical load), while the second case considers an asymmetry in the variances of the two random variables (anisotropic distribution). The parameters used in the simulation are reported in table 4.1. The shapes obtained by the execution of 200 iterations of the *null space optimization* algorithm for both cases are reported in fig. 4.3, and the convergences of the objective and the constraint functions in fig. 4.4. All simulations have been performed on a Virtualbox virtual machine Linux with 1GB of dedicated memory, installed on a Dell PC equipped with a 2.80 GHz Intel i7 processor. The numerical results are reported in table 4.2.



(a) Isotropic mechanical load

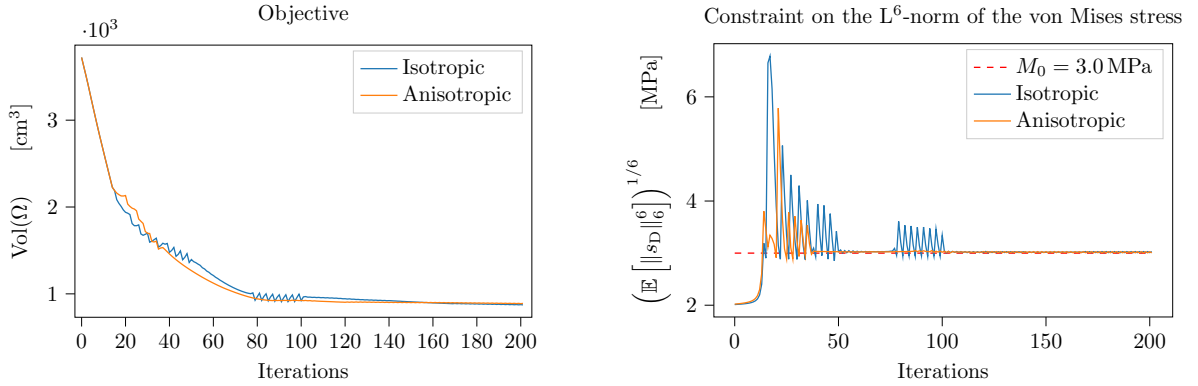
(b) Anisotropic mechanical load

Figure 4.3: Optimal shapes for problem (4.30) under isotropic and anisotropic mechanical loads applied on Γ_{N}.

From the observation of fig. 4.3 and fig. 4.4 we remark firstly the efficiency of the *null space optimization* algorithm in the solution of the constrained optimization problem (4.30). Indeed, the value of the objective functional is decreasing (see fig. 4.4a). As seen in fig. 4.4b, the constraint on the expectation of $\mathcal{H}_6$ is saturated in less than 50 iterations for the anisotropic case. In the isotropic case, we observe some oscillations in the constraint saturation around iteration 80, which are due to a change in the topology around that step of the optimization.

Heigth of the domain D		12.0 cm	
Radius of the cylinder D		12.0 cm	
Region Γ_N			
inner radius		4.0 cm	
outer radius		6.0 cm	
Region Γ_D			
thickness		2.0 cm	
distance from the edge of D		1.0 cm	
Mesh size parameters			
minimal element size	`hmin`	0.4	
maximal element size	`hmax`	0.8	
gradation value	`hgrad`	1.3	
Elastic coefficients			
Young's modulus	E	15 MPa	
Poisson's ration	ν	0.35	
Ersatz material coefficient	εers	10^{-3}	
Threshold	M_0	3.0 MPa	
Variances of the random variables		**isotropic**	**anisotropic**
variance of X_1	σ_1^2	2.5	1.0
variance of X_2	σ_2^2	2.5	4.0

Table 4.1: Numerical parameters for problem (4.30) for the cases of random variables with equal and with different variances.



(a) Convergence of the volume of Ω, as the objective

(b) Convergence of the constraint $\mathbb{E}\left[\mathcal{H}_6(\Omega, \mathbf{u}_{\Omega,\mathbf{g}})\right] - M_0^6$

Figure 4.4: Convergence of the objective and constraint of problem (4.30).

	Isotropic case	**Anisotropic case**
Number of iterations	200	200
Execution time	129 minutes	148 minutes
Final volume Vol(Ω)	872.93 cm³	890.30 cm³
Normalized saturation of the constraint $(\mathbb{E}\left[\mathcal{H}_6\right] - M_0^6)/M_0^6$	0.03002	0.005351

Table 4.2: Numerical results of the solution of problem (4.30) for an isotropic and anisotropic mechanical load.

The shapes of fig. 4.3 show that a branched structure presents the minimal volume ensuring enough resistance with respect to random mechanical loads. Moreover, if the direction of the mechanical load $\mathbf{g}(\cdot)$ is not uniformly distributed in the interval $[0, 2\pi]$, the branches tend to align parallel to the most probable direction of the load (see fig. 4.3b). Finally, we remark that the constraint imposed in problem (4.30) is a quite conservative estimate for the expected value of the L^6-norm of the von Mises stress. Thanks to the inequality equation (4.26) and the fact that the optimal shapes respect the constraint $\mathbb{E}\left[\mathcal{H}_6\right] \leq M_0^6$, we deduce that the average of the L^6-norm of the von Mises stress in the structures is actually less than the chosen threshold M_0.

4.5 Optimization under constraints on expectation and variance of a quadratic functional

4.5.1 Expression of the variance of the mechanical compliance

The technique presented in 4.3.2 can be applied to compute the shape derivative of the variance of a quadratic functional. As an example, we consider the optimization of a 2D bridge-like structure with respect to the expectation and the variance of the mechanical compliance. We recall that the compliance of a shape Ω is defined in section 2.1.2 as the work of the external forces $\mathbf{g}$ acting on Ω and it can be expressed as:

$$\widehat{\mathcal{C}}(\Omega) = \mathcal{C}\left(\Omega, \mathbf{u}_\Omega\right) = \int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}_\Omega\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}_\Omega\right) \mathrm{d}\mathbf{x} = C_\Omega(\mathbf{u}_\Omega, \mathbf{u}_\Omega), \tag{4.31}$$

where $\mathbf{u}_\Omega$ is the displacement computed as solution to the elasticity equation under the application of the load $\mathbf{g}$, and $C_\Omega : \mathrm{H}^1\left(\Omega\right)^d \times \mathrm{H}^1\left(\Omega\right)^d \to \mathbb{R}$ is a continuous bilinear functional. From its expression, and from the fact that the elasticity equation is linear, we recognize that the compliance is a quadratic functional of the applied load $\mathbf{g}$.

We suppose that the structure Ω is enclosed in a square computational domain D of size 1.0×1.0, its lower side Γ_D is clamped, and a random uniform load $\mathbf{g}$ is applied on the upper side Γ_N (see fig. 4.6a). We assume the random load $\mathbf{g}$ to have the following structure:

$$\mathbf{g}(\omega) = \mathbf{g}_0 + X(\omega)\,\mathbf{g}_x + Y(\omega)\,\mathbf{g}_y, \qquad \text{for all } \omega \in \mathcal{O} \tag{4.32}$$

where $X, Y \in \mathrm{L}^4\left(\mathcal{O}, \mathbb{P}; \mathbb{R}\right)$ are real-valued random variables, and $\mathbf{g}_0 = (0, -g_0)$, $\mathbf{g}_x = (g_x, 0)$, and $\mathbf{g}_y = (0, -g_y)$. Under such hypotheses, the variance of the compliance can be written as

$$\begin{aligned}
\mathrm{Var}\left[\mathcal{C}\left(\Omega, \mathbf{u}_\Omega\right)\right] =& \mathrm{Var}\left[C_\Omega(\mathbf{u}_\Omega, \mathbf{u}_\Omega\right] = \mathbb{E}\left[X^4\right] C_\Omega(\mathbf{u}_x, \mathbf{u}_x)^2 + 4\mathbb{E}\left[X^3\,Y\right] C_\Omega(\mathbf{u}_x, \mathbf{u}_x)\, C_\Omega(\mathbf{u}_x, \mathbf{u}_y) \\
&+ \mathbb{E}\left[X^2\,Y^2\right]\left(2\, C_\Omega(\mathbf{u}_x, \mathbf{u}_x)\, C_\Omega(\mathbf{u}_y, \mathbf{u}_y) + 4\, C_\Omega(\mathbf{u}_x, \mathbf{u}_y)^2\right) \\
&+ 4\mathbb{E}\left[X\,Y^3\right] C_\Omega(\mathbf{u}_x, \mathbf{u}_y)\, C_\Omega(\mathbf{u}_y, \mathbf{u}_y) + \mathbb{E}\left[Y^4\right] C_\Omega(\mathbf{u}_y, \mathbf{u}_y)^2 \\
&- \left(\mathbb{E}\left[X^2\right] C_\Omega(\mathbf{u}_x, \mathbf{u}_x) + 2\,\mathbb{E}\left[X\,Y\right] C_\Omega(\mathbf{u}_x, \mathbf{u}_y) + \mathbb{E}\left[Y^2\right] C_\Omega(\mathbf{u}_y, \mathbf{u}_y)\right)^2.
\end{aligned} \tag{4.33}$$

Given $\beta, \alpha \in [0, \frac{\pi}{2})$, we define the following random variables:

$$\begin{aligned}
X_\alpha &= T \sin\alpha + N_X \cos\alpha, \\
Y_{\alpha,\beta} &= \frac{\sin\beta}{\sqrt{\mathrm{Var}\left[X_\alpha^2\right]}}\left(X_\alpha^2 - \mathbb{E}\left[X_\alpha^2\right]\right) + N_Y \cos\beta,
\end{aligned} \tag{4.34}$$

where T, N_X, and N_Y are independent random variables such that T is uniformly distributed on the set $\{-1, 1\}$, while N_X, and N_Y follow a standard Gaussian distribution. The densities of the variables X_α and $Y_{\alpha,\beta}$ for different values of β and α are represented in fig. 4.5. It can be remarked that, unless $\beta = 0$, X_α and $Y_{\alpha,\beta}$ are not independent from one another. However, for any choice of $\beta, \alpha \in [0, \frac{\pi}{2})$, they are centered, normalized and uncorrelated, that is:

$$\mathbb{E}\left[X_\alpha\right] = \mathbb{E}\left[Y_{\alpha,\beta}\right] = 0, \qquad \mathbb{E}\left[X_\alpha^2\right] = \mathbb{E}\left[Y_{\alpha,\beta}^2\right] = 1, \qquad \mathbb{E}\left[X_\alpha\, Y_{\alpha,\beta}\right] = 0.$$

Therefore, solving different shape optimization problems on structures charged by random mechanical loads with the structure (4.32) for different parameters α and β allows to highlight the importance of the fourth order moment of the random variables.

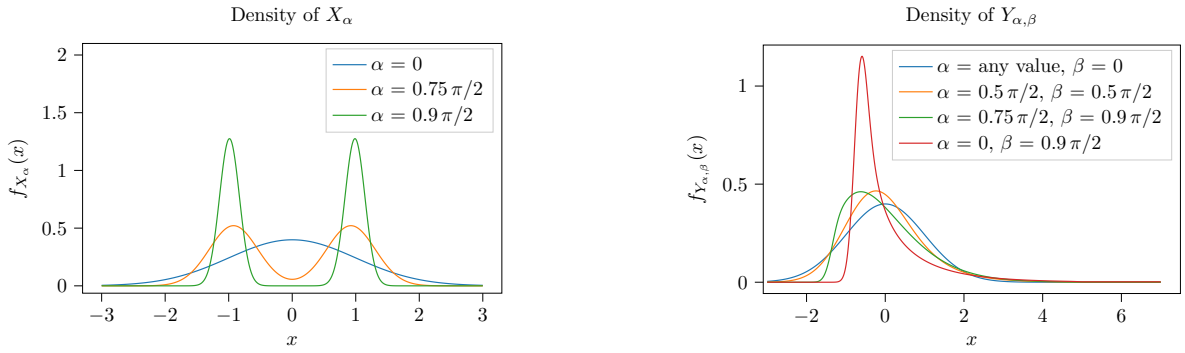


Figure 4.5: Probability densities of the random variables X_α and $Y_{\alpha,\beta}$ for different values of the parameters $\alpha, \beta \in [0, \frac{\pi}{2})$ as introduced in (4.34).

A deterministic expression for the variance of a generic cost functional has been proposed in [11] for the case of small perturbations, by using a linearization procedure. However, equation (4.33) provides a deterministic expression for the variance of the compliance without any assumption on the size of the uncertainties, and can be differentiated thanks to proposition 4.18.

4.5.2 An optimization problem for the variance of the compliance

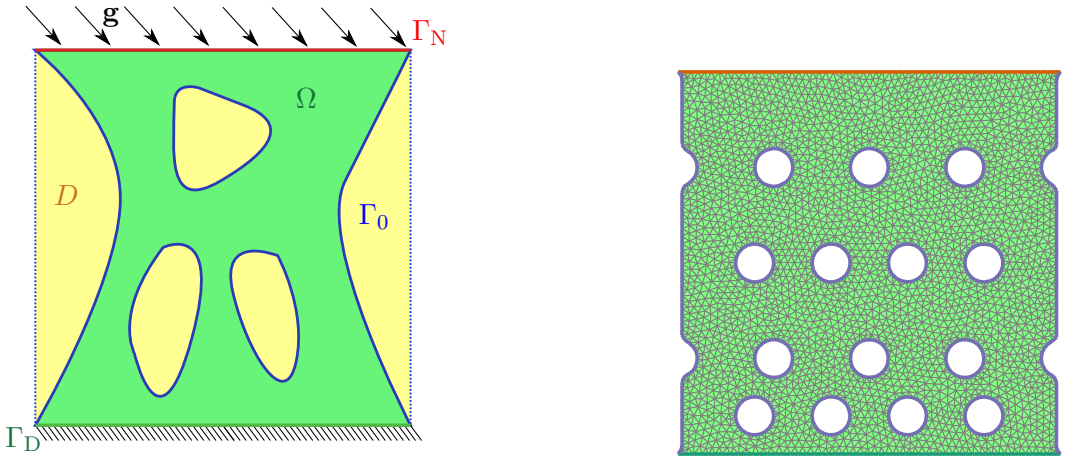
Given the setting outlined in section 4.5.1, we consider the following optimization problem:

$$\left|\begin{array}{l}
\text{Find } \Omega \in \mathcal{S}_{\text{adm}} \text{ minimizing } \operatorname{Vol}(\Omega), \\
\text{where, for all } \omega \in \mathcal{O}, \text{ the state } \mathbf{u}_\Omega \in \mathrm{H}^1(\Omega)^2 \text{ solves:} \\
\qquad \left\{\begin{array}{ll}
-\operatorname{div}\boldsymbol{\sigma}(\mathbf{u}_\Omega) = 0 & \text{in } \Omega, \\
\boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\,\mathbf{n} = \mathbf{g}(\omega) & \text{on } \Gamma_{\mathrm{N}}, \\
\boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\,\mathbf{n} = 0 & \text{on } \Gamma_0, \\
\mathbf{u}_\Omega(\omega) = 0 & \text{on } \Gamma_{\mathrm{D}}.
\end{array}\right. \\
\text{and the following constraint holds:} \\
\qquad \mathbb{E}\left[C_\Omega(\mathbf{u}_\Omega(\omega), \mathbf{u}_\Omega(\omega))\right] \leq M_0, \\
\qquad \operatorname{Var}\left[C_\Omega(\mathbf{u}_\Omega(\omega), \mathbf{u}_\Omega(\omega))\right] \leq M_1.
\end{array}\right. \tag{4.35}$$

The terms M_0 and M_1 are thresholds for the expectation and the variance of the compliance not to be exceeded.

At first, we remark that, without the constraint on the variance, the solution of problem (4.35) would be the same for any choice of $\beta, \alpha \in [0, \frac{\pi}{2})$. Indeed, as stated in 4.14, the value of $\mathbb{E}\left[C_\Omega(\mathbf{u}_\Omega(\omega), \mathbf{u}_\Omega(\omega))\right]$ is a function of $\mathrm{Cor}_2(\mathbf{g}, \mathbf{g})$, which depends only on the first two moments of the variables X_α and $Y_{\alpha,\beta}$. It turns out that X_α and $Y_{\alpha,\beta}$ share the same expected value, variance and correlation, for any $\beta, \alpha \in [0, \frac{\pi}{2})$.

The optimization is performed numerically by using the *nullspace optimization* algorithm on an adaptive 2D mesh, as in [114] using the *mmg* platform for the mesh adaptation [89]. The results of two simulations are presented, under different parameters α and β. In both cases, we used as initial condition the structure shown in fig. 4.6b. The numerical parameters are listed in table 4.3, and the results in table 4.4. Once again, the simulations have been done on a Dell PC with a 2.80 GHz Intel i7 processor, on a Virtualbox virtual machine Linux with 1GB of dedicated memory.



(a) Test case

(b) Initial condition

Figure 4.6: Representation of the test case described in section 4.5.1 and initial condition.

The final results of the optimization, the evolution of the volume and of the expectation and variance of the compliance in the two cases are shown in fig. 4.7, fig. 4.8, and fig. 4.9.

At first, we remark the similarity in the optimal structures for the two problems, represented in fig. 4.7. However, the structure in fig. 4.7a is thicker than fig. 4.7b, and the final volume of the solution of case 1 is 31% higher than the volume occupied by the solution of case 2. Such a gap is explained by the fact that the variance of the compliance is significantly different in the two cases, as shown by fig. 4.9b. Indeed, in the first case, the constraint on the variance is saturated first, while in the second case the variance stays small throughout the optimization and the constraint on the expected value is saturated instead.

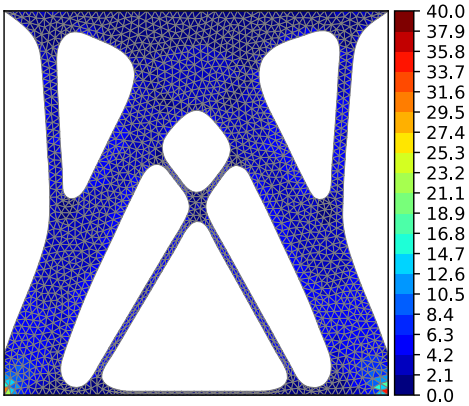
This example underlines the importance of the high order moments of the uncertainties in the domain of robust optimization, in particular when the random variables describing the boundary condition show a strong dependence from each other.

4.6 Conclusions and perspectives

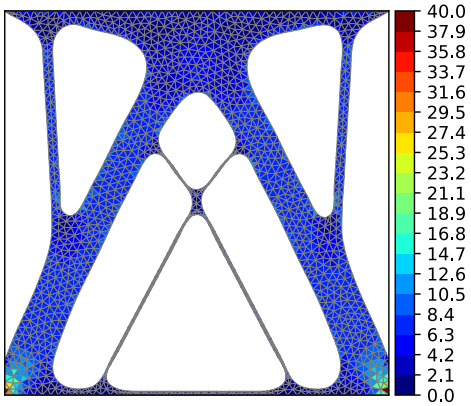
This chapter focused on a procedure of shape optimization of polynomial functionals, where the external load applied to the structure is subject to uncertainties. Particular attention has been

Geometry of the structure	
heigth of the domain	1.0
length of the domain	1.0
Mesh size parameters	
minimal element size `hmin`	0.01
maximal element size `hmax`	0.02
gradation value `hgrad`	0.5
Elastic coefficients	
Young's modulus E	15
Poisson's ration ν	0.35
Mechanical loads	
Fixed load $\mathbf{g}_0$	1.2
Horizontal term $\mathbf{g}_x$	1.0
Vertical term $\mathbf{g}_y$	0.3
Thresholds for the inequality constraints	
Treshold for the expected value M_0	2.0
Treshold for the variance M_1	3.0625

Table 4.3: Numerical parameters for problem (4.35).



(a) Optimal structure for the choice of parameters $\alpha = 0.0$ and $\beta = 0.95\frac{\pi}{2}$ (case 1)



(b) Optimal structure for the choice of parameters $\alpha = 0.95\frac{\pi}{2}$ and $\beta = 0.0$ (case 2)

Figure 4.7: Solution of the shape optimization problem (4.35) for two choices of α and β. The color scale represents the expectation of the concentration of the elastic energy $\mathbb{E}\left[\boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\varepsilon}\left(\mathbf{u}\right)\right]$.

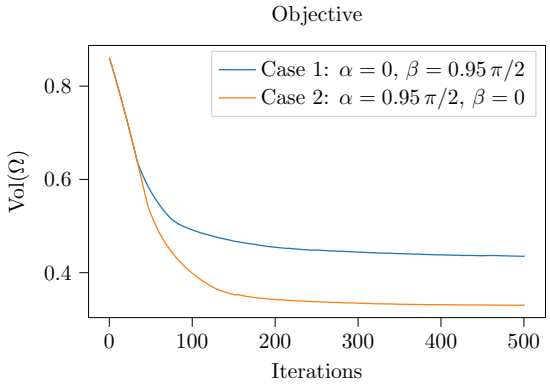


Figure 4.8: Volume of the structure during the optimization problem for two choices of the parameters α and β.

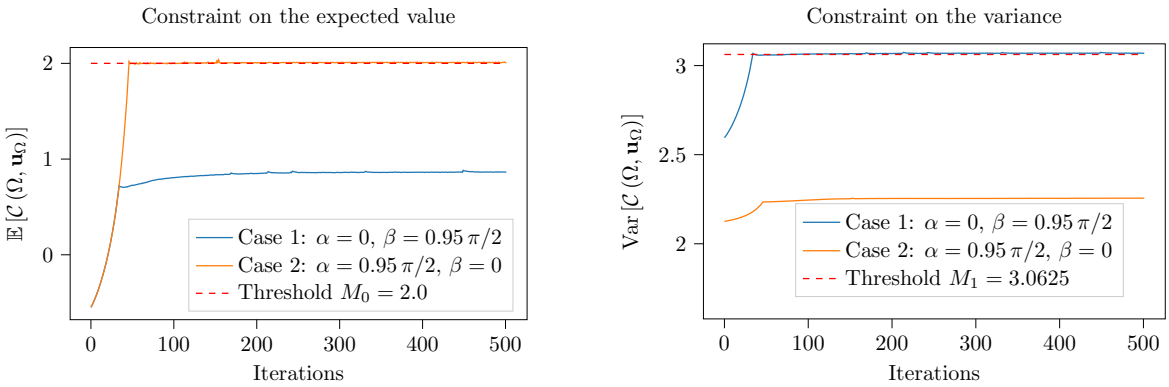


Figure 4.9: Values of the expectation and the variance of the mechanical compliance throughout the optimization.

	Case 1	**Case 2**
Parameters for X_α and $Y_{\alpha,\beta}$		
Parameter α	0.0	$0.95\,\frac{\pi}{2}$
Parameter β	$0.95\,\frac{\pi}{2}$	0.0
Execution of the optimization		
Duration of the optimization	25 min 49 s	30 min 48 s
Number of iterations	500	500
Final volume Vol(Ω)	0.435348	0.330348
Saturation of the constraints		
on the expected value $\mathbb{E}[C_{\Omega\Omega}] - M_0$	−0.56886	0.004041
on the variance $\mathrm{Var}[C_{\Omega\Omega}] - M_1$	0.03952	−2.94801

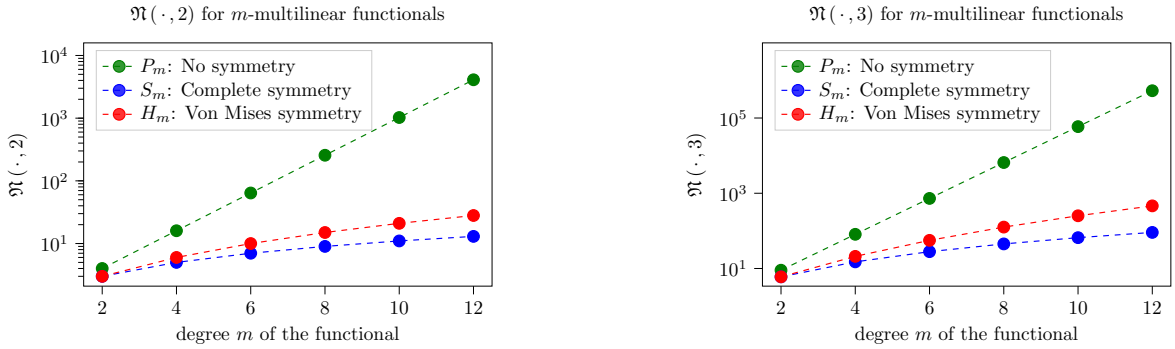
Table 4.4: Results of the numerical solution of problem (4.35) for two choices of α and β.

payed to the optimization of linear elastic structures, and we adopted the level-set approach to topology optimization. The present section proposes an extension of the technique proposed in [80] to the case of continuous multilinear functionals, and relies on the linearization properties of the tensor product between elements of a Banach space.

A significant obstacle in the application of this method is the number of terms appearing in the sums of equation (4.17) (for the computation of the functional of interest), and equation (4.18) (for its derivative). Let us recall the definition of $\mathfrak{N}(P, N)$ introduced at the end of section 4.3.3 as the minimal number of terms that are necessary to compute $\mathbb{E}[P(\mathbf{u}, \dots, \mathbf{u})]$ and its derivative, where P is a m-multilinear functional, and $\mathbf{u}$ is described by N random variables. Let us consider three different bounded m-multilinear functionals: a generic functional P, a functional S which is completely symmetric in its arguments, and the von Mises functional G defined in equation (4.25). We recall that in section 4.3.3 and in section 4.4.1 we found the following expressions for the number of terms necessary to compute the expectations of such functionals

$$\mathfrak{N}(P, N) = N^m, \quad \mathfrak{N}(S, N) = \binom{N+m-1}{m} \text{ and } \mathfrak{N}(G, N) = \binom{\frac{N(N+1)}{2} + \frac{m}{2} - 1}{\frac{m}{2}}. \tag{4.36}$$

As represented in fig. 4.10, the number of terms to be computed increases rapidly with the degree m of the multilinear functional, even if the number of random variables N is limited to 2 or 3. Naturally, the presence of symmetries in the multilinear mapping greatly reduces the number of terms to be computed, but the problem can still become too complex if the degree m required is too high.



(a) Case of $N = 2$ random variables

(b) Case of $N = 3$ random variables

Figure 4.10: Evaluation of the number of terms in the expression of the expectation of three m-multilinear functionals showing different symmetries among their arguments, when the stochastic aspect is modeled by two or three random variables.

As remedy to this issue, we suggest to exploit the symmetric nature of the correlation tensor, and study the application of some techniques of tensor decomposition. One promising solution consists in the approximation of the discretized correlation tensor as a sum of tensor of rank 1, by using the CP-decomposition. This technique and other kinds of tensor decomposition are detailed in [155, 154, 75], and have been implemented in Python libraries as TensorLy [158]. However, its interpretation and applicability in the field of shape optimization are still to be investigated.

Chapter 5

Shape optimization of the probability to exceed a threshold

Portions of this chapters have been submitted for publication as a journal paper under the title *Shape optimization under constraints on the probability of a quadratic functional to exceed a given threshold*, written in collaboration with Marc Dambrine, Helmut Harbrecht, and Jérôme Maynadier.

Contents

5.1 Introduction

In this chapter, we are once again interested in the solution of a constrained shape optimization problem on a set of mechanical structures subject to a random mechanical loading $\mathbf{g} = \mathbf{g}(\omega)$. Unlike in chapters 3 and 4, we restrict our attention to a constraint functional $\mathcal{Q}(\mathbf{g}, \Omega)$ depending quadratically on the uncertain parameter $\mathbf{g}$. Among such functionals we count the mechanical compliance and the L^2-norm of the von Mises stress. The objective is the identification of the structure Ω with the smallest volume for which the probability of failure $\mathbb{P}[\mathcal{Q}(\mathbf{g}, \Omega) > \tau]$ does not exceed a prescribed threshold.

The problem under consideration is known in the literature as Reliability-Based Topology optimization [121], and is known to be computationally hard as the probability of failure defines a quantity of interest which is not smooth with respect to the random parameter ω. Numerous approaches to reliability-based parametric optimization applied to structure optimization are cited in the manual [69]. The authors of [34, 180, 147, 152] use a density-based approach to represent the structure to be optimized, and estimate the probability of failure using a First and a Second Order Reliability Method [134]. In [151, 179] the authors use a polynomial chaos technique to estimate the probability of failure and compute its sensitivity with respect to a set of parameters characterizing the structure. A different technique is proposed in [76, 11], where Reliability-Based Topology Optimization problems have been tackled by approximating the non-smooth functional by a smooth one.

In the present setting of a quadratic shape functional, we will show that the region, where $\mathcal{Q}(\mathbf{g}, \Omega) > \tau$ holds, is the exterior of an ellipsoid with respect to the stochastic parameter ω. We will exploit this fact in order to compute the shape derivative of the problem under consideration and to derive an efficient, deterministic shape optimization algorithm. Unlike in the approach proposed in [11], we make no assumptions on the amplitude of the perturbation. Moreover, in the case of Gaussian perturbations, the computation of the probability of failure and its sensitivity with respect to the shape are computed by suitable formulas relying on recurrence relations, without the need of computationally expensive sampling techniques.

The rest of this chapter is structured as follows. In section 5.2, we introduce the model problem and compute the shape functional and its shape gradient. Section 5.4 is then dedicated to our showcase, where we suppose that the loading $\mathbf{g} = \mathbf{g}(\omega)$ is a Gaussian random field. We develop a suitable quadrature formula which can be used to numerically compute the shape functional and the associated shape gradient. Then, in 5.5 we present the details of the algorithms used for the numerical simulations, and two examples are compiled in 5.6 in order to demonstrate the feasibility of the present approach. In particular, we compare the method presented in section 5.4 with the generic expression of the shape derivatives computed in section 5.2 in the case of centered Gaussian distributions, using the quadrature formulas of appendix C. The coefficients for the evaluation of the probability of failure and their shape derivatives are computed in python using the algorithms detailed in appendix D.

5.2 The shape optimization problem

5.2.1 Problem statement

Let us consider a family of Lipschitz continuous admissible domains $\mathcal{S}_{\text{adm}}$ in $\mathbb{R}^d$ (for $d = 2$ or 3) sharing the portions Γ_{N} and Γ_{D}, which we suppose to be disjoint. For each $\Omega \in \mathcal{S}_{\text{adm}}$, we denote $\Gamma_0 = \partial\Omega \backslash (\Gamma_{\text{N}} \cup \Gamma_{\text{D}})$ the optimizable portion of the boundary. We suppose that the structure to be optimized is made up of a linear elastic material, characterized by the Lamé parameters λ and μ,

and is clamped on Γ_{D}. Let further $(\mathcal{O}, \mathcal{A}, \mathbb{P})$ be a probability space, where $\mathcal{A} \subset 2^{\mathcal{O}}$ is a σ-algebra on $\mathcal{O}$ and $\mathbb{P}$ is a probability measure. A random mechanical load $\mathbf{g} \in \mathrm{L}^2\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})^d\right)$ is applied on the portion Γ_{N} of the boundary. In particular, we suppose that $\mathbf{g}$ can be written in terms of a deterministic term $\overline{\mathbf{g}}_0$ and a finite number N of random terms in accordance with

$$\mathbf{g}(\omega) = \overline{\mathbf{g}}_0 + \overline{\mathbf{g}}_1 X_1(\omega) + \ldots + \overline{\mathbf{g}}_N X_N(\omega) \qquad \text{for almost all } \omega \in \mathcal{O}, \tag{5.1}$$

where $X_1, \ldots, X_N \in \mathrm{L}^2(\mathcal{O}, \mathbb{P}; \mathbb{R})$ are independent, real-valued random variables, and the terms $\overline{\mathbf{g}}_0, \ldots, \overline{\mathbf{g}}_N$ belong to $\mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})^d$. Then, for almost any event $\omega \in \mathcal{O}$, the displacement $\mathbf{u}_\Omega(\omega) \in \mathrm{H}^1(\Omega)^d$ is the solution of the linear elasticity system

$$\left\{\begin{array}{rcll} -\operatorname{div}\boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega)) &=& \mathbf{0} & \text{in } \Omega, \\ \boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\mathbf{n} &=& \mathbf{0} & \text{on } \Gamma_0, \\ \boldsymbol{\sigma}(\mathbf{u}_\Omega(\omega))\mathbf{n} &=& \mathbf{g}(\omega) & \text{on } \Gamma_{\mathrm{N}}, \\ \mathbf{u}_\Omega(\omega) &=& \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}. \end{array}\right. \tag{5.2}$$

In this chapter, we suppose that the functional of interest for the constraint $\mathcal{Q}(\mathbf{g}, \Omega)$ is a quadratic function of the displacement $\mathbf{u}_\Omega$, and thus can be written as

$$\mathcal{Q}(\mathbf{g}, \Omega) = \langle \mathbf{u}_\Omega, Q\,\mathbf{u}_\Omega \rangle_{\mathrm{H}^1(\Omega)},$$

where Q is a self-adjoint positive definite endomorphism in $\mathrm{H}^1(\Omega)^d$. We consider the shape optimization problem

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\mathrm{adm}} \\ \text{minimizing } \operatorname{Vol}(\Omega) \\ \text{under the constraint} \\ \qquad \mathbb{P}\left[\langle \mathbf{u}_\Omega, Q\,\mathbf{u}_\Omega \rangle_{\mathrm{H}^1(\Omega)} > \tau\right] \leq \bar{p}, \\ \text{where the state } \mathbf{u}_\Omega(\omega) \\ \text{satisfies the state equation (5.2) for almost all } \omega \in \mathcal{O}, \end{array}\right. \tag{5.3}$$

where $\mathcal{S}_{\mathrm{adm}}$ is a class of Lipschitz continuous and uniformly bounded domains in $\mathbb{R}^d$. Note that the safety criterion $\mathcal{Q}(\mathbf{g}, \Omega)$ is supposed to be a quadratic functional of the displacement $\mathbf{u}_\Omega$. As we intend to adopt the moving boundary approach introduced in section 1.2, we require that, for any $\mathbf{g} \in \mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})^d$, the mapping $\Omega \mapsto \mathcal{Q}(\mathbf{g}, \Omega)$ is differentiable with respect to the shape.

5.2.2 Properties of the safety criterion

We shall highlight the dependency of the constraint $\mathbb{P}\left[\langle \mathbf{u}_\Omega, Q\,\mathbf{u}_\Omega \rangle_{\mathrm{H}^1(\Omega)} > \tau\right]$ from the random variables $X_1, \ldots, X_N$ appearing in the definition (5.1) of the mechanical load. For all $i \in \{1, \ldots, N\}$, we define the displacement $\mathbf{u}_{\Omega,i} \in \mathrm{H}^1(\Omega)^d$ as the solution of the following deterministic elasticity problem

$$\left\{\begin{array}{rcll} -\operatorname{div}\boldsymbol{\sigma}(\mathbf{u}_{\Omega,i}) &=& \mathbf{0} & \text{in } \Omega, \\ \boldsymbol{\sigma}(\mathbf{u}_{\Omega,i})\mathbf{n} &=& \mathbf{0} & \text{on } \Gamma_0, \\ \boldsymbol{\sigma}(\mathbf{u}_{\Omega,i})\mathbf{n} &=& \overline{\mathbf{g}}_i & \text{on } \Gamma_{\mathrm{N}}, \\ \mathbf{u}_{\Omega,i} &=& \mathbf{0} & \text{on } \Gamma_{\mathrm{D}}, \end{array}\right. \tag{5.4}$$

where $\boldsymbol{\sigma}(\mathbf{u}) = \mathbb{C}\nabla\mathbf{u}$ identifies the Cauchy stress tensor for the displacement $\mathbf{u} \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)$. Thanks to the linearity of the state equation (5.2), the displacement $\mathbf{u}_\Omega \in \mathrm{L}^2(\mathcal{O}, \mathbb{P}; \mathrm{H}^1(\Omega))$ can be written as a sum of N terms, depending from the same random variables as in (5.1)

$$\mathbf{u}_\Omega(\omega) = \mathbf{u}_{\Omega,0} + \mathbf{u}_{\Omega,1}X_1(\omega) + \ldots + \mathbf{u}_{\Omega,N}X_N(\omega) \qquad \text{for almost all } \omega \in \mathcal{O}. \tag{5.5}$$

The safety functional is quadratic with respect to the displacement, and we can express it as a quadratic function $\Psi_\Omega : \mathbb{R}^N \to \mathbb{R}$ of the random vector $\mathbf{X} = (X_1, \ldots, X_N) \in \mathrm{L}^2\left(\mathcal{O}, \mathbb{P}; \mathbb{R}^N\right)$ as

$$\mathcal{Q}(\mathbf{g}(\omega), \Omega) = \Psi_\Omega(\mathbf{X}(\omega)) = \mathbf{X}(\omega)^\mathrm{T}\mathbf{M}_\Omega\,\mathbf{X}(\omega) + 2\mathbf{b}_\Omega{}^\mathrm{T}\mathbf{X}(\omega) + c_\Omega, \tag{5.6}$$

for almost all $\omega \in \mathcal{O}$. The symmetric matrix $\mathbf{M}_\Omega \in \mathbb{R}^{N\times N}$, the vector $\mathbf{b}_\Omega \in \mathbb{R}^N$, and the scalar c_Ω are functions of the displacements $\mathbf{u}_{\Omega,1}, \ldots, \mathbf{u}_{\Omega,N}$, and are defined as

- $[\mathbf{M}_\Omega]_{i,j} = \langle \mathbf{u}_{\Omega,i}, Q\,\mathbf{u}_{\Omega,j}\rangle_{\mathrm{H}^1(\Omega)}$ for all $i, j \in \{1, \ldots, N\}$;
- $[\mathbf{b}_\Omega]_k = \langle \mathbf{u}_{\Omega,0}, Q\,\mathbf{u}_{\Omega,k}\rangle_{\mathrm{H}^1(\Omega)}$ for all $k \in \{1, \ldots, N\}$;
- $c_\Omega = \langle \mathbf{u}_{\Omega,0}, Q\,\mathbf{u}_{\Omega,0}\rangle_{\mathrm{H}^1(\Omega)}$.

Since Q is a self-adjoint positive definite operator, the matrix $\mathbf{M}_\Omega$ is symmetric, having N eigenvalues $\lambda_{\Omega,1}, \ldots, \lambda_{\Omega,N}$ that are real and strictly positive. We denote $\mathbf{D}_\Omega = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$ the matrix of eigenvalues, and $\mathbf{U}_\Omega \in \mathbb{R}^{N\times N}$ the orthogonal matrix of eigenvectors such that $\mathbf{M}_\Omega = \mathbf{U}_\Omega\mathbf{D}_\Omega\mathbf{U}_\Omega{}^\mathrm{T}$.

Let us consider the (deterministic) subset of $\mathbb{R}^N$ $\mathcal{E}_{\Psi_\Omega,\tau}$ containing all the realizations of the random vector $\mathbf{X}$ for which the constraint is satisfied:

$$\mathcal{E}_{\Psi_\Omega,\tau} = \left\{\mathbf{x} \in \mathbb{R}^N \,:\, \Psi_\Omega(\mathbf{x}) \le \tau\right\}. \tag{5.7}$$

We denote $\widetilde{\tau}_\Omega$ the following quantity:

$$\widetilde{\tau}_\Omega = \tau - \left(c_\Omega - \mathbf{b}_\Omega{}^\mathrm{T}\mathbf{M}_\Omega{}^{-1}\mathbf{b}_\Omega\right). \tag{5.8}$$

Given the properties of the quadratic function Ψ_Ω and assuming that $\widetilde{\tau}_\Omega > 0$, we recognize that $\mathcal{E}_{\Psi_\Omega,\tau}$ is an ellipsoid in $\mathbb{R}^N$, centered in $-\mathbf{M}_\Omega{}^{-1}\mathbf{b}_\Omega$, and whose semi-axes are oriented as the eigenvectors of $\mathbf{M}_\Omega$ and have length $r_1^{\Omega,\tau}, \ldots, r_N^{\Omega,\tau}$:

$$r_i^{\Omega,\tau} = \sqrt{\widetilde{\tau}_\Omega/\lambda_{\Omega,i}} \qquad \text{for all } i \in \{1, \ldots, N\}. \tag{5.9}$$

However, if $\widetilde{\tau}_\Omega < 0$, we have that $\mathcal{E}_{\Psi_\Omega,\tau} = \emptyset$, and the constraint cannot be satisfied for any $\bar{p} < 1$.

For the sake of clarity, we introduce the shape functional $\Phi : \mathcal{S}_\mathrm{adm} \to \mathbb{R}$ defined as the probability of the constraint to be satisfied:

$$\Phi(\Omega) = \mathbb{P}[\mathcal{Q}(\mathbf{g}, \Omega) \le \tau] = 1 - \mathbb{P}[\mathcal{Q}(\mathbf{g}, \Omega) > \tau].$$

The inequality constraint in problem (5.3) can be written alternatively as $\Phi(\Omega) \ge 1 - \bar{p}$. Especially, $\Phi(\Omega)$ can be expressed by means of the probability for the random vector $\mathbf{X}$ to belong to the ellipsoid $\mathcal{E}_{\Psi_\Omega,\tau}$

$$\Phi(\Omega) = \mathbb{P}[\Psi_\Omega(\mathbf{X}) \le \tau] = \mathbb{P}[\mathbf{X} \in \mathcal{E}_{\Psi_\Omega,\tau}].$$

Therefore, $\Phi(\Omega)$ can be interpreted as the volume of the ellipsoid $\mathcal{E}_{\Psi_\Omega,\tau}$ with respect to the probability measure $\mathbb{P}_\mathbf{X}$ induced by the random variable $\mathbf{X}$:

$$\Phi(\Omega) = \mathbb{P}[\mathbf{X} \in \mathcal{E}_{\Psi_\Omega,\tau}] = \mathbb{P}_\mathbf{X}(\mathcal{E}_{\Psi_\Omega,\tau}) = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} 1\,\mathrm{d}\mathbb{P}_\mathbf{X}(\mathbf{x}). \tag{5.10}$$

5.3 Sensitivity of the exceeding probability

5.3.1 General formula

In order to solve problem (5.3) using a gradient-based optimization algorithm, we have to compute an expression for $\Phi(\Omega)$ and for its shape derivative $\mathrm{D}\Phi(\Omega)$. To this end, let us suppose that the random vector $\mathbf{X}$ admits a probability density function $f : \mathbb{R}^N \to \mathbb{R}^+$, such that $f \in \mathrm{W}^{1,1}\left(\mathbb{R}^N\right)$. Then, in view of (5.10), the quantity $\Phi(\Omega)$ can be written as:

$$\Phi(\Omega) = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{5.11}$$

Moreover, we suppose that all entries of $\mathbf{M}_\Omega$ and $\mathbf{b}_\Omega$, as well as c_Ω are differentiable with respect to the shape, and we denote their shape derivatives evaluated in $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}(\Omega)^d$ by $\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})$, $\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta})$, and $\mathrm{D}c_\Omega(\boldsymbol{\theta})$, respectively.

We recognize in (5.11) the expression of the integral of a constant function over a variable domain $\mathcal{E}_{\Psi_\Omega,\tau}$. Let $\boldsymbol{\xi} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^N;\mathbb{R}^N\right)$ be a Lipschitz continuous deformation field in $\mathbb{R}^N$. Then, we can compute the derivative of the mapping $\boldsymbol{\xi} \mapsto \mathbb{P}\left[\mathbf{X} \in (\mathbb{I}+\boldsymbol{\xi})\mathcal{E}_{\Psi_\Omega,\tau}\right]$ thanks to the usual shape differentiation techniques (see [138, Eq. (5.24)]). Moreover, since $\mathcal{E}_{\Psi_\Omega,\tau}$ is an ellipsoid and supposing that $\boldsymbol{\xi}$ is also $\mathcal{C}^1$, we can apply Hadamard's regularity theorem (see [138, Proposition 5.9.1]) and write

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{\boldsymbol{\xi}}}\mathbb{P}\left[\mathbf{X} \in (\mathbb{I}+\tilde{\boldsymbol{\xi}})\mathcal{E}_{\Psi_\Omega,\tau}\right]\Big|_{\tilde{\boldsymbol{\xi}}=0}(\boldsymbol{\xi}) = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} \mathrm{div}\boldsymbol{\xi}(\mathbf{x})\, f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\partial\mathcal{E}_{\Psi_\Omega,\tau}} f(\mathbf{s})\,(\boldsymbol{\xi}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s}))\,\mathrm{d}\mathbf{s}. \tag{5.12}$$

Here, for all $\mathbf{s} \in \partial\mathcal{E}_{\Psi_\Omega,\tau}$, $\mathbf{n}(\mathbf{s}) \in \mathbb{R}^N$ is the unitary vector orthogonal to $\partial\mathcal{E}_{\Psi_\Omega,\tau}$ in $\mathbf{s}$.

Lemma 5.1. *Let us consider an admissible domain $\Omega \in \mathcal{S}_{\mathrm{adm}}$ and a regular enough displacement field $\boldsymbol{\theta} \in \mathcal{C}^1 \cap \mathrm{W}^{1,\infty}(\Omega)^d$. We denote $\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} \in \mathbb{R}^{N\times N}$ and $\mathbf{r}_{\Omega,\boldsymbol{\theta}} \in \mathbb{R}^N$ the matrix and the vector respectively defined as*

$$\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} = \frac{\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta})}{2\widetilde{\tau}_\Omega}\mathbb{I} - \frac{1}{2}{\mathbf{M}_\Omega}^{-1}\,\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta}); \tag{5.13}$$

$$\mathbf{r}_{\Omega,\boldsymbol{\theta}} = -{\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta}) + \left(\frac{\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta})}{2\widetilde{\tau}_\Omega}\mathbb{I} + \frac{1}{2}{\mathbf{M}_\Omega}^{-1}\,\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right){\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega, \tag{5.14}$$

where $\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta})$ has the expression

$$\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta}) = -\mathrm{D}c_\Omega(\boldsymbol{\theta}) - {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta}){\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega + {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta}). \tag{5.15}$$

Then, $\boldsymbol{\xi}^{\boldsymbol{\theta}} : \mathbf{x} \mapsto \boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}\,\mathbf{x} + \mathbf{r}_{\Omega,\boldsymbol{\theta}}$ is a $\mathcal{C}^1$ Lipschitz continuous displacement field on $\mathbb{R}^N$ such that the shape derivative of $\Phi(\cdot)$ in Ω can be written in its volumic and surface forms as

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} \mathrm{div}\Big(f(\mathbf{x})\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x})\Big)\,\mathrm{d}\mathbf{x} = \int_{\partial\mathcal{E}_{\Psi_\Omega,\tau}} f(\mathbf{s})\,(\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{s})\cdot\mathbf{n}(\mathbf{s}))\,\mathrm{d}\mathbf{s}. \tag{5.16}$$

Proof. Let $\delta > 0$ be such that, for any $t \in [0,\delta]$, $\widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega} > 0$. We consider the following dynamical system:

$$\left\{\begin{array}{rcll} \dot{\mathbf{x}}(t;\bar{\mathbf{x}}) & = & \boldsymbol{\Xi}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega,\boldsymbol{\theta}}\mathbf{x}(t;\bar{\mathbf{x}}) + \mathbf{r}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega,\boldsymbol{\theta}} & \text{for } t \in [0,\delta], \bar{\mathbf{x}} \in \mathbb{R}^N, \\ \mathbf{x}(0;\bar{\mathbf{x}}) & = & \bar{\mathbf{x}} & \text{for } \bar{\mathbf{x}} \in \mathbb{R}^N. \end{array}\right. \tag{5.17}$$

We set

$$\mathbf{y}(t,\theta,\bar{\mathbf{x}}) := \mathbf{x}(t;\bar{\mathbf{x}}) + {\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}}^{-1}\mathbf{b}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}$$

and remark that the quantity defined as

$$\mathfrak{F}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\big(\mathbf{y}(t,\theta,\bar{\mathbf{x}})\big) = \frac{\mathbf{y}(t,\theta,\bar{\mathbf{x}})^{\mathrm{T}}\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\mathbf{y}(t,\theta,\bar{\mathbf{x}})}{\widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}}$$

is constant along the trajectories. Indeed, by expressions (5.13), (5.14), and (5.17), it holds

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{F}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\big(\mathbf{y}(t,\theta,\bar{\mathbf{x}})\big) = \widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}^{-2}\Big[&- \mathrm{D}\widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}(\boldsymbol{\theta})\ \mathbf{y}(t,\theta,\bar{\mathbf{x}})^{\mathrm{T}}\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\mathbf{y}(t,\theta,\bar{\mathbf{x}}) \\ &+ \widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\Big(\mathbf{y}(t,\theta,\bar{\mathbf{x}})^{\mathrm{T}}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\mathbf{y}(t,\theta,\bar{\mathbf{x}}) + 2\mathbf{y}(t,\theta,\bar{\mathbf{x}})^{\mathrm{T}} \\ &\times\Big(\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\dot{\mathbf{x}}(t;\bar{\mathbf{x}}) - \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}{\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}}^{-1}\mathbf{b}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega} + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{b}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}\Big)\Big)\Big] = 0.\end{aligned}$$

Moreover, for any $t \in [1,\delta]$, the inequality $\mathfrak{F}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}(\mathbf{x}) \leq 1$ defines the same ellipsoid $\mathcal{E}_{\Psi_{(\mathbb{I}+t\boldsymbol{\theta})\Omega},\tau}$ as the inequality $\Psi_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}(\mathbf{x}) \leq \tau$. Therefore, the deformation $\mathbf{x} \mapsto (\mathbb{I}+\mathcal{F}_t)\,\mathbf{x}$ gives the identity $\mathcal{E}_{\Psi_{(\mathbb{I}+t\boldsymbol{\theta})\Omega},\tau} = (\mathbb{I}+\mathcal{F}_t)\,\mathcal{E}_{\Psi_\Omega,\tau}$, where $\mathcal{F}_t : \mathbb{R}^N \to \mathbb{R}^N$ is defined as $\mathcal{F}_t\,\mathbf{x} = \int_0^t \dot{\mathbf{x}}(s;\mathbf{x})\mathrm{d}s$ for $t \in [0,\delta]$.

We recall that, for any differentiable shape functional F and Lipschitz continuous domain $D \in \mathbb{R}^N$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}F\Big(\big(\mathbb{I}+\boldsymbol{\xi}(t)\big)D\Big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}D}F(D)\left(\boldsymbol{\xi}'(0)\right), \tag{5.18}$$

provided that $\boldsymbol{\xi} : [0,\delta] \to \mathrm{W}^{1,\infty}\left(\mathbb{R}^N;\mathbb{R}^N\right)$ is a differentiable mapping that vanishes in $t=0$. Therefore, since $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_t\big|_{t=0} = \dot{\mathbf{x}}(0,\mathbf{x}) = \boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}\mathbf{x} + \mathbf{r}_{\Omega,\boldsymbol{\theta}} = \boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x})$, we conclude that

$$\begin{aligned}\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) &= \frac{\mathrm{d}}{\mathrm{d}t}\Phi\left(\left(\mathbb{I}+t\boldsymbol{\theta}\right)\Omega\right)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{E}_{\mathbf{M}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega},\tau}} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{(\mathbb{I}+\mathcal{F}_t)\mathcal{E}_{\Psi_\Omega,\tau}} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\Big|_{t=0} = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} \mathrm{div}\Big(f(\mathbf{x})\,\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x})\Big)\,\mathrm{d}\mathbf{x} \\ &= \int_{\partial\mathcal{E}_{\Psi_\Omega,\tau}} f(\mathbf{s})\left(\mathbf{n}(\mathbf{s})\cdot\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{s})\right)\mathrm{d}\mathbf{s}.\end{aligned}$$

□

A first remark on the result of lemma 5.1 is that, since $\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x})$ is a linear function of $\boldsymbol{\theta}$, the expression we found is a Fréchet derivative of the functional $\Phi(\cdot)$. A second observation concerns the expression of the derivative as a surface integral on a variable ellipsoid. For numerical reasons, it might be more interesting to reformulate the integral as one on a fixed domain. Thus, we can use the volumic expression of the shape derivative to write (5.16) as an integral on the unitary N-sphere, as is done in the following proposition.

Proposition 5.2. *Under the hypotheses of lemma 5.1, the shape derivative of the functional $\Phi(\cdot)$ in Ω can be written as an integral on the unit N-sphere $\mathbb{S}_{N-1}$ in accordance with*

$$\begin{aligned}\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) =& \sqrt{\frac{\widetilde{\tau}_\Omega^N}{\det\mathbf{M}_\Omega}}\int_{\mathbb{S}_{N-1}} f\left(\sqrt{\widetilde{\tau}_\Omega}{\mathbf{M}_\Omega}^{-1/2}\mathbf{s} - {\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right) \\ &\times\left(\left(\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}{\mathbf{M}_\Omega}^{-1/2}\mathbf{s} + \frac{1}{\sqrt{\widetilde{\tau}_\Omega}}\left(\mathbf{r}_{\Omega,\boldsymbol{\theta}} - \boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}{\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right)\right)\cdot\left({\mathbf{M}_\Omega}^{1/2}\mathbf{s}\right)\right)\mathrm{d}\mathbf{s}.\end{aligned} \tag{5.19}$$

Proof. In order to prove (5.19), we consider the expression of the shape derivative given by lemma 5.1 and apply the change of variables such that $\mathbf{y} = \frac{1}{\sqrt{\widetilde{\tau}_\Omega}} \mathbf{M}_\Omega{}^{1/2} \left(\mathbf{x} + \mathbf{M}_\Omega{}^{-1} \mathbf{b}_\Omega \right)$, mapping $\mathcal{E}_{\Psi_\Omega,\tau}$ to the unit ball $\mathbb{B}_N$. We recall that, for any function $\mathbf{f} : \mathbb{R}^N \to \mathbb{R}^N$ that is $\mathcal{C}^1(\mathcal{A})$ in a given open subset $\mathcal{A}$ of $\mathbb{R}^N$, the expression of the divergence with respect to the variable $\mathbf{y}$ can be written as

$$\operatorname{div} \mathbf{f}(\mathbf{x}) = \frac{1}{\sqrt{\widetilde{\tau}_\Omega}} \operatorname{div}_{\mathbf{y}} \Big(\mathbf{M}_\Omega{}^{1/2} \, \mathbf{f} \left(\sqrt{\widetilde{\tau}_\Omega} \mathbf{M}_\Omega{}^{-1/2} \mathbf{y} - \mathbf{M}_\Omega{}^{-1} \mathbf{b}_\Omega \right) \Big).$$

By considering the expression of the displacement field $\boldsymbol{\xi}^{\boldsymbol{\theta}} : \mathbb{R}^N \to \mathbb{R}^N$ as $\boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x}) = \boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} \, \mathbf{x} + \mathbf{r}_{\Omega,\boldsymbol{\theta}}$, where $\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}$ and $\mathbf{r}_{\Omega,\boldsymbol{\theta}}$ are defined in equation (5.13) and equation (5.14), we get

$$\begin{aligned}
\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) &= \int_{\mathcal{E}_{\Psi_\Omega,\tau}} \operatorname{div} \Big(f(\mathbf{x}) \, \boldsymbol{\xi}^{\boldsymbol{\theta}}(\mathbf{x}) \Big) \, \mathrm{d}\mathbf{x} \\
&= \int_{\mathcal{E}_{\Psi_\Omega,\tau}} \operatorname{div} (f(\mathbf{x}) \, (\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} \, \mathbf{x} + \mathbf{r}_{\Omega,\boldsymbol{\theta}})) \, \mathrm{d}\mathbf{x} \\
&= \sqrt{\frac{\widetilde{\tau}_\Omega^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} \operatorname{div}_{\mathbf{y}} \Bigg(\Bigg(f \left(\sqrt{\widetilde{\tau}_\Omega} \mathbf{M}_\Omega{}^{-1/2} \mathbf{y} - \mathbf{M}_\Omega{}^{-1} \mathbf{b}_\Omega \right) \\
&\quad \times \mathbf{M}_\Omega{}^{1/2} \left(\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} \mathbf{M}_\Omega{}^{-1/2} \mathbf{y} + \frac{1}{\sqrt{\widetilde{\tau}_\Omega}} \left(\mathbf{r}_{\Omega,\boldsymbol{\theta}} - \boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}} \mathbf{M}_\Omega{}^{-1} \mathbf{b}_\Omega \right) \right) \Bigg) \Bigg) \, \mathrm{d}\mathbf{y}.
\end{aligned} \tag{5.20}$$

Observing that the normal vector on the unit sphere $\mathbb{S}_{N-1}$ in any point $\mathbf{s}$ coincides with the vector $\mathbf{s}$ itself, the expression (5.20) can be written as an integral on the sphere $\partial\mathbb{B}_N$ according to (5.19). □

It is possible to demonstrate proposition 5.2 without the use of lemma 5.1 and relying only on change of variables, integration by parts, and Jacobi's formula for the derivative of the determinant of an invertible matrix [174]. The alternative proof is reported in appendix B.

The expression of the derivative of $\Phi(\cdot)$ as found in proposition 5.2 is valid only if the random vector $\mathbf{X}$ admits a $\mathcal{C}^1$ density function $f(\cdot)$ in an open neighborhood of the ellipsoid $\mathcal{E}_{\Psi_\Omega,\tau}$. However, if the sensitivity of $\Phi(\cdot)$ is computed as part of a shape optimization procedure, such assumption should be verified for all shapes obtained during the execution of the algorithm. Therefore, it is crucial that the density $f(\cdot)$ is $\mathcal{C}^1$ in an open subset of $\mathbb{R}^N$ containing all the ellipsoids corresponding to $\Omega_0, \ldots, \Omega_{n_{\max}}$. Such condition might be unrealistic if the density f is not $\mathcal{C}^1$ on the entire space $\mathbb{R}^N$ which happens if it is compactly supported like the uniform distribution. This issue is investigated more thoroughly in section 5.3.2 for the particular case of the uniform probability distribution.

The expression (5.19) can be reformulated in order to highlight the terms depending on the argument of the shape derivative $\boldsymbol{\theta}$. We denote $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ the canonical basis of $\mathbb{R}^N$, and we consider a basis $\{\mathbf{B}^{i,j}\}_{0 \leq i \leq j \leq N}^{\infty}$ for the space of $N \times N$ symmetric matrices such that

$$\left[\mathbf{B}^{i,j}\right]_{k,\ell} = \begin{cases} \beta_{i,j}, & \text{if } k = i, \ \ell = j, \\ \beta_{i,j}, & \text{if } k = j, \ \ell = i, \\ 0, & \text{otherwise,} \end{cases} \qquad \beta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 1/\sqrt{2}, & \text{if } i \neq j. \end{cases} \tag{5.21}$$

Thus, the shape derivative of $\Phi(\cdot)$ in Ω becomes

$$\begin{aligned}\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = \sum_{1\le i\le j\le N}\Bigg(&\left(\mathbf{M}_\Omega{}^{1/2}\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}\mathbf{M}_\Omega{}^{-1/2}\right):\mathbf{B}^{i,j}\\ &\times\int_{\mathbb{S}_{N-1}}\sqrt{\frac{\widetilde{\tau}_\Omega^N}{\det\mathbf{M}_\Omega}}f\left(\sqrt{\widetilde{\tau}_\Omega}\mathbf{M}_\Omega{}^{-1/2}\mathbf{s}-\mathbf{M}_\Omega{}^{-1}\mathbf{b}_\Omega\right)s_is_j\,\mathrm{d}\mathbf{s}\Bigg)\\ +\sum_{k=1}^N\Bigg(&\left(\mathbf{r}_{\Omega,\boldsymbol{\theta}}-\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}\mathbf{M}_\Omega{}^{-1}\mathbf{b}_\Omega\right)\cdot\mathbf{e}_k\int_{\mathbb{S}_{N-1}}\sqrt{\frac{\widetilde{\tau}_\Omega^{N-1}}{\det\mathbf{M}_\Omega}}f\left(\sqrt{\widetilde{\tau}_\Omega}\mathbf{M}_\Omega{}^{-1/2}\mathbf{s}-\mathbf{M}_\Omega{}^{-1}\mathbf{b}_\Omega\right)s_k\,\mathrm{d}\mathbf{s}\Bigg).\end{aligned} \tag{5.22}$$

For problems where $\overline{\mathbf{g}}_0=\mathbf{0}$, equation equation (5.22) simplifies to

$$\begin{aligned}\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{1\le i\le j\le N}\Bigg(&\left(\mathbf{M}_\Omega{}^{-1/2}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{M}_\Omega{}^{-1/2}\right):\mathbf{B}^{i,j}\\ &\times\int_{\mathbb{S}_{N-1}}\sqrt{\frac{\widetilde{\tau}_\Omega^N}{\det\mathbf{M}_\Omega}}f\left(\sqrt{\widetilde{\tau}_\Omega}\mathbf{M}_\Omega{}^{-1/2}\mathbf{s}\right)s_is_j\,\mathrm{d}\mathbf{s}\Bigg)\end{aligned} \tag{5.23}$$

The expression (5.22) of the shape derivative of $\Phi(\Omega)$ requires the computation of all the entries of $\boldsymbol{\Xi}_{\Omega,\boldsymbol{\theta}}$ and $\mathbf{r}_{\Omega,\boldsymbol{\theta}}$ (which are functions of $\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})$, $\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta})$, and $\mathrm{D}c_\Omega(\boldsymbol{\theta})$), as well as $N(N+3)/2$ integrals on $\mathbb{S}_{N-1}$. The evaluation of said integrals can be done by applying suitable quadrature formulas on $\mathbb{S}_{N-1}$, which of course might be quite expensive if the number N of random variables is large.

5.3.2 Remarks about the uniform distribution

Let us consider the random variables $X_1,\dots,X_N$ to be independent and uniformly distributed in the interval $[-\frac{1}{2},\frac{1}{2}]$. For the sake of simplicity we suppose also that $\overline{\mathbf{g}}_0=\mathbf{0}$. If we denote $\mathbb{T}_N=[-\frac{1}{2},\frac{1}{2}]\times\ldots\times[-\frac{1}{2},\frac{1}{2}]$, the vector $\mathbf{X}$ follows the uniform distribution $\mathbf{X}\sim\mathcal{U}(\mathbb{T}_N)$. In such case, for a given admissible domain $\Omega\in\mathcal{S}_{\mathrm{adm}}$, the probability of the quantity $\mathcal{Q}(\mathbf{u}_{\Omega,\mathbf{g}}(\cdot),\Omega)$ to be below the threshold τ has a simple expression:

$$\Phi(\Omega)=\mathbb{P}\left[\mathcal{Q}(\mathbf{u}_{\Omega,\mathbf{g}},\Omega)\le\tau\right]=\frac{\mathrm{Vol}\left(\mathcal{E}_{\mathbf{M}_\Omega,\tau}\cap\mathbb{T}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)}. \tag{5.24}$$

Equation (5.24) has an explicit expression if the ellipsoid $\mathcal{E}_{\mathbf{M}_\Omega,\tau}$ is included into the box $\mathbb{T}_N$. Indeed, the volume of an N-dimensional ellipsoid $\mathcal{E}$ with semiaxes $a_1,\dots,a_N$ is:

$$\mathrm{Vol}\left(\mathcal{E}\right)=\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)}\prod_{i=1}^N a_i. \tag{5.25}$$

Given the expression (5.9) for the length of the semiaxes of $\mathcal{E}_{\mathbf{M}_\Omega,\tau}$, in the cases where $\mathcal{E}_{\mathbf{M}_\Omega,\tau}\subset\mathbb{T}_N$, the quantity $\Phi(\Omega)$ can be expressed as:

$$\Phi(\Omega)=\frac{\mathrm{Vol}\left(\mathcal{E}_{\mathbf{M}_\Omega,\tau}\cap\mathbb{T}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)}=\frac{\mathrm{Vol}\left(\mathcal{E}_{\mathbf{M}_\Omega,\tau}\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)}=\frac{\pi^{N/2}\,\tau^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)}\left(\det\mathbf{M}_\Omega\right)^{-1/2}. \tag{5.26}$$

We suppose that all the components of the $N \times N$ matrix $\mathbf{M}_\Omega$ are differentiable with respect to the domain Ω, and we denote $\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})$ the $N \times N$ matrix such that:

$$[\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})]_{ij} = \mathrm{D}\langle \mathbf{u}_\Omega i, Q\, \mathbf{u}_\Omega j\rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}) \qquad \text{for all } i,j \in \{1,\dots,N\}.$$

By applying the expression (5.22), the shape derivative $\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta})$ can be written as:

$$\begin{aligned} \mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) &= -\frac{1}{2}\left(\frac{\pi^{N/2}\,\tau^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)}\right)(\det \mathbf{M}_\Omega)^{-1/2}\,\mathrm{tr}\left(\mathbf{M}_\Omega{}^{-1}\,\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right) \\ &= -\frac{1}{2}\Phi(\Omega)\,\mathrm{tr}\left(\mathbf{M}_\Omega{}^{-1}\,\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right). \end{aligned} \tag{5.27}$$

Unfortunately, the relatively simple expression (5.27) applies only when the hypothesis $\mathcal{E}_{\mathbf{M}_\Omega,\tau} \subset \mathbb{T}_N$ is verified. Let $\widehat{\mathbb{B}}_N$ be the N-dimensional ball of radius 0.5 centered in $\mathbf{0}$, and ν_N the ratio between the volumes of the ball $\widehat{\mathbb{B}}_N$ and the box $\mathbb{T}_N$:

$$\nu_N = \frac{\mathrm{Vol}\left(\widehat{\mathbb{B}}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)} = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)}\left(\frac{1}{2}\right)^N.$$

Since $\widehat{\mathbb{B}}_N$ is the biggest ellipsoid included in $\mathbb{T}_N$, a necessary, but not sufficient, condition for the hypothesis $\mathcal{E}_{\mathbf{M}_\Omega,\tau} \subset \mathbb{T}_N$ is that

$$\Phi(\Omega) = \frac{\mathrm{Vol}\left(\mathcal{E}_{\mathbf{M}_\Omega,\tau} \cap \mathbb{T}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)} = \frac{\mathrm{Vol}\left(\mathcal{E}_{\mathbf{M}_\Omega,\tau}\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)} \leq \frac{\mathrm{Vol}\left(\widehat{\mathbb{B}}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)} = \nu_N.$$

Therefore, in order to be able to use the expression (5.27), the threshold $\bar{p}$ introduced in problem (5.3) must satisfy the following inequality:

$$\bar{p} \geq \mathbb{P}\left[\mathcal{Q}\left(\mathbf{u}_{\Omega,\mathbf{g}}(\cdot),\Omega\right) > \tau\right] = 1 - \Phi(\Omega) \geq 1 - \nu_N. \tag{5.28}$$

The solution of problem (5.3) is significant when the threshold $\bar{p}$ is sufficiently small (around 10^{-1}, or 10^{-2}, or less). Unfortunately, such problems are not compatible with the inequality (5.28), especially when the number N of random components is bigger than 4, as shown in table 5.1.

N	$\mathrm{Vol}\left(\widehat{\mathbb{B}}_N\right)$	$= \frac{\mathrm{Vol}\left(\widehat{\mathbb{B}}_N\right)}{\mathrm{Vol}\left(\mathbb{T}_N\right)} = \nu_N$	$\mathbf{1 - \nu_N}$
2	$3.1416 \times (0.5)^2$	$= 0.7853$	**0.2146**
3	$4.1888 \times (0.5)^3$	$= 0.5236$	**0.4764**
4	$4.9348 \times (0.5)^4$	$= 0.3084$	**0.6916**
5	$5.2638 \times (0.5)^5$	$= 0.1645$	**0.8355**
6	$5.1677 \times (0.5)^6$	$= 0.0807$	**0.9193**
8	$4.0587 \times (0.5)^8$	$= 0.0159$	**0.9841**
10	$2.5502 \times (0.5)^{10}$	$= 0.0025$	**0.9975**

Table 5.1: Ratio of the volume of the unit sphere $\widehat{\mathbb{B}}_N$ and the volume of the box $\mathbb{T}_N$, and maximum value of $\bar{p}$ for equation (5.26) and equation (5.27) to be valid, as function of the number N of random variables.

5.3.3 Shape derivative for centered Gaussian perturbations

In this section we aim to apply the formula proposed in (5.22) to the case of centered Gaussian mechanical loads. In particular, in order to avoid the numerical computation of the integrals on the N-ball and N-sphere that are necessary to compute $\Phi(\Omega)$ and its shape derivative, we apply the quadrature formulas presented in appendix C.

Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$ be a standard Gaussian multivariate random variable with N components, and $\mathbf{g} \in \mathrm{L}^2\left(\mathcal{O}, \mathbb{P}; \mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})^d\right)$ be a centered random load such that, for almost any event $\omega \in \mathcal{O}$,

$$\mathbf{g}(\omega) = \overline{\mathbf{g}}_1 X_1(\omega) + \ldots + \overline{\mathbf{g}}_N X_N(\omega). \tag{5.29}$$

The random variable $\mathbf{X}$ has the following probability density function

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}}. \tag{5.30}$$

By equation (5.11) and by the change of variable $\mathbf{x} = \frac{1}{\sqrt{\tau}}{\mathbf{M}_\Omega}^{1/2}\mathbf{y}$, we express the functional $\Phi(\Omega) = \mathbb{P}[\mathbf{X} \in \mathcal{E}_{\Psi_\Omega,\tau}]$ as

$$\Phi(\Omega) = \int_{\mathcal{E}_{\Psi_\Omega,\tau}} f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{\tau^{N/2}}{(2\pi)^{N/2}\sqrt{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} \exp\left(-\frac{\tau}{2}\mathbf{y}{\mathbf{M}_\Omega}^{-1}\mathbf{y}\right)\mathrm{d}\mathbf{x}. \tag{5.31}$$

The expression (5.31) can be formulated using the results of appendix C.

Proposition 5.3. *Under the hypotheses of section 5.3.3, the quantity $\Phi(\Omega)$ can be computed as*

$$\Phi(\Omega) = \frac{(\lambda_N)^{N/2}}{2\,(\pi)^{N/2}\sqrt{\det \mathbf{M}_\Omega}} \sum_{k=0}^{\infty} \frac{1}{k!}\gamma\left(\frac{N}{2}+k, \frac{\tau}{2\lambda_N}\right) B_k, \tag{5.32}$$

where λ_N is the smallest eigenvalue of $\mathbf{M}_\Omega$, $\gamma(\cdot,\cdot)$ denotes the lower incomplete gamma function, and the terms $\{B_k\}_{k=0}^{\infty}$ are found by the following recursive relation

$$\begin{cases} B_k = \sum_{j=2}^{N}\left(1-\frac{\lambda_j}{\lambda_1}\right)A_k^j, \\ A_k^1 = \frac{2\pi^{N/2}}{N(N+2k)\,\Gamma\left(\frac{N}{2}\right)}, \\ A_k^i = \frac{2\pi^{N/2}}{N(N+2k)\,\Gamma\left(\frac{N}{2}\right)} + \frac{2k}{N+2k}\left(1-\frac{\lambda_N}{\lambda_i}\right)A_{k-1}^i \qquad \text{for } 2 \le i \le N. \end{cases} \tag{5.33}$$

The recursive relation (5.33) *is initialized by:*

$$\begin{cases} B_0 = |\mathbb{S}_{N-1}| = \frac{2\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}, \\ A_0^i = \frac{|\mathbb{S}_{N-1}|}{N} = \frac{2\pi^{N/2}}{N\,\Gamma\left(\frac{N}{2}\right)} \qquad \text{for } 1 \le i \le N. \end{cases} \tag{5.34}$$

Proof. We consider the expression (5.31) for $\Phi(\Omega)$, and we focus on the integral term. Thanks to the spherical symmetry of $\mathbb{B}_N$, we can diagonalize $\mathbf{M}_\Omega$ and replace the integral with

$$\int_{\mathbb{B}_N} \exp\left(-\frac{\tau}{2}\mathbf{y}{\mathbf{M}_\Omega}^{-1}\mathbf{y}\right)\mathrm{d}\mathbf{x} = \int_{\mathbb{B}_N} \exp\left(-\mathbf{y}^{\mathrm{T}}\widetilde{\mathbf{D}_\Omega}\mathbf{y}\right)\mathrm{d}\mathbf{x}, \tag{5.35}$$

where $\widetilde{\mathbf{D}_\Omega} = \operatorname{diag}\left\{\frac{\tau}{2\lambda_N}, \ldots, \frac{\tau}{2\lambda_N}\right\}$, and $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of $\mathbf{M}_\Omega$ in decreasing order. The thesis follows from the expansion of the integral (5.35) into a series using proposition C.21, where the terms $\{A_k^i\}_{k=0}^\infty$ and $\{B_k\}_{k=0}^\infty$ denote, for all $i \in \{1, \ldots, N\}$ and $k \geq 0$,

$$A_k^i = \int_{\mathbb{S}_{N-1}} \left(\mathbf{s}^{\mathrm{T}}(\mathbb{I} - \widetilde{\mathbf{D}_\Omega})\mathbf{s}\right)^k s_i^2 \, \mathrm{d}\mathbf{s} \qquad \text{and} \qquad B_k = \int_{\mathbb{S}_{N-1}} \left(\mathbf{s}^{\mathrm{T}}(\mathbb{I} - \widetilde{\mathbf{D}_\Omega})\mathbf{s}\right)^k \mathrm{d}\mathbf{s}. \tag{5.36}$$

□

A similar expression can be obtained for the shape derivative, as shown by the following result.

Proposition 5.4. *Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a Lipschitz continuous domain in $\mathbb{R}^d$. For all $i \in \{1, \ldots, N\}$ we define the load $\widetilde{\mathbf{g}_i} \in \mathrm{H}^{-1/2}(\Gamma_{\mathrm{N}})^d$ as*

$$\widetilde{\mathbf{g}_i} = \overline{\mathbf{g}}_1 {\mathbf{v}^{i,\Omega}}_1 + \ldots + \overline{\mathbf{g}}_N {\mathbf{v}^{i,\Omega}}_N,$$

where $\mathbf{v}^{i,\Omega}$ is the eigenvector of $\mathbf{M}_\Omega$ corresponding to the eigenvalue λ_i, and ${\mathbf{v}^{i,\Omega}}_j$ its j-th component. Moreover, we denote $\widetilde{\mathbf{u}_{\Omega,i}}$ the displacement obtained by solving the state equation (5.2) for the load $\widetilde{\mathbf{g}_i}$. Under the same hypotheses and notations of proposition 5.3, the shape derivative of $\Phi(\cdot)$ evaluated in $\Omega \in \mathcal{S}_{\mathrm{adm}}$ can be expressed as

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \sum_{i=1}^N \left(\frac{e^{-\frac{\tau}{2\lambda_N}}}{\lambda_i} \mathrm{D}\langle \widetilde{\mathbf{u}_{\Omega,i}}, Q\, \widetilde{\mathbf{u}_{\Omega,i}} \rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}) \sum_{k=0}^\infty \frac{A_k^i}{k!} \left(\frac{\tau}{2\lambda_N}\right)^k \right). \tag{5.37}$$

Proof. We start by considering the expression (5.19) for the case where $\mathbf{X}$ is a centered Gaussian random vector.

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \int_{\mathbb{S}_{N-1}} e^{-\frac{\tau}{2}\mathbf{s}^{\mathrm{T}}\mathbf{M}_\Omega\mathbf{s}} \left(\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta})\mathbf{M}_\Omega{}^{-1/2}\mathbf{s}\right) \cdot \left(\mathbf{M}_\Omega{}^{-1/2}\mathbf{s}\right) \mathrm{d}\mathbf{s}.$$

We diagonalize the matrix $\mathbf{M}_\Omega$ as $\mathbf{M}_\Omega = \mathbf{U}_\Omega \mathbf{D}_\Omega \mathbf{U}_\Omega{}^{\mathrm{T}}$ and apply the change of variables $\mathbf{s} \mapsto \widetilde{\mathbf{s}} = \mathbf{U}_\Omega{}^{\mathrm{T}}\mathbf{s}$. Since $\mathbf{U}_\Omega$ is an orthogonal matrix, the domain of integration after the change of variable stays the same. As in proposition 5.3, we denote $\widetilde{\mathbf{D}_\Omega} = \frac{\tau}{2}\mathbf{D}_\Omega = \operatorname{diag}\left\{\frac{\tau}{2\lambda_N}, \ldots, \frac{\tau}{2\lambda_N}\right\}$

$$\begin{aligned} \mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) &= -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \int_{\mathbb{S}_{N-1}} e^{-\frac{\tau}{2}\mathbf{s}^{\mathrm{T}}\mathbf{U}_\Omega\mathbf{D}_\Omega\mathbf{U}_\Omega{}^{\mathrm{T}}\mathbf{s}} \\ &\quad \times \left(\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\mathbf{D}_\Omega{}^{-1/2}\mathbf{U}_\Omega{}^{\mathrm{T}}\mathbf{s}\right) \cdot \left(\mathbf{U}_\Omega\mathbf{D}_\Omega{}^{-1/2}\mathbf{U}_\Omega{}^{\mathrm{T}}\mathbf{s}\right) \mathrm{d}\mathbf{s} \\ &= -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \int_{\mathbb{S}_{N-1}} e^{-\widetilde{\mathbf{s}}^{\mathrm{T}}\widetilde{\mathbf{D}_\Omega}\widetilde{\mathbf{s}}} \left(\left(\mathbf{U}_\Omega{}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right)\mathbf{D}_\Omega{}^{-1/2}\widetilde{\mathbf{s}}\right) \cdot \left(\mathbf{D}_\Omega{}^{-1/2}\widetilde{\mathbf{s}}\right) \mathrm{d}\widetilde{\mathbf{s}}. \end{aligned} \tag{5.38}$$

We decompose the symmetric matrix $\left(\mathbf{U}_\Omega{}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right)$ on the basis of symmetric matrices $\{\mathbf{B}^{i,j}\}_{1\leq i\leq j\leq N}^\infty$ introduced in (5.21)

$$\mathbf{U}_\Omega{}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega = \sum_{1\leq i\leq j\leq N} \left(\left(\mathbf{U}_\Omega{}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right) : \mathbf{B}^{i,j}\right)\mathbf{B}^{i,j}.$$

Therefore, equation (5.38) becomes

$$\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \sum_{1\leq i\leq j\leq N} \left(\left({\mathbf{U}_\Omega}^{\mathrm{T}} \mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right) : \mathbf{B}^{i,j}\right) \int_{\mathbb{S}_{N-1}} \frac{e^{-\widetilde{\mathbf{s}}^{\mathrm{T}}\widetilde{\mathbf{D}_\Omega}\widetilde{\mathbf{s}}}\widetilde{s}_i\,\widetilde{s}_j}{\sqrt{\lambda_i\lambda_j}}\,\mathrm{d}\widetilde{\mathbf{s}}. \quad (5.39)$$

In order to conclude the proof we remark three properties of the different terms in (5.39). First, we consider the quadratic nature of the mapping $\mathbf{u}_\Omega \mapsto \langle \mathbf{u}_\Omega, Q\,\mathbf{u}_\Omega\rangle_{\mathrm{H}^1(\Omega)}$, the linearity of $\mathbf{g} \mapsto \mathbf{u}_\Omega$, and the definition of the matrix $\mathbf{M}_\Omega$. For any $\mathbf{b}, \mathbf{c} \in \mathbb{R}^N$ we denote $\widetilde{\mathbf{g}^{\mathbf{b}}}, \widetilde{\mathbf{g}^{\mathbf{b}}} \in \mathrm{H}^{-1/2}\left(\Gamma_{\mathrm{N}}\right)^d$ the loads

$$\widetilde{\mathbf{g}^{\mathbf{b}}} = \overline{\mathbf{g}}_1 b_1 + \ldots + \overline{\mathbf{g}}_N b_N, \qquad \textit{and} \qquad \widetilde{\mathbf{g}^{\mathbf{c}}} = \overline{\mathbf{g}}_1 c_1 + \ldots + \overline{\mathbf{g}}_N c_N,$$

and $\widetilde{\mathbf{u}_\Omega{}^{\mathbf{b}}}, \widetilde{\mathbf{u}_\Omega{}^{\mathbf{c}}} \in \mathrm{H}^1\left(\Omega\right)^d$ the displacements obtained by solving the state equation (5.2) for $\widetilde{\mathbf{g}^{\mathbf{b}}}$ and $\widetilde{\mathbf{g}^{\mathbf{c}}}$ respectively. Then

$$\mathbf{b}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{c} = \mathrm{D}\left\langle \widetilde{\mathbf{u}_\Omega{}^{\mathbf{b}}}, Q\,\widetilde{\mathbf{u}_\Omega{}^{\mathbf{c}}}\right\rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}).$$

Therefore

$$\left[{\mathbf{U}_\Omega}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right]_{ij} = \mathrm{D}\langle \widetilde{\mathbf{u}_{\Omega,i}}, Q\,\widetilde{\mathbf{u}_{\Omega,j}}\rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}).$$

A second results concerns the terms inside the integral in equation (5.39). Indeed, thanks to lemma C.18 all terms where $i \neq j$ are equal to zero, and only the N terms of the sum for which $i = j$ remain

$$\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det \mathbf{M}_\Omega}} \sum_{i=1}^{N} \mathrm{D}\langle \widetilde{\mathbf{u}_{\Omega,i}}, Q\,\widetilde{\mathbf{u}_{\Omega,i}}\rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}) \int_{\mathbb{S}_{N-1}} e^{-\widetilde{\mathbf{s}}^{\mathrm{T}}\widetilde{\mathbf{D}_\Omega}\widetilde{\mathbf{s}}}\frac{\widetilde{s}_i^2}{\lambda_i}\,\mathrm{d}\widetilde{\mathbf{s}}. \quad (5.40)$$

Finally, proposition C.17 allows to express the integrals in (5.40) in the form of an infinite series as

$$\int_{\mathbb{S}_{N-1}} e^{-\widetilde{\mathbf{s}}^{\mathrm{T}}\widetilde{\mathbf{D}_\Omega}\widetilde{\mathbf{s}}}\frac{\widetilde{s}_i^2}{\lambda_i}\,\mathrm{d}\widetilde{\mathbf{s}} = \sum_{k=0}^{\infty} A_k^i, \quad (5.41)$$

for any $i \in \{1, \ldots, N\}$. The terms $\{A_k^i\}_{k=0}^{\infty}$ are introduced in (5.36) and follow the recursive relation (5.33) - (5.34). In conclusion, we obtain that the shape derivative of $\Phi\left(\Omega\right)$ evaluated $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\Omega\right)^d$ can be expressed as

$$\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det\,\mathbf{M}_\Omega}} \sum_{i=1}^{N} \left(\frac{e^{-\frac{\tau}{2\lambda_N}}}{\lambda_i} \mathrm{D}\langle \widetilde{\mathbf{u}_{\Omega,i}}, Q\,\widetilde{\mathbf{u}_{\Omega,i}}\rangle_{\mathrm{H}^1(\Omega)}(\boldsymbol{\theta}) \sum_{k=0}^{\infty} \frac{A_k^i}{k!} \left(\frac{\tau}{2\lambda_N}\right)^k \right).$$

□

5.4 The generalized noncentral chi-squared distribution

5.4.1 Series expansion of the cumulative distribution function

Let $\mathbf{X} \sim \mathcal{N}\left(\boldsymbol{\mu}, \Sigma\right)$ be a Gaussian random vector with N components, mean $\boldsymbol{\mu}$ and covariance matrix Σ, and let $\mathbf{D} = \mathrm{diag}\left\{\lambda_1, \ldots, \lambda_N\right\}$ be a positive definite diagonal matrix. Let T be the random variable defined as follows:

$$T = \mathbf{X}^{\mathrm{T}}\,\mathbf{D}\,\mathbf{X} = \lambda_1 X_1^2 + \ldots + \lambda_N X_N^2. \quad (5.42)$$

Without loss of generality, we suppose that the covariance matrix of the Gaussian random vector $\mathbf{X}$ is the identity matrix: $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbb{I})$. In such case, each random variable X_i^2 follows a noncentral chi-squared distribution with one degree of freedom and non-centrality parameter μ_i^2. The random variable T is said to follow a generalized non-central chi-squared distribution

$$T \sim \widetilde{\chi^2}(\mathbf{1}; \boldsymbol{\mu} \odot \boldsymbol{\mu}; \boldsymbol{\lambda}), \tag{5.43}$$

where $\mathbf{1} = [1, \ldots, 1]$ is the vector of the degrees of freedom, $\boldsymbol{\mu} \odot \boldsymbol{\mu} = [\mu_1^2, \ldots, \mu_N^2]$ is the vector of noncentrality parameters (the symbol "$\odot$" represent the elementwise product), and $\boldsymbol{\lambda} = \operatorname{diag}\{\mathbf{D}\}$ is the vector of the weights of the random variables $X_1, \ldots, X_N$.

The characterization of the cumulative distribution function F_T of the random variable T has been studied analytically in [208, 214]. The results of these articles have led to the development of several algorithms for the numerical computation of the quantiles of T. Sequential methods that provide an estimate for the truncation error include the algorithms developed by Imhof [142], Davies [95, 96], and Farebrother [111], who refined the result obtained by Sheil and O'Muircheartaigh in [231]. If the number N of random variables is large, faster but less accurate approximations should be considered. Among such techniques we mention Kuonen's method [159], which is based on a saddlepoint approximation of the distribution of T, the approach based on the leading eigenvalues developed by Lumley et al. in [171], and the several approaches based on the computation of the stochastic moments of the random variable T like the methods deveolped by Liu–Tang–Zhang [166], Satterthwaite–Welch [219], Hall-Buckley–Eagleson [132, 55], and Lindsay–Pilla–Basak [164]. Further information on the comparison between the different methods can be found in [105, 52, 66].

In this section, we present the results of Ruben [214], where, for any threshold $\tau > 0$, the quantity $F_T(\tau)$ is expressed in terms of a series of cumulative distribution functions of centered chi-squared random variables [214, Theorem 1]. The coefficients of the decomposition are defined by a recurrence relation. Moreover, an upper bound on the truncation error of the series is provided.

Theorem 5.5 (Decomposition of $F_T(\tau)$ by chi-squared random variables)**.** *Let T be a real-valued random variable defined as in* (5.42)*. Then, for any choice of $\beta > 0$, the quantity $F_T(\tau) = \mathbb{P}[T \leq \tau]$ can be expressed as*

$$F_T(\tau) = \sum_{k=0}^{\infty} \gamma_k F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right). \tag{5.44}$$

The weights $\{\gamma_k\}_{k=0}^{\infty}$ are computed by using the recurrence relation

$$\gamma_0 = e^{-\frac{1}{2}\|\boldsymbol{\mu}\|^2} \beta^{N/2} \det(\mathbf{D})^{-1/2} \quad \text{and} \quad \gamma_k = \frac{1}{2k} \sum_{\ell=0}^{k-1} g_{k-\ell}\gamma_\ell \text{ for } k \geq 1, \tag{5.45}$$

where the coefficients $\{g_k\}_{k=1}^{\infty}$ are defined in accordance with

$$g_k = \sum_{i=1}^{N} \left(1 - \frac{\beta}{\lambda_i}\right)^{k-1} \left(1 + (k\mu_i^2 - 1)\frac{\beta}{\lambda_i}\right). \tag{5.46}$$

In particular, if $0 < \beta < \min_{i \in \{1,\ldots,N\}} \{\lambda_1, \ldots, \lambda_N\}$, the series (5.44) *is a mixture representation, meaning that all coefficients γ_k are non-negative and $\sum_{k=0}^{\infty} \gamma_k = 1$.*

This result is stated and proven in [214, Theorem 1], while the condition of the mixture representation is stated in [214, Section 5]. Note that [214] provides also an explicit expression for the coefficients $\{\gamma_k\}_{k=0}^{\infty}$ which can be used to prove the uniform convergence of the series (5.44) for any choice of $\beta > 0$ and for any finite value of the threshold $0 \leq \tau < \infty$. Especially, analogous results apply also to the probability density function of T.

Corollary 5.6. *If $0 < \beta < \min_{i\in\{1,\dots,N\}} \{\lambda_1, \dots, \lambda_N\}$, for any $\tau > 0$, the following expression for the probability density function of T holds:*

$$f_T(\tau) = \sum_{k=0}^{\infty} \gamma_k f_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right).$$

If the mixture representation holds (that is if $0 < \beta < \min\{\lambda_1, \dots, \lambda_N\}$), it is possible to establish the following upper bound on the truncation error of the series (5.44).

Proposition 5.7. *If $0 < \beta < \min\{\lambda_1, \dots, \lambda_N\}$ and the hypotheses of theorem 5.5 hold, then*

$$\left| F_T(\tau) - \sum_{k=0}^{n} \gamma_k F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right) \right| \leq \left(1 - \sum_{k=0}^{n} \gamma_k\right) F_{\chi^2(2n+2+N)}\left(\frac{\tau}{\beta}\right) \tag{5.47}$$

for all $0 < \tau < \infty$ and any positive integer n.

Proof. One readily verifies that $F_{\chi^2(m)}(\tau) < F_{\chi^2(n)}(\tau)$ for any pair of integers $m > n$ and any $\tau > 0$ fixed. Therefore, the sequence $\left\{F_{\chi^2(2k+N+2)}\left(\frac{\tau}{\beta}\right)\right\}_{k=0}^{\infty}$ is decreasing whenever τ/β is fixed. Thus, we conclude

$$\begin{aligned}\left| F_T(\tau) - \sum_{k=0}^{n} \gamma_k F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right) \right| &= \left| \sum_{k=n+1}^{\infty} \gamma_k F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right) \right| \\ \leq F_{\chi^2(2n+N+2)}\left(\frac{\tau}{\beta}\right) \sum_{k=n+1}^{\infty} \gamma_k &= \left(1 - \sum_{k=0}^{n} \gamma_k\right) F_{\chi^2(2n+2+N)}\left(\frac{\tau}{\beta}\right).\end{aligned}$$

□

5.4.2 Differentiating the probability to exceed a threshold

Let τ be a positive constant, and let us consider the following mappings:

- $\mathbf{M} : [0, \delta] \to \mathbb{R}^{N\times N}$ associating to any $t \in [0, \delta]$ a positive definite symmetric matrix;
- $\mathbf{b} : [0, \delta] \to \mathbb{R}^N$;
- $c : [0, \delta] \to \mathbb{R}$.

We assume that these three functions are all $\mathcal{C}^1$, and we denote by Ψ_t the quadratic form defined on $\mathbb{R}^N$ given by

$$\Psi_t : \mathbf{x} \mapsto \mathbf{x}^{\mathrm{T}}\mathbf{M}(t)\mathbf{x} + 2^{\mathrm{T}}\mathbf{b}(t)\mathbf{x} + c(t). \tag{5.48}$$

We suppose that $\Psi_t(\mathbf{x}) > 0$ and that $\tau > c(t) - \mathbf{b}^{\mathrm{T}}(t)\mathbf{M}^{-1}(t)\mathbf{b}(t) = \Psi_t\left(-\mathbf{M}^{-1}(t)\mathbf{b}(t)\right)$ holds for all $t \in [0, \delta]$ and $\mathbf{x} \in \mathbb{R}^N$.

Let $\mathbf{X} \sim \mathcal{N}(\mathbf{h}, \mathbb{I})$ be a Gaussian random vector, where $\mathbf{h} \in \mathbb{R}^N$ is constant and $\mathbb{I}$ is the $N \times N$ identity matrix. We are interested in differentiating the cumulative distribution function of the random variable $\Psi_t(\mathbf{X})$ with respect to the parameter t. In order to do so, we prove the following lemma about the derivative of the cumulative distribution of a generalized $\widetilde{\chi^2}$ random variable.

Lemma 5.8. *Let us consider two $\mathcal{C}^1$ vector-valued functions $\boldsymbol{\mu}, \boldsymbol{\lambda} : [0, \delta] \to \mathbb{R}^N$ such that, for all $t \in [0, \delta]$, all components of $\boldsymbol{\lambda}(t)$ are strictly larger than a positive constant β independent from t. For all $t \in [0, \delta]$, let $T(t)$ be a random variable with the following generalized chi-squared distribution:*

$$T(t) \sim \widetilde{\chi^2}\left(\mathbf{1}; \boldsymbol{\mu}(t) \odot \boldsymbol{\mu}(t); \boldsymbol{\lambda}(t)\right), \tag{5.49}$$

Due to theorem 5.5, its cumulative distribution function evaluated in τ can be expressed as

$$F_{T(t)}(\tau) = \sum_{k=0}^{\infty} \gamma_k(t) F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right). \tag{5.50}$$

Then, the coefficients $\gamma_k(t)$ of the respective cumulative distribution function (5.44)evaluated in τ are differentiable with respect to t for all $t \in [0, \delta]$ and all $k \in \mathbb{N}$, and their derivative is

$$\gamma_k'(t) = \boldsymbol{\lambda}'(t) \cdot \mathbf{p}^k + \boldsymbol{\mu}'(t) \cdot \mathbf{q}^k.$$

Herein, the terms $\mathbf{p}^k = [p_1^k, \ldots, p_N^k]^{\mathrm{T}}$ and $\mathbf{q}^k = [q_1^k, \ldots, q_N^k]^{\mathrm{T}}$, and $\mathbf{D}^j$ are defined as follows for any $j \in \{1, \ldots, N\}$ and $k \geq 0$:

- $p_j^0 = -\frac{\gamma_0}{2\lambda_j}$ *and* $p_j^k = \frac{1}{2k} \sum_{\ell=0}^{k-1} \left(\nu_j^{k-\ell} \gamma_\ell + p_j^\ell g_{k-\ell}\right)$ *for $k \geq 1$;*
- $q_j^0 = 0$ *and* $q_j^k = \frac{1}{2k} \sum_{\ell=0}^{k-1} \left(\kappa_j^{k-\ell} \gamma_\ell + q_j^\ell g_{k-\ell}\right)$ *for $k \geq 1$;*
- $\nu_j^1 = \frac{\beta}{\lambda_j^2}(1 - \mu_j^2)$ *and* $\nu_j^k = \frac{\beta}{\lambda_j^2}(1 - \frac{\beta}{\lambda_j})^{k-2}\left[(k-1)(1 + \frac{\beta}{\lambda_j}(k\mu_j^2 - 1)) + (1 - \frac{\beta}{\lambda_j})(1 - k\mu_j^2)\right]$ *for $k \geq 1$;*
- $\kappa_j^k = 2k\mu_j \frac{\beta}{\lambda_j}(1 - \frac{\beta}{\lambda_j})^{k-1}$ *for $k \geq 1$.*

Proof. According to theorem 5.5, the coefficients γ_k are defined as in (5.45), where the coefficients g_k are given by:

$$g_k = \sum_{j=1}^{N} \left(1 - \frac{\beta}{\lambda_j}\right)^{k-1} \left(1 + (k\mu_j(t)^2 - 1)\frac{\beta}{\lambda_j(t)}\right). \tag{5.51}$$

Differentiating equation (5.51), we obtain

$$g_1'(t) = \sum_{j=1}^{N} \left(2\frac{h_i\beta}{\lambda_j}\mu_j'(t) - (h_j^2 - 1)\frac{\beta}{\lambda_j^2}\lambda_j'(t)\right) = \sum_{j=1}^{N} \left(\kappa_j^1 \mu_j'(t) + \nu_j^1 \lambda_j'(t)\right)$$

and for $k > 1$

$$\begin{aligned} g_k'(t) = \sum_{j=1}^{N} \Bigg[&\left(1 - \frac{\beta}{\lambda_j}\right)^{k-2} \left((k-1)\frac{\beta}{\lambda_j^2}\left(1 + (k\mu_j^2 - 1)\frac{\beta}{\lambda_j}\right)\right)\lambda_j'(t) \\ &+ \left(1 - \frac{\beta}{\lambda_j}\right)\left(2k\frac{\mu_j\beta}{\lambda_j}\mu_j'(t) - \left((k\mu_j^2 - 1)\frac{\beta}{\lambda_j(t)^2}\right)\lambda_j'(t)\right)\Bigg] = \sum_{j=1}^{N} \left(\kappa_j^k \mu_j'(t) + \nu_j^k \lambda_j'(t)\right). \end{aligned}$$

The assertion follows by differentiating the definitions of γ_k, found in (5.45), and using the expression above for the derivatives of g_k. □

Proposition 5.9. *Let $\Psi_t : \mathbb{R}^N \to \mathbb{R}$ be defined as in (5.48) for $t \in [0,\delta]$, let $\mathbf{X} \sim \mathcal{N}(\mathbf{h}, \mathbb{I})$ be a Gaussian vector, and let τ be a positive constant. We assume that $\tau > c(t) - \mathbf{b}^{\mathrm{T}}(t)\mathbf{M}^{-1}(t)\mathbf{b}(t)$ for all $t \in [0,\delta]$, and that all eigenvalues of $\mathbf{M}(t)$ $\lambda_1(t), \dots, \lambda_N(t)$ are pairwise distinct and larger than a strictly positive constant $\beta > 0$. We introduce the following notation:*

- *$\mathbf{Y}(t) \in \mathrm{L}^2(\mathcal{O}, \mathbb{P})^N$ is the random variable defined as $\mathbf{Y}(t) = \mathbf{X} + \mathbf{M}^{-1}(t)\mathbf{b}(t)$, so that its law is $\mathbf{Y}(t) \sim \mathcal{N}(\mathbf{h} + \mathbf{M}^{-1}(t)\mathbf{b}(t), \mathbb{I})$;*
- *for all $t \in [0,\delta]$, we denote $T(t)$ the random variable $T(t) = \mathbf{Y}^{\mathrm{T}}(t)\mathbf{M}(t)\mathbf{Y}(t)$;*
- *$\widetilde{\tau} : [0,\delta] \to \mathbb{R}$ mapping $t \mapsto \tau - c(t) + \mathbf{b}(t)^{\mathrm{T}}\mathbf{M}^{-1}(t) + \mathbf{b}(t)$;*
- *$\mathbf{M}(t)$ is diagonalized as $\mathbf{M}(t) = \mathbf{Q}(t)\mathbf{D}(t)\mathbf{Q}^{\mathrm{T}}(t)$, where $\mathbf{Q}(t) = [\mathbf{v}^1|, \dots, |\mathbf{v}^N]$ is an orthogonal matrix, and $\mathbf{D}(t) = \mathrm{diag}\{\boldsymbol{\lambda}(t)\} = \mathrm{diag}\{\lambda_1(t), \dots, \lambda_N(t)\}$;*
- *$\boldsymbol{\mu} : [0,\delta] \to \mathbb{R}^N$ such that $\boldsymbol{\mu}(t) = \mathbf{Q}^{\mathrm{T}}(t)\mathbf{h} + \mathbf{Q}^{\mathrm{T}}(t)\mathbf{M}^{-1}(t)\mathbf{b}(t)$.*

Then, for any $t \in [0,\delta]$, $\mathbf{Y}(t)$ is a normalized Gaussian random variable centered in $\boldsymbol{\mu}(t)$, and $T(t)$ has the following chi-squared distribution:

$$T(t) \sim \widetilde{\chi^2}(\mathbf{1}; \boldsymbol{\mu}(t) \odot \boldsymbol{\mu}(t); \boldsymbol{\lambda}). \tag{5.52}$$

Moreover, for all $t \in [0,\delta]$, the following identity between the values of the cumulative distribution functions of $\Psi_t(\mathbf{X})$ and $T(t)$ holds:

$$F_{\Psi_t(\mathbf{X})}(\tau) = F_{T(t)}(\widetilde{\tau}(t)). \tag{5.53}$$

Finally, the mapping $t \mapsto F_{\Psi_t(\mathbf{X})}(\tau)$ is differentiable and its derivative can be written as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}F_{\Psi_t(\mathbf{X})}(\tau) = &\left(\sum_{k=0}^{\infty} \mathbf{p}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right) \cdot \boldsymbol{\lambda}'(t) \\ &+ \left(\sum_{k=0}^{\infty} \mathbf{q}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right) \cdot \boldsymbol{\mu}'(t) + \frac{1}{\beta}\left(\sum_{k=0}^{\infty} \gamma_k f_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right)\widetilde{\tau}'(t).\end{aligned} \tag{5.54}$$

Here, for all $n \in \mathbb{N}$, $f_{\chi^2(n)}$ is the density of a chi-squared random variable with n degrees of freedom. The components of $\mathbf{p}^k$ and $\mathbf{q}^k$ are the coefficients appearing in the decomposition of $F_{T(t)}(\widetilde{\tau}(t))$ expressed as in lemma 5.8, while the derivatives of $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\widetilde{\tau}$ are:

$$\boldsymbol{\lambda}'(t) = \mathrm{diag}\left\{\mathbf{Q}^{\mathrm{T}}(t)\mathbf{M}'(t)\mathbf{Q}(t)\right\}; \tag{5.55}$$

$$\begin{aligned}\mu_i'(t) = &\sum_{j \neq i}\left(\frac{1}{\lambda_i - \lambda_i}\left(\mathbf{v}^{i\mathrm{T}}\mathbf{M}'(t)\mathbf{v}^j\right)\left(\mathbf{v}^{j\mathrm{T}}\left(\mathbf{h} + \mathbf{M}^{-1}(t)\mathbf{b}(t)\right)\right)\right) \\ &+ \mathbf{v}^{i\mathrm{T}}\left(\mathbf{M}^{-1}(t)\mathbf{b}'(t) + \mathbf{M}^{-1}(t)\mathbf{M}'(t)\mathbf{M}^{-1}(t)\mathbf{b}(t)\right) \quad \text{for all } i \in \{1, \dots, N\};\end{aligned} \tag{5.56}$$

$$\widetilde{\tau}'(t) = -\frac{\mathrm{d}}{\mathrm{d}t}c(t) - \mathbf{b}^{\mathrm{T}}(t)\mathbf{M}^{-1}(t)\mathbf{M}'(t)\mathbf{M}^{-1}(t)\mathbf{b}(t) + 2\mathbf{b}^{\mathrm{T}}(t)\mathbf{M}^{-1}(t)\mathbf{b}'(t). \tag{5.57}$$

Proof. The identity (5.53) follows from

$$\begin{aligned}F_{\Psi_t(\mathbf{X})}(\tau) &= \mathbb{P}[\Psi_t(\mathbf{X}) \leq \tau] = \mathbb{P}\left[\mathbf{X}^{\mathrm{T}}\mathbf{M}(t)\mathbf{X} + 2\mathbf{b}(t)^{\mathrm{T}}\mathbf{X} + c(t) \leq \tau\right] \\ &= \mathbb{P}\left[\left(\mathbf{X} + \mathbf{M}^{-1}(t)\mathbf{b}(t)\right)^{\mathrm{T}}\mathbf{M}(t)\left(\mathbf{X} + \mathbf{M}^{-1}(t)\mathbf{b}(t)\right) \leq \tau - c(t) + \mathbf{b}(t)^{\mathrm{T}}\mathbf{M}^{-1}(t) + \mathbf{b}(t)\right] \\ &= \mathbb{P}[T(t) \leq \widetilde{\tau}(t)] = F_{T(t)}(\widetilde{\tau}(t)).\end{aligned}$$

We prove next the differentiability of $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\widetilde{\tau}$ and equations (5.55), (5.56), and (5.57). Equation (5.55) is deduced directly from [173, Equation (4)]. Equation (5.56) can be proven by using [173, Equation (5)] on the derivative of the eigenvector of a symmetric matrix with distinct eigenvalues

$$\mathbf{v}^{i\prime}(t) = (\lambda_i \mathbb{I} - \mathbf{M}(t))^{+} \mathbf{M}'(t)\mathbf{v}^i(t) = \sum_{j\neq i} \frac{1}{\lambda_i - \lambda_j} \left(\mathbf{v}^{j\mathrm{T}} \mathbf{M}'(t)\mathbf{v}^i\right) \mathbf{v}^j,$$

where the symbol "+" denotes the Moore-Penrose inverse. Indeed, using the properties of the Moore-Penrose inverse, we arrive at

$$(\lambda_i \mathbb{I} - \mathbf{M})^{+} = (\mathbf{Q}\,(\lambda_i \mathbb{I} - \mathbf{D})\,\mathbf{Q}^{\mathrm{T}})^{+} = \mathbf{Q}(t)\ \mathrm{diag}\left\{\mathbf{D}^i(t)\right\}\ \mathbf{Q}(t)^{\mathrm{T}}.$$

Herein, for all $i, j \in \{1, \ldots, N\}$, $\mathbf{D}^i(t) = [d_1^i(t), \ldots, d_N^i(t)]^{\mathrm{T}}$ with $d_i^i = 0$ and $d_j^i = \frac{1}{\lambda_i(t) - \lambda_j(t)}$ if $i \neq j$. Since $\mu_i(t) = \mathbf{v}^{i\mathrm{T}} \mathbf{M}^{-1}(t)\mathbf{b}(t)$ for all $1 \leq i \leq N$, we deduce

$$\mu_i'(t) = \mathbf{v}^{i\prime}(t)^{\mathrm{T}} \mathbf{M}^{-1}(t)\mathbf{b}(t) + \mathbf{v}^i(t)^{\mathrm{T}} \mathbf{M}^{-1}(t)\mathbf{M}'(t)\mathbf{M}^{-1}(t)\mathbf{b}(t)\mathbf{M}^{-1}(t)\mathbf{b}'(t),$$

which is equivalent to (5.56). Next, equation (5.57) can be computed directly by applying the chain rule on the definition (5.57) of $\widetilde{\tau}$.

Finally, in order to prove the expression (5.54) of the derivative of $F_{\Psi_t(T)}(\tau)$, we consider the identity (5.53) and the result of theorem 5.5 to write

$$F_{\Psi_t(\mathbf{X})}(\tau) = F_{T(t)}(\widetilde{\tau}(t)) = \sum_{k=0}^{\infty} \gamma_k(t) F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right).$$

By differentiating both sides with respect to t, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} F_{\Psi_t(\mathbf{X})}(\tau) = \frac{\partial}{\partial t_1} F_{T(t_1)}(\widetilde{\tau}(t))\Big|_{t_1=t} + \frac{\partial}{\partial t_2} F_{T(t)}(\widetilde{\tau}(t_2))\Big|_{t_2=t}. \tag{5.58}$$

We treat the two terms on the right-hand side of equation (5.58) separately.

In order to evaluate the first term, we aim to prove the uniform convergence of the series $\sum_{k=0}^{\infty} \mathbf{p}^k \cdot \boldsymbol{\lambda}'(t) F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)$ and $\sum_{k=0}^{\infty} \mathbf{q}^k \cdot \boldsymbol{\mu}'(t) F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)$. We start proving by induction the inequalities

$$\left|p_j^k\right| \leq \eta_k \gamma_k \qquad \text{and} \qquad \left|q_j^k\right| \leq \zeta_k \gamma_k \text{ for all } j \in \{1, \ldots, N\},\ k \geq 0, \tag{5.59}$$

where η_k and ζ_k are defined for $k \geq 0$ as

$$\begin{aligned} \eta_k &= \max_{1\leq i\leq N} \left\{\frac{1}{2\lambda_i}\right\} + \frac{k(k+1)}{2} \max_{1\leq i\leq N} \left\{\frac{\beta(h_i^2+3)}{\lambda_i^2\left(1-\frac{\beta}{\lambda_i}\right)}\right\}, \\ \zeta_k &= \frac{k(k+1)}{2} \max_{1\leq i\leq N} \left\{\frac{2\beta\,|h_i|}{\lambda_i^2\left(1-\frac{\beta}{\lambda_i}\right)}\right\}. \end{aligned} \tag{5.60}$$

For $k = 0$, the inequalities in (5.59) are satisfied. Let us therefore suppose that they are valid for the step $k-1$ and prove that they hold for the step k. Thanks to the fact that

$0 < \beta < \min_{i\in\{1,\dots,N\}} \{\lambda_1, \dots, \lambda_N\}$, we have for all $k \geq 1$ that

$$\begin{aligned}
\left|\nu_j^k\right| &\leq \frac{\beta}{\lambda_j^2} \frac{(k-1)}{\left(1-\frac{\beta}{\lambda_j}\right)} \left(1-\frac{\beta}{\lambda_j}\right)^{k-1} \left(1+(kh_j^2-1)\frac{\beta}{\lambda_j^2}\right) + \frac{\beta}{\lambda_j^2}\left(1-\frac{\beta}{\lambda_j}\right)^{k-1} \left|kh_j^2-1\right| \\
&\leq \frac{\beta}{\lambda_j^2} g_k \left(\frac{k-1}{\left(1-\frac{\beta}{\lambda_j}\right)} + \frac{\left|kh_j^2-1\right|}{\left(1-\frac{\beta}{\lambda_j}+\frac{\beta}{\lambda_j}kh_j^2\right)} \right) \leq \frac{\beta}{\lambda_j^2} g_k \frac{k-1+\left|1-kh_j^2\right|}{\left(1-\frac{\beta}{\lambda_j}\right)} \\
&\leq k\, g_k\, \beta \max_{i\in\{1,\dots,N\}} \left\{ \frac{1+h_j^2+2/k}{\lambda_j^2\left(1-\frac{\beta}{\lambda_j}\right)} \right\} \leq k\, g_k \max_{i\in\{1,\dots,N\}} \left\{ \frac{\beta(h_j^2+3)}{\lambda_j^2\left(1-\frac{\beta}{\lambda_j}\right)} \right\}
\end{aligned}$$

and

$$\left|\kappa_j^k\right| \leq \frac{2k\,|h_j|\,\beta}{\lambda_j}\left(1-\frac{\beta}{\lambda_j}\right)^{k-1} \frac{\left(1+(kh_j^2-1)\frac{\beta}{\lambda_j^2}\right)}{\left(1-\frac{\beta}{\lambda_j^2}+kh_j^2\frac{\beta}{\lambda_j^2}\right)} \leq \frac{2kg_k\,|h_j|\,\beta}{\lambda_j\left(1-\frac{\beta}{\lambda_j}\right)} \leq \max_{i\in\{1,\dots,N\}} \left\{ \frac{2k\,g_k\,\beta\,|h_j|}{\lambda_j\left(1-\frac{\beta}{\lambda_j}\right)} \right\}.$$

In view of such upper bounds and since the sequences $\{\eta_k\}_{k=0}^{\infty}$ and $\{\zeta_k\}_{k=0}^{\infty}$ defined in (5.60) are strictly increasing, we arrive at

$$\begin{aligned}
\left|p_j^k\right| &= \left|\frac{1}{2k}\sum_{\ell=0}^{k-1}\left(\nu_j^{k-\ell}\gamma_\ell + p_j^\ell g_{k-\ell}\right)\right| \leq \frac{1}{2k}\sum_{\ell=0}^{k-1}\left|\nu_j^{k-\ell}\right|\gamma_\ell + \frac{1}{2k}\sum_{\ell=0}^{k-1}\left|p_j^\ell\right| g_{k-\ell} \\
&\leq \max_{i\in\{1,\dots,N\}}\left\{\frac{\beta(h_j^2+3)}{\lambda_j^2\left(1-\frac{\beta}{\lambda_j}\right)}\right\}\frac{1}{2k}\sum_{\ell=0}^{k-1}(k-\ell)g_{k-\ell}\gamma_\ell + \frac{1}{2k}\sum_{\ell=0}^{k-1}\eta_\ell\gamma_\ell g_{k-\ell} \\
&\leq \max_{i\in\{1,\dots,N\}}\left\{\frac{\beta(h_j^2+3)}{\lambda_j^2\left(1-\frac{\beta}{\lambda_j}\right)}\right\}\frac{k}{2k}\sum_{\ell=0}^{k-1} g_{k-\ell}\gamma_\ell + \frac{1}{2k}\eta_{k-1}\sum_{\ell=0}^{k-1}\gamma_\ell g_{k-\ell} \\
&= \left(k\max_{i\in\{1,\dots,N\}}\left\{\frac{\beta(h_j^2+3)}{\lambda_j^2\left(1-\frac{\beta}{\lambda_j}\right)}\right\} + \eta_{k-1}\right)\gamma_k = \eta_k\,\gamma_k,
\end{aligned}$$

and

$$\begin{aligned}
\left|q_j^k\right| &= \left|\frac{1}{2k}\sum_{\ell=0}^{k-1}\left(\kappa_j^{k-\ell}\gamma_\ell + q_j^\ell g_{k-\ell}\right)\right| \leq \frac{1}{2k}\sum_{\ell=0}^{k-1}\left|\kappa_j^{k-\ell}\right|\gamma_\ell + \frac{1}{2k}\sum_{\ell=0}^{k-1}\left|q_j^\ell\right| g_{k-\ell} \\
&\leq \max_{i\in\{1,\dots,N\}}\left\{\frac{2\beta\,|h_j|}{\lambda_j\left(1-\frac{\beta}{\lambda_j}\right)}\right\}\frac{1}{2k}\sum_{\ell=0}^{k-1}(k-\ell)g_{k-\ell}\gamma_\ell + \frac{1}{2k}\sum_{\ell=0}^{k-1}\zeta_\ell\gamma_\ell g_{k-\ell} \\
&\leq \max_{i\in\{1,\dots,N\}}\left\{\frac{2\beta\,|h_j|}{\lambda_j\left(1-\frac{\beta}{\lambda_j}\right)}\right\}\frac{k}{2k}\sum_{\ell=0}^{k-1} g_{k-\ell}\gamma_\ell + \frac{1}{2k}\zeta_{k-1}\sum_{\ell=0}^{k-1}\gamma_\ell g_{k-\ell} \\
&= \left(k\max_{i\in\{1,\dots,N\}}\left\{\frac{2\beta\,|h_j|}{\lambda_j\left(1-\frac{\beta}{\lambda_j}\right)}\right\} + \zeta_{k-1}\right)\gamma_k = \zeta_k\,\gamma_k.
\end{aligned}$$

In order to prove the uniform convergence of the series of (5.59), we use two results from [214]. The first one is presented as [214, Equation (4.14)] and states that

$$\gamma_k \leq \gamma_0 \frac{\Gamma\left(\frac{N}{2}+k\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{\nu^k}{k!} \tag{5.61}$$

for any $k \geq 0$, where ν is a positive constant depending on β, $\boldsymbol{\lambda}(t)$, and $\boldsymbol{\mu}(t)$. The second result is [214, Lemma 4] and states that the series

$$\sum_{k=0}^{\infty} \frac{\Gamma\left(\frac{N}{2}+k\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{\widetilde{\nu}^k}{k!} F_{\chi^2(2k+N)}(x) \tag{5.62}$$

is uniformly convergent (and therefore absolutely convergent) for any positive and finite $\widetilde{\nu}$ and $\bar{x}$ on the interval $[-\infty, \bar{x}]$. Thus, we can introduce the quantities ρ_1, ρ_2, σ_1 and σ_2 with the property

$$\eta_k \leq \rho_1 \sigma_1^k \qquad \text{and} \qquad \zeta_k \leq \rho_2 \sigma_2^k \text{ for all } k \geq 0. \tag{5.63}$$

A suitable choice is given by

$$\begin{aligned} \rho_1 &= \max\nolimits_{1\leq i\leq N} \left\{ \frac{1}{2\lambda_i} \right\}, & \rho_2 &= 1, \\ \sigma_1 &= \max\nolimits_{1\leq i\leq N} \left\{ \frac{\beta(h_i^2+3)}{\lambda_i^2\left(1-\frac{\beta}{\lambda_i}\right)} \right\}, & \sigma_2 &= \max\nolimits_{1\leq i\leq N} \left\{ \frac{2\beta\,|h_i|}{\lambda_i^2\left(1-\frac{\beta}{\lambda_i}\right)} \right\}. \end{aligned} \tag{5.64}$$

Using the bounds from (5.59) and the two results from [214] stated above, we remark that the first and second series in (5.54) are absolutely convergent, since

$$\begin{aligned} &\sum_{k=0}^{\infty} \left| p_j^k \right| F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) \leq \sum_{k=0}^{\infty} \eta_k \gamma_k F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) \\ &\leq \sum_{k=0}^{\infty} \rho_1 \gamma_0 \frac{\Gamma\left(\frac{N}{2}+k\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{(\sigma_1\nu)^k}{k!} F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) < \infty, \end{aligned}$$

and

$$\begin{aligned} &\sum_{k=0}^{\infty} \left| q_j^k \right| F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) \leq \sum_{k=0}^{\infty} \zeta_k \gamma_k F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) \\ &\leq \sum_{k=0}^{\infty} \rho_2 \gamma_0 \frac{\Gamma\left(\frac{N}{2}+k\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{(\sigma_2\nu)^k}{k!} F_{\chi^2(N+2k)}\left(\frac{\tau}{\beta}\right) < \infty. \end{aligned}$$

Thus, the series $\sum_{k=0}^{\infty} \mathbf{p}^k \cdot \boldsymbol{\lambda}'(t) F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)$ and $\sum_{k=0}^{\infty} \mathbf{q}^k \cdot \boldsymbol{\mu}'(t) F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)$ are absolutely convergent and, hence, uniformly convergent by the Weierstrass criterion (see e.g. [215, Theorem 7.10]). Consequently, it is possible to swap the summation and the derivative for the first term

of (5.58) (see [215, Theorem 7.17]) and we obtain

$$\begin{aligned}
&\sum_{k=0}^{\infty}\left(\mathbf{p}^k\cdot\boldsymbol{\lambda}'(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right)+\sum_{k=0}^{\infty}\left(\mathbf{q}^k\cdot\boldsymbol{\mu}'(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right)\\
=&\sum_{k=0}^{\infty}\left(\mathbf{p}^k\cdot\boldsymbol{\lambda}'(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)+\mathbf{q}^k\cdot\boldsymbol{\mu}'(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)\right) \qquad (5.65)\\
=&\sum_{k=0}^{\infty}\gamma_k'(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)=\frac{\partial}{\partial t_1}\sum_{k=0}^{\infty}\gamma_k(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)=\frac{\partial}{\partial t_1}F_{T(t_1)}\left(\widetilde{\tau}(t)\right)\Big|_{t_1=t}.
\end{aligned}$$

We pass to the second term of (5.58). Since the generalized chi-squared distribution of $T(t)$ is continuous in $\mathbb{R}^+$, for any $\tau^*>0$ the quantity $f_{T(t)}(\tau^*)$ exists and is finite for any $\tau^*>0$. Moreover, thanks to theorem 5.5 and corollary 5.6, $f_{T(t)}(\tau^*)=\sum_{k=0}^{\infty}\gamma_k(t)f_{\chi^2(2k+N)}\left(\frac{\tau^*}{\beta}\right)$. The set $\mathfrak{T}=\{\widetilde{\tau}(t)\ :\ t\in[0,\delta]\}$ is compact, so the series converges point wise, and all of its terms are positive, the series $\sum_{k=0}^{\infty}\gamma_k(t)f_{\chi^2(2k+N)}\left(\frac{\tau^*}{\beta}\right)$ is uniformly convergent on $\mathfrak{T}$ [215, Theorem 7.13]. Hence, thanks to the absolute continuity of $\widetilde{\tau}'(t)$ for all $t\in[0,\delta]$, we have

$$\begin{aligned}
\frac{\widetilde{\tau}'(t)}{\beta}\sum_{k=0}^{\infty}\gamma_k(t)f_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t)}{\beta}\right)&=\sum_{k=0}^{\infty}\frac{\partial}{\partial t_2}\left(\gamma_k(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t_2)}{\beta}\right)\right)\Big|_{t_2=t}\\
=\frac{\partial}{\partial t_2}\left(\sum_{k=0}^{\infty}\gamma_k(t)F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}(t_2)}{\beta}\right)\right)\Big|_{t_2=t}&=\frac{\partial}{\partial t_2}F_{T(t)}\left(\widetilde{\tau}(t_2)\right)\Big|_{t_2=t}.
\end{aligned} \qquad (5.66)$$

In conclusion, the combination of the equations (5.58), (5.65), and (5.66) proves the expression (5.54) for the derivative of the cumulative distribution function of $\Psi_t(\mathbf{X})$. □

5.4.3 Shape optimization under Gaussian perturbations

Let us consider once again the shape optimization problem (5.3). Using the notations of Section 5.2, we suppose that the random vector $\mathbf{X}$ follows a Gaussian distribution with mean $\mathbf{h}=[h_1,\ldots,h_N]^{\mathrm{T}}$ and, without loss of generality, covariance matrix equal to the identity.

If the vector $\mathbf{h}$ or the deterministic load $\mathbf{g}_0$ are large enough, the uncertain component can be seen as a small random perturbation around a deterministic load $\overline{\mathbf{g}}=\mathbf{g}_0+\overline{\mathbf{g}}_1h_1+\ldots+\overline{\mathbf{g}}_Nh_N$, and the shape derivative can be computed as in [11, Section 4.2.3]. Otherwise, if the mechanical loads are centered on 0 or the uncertainties are wide enough not to be treated as small perturbations, a different method should be considered. If the probability density $f_{\mathbf{X}}$ of the uncertainties is known, the technique detailed in Section 5.3.1 can be applied. However, if the number of random variables involved in the modelization of the uncertainties is significant, the computation of the integrals on the N-ball and the N-sphere can be challenging.

Since we suppose that $\mathbf{X}$ follows a Gaussian distribution, by considering the diagonalization of the matrix $\mathbf{M}_\Omega=\mathbf{U}_\Omega\mathbf{D}_\Omega{\mathbf{U}_\Omega}^{\mathrm{T}}$, we can use corollary 5.6 and proposition 5.9 to express $\Phi(\Omega)=\mathbb{P}\left[\Psi_\Omega(\mathbf{X})\leq\tau\right]$ as the cumulative distribution function of a generalized chi-squared random variable, and compute the shape derivative of $\Phi(\cdot)$ in $\Omega\in\mathcal{S}_{\mathrm{adm}}$.

Proposition 5.10. *Let $\mathbf{X}\sim\mathcal{N}(\boldsymbol{\mu},\mathbb{I})$ be a Gaussian random vector in $\mathbb{R}^N$, $\Omega\in\mathcal{S}_{\mathrm{adm}}$ a Lipschitz continuous domain in $\mathbb{R}^2$ or $\mathbb{R}^3$, and $\tau\in\mathbb{R}^+$ a strictly positive threshold. The quantities $\mathbf{M}_\Omega\in\mathbb{R}^{N\times N}$, $\mathbf{b}_\Omega\in\mathbb{R}^N$, and $c_\Omega\in\mathbb{R}$ are functions of the domain $\Omega\in\mathcal{S}_{\mathrm{adm}}$, and are defined as in Section 5.2.2, and we suppose that $\widetilde{\tau}_\Omega$, defined as in (5.8), is strictly positive for all $\Omega\in\mathcal{S}_{\mathrm{adm}}$. In addition, we suppose that the mappings $\Omega\mapsto[\mathbf{M}_\Omega]_{i,j}$, $\Omega\mapsto[\mathbf{b}_\Omega]_i$, and $\Omega\mapsto c_\Omega$ admit a shape*

derivative at Ω for all $i, j \in \{1, \dots, N\}$ and that all eigenvalues of $\mathbf{M}_\Omega$ are distinct, strictly positive, and larger than a positive constant β independent from Ω.

Then, $\Phi(\Omega)$ can be written as the cumulative distribution function as $\Phi(\Omega) = F_{T_\Omega}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)$, where T_Ω is a random variable such that

$$T_\Omega \sim \widetilde{\chi^2}\left(\mathbf{1}; \boldsymbol{\mu}_\Omega \odot \boldsymbol{\mu}_\Omega; \boldsymbol{\lambda}_\Omega\right)$$

with $\boldsymbol{\lambda}_\Omega$ being the vector of the eigenvalues of $\mathbf{M}_\Omega$ and $\boldsymbol{\mu}_\Omega = \left(\mathbf{h} + {\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right)$. Moreover, $\Phi(\cdot)$ is shape-differentiable at Ω, and its derivative can be expressed as

$$\begin{aligned} \mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) &= \left(\sum_{k=0}^{\infty} \mathbf{p}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right) \cdot \mathrm{D}\boldsymbol{\lambda}_\Omega(\boldsymbol{\theta}) \\ &+ \left(\sum_{k=0}^{\infty} \mathbf{q}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right) \cdot \mathrm{D}\boldsymbol{\mu}_\Omega(\boldsymbol{\theta}) + \frac{1}{\beta}\left(\sum_{k=0}^{\infty} \gamma_k f_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right) \mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta}). \end{aligned} \tag{5.67}$$

Once again, the components of $\mathbf{p}^k$ and $\mathbf{q}^k$ are the coefficients appearing in the decomposition of $F_{T_\Omega}(\widetilde{\tau}_\Omega)$ expressed as in lemma 5.8, while $\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta})$ is as in equation (5.15), and the shape derivatives of $\boldsymbol{\lambda}_\Omega$, $\boldsymbol{\mu}_\Omega$ are

$$\begin{aligned} \mathrm{D}\boldsymbol{\lambda}_\Omega(\boldsymbol{\theta}) &= \mathrm{diag}\left\{{\mathbf{U}_\Omega}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right\}; \\ \mathrm{D}\mu_{\Omega i}(\boldsymbol{\theta}) &= \sum_{j \neq i}\left(\frac{1}{\lambda_{\Omega,i} - \lambda_{\Omega,j}}\left({\mathbf{v}^i}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{v}^j\right)\left({\mathbf{v}^j}^{\mathrm{T}}\left(\mathbf{h} + {\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right)\right)\right) \\ &+ {\mathbf{v}^i}^{\mathrm{T}}\left({\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta}) + {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta}){\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right) \qquad \textit{for all } i \in \{1, \dots, N\}. \end{aligned}$$

Proof. The proof of the identity $\Phi(\Omega) = F_{T_\Omega}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)$ is analogous to the proof of equation (5.53) in proposition 5.9. In order to compute the shape derivative of $\Phi(\cdot)$ at Ω, we recall that the identity (5.18) holds for any differentiable shape functional $\mathcal{S}_{\mathrm{adm}} \to \mathbb{R}$ any Lipschitz continuous domain, and any mapping $\boldsymbol{\xi} : [0, \delta] \to \mathrm{W}^{1,\infty}\left(\mathbb{R}^d; \mathbb{R}^d\right)$. Thus, taking as deformation field $\boldsymbol{\xi}(t) = t\boldsymbol{\theta}$, we have

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi\left((\mathbb{I} + t\boldsymbol{\theta})\Omega\right)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}F_{T_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}}(\widetilde{\tau}_\Omega)\Big|_{t=0}.$$

We denote $T(t) = T_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}$, $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}$, $\boldsymbol{\mu}(t) = \boldsymbol{\mu}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}$, and $\widetilde{\tau}(t) = \widetilde{\tau}_{(\mathbb{I}+t\boldsymbol{\theta})\Omega}$. Equation (5.67) and the expressions of the shape derivatives of $\boldsymbol{\lambda}_\Omega$, $\boldsymbol{\mu}_\Omega$ and $\widetilde{\tau}_\Omega$ are found using proposition 5.9 and the identity (5.18). □

5.5 Presentation of the algorithm

5.5.1 Generalities on the algorithm

The theoretical results stated in the previous section have been applied to the shape optimization of a cantilever (in section 5.6.1) and a bridge-like structure (in section 5.6.2). In both examples, we considered the structure to be composed by an isotropic linear elastic material, subject to random mechanical loads. For the two structures, we aimed to minimize their mass under constraints on the probability of the compliance to exceed a threshold. We recall from

section 2.1.2 that the compliance of an elastic structure Ω is defined as the work of the external mechanical load $\mathbf{g}$ and can be expressed as a quadratic function of the displacement $\mathbf{u}_\Omega$ as

$$\mathcal{C}(\Omega, \mathbf{u}_\Omega) = \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{u}_\Omega \, \mathrm{d}\mathbf{s} = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_\Omega) : \boldsymbol{\varepsilon}(\mathbf{u}_\Omega) \, \mathrm{d}\mathbf{x}. \tag{5.68}$$

The problems considered in the following can be resumed by the following structure:

$$\left|\begin{array}{l} \text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\ \text{ minimizing } \mathrm{Vol}(\Omega) \\ \text{ under the constraint} \\ \qquad \Phi(\Omega) = \mathbb{P}\left[\mathcal{C}(\Omega, \mathbf{u}_\Omega(\omega)) > \tau\right] \leq \bar{p}, \\ \text{where the state } \mathbf{u}_\Omega \text{ satisfies the state equation (5.2) for almost all } \omega \in \mathcal{O} \\ \text{with } \mathbf{g} \in \mathrm{L}^2\left(\mathcal{O}, \mathbb{P}; \mathrm{L}^2(\Gamma_N)\right) \text{ such that} \\ \qquad \mathbf{g}(\omega) = \overline{\mathbf{g}}_0 + \overline{\mathbf{g}}_1 X_1(\omega) + \ldots + \overline{\mathbf{g}}_N X_N(\omega) \qquad \text{for almost all } \omega \in \mathcal{O}, \text{ and} \\ \qquad \mathbf{X} = (X_1, \ldots, X_N)^{\mathrm{T}} \sim \mathcal{N}(\mathbf{h}, \mathbb{I}). \end{array}\right. \tag{5.69}$$

All simulations have been performed under the python-based *sotuto* platform proposed by Dapogny and Feppon in [91] and presented in section 1.3. The computation of the elastic displacements and the adjoint states has been performed using the finite-element solver *FreeFem++* [136]. We represented the domains by the means of conforming meshes obtained using the implicit-domain remeshing tool of *mmg* [89], coupled to the level-set representation of the shapes [12, 247]. The advection of the level-set function is handled by the *advect* library [56], while the computation of the signed distance function is performed by *mshdist* [92]. The simulations have been ran on a Virtualbox virtual machine Linux with 1GB of dedicated memory, installed on a Dell PC equipped with a 2.80 GHz Intel i7 processor.

For all structures we considered two cases. In the first one we suppose $\mathbf{X}$ to be a centered Gaussian random variable, and we adopt the method described in section 5.3.3. In the second case we consider $\mathbf{X}$ to be a non-centered Gaussian vector, and we apply the results of section 5.4.3.

In the rest of this section we present the details of the numerical implementation of the to methods. Let $D \in \mathbb{R}^d$ be a computational domain such that any admissible shape Ω is included into D. For both situations, at each step n of the optimization algorithm we have a mesh $\mathcal{T}_{D,\Omega}$ covering D where the domain Ω is represented explicitly. At first, we solve numerically the state equation (5.4) for the N different mechanical loads $\overline{\mathbf{g}}_1, \ldots, \overline{\mathbf{g}}_N$, obtaining the displacement fields $\mathbf{u}_{\Omega,1}, \ldots, \mathbf{u}_{\Omega,N}$. Next, we compute all entries of the matrix $\mathbf{M}_\Omega$ as

$$[\mathbf{M}_\Omega]_{ij} = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_{\Omega,i}) : \boldsymbol{\varepsilon}(\mathbf{u}_{\Omega,j}) \, \mathrm{d}\mathbf{x}.$$

The matrix $\mathbf{M}_\Omega$ is symmetric and positive definite and is diagonalized as $\mathbf{M}_\Omega = \mathbf{U}_\Omega \mathbf{D}_\Omega \mathbf{Q}_\Omega^{\mathrm{T}}$, where $\mathbf{U}_\Omega$ is the orthogonal matrix of the eigenvectors, and $\mathbf{D}_\Omega = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$ the diagonal matrix of the eigenvalues. If the term $\overline{\mathbf{g}}_0$ is not $\mathbf{0}$, the quantities $\mathbf{b}_\Omega \in \mathbb{R}^N$ and $c_\Omega \in \mathbb{R}$ are computed as

$$[\mathbf{b}_\Omega]_i = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_{\Omega,i}) : \boldsymbol{\varepsilon}(\mathbf{u}_{\Omega,0}) \, \mathrm{d}\mathbf{x} \qquad \text{and} \qquad c_\Omega = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_{\Omega,0}) : \boldsymbol{\varepsilon}(\mathbf{u}_{\Omega,0}) \, \mathrm{d}\mathbf{x}.$$

The computation of $\Phi(\Omega)$ and its shape derivative $\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta})$ is done differently in the two methods, as detailed in the rest of this section.

5.5.2 Centered Gaussian perturbations

In the case where $\overline{\mathbf{g}}_0 = \mathbf{0}$ and $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$ is a centered Gaussian vector, the computation of $\Phi(\Omega)$ relies on proposition 5.3. Let us consider a tolerance $\varepsilon > 0$ for the numerical computation of $\Phi(\Omega)$. The first step is the computation of the coefficients A_k^i and B_k for $i \in \{1, \ldots, N\}$ according to the recursive relation (5.33) - (5.34), up to a sufficiently large index $\bar{k}_\varepsilon$. The probability $\Phi(\Omega)$ is approximated thanks to proposition 5.3 as

$$\Phi(\Omega) = \frac{(\lambda_N)^{N/2}}{2\,(\pi)^{N/2}\sqrt{\det \mathbf{M}_\Omega}} \sum_{k=0}^{\bar{k}_\varepsilon} \frac{1}{k!} \gamma\left(\frac{N}{2} + k, \frac{\tau}{2\lambda_N}\right) B_k.$$

According to proposition C.21, in order to compute $\Phi(\Omega)$ up to the tolerance ε it is necessary for $\bar{k}_\varepsilon$ to satisfy the following condition

$$\frac{\left(\frac{1}{\lambda_N} - \frac{1}{\lambda_1}\right)^{\bar{k}_\varepsilon - 1}}{(N + 2(\bar{k}_\varepsilon + 1))(\bar{k}_\varepsilon + 1)!} \leq \frac{2\varepsilon\, e^{\frac{\tau}{2\lambda_1}}}{\tau\,|\mathbb{S}_{N-1}|}.$$

Let $\mathbf{v}^1, \ldots, \mathbf{v}^N$ be the N unitary eigenvectors of $\mathbf{M}_\Omega$. We compute the displacements $\widetilde{\mathbf{u}_{\Omega,1}}, \ldots, \widetilde{\mathbf{u}_{\Omega,N}}$ as $\mathbf{u}_{\Omega,i} = \sum_{j=1}^N \mathbf{v}^i{}_j \mathbf{u}_{\Omega,j}$ thanks to the linearity of the state equation. In order to compute the shape derivative $\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta})$ it is necessary to compute N shape derivatives $\mathrm{D}\mathcal{C}(\Omega, \widetilde{\mathbf{u}_{\Omega,i}})(\boldsymbol{\theta})$. Then, we can apply proposition 5.4 and use the coefficients $A_0^i, \ldots, A_{\bar{k}_\varepsilon}^i$ computed earlier and get

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = -\frac{\left(\frac{\tau}{2\pi}\right)^{N/2}}{2\sqrt{\det\,\mathbf{M}_\Omega}} \sum_{i=1}^N \left(\frac{e^{-\frac{\tau}{2\lambda_N}}}{\lambda_i} \mathrm{D}\mathcal{C}(\Omega, \widetilde{\mathbf{u}_{\Omega,i}})(\boldsymbol{\theta}) \sum_{k=0}^{\bar{k}_\varepsilon} \frac{A_k^i}{k!} \left(\frac{\tau}{2\lambda_N}\right)^k \right).$$

5.5.3 Non-Centered Gaussian perturbations

If $\overline{\mathbf{g}}_0$ is not zero or if $\mathbf{X}$ is a Gaussian vector centered on $\mathbf{h} \neq \mathbf{0}$, we use the results of 5.4 to compute $\Phi(\Omega)$ up to an arbitrary precision, as well as its shape derivative. The first step consists in computing the vectors $\boldsymbol{\lambda}_\Omega$ and $\boldsymbol{\mu}_\Omega$ as

$$\boldsymbol{\lambda}_\Omega = [\lambda_1, \ldots, \lambda_N]^{\mathrm{T}} \qquad \text{and} \qquad \boldsymbol{\mu}_\Omega = \mathbf{h} + (\mathbf{M}_\Omega)^{-1}\mathbf{b}_\Omega.$$

Moreover, we denote β_n and $\widetilde{\tau}_\Omega$ the quantities

$$\beta_n = \frac{1}{2}\min(\lambda_1, \ldots, \lambda_N) \qquad \text{and} \qquad \widetilde{\tau}_\Omega = \tau - \left(c_\Omega - \mathbf{b}_\Omega{}^{\mathrm{T}}(\mathbf{M}_\Omega)^{-1}\mathbf{b}_\Omega\right).$$

Thanks to proposition 5.10 and theorem 5.5, $\Phi(\Omega) = F_T\left(\frac{\widetilde{\tau}_\Omega}{\beta_n}\right)$, where T is a generalized non-central chi-squared random variable such that

$$T \sim \widetilde{\chi^2}(\mathbf{1}; \boldsymbol{\mu}_\Omega \odot \boldsymbol{\mu}_\Omega; \boldsymbol{\lambda}_\Omega).$$

Let us consider a tolerance $\varepsilon > 0$. In order to compute $\Phi(\Omega)$ up to a precision ε we need to compute the coefficients $\gamma_0, \ldots, \gamma_{\bar{k}_\varepsilon}$ and $g_1, \ldots, g_{\bar{k}_\varepsilon}$ according to the following recursive relation

$$\left\{ \begin{array}{rcl} \gamma_0 & = & e^{-\frac{1}{2}\|\boldsymbol{\mu}_\Omega\|^2} \beta_n^{N/2} \left(\det \mathbf{D}_\Omega\right)^{-1/2}, \\ \gamma_k & = & \sum_{\ell=0}^{k-1} g_{k-\ell}\gamma_\ell \end{array} \right. \qquad \text{and} \qquad g_k = \sum_{i=1}^N \left(1 - \frac{\beta_n}{\lambda_i}\right)^{k-1}. \tag{5.70}$$

The number $\bar{k}_\varepsilon$ of terms to be considered satisfies the following inequality

$$\left(1 - \sum_{k=0}^{\bar{k}_\varepsilon} \gamma_k\right) F_{\chi^2\left(2\bar{k}_\varepsilon+N+2\right)}\left(\frac{\widetilde{\tau}_\Omega}{\beta_n}\right) < \varepsilon.$$

Having computed the coefficients $\gamma_0, \ldots, \gamma_{\bar{k}_\varepsilon}$, the probability $\Phi(\Omega)$ can be computed up to a precisionε as

$$\Phi(\Omega) = \sum_{k=0}^{\bar{k}_\varepsilon} \gamma_k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta_n}\right). \tag{5.71}$$

The computation of the shape derivative of $\Phi(\Omega)$ requires the evaluation of the coefficients $\mathbf{p}^1, \ldots, \mathbf{p}^{\bar{k}_\varepsilon} \in \mathbb{R}^N$ and $\mathbf{q}^1, \ldots, \mathbf{q}^{\bar{k}_\varepsilon} \in \mathbb{R}^N$, as well as the derivatives of $\widetilde{\tau}_\Omega$, $\boldsymbol{\lambda}_\Omega$, and $\boldsymbol{\mu}_\Omega$. According to equation (5.15) and proposition 5.10 such derivatives can be written in terms of the derivatives of $\mathbf{M}_\Omega$, $\mathbf{b}_\Omega$, and c_Ω as

- $\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta}) = -\mathrm{D}c_\Omega(\boldsymbol{\theta}) - {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta}){\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega + {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta})$;
- $\mathrm{D}\boldsymbol{\lambda}_\Omega(\boldsymbol{\theta}) = \mathrm{diag}\left\{{\mathbf{U}_\Omega}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{U}_\Omega\right\}$;
- $\mathrm{D}\mu_{\Omega i}(\boldsymbol{\theta}) = \sum_{j\neq i}\left(\frac{1}{\lambda_{\Omega,i}-\lambda_{\Omega,i}}\left({\mathbf{v}^i}^{\mathrm{T}}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\mathbf{v}^j\right)\left({\mathbf{v}^j}^{\mathrm{T}}\left(\mathbf{h} + {\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right)\right)\right)$
  $+{\mathbf{v}^i}^{\mathrm{T}}\left({\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{b}_\Omega(\boldsymbol{\theta}) + {\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta}){\mathbf{M}_\Omega}^{-1}\mathbf{b}_\Omega\right)$ for all $i \in \{1, \ldots, N\}$.

The vectors $\mathbf{p}^1, \ldots, \mathbf{p}^{\bar{k}_\varepsilon} \in \mathbb{R}^N$ and $\mathbf{q}^1, \ldots, \mathbf{q}^{\bar{k}_\varepsilon} \in \mathbb{R}^N$ are computed component by component according to lemma 5.8 as

- $p_j^0 = -\frac{\gamma_0}{2\lambda_j}$ and $p_j^k = \frac{1}{2k}\sum_{\ell=0}^{k-1}\left(\nu_j^{k-\ell}\gamma_\ell + p_j^\ell g_{k-\ell}\right)$ for $1 \leq k \leq \bar{k}_\varepsilon$;
- $q_j^0 = 0$ and $q_j^k = \frac{1}{2k}\sum_{\ell=0}^{k-1}\left(\kappa_j^{k-\ell}\gamma_\ell + q_j^\ell g_{k-\ell}\right)$ for $1 \leq k \leq \bar{k}_\varepsilon$;
- $\nu_j^1 = \frac{\beta}{\lambda_j^2}(1-\mu_j^2)$ and $\nu_j^k = \frac{\beta}{\lambda_j^2}(1-\frac{\beta}{\lambda_j})^{k-2}\left[(k-1)(1+\frac{\beta}{\lambda_j}(k\mu_j^2-1)) + (1-\frac{\beta}{\lambda_j})(1-k\mu_j^2)\right]$ for $1 \leq k \leq \bar{k}_\varepsilon$;
- $\kappa_j^k = 2k\mu_j\frac{\beta}{\lambda_j}(1-\frac{\beta}{\lambda_j})^{k-1}$ for $1 \leq k \leq \bar{k}_\varepsilon$.

After the computation of all these terms, the shape derivative of $\Phi(\Omega)$ is computed according to proposition 5.10 as

$$\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) = \left(\sum_{k=0}^{\infty}\mathbf{p}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right)\cdot\mathrm{D}\boldsymbol{\lambda}_\Omega(\boldsymbol{\theta})$$
$$+\left(\sum_{k=0}^{\infty}\mathbf{q}^k F_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right)\cdot\mathrm{D}\boldsymbol{\mu}_\Omega(\boldsymbol{\theta}) + \frac{1}{\beta}\left(\sum_{k=0}^{\infty}\gamma_k f_{\chi^2(2k+N)}\left(\frac{\widetilde{\tau}_\Omega}{\beta}\right)\right)\mathrm{D}\widetilde{\tau}_\Omega(\boldsymbol{\theta}).$$

5.6 Numerical simulations

5.6.1 Optimization of a 3D cantilever

We consider Ω to be the cantilever structure represented as seen in fig. 5.1, subject to an uncertain mechanical load $\mathbf{g}$ perpendicular to the main axis of the cantilever. The load is applied on the

region of the boundary denoted by Γ_{N}, while the structure is clamped on the four corner regions marked as Γ_{D}. We suppose that the cantilever has a square cross section with side length ℓ_s, and its length along the x axis is ℓ_x. Moreover, we consider the structure to made up of an elastic material characterized by a Young's modulus E and a Poisson's ratio ν. We consider the uncertain load to have the structure

$$\mathbf{g}(\omega) = \overline{g}_x X_x(\omega)\mathbf{e}_x + \overline{g}_y X_y(\omega)\mathbf{e}_y + \left(\overline{g}_0 + \overline{g}_z X_z(\omega)\right)\mathbf{e}_z, \tag{5.72}$$

where X_x, X_y and X_z are real valued Gaussian random variables, $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ is the canonical basis of $\mathbb{R}^3$, and $\overline{g}_x$, $\overline{g}_y$, $\overline{g}_z$ and $\overline{g}_0$ are deterministic forces. The geometric and material properties of the structure are collected in table 5.2.

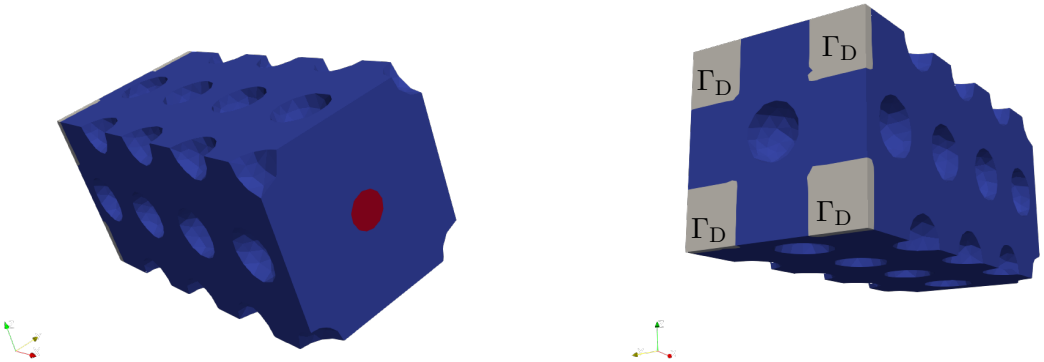


Figure 5.1: Structure of the 3D cantilever. The region Γ_{N} where the random load is applied is marked in red, while the clamping region Γ_{D} is highlighted in grey.

We considered two different cases. In **case A**, we consider a random load $\mathbf{g}_A$ orthogonal to the main axis of the cantilever. In **case B**, the stochastic term in the direction y in the load $\mathbf{g}_B$ is replaced by a random traction-compression force parallel to the main axis x. For both cases we optimized the structure against centered Gaussian perturbations using the method of section 5.5.2 and against Gaussian perturbations with a deterministic load $\mathbf{g}_0$ using the method of section 5.5.3. The results of the noncentered optimization problems are compared to a fully deterministic problem where $\mathbf{g}_D = \overline{g}_0\mathbf{e}_z$ is the only load applied to Γ_{N}, and the constraint $\Phi(\Omega) \leq \tau$ of the optimization problem (5.69) is replaced by

$$\mathcal{C}(\Omega, \mathbf{u}_\Omega) \leq \tau\,.$$

The results of the optimization problems under centered loads are reported in table 5.3, and the resulting optimal shapes are presented in fig. 5.2 and fig. 5.3. The trends of the objective function and the constraint are shown in fig. 5.4a and fig. 5.4b respectively.

By comparing the optimal shapes presented in fig. 5.2 and fig. 5.3 we remark that both have a hollow structure, allowing to distribute the random mechanical load from Γ_{N} to the four supports of Γ_{D}. The graphs of fig. 5.4 show that the objective function decreases progressively with the number of iterations (fig. 5.4a), and the constraint on the probability of the compliance to exceed a threshold τ is satisfied (fig. 5.4b). An examination of table 5.3 confirms that the

Geometry of the structure		
cross section length	ℓ_s	1.0 cm
longitudinal length	ℓ_x	2.0 cm
sidelength of Γ_{D}		0.3 cm
radius of Γ_{N}		0.1 cm
Elastic coefficients		
Young's modulus	E	200 MPa
Poisson's ratio	ν	0.3
Mechanical loads		
compression load	$\overline{g}_x$	10 kPa
horizontal load	$\overline{g}_y$	10 kPa
vertical load	$\overline{g}_z$	10 kPa
Mesh size parameters		
minimal mesh size	`hmin`	0.025 cm
maximal mesh size	`hmin`	0.10 cm
Thresholds for the inequality constraints		
threshold on the compliance	τ	3.3×10^{-2} MPa cm^3
bound on the probability of failure	$\bar{p}$	1.0%

Table 5.2: Numerical data concerning the geometry and the mechanics of the cantilever structure of fig. 5.1.

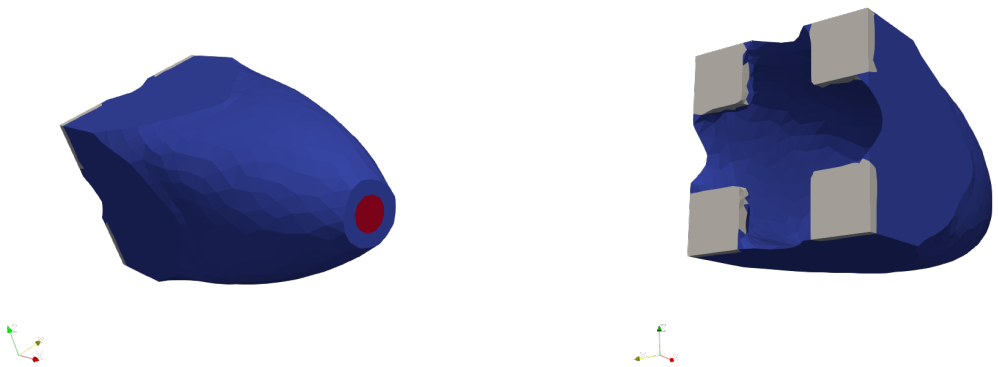


Figure 5.2: Optimal shape for **case A**, where the applied load is $\mathbf{g}_A(\omega) = \overline{g}_y X_y(\omega)\mathbf{e}_y + \overline{g}_z X_z(\omega)\mathbf{e}_z$.

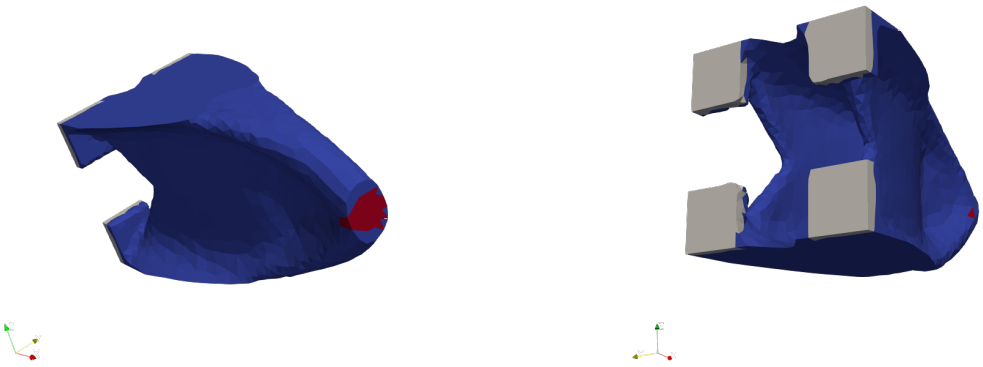
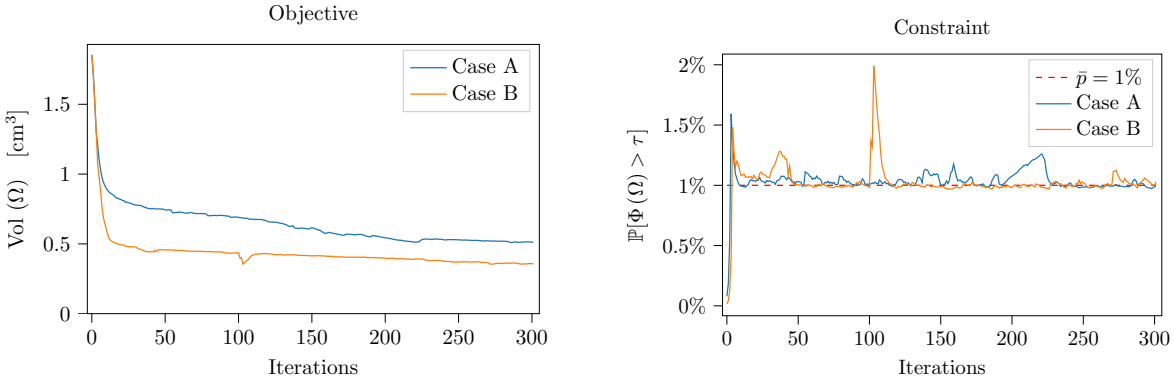


Figure 5.3: Optimal shape for **case B**, where the applied load is $\mathbf{g}_B(\omega) = \overline{g}_x X_x(\omega)\mathbf{e}_x + \overline{g}_z X_z(\omega)\mathbf{e}_z$.

Centered loads	**case A**	**case B**
Number of iterations	300	300
Execution time	141 min 26 s	158 min 56 s
Final volume $\mathrm{Vol}\,(\Omega_{\mathrm{opt}})$	$0.512\,\mathrm{cm}^3$	$0.355\,\mathrm{cm}^3$
$\mathbb{P}\left[\mathcal{C}\left(\Omega, \mathbf{u}_\Omega(\omega)\right) > \tau\right]$		
Excess probability under load $\mathbf{g}_A$	0.985 %	24.701 %
Excess probability under load $\mathbf{g}_B$	0.258 %	1.032 %

Table 5.3: Numerical results for the optimization of the volume of a cantilever subject to centered Gaussian mechanical loads under constraint on the probability of the compliance to exceed a threshold τ.



(a) Evolution of the objective function.

(b) Evolution of the constraint.

Figure 5.4: Convergence of the objective and the constraints for the cantilever problems under centered perturbations.

approach detailed in section 5.5.2 to differentiate $\Phi(\Omega) = \mathbb{P}[\mathcal{C}(\Omega, \mathbf{u}_\Omega(\omega)) > \tau]$ with respect to shape is correct. Indeed, the computation of $\Phi(\Omega)$ for the two optimized structures under cases **A** and **B** respectively yields results close to the upper bound $\bar{p} = 1.0\%$ on the probability both times (0.985% and 1.032%). By subjecting the two structures to the other loading condition we remark that the structure optimized for **case A** complies with the constraint imposed by the loadings of **case B**, but the reciprocate does not hold (the probability of failure being 0.258% and 24.701% respectively).

A notable consideration emerging from table 5.3 concerns the duration of the optimization. The execution time of the simulations discussed in this section is comparable to the time to solve a deterministic problem. Indeed, the largest fraction of the duration of each iteration comprises the solution of the finite element problems and the mesh adaptation.

The optimal shape for the noncentered cases are presented in fig. 5.2 (for **case A**) and fig. 5.3 (for **case B**), while the optimal shape for the deterministic problem is in fig. 5.7. The results for the problems with noncentered problems are compared in table 5.4 with the outcome of the deterministic problem. The decrease of the objective function in the three problems is shown in fig. 5.8a, and the trend of the constraint for **case A** and **case B** is reported in fig. 5.8b.

Noncentered loads	**case A**	**case B**	**Deterministic case**
Execution			
Number of iterations	500	500	348
Execution time	152 min 32 s	177 min 49 s	114 min 15 s
Final volume $\mathrm{Vol}(\Omega_{\mathrm{opt}})$	$0.4605\,\mathrm{cm}^3$	$0.4103\,\mathrm{cm}^3$	$0.0573\,\mathrm{cm}^3$
$\mathbb{P}[\mathcal{Q}(\mathbf{u}_\Omega(\omega), \Omega) > \tau]$			
Excess probability under load $\mathbf{g}_A$	0.996 %	4.005 %	59.579 %
Excess probability under load $\mathbf{g}_B$	4.726 %	0.991 %	88.293 %

Table 5.4: Numerical results for the optimization of the volume of a cantilever subject to noncentered uncertain mechanical loads under constraint on the probability of the compliance to exceed a threshold τ.

By comparing fig. 5.5 and fig. 5.6, we observe that the optimal solutions for **case A** and **case B** are quite similar to the ones optimized for centered loads, being convex hulls. The main difference being the slight reinforcement on the z direction. In contrast, the solution of the deterministic problem presented in fig. 5.7 is radically different, showing a thin branched structure. Such difference can be explained by the fact that, on average, the cantilever is subject to a stronger mechanical load in **case A** and **case B**, therefore the corresponding optimal structures ought to be more robust in order to satisfy the constraint on the probability for the compliance to exceed the threshold τ.

Another notable difference between the deterministic and the uncertain cases concerns the speed of convergence. Figure 5.8a shows that the volume of the cantilever in the deterministic problem converges much faster than the simulations of **case A** and **case B**. Moreover, in the deterministic case, the optimization algorithm reaches a satisfying result and stops after 349 iterations, while the rate of convergence is much slower for **case A** and **case B**. Difficulties in the convergence of the cantilever structure discussed here have also been observed in [112, Section 6.2.1].

Finally, we remark that the shapes resulting from the solution of for **case A** and **case B** comply with the constraint on the probability of failure, as shown in table 5.4. The observance

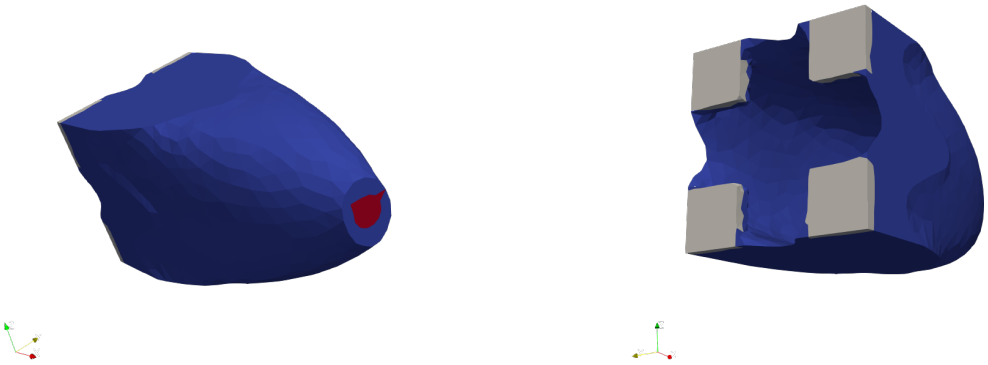


Figure 5.5: Optimal shape for **case A**, where the applied load is $\mathbf{g}_A(\omega) = \overline{g}_y X_y(\omega)\mathbf{e}_y + (\overline{g}_0 + \overline{g}_z X_z(\omega))\,\mathbf{e}_z$.

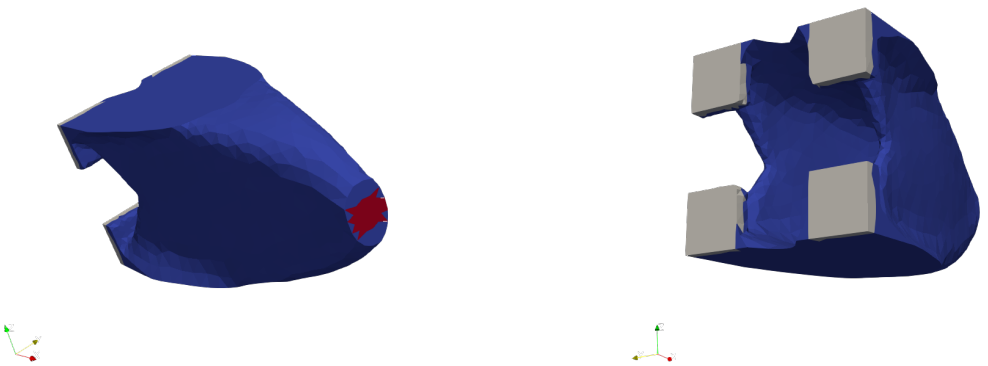


Figure 5.6: Optimal shape for **case B**, where the applied load is $\mathbf{g}_B(\omega) = \overline{g}_x X_x(\omega)\mathbf{e}_x + (\overline{g}_0 + \overline{g}_z X_z(\omega))\,\mathbf{e}_z$.

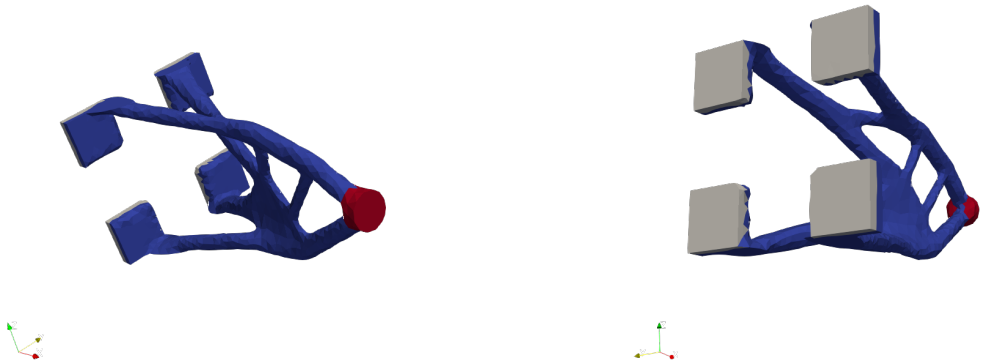
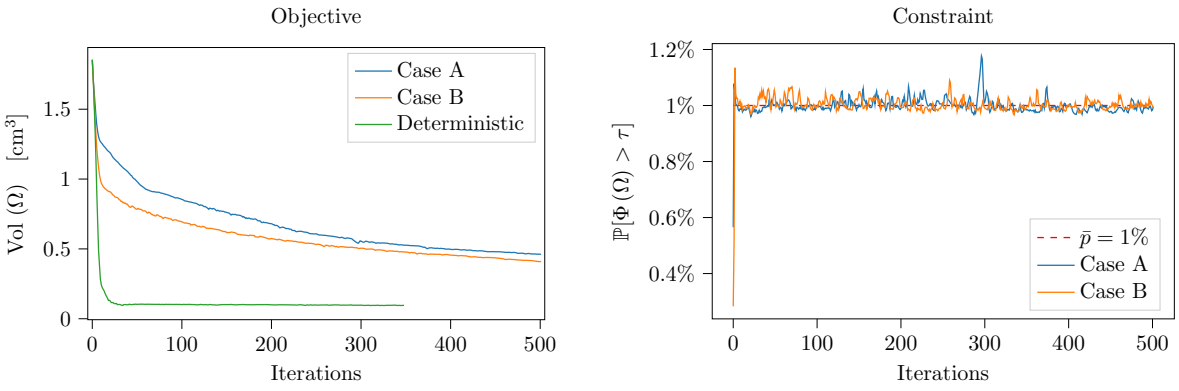


Figure 5.7: Optimal shape for the **deterministic case**, where the mechanical load applied is $\mathbf{g}_D = \overline{g}_0 \mathbf{e}_z$.



(a) Evolution of the objective function.

(b) Evolution of the constraint.

Figure 5.8: Convergence of the objective and the constraints for the cantilever problems under noncentered perturbations.

of the constraint, the decrease of the objective functional, and the radically different result with respect to the deterministic case justify the use of the *null space optimization* algorithm for the solution of problem (5.69), and the suitability of the approach of section 5.4 for the expression of $\Phi(\Omega)$ and its shape derivative.

5.6.2 Optimization of a 3D bridge

As a second example, we consider the optimization of the bridge structure found in fig. 5.9. The structure is clamped on the lower surface on its four corners, marked in grey in the picture. The pinned region, where Dirichlet boundary conditions on the displacement are applied, is denoted Γ_D. The upper face of the bridge is divided into five sections $\Gamma_N^1, \dots, \Gamma_N^5$ of equal size. On each section Γ_N^i, a random load $\mathbf{g}_i \in L^2(\mathcal{O}, \mathbb{P}; L^2(\Gamma_N^i))$ is applied. We suppose that the loads are oriented vertically (that is along the z axis), independent from one another, and such that

$$\mathbf{g}_i(\omega) = -\overline{g}_i X_i(\omega)\mathbf{e}_z \qquad \text{on } \Gamma_N{}^i \tag{5.73}$$

for all $i \in \{1, \dots, 5\}$, where $\overline{g}_i\mathbf{e}_z$ is a deterministic vertical pressure and X_i a Gaussian random variable. The numerical parameters describing the geometry and the mechanical properties of the bridge are reported in table 5.5.

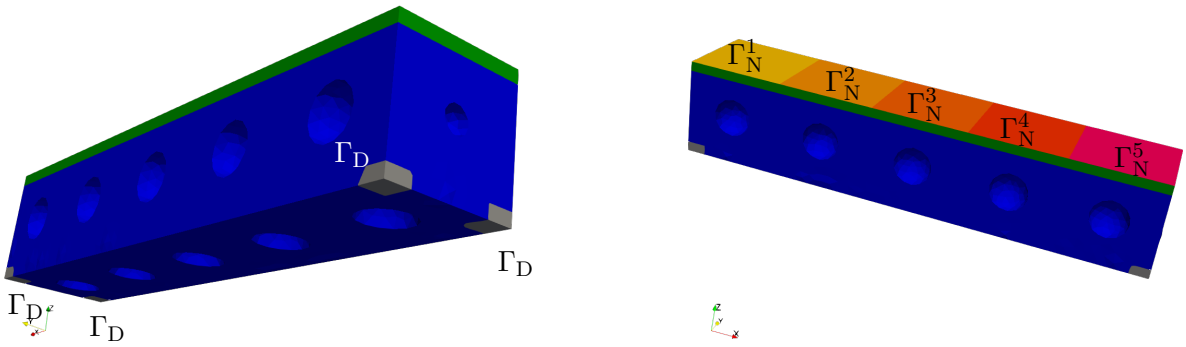


Figure 5.9: Structure of the bridge. The non-optimizable supports of the bridge are marked in grey and their lower surface Γ_D is where Dirichlet are applied. The green block is non-optimizable as well, and on its upper surface five random mechanical loads are applied on the sections $\Gamma_N^1, \dots, \Gamma_N^5$.

We suppose that $\mathbf{X} = [X_1, \dots, X_5]$ is a Gaussian random vector with covariance matrix equal to the identity. Once again, we consider two cases. In the first one the random variables X_i are centered in 0, and the techniques of section 5.5.2 are used to evaluate the constraint functional and its derivative. In the second case we suppose that all random variables X_i to have a mean equal to -1.0, implying that an average compression load of 1.0 MPa is applied on each of the five sections of the bridge. We consider the shape shown in fig. 5.9 as initial condition. The optimized shapes for the centered and noncentered cases are reported in fig. 5.10 and fig. 5.11 respectively. The trends of the objective and the constraint are presented in fig. 5.12a and fig. 5.12b, and the numerical results for both problems are collected in table 5.6.

Geometry of the structure		
Longitudinal length	ℓ_x	5.0 cm
Cross section length	ℓ_y	1.0 cm
Height	ℓ_z	1.0 cm
Sidelength of Γ_{D}		0.2 cm
Sidelength of each Γ_{N}^i		1.0 cm
Mesh size parameters		
Minimal mesh size	`hmin`	0.10 cm
Maximal mesh size	`hmax`	0.05 cm
Elastic coefficients		
Young's modulus	E	200 MPa
Poisson's ratio	ν	0.3
Mechanical loads		
Vertical load	$\overline{g}_i$	1 MPa
Thresholds for the inequality constraints		
Threshold on the compliance	τ	1×10^{-1} MPa cm^3
Bound on the probability of failure	$\bar{p}$	1.0 %

Table 5.5: Numerical data concerning the geometry and the mechanics of the bridge of fig. 5.9.

As for the cantilever in section 5.6.1, these results illustrate that the constraint on the probability of failure is upheld for the centered and for the noncentered case, validating both methods. Moreover, fig. 5.12a shows that the convergence of the objective function is faster for the bridge than the cantilever. Comparing the shapes presented in fig. 5.10 and fig. 5.11 we remark that structure optimized for loads centered in 0 is lighter and thinner than the structure optimized for a load whose expected value is not equal to zero. Such result can be expected since, on average, the pressure exerted on the first structure is less than the pressure acting on the second one, and a thinner structure is enough to ensure the enforcement of the constraint on the probability for the compliance to exceed the threshold τ.

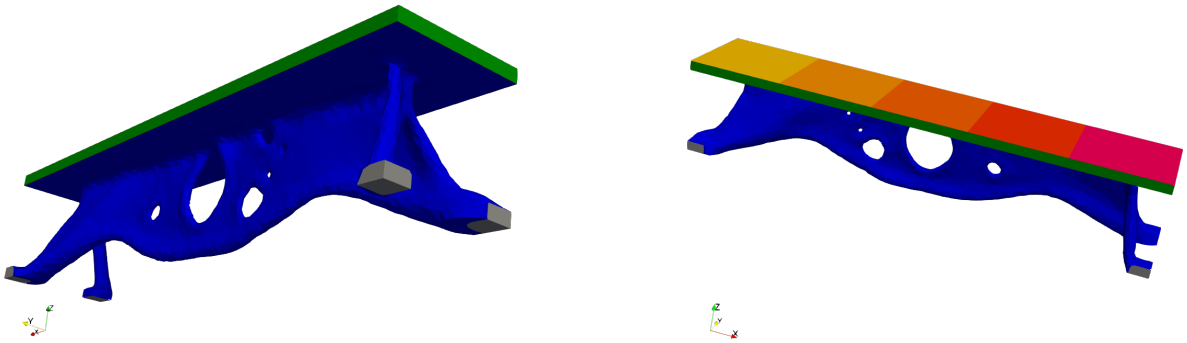


Figure 5.10: Optimal shape for the bridge subject to vertical loads centered on 0.

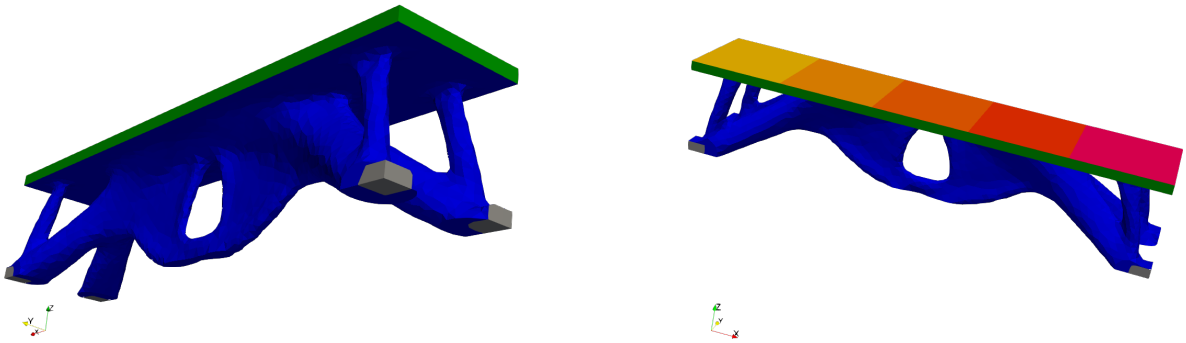
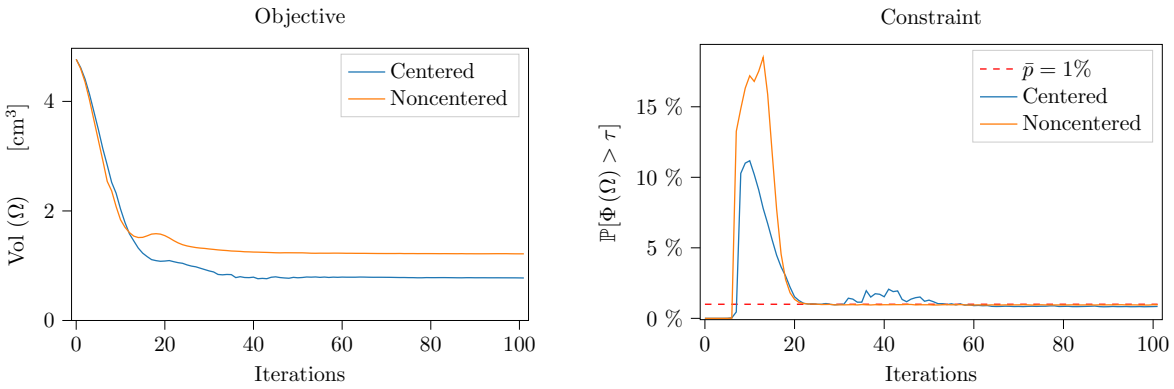


Figure 5.11: Optimal shape for the bridge subject to vertical loads, with an average compression of 1.0 MPa.



(a) Evolution of the objective function.

(b) Evolution of the constraint.

Figure 5.12: Convergence of the objective and the constraints for the optimization problem of the 3D bridge.

Bridge optimization		**Centered**	**Noncentered**
Execution			
Number of iterations		100	100
Execution time	[min]	138	127
Results			
Final volume $\mathrm{Vol}\,(\Omega_{\mathrm{opt}})$	[cm^3]	0.776	1.217
Excess probability $\mathbb{P}\left[\mathcal{Q}\left(\mathbf{u}_\Omega(\omega), \Omega\right) > \tau\right]$		0.979%	0.961%

Table 5.6: Numerical results for the optimization of the volume of a bridge subject to five uncertain mechanical loads on its upper surface.

5.7 Conclusions and perspectives

In this chapter we presented a method to solve Reliability-Based Topology Optimization problems for elastic structures using a level-set representation and Hadamard's approach to shape derivatives. We limited our study to the differentiation of quadratic functionals of the displacement. After having stated the optimization problem and the properties of the safety criterion, we define a shape functional $\Phi(\cdot)$ measuring the probability to violate the criterion when the random external loads follow a given probability distribution. We assumed also that the random behavior of the external loads can me parametrized by a finite number N of random variables. We computed the shape derivative of such functional by interpreting the set of situations when the criterion is satisfied as an ellipsoid in $\mathbb{R}^N$ and integrating suitable functions of the probability density on its surface.

We aimed to avoid the numerical computation of the integral of the density function on the surface of the N-dimensional ellipse, which can be computationally expensive and requires great accuracy if the tolerance on the probability to exceed the safety criterion is low. Therefore, we tested numerically the expression of the shape derivative using suitable quadrature formulas for centered Gaussian distribution, as presented in appendix C. In section 5.4 we provided a different method to compute and differentiate the functional $\Phi(\cdot)$ in the case where the uncertainties can be parametrized as non-centered Gaussian random variables. In order to do so, we applied the series expansion of the cumulative distribution function of generalized chi-squared random variables as proven by Ruben in [214]. The reliability of both methods has been tested in 5.6 for the optimization of two structures, subject to a constraint on the probability for the mechanical compliance to exceed a threshold. One of the main advantages of this approach consists in the fact that the computation of integrals on high dimensional domains is avoided. Instead an analytic expression for $\Phi(\Omega)$ and its derivative is used, which is extremely fast to evaluate numerically, and can be computed up to an arbitrary precision.

The approaches detailed in this chapter are well-adapted for quadratic functionals of the state of the optimization problem, like the mechanical compliance or the L^2-norm of the stress. However, the study of polynomial functionals of higher order is of particular interest, since they can better approximate quantities like the maximal concentration of mechanical stress in the structure. Thus, a future work could focus on the extension to more complex constraints.

A concept strictly related to the probability for a quantity $\mathcal{Q}(\mathbf{g}, \Omega)$ to exceed a threshold τ is the notion of *Value at Risk*. Let us consider a quantity $\bar{p} \in (0, 1)$. The *Value at Risk* at level $\bar{p}$ for the random variable $\mathcal{Q}(\mathbf{g}, \Omega)$ is defined as the minimal $V \in \mathbb{R}$ such that the probability of $\mathcal{Q}(\mathbf{g}, \Omega)$ to exceed V is $1 - \bar{p}$

$$\mathbb{P}\left[\mathcal{Q}(\mathbf{g}, \Omega) > \mathrm{VaR}_{\bar{p}}\right] \geq 1 - \bar{p}.$$

By definition, imposing that the value at risk at level $\bar{p}$ should not exceed a threshold τ is equivalent to impose that the probability for $\mathcal{Q}(\mathbf{g}, \Omega)$ to exceed τ should be below $\bar{p}$. However, being able to differentiate $\mathrm{VaR}_{\bar{p}}$ with respect to the shape would allow to consider it as the objective in an optimization problem.

In this chapter we studied the differentiation of the failure probability for quadratic functionals of the displacement, without any assumption on the size of the uncertainties. In [11], Allaire and Dapogny proposed a linearization approach to estimate and differentiate the failure probability of a generic functional $\mathcal{H}(\mathbf{g}, \Omega)$ at least differentiable with respect to $\mathbf{g}$. The present method can be adapted to replace the linearized functional with a quadratic approximation, under the hypotheses of small perturbations and twice differentiability of $\mathcal{H}(\mathbf{g}, \Omega)$ with respect to $\mathbf{g}$.

Conclusions and perspectives

Innovative techniques of shape and topology optimization have found their application in a large variety of fields, ranging from structural mechanics to aerodynamics, and from the design of components in microelectronics to the study of complex lattice structures for composite materials. However, in any industrial context, the performance of a structure should not be measured only with respect to its behavior in a baseline situation, but also on its robustness with respect to altered external conditions or perturbations in its manufacturing.

The main objective of this thesis has been the study of different approaches to the inclusion of uncertainties in the optimization of elastic structures. In particular, we focused on the case where the uncertainties lie in the mechanical loads applied to the structure. We considered shape optimization problems where the objective is the minimization of the volume under constraints dependent on the uncertain loads.

In the introductory chapter 1 we presented the main tools and concepts used in this thesis. We started from general recalls on optimization problems and gradient-based algorithms to solve them. Next, we presented Hadamard's method to differentiate a function with respect to the shape and we provided the expression of some generic functionals. Then, we focused on the differentiation of shape functionals in an integral form, whose value depends on the solution of an elliptic PDE. We provided a general formula for the shape derivative of this kind of functionals, and we proved it by computing the Lagrangian derivative of the solution of the PDE, and by Céa's fast derivation method. Finally, we discussed in details the different aspects of the optimization procedure adopted in this work.

In the first section of chapter 2 we recalled the equations of linear elasticity and the definition of the mechanical compliance and the von Mises stress. In section 2.2 we considered the optimization of an elastic structure subject to a time-dependent thermal field, under a constraint on the mechanical compliance. The constraint functional chosen in our example takes into account both the average of the compliance in the time interval of the simulation, and its value at the final instant. In section 2.2.3 we provided an expression for the shape derivative of the constraint functional. We showed also that the adjoint to the temperature solves a differential equation backwards in time, and that the weak coupling of the adjoints is reversed with respect to the coupling of the temperature and displacement fields.

These results ought to be considered as part of an ongoing project aimed at the study of shape optimization problems of elastic structures subject to uncertain and time-variant temperature fields. Another direction of research on the subject could be the study the sensitivity of the thermal compliance with respect to the different parameters of the problem. Finally, the model can be enriched by considering other phenomena like the occurring of thermal radiation or the dependence of the material properties from the temperature of the structure.

The chapters in part II presented three different ways to take into account uncertainties

and perturbations on the external loads while solving shape optimization problems. All shape optimization problems considered in part II had the following structure

$$
\left|\begin{array}{l}
\text{Find the admissible shape } \Omega \in \mathcal{S}_{\text{adm}} \\
\text{minimizing the volume } \operatorname{Vol}(\Omega) \\
\text{under the constraint } \mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] \leq \tau, \\
\text{where the displacement } \mathbf{u}_\Omega(\omega) \in \mathrm{H}^1_{\Gamma_\mathrm{D}}(\Omega)^d \\
\text{solves the elasticity equation} \\
\quad \left\{\begin{array}{rcll}
-\operatorname{div}\left(\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\right) & = & \mathbf{f} & \text{in } \Omega, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{g}(\omega) & \text{on } \Gamma_\mathrm{N}, \\
\boldsymbol{\sigma}\left(\mathbf{u}_\Omega(\omega)\right)\mathbf{n} & = & \mathbf{0} & \text{on } \Gamma_0, \\
\mathbf{u}_\Omega(\omega) & = & \mathbf{0} & \text{on } \Gamma_\mathrm{D},
\end{array}\right. \\
\text{for almost all event } \omega \in \mathcal{O}.
\end{array}\right.
$$

The operator $\mathcal{F}[\,\cdot\,]$ associates a deterministic quantity to the random variable $H(\mathbf{u}_\Omega, \Omega)$, and different approaches lead to different choices of $\mathcal{F}[\,\cdot\,]$. A common feature to all techniques considered in part II is the avoidance of sampling methods to estimate the impact of the uncertainties on the objective and constraint functionals, since they can be extremely costly from a computational point of view.

In chapter 3, we studied optimization problems under a constraint on the worst-case scenario, so that the constraint reads $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right] = \operatorname{ess\,sup}_{\omega\in\mathcal{O}} H(\mathbf{u}_\Omega(\omega), \Omega)$. We proposed two different techniques. In both cases we supposed that the external loads could be parametrized as elements of a finite-dimensional compact and convex set $\mathcal{G}$.

The first method, introduced in section 3.2, applies only to convex functionals of the uncertain parameters, and consists in approximating $\mathcal{G}$ by a polyhedron $\mathcal{G}_N$, thus transforming the original optimization problem into a problem with multiple deterministic constraints. The *null space optimization* algorithm is well suited to the solution of multi-constrained problems, as shown in [114] and in the simulations of section 3.4. Moreover, we have been able to show numerically and theoretically the convergence of the approximated solution with N constraints towards the exact solution, when N tends to infinity and $\mathcal{G}_N$ converges towards $\mathcal{G}$ with respect to the Hausdorff distance. However, the number of vertices required for the polyhedron $\mathcal{G}_N$ to properly represent $\mathcal{G}$ increases rapidly with the dimension of $\mathcal{G}$. Therefore, this method is not viable if dimension of the set of admissible loads is too high.

The second technique, discussed in section 3.3, relies on the computation of one element of the subdifferential $\partial H(\mathbf{u}_\Omega, \Omega)$. The numerical simulations of section 3.4 show that the approach based on the subdifferential is faster than the method based on the polyhedral approximation of $\mathcal{G}$, since it requires fewer computations of shape derivatives. However, we remarked that the technique of the subdifferential suffers from wide oscillations impairing the convergence when the constraint functional is maximized by multiple elements of $\mathcal{G}$. Further research could focus on the characterization of the subdifferential of the constraint and on the design of an optimization algorithm capable to take into account multiple elements of $\partial H(\mathbf{u}_\Omega, \Omega)$.

RTO are the main focus of chapter 4, meaning that $\mathcal{F}\left[H(\mathbf{u}_\Omega, \Omega)\right]$ has been taken as the expectation of $H(\mathbf{u}_\Omega, \Omega)$. Our main objective has been the extension of the results obtained by Dambrine, Dapogny and Harbrecht in [80] to polynomial functions of degree m of the uncertain

loads. We were able to compute a tensor formulation for the optimization problem by introducing a suitable correlation tensor of order m, proving that the expectation of an m-multilinear function depends on the m first stochastic moments of the perturbation. However, such expression is impossible to solve numerically as a consequence of the high dimensionality of the tensor space. Therefore, we limited our study to perturbations of finite rank N.

In section 4.4 we tested our approach for the optimization of a structure under a constraint on the L^6-norm of the von Mises stress, and in section 4.5 we showed how this method can be used to compute the shape derivative of the variance of a quadratic functional. Similarly to the first approach of chapter 3, this method suffers from the curse of dimensionality. Indeed, the number of components the shape derivative grows exponentially with the rank N of the perturbation due to the presence of numerous cross terms. One possible approach to this issue is the application of tensor decomposition techniques to the correlation tensor, in order to isolate the components responsible for the largest part of the constraint functional. This method would allow to reduce the number of components to take into account in the expression of the shape derivative by accepting a systematic error in the evaluation of the constraint.

A different strategy to be investigated consists in associating a weight to each of the N components of the perturbation. Since all terms of the constraint functional are obtained by combining m components of the random load, each one can be associated to a weight defined as the product of the weights of its parts. By limiting the computation of the constraint functional (as well as its derivative) to the terms with a weight larger than a given tolerance, it should be possible to reduce the complexity of the expression while controlling the error.

Finally, in chapter 5 we focused on RBTO problems for quadratic functionals of the state, where the applied loads are subject to perturbations of finite rank N. This approach is particularly interesting if the aim of the optimization is to be able to control the failure probability of a given structure. The expression of $\mathcal{F}[\,\cdot\,]$ chosen for this last case has been $\mathcal{F}[H(\mathbf{u}_\Omega, \Omega)] = \mathbb{P}[H(\mathbf{u}_\Omega, \Omega) > \tau]$.

In section 5.3 we provided a generic formula for the derivative of the failure probability for the case where the random variable describing the uncertain loads admits a density function. The coefficients of the several components of the shape derivative have been computed as integrals of the probability density function on the surface of an N-dimensional ellipsoid. In section 5.4 we provided an alternative technique to compute and differentiate the failure probability for the case of Gaussian perturbations, avoiding numerical integration in large dimension. This new approach relied on the series expansion of the cumulative distribution function of a non-central chi-squared random variable proven by Ruben [214]. This method scaled well with the rank of the perturbations since the shape derivative of the failure probability can be written as the weighted sum of just $(N+1)^2$ derivatives. The weights of each term have been computed by an inexpensive recursive relation.

The simulations of section 5.6 show the efficacy and the accuracy of both approaches in multiple cases. As an example, we cite the comparison between the results of section 3.4.1 and section 5.6.1. In both cases we aimed to minimize the volume of a 3D cantilever under a constraint on the mechanical compliance. In the first case the constraint applied on the worst-case scenario, while in the second we imposed that the event of the compliance exceeding a threshold τ should happen with a probability lower than 1%. The two approaches yielded very similar results, but the probabilistic approach has been slightly faster and its performance would not be impaired by the addition of further uncertain loads. Possible directions of further research in the RBTO approach consist in the search for results analogous to the ones of section 5.3 for functionals other than quadratic. Other possible developments include the computation of the

shape derivative of the *Value at Risk*, and the extension of the linearization approach of Allaire and Dapogny [11] to include a second order approximation of the functional of interest.

The main objective of the present thesis has been the study of different approaches to address the presence of uncertainties in the solution of shape and topology optimization problems. The methods proposed in this work have been presented and tested for elastic structures. However, their scope is larger than the context of linear elasticity, and future works on this subject may focus on the implementation of the aforementioned techniques in a wider array of situations. As stated in chapter 2, some studies have already been launched concerning the application of RTO techniques to elastic structures subject to time-dependent and uncertain thermal perturbations.

All methods discussed in part II have been analyzed from the theoretical point of view, and tested on academic examples. However, in engineering applications, the constraints on uncertain functionals would have to concur with other kinds of constraints concerning various aspects of the structure, as its manufacturability, or the interactions with its surroundings. A natural extension of this work is the application of the methods detailed in this work to structures with an industrial interest, and the development of numerical tools to deal with the increased complexity of the problem. Some preliminary study on the application of the aforementioned techniques to mechanical pieces is currently ongoing at Safran Helicopter Engines.

Appendix

Appendix A

Some shape calculus formulas

In this appendix we recall some formulas and propositions which have been used throughout this thesis for the computation of shape derivatives. For the entirety of this appendix we will suppose that Ω is a Lipschitz continuous domain in $\mathbb{R}^d$, and $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ a Lipschitz continuous displacement field such that $\|\boldsymbol{\theta}\|_{1,\infty} < 1$.

Contents

A.1 Shape derivatives on variable domains

In this section we recall the results of proposition 1.8 and proposition 1.12.

Proposition A.11 (Shape derivative of a volume function)**.** *Let $f \in \mathrm{W}^{1,1}\left(\mathbb{R}^d\right)$ be a real valued function, and $\mathcal{S}_{\mathrm{adm}}$ a class of Lipschitz continuous domains in $\mathbb{R}^d$. Let $\Omega \in \mathcal{S}_{\mathrm{adm}}$ be a domain such that, for any admissible displacement field $\boldsymbol{\theta} \in \Theta_{adm}$ the perturbed domain $\Omega_{\boldsymbol{\theta}}$ belongs to $\mathcal{S}_{\mathrm{adm}}$. Then, the shape functional $\Phi : \Omega \mapsto \int_\Omega f(\mathbf{x})\mathrm{d}\mathbf{x}$ is shape differentiable in Ω and its derivative can be written as*

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_\Omega \mathrm{div}\left(\widehat{\boldsymbol{\theta}}(\mathbf{x})f(\mathbf{x})\right)\,\mathrm{d}\mathbf{x}.$$

If, moreover, Ω has a Lipschitz continuous boundary, the shape derivative of Φ can be also expressed as

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_{\partial\Omega}\left(\mathbf{n}(\mathbf{s})\cdot\widehat{\boldsymbol{\theta}}(\mathbf{s})\right)f(\mathbf{s})\,\mathrm{d}\mathbf{s}.$$

This result is reported and proven in section 1.2.1 as proposition 1.8.

Proposition A.12 (Shape derivative of a boundary functional)**.** *Let Ω be a $\mathcal{C}^2$ domain with Lipschitz continuous boundary, and $g \in \mathrm{W}^{2,1}\left(\mathbb{R}^d\right)$ a real valued function. Then, the shape*

functional $\Phi : \Omega \mapsto \int_{\partial\Omega} g(\mathbf{s})\mathrm{d}\mathbf{s}$ *is shape differentiable in* Ω*, and its derivative is expressed as*

$$\mathrm{D}\Phi(\Omega)(\widehat{\boldsymbol{\theta}}) = \int_{\partial\Omega} \left(\nabla g \cdot \widehat{\boldsymbol{\theta}} + g \operatorname{div}_\Gamma \widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{s}. = \int_{\partial\Omega} \left(\mathbf{n}(\mathbf{s}) \cdot \widehat{\boldsymbol{\theta}}(\mathbf{s})\right) \left(\frac{\partial g}{\partial \mathbf{n}} + Hg\right) \mathrm{d}\mathbf{s}.$$

The operator $\operatorname{div}_\Gamma$ denotes the tangential divergence defined in definition 1.9. The quantity $H : \partial\Omega \to \mathbb{R}$ is the mean curvature of the surface, and is introduced in definition 1.10. The aforementioned result is reported and proven in [6, Proposition 6.24] and [138, Proposition 5.4.18].

A.2 Differential operators under a change of variables

Here we list the expression of different operators under a change of variables. Let us consider a Lipschitz continuous vector field $\boldsymbol{\theta} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$. The mapping defined as $\Phi_{\boldsymbol{\theta}} : \mathbf{x} \mapsto \mathbf{y} = (\mathbf{I} + \boldsymbol{\theta}(\mathbf{x}))\mathbf{x}$ is a diffeomorphism in $\mathbb{R}^d$. In this section we denote by an index "$\mathbf{y}$" the differential operators computed with respect to the perturbed reference system.

- **Gradient of a scalar function.** Let us consider $\varphi \in \mathcal{C}^1\left(\mathbb{R}^d\right)$ a continuous scalar function. Then
$$\nabla_{\mathbf{y}} \left(\varphi \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) = (\mathbb{I} + \nabla\boldsymbol{\theta})^{-\mathrm{T}} \nabla\varphi$$
where $\nabla\varphi$ is a column vector, and $\nabla\boldsymbol{\theta} = [\nabla\boldsymbol{\theta}_1| \cdots |\nabla\boldsymbol{\theta}_N]^{\mathrm{T}}$.

- **Product of gradients.** Let $\varphi, \psi \in \mathcal{C}^1\left(\mathbb{R}^d\right)$ be two continuous scalar functions. Then
$$\nabla_{\mathbf{y}} \left(\varphi \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) \cdot \nabla_{\mathbf{y}} \left(\psi \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) = (\mathbf{A}_{\boldsymbol{\theta}} \nabla\varphi) \cdot \nabla\psi,$$
where the matrix $\mathbf{A}_{\boldsymbol{\theta}} \in \mathbb{R}^{d\times d}$ is defined as
$$\mathbf{A}_{\boldsymbol{\theta}} = (\mathbb{I} + \nabla\boldsymbol{\theta})^{-1} (\mathbb{I} + \nabla\boldsymbol{\theta})^{-\mathrm{T}}.$$

- **Gradient of a vector function.** Let $\mathbf{v} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$ be a smooth vector field. Then
$$\nabla_{\mathbf{y}} \left(\mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) = \nabla\mathbf{v}\, (\mathbb{I} + \nabla\boldsymbol{\theta})^{-1},$$
where $\nabla\mathbf{v} = [\nabla\mathbf{v}_1| \cdots |\nabla\mathbf{v}_N]^{\mathrm{T}}$.

- **Divergence of a vector.** We consider a smooth vector field $\mathbf{v} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$. Then
$$\operatorname{div}_{\mathbf{y}} \left(\mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) = (\mathbb{I} + \nabla\boldsymbol{\theta})^{-\mathrm{T}} : \nabla\mathbf{v}.$$

- **Product of gradients via an elasticity tensor.** Let $\mathbf{v}, \mathbf{w} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$ be smooth vector fields, and $\mathbb{C}$ a fourth order tensor satisfying the assumptions of section 2.1.1. Then
$$\mathbb{C}\nabla_{\mathbf{y}} \left(\mathbf{v} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) : \nabla_{\mathbf{y}} \left(\mathbf{w} \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}\right) = \mathbb{C} \left(\nabla\mathbf{v}(\mathbb{I} + \nabla\boldsymbol{\theta})^{-1}\right) : \left(\nabla\mathbf{w}(\mathbb{I} + \nabla\boldsymbol{\theta})^{-1}\right).$$

- **Volume integral.** Let $\varphi \in \mathrm{L}^1(\Omega)$ be an integrable real function, and $\Omega_{\boldsymbol{\theta}} = \Phi_{\boldsymbol{\theta}}(\Omega)$ a deformed Lipschitz continuous domain. Then

$$\int_{\Omega_{\boldsymbol{\theta}}} \varphi \circ (\mathbf{I} + \boldsymbol{\theta})^{-1}(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{\Omega} \varphi(\mathbf{x})\ |\det(\mathbb{I} + \nabla\boldsymbol{\theta})|\ \mathrm{d}\mathbf{x}.$$

- **Surface integral.** Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a $\mathcal{C}^1$ surface, and $\psi \in \mathrm{L}^1(\partial\Omega)$ a real-valued function defined on the surface of Ω. Once again, we denote $\Omega_{\boldsymbol{\theta}} = \Phi_{\boldsymbol{\theta}}(\Omega)$ the domain obtained by the deformation $\Phi_{\boldsymbol{\theta}}$, and $\partial\Omega_{\boldsymbol{\theta}}$ its surface. Then

$$\int_{\partial\Omega_{\boldsymbol{\theta}}} \psi \circ (\mathbf{I} + \boldsymbol{\theta})^{-1} = \int_{\partial\Omega} \psi(\mathbf{s})\, \mathrm{Jac}_{\partial\Omega}(\mathbf{I} + \boldsymbol{\theta})\ \mathrm{d}\mathbf{s},$$

  where the Jacobian term is defined on each point $\mathbf{s}$ of $\partial\Omega_{\boldsymbol{\theta}}$ as

$$\mathrm{Jac}_{\partial\Omega}(\mathbf{I} + \boldsymbol{\theta})(\mathbf{s}) = \left\| (\mathbb{I} + \nabla\boldsymbol{\theta})^{-\mathrm{T}}\, \mathbf{n}(\mathbf{s}) \right\| |\det(\mathbb{I} + \nabla\boldsymbol{\theta})|\,.$$

A.3 Shape derivatives of common expressions

Let $\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ be Lipschitz continuous vector fields defined on $\mathbb{R}^d$. We collect in this section the expression of the derivatives with respect to $\boldsymbol{\theta}$ of quantities appearing frequently in the computation of Hadamard shape derivatives.

- $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\left((\mathbb{I} + \nabla\boldsymbol{\theta})^{-1}\right)\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = -\nabla\widehat{\boldsymbol{\theta}}$.
- $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\left(|\det(\mathbb{I} + \nabla\boldsymbol{\theta})|\right)\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = -\nabla\cdot\widehat{\boldsymbol{\theta}}$ is the derivative of the Jacobian term appearing in the expression of a volume integral under change of coordinates.
- $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\left(\mathrm{Jac}_{\partial\Omega}(\mathbf{I} + \boldsymbol{\theta})\right)\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = \mathrm{div}_{\partial\Omega}\,\widehat{\boldsymbol{\theta}} = \mathrm{div}\,\widehat{\boldsymbol{\theta}} - (\nabla\widehat{\boldsymbol{\theta}}\,\mathbf{n})\cdot\mathbf{n}$ is the derivative of the Jacobian for surface integrals under change of coordinates, and $\mathbf{n}$ identifies the unit vector normal to the surface. This identity is proven in [138, Lemma 5.4.15].
- $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\left(\mathbb{C}\left(\nabla\mathbf{v}(\mathbb{I} + \nabla\boldsymbol{\theta})^{-1}\right) : \left(\nabla\mathbf{w}(\mathbb{I} + \nabla\boldsymbol{\theta})^{-1}\right)\right)\Big|_{\boldsymbol{\theta}=0}(\widehat{\boldsymbol{\theta}}) = -\mathbb{C}\nabla\mathbf{w} : (\nabla\mathbf{v}\,\nabla\widehat{\boldsymbol{\theta}}) - \mathbb{C}\nabla\mathbf{v} : (\nabla\mathbf{w}\,\nabla\widehat{\boldsymbol{\theta}})$, where $\mathbf{v}, \mathbf{u} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$ are smooth vector fields, and $\mathbb{C}$ is a fourth order tensor respecting the constraints of Hooke's elasticity as listed in section 2.1.1.

A.4 Integration by parts

Let $\Omega \subset \mathbb{R}^d$ be a $\mathcal{C}^1$ domain, and let us denote $\mathbf{n}(\mathbf{s})$ the unitary vector normal to the surface of the domain in $\mathbf{s} \in \partial\Omega$. As proven by Hadamard's structure theorem (see theorem 1.7), the shape derivative of a differentiable function on a $\mathcal{C}^1$ domain $\Omega \subset \mathbb{R}^d$ evaluated in $\widehat{\boldsymbol{\theta}} \in \mathrm{W}^{1,\infty}\left(\mathbb{R}^d\right)^d$ depends only on the normal component of $\widehat{\boldsymbol{\theta}}$ on $\partial\Omega$. For each $\mathbf{s} \in \partial\Omega$, we denote $\widehat{\boldsymbol{\theta}}_\tau(\mathbf{s})$ the component of $\widehat{\boldsymbol{\theta}}(\mathbf{s})$ tangent to $\partial\Omega$ in $\mathbf{s}$. Thus, on the surface of the domain Ω, the vector $\widehat{\boldsymbol{\theta}}(\mathbf{s})$ can be decomposed as $\widehat{\boldsymbol{\theta}}_\tau + \mathbf{n}(\widehat{\boldsymbol{\theta}}\cdot\mathbf{n})$.

In this section we report the results of the integrations by parts of some expressions appearing in the computation of shape derivatives. In particular, we express each integral as the sum of a volume term, and two surface terms depending on the tangent and normal components of $\widehat{\boldsymbol{\theta}}$ respectively.

- Let $\varphi, \psi \in \mathcal{C}^1(\Omega)$ be two continuous and differentiable functions. Then

$$\int_\Omega \varphi(\mathbf{x}) \operatorname{div}\left(\psi\,\widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{x} = -\int_\Omega (\widehat{\boldsymbol{\theta}} \cdot \nabla\varphi)\,\psi(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} \varphi(\mathbf{s})\,\psi(\mathbf{s})\,(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{s}.$$

- Let $\varphi, \psi \in \mathcal{C}^2(\Omega)$ be two continuous and twice differentiable functions. Then

$$\begin{aligned}\int_\Omega \left(\nabla\widehat{\boldsymbol{\theta}}\,\nabla\varphi\right) \cdot \nabla\psi\,\mathrm{d}\mathbf{x} = -\int_\Omega \widehat{\boldsymbol{\theta}} \cdot \operatorname{div}\left(\nabla\psi \otimes \nabla\varphi\right) \mathrm{d}\mathbf{x} \\ + \int_{\partial\Omega} \frac{\partial\varphi}{\partial\mathbf{n}} \left(\nabla\psi \cdot \widehat{\boldsymbol{\theta}}_\tau\right) \mathrm{d}\mathbf{s} + \int_{\partial\Omega} \frac{\partial\varphi}{\partial\mathbf{n}} \frac{\partial\psi}{\partial\mathbf{n}} \,(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{s}.\end{aligned}$$

- Let us consider two vector-valued functions $\mathbf{f}, \mathbf{g} \in \mathcal{C}^2\left(\mathbb{R}^d\right)^d$ and a fourth-order tensor $\mathbb{C}$. Then

$$\begin{aligned}\int_\Omega \left(\mathbb{C}\nabla\mathbf{f} : (\nabla\mathbf{g}\nabla\widehat{\boldsymbol{\theta}})\right) \mathrm{d}\mathbf{x} = -\int_\Omega \widehat{\boldsymbol{\theta}} \cdot \operatorname{div}\left(\nabla\mathbf{g}^{\mathrm{T}}\,\mathbb{C}\nabla\mathbf{f}\right) \mathrm{d}\mathbf{x} \\ + \int_{\partial\Omega} (\mathbb{C}\nabla\mathbf{f}\,\mathbf{n}) \cdot (\nabla\mathbf{g}^{\mathrm{T}}\,\widehat{\boldsymbol{\theta}}_\tau)\,\mathrm{d}\mathbf{s} + \int_{\partial\Omega} (\mathbb{C}\nabla\mathbf{f}\,\mathbf{n}) \cdot \frac{\partial\mathbf{g}}{\partial\mathbf{n}}\,(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{s}.\end{aligned}$$

- Let us consider a $\mathcal{C}^2$ vector valued function $\mathbf{f}$, and a scalar function $\varphi \in \mathcal{C}^1\left(\mathbb{R}^d\right)$. Then

$$\begin{aligned}\int_\Omega \varphi(\mathbf{x}) \left(\nabla\mathbf{f}^{\mathrm{T}} : \nabla\widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{x} = -\int_\Omega \widehat{\boldsymbol{\theta}} \cdot \operatorname{div}\left(\varphi\,\mathbf{f}^{\mathrm{T}}\right) \mathrm{d}\mathbf{x} \\ + \int_{\partial\Omega} \left(\widehat{\boldsymbol{\theta}}_\tau \cdot \frac{\partial\mathbf{f}}{\partial\mathbf{n}}\right) \varphi\,\mathrm{d}\mathbf{s} + \int_{\partial\Omega} \left(\mathbf{n} \cdot \frac{\partial\mathbf{f}}{\partial\mathbf{n}}\right) \varphi\,(\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{s}.\end{aligned}$$

- Let $\mathbf{f}, \mathbf{g} \in \mathcal{C}^1\left(\mathbb{R}^d\right)^d$ be two continuous and differentiable vector functions. Then

$$\int_\Omega \mathbf{g}\operatorname{div}\left(\mathbf{f} \otimes \widehat{\boldsymbol{\theta}}\right) \mathrm{d}\mathbf{x} = -\int_\Omega \widehat{\boldsymbol{\theta}} \cdot \left(\nabla\mathbf{g}^{\mathrm{T}}\mathbf{f}\right) \mathrm{d}\mathbf{x} + \int_{\partial\Omega} (\mathbf{g} \cdot \mathbf{f})\ (\widehat{\boldsymbol{\theta}} \cdot \mathbf{n})\,\mathrm{d}\mathbf{s}.$$

Appendix B

Alternative proof of proposition 5.2

In this annex we provide an alternative proof of proposition 5.2 which does not rely on lemma 5.1.

Alternative proof of proposition 5.2. As first step, we consider the expression 5.11 of the functional $\Phi(\cdot)$, and we apply the change of variables $\mathbf{y} = \frac{1}{\sqrt{\tau}}{\mathbf{M}_\Omega}^{1/2}\mathbf{x}$ in order to express $\Phi(\cdot)$ on the unit sphere:

$$\Phi(\Omega) = \int_{\mathcal{E}_{\mathbf{M}_\Omega,\tau}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\,\mathrm{d}\mathbf{y}. \tag{B.1}$$

The shape derivative of equation (B.1) can be written as the sum of two terms, that we denote $A(\boldsymbol{\theta})$ and $B(\boldsymbol{\theta})$.

$$\begin{aligned}\mathrm{D}\Phi(\Omega)(\boldsymbol{\theta}) &= \mathrm{D}\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\mathrm{d}\mathbf{y}(\boldsymbol{\theta}) \\ = \underbrace{\mathrm{D}\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}}(\boldsymbol{\theta}) \int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\,\mathrm{d}\mathbf{y}}_{A(\boldsymbol{\theta})} &+ \underbrace{\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} \mathrm{D}f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)(\boldsymbol{\theta})\mathrm{d}\mathbf{y}}_{B(\boldsymbol{\theta})}.\end{aligned}$$

We develop $A(\boldsymbol{\theta})$ and $B(\boldsymbol{\theta})$ separately. In order to express $A(\boldsymbol{\theta})$, we apply Jacobi's formula for the determinant of an invertible matrix $\mathbf{M}$ [174]):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\det \mathbf{M}(t)\right] = \left(\det \mathbf{M}(t)\right) \mathrm{tr}\left(\mathbf{M}^{-1}(t)\,\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}(t)\right). \tag{B.2}$$

Therefore, we have

$$\begin{aligned}A(\boldsymbol{\theta}) &= \tau^{N/2}\left(\int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\,\mathrm{d}\mathbf{y}\right)\mathrm{D}(\det \mathbf{M}_\Omega)^{-1/2}(\boldsymbol{\theta}) \\ &= -\frac{\tau^{N/2}}{2}\left(\int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\,\mathrm{d}\mathbf{y}\right)(\det \mathbf{M}_\Omega)^{-3/2}\,\mathrm{D}(\det \mathbf{M}_\Omega)(\boldsymbol{\theta}) \\ &= -\frac{1}{2\left(\det \mathbf{M}_\Omega\right)}\Phi(\Omega)\,\mathrm{tr}\left({\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right) = -\frac{1}{2}\Phi(\Omega)\,\mathrm{tr}\left({\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right).\end{aligned} \tag{B.3}$$

For the term $B(\boldsymbol{\theta})$, we remark the following identity:

$$\begin{aligned}&\mathrm{div}_{\mathbf{y}}\left(f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right){\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\mathbf{y}\right) \\ =&\sqrt{\tau}\left(\nabla_{\mathbf{y}} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\right)\cdot\left(\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\mathbf{y}\right) \\ &+ f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right)\mathrm{tr}\left({\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\right).\end{aligned} \tag{B.4}$$

Thus, using equation (B.4) and Green's formula, and observing that on the surface of the unit sphere $\mathbb{S}_{N-1}$ the unit normal vector in $\mathbf{y} \in \mathbb{S}_{N-1}$ is $\mathbf{y}$ itself, we can decompose $B(\boldsymbol{\theta})$ in two terms:

$$\begin{aligned} B(\boldsymbol{\theta}) =& \sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} \nabla_{\mathbf{y}} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right) \cdot \left(\sqrt{\tau}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\mathbf{y}\right) \mathrm{d}\mathbf{y} \\ =& \underbrace{\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{S}_{N-1}} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{s}\right) \left({\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\mathbf{s}\right) \cdot \mathbf{s}\, \mathrm{d}\mathbf{s}}_{B_1(\boldsymbol{\theta})} \\ & \underbrace{-\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}}\, \mathrm{tr}\left({\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\right) \int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right) \mathrm{d}\mathbf{y}}_{B_2(\boldsymbol{\theta})}. \end{aligned} \tag{B.5}$$

In the expressions of $B_1(\boldsymbol{\theta})$ and $B_2(\boldsymbol{\theta})$ in (B.5), the term ${\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})$ appears. In order to characterize this term, we consider the identity $\mathbb{I} = {\mathbf{M}_\Omega}^{-1/2}\mathbf{M}_\Omega\, {\mathbf{M}_\Omega}^{-1/2}$ and we differentiate it with respect to the shape:

$$\begin{aligned} \mathbf{0} =& \mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\mathbf{M}_\Omega\, {\mathbf{M}_\Omega}^{-1/2} + {\mathbf{M}_\Omega}^{-1/2}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\, {\mathbf{M}_\Omega}^{-1/2} \\ & + {\mathbf{M}_\Omega}^{-1/2}\mathbf{M}_\Omega\, \mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta}). \end{aligned}$$

Considering that $\mathbf{M}_\Omega$ is a symmetric matrix, and that ${\mathbf{M}_\Omega}^{-1/2}$ is symmetric as well, we deduce:

$$ {\mathbf{M}_\Omega}^{1/2}\, \mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta}) = -\frac{1}{2}{\mathbf{M}_\Omega}^{-1/2}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\, {\mathbf{M}_\Omega}^{-1/2}. \tag{B.6}$$

Using equation (B.6) and the fact that, if $\mathbf{F}$ and $\mathbf{G}$ are square matrices, $\mathrm{tr}(\mathbf{FG}) = \mathrm{tr}(\mathbf{GF})$, we remark that:

$$\begin{aligned} B_2(\boldsymbol{\theta}) =& -\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{B}_N} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{y}\right) \mathrm{d}\mathbf{y}\, \mathrm{tr}\left({\mathbf{M}_\Omega}^{1/2}\mathrm{D}{\mathbf{M}_\Omega}^{-1/2}(\boldsymbol{\theta})\right) \\ =& \frac{1}{2}\Phi\left(\Omega\right) \mathrm{tr}\left({\mathbf{M}_\Omega}^{-1/2}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\, {\mathbf{M}_\Omega}^{-1/2}\right) = \frac{1}{2}\Phi\left(\Omega\right) \mathrm{tr}\left({\mathbf{M}_\Omega}^{-1}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\right) = -A(\boldsymbol{\theta}). \end{aligned}$$

Therefore, the shape derivative of $\Phi\left(\cdot\right)$ in Ω consists only in the term $B_1(\boldsymbol{\theta})$

$$\mathrm{D}\Phi\left(\Omega\right)(\boldsymbol{\theta}) = -\frac{1}{2}\sqrt{\frac{\tau^N}{\det \mathbf{M}_\Omega}} \int_{\mathbb{S}_{N-1}} f\left(\sqrt{\tau}{\mathbf{M}_\Omega}^{-1/2}\mathbf{s}\right) \left({\mathbf{M}_\Omega}^{-1/2}\mathrm{D}\mathbf{M}_\Omega(\boldsymbol{\theta})\, {\mathbf{M}_\Omega}^{-1/2}\mathbf{s}\right) \cdot \mathbf{s}\mathrm{d}\mathbf{s}. \tag{B.7}$$

Comparing the expressions 5.19 and (B.7), we conclude that the two approaches to compute the shape derivative of $\Phi\left(\cdot\right)$ are equivalent. □

Appendix C

Quadrature formulas for the centered Gaussian distribution

In this appendix we present some quadrature formulas to integrate multivariate Gaussian densities on the unit sphere or the unit ball in an arbitrary number of dimensions N. We propose recursive relations to express such integrals, which scale well with the dimension N of the Gaussian vector.

Contents

C.1 Notations and definitions

Let N be a positive integer. We denote $\mathbb{B}_N$ the unit ball in $\mathbb{R}^N$ centered in the origin, and $\mathbb{S}_{N-1} = \partial\mathbb{B}_N$ the N-sphere

$$\begin{aligned} \mathbb{B}_N &= \left\{\mathbf{x} \in \mathbb{R}^N \,:\, \|\mathbf{x}\| \leq 1\right\}, \\ \mathbb{S}_{N-1} &= \left\{\mathbf{x} \in \mathbb{R}^N \,:\, \|\mathbf{x}\| = 1\right\}. \end{aligned}$$

The space of real valued square integrable functions on the unit sphere is the space $\mathrm{L}^2(\mathbb{S}_{N-1})$, which is a Hilbert space with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathbb{S}_{N-1}} : \mathrm{L}^2(\mathbb{S}_{N-1}) \times \mathrm{L}^2(\mathbb{S}_{N-1}) \to \mathbb{R}$ such that

$$\langle \phi, \psi \rangle_{\mathbb{S}_{N-1}} = \int_{\mathbb{S}_{N-1}} \phi(\mathbf{s})\ \psi(\mathbf{s})\, \mathrm{d}\mathbf{s} \qquad \text{for any choice of } \phi, \psi \in \mathrm{L}^2(\mathbb{S}_{N-1}).$$

The main objective of this section is the analytic computation of integrals of the quadratic form $\mathbf{x} \mapsto \left(\mathbf{x}^{\mathrm{T}}\mathbf{M}\,\mathbf{x}\right)$ on $\mathbb{B}_N$ and on $\mathbb{S}_{N-1}$, where $\mathbf{M}$ is an $N \times N$ real symmetric matrix. By extension, the knowledge of the expression of such integrals would allow to compute the integral on the ball or on the sphere of any function of the type $\mathbf{x} \mapsto f\left(\mathbf{x}^{\mathrm{T}}\mathbf{M}\,\mathbf{x}\right)$, where $f : \mathbb{B}_N \to \mathbb{R}$ is analytic. A first remark on such integrals is the fact that, given the symmetry of $\mathbb{B}_N$ and $\mathbb{S}_{N-1}$, the value of the integral of $f\left(\mathbf{x}^{\mathrm{T}}\mathbf{M}\,\mathbf{x}\right)$ on either the ball or the sphere is independent from the choice of the cartesian reference, and depends only on the eigenvalues of $\mathbf{M}$. Therefore, without

loss of generality and for the entirety of the current section, we focus on quadratic functions in the form $\mathbf{x} \mapsto \left(\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}\right)$ where $\mathbf{D} = \operatorname{diag}\{\lambda_1, \ldots, \lambda_N\}$ is an $N \times N$ diagonal matrix.

In this section we propose an iterative method to compute integrals of integers power of a quadratic function on the N-sphere and on the N-ball. Such technique does not require the numerical computation of high-dimensional integrals, but relies on a recursive formula and its complexity increases linearly with the dimension N.

C.2 Integrals on the sphere

We are interested in the computation of the integrals on the unit N-sphere $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), 1\right\rangle_{\mathbb{S}_{N-1}}$ and $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ for any $i = 1, \ldots, N$ and for $f(\cdot)$ regular on $\mathbb{B}_N$. Our method relies on a Taylor expansion of $f(\cdot)$. The coefficients of the series are computed thanks to a recursive relation, and we provide an upper bound for the truncation error. The computation of the recursive relation relies on the properties of the spherical Laplace operator $\Delta_{\mathbb{S}_{N-1}} : \mathcal{C}^2(\mathbb{S}_{N-1}) \to \mathcal{C}^0(\mathbb{S}_{N-1})$, that is the Laplace-Beltrami operator on the sphere $\mathbb{S}_{N-1}$, considered as a Riemannian manifold. The Laplace-Beltrami operator on $\mathbb{S}_{N-1}$ is defined for any function $\phi \in \mathcal{C}^2(\mathbb{S}_{N-1})$ as follows

$$\Delta_{\mathbb{S}_{N-1}}\phi = \operatorname{div}_{\mathbb{S}_{N-1}}(\nabla_{\mathbb{S}_{N-1}}\phi), \tag{C.1}$$

where $\operatorname{div}_{\mathbb{S}_{N-1}}$ and $\nabla_{\mathbb{S}_{N-1}}$ represent the divergence and gradient operator on the manifold $\mathbb{S}_{N-1}$ respectively, as introduced in definition 1.9.

Given a $\mathcal{C}^2(\mathcal{A}_N)$ extension $\widetilde{\phi}$ of $\phi \in \mathcal{C}^2(\mathbb{S}_{N-1})$ to an open neighborhood $\mathcal{A}_N \subset \mathbb{R}^N$ of $\mathbb{S}_{N-1}$, it is possible to compute $\Delta_{\mathbb{S}_{N-1}}\phi$ thanks to the following result, derived from [138, Definition 5.4.11 and Proposition 5.4.12]

$$\Delta_{\mathbb{S}_{N-1}}\phi = \operatorname{div}_{\mathbb{S}_{N-1}}(\nabla_{\mathbb{S}_{N-1}}\phi) = \Delta\phi - (N-1)\frac{\partial\widetilde{\phi}}{\partial r}\bigg|_{r=1} - \frac{\partial^2\widetilde{\phi}}{\partial r^2}\bigg|_{r=1}. \tag{C.2}$$

Some important properties of the operator $\Delta_{\mathbb{S}_{N-1}}$ used subsequently are compiled in the following result.

Theorem C.13. *The Laplace-Beltrami operator $\Delta_{\mathbb{S}_{N-1}}$ is self-adjoint with respect to the scalar product $\langle\cdot,\cdot\rangle_{\mathbb{S}_{N-1}}$. The spectrum of the operator $\Delta_{\mathbb{S}_{N-1}}$ is the sequence $\{\ell_k\}_{k=0}^{\infty}$, where, for all $k \geq 0$, $\ell_k = -k(k+N-2)$, and each ℓ_k has multiplicity $\frac{2k+N-2}{k}\binom{N+k-3}{k-1}$.*

The proof of the self-adjoint nature of the operator $\Delta_{\mathbb{S}_{N-1}}$ can be derived using [138, Lemma 5.4.10], or deduced from the result presented in [236] for a generic complete Riemannian manifold. The result about the eigenvalues of $\Delta_{\mathbb{S}_{N-1}}$, as well as other results on spherical harmonics, can be found in [108].

Proposition C.14. *Let $\mathbf{D} = \operatorname{diag}\{\lambda_1, \ldots, \lambda_N\}$ be a $N \times N$ diagonal matrix. Then, for any $k \geq 1$, the quantities $\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}$ and $\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ satisfy the following recursive relation:*

$$\begin{cases} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}} &= \sum_{j=1}^{N}\lambda_j\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_j^2\right\rangle_{\mathbb{S}_{N-1}} \\ \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} &= \frac{1}{N+2k}\left(\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, 1\right\rangle_{\mathbb{S}_{N-1}}\right. \\ &+ \left. 2k\lambda_i\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}}\right) \qquad \text{for } 1 \leq i \leq N. \end{cases} \tag{C.3}$$

The initial value of the recurrence (C.3) *is given by:*

$$\begin{cases} \langle 1,1\rangle_{\mathbb{S}_{N-1}} &= |\mathbb{S}_{N-1}| = \frac{2\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}, \\ \langle 1, s_i^2\rangle_{\mathbb{S}_{N-1}} &= \frac{|\mathbb{S}_{N-1}|}{N} = \frac{\langle 1,1\rangle_{\mathbb{S}_{N-1}}}{N} \qquad \textit{for } 1 \le i \le N. \end{cases} \tag{C.4}$$

Proof. We start by proving the identities (C.4) on the starting point of the recurrence using the formula found in [118]. Let us consider $\boldsymbol{\alpha} \in \mathbb{N}^N$, and $P_{\boldsymbol{\alpha}}$ be a monomial defined on $\mathbb{R}^N$ in the form $P_{\boldsymbol{\alpha}}(\mathbf{x}) = \mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \dots x_N^{\alpha_N}$. Then, the following identity holds:

$$\int_{\mathbb{S}_{N-1}} P_{\boldsymbol{\alpha}}(\mathbf{s})\,\mathrm{d}\mathbf{s} = \begin{cases} 0 & \text{if some } \alpha_j \text{ is odd,} \\ 2\frac{\Gamma(\beta_1)\dots\Gamma(\beta_N)}{\Gamma(\beta_1+\dots+\beta_N)} & \text{if all } \alpha_j \text{ are even,} \end{cases}$$

where $\beta_j = \frac{1}{2}\alpha_j + 1$ for all $j = 1, \dots, N$.

Now we can prove the first equation of (C.3).

$$\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} = \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, 1 \right\rangle_{\mathbb{S}_{N-1}} = \sum_{j=1}^{N} \lambda_j \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_j^2 \right\rangle_{\mathbb{S}_{N-1}}.$$

In order to prove the second equation of (C.3) we have to consider the following two expressions:

$$\Delta_{\mathbb{S}_{N-1}} s_i = -(N-1)s_1, \tag{C.5a}$$
$$\Delta_{\mathbb{S}_{N-1}} s_i^2 = 2 - 2Ns_i^2, \tag{C.5b}$$

as well as the identities valid for all $\phi, \psi \in \mathcal{C}^2(\mathbb{S}_{N-1})$:

$$\Delta_{\mathbb{S}_{N-1}}\phi\psi = \mathrm{div}_{\mathbb{S}_{N-1}}\left(\nabla_{\mathbb{S}_{N-1}}\phi\psi\right) = \phi\Delta_{\mathbb{S}_{N-1}}\psi + \psi\Delta_{\mathbb{S}_{N-1}}\phi + 2\nabla_{\mathbb{S}_{N-1}}\phi\cdot\nabla_{\mathbb{S}_{N-1}}\psi, \tag{C.6a}$$
$$\Delta_{\mathbb{S}_{N-1}}\phi^k = \mathrm{div}_{\mathbb{S}_{N-1}}\left(\nabla_{\mathbb{S}_{N-1}}\phi^k\right) = k\phi^{k-1}\Delta_{\mathbb{S}_{N-1}}\phi + k(k-1)\phi^{k-2}\left\|\nabla_{\mathbb{S}_{N-1}}\phi\right\|^2. \tag{C.6b}$$

Moreover, thanks to equation (C.2) and to the definition of the tangential gradient, we also know that:

$$\begin{aligned} \nabla_{\mathbb{S}_{N-1}}\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s} &= \nabla\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x} - \mathbf{s}\frac{\partial \mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}{\partial r}\Big|_{r=1} &&= 2\mathbf{D}\mathbf{s} - 2(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})\mathbf{s}, \\ \Delta_{\mathbb{S}_{N-1}}\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s} &= \Delta\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x} - (N-1)\frac{\partial \mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}{\partial r}\Big|_{r=1} - \frac{\partial^2 \mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}{\partial r^2}\Big|_{r=1} &&= 2\,\mathrm{tr}\,\mathbf{D} - 2N\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}, \\ \nabla_{\mathbb{S}_{N-1}}s_i &= \nabla_{\mathbb{S}_{N-1}}(\mathbf{e}_i\cdot\mathbf{s}) \;=\; \nabla(\mathbf{e}_i\cdot\mathbf{x}) - \mathbf{s}\frac{\partial r(\mathbf{e}_i\cdot\mathbf{s})}{\partial r}\Big|_{r=1} &&= \mathbf{e}_i - s_i\mathbf{s}, \end{aligned}$$

where $\mathbf{x}$ belongs to a neighborhood of $\mathbb{S}_{N-1} \subset \mathbb{R}^N$, $\mathbf{s} = \mathbf{x}/\left\|\mathbf{x}\right\|$, and $\mathbf{e}_i \in \mathbb{R}^N$ is the i-th element of the canonical basis of $\mathbb{R}^N$. Therefore, we get the following expressions

$$\begin{aligned} &\Delta_{\mathbb{S}_{N-1}}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k = k\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}\Delta_{\mathbb{S}_{N-1}}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right) + k(k-1)\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left\|\nabla_{\mathbb{S}_{N-1}}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)\right\|^2 \\ =&2k\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}\left(\mathrm{tr}\,\mathbf{D} - N(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})\right) + 4k(k-1)\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s} - \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^2\right) \\ =&4\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(k(k-1)\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right)\right) + 2\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}k\,\mathrm{tr}\,\mathbf{D} - 2\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k(kN + 2k(k-1)); \end{aligned} \tag{C.7}$$

$$\begin{aligned}\Delta_{\mathbb{S}_{N-1}}\left(\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k s_i\right) &= s_i\Delta_{\mathbb{S}_{N-1}}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k + \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k \Delta_{\mathbb{S}_{N-1}} s_i + 2\nabla_{\mathbb{S}_{N-1}}(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})\cdot\nabla_{\mathbb{S}_{N-1}} s_i \\ =&4s_i\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(k(k-1)\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right)\right) + 2s_i\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}(k\,\mathrm{tr}\,\mathbf{D} + 2k\lambda_i) \\ &- s_i\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k\left(2kN + 4k^2 - N + 1\right).\end{aligned} \tag{C.8}$$

Using equation (C.5a) and the self-adjoint nature of the Laplace-Beltrami operator we get:

$$\begin{aligned}&-(N-1)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k s_i, s_i\right\rangle_{\mathbb{S}_{N-1}} = \left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k s_i, \Delta_{\mathbb{S}_{N-1}} s_i\right\rangle_{\mathbb{S}_{N-1}} \\ =&\left\langle\Delta_{\mathbb{S}_{N-1}}\left(\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k s_i\right), s_i\right\rangle_{\mathbb{S}_{N-1}} = 4k(k-1)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right), s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ &+2k\left(\mathrm{tr}\,\mathbf{D} + 2\lambda_i\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} - \left(2kN + 4k^2 + N - 1\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}\end{aligned}$$

Thus

$$\begin{aligned}2k(k-1)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right), s_i^2\right\rangle_{\mathbb{S}_{N-1}} &+ k\left(\mathrm{tr}\,\mathbf{D} + 2\lambda_i\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ -k\left(N + 2k\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} &= 0.\end{aligned} \tag{C.9}$$

We perform the same procedure on (C.5b)

$$\begin{aligned}&2\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}} - 2N\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} = \left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, \Delta_{\mathbb{S}_{N-1}} s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ =&\left\langle\Delta_{\mathbb{S}_{N-1}}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} = 4k(k-1)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right), s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ &+2k\,\mathrm{tr}\,\mathbf{D}\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} - 2k(N+2k-2)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}.\end{aligned}$$

Therefore

$$\begin{aligned}2k(k-1)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-2}\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}^2\,\mathbf{s}\right), s_i^2\right\rangle_{\mathbb{S}_{N-1}} &+ k\,\mathrm{tr}\,\mathbf{D}\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ -k\left(kN + 2k^2 - 2k + N\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} &= 0.\end{aligned} \tag{C.10}$$

By subtracting (C.10) from (C.9) we get

$$(N+2k)\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} - 2k\lambda_i\left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} - \left\langle\left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}} = 0, \tag{C.11}$$

proving the second identity of (C.3). □

Given the result of proposition C.14, we state a result about the development of the quantities $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), 1\right\rangle_{\mathbb{S}_{N-1}}$ and $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ as infinite sums for f sufficiently regular on $\mathbb{B}_N$, providing also an estimate on the error intoduced truncating the sum after n terms.

Proposition C.15. *Let* $\mathbf{D} = \operatorname{diag}\{\lambda_1, \dots, \lambda_N\}$ *be a diagonal matrix. We consider* $f : \mathbb{R} \to \mathbb{R}$ *to be an analytic real valued function admitting an extension* $\widetilde{f} : \mathbb{C} \to \mathbb{C}$ *such that* $\widetilde{f}$ *is holomorphic on a disc centered in* 0 *with radius* $r > |\lambda_j|$ *for all* $1 \leq j \leq N$*. Then, the following expansions in infinite series hold for all* $1 \leq i \leq N$

$$\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), 1 \right\rangle_{\mathbb{S}_{N-1}} = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}, \tag{C.12a}$$

$$\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), s_i^2 \right\rangle_{\mathbb{S}_{N-1}} = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2 \right\rangle_{\mathbb{S}_{N-1}}. \tag{C.12b}$$

Moreover, if the sums in the expressions (C.12a) *and* (C.12a) *are computed up to the first* n *terms, the truncation errors committed can both be bounded by the same quantity* $E_f^{\mathrm{trunc}}(n)$ *defined as*

$$E_f^{\mathrm{trunc}}(n) = |\mathbb{S}_{N-1}| \, \frac{M_{f,n+1}}{(n+1)!} \Lambda^{n+1}, \tag{C.13}$$

where $\Lambda = \max\{|\lambda_i| \ : \ i = 1, \dots, N\}$ *and* $M_{f,n+1} \geq f^{(n+1)}(c)$ *for all* $c \in \operatorname{hull}\{0, \lambda_1, \dots, \lambda_N\}$*.*

Proof. Since f is supposed to be holomorphic on the disc centered in 0 with radius r, for any $z \in \mathbb{C}$ such that $|z| < r$, the series $\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} z^k$ is absolutely convergent and converges to $f(z)$. Since, for all $\mathbf{s} \in \mathbb{S}_{N-1}$, $\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s} < r$, we can use Fubini's theorem to swap the integral and the sum. Therefore, we prove the identity (C.12a) as:

$$\begin{aligned}
\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), 1 \right\rangle_{\mathbb{S}_{N-1}} &= \left\langle \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} = \int_{\mathbb{S}_{N-1}} \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k \mathrm{d}\mathbf{s} \\
&= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \int_{\mathbb{S}_{N-1}} \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k \mathrm{d}\mathbf{s} = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}.
\end{aligned}$$

In order to derive the error estimate, we start by remarking that, if $\mathbf{x} \in \mathbb{S}_{N-1}$, then $(\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}) \in \operatorname{hull}\{0, \lambda_1, \dots, \lambda_N\}$. Therefore, using the Lagrange form of the remainder for the Taylor series [25, Section 7.7] and considering $n \in \mathbb{N}$ fixed, we get:

$$\begin{aligned}
\left| \sum_{k=n+1}^{\infty} \frac{f^{(k)}(0)}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| &\leq |\mathbb{S}_{N-1}| \left| \sum_{k=n+1}^{\infty} \frac{f^{(k)}(0)}{k!} |x|^k \right| \\
&\leq |\mathbb{S}_{N-1}| \, \frac{M_{f,n+1}}{(n+1)!} |x|^{n+1} \leq |\mathbb{S}_{N-1}| \, \frac{M_{f,n+1}}{(n+1)!} \Lambda^{n+1},
\end{aligned} \tag{C.14}$$

for any $x \in \operatorname{hull}\{\lambda_1, \dots, \lambda_N\}$ and where $M_{f,n+1} \geq \left| f^{(n+1)}(x) \right|$ for all $c \in \operatorname{hull}\{0, \lambda_1, \dots, \lambda_N\}$.

In order to find the identity (C.12b) we perform the same procedure on $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), s_i^2 \right\rangle_{\mathbb{S}_{N-1}}$. □

By applying proposition C.15 to the function $x \mapsto e^{-x}$ we get the following result.

Corollary C.16. *For any diagonal matrix* $\mathbf{D} = \operatorname{diag}\{\lambda_1, \dots, \lambda_N\}$*, it is possible to develop the*

quantities $\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, 1\right\rangle_{\mathbb{S}_{N-1}}$ *and* $\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ *as follows*

$$\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, 1\right\rangle_{\mathbb{S}_{N-1}} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}, \tag{C.15a}$$

$$\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}. \tag{C.15b}$$

The truncation error for (C.15a) *and* (C.15b) *fixed* $n \in \mathbb{N}$*, has as upper bound the quantity* $E^{\mathrm{trunc}}(n)$ *defined as:*

$$E^{\mathrm{trunc}}(n) = |\mathbb{S}_{N-1}| \, \frac{e^{-\min(0,\lambda_{\min})}}{(n+1)!} \Lambda^{n+1}. \tag{C.16}$$

The procedure detailed in proposition C.15 allows to compute the terms $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), 1\right\rangle_{\mathbb{S}_{N-1}}$ and $\left\langle f(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}), s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ without performing neither a high-dimensional numerical integration on $\mathbb{S}_{N-1}$ nor a Monte Carlo simulation. Moreover, given the structure of the recursive relation (C.3), the complexity of the procedure is linear in N, thus it is free from the curse of dimensionality.

If all eigenvalues $\lambda_1, \ldots, \lambda_N$ are strictly positive, the expressions (C.15a) and (C.15b) appear as infinite sums of alternating terms. On the one hand, alternating series provide a simple method to estimate the error once the series of the coefficients is monotonically decreasing (Leibniz's rule, presented as in [25, Theorem 10.14]). On the other hand, the terms of the alternating series might become large before definitely converging towards zero, and an alternating sum of large numbers can result in considerable numerical errors due to floating point arithmetic (see [125]). Fortunately, there exists an equivalent formulation of the series (C.15a) and (C.15b) which is more adapted to numerical computations, as stated by the following result.

Proposition C.17. *Let* $\mathbf{D} = \mathrm{diag}\,\{\lambda_1, \ldots, \lambda_N\}$ *be a positive definite diagonal matrix where* $\lambda_1 \geq \ldots \geq \lambda_N > 0$*. We denote* $\widetilde{\mathbf{D}} = \mathbf{D}/\lambda_1$ *the rescaled diagonal matrix. Then, the following identities hold for all* $1 \leq i \leq N$

$$\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, 1\right\rangle_{\mathbb{S}_{N-1}} = e^{-\lambda_1} \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}, \tag{C.17a}$$

$$\left\langle e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} = e^{-\lambda_1} \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}, \tag{C.17b}$$

The quantities $\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}$ *and* $\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}}$ *fulfill the following recursive relation for all* $k \geq 1$ *and* $1 \leq i \leq N$

$$\begin{cases} \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}} = \sum_{j=2}^{N} \left(1 - \frac{\lambda_j}{\lambda_1}\right) \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k-1}, s_j^2\right\rangle_{\mathbb{S}_{N-1}}, \\ \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, s_1^2\right\rangle_{\mathbb{S}_{N-1}} = \frac{\left\langle 1, s_i^2\right\rangle_{\mathbb{S}_{N-1}}}{N+2k}, \\ \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} = \frac{\left\langle 1, s_i^2\right\rangle_{\mathbb{S}_{N-1}}}{N+2k} + \frac{2k}{N+2k}\left(1 - \frac{\lambda_i}{\lambda_1}\right) \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k-1}, s_i^2\right\rangle_{\mathbb{S}_{N-1}} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } 2 \leq i \leq N. \end{cases} \tag{C.18}$$

The recursive relation (C.18) *is initialized by:*

$$\begin{cases} \langle 1,1\rangle_{\mathbb{S}_{N-1}} &= |\mathbb{S}_{N-1}| = \frac{2\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}, \\ \langle 1, s_i^2\rangle_{\mathbb{S}_{N-1}} &= \frac{|\mathbb{S}_{N-1}|}{N} = \frac{\langle 1,1\rangle_{\mathbb{S}_{N-1}}}{N} \qquad \textit{for } 1 \leq i \leq N. \end{cases} \tag{C.19}$$

If the series in the expressions (C.17a) *and* (C.17b) *are approximated by the sum of their first n terms, the truncation errors can once again be bounded by the same quantity* $E_{\text{resc}}^{\text{trunc}}(n)$ *defined as*

$$E_{\text{resc}}^{\text{trunc}}(n) = e^{-\lambda_N} \frac{|\mathbb{S}_{N-1}|}{N} \frac{(\lambda_1 - \lambda_N)^{n+1}}{(n+1)!}. \tag{C.20}$$

Proof. As first step, we modify the expression of the function $\mathbf{s} \mapsto e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})}$ in order to isolate the factor $e^{-\lambda_1}$.

$$e^{-(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s})} = e^{-\lambda_1(\mathbf{s}^{\mathrm{T}}\widetilde{\mathbf{D}}\,\mathbf{s})} = e^{\lambda_1(-\mathbf{s}^{\mathrm{T}}\mathbf{s}+\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s})} = e^{-\lambda_1}\, e^{\lambda_1\ \mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}},$$

for any $\mathbf{s} \in \mathbb{S}_{N-1}$.

Since the matrix $-\lambda_1\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)$ is diagonal, we can apply corollary C.16 and derive the expressions (C.17a) and (C.17b) from (C.15a) and (C.15b) respectively. The recursive relation (C.18) and the initialization (C.19) derive from the application of proposition C.14, remarking that $\left(\mathbb{I}-\widetilde{\mathbf{D}}\right) = \operatorname{diag}\left\{0, \left(1-\frac{\lambda_2}{\lambda_1}\right), \ldots, \left(1-\frac{\lambda_N}{\lambda_1}\right)\right\}$.

Finally, in order to estimate the truncation error, we remark that, for all $k \geq 0$

$$0 < e^{-\lambda_1}\frac{\lambda_1^k}{k!}\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, s_i^2\right\rangle_{\mathbb{S}_{N-1}} \leq e^{-\lambda_1}\frac{\lambda_1^k}{k!}\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}.$$

Therefore, any upper bound for the truncation error on the series (C.17a) holds also for the truncation error on (C.17b). Next, we bound the truncation error, exploiting the uniform convergence of the series (C.17a) to swap the sum and the integral defining the scalar product.

$$\begin{aligned} &\left|\left\langle e^{-\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}}, 1\right\rangle_{\mathbb{S}_{N-1}} - e^{-\lambda_1}\sum_{k=0}^{n}\frac{\lambda_1^k}{k!}\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}\right| \\ &= e^{-\lambda_1}\left|\left\langle e^{\lambda_1\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}} - \sum_{k=0}^{n}\frac{\lambda_1^k}{k!}\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}\right| \\ &\leq e^{-\lambda_1}\left\langle \left|e^{\lambda_1\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}} - \sum_{k=0}^{n}\frac{\lambda_1^k}{k!}\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k\right|, 1\right\rangle_{\mathbb{S}_{N-1}}. \end{aligned}$$

Using the error estimate deriving from Lagrange's form of the remainder as in proposition C.15, and remarking that $\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s} \in \left[0, 1-\frac{\lambda_N}{\lambda_1}\right]$ for any $\mathbf{s} \in \mathbb{S}_{N-1}$, we find the estimate

$$\left|e^{\lambda_1\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}} - \sum_{k=0}^{n}\frac{\lambda_1^k}{k!}\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k\right| \leq \frac{M}{(n+1)!}\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k,$$

where $M \geq \left|\frac{\mathrm{d}^{n+1}}{\mathrm{d}t^{n+1}} e^{-\lambda_{\max} t}\right|$ for all $t \in \left[0, 1 - \frac{\lambda_N}{\lambda_N}\right]$. Since $\left|\frac{\mathrm{d}^{n+1}}{\mathrm{d}t^{n+1}} e^{-\lambda_1 t}\right| \leq \lambda_1^{n+1} e^{\lambda_1 - \lambda_N}$ for all $t \in \left[0, 1 - \frac{\lambda_N}{\lambda_1}\right]$, we get

$$
\begin{aligned}
& e^{-\lambda_1} \left\langle \left| e^{\lambda_1 \mathbf{s}^{\mathrm{T}} \left(\mathbb{I} - \widetilde{\mathbf{D}}\right) \mathbf{s}} - \sum_{k=0}^{n} \frac{\lambda_1^k}{k!} \left(\mathbf{s}^{\mathrm{T}} \left(\mathbb{I} - \widetilde{\mathbf{D}}\right) \mathbf{s}\right)^k \right|, 1 \right\rangle_{\mathbb{S}_{N-1}} \\
\leq & e^{-\lambda_1} \frac{\lambda_1^{n+1}}{(n+1)!} e^{\lambda_1 - \lambda_N} \left\langle \left(\mathbf{s}^{\mathrm{T}} \left(\mathbb{I} - \widetilde{\mathbf{D}}\right) \mathbf{s}\right)^{k+1}, 1 \right\rangle_{\mathbb{S}_{N-1}} \\
\leq & e^{-\lambda_N} \frac{\lambda_1^{n+1}}{(n+1)!} \frac{|\mathbb{S}_{N-1}|}{N} \left(1 - \frac{\lambda_N}{\lambda_1}\right)^{n+1} = e^{-\lambda_N} \frac{(\lambda_1 - \lambda_N)^{n+1}}{(n+1)!} \frac{|\mathbb{S}_{N-1}|}{N}.
\end{aligned}
$$

□

We conclude this section with a result on the inner products $\left\langle \left(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s}\right)^k, s_i s_j \right\rangle_{\mathbb{S}_{N-1}}$.

Lemma C.18. *Let $\mathbf{D} \in \mathbb{R}^{N \times N}$ be a symmetric matrix and $f : \mathbb{R}^N \to \mathbb{R}$ a function such that $\left(\mathbf{s} \mapsto f(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s})\right) \in \mathrm{L}^2(\mathbb{S}_{N-1})$. Then, for any pair $i, j \in \{1, \ldots, N\}$ such that $i \neq j$,*

$$
\left\langle \left(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s}\right), s_i s_j \right\rangle_{\mathbb{S}_{N-1}} = 0.
$$

Proof. Let us consider the change of variables $\mathbf{s} \mapsto \widetilde{\mathbf{s}}$ such that $s_\ell = \widetilde{s}_\ell$ if $\ell \neq j$ and $s_j = -\widetilde{s}_j$. Then

$$
\int_{\mathbb{S}_{N-1}} f(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s}) s_i\, s_j \, \mathrm{d}\mathbf{s} = - \int_{\mathbb{S}_{N-1}} f(\widetilde{\mathbf{s}}^{\mathrm{T}} \mathbf{D} \widetilde{\mathbf{s}}) \widetilde{s}_i\, \widetilde{s}_j \, \mathrm{d}\widetilde{\mathbf{s}}. \tag{C.21}
$$

Therefore $\left\langle \left(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s}\right), s_i s_j \right\rangle_{\mathbb{S}_{N-1}} = 0$. □

C.3 Integrals on the unit ball

The integration of the exponential of a quadratic form of $\mathbf{x}$ on the unit sphere is a classical problem in probability theory. Let us consider a random vector $\mathbf{X} = (X_1, \ldots, X_N)^{\mathrm{T}} \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$, and we denote T the random variable

$$
T = \mathbf{X}^{\mathrm{T}}\ \mathbf{D}\ \mathbf{X} = \lambda_1 X_1^2 + \ldots + \lambda_N X_N^2. \tag{C.22}
$$

If $\lambda_1 = \ldots = \lambda_N = 1$, the integral of $\exp\left(\tau\ \mathbf{x}^{\mathrm{T}} \mathbf{D}\, \mathbf{x}\right)$ is strictly related to the cumulative distribution function of a χ^2 random variable with N degrees of freedom. Indeed:

$$
\frac{(\tau)^{N/2}}{(2\pi)^{N/2}} \int_{\mathbb{B}_N} e^{-\frac{\sqrt{\tau}}{2}\|\mathbf{x}\|^2} \, \mathrm{d}\mathbf{x} = \mathbb{P}\left[\frac{1}{\tau}\mathbf{X} \in \mathbb{B}_N\right] = \mathbb{P}\left[T \leq \tau\right] = F_{\chi^2(N)}(\tau) \qquad \text{for all } \tau \geq 0.
$$

In this section we focus on the case where $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ is a Gaussian random vector whose mean $\boldsymbol{\mu}$ is zero, and we propose an iterative method to compute the integral of the function $e^{-\mathbf{x}^{\mathrm{T}} \mathbf{D}\, \mathbf{x}}$ on the unit ball $\mathbb{B}_N$ based on the terms $\left\langle \left(\mathbf{s}^{\mathrm{T}} \mathbf{D}\, \mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}$ defined in section C.2. Such technique does not require on the numerical computation of high-dimensional integrals, but relies on a recursive formula and its complexity increases linearly with the number N of random variables involved.

Proposition C.19. *Let* $\mathbf{D} = \operatorname{diag}\{\lambda_1, \dots, \lambda_N\}$ *be a diagonal matrix. Then, the following identity holds:*

$$\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}, \tag{C.23}$$

where the quantities $\left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}$ *satisfy the recursive relation of proposition C.14.*

Moreover, if the sum (C.23) *is stopped after* n *terms, the following estimate on the truncation error holds:*

$$\left| \int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} - \sum_{k=0}^{n} \frac{(-1)^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \leq \frac{e^{-\min(0,\lambda_{\min})}\,|\mathbb{S}_{N-1}|\,\Lambda^{n+1}}{(N+2(n+1))\,(n+1)!}, \tag{C.24}$$

where $\lambda_{\min} = \min\{\lambda_1, \dots, \lambda_N\}$*, and* $\Lambda = \max\{|\lambda_j| \,:\, j \in \{1, \dots, N\}\}$.

Proof. The proof relies on the expression of the left-hand side of (C.23) in spherical coordinates and the absolute convergence of the Taylor expansion of the exponential.

$$\begin{aligned}
\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} &= \int_0^1 r^{N-1} \left(\int_{\mathbb{S}_{N-1}} e^{-r^2\,\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}}\,\mathrm{d}\mathbf{s} \right) \mathrm{d}r \\
&= \int_0^1 r^{N-1} \left(\int_{\mathbb{S}_{N-1}} \sum_{k=0}^{\infty} \frac{(-1)^k\,r^k}{k!} \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k \mathrm{d}\mathbf{s} \right) \mathrm{d}r \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \int_0^1 r^{N+2k-1}\,\mathrm{d}r \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}}.
\end{aligned}$$

In order to prove the truncation error estimate (C.24), we consider once again the expression of the remainder in the Lagrange form as in the proof of proposition C.15. Indeed, knowing that, for all $\mathbf{s} \in \mathbb{S}_{N-1}$ we have $\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s} \in [\lambda_{\min}, \lambda_{\max}]$, and that $\sup_{t \in \operatorname{hull}\{0,\lambda_1,\dots,\lambda_N\}} \left| \frac{\mathrm{d}^{n+1}}{\mathrm{d}t^{n+1}} e^{-t} \right| = e^{-\min(0,\lambda_{\min})}$, we can deduce that:

$$\begin{aligned}
&\left| \int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} - \sum_{k=0}^{\infty} \frac{(-1)^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
=& \left| \sum_{k=n+1}^{\infty} \frac{(-1)^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
\leq& \frac{1}{N+2(n+1)} \left| \left\langle \sum_{k=n+1}^{\infty} \frac{(-1)^k \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k}{k!}, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
=& \frac{1}{N+2(n+1)} \left| \left\langle e^{-\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}} - \sum_{k=0}^{n} \frac{(-1)^k \left(\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}\right)^k}{k!}, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
\leq& \frac{e^{-\min(0,\lambda_{\min})}\,|\mathbb{S}_{N-1}|\,\Lambda^{n+1}}{(N+2(n+1))\,(n+1)!}.
\end{aligned}$$

□

As we remarked for the result of corollary C.16, if $\mathbf{D}$ is a positive definite matrix, the alternating series (C.23) can be subject to numerical errors deriving from floating point arithmetics. Therefore, we propose a different expansion of the integral $\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\mathrm{d}\mathbf{x}$ as an infinite sum of positive terms, relying on the results of proposition C.17.

Before providing the series expansion of the integral, we state a result on the lower incomplete gamma function γ.

Lemma C.20. *We recall the definition of the lower incomplete gamma function* $\gamma : (\mathbb{R}^+ \setminus \{0\}) \times \mathbb{R}^+ \to \mathbb{R}$*:*

$$\gamma(a,x) = \int_0^x t^{a-1}e^{-t}\,\mathrm{d}t. \tag{C.25}$$

Then, we have the following identity for any integer $\kappa > 0$*:*

$$\int_0^x r^{\kappa}e^{-r^2}\,\mathrm{d}r = \frac{1}{2}\,\gamma\left(\frac{\kappa+1}{2}, x^2\right). \tag{C.26}$$

Moreover, the terms $\gamma(\frac{\kappa}{2}, x)$ *satisfy the following recursive relation for all* $x \geq 0$*:*

$$\begin{cases} \gamma\left(\frac{1}{2}, x\right) &= \sqrt{\pi}\,\mathrm{erf}(x) \\ \gamma(1,x) &= 1 - e^{-x} \\ \gamma\left(\frac{\kappa}{2}, x\right) &= \left(\frac{\kappa}{2}-1\right)\,\gamma\left(\frac{\kappa}{2}-1, x\right) - x^{\frac{\kappa}{2}-1}e^{-x} \qquad \textit{for } \kappa > 2, \end{cases} \tag{C.27}$$

where $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,\mathrm{d}t$ *is the error function evaluated in* x*.*

Proof. Equation (C.26) can be deduced by the change of variable $t = r^2$:

$$\int_0^x r^{\kappa}e^{-r^2}\mathrm{d}r = \frac{1}{2}\int_0^{x^2} t^{\frac{\kappa-1}{2}}e^{-t}\mathrm{d}t = \frac{1}{2}\int_0^{x^2} t^{\frac{\kappa+1}{2}-1}e^{-t}\mathrm{d}t = \frac{1}{2}\,\gamma\left(\frac{\kappa+1}{2}, x^2\right).$$

In order to prove the first two identities in (C.27) we use the expression (C.26) and the definition (C.25) respectively. Indeed:

$$\begin{aligned} \gamma\left(\tfrac{1}{2}, x\right) &= 2\int_0^x e^{-r^2}\,\mathrm{d}r = \sqrt{\pi}\,\mathrm{erf}(x); \\ \gamma(1,x) &= \int_0^x e^{-r}\,\mathrm{d}r = [-e^{-r}]_0^x = 1 - e^{-x}. \end{aligned}$$

The last identity in (C.27) comes from an integration by parts of (C.25):

$$\begin{aligned} \gamma\left(\frac{\kappa}{2}, x\right) &= \int_0^x t^{\frac{\kappa}{2}-1}e^{-t}\,\mathrm{d}t = \left[-t^{\frac{\kappa}{2}-1}e^{-t}\right]_0^x + \left(\frac{\kappa}{2}-1\right)\int_0^x t^{\frac{\kappa}{2}-2}e^{-t} \\ &= \left(\frac{\kappa}{2}-1\right)\,\gamma\left(\frac{\kappa}{2}-1, x\right) - x^{\frac{\kappa}{2}-1}e^{-x}. \end{aligned}$$

□

Proposition C.21. *Let* $\mathbf{D} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$ *be a positive definite diagonal matrix where* $\lambda_1 \geq \ldots \geq \lambda_N > 0$*, and* $\widetilde{\mathbf{D}} = \mathbf{D}/\lambda_1$ *the diagonal matrix rescaled with respect to its largest eigenvalue. Then, the integral on the unit ball* $\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x}$ *can be expressed as follows:*

$$\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} = \frac{\lambda_1^{-N/2}}{2}\sum_{k=0}^{\infty}\frac{1}{k!}\gamma\left(\frac{N}{2}+k, \lambda_1\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I} - \widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1\right\rangle_{\mathbb{S}_{N-1}}, \tag{C.28}$$

where, for all $k \geq 0$, the terms $\left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k}, 1\right\rangle_{\mathbb{S}_{N-1}}$ *are computed by the recursive relation* (C.18) *in proposition C.17, while the terms* $\gamma\left(\frac{N-1}{2}-1, 1\right)$ *are evaluated as in lemma C.20.*

If only the first n terms of (C.28) *are considered, the truncation error can be bounded as:*

$$\left|\frac{\lambda_1^{-N/2}}{2}\sum_{k=n+1}^{\infty}\frac{1}{k!}\,\gamma\left(\frac{N}{2}+k,\lambda_1\right)\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\right| \leq |\mathbb{S}_{N-1}|\,\frac{(\lambda_1-\lambda_N)^{n+1}\,e^{-\lambda_N}}{(N+2(n+1))\;(n+1)!}. \tag{C.29}$$

Proof. At first, we write the term $e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}$ in spherical coordinates, isolating the matrix $\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)$ in the argument o the exponential:

$$e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}} = e^{-\lambda_1\|\mathbf{x}\|^2}e^{\lambda_1\,\mathbf{x}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{x}} = e^{-r^2\lambda_1\|\mathbf{s}\|^2}e^{\lambda_1 r^2\,\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}}.$$

Then, we compute the integral $\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\mathrm{d}\mathbf{x}$ in spherical coordinates.

$$\begin{aligned}
\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\mathrm{d}\mathbf{x} &= \int_{\mathbb{B}_N} e^{-\lambda_1\|\mathbf{x}\|^2}e^{\lambda_1\,\mathbf{x}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{x}}\mathrm{d}\mathbf{x}\\
&= \int_0^1 r^{N-1}e^{-r^2\lambda_1}\left(\int_{\mathbb{S}_{N-1}} e^{\lambda_1 r^2\,\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}}\,\mathrm{d}\mathbf{s}\right)\mathrm{d}r\\
&= \int_0^1\left(\sum_{k=0}^{\infty}\frac{\lambda_1^k}{k!}e^{-r^2\lambda_1}r^{N+2k-1}\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\right)\mathrm{d}r.
\end{aligned}$$

Since all the terms of the sum are positive, we can swap the integral and the sum thanks to Tonelli's theorem. Therefore, using the change of variables $t = r\sqrt{\lambda_1}$ and (C.26), we get:

$$\begin{aligned}
\int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\mathrm{d}\mathbf{x} &= \sum_{k=0}^{\infty}\frac{\lambda_1^k}{k!}\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\int_0^1 r^{N+2k-1}e^{-r^2\lambda_1}\mathrm{d}r\\
&= \sum_{k=0}^{\infty}\frac{\lambda_1^k}{k!}\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\int_0^{\sqrt{\lambda_1}}\frac{t^{N+2k-1}}{\lambda_1^{k+N/2}}e^{-t^2}\,\mathrm{d}t\\
&= \lambda_1^{-N/2}\sum_{k=0}^{\infty}\frac{1}{k!}\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\int_0^{\sqrt{\lambda_1}}e^{-t^2}t^{N+2k-1}\mathrm{d}t\\
&= \frac{\lambda_1^{-N/2}}{2}\sum_{k=0}^{\infty}\frac{1}{k!}\left\langle\left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^{k},1\right\rangle_{\mathbb{S}_{N-1}}\gamma\left(k+\frac{N}{2},\lambda_1\right).
\end{aligned}$$

In order to provide the error estimate (C.29) we remark that:

$$\begin{aligned}
0 < \gamma\left(k+\frac{N}{2},\lambda_1\right) &= \lambda_1^{k+N/2}\int_0^{\sqrt{\lambda_1}}\frac{t^{N+2k-1}}{\lambda_1^{k+N/2}}e^{-t^2}\,\mathrm{d}t\\
&= \lambda_1^{k+N/2}\int_0^1 r^{N+2k-1}e^{-\lambda_1 r^2}\,\mathrm{d}t = \lambda_1^{k+N/2}\frac{e^{-\lambda_1}}{N+2k}.
\end{aligned}$$

Therefore, similarly to the proof of the error estimate in proposition C.17, we get:

$$
\begin{aligned}
&\left| \int_{\mathbb{B}_N} e^{-\mathbf{x}^{\mathrm{T}}\mathbf{D}\,\mathbf{x}}\,\mathrm{d}\mathbf{x} - \frac{\lambda_1^{-N/2}}{2} \sum_{k=0}^{n} \frac{1}{k!} \gamma\left(\frac{N}{2}+k, \lambda_1\right) \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
&\leq \left| \sum_{k=n+1}^{\infty} \frac{\lambda_1^k}{k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \int_0^1 r^{N+2k-1} e^{-\lambda_1 r^2}\,\mathrm{d}t \right| \\
&\leq e^{-\lambda_1} \left| \sum_{k=n+1}^{\infty} \frac{\lambda_1^k}{(N+2k)\,k!} \left\langle \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
&\leq \frac{e^{-\lambda_1}}{2(n+1)+N} \left| \left\langle e^{\lambda_1(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s})} - \sum_{k=0}^{n} \frac{\lambda_1^k}{k!} \left(\mathbf{s}^{\mathrm{T}}\left(\mathbb{I}-\widetilde{\mathbf{D}}\right)\mathbf{s}\right)^k, 1 \right\rangle_{\mathbb{S}_{N-1}} \right| \\
&\leq \frac{e^{-\lambda_1}}{2(n+1)+N} \left|\mathbb{S}_{N-1}\right| \frac{e^{\lambda_1-\lambda_N}}{(n+1)!} (\lambda_1-\lambda_N)^{n+1} \\
&= \frac{e^{-\lambda_N}}{(n+1)!\,(2(n+1)+N)} \left|\mathbb{S}_{N-1}\right| (\lambda_1-\lambda_N)^{n+1}.
\end{aligned}
$$

□

We have expressed the quantities $\int_{\mathbb{S}_{N-1}} e^{-\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}}\mathrm{d}\mathbf{s}$, $\int_{\mathbb{S}_{N-1}} e^{-\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}} s_i^2\,\mathrm{d}\mathbf{s}$ and $\int_{\mathbb{B}_N} e^{-\mathbf{s}^{\mathrm{T}}\mathbf{D}\,\mathbf{s}}\mathrm{d}\mathbf{s}$ as infinite series for a positive definite diagonal matrix $\mathbf{D}$. The coefficients of the series are all positive and depend from a single recursive relation (C.18), initialized by (C.19). Moreover, for each development in series we provided an estimate for the truncation error after n terms.

It should be remarked that, for numerical applications, there exist algorithms computing the lower incomplete gamma function appearing in equation (C.28) that are more stable than the recursive relation (C.27) [20, 120, 244].

Appendix D

Python algorithms for the CDF of a generalized chi-square

In this appendix we report the algorithms used to compute numerically the Cumulative Distribution Function (CDF) of a generalized chi-squared random variable as well as its sensitivity with respect to the parameters of its distribution. The shape optimization methods in which these algorithms are applied are detailed in chapter 5.

Let us consider N to be a positive integer, $\mathcal{O}$ a space of events, $\mathcal{A} \subset 2^{\mathcal{O}}$ a σ-algebra, and $\mathbb{P}$ a probability measure. Let $\mathbf{g} \in \mathbb{N}^N$ be a vector of N strictly positive integers, and $\mathbf{w}(\cdot)$ and $\mathbf{e}(\cdot)$ two vectors in $\mathbb{R}^N$ depending from a parameter $t \in [-\delta, \delta]$, with $\delta > 0$ and such that they are differentiable in $t = 0$. We suppose that all components of $\mathbf{w}(t)$ are strictly positive for all admissible t. For any $t \in [-\delta, \delta]$, we consider the random variable $T(t)$ to follow a generalized chi-squared distribution, where $\mathbf{g}$ is the vector of the degrees of freedom, $\mathbf{w}$ the vector of weights on the non-central chi-squared components, and $\mathbf{e} \odot \mathbf{e}$ the vector of noncentrality parameters.

The random variable $T(t) \in \mathrm{L}^2(\mathcal{O}, \mathbb{P}; \mathbb{R})$ can be interpreted as a weighted sum of N independent random variables as

$$T(t) = w_1(t)Y_1(t) + \ldots + w_N(t)Y_N(t),$$

where each $Y_i(t)$ follows a noncentral chi-squared distribution with g_i degrees of freedom and noncentrality parameter d_i^2. Moreover, $T(t)$ can be seen as the sum of $|\mathbf{g}| = \sum_{j=1}^{N} g_j$ independent Gaussian random variables squared

$$T(t) = \left(\sum_{k_1=1}^{g_1} X_{1,k}^2(t) \right) + \ldots + \left(\sum_{k_N=1}^{g_N} X_{N,k}^2(t) \right),$$

where, for each $i \in \{1, \ldots, N\}$ and $k_i \in \{1, \ldots, g_i\}$, $X_{1,k}(t) \sim \mathcal{N}(d_i(t), w_i(t))$.

Let $\tau > 0$ be a positive constant. In order to ease the notation, we denote $T = T(0)$ the random variable corresponding to $t = 0$, as well as $\mathbf{w} = \mathbf{w}(0)$ and $\mathbf{e} = \mathbf{e}(0)$. We are interested in the computation of the CDF $F_{T(t)}(\tau)$, as well as its derivative with respect to t in $t = 0$.

D.1 Centered case

The algorithm presented in this section can be applied only for cases where $\mathbf{e} = \mathbf{0}$, that is when $T(t)$ is a combination of standard chi-squared random variables. Without loss of generality, we assume that w_N is the minimal component of the vector of weights $\mathbf{w}$. As proven in

proposition 5.3, the CDF of T can be expressed as

$$F_T(\tau) = \frac{(w_N)^{N/2}}{2(\pi)^{N/2}\sqrt{\prod_{i=1}^{N} w_i}} \sum_{k=0}^{\infty} \frac{1}{k!}\gamma\left(\frac{N}{2}+k, \frac{\tau}{2w_N}\right) B_k,$$

where the terms B_k are defined by the recursive relation equation (5.33). Similarly, proposition 5.4 proves that the derivative of $t \mapsto F_{T(t)}(\tau)$ in $t = 0$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} F_T(\tau)\Big|_{t=0} = -\frac{\exp\left(\frac{\tau}{2w_N}\right)}{2\sqrt{\prod_{i=1}^{N} w_i}} \left(\frac{\tau}{2\pi}\right)^{N/2} \sum_{i=1}^{N} \frac{R_i}{w_i} w_i'(t).$$

The CDF of T in τ as well as its derivative $\frac{\mathrm{d}}{\mathrm{d}t} F_T(\tau)\Big|_{t=0}$ can be computed using the expressions above, by the following python code.

```
class IntegrateCenteredGaussian:
    def __init__(self, weights, derivative_weights=None, max_terms = 1000):
            self.weights = weights
        self.N = self.weights.size
        self.max_terms = max_terms

        if derivative_weights is None:
            self.derivative_weights = np.zeros(self.N)
        else:
            self.derivative_weights = derivative_weights


    def n_iterations(self, threshold, precision):
        max_w = np.max(self.weights)
        min_w = np.min(self.weights)
        critical_factor = threshold*0.5 \
            * (1/min_w - 1/max_w)
        factor = critical_factor
        for n in range(0, self.max_terms):
            err = np.exp(-threshold/(2*max_w))*(2*np.pi**(self.N/2))\
                /gamma(self.N/2) * (1/self.N * factor)
            factor = factor * critical_factor/(n+2)
            if (err < precision and n>10) or (n == self.max_terms-1):
                return n
        return self.max_terms

    def computeCoefficients(self, threshold, precision) :
        min_w = np.min(self.weights)
        n_iterations = self.n_iterations(threshold, precision)
        Bs = np.zeros((n_iterations+1,))
        As = np.zeros((self.N, n_iterations+1))
        Bs[0] = (2*np.pi**(self.N/2))/gamma(self.N/2)
        for ii in range(self.N):
```

```
            As[ii][0] = Bs[0]/self.N
        for k in range(n_iterations):
            for ii in range(self.N):
                Bs[k+1] += (1-min_w/self.weights[ii])\
                    *As[ii][k]
            for ii in range(self.N):
                As[ii][k+1] = (Bs[k+1] + 2*(k+1)\
                               *(1-min_w/self.weights[ii])\
                               *As[ii][k])/(self.N + 2*(k+1))
        return As, Bs

    def computeWeightsRi(self, coeffsAs, threshold) :
        min_w = np.min(self.weights)
        critical_factor = threshold*0.5 * (1/min_w)
        Niter = coeffsAs.shape[1]
        Ris = np.zeros((self.N,))
        for ii in range(self.N):
            factor = 1.0
            for k in range(Niter):
                Ris[ii] += factor*coeffsAs[ii][k]
                factor = factor*critical_factor/(k+1)
        return Ris

    def cdf(self, threshold, precision = 1e-5):
        prod_ws = np.prod(self.weights)
        min_w = np.min(self.weights)
        _, coeffsBs = self.computeCoefficients(threshold,precision)
        eigmax = threshold/(2*min_w)
        sumval = 0.0
        coeffInt = 0.5*min_w**(self.N/2)\
            /(np.pi**(self.N/2) * np.sqrt(prod_ws))
        for k in reversed(range(coeffsBs.size)):
          if self.N == 1:
            val = coeffsBs[k]*gammainc(1/2+k, eigmax)\
                *(beta(k+1/2, 1/2))/np.sqrt(np.pi)
          elif self.N == 2:
            val = coeffsBs[k]*gammainc(self.N/2+k, eigmax)
          else:
            val = coeffsBs[k]*gamma(self.N/2-1)\
                *gammainc(self.N/2+k, eigmax)/((beta(k+1, self.N/2-1)))
          sumval += val
        estcdf = sumval*coeffInt
        return estcdf

    def derivative_cdf(self, threshold, precision = 1e-5):
        prod_ws = np.prod(self.weights)
        min_w = np.min(self.weights)
        coeffsAs, _ = self.computeCoefficients(threshold, precision)
```

```
        derivative = 0.0
        coeffK = 0.5*(threshold/(2*np.pi))**(self.N/2)\
            *np.exp(-0.5*threshold/min_w)/(np.sqrt(prod_ws))
        Ris = self.computeWeightsRi(coeffsAs, threshold)
        for ii in range(self.N):
            derivative -= (self.dev_ws[ii] * Ris[ii])/(self.weights[ii])
        return derivative * coeffK
```

D.2 General case

If the vector $\mathbf{e}$ is not identically zero, we can adopt the technique detailed in section 5.4. Thanks to the results of Ruben [214], the CDF of $T(t)$ can be expressed by

$$F_{T(t)}(\tau) = \sum_{k=0}^{\infty} \gamma_k(t) F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right),$$

where $\beta > 0$ is a positive parameter independent from t. The derivative of $t \mapsto F_{T(t)}(\tau)$ in $t = 0$ can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} F_T(\tau)\Big|_{t=0} = \sum_{k=0}^{\infty} \gamma_k'(0) F_{\chi^2(2k+N)}\left(\frac{\tau}{\beta}\right),$$

where

$$\gamma_k'(0) = \mathbf{w}'(0) \cdot \mathbf{p}^k + \mathbf{e}'(0) \cdot \mathbf{q}^k$$

as proven by lemma 5.8.

The quantities $F_T(\tau)$, $\frac{\mathrm{d}}{\mathrm{d}t} F_T(\tau)\Big|_{t=0}$, and the coefficients γ_k and their derivatives can be computed by the python code below.

```
class IntegrateGenericGaussian(IntegrateGaussian):
    def __init__(self, weights, noncentrals=None, derivative_weights=None, ndof=None,
                 derivative_noncentrals=None, max_terms = 1000, beta=None):
        self.weights = weights
        self.N = self.weights.size
        self.max_terms = max_terms
        if noncentrals is None:
            self.noncentrals = np.zeros(self.N)
        else:
            self.noncentrals = noncentrals

        if ndof is None:
            self.ndof = np.ones(self.N)
        else:
            self.ndof = ndof

        if derivative_weights is None:
            self.derivative_weights = np.zeros(self.N)
        else:
            self.derivative_weights = derivative_weights
```

```
        if derivative_noncentrals is None:
            self.derivative_noncentrals = np.zeros(self.N)
        else:
            self.derivative_noncentrals = derivative_noncentrals

        if beta is None:
            self.beta = 0.9 * np.min(weights)
        else:
            self.beta = beta


    def getcoeffs(self, threshold, precision):
        coeffspwr = [self.weights[j]**self.ndof[j] for j in range(self.N)]
        noncentralsSquared = np.asarray([nc**2 for nc in self.noncentrals])
        gvec = np.zeros(self.max_terms)
        nterms = 0
        error = precision+1
        noncentrstot = np.sum(noncentralsSquared)
        ndofstot = np.sum(self.ndof)
        coeffs = np.zeros(self.max_terms)
        errs = np.zeros(self.max_terms)

        ### Auxiliary functions
        def computegterm(k):
            if k == 0:
                g = 0.0
                for j in range(self.N):
                    g += self.ndof[j] + self.beta/self.weights[j] \
                        * (noncentralsSquared[j] - self.ndof[j])
            else:
                g = 0.0
                for j in range(self.N):
                    g += (self.ndof[j]*(1.0-self.beta/self.weights[j])**(k+1)+
                          (k+1)*noncentralsSquared[j]*self.beta\
                              /self.weights[j]\
                              *(1.0-self.beta/self.weights[j])**(k))
            return g

        def computecoeffsterm(k):
            if k == 0:
                return np.exp(-0.5*noncentrstot)\
                    *np.sqrt(self.beta**ndofstot/np.prod(coeffspwrs))
            else:
                convterm = np.dot(coeffs[:k], list(reversed(gvec[:k])))
            return 0.5*convterm/k
```

```
        def computeerror(k, coeffs):
            sumcoeffs = np.sum(list(reversed(coeffs[0:k])))
            return (1.0 - sumcoeffs)*ncx2.cdf(x=threshold/self.beta\
                                , df=ndofstot+2*k,  nc=noncentrstot)
        ### Execution
        while(nterms < self.max_terms and error > precision and error > 0):
            gvec[nterms] = computegterm(nterms)
            coeffs[nterms] = computecoeffsterm(nterms)
            error = computeerror(nterms, coeffs)
            errs[nterms] = error
            nterms += 1
        return coeffs[0:nterms], gvec[0:nterms]

    def cdf(self, threshold, precision = 1e-5):
        coeffsc, gvec = self.getcoeffs(threshold, precision)
        cdfsthresholdcd = np.array([chi2.cdf(x=threshold/self.beta,\
                        df=self.N+2*k) for k in range(coeffsc.size)])
        estcdf = coeffsc.dot(cdfsthresholdcd)
        return estcdf

    def sf(self, threshold, precision = 1e-5):
        coeffsc, beta, gvec = self.getcoeffs(threshold, precision)
        sfsthresholdcd = np.array([chi2.sf(x=threshold/self.beta,\
                        df=self.N+2*k) for k in range(coeffsc.size)])
        estsfcd = coeffsc.dot(sfsthresholdcd)
        return  estsfcd

    def compute_terms_derivative_p(self, threshold, precision):
        gamma_coeffs, g_coeffs = self.getcoeffs(threshold, precision)
        N_term_expansion = gamma_coeffs.size
        p_coeffs = np.zeros([self.N, N_term_expansion])
        for jj in range(self.N):
            p_j = np.zeros(N_term_expansion)
            a_j = np.zeros(N_term_expansion)
            for kk in range(N_term_expansion):
                if kk == 0:
                    p_j[kk] = -gamma_coeffs[0]/(2.0*self.weights[jj])
                elif kk == 1:
                    a_j[0] = (self.beta/self.weights[jj]**2)*(
                        (1.0-self.noncentrals[jj]**2))
                    p_j[1] = (gamma_coeffs[0]* a_j[0] \
                              + g_coeffs[0]* p_j[0])/2
                elif kk == 2:
                    a_j[1] = (self.beta/self.weights[jj]**2)\
                        *(1+ (2*self.noncentrals[jj]**2 - 1)\
                          *(2*self.beta/self.weights[jj]-1))
                    p_j[2] = (np.dot(gamma_coeffs[:kk], \
                                     list(reversed(a_j[:kk])))\
```

```
                            + np.dot(g_coeffs[:kk],\
                                    list(reversed(p_j[:kk])))/(2*kk)
            else:
                a_j[kk-1] = ((1.-self.beta/self.weights[jj])**(kk-2)*\
                             self.beta/self.weights[jj]**2\
                             *((kk-1+ (kk*self.noncentrals[jj]**2 - 1)\
                                *(kk*self.beta/self.weights[jj]-1))))
                p_j[kk] = (np.dot(gamma_coeffs[:kk], \
                                  list(reversed(a_j[:kk]))) +
                           np.dot(g_coeffs[:kk], \
                                  list(reversed(p_j[:kk])))/(2*kk)
        p_coeffs[jj] = p_j
    return p_coeffs

def compute_terms_derivative_q(self, threshold, precision):
    gamma_coeffs, g_coeffs = self.getcoeffs(threshold, precision)
    N_term_expansion = gamma_coeffs.size
    q_coeffs = np.zeros([self.N, N_term_expansion])
    for jj in range(self.N):
        q_j = np.zeros(N_term_expansion)
        b_j = np.zeros(N_term_expansion)
        for kk in range(N_term_expansion):
            if kk == 0:
                q_j[kk] = 0.0
            elif kk == 1:
                b_j[0] = self.beta/self.weights[jj]\
                    *(2 * self.noncentrals[jj])
                q_j[1] = (gamma_coeffs[0]* b_j[0] \
                          +g_coeffs[0]* q_j[0])/2
            elif kk == 2:
                b_j[1] = self.beta/self.weights[jj]**2*(
                          4* self.weights[jj] * self.noncentrals[jj]*
                          (1.-self.beta/self.weights[jj]))
                q_j[2] = (np.dot(gamma_coeffs[:kk], \
                                 list(reversed(b_j[:kk]))) +
                          np.dot(g_coeffs[:kk],\
                                 list(reversed(q_j[:kk])))/(2*kk)
            else:
                b_j[kk-1] = ((1.-self.beta/self.weights[jj])**(kk-2)*
                  self.beta/self.weights[jj]**2*(
                  2*kk* self.weights[jj] * self.noncentrals[jj]*
                      (1.-self.beta/self.weights[jj])))
                q_j[kk] = (np.dot(gamma_coeffs[:kk],\
                                  list(reversed(b_j[:kk]))) +
                       np.dot(g_coeffs[:kk],\
                              list(reversed(q_j[:kk])))/(2*kk)
        q_coeffs[jj] = q_j
    return q_coeffs
```

```
    def derivative_cdf(self, threshold, precision = 1e-5):
        p_coeffs = self.compute_terms_derivative_p(threshold, precision)
        q_coeffs = self.compute_terms_derivative_q(threshold, precision)
        deriv = 0.0
        print()
        for k in range(p_coeffs.shape[0]):
            deriv -= (p_coeffs.T[k].dot(self.dev_ws)
                      + q_coeffs.T[k].dot(self.derivative_noncentrals))\
                    *chi2.cdf(x=threshold/self.beta, df=self.N+2*k)

        return deriv
```

List of Figures

List of Tables

Bibliography

[1] N. Aage and B. S. Lazarov, *Parallel framework for topology optimization using the method of moving asymptotes*, Structural and Multidisciplinary Optimization, 47 (2013), pp. 493–505.

[2] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Elsevier, June 2003.

[3] S. Adly, L. Bourdin, F. Caubet, and A. Jacob de Cordemoy, *Shape Optimization for Variational Inequalities: The Scalar Tresca Friction Problem*, SIAM Journal on Optimization, (2023), pp. 2512–2541.

[4] J. Alexandersen and C. S. Andreasen, *A Review of Topology Optimisation for Fluid-Based Problems*, Fluids, 5 (2020), p. 29.

[5] G. Allaire, *Shape Optimization by the Homogenization Method*, vol. 146 of Applied Mathematical Sciences, Springer, New York, NY, 2002.

[6] G. Allaire, *Conception optimale de structures*, vol. 58 of Mathématiques & applications, Springer Berlin Heidelberg, Berlin, 2007.

[7] G. Allaire, M. Bihr, and B. Bogosel, *Support optimization in additive manufacturing for geometric and thermo-mechanical constraints*, Structural and Multidisciplinary Optimization, 61 (2020), pp. 2377–2399.

[8] G. Allaire, M. Bihr, B. Bogosel, and M. Godoy, *Accessibility constraints in structural optimization via distance functions*, Journal of Computational Physics, 484 (2023), p. 112083.

[9] G. Allaire and B. Bogosel, *Optimizing supports for additive manufacturing*, Structural and Multidisciplinary Optimization, 58 (2018), pp. 2493–2515.

[10] G. Allaire and C. Dapogny, *A linearized approach to worst-case design in parametric and geometric shape optimization*, Mathematical Models and Methods in Applied Sciences, 24 (2014), pp. 2199–2257.

[11] G. Allaire and C. Dapogny, *A deterministic approximation method in shape optimization under random uncertainties*, SMAI Journal of Computational Mathematics, 1 (2015), pp. 83–143.

[12] G. Allaire, C. Dapogny, and P. Frey, *A mesh evolution algorithm based on the level set method for geometry and topology optimization*, Structural and Multidisciplinary Optimization, 48 (2013), pp. 711–715.

[13] G. Allaire, C. Dapogny, and P. Frey, *Shape optimization with a level set based mesh evolution method*, Computer Methods in Applied Mechanics and Engineering, 282 (2014), pp. 22–53.

[14] G. Allaire, C. Dapogny, and F. Jouve, *Shape and Topology Optimization*, vol. 22, Elsevier, 2021.

[15] G. Allaire and L. Jakabčin, *Taking into account thermal residual stresses in topology optimization of structures built by additive manufacturing*, Mathematical Models and Methods in Applied Sciences, 28 (2018), pp. 2313–2366.

[16] G. Allaire and F. Jouve, *A level-set method for vibration and multiple loads structural optimization*, Computer Methods in Applied Mechanics and Engineering, 194 (2005), pp. 3269–3290.

[17] G. Allaire and F. Jouve, *Minimum stress optimal design with the level set method*, Engineering Analysis with Boundary Elements, 32 (2008), pp. 909–918.

[18] G. Allaire, F. Jouve, and G. Michailidis, *Thickness control in structural optimization via a level set method*, Structural and Multidisciplinary Optimization, 53 (2016), pp. 1349–1382.

[19] G. Allaire, F. Jouve, and A.-M. Toader, *Structural optimization using sensitivity analysis and a level-set method*, Journal of Computational Physics, 194 (2004), pp. 363–393.

[20] G. Allasia and R. Besenghi, *Numerical calculation of incomplete gamma functions by the trapezoidal rule*, Numerische Mathematik, 50 (1986), pp. 419–428.

[21] S. Amstutz and H. Andrä, *A new algorithm for topology optimization using a level-set method*, Journal of Computational Physics, 216 (2006), pp. 573–588.

[22] S. Amstutz and M. Ciligot-Travain, *A notion of compliance robustness in topology optimization*, ESAIM: Control, Optimisation and Calculus of Variations, 22 (2016), pp. 64–87.

[23] S. Amstutz and A. A. Novotny, *Topological optimization of structures subject to Von Mises stress constraints*, Structural and Multidisciplinary Optimization, 41 (2010), pp. 407–420.

[24] M. S. Andersen, J. Dahl, and L. Vanderberghe, *CVXOPT: A Python package for convex optimization. Version 1.1.6*, 2013.

[25] T. M. Apostol, *Calculus*, New York : J. Wiley, 1967.

[26] A. Asadpoure, M. Tootkaboni, and J. K. Guest, *Robust topology optimization of structures with uncertainties in stiffness – Application to truss structures*, Computers & Structures, 89 (2011), pp. 1131–1141.

[27] H. Attouch and J. Bolte, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, Mathematical Programming, 116 (2009), pp. 5–16.

[28] H. Attouch, R. Lucchetti, and R. J. B. Wets, *The topology of the ρ-Hausdorff distance*, Annali di Matematica Pura ed Applicata, 160 (1991), pp. 303–320.

[29] H. Attouch and R. J.-B. Wets, *Quantitative Stability of Variational Systems: I. The Epigraphical Distance*, Transactions of the American Mathematical Society, 328 (1991), pp. 695–729.

[30] P. Atwal, S. Conti, B. Geihe, M. Pach, M. Rumpf, and R. Schultz, *On Shape Optimization with Stochastic Loadings*, in Constrained Optimization and Optimal Control for Partial Differential Equations, International Series of Numerical Mathematics, Springer, Basel, 2012, pp. 215–243.

[31] G. Augusti, J. B. Martin, and W. Prager, *On the Decomposition of Stress and Strain Tensors into Spherical and Deviatoric Parts*, Proceedings of the National Academy of Sciences of the United States of America, 63 (1969), pp. 239–241.

[32] J.-F. Aujol and G. Aubert, *Signed distance functions and viscosity solutions of discontinuous Hamilton-Jacobi Equations*, report, INRIA, July 2002.

[33] H. Azegami and Z. C. Wu, *Domain Optimization Analysis in Linear Elastic Problems: Approach Using Traction Method*, JSME international journal. Ser. A, Mechanics and material engineering, 39 (1996), pp. 272–278.

[34] K.-R. Bae and S. Wang, *Reliability-Based Topology Optimization*, 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, (2002).

[35] S. Bak, J. McLaughlin, and D. Renzi, *Some Improvements for the Fast Sweeping Method*, SIAM Journal on Scientific Computing, 32 (2010), pp. 2853–2874.

[36] T. Baker and P. Cavallo, *Dynamic adaptation for deforming tetrahedral meshes*, in 14th Computational Fluid Dynamics Conference, Fluid Dynamics and Co-located Conferences, American Institute of Aeronautics and Astronautics, Nov. 1999.

[37] G. Balarac, F. Basile, P. Bénard, F. Bordeu, J.-B. Chapelier, L. Cirrottola, G. Caumon, C. Dapogny, P. Frey, A. Froehly, G. Ghigliotti, R. Laraufie, G. Lartigue, C. Legentil, R. Mercier, V. Moureau, C. Nardoni, S. Pertant, and M. Zakari, *Tetrahedral remeshing in the context of large-scale numerical simulation and high performance computing*, MathematicS In Action, 11 (2022), pp. 129–164.

[38] I. Bárány and Z. Füredi, *Approximation of the Sphere by Polytopes having Few Vertices*, Proceedings of the American Mathematical Society, 102 (1988), pp. 651–659.

[39] C. Barbarosie, A.-M. Toader, and S. Lopes, *A gradient-type algorithm for constrained optimization with application to microstructure optimization*, Discrete & Continuous Dynamical Systems - B, 25 (2020), p. 1729.

[40] C. Bardos, *Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels ; théorèmes d'approximation ; application à l'équation de transport*, Annales scientifiques de l'École Normale Supérieure, 3 (1970), pp. 185–233.

[41] N. Barral and F. Alauzet, *Three-dimensional CFD simulations with large displacement of the geometries using a connectivity-change moving mesh approach*, Engineering with Computers, 35 (2019), pp. 397–422.

[42] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, in Robust Optimization, Princeton University Press, Aug. 2009.

[43] A. Ben-Tal and A. Nemirovski, *Robust optimization – methodology and applications*, Mathematical Programming, 92 (2002), pp. 453–480.

[44] M. P. Bendsøe and N. Kikuchi, *Generating optimal topologies in structural design using a homogenization method*, Computer Methods in Applied Mechanics and Engineering, 71 (1988), pp. 197–224.

[45] M. P. Bendsøe and O. Sigmund, *Topology Optimization. Theory, Methods, and Applications. 2nd Ed., Corrected Printing*, Springer-Verlag, Jan. 2004.

[46] Martin. P. Bendsøe, *Optimal shape design as a material distribution problem*, Structural optimization, 1 (1989), pp. 193–202.

[47] Martin. P. Bendsøe and O. Sigmund, *Material interpolation schemes in topology optimization*, Archive of Applied Mechanics, 69 (1999), pp. 635–654.

[48] G. Bertolino and M. Montemurro, *Two-scale topology optimisation of cellular materials under mixed boundary conditions*, International Journal of Mechanical Sciences, 216 (2022), p. 106961.

[49] G. Bertolino, M. Montemurro, N. Perry, and F. Pourroy, *An Efficient Hybrid Optimization Strategy for Surface Reconstruction*, Computer Graphics Forum, 40 (2021), pp. 215–241.

[50] R. Bey, J.-P. Loréac, and M. Moussaoui, *Singularities of the solution of a mixed problem for a general second order elliptic equation and boundary stabilizationof the wave equation*, Journal de Mathématiques Pures et Appliquées, 78 (1999), pp. 1043–1067.

[51] E. G. Birgin and J. M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Fundamentals of Algorithms, Society for Industrial and Applied Mathematics, May 2014.

[52] D. A. Bodenham and N. M. Adams, *A comparison of efficient approximations for a weighted sum of chi-squared random variables*, Statistics and Computing, 26 (2016), pp. 917–928.

[53] T. E. Bruns, *Topology optimization by penalty (TOP) method*, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 4430–4443.

[54] D. Büche, *Robust Compressor Optimization by Evolutionary Algorithms*, in Uncertainty Management for Robust Industrial Design in Aeronautics : Findings and Best Practice Collected During UMRIDA, a Collaborative Research Project (2013–2016) Funded by the European Union, Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer International Publishing, Cham, 2019, pp. 629–645.

[55] M. Buckley and G. Eagleson, *An Approximation to the Distribution of Quadratic Forms in Normal Random Variables*, Australian Journal of Statistics, 30A (1988), pp. 150–159.

[56] C. T. T. Bui, C. Dapogny, and P. Frey, *An accurate anisotropic adaptation method for solving the level set advection equation*, International Journal for Numerical Methods in Fluids, 70 (2012), pp. 899–922.

[57] D. Bump, *Lie Groups*, vol. 225 of Graduate Texts in Mathematics, Springer, New York, NY, 2004.

[58] M. Burger, *A framework for the construction of level set methods for shape optimization and reconstruction*, Interfaces and Free Boundaries, 5 (2003), pp. 301–329.

[59] E. Burman, S. Claus, P. Hansbo, M. G. Larson, and A. Massing, *CutFEM: Discretizing geometry and partial differential equations*, International Journal for Numerical Methods in Engineering, 104 (2015), pp. 472–501.

[60] F. Caubet, M. Dambrine, and R. Mahadevan, *Shape Sensitivity of Eigenvalue Functionals for Scalar Problems: Computing the Semi-derivative of a Minimum*, Applied Mathematics & Optimization, 86 (2022), p. 10.

[61] J. Céa, *Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût*, ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique, 20 (1986), pp. 371–402.

[62] M. Ceze and K. J. Fidkowski, *Constrained pseudo-transient continuation*, International Journal for Numerical Methods in Engineering, 102 (2015), pp. 1683–1703.

[63] S. Chen and W. Chen, *A new level-set based approach to shape and topology optimization under geometric uncertainty*, Structural and Multidisciplinary Optimization, 44 (2011), pp. 1–18.

[64] S. Chen, W. Chen, and S. Lee, *Level set based robust shape and topology optimization under random field uncertainties*, Structural and Multidisciplinary Optimization, 41 (2010), pp. 507–524.

[65] S. Chen, M. Y. Wang, and A. Q. Liu, *Shape feature control in structural topology optimization*, Computer-Aided Design, 40 (2008), pp. 951–962.

[66] T. Chen and T. Lumley, *Numerical evaluation of methods approximating the distribution of a large quadratic form in normal variables*, Computational Statistics & Data Analysis, 139 (2019), pp. 75–81.

[67] D. Chenais, *On the existence of a solution in a domain identification problem*, Journal of Mathematical Analysis and Applications, 52 (1975), pp. 189–219.

[68] E. Cherkaev and A. Cherkaev, *Principal Compliance and Robust Optimal Design*, Journal of Elasticity, 72 (2003), pp. 71–98.

[69] S.-K. Choi, R. A. Canfield, and R. V. Grandhi, *Reliability-Based Structural Design*, Springer London, 2007.

[70] D. L. Chopp, *Computing Minimal Surfaces via Level Set Curvature Flow*, Journal of Computational Physics, 106 (1993), pp. 77–91.

[71] A. N. Christiansen, J. A. Bærentzen, M. Nobel-Jørgensen, N. Aage, and O. Sigmund, *Combined shape and topology optimization of 3D structures*, Computers & Graphics, 46 (2015), pp. 25–35.

[72] A. N. Christiansen, M. Nobel-Jørgensen, N. Aage, O. Sigmund, and J. A. Bærentzen, *Topology optimization using an explicit interface representation*, Structural and Multidisciplinary Optimization, 49 (2014), pp. 387–399.

[73] P. G. Ciarlet, *On Korn's inequality*, Chinese Annals of Mathematics, Series B, 31 (2010), pp. 607–618.

[74] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1990.

[75] P. Comon, X. Luciani, and A. L. F. de Almeida, *Tensor decompositions, alternating least squares and other tales*, Journal of Chemometrics, 23 (2009), pp. 393–405.

[76] S. Conti, H. Held, M. Pach, M. Rumpf, and R. Schultz, *Shape Optimization Under Uncertainty—A Stochastic Programming Perspective*, SIAM Journal on Optimization, 19 (2009), pp. 1610–1632.

[77] D. Cortellessa, N. Ferro, S. Perotto, and S. Micheletti, *Enhancing level set-based topology optimization with anisotropic graded meshes*, Applied Mathematics and Computation, 447 (2023), p. 127903.

[78] M. G. Crandall and P.-L. Lions, *Viscosity Solutions of Hamilton-Jacobi Equations*, Transactions of the American Mathematical Society, 277 (1983), pp. 1–42.

[79] G. A. da Silva and E. L. Cardoso, *Topology optimization of continuum structures subjected to uncertainties in material properties*, International Journal for Numerical Methods in Engineering, 106 (2016), pp. 192–212.

[80] M. Dambrine, C. Dapogny, and H. Harbrecht, *Shape Optimization for Quadratic Functionals and States with Random Right-Hand Sides*, SIAM Journal on Control and Optimization, 53 (2015), pp. 3081–3103.

[81] M. Dambrine, H. Harbrecht, and B. Puig, *Computing quantities of interest for random domains with second order shape sensitivity analysis*, ESAIM: Mathematical Modelling and Numerical Analysis, 49 (2015), pp. 1285–1302.

[82] M. Dambrine, H. Harbrecht, and B. Puig, *Incorporating knowledge on the measurement noise in electrical impedance tomography*, ESAIM: Control, Optimisation and Calculus of Variations, 25 (2019), p. 84.

[83] M. Dambrine and V. Karnaev, *Robust obstacle reconstruction in an elastic medium*, Discrete and Continuous Dynamical Systems - B, (2024), pp. 151–173.

[84] M. Dambrine and D. Kateb, *On the Ersatz material approximation in level-set methods*, ESAIM: Control, Optimisation and Calculus of Variations, 16 (2010), pp. 618–634.

[85] M. Dambrine, D. Kateb, and J. Lamboley, *An extremal eigenvalue problem for the Wentzell–Laplace operator*, Annales de l'Institut Henri Poincaré C, Analyse non linéaire, 33 (2016), pp. 409–450.

[86] M. Dambrine and J. Lamboley, *Stability in shape optimization with second variation*, Journal of Differential Equations, 267 (2019), pp. 3009–3045.

[87] J. M. Danskin, *The Theory of Max-Min, with Applications*, SIAM Journal on Applied Mathematics, 14 (1966), pp. 641–664.

[88] C. Dapogny, *Shape Optimization, Level Set Methods on Unstructured Meshes and Mesh Evolution*, these de doctorat, Paris 6, Jan. 2013.

[89] C. Dapogny, C. Dobrzynski, and P. Frey, *Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems*, Journal of Computational Physics, 262 (2014), pp. 358–378.

[90] C. Dapogny, A. Faure, G. Michailidis, G. Allaire, A. Couvelas, and R. Estevez, *Geometric constraints for shape and topology optimization in architectural design*, Computational Mechanics, 59 (2017), pp. 933–965.

[91] C. Dapogny and F. Feppon, *Shape optimization using a level set based mesh evolution method: An overview and tutorial*, Comptes Rendus. Mathématique, 361 (2023), pp. 1267–1332.

[92] C. Dapogny and P. Frey, *Computation of the signed distance function to a discrete contour on adapted triangulation*, Calcolo, 49 (2012), pp. 193–219.

[93] C. Dapogny, F. Iutzeler, A. Meda, and B. Thibert, *Entropy-regularized Wasserstein distributionally robust shape and topology optimization*, Structural and Multidisciplinary Optimization, 66 (2023).

[94] C. Dapogny, N. Lebbe, and E. Oudet, *Optimization of the shape of regions supporting boundary conditions*, Numerische Mathematik, 146 (2020), pp. 51–104.

[95] R. B. Davies, *Numerical Inversion of a Characteristic Function*, Biometrika, 60 (1973), pp. 415–417.

[96] R. B. Davies, *Algorithm AS 155: The Distribution of a Linear Combination of χ^2 Random Variables*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 29 (1980), pp. 323–333.

[97] S. De, J. Hampton, K. Maute, and A. Doostan, *Topology optimization under uncertainty using a stochastic gradient-based approach*, Structural and Multidisciplinary Optimization, 62 (2020), pp. 2255–2278.

[98] F. de Gournay, *Velocity Extension for the Level-set Method and Multiple Eigenvalues in Shape Optimization*, SIAM Journal on Control and Optimization, 45 (2006), pp. 343–367.

[99] F. de Gournay, G. Allaire, and F. Jouve, *Shape and topology optimization of the robust compliance via the level set method*, ESAIM: Control, Optimisation and Calculus of Variations, 14 (2008), pp. 43–70.

[100] M. C. Delfour and J.-P. Zolesio, *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization, Second Edition*, Advances in Design and Control, SIAM, Philadelphia, siam ed., Jan. 2011.

[101] J. Desai, *Topology Optimization in Contact, Plasticity, and Fracture Mechanics Using a Level-Set Method*, PhD thesis, Universite Paris Diderot-Paris VII; Ecole polytechnique, Sept. 2021.

[102] J. Desai, G. Allaire, F. Jouve, and C. Mang, *Topology optimization in quasi-static plasticity with hardening using a level-set method*, Structural and Multidisciplinary Optimization, 64 (2021), pp. 3163–3191.

[103] L. Dieci and L. Lopez, *A survey of numerical methods for IVPs of ODEs with discontinuous right-hand side*, Journal of Computational and Applied Mathematics, 236 (2012), pp. 3967–3991.

[104] M. P. do Carmo, *Riemannian Geometry*, Mathematics: Theory & Applications, Birkhäuser Boston, MA, 1 ed., 1992.

[105] P. Duchesne and P. Lafaye De Micheaux, *Computing the distribution of quadratic forms: Further comparisons between the Liu–Tang–Zhang approximation and exact methods*, Computational Statistics & Data Analysis, 54 (2010), pp. 858–862.

[106] P. D. Dunning and H. A. Kim, *Robust Topology Optimization: Minimization of Expected and Variance of Compliance*, AIAA Journal, 51 (2013), pp. 2656–2664.

[107] P. Duysinx, L. Van Miegroet, T. Jacobs, and C. Fleury, *Generalized Shape Optimization Using X-FEM and Level Set Methods*, in IUTAM Symposium on Topological Design Optimization of Structures, Machines and Materials, Solid Mechanics and Its Applications, Dordrecht, 2006, Springer Netherlands, pp. 23–32.

[108] C. Efthimiou and C. Frye, *Spherical Harmonics in p Dimensions*, World Scientific, May 2014.

[109] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, *Design of robust and efficient photonic switches using topology optimization*, Photonics and Nanostructures - Fundamentals and Applications, 10 (2012), pp. 153–165.

[110] L. Evans, *Partial Differential Equations, Second Edition*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Mar. 2010.

[111] R. W. Farebrother, *Algorithm AS 204: The Distribution of a Positive Linear Combination of χ^2 Random Variables*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 33 (1984), pp. 332–339.

[112] F. Feppon, *Optimisation Topologique de Systèmes Multiphysiques*, PhD thesis, Université Paris Saclay (COmUE), Dec. 2019.

[113] F. Feppon, G. Allaire, F. Bordeu, J. Cortial, and C. Dapogny, *Shape optimization of a coupled thermal fluid–structure problem in a level set mesh evolution framework*, SeMA Journal, 76 (2019), pp. 413–458.

[114] F. Feppon, G. Allaire, and C. Dapogny, *Null space gradient flows for constrained optimization with applications to shape optimization*, ESAIM: Control, Optimisation and Calculus of Variations, 26 (2020), p. 90.

[115] F. Feppon, G. Allaire, and C. Dapogny, *A variational formulation for computing shape derivatives of geometric constraints along rays*, ESAIM: Mathematical Modelling and Numerical Analysis, 54 (2020), pp. 181–228.

[116] G. Fichera, *Existence Theorems in Linear and Semi-Linear Elasticity*, ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik, 54 (1974), pp. 24–36.

[117] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Ltd, 2000.

[118] G. B. Folland, *How to Integrate A Polynomial Over A Sphere*, The American Mathematical Monthly, 108 (2001), pp. 446–448.

[119] N. Forcadel, C. Le Guyader, and C. Gout, *Generalized fast marching method: Applications to image segmentation*, Numerical Algorithms, 48 (2008), pp. 189–211.

[120] H. Früchtl and P. Otto, *A new algorithm for the evaluation of the incomplete gamma function on vector computers*, ACM Transactions on Mathematical Software, 20 (1994), pp. 436–446.

[121] K. Gao, D. M. Do, S. Chu, G. Wu, H. A. Kim, and C. A. Featherston, *Robust topology optimization of structures under uncertain propagation of imprecise stochastic-based uncertain field*, Thin-Walled Structures, 175 (2022), p. 109238.

[122] P. Geoffroy-Donders, *Homogenization Method for Topology Optimization of Structures Built with Lattice Materials*, PhD thesis, Ecole Polytechnique, Dec. 2018.

[123] L. V. Gibiansky and A. V. Cherkaev, *Design of Composite Plates of Extremal Rigidity*, in Topics in the Mathematical Modelling of Composite Materials, Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, Boston, MA, 1997, pp. 95–137.

[124] L. V. Gibiansky and A. V. Cherkaev, *Microstructures of Composites of Extremal Rigidity and Exact Bounds on the Associated Energy Density*, in Topics in the Mathematical Modelling of Composite Materials, A. V. Cherkaev and R. Kohn, eds., Modern Birkhäuser Classics, Springer International Publishing, Cham, 2018, pp. 273–317.

[125] D. Goldberg, *What every computer scientist should know about floating-point arithmetic*, ACM Computing Surveys, 23 (1991), pp. 5–48.

[126] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, Advances in Design and Control, Society for Industrial and Applied Mathematics, Jan. 2002.

[127] L.-X. Guo and J.-Y. Yin, *Finite element analysis and design of an interspinous device using topology optimization*, Medical & Biological Engineering & Computing, 57 (2019), pp. 89–98.

[128] X. Guo, W. Zhang, and L. Zhang, *Robust structural topology optimization considering boundary uncertainties*, Computer Methods in Applied Mechanics and Engineering, 253 (2013), pp. 356–368.

[129] A. Habbal, *Nonsmooth Shape Optimization Applied to Linear Acoustics*, SIAM Journal on Optimization, 8 (1998), pp. 989–1006.

[130] R. Haber, M. P. Bendsøe, and C. Jog, *A new Approach to Variable-Topology Shape Design Using a Constraint on the Perimeter*, Structural optimization (Print), 11 (1996), pp. 1–12.

[131] J. Hadamard, *Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées*, Impr. nationale, Paris, 1908.

[132] P. Hall, *Chi Squared Approximations to the Distribution of a Sum of Independent Random Variables*, The Annals of Probability, 11 (1983), pp. 1028–1036.

[133] R. Haller-Dintelmann, C. Meyer, J. Rehberg, and A. Schiela, *Hölder Continuity and Optimal Control for Nonsmooth Elliptic Problems*, Applied Mathematics and Optimization, 60 (2009), pp. 397–428.

[134] A. M. Hasofer and N. C. Lind, *Exact and Invariant Second-Moment Code Format*, Journal of the Engineering Mechanics Division, 100 (1974), pp. 111–121.

[135] D. He and S. Liu, *BESO method for topology optimization of structures with high efficiency of heat dissipation*, International Journal for Simulation and Multidisciplinary Design Optimization, 2 (2008), pp. 43–48.

[136] F. Hecht, *New development in freefem++*, Journal of Numerical Mathematics, 20 (2012), pp. 1–14.

[137] H. Hencky, *Zur Theorie plastischer Deformationen und der hierdurch im Material hervorgerufenen Nachspannungen*, ZAMM - Journal of Applied Mathematics and Mechanics/ Zeitschrift für Angewandte Mathematik und Mechanik, 4 (1924), pp. 323–334.

[138] A. Henrot and M. Pierre, *Shape Variation and Optimization: A Geometrical Analysis*, no. 28 in Tracts in Mathematics, European Mathematical Society, Zurich, 2018.

[139] C. Hirsch, D. Wunsch, J. Szumbarski, Ł. Łaniewski-Wołłk, and J. Pons-Prats, eds., *Uncertainty Management for Robust Industrial Design in Aeronautics: Findings and Best Practice Collected During UMRIDA, a Collaborative Research Project (2013–2016) Funded by the European Union*, vol. 140 of Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer International Publishing, Cham, 2019.

[140] X. Huang and Y. M. Xie, *Bi-directional evolutionary topology optimization of continuum structures with one or multiple materials*, Computational Mechanics, 43 (2009), pp. 393–401.

[141] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach Spaces: Volume I: Martingales and Littlewood-Paley Theory*, Ergebnisse Der Mathematik Und Ihrer Grenzgebiete. 3. Folge, Springer, Cham, 2016.

[142] J.-P. Imhof, *Computing the distribution of quadratic forms in normal variables*, Biometrika, 48 (1961), pp. 419–426.

[143] A. Jacob de Cordemoy, *Sensitivity Analysis and Optimization for Mechanical Contact Problems*, PhD thesis, Université de Pau et des Pays de l'Adour, July 2023.

[144] K. A. James and H. Waisman, *Layout design of a bi-stable cardiovascular stent using topology optimization*, Computer Methods in Applied Mechanics and Engineering, 305 (2016), pp. 869–890.

[145] R. M. Jones, *Deformation Theory of Plasticity*, Bull Ridge Corporation, Blacksburg, Virginia, 2008.

[146] J. Jonsmann, O. Sigmund, and S. Bouwstra, *Compliant electro-thermal microactuators*, in Technical Digest. IEEE International MEMS 99 Conference. Twelfth IEEE International Conference on Micro Electro Mechanical Systems (Cat. No.99CH36291), Jan. 1999, pp. 588–593.

[147] H.-S. Jung and S. Cho, *Reliability-based topology optimization of geometrically nonlinear structures with loading and material uncertainties*, Finite Elements in Analysis and Design, 41 (2004), pp. 311–331.

[148] R. V. Kadison and J. R. Ringrose, *Elementary Theory: Fundamentals of the Theory of Operator Algebras*, vol. I, Elsevier, New York, 1983.

[149] S. Katoch, S. S. Chauhan, and V. Kumar, *A review on genetic algorithm: Past, present, and future*, Multimedia Tools and Applications, 80 (2021), pp. 8091–8126.

[150] C. T. Kelley and D. E. Keyes, *Convergence Analysis of Pseudo-Transient Continuation*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 508–523.

[151] V. Keshavarzzadeh, H. Meidani, and D. A. Tortorelli, *Gradient based design optimization under uncertainty via stochastic expansion methods*, Computer Methods in Applied Mechanics and Engineering, 306 (2016), pp. 47–76.

[152] G. Kharmanda, N. Olhoff, A. Mohamed, and M. Lemaire, *Reliability-based topology optimization*, Structural and Multidisciplinary Optimization, 26 (2004), pp. 295–307.

[153] R. V. Kohn and G. Strang, *Optimal design and relaxation of variational problems, I, II, and III*, Communications on Pure and Applied Mathematics, 39 (1986), pp. 113–137.

[154] T. G. Kolda, *Numerical optimization for symmetric tensor decomposition*, Mathematical Programming, 151 (2015), pp. 225–248.

[155] T. G. Kolda and B. W. Bader, *Tensor Decompositions and Applications*, SIAM Review, 51 (2009), pp. 455–500.

[156] P. Kolvenbach, O. Lass, and S. Ulbrich, *An approach for robust PDE-constrained optimization with application to shape optimization of electrical engines and of dynamic elastic structures under uncertainty*, Optimization and Engineering, 19 (2018), pp. 697–731.

[157] V. A. Kondrat'ev and O. A. Oleinik, *Boundary-value problems for the system of elasticity theory in unbounded domains. Korn's inequalities*, Russian Mathematical Surveys, 43 (1988), p. 65.

[158] J. Kossaifi, Y. Panagakis, A. Anandkumar, and M. Pantic, *TensorLy: Tensor Learning in Python*, Journal of Machine Learning Research, 20 (2019), pp. 1–6.

[159] D. Kuonen, *Saddlepoint Approximations for Distributions of Quadratic Forms in Normal Variables*, Biometrika, 86 (1999), pp. 929–935.

[160] S. Lang, *Fundamentals of Differential Geometry*, vol. 191 of Graduate Texts in Mathematics, Springer, New York, NY, 1999.

[161] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, *Boundary control and shape optimization for the robust design of bypass anastomoses under uncertainty*, ESAIM: Mathematical Modelling and Numerical Analysis, 47 (2013), pp. 1107–1131.

[162] B. S. Lazarov, M. Schevenels, and O. Sigmund, *Topology optimization with geometric uncertainties by perturbation techniques*, International Journal for Numerical Methods in Engineering, 90 (2012), pp. 1321–1336.

[163] N. Lebbe, C. Dapogny, E. Oudet, K. Hassan, and A. Gliere, *Robust shape and topology optimization of nanophotonic devices using the level set method*, Journal of Computational Physics, 395 (2019), pp. 710–746.

[164] B. G. Lindsay, R. S. Pilla, and P. Basak, *Moment-Based Approximations of Distributions Using Mixtures: Theory and Applications*, Annals of the Institute of Statistical Mathematics, 52 (2000), pp. 215–230.

[165] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Springer, Berlin, Heidelberg, 1972.

[166] H. Liu, Y. Tang, and H. H. Zhang, *A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables*, Computational Statistics & Data Analysis, 53 (2009), pp. 853–856.

[167] J. Liu and Y. Ma, *A survey of manufacturing oriented topology optimization methods*, Advances in Engineering Software, 100 (2016), pp. 161–175.

[168] S. C. H. Lu and K. S. Pister, *Decomposition of deformation and representation of the free energy function for isotropic thermoelastic solids*, International Journal of Solids and Structures, 11 (1975), pp. 927–934.

[169] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, Jan. 1969.

[170] D. Luft, V. H. Schulz, and K. Welker, *Efficient Techniques for Shape Optimization with Variational Inequalities Using Adjoints*, SIAM Journal on Optimization, 30 (2020), pp. 1922–1953.

[171] T. Lumley, J. Brody, G. Peloso, A. Morrison, and K. Rice, *FastSKAT: Sequence kernel association tests for very large sets of markers*, Genetic Epidemiology, 42 (2018), pp. 516–527.

[172] K. A. Lurie, A. V. Cherkaev, and A. V. Fedorov, *Regularization of optimal design problems for bars and plates, parts I and II*, Journal of Optimization Theory and Applications, 37 (1982), pp. 499–522.

[173] J. R. Magnus, *On Differentiating Eigenvalues and Eigenvectors*, Econometric Theory, 1 (1985), pp. 179–191.

[174] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley Series in Probability and Statistics, John Wiley & Sons, third edition (2007) ed., 1999.

[175] S. Mantovani, S. G. Barbieri, M. Giacopini, A. Croce, A. Sola, and E. Bassoli, *Synergy between topology optimization and additive manufacturing in the automotive field*, Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 235 (2021), pp. 555–567.

[176] A. Manzoni, A. Quarteroni, and S. Salsa, *Optimal Control of Partial Differential Equations: Analysis, Approximation, and Applications*, vol. 207 of Applied Mathematical Sciences, Springer International Publishing, Cham, 2021.

[177] G. Marck, M. Nemer, and J.-L. Harion, *Topology Optimization of Heat and Mass Transfer Problems: Laminar Flow*, Numerical Heat Transfer, Part B: Fundamentals, 63 (2013), pp. 508–539.

[178] J. Martínez, *A note on the theoretical convergence properties of the SIMP method*, Structural and Multidisciplinary Optimization, 29 (2005), pp. 319–323.

[179] J. Martínez-Frutos, D. Herrero-Pérez, M. Kessler, and F. Periago, *Risk-averse structural topology optimization under random fields using stochastic expansion methods*, Computer Methods in Applied Mechanics and Engineering, 330 (2018), pp. 180–206.

[180] K. Maute and D. M. Frangopol, *Reliability-based design of MEMS mechanisms by topology optimization*, Computers & Structures, 81 (2003), pp. 813–824.

[181] G. Michailidis, *Manufacturing Constraints and Multi-Phase Shape and Topology Optimization via a Level-Set Method*, PhD thesis, Ecole Polytechnique X, Jan. 2014.

[182] M. K. Misztal and J. A. Bærentzen, *Topology-adaptive interface tracking using the deformable simplicial complex*, ACM Transactions on Graphics, 31 (2012), pp. 24:1–24:12.

[183] F. Mitjana, *Optimisation Topologique de Structures Sous Contraintes de Flambage*, these de doctorat, Toulouse 3, June 2018.

[184] B. Mohammadi and O. Pironneau, *Applied Shape Optimization for Fluids*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, New York, second edition, second edition ed., Sept. 2009.

[185] M. Montemurro, G. Bertolino, and T. Roiné, *A general multi-scale topology optimisation method for lightweight lattice structures obtained through additive manufacturing technology*, Composite Structures, 258 (2021), p. 113360.

[186] M. Montemurro, A. Vincenti, and P. Vannucci, *A Two-Level Procedure for the Global Optimum Design of Composite Modular Structures—Application to the Design of an Aircraft Wing*, Journal of Optimization Theory and Applications, 155 (2012), pp. 1–23.

[187] F. Murat and L. Tartar, *Calculus of Variations and Homogenization*, in Les méthodes de l'homogénéisation : Théorie et applications en physique, vol. 57 of Coll. Dir. Etudes et Recherches EDF, Eyrolles, Paris, 1985, pp. 319–369.

[188] C. Nardoni, D. Danan, C. Mang, F. Bordeu, and J. Cortial, *A R&D Software Platform for Shape and Topology Optimization Using Body-Fitted Meshes*, in Mesh Generation and Adaptation: Cutting-Edge Techniques, R. Sevilla, S. Perotto, and K. Morgan, eds., SEMA SIMAI Springer Series, Springer International Publishing, Cham, 2022, pp. 23–39.

[189] R. Nigro, D. Wunsch, G. Coussement, and C. Hirsch, *Robust Design in Turbomachinery Applications*, in Uncertainty Management for Robust Industrial Design in Aeronautics, Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer International Publishing, Cham, 2019, pp. 495–511.

[190] D. Noll, *Convergence of Non-smooth Descent Methods Using the Kurdyka–Łojasiewicz Inequality*, Journal of Optimization Theory and Applications, 160 (2014), pp. 553–572.

[191] A. A. Novotny and J. Sokołowski, *Topological Derivatives in Shape Optimization*, Interaction of Mechanics and Mathematics, Springer, Berlin, Heidelberg, 2013.

[192] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, vol. 153 of Applied Mathematical Sciences, Springer, New York, NY, 2003.

[193] S. Osher and J. A. Sethian, *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations*, Journal of Computational Physics, 79 (1988), pp. 12–49.

[194] S. J. Osher and F. Santosa, *Level Set Methods for Optimization Problems Involving Geometry and Constraints: I. Frequencies of a Two-Density Inhomogeneous Drum*, Journal of Computational Physics, 171 (2001), pp. 272–288.

[195] C. Othmer, *A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows*, International Journal for Numerical Methods in Fluids, 58 (2008), pp. 861–877.

[196] S. L. Padula, C. R. Gumbert, and W. Li, *Aerospace applications of optimization under uncertainty*, Optimization and Engineering, 7 (2006), pp. 317–328.

[197] O. Pantz and K. Trabelsi, *A Post-Treatment of the Homogenization Method for Shape Optimization*, SIAM Journal on Control and Optimization, 47 (2008), pp. 1380–1398.

[198] O. Pironneau, *Optimal Shape Design for Elliptic Systems*, Springer, New York, 1984.

[199] J. Plocher and A. Panesar, *Review on design and structural optimisation in additive manufacturing: Towards next-generation lightweight structures*, Materials & Design, 183 (2019), p. 108164.

[200] F. A. Potra and S. J. Wright, *Interior-point methods*, Journal of Computational and Applied Mathematics, 124 (2000), pp. 281–302.

[201] P. Putek, E. J. W. ter Maten, M. Günther, A. Bartel, R. Pulch, P. Meuris, and W. Schoenmaker, *Robust Shape Optimization under Uncertainties in Device Materials, Geometry and Boundary Conditions*, in Nanoelectronic Coupled Problems Solutions, Mathematics in Industry, Springer International Publishing, Cham, 2019, pp. 223–260.

[202] J. Qian, Y.-T. Zhang, and H.-K. Zhao, *Fast Sweeping Methods for Eikonal Equations on Triangular Meshes*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 83–107.

[203] O. Querin, G. Steven, and Y. Xie, *Evolutionary structural optimisation (ESO) using a bidirectional algorithm*, Engineering Computations, 15 (1998), pp. 1031–1048.

[204] L. Rakotondrainibe, *Topology Optimization of Connections in Mechanical Systems*, PhD thesis, Institut Polytechnique de Paris, Dec. 2020.

[205] L. Rakotondrainibe, G. Allaire, and P. Orval, *Topology optimization of connections in mechanical systems*, Structural and Multidisciplinary Optimization, 61 (2020), pp. 2253–2269.

[206] M. Reed and B. Simon, eds., *Methods of Modern Mathematical Physics, Volume I: Functional Analysis*, Academic Press, Jan. 1972.

[207] T. P. Ribeiro, L. F. A. Bernardo, and J. M. A. Andrade, *Topology Optimisation in Structural Steel Design for Additive Manufacturing*, Applied Sciences, 11 (2021), p. 2112.

[208] H. Robbins and E. J. G. Pitman, *Application of the Method of Mixtures to Quadratic Forms in Normal Variates*, The Annals of Mathematical Statistics, 20 (1949), pp. 552–560.

[209] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, Princeton University Press, 1970.

[210] T. Roiné, M. Montemurro, and J. Pailhès, *Stress-based topology optimization through non-uniform rational basis spline hyper-surfaces*, Mechanics of Advanced Materials and Structures, 258 (2021).

[211] S. A. L. Rostami and A. Ghoddosian, *Robust Topology Optimization under Load and Geometry Uncertainties by Using New Sparse Grid Collocation Method*, Periodica Polytechnica Civil Engineering, 63 (2019), pp. 898–907.

[212] S. A. L. Rostami, M. Li, A. Kolahdooz, H. Chung, and J. Zhang, *Robust Topology Optimization of Continuum Structures under the Hybrid Uncertainties: A Comparative Study*, Periodica Polytechnica Civil Engineering, 67 (2023), pp. 637–645.

[213] G. I. N. Rozvany and T. Lewiński, *Topology Optimization in Structural and Continuum Mechanics*, vol. 549 of CISM International Centre for Mechanical Sciences, Springer, Vienna, 2014.

[214] H. Ruben, *Probability Content of Regions Under Spherical Normal Distributions, IV: The Distribution of Homogeneous and Non-Homogeneous Quadratic Functions of Normal Variables*, The Annals of Mathematical Statistics, 33 (1962), pp. 542–570.

[215] W. Rudin, *Principles of Mathematical Analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill, New York, third edition ed., 1976.

[216] W. Rudin, *Real and Complex Analysis, 3rd Ed.*, McGraw-Hill, Inc., USA, 1987.

[217] R. A. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer Monographs in Mathematics, Springer, London, 2002.

[218] S. Salsa, *Partial Differential Equations in Action: From Modeling To Theory*, UNITEXT, Springer International Publishing AG, Cham, 3 ed., Apr. 2015.

[219] F. E. Satterthwaite, *An Approximate Distribution of Estimates of Variance Components*, Biometrics Bulletin, 2 (1946), pp. 110–114.

[220] C. Schillings, S. Schmidt, and V. Schulz, *Efficient shape optimization for certain and uncertain aerodynamic design*, Computers & Fluids, 46 (2011), pp. 78–87.

[221] J. Schropp and I. Singer, *A dynamical systems approach to constrained minimization*, Numerical Functional Analysis and Optimization, 21 (2000), pp. 537–551.

[222] V. Schulz and C. Schillings, *Optimal Aerodynamic Design under Uncertainty*, in Management and Minimisation of Uncertainties and Errors in Numerical Aerodynamics: Results of the German Collaborative Project MUNA, B. Eisfeld, H. Barnewitz, W. Fritz, and F. Thiele, eds., Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer, Berlin, Heidelberg, 2013, pp. 297–338.

[223] V. H. Schulz, *A Riemannian View on Shape Optimization*, Foundations of Computational Mathematics, 14 (2014), pp. 483–501.

[224] V. H. Schulz, M. Siebenborn, and K. Welker, *Efficient PDE Constrained Shape Optimization Based on Steklov–Poincaré-Type Metrics*, SIAM Journal on Optimization, 26 (2016), pp. 2800–2819.

[225] C. Schwab and R.-A. Todor, *Sparse finite elements for elliptic problems with stochastic loading*, Numerische Mathematik, 95 (2003), pp. 707–734.

[226] Ch. Schwab and R. Todor, *Sparse Finite Elements for Stochastic Elliptic Problems – Higher Order Moments*, Computing, 71 (2003), pp. 43–63.

[227] M. Sehmi, J. Christensen, C. Bastien, and S. Kanarachos, *Review of topology optimisation refinement processes for sheet metal manufacturing in the automotive industry*, Structural and Multidisciplinary Optimization, 58 (2018), pp. 305–330.

[228] J. A. Sethian, *A Fast Marching Level Set Method for Monotonically Advancing Fronts*, Proceedings of the National Academy of Sciences of the United States of America, 93 (1996), pp. 1591–1595.

[229] J. A. Sethian, *Fast Marching Methods*, SIAM Review, 41 (1999), pp. 199–235.

[230] J. A. Sethian and A. Wiegmann, *Structural Boundary Design via Level Set and Immersed Interface Methods*, Journal of Computational Physics, 163 (2000), pp. 489–528.

[231] J. Sheil and I. O'Muircheartaigh, *Algorithm AS 106: The Distribution of Non - Negative Quadratic Forms in Normal Variables*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 26 (1977), pp. 92–98.

[232] W. S. Slaughter, *The Linearized Theory of Elasticity*, Birkhäuser, Boston, 2002.

[233] J. Sokołowski and A. Żochowski, *Topological Derivatives of Shape Functionals for Elasticity Systems*, in Optimal Control of Complex Structures, ISNM International Series of Numerical Mathematics, Basel, 2002, Birkhäuser, pp. 231–244.

[234] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, *OSQP: An operator splitting solver for quadratic programs*, Mathematical Programming Computation, 12 (2020), pp. 637–672.

[235] J. Strain, *Fast Tree-Based Redistancing for Level Set Computations*, Journal of Computational Physics, 152 (1999), pp. 664–686.

[236] R. S. Strichartz, *Analysis of the Laplacian on the complete Riemannian manifold*, Journal of Functional Analysis, 52 (1983), pp. 48–79.

[237] L. L. Stromberg, A. Beghini, W. F. Baker, and G. H. Paulino, *Application of layout and topology optimization using pattern gradation for the conceptual design of buildings*, Structural and Multidisciplinary Optimization, 43 (2011), pp. 165–180.

[238] V. Subramaniam, *Topology Optimization of Conjugated Heat Transfer Devices : Experimental and Numerical Investigation*, PhD thesis, Ecole nationale supérieure Mines-Télécom Lille Douai, Dec. 2018.

[239] Y. Sun, S. Yao, and J. Alexandersen, *Topography optimisation using a reduced-dimensional model for convective heat transfer between plates with varying channel height and constant temperature*, Structural and Multidisciplinary Optimization, 66 (2023), p. 206.

[240] K. Svanberg, *The method of moving asymptotes—a new method for structural optimization*, International Journal for Numerical Methods in Engineering, 24 (1987), pp. 359–373.

[241] Y. Tang, A. Kurtz, and Y. F. Zhao, *Bidirectional Evolutionary Structural Optimization (BESO) based design method for lattice structure to be fabricated by additive manufacturing*, Computer-Aided Design, 69 (2015), pp. 91–101.

[242] L. Tartar, *Problemes de Controle des Coefficients Dans des Equations aux Derivees Partielles*, in Control Theory, Numerical Methods and Computer Systems Modelling, Berlin, Heidelberg, 1975, Springer, pp. 420–426.

[243] R. Tawk, *Optimisation Topologique Des Transferts de Masse et de Chaleur En Écoulement Bi-Fluide Laminaire : Application Aux Échangeurs de Chaleur*, these de doctorat, Paris Sciences et Lettres (ComUE), June 2018.

[244] N. M. Temme, *A Set of Algorithms for the Incomplete Gamma Functions*, Probability in the Engineering and Informational Sciences, 8 (1994), pp. 291–307.

[245] M. Tootkaboni, A. Asadpoure, and J. K. Guest, *Topology optimization of continuum structures under uncertainty – A Polynomial Chaos approach*, Computer Methods in Applied Mechanics and Engineering, 201–204 (2012), pp. 263–275.

[246] Y.-C. Tsai, F.-B. Liu, and P.-T. Shen, *Investigations of the pressure drop and flow distribution in a chevron-type plate heat exchanger*, International Communications in Heat and Mass Transfer, 36 (2009), pp. 574–578.

[247] N. P. van Dijk, K. Maute, M. Langelaar, and F. van Keulen, *Level-set methods for structural topology optimization: A review*, Structural and Multidisciplinary Optimization, 48 (2013), pp. 437–472.

[248] J.-L. Vie, *Second-Order Derivatives for Shape Optimization with a Level-Set Method*, PhD thesis, Université Paris-Est, Dec. 2016.

[249] C. H. Villanueva and K. Maute, *Density and level set-XFEM schemes for topology optimization of 3-D structures*, Computational Mechanics, 54 (2014), pp. 133–150.

[250] C. H. Villanueva and K. Maute, *CutFEM topology optimization of 3D laminar incompressible flow problems*, Computer Methods in Applied Mechanics and Engineering, 320 (2017), pp. 444–473.

[251] R. von Mises, *Mechanik der festen Körper im plastisch - deformablen Zustand*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1913 (1913), pp. 582–592.

[252] A. Wächter and L. T. Biegler, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Mathematical Programming, 106 (2006), pp. 25–57.

[253] M. Y. Wang, X. Wang, and D. Guo, *A level set method for structural topology optimization*, Computer Methods in Applied Mechanics and Engineering, 192 (2003), pp. 227–246.

[254] A. Wiegmann and K. P. Bube, *The Explicit-Jump Immersed Interface Method: Finite Difference Methods for PDES with Piecewise Smooth Solutions*, SIAM Journal on Numerical Analysis, 37 (2000), pp. 827–862.

[255] Q. Xia and M. Y. Wang, *Topology optimization of thermoelastic structures using level set method*, Computational Mechanics, 42 (2008), pp. 837–857.

[256] Y. M. Xie and G. P. Steven, *Evolutionary Structural Optimization*, Springer, London, 1997.

[257] Y. M. Xie, Z. H. Zuo, X. Huang, T. Black, and P. Felicetti, *Application of Topological Optimisation Technology to Bridge Design*, Structural Engineering International, 24 (2014), pp. 185–191.

[258] H. Yamashita, *A differential equation approach to nonlinear programming*, Mathematical Programming, 18 (1980), pp. 155–168.

[259] M. Yoshimura, K. Shimoyama, T. Misaka, and S. Obayashi, *Topology optimization of fluid problems using genetic algorithm assisted by the Kriging model*, International Journal for Numerical Methods in Engineering, 109 (2017), pp. 514–532.

[260] T. Zegard and G. H. Paulino, *Bridging topology optimization and additive manufacturing*, Structural and Multidisciplinary Optimization, 53 (2016), pp. 175–192.

[261] S.-E. Zerrouq, *Robust Shape Optimization for Solid Mechanics and Fluid Mechanics*, PhD thesis, Université de Pau et des Pays de l'Adour, Dec. 2022.

[262] S.-Q. Zhang, B.-B. Xu, Z.-H. Gao, G.-H. Jiang, Y.-T. Zheng, H.-Y. Liu, W.-W. Jiang, K. Yang, and X.-W. Gao, *Topology optimization of heat transfer and elastic problems based on element differential method*, Engineering Analysis with Boundary Elements, 149 (2023), pp. 190–202.

[263] H. Zhao, *A fast sweeping method for Eikonal equations*, Mathematics of Computation, 74 (2005), pp. 603–627.

[264] S. Zhou and Q. Li, *A variational level set method for the topology optimization of steady-state Navier–Stokes flow*, Journal of Computational Physics, 227 (2008), pp. 10178–10195.

ÉCOLE DOCTORALE :
École doctorale des sciences exactes et leurs applications

LABORATOIRE :
Laboratoire de Mathématiques et de leurs Applications

Giulio Gargantini

giulio.gargantini@univ-pau.fr

Université de Pau et des Pays de l'Adour
Avenue de l'Université
BP 576
64012 Pau Cedex

www.univ-pau.fr