Modeling and Data Assimilation for Ultrasonic Guided Waves in Structural Health Monitoring under Operational Loading Conditions

André Dalmora

To cite this version:

André Dalmora. Modeling and Data Assimilation for Ultrasonic Guided Waves in Structural Health Monitoring under Operational Loading Conditions. Modeling and Simulation. Institut Polytechnique de Paris, 2023. English. NNT : 2023IPPAX137 . tel-04453563

HAL Id: tel-04453563
https://theses.hal.science/tel-04453563v1

Submitted on 12 Feb 2024

INSTITUT
POLYTECHNIQUE
DE PARIS

NNT : 2023IPPAX137

Thèse de doctorat

Modeling and Data Assimilation for Ultrasonic Guided Waves in *Structural Health Monitoring* under Operational Loading Conditions

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : mécanique des fluides et des solides, acoustique

Thèse présentée et soutenue à Gif-sur-Yvette, le 30 Novembre 2023, par

ANDRÉ DALMORA

Composition du Jury :

Marc Bonnet Directeur de recherche, CNRS	Président
Ludovic Chamoin Professeur, ENS Paris-Saclay	Rapporteur
Bruno Lombard Directeur de recherche, CNRS	Rapporteur
Maya de Buhan Ingénieure de recherche, Safran Tech	Examinatrice
Peter Huthwaite Reader, Imperial College London	Examinateur
Alexandre Imperiale Ingénieur chercheur, CEA List	Co-directeur de thèse
Sébastien Imperiale Chargé de recherche, Inria	Co-encadrant de thèse
Philippe Moireau Directeur de recherche, Inria	Directeur de thèse

à Claudirce B. e Renato D.,

Remerciements

Cette thèse a été faite dans un environnement riche et avec des personnes passionnées par leur sujet de travail. Dans ce contexte, je tiens d'abord à te remercier, Philippe Moireau. Tes vastes connaissances m'ont conduit à des compréhensions sous différentes perspectives et je suis reconnaissant de ta passion pour la science et pour les differents domaines traités ici. Ta rigueur scientifique me servira toujours comme inspiration. Merci également pour tes efforts visant à diriger cette thèse en veillant à ce que je tire le meilleur parti de mes capacités, ainsi que de l'environnement dans lequel je me trouvais. Alexandre Imperiale, je tiens aussi à t'exprimer ma profonde gratitude pour ton fort et constant investissement dans l'encadrement depuis mes premiers jours de thèse. Ta capacité unique à décortiquer des concepts complexes et à les rendre compréhensibles m'a permis de maîtriser des sujets divers et de les travailler pour une bonne qualité de cette thèse. Je te remercie également, Sébastien Imperiale, pour ta pédagogie et le partage de tes connaissances en mathématiques, qui ont été fondamentales pour m'aider à faire le lien entre les concepts travaillés au cours de ma thèse. Tes propositions pertinentes m'ont notamment aidé dans ma compréhension en analyse fonctionnelle et acoustoélasticité. Enfin et surtout, c'était un plaisir de vous voir expliquer, développer et discuter des concepts sur le tableau (blanc ou noir) de manière particulièrement captivante.

Merci à Bruno Lombard d'avoir accepté d'être rapporteur de mon manuscrit, pour ta lecture attentive et pour toutes discussions au cours de ces trois années, qui m'ont été précieuses pour ma trajectoire. Je remercie également Ludovic Chamoin d'avoir accepté d'être rapporteur de mon manuscrit, avec une lecture attentive et les perspectives apportées par votre expertise. Je remercie Marc Bonet d'avoir présidé de manière impeccable ma soutenance et pour des discussions apportées. Merci à Maya De Buhan et à Peter Huthwaite d'avoir fait partie de mon jury et aussi d'avoir apporté vos questions et propositions intéressantes.

Pendant ces trois années, j'ai eu l'énorme chance de faire partie des environnements chaleureux et culturellement riches. D'un côté du plateau de Saclay, ce fut un plaisir de faire partie du LSMA, au DIN, CEA List; Merci, Marie Palla, pour ton amitié et compagnie au cours de ces 3 années. Je suis vraiment ravi d'avoir partagé le bureau et le parcours de thèse avec toi! ; Un merci chaleureux à Nouhayla Khalid pour ta compagnie et irremplaçable presence. Tu es une personne très spéciale et avec qui j'ai beaucoup apprécié discuter et d'avoir à mes côtés :) ; Gerardo (inho) Granados, pour ton amitié, ton humeur inégalable et ton soutien dans tous les moments (des événements festifs, notamment); Amond (ami) Allouko, pour ton support et ton amitié depuis notre arrivée en France, le même jour! ; Victor Bussy pour ton animation et spontanéité, ainsi que notre voyage incroyable avec Camille Ly et Emiel Hassefras, les mecs!; Romain Vo (alecrim, haha) pour ton amitié, ton support et ton esprit amusant; à Vinduja Vasanthan pour ton écoute, tes origamis geniaux et ta personalité unique. Merci également à Vivek Nerlikar, Celestine Angla, Jordan Barras, Clément Fisher, Valentin Serey, Maxance Marmonier, Flavien Agon, Alexandre Charau, Imanol Setoain, Laureen Guitard, entre d'autres collègues pour la bonne compagnie et pour les activités ensemble. De l'autre côté du plateau, ce fut un plaisir de faire partie de l'équipe M3DISIM à l'Inria. Il ne manquait pas d'activités sportives. Et d'autres pas trop... haha. Je vous remercie; Mathieu Barré et Tiphaine Delaunay pour votre humeur unique et votre soutien dans ce parcours partagé; Jérôme Diaz pour ta compagnie dans les meilleurs concerts (!!), pour nos discussions et pour ton (très importante) assistance au cours de ma thèse; François Kimmig pour ta sympathie et tes connaissances qui m'apportaient toujours de nouvelles choses à apprendre; L'esprit italien qui était très present, vivant et amusant, avec Jéssica Manganotti et Giulia Merlini qui me manqueront fortement. C'était un plaisir d'être là avec vous; Alice Peyraut (,) pour ta bonne humeur, ta vivacité (!) et ton excellence en coach sportif; Zineb Ramiche et notre partenariat Brarrocos; Je tiens également à remercier Vincenzo Zarra, Gaël Le Ruz, Chloé Giraudet, Mahdi Manoochehrtayebi et Nagham Chibli pour être present dans plusieurs moments de partage et bonheur.

It's impossible to stop there, being part of the european project GW4SHM made me get in touch with amazing people, go to exquisite places and participate to interesting events. It was a pleasure to meet all the people involved in it; I would like to thank you ; Yevgeniya Lugovtsova for greatly managing our project and for all our interesting scientific (and non-scientific) discussions; Daniel Lozano, lover of brazilian soup and such a nice company to have in small towns; Emiel Hassefras for all shared moments, LeKlub (and alikes) and our amazing trip in California; Omar Rasgado, the best latino DJ in Europe; Ahmed Bayoumi for your great and welcoming energy; Panpan Xu for your kindness and for welcoming me at ICL; Aadhik Asokkumar, it was an honnor to have the moments registered by you; Thanks also for the valuable dialogues, foods and drinks together: Blaž Brence, Ferda Gül, Masoud Mohammadgholiha, Enes Savli, Mohsen Barzegar and Mateen Qadri; I would like also to thank all project PIs with whom I had the pleasure of discussing and learning more about several different subjects; You all made my three years rich in knowing different cultures and points of view.

L'ambiance était vraiment incroyable avec vous ! Je vous remercie pour votre amitié, votre soutien, nos voyages, votre présence et des moments au-delà du spécial.

Agradeço à minha mãe Claudirce Barbieri pelo carinho, profundidade e dedicação em minha criação ao longo da minha vida. Agradeço à meu pai Renato Dalmora por sua coragem, seu caracter e por me proporcionar uma base de confiança onde pude me desenvolver. Às minhas amadas irmãs Ana Cláudia Dalmora Büll e Renata Dalmora por terem criados caminhos novos e sempre me incentivarem em meus estudos. Ao meu sobrinho Leonardo por sua personalidade, alegria e carinho que me são fontes de motivação. Queria também deixar um grande abraço aos meus amigos Strauss Vidrich, Adriano Garcia e Vinicius Neves que mesmo de longe sempre se fizeram presentes e importantes em minha vida.

Résumé (Français)

Récemment, avec les avancées technologiques et le vieillissement des structures, l'intérêt pour l'évaluation de l'intégrité et de la durée de vie utile restante des structures s'est accru. Outre les raisons de sécurité, une évaluation efficace et fréquente permet d'éviter une utilisation inadéquat et le remplacement de composants structurels sains. Parmi les structures pour lesquelles ce type d'évaluation est essentiel pour éviter les accidents, on peut citer les ponts, les centrales nucléaires, les avions, les pipelines, etc. Dans ces cas, les défaillances structurelles peuvent entraîner des décès et avoir des répercussions considérables sur l'environnement. Pour des raisons réglementaires, ces structures sont généralement inspectées fréquemment pour détecter tout signe de fatigue ou de défaillance potentielle. Certaines méthodes peuvent potentiellement rendre la structure inopérante de manière permanente, mais nous nous intéresserons ici aux méthodes non destructives, appelées *contrôle non destructif* (CND). Plusieurs méthodes différentes sont employées et validées pour ces tâches, faisant appel à différents phénomènes physiques, dont les principaux sont les ondes élastiques (acoustique, ultrasons, vibrations structurelles) et l'électromagnétisme (lumière visible, champs magnétiques, rayons X).

Parmi les méthodes de CND, les ondes élastiques de faible amplitude peuvent pénétrer de manière non destructive dans les solides et être ensuite utilisées pour extraire les caractéristiques de la structure. Les ondes élastiques à haute fréquence, également appelées ultrasons, sont largement utilisées dans ce contexte [Workman, Kishoni, and Moore, 2007], lorsque l'on souhaite détecter des défauts potentiels, des contraintes, des fissures, de la fatigue ou des propriétés du matériau. L'accès aux caractéristiques des matériaux permet d'évaluer les dangers et de prédire la durée de vie restante des structures. L'interaction des ondes avec les frontières de la structure peut donner lieu à des ondes guidées. Lorsque la propagation est confinée entre des frontières parallèles, interférant et formant des modes de propagation, nous avons les ondes de Lamb [Lamb, 1917]. Ce type d'ondes peut se propager sur de plus longues distances. Les ondes de Lamb sont dispersives et multimodales, c'est-à-dire qu'elles présentent différents modes de propagation par fréquence avec des caractéristiques de dispersion et des vitesses d'onde différentes. Le fait que plusieurs modes, se propageant à des vitesses différentes, soient enregistrés dans un signal nécessite des techniques spéciales de traitement du signal et une distinction complète de ces modes est souvent difficile. Bien que le travail effectué dans cette thèse puisse être appliqué aux ondes élastiques de volume ou guidées, nous nous concentrons principalement sur les ondes de Lamb.

L'utilisation de techniques de CND peut rendre la structure temporairement inopérante, ce qui entraîne des coûts de maintenance. Pour atténuer ce type de problème, les récents progrès technologiques, tels que le développement de systèmes intégrés, l'électronique à faible consommation d'énergie, le traitement des signaux et les capteurs miniaturisés, favorisent le développement de systèmes de surveillance *in situ* pour les structures. Ce domaine de l'ingénierie est appelé *contrôle de santé intégré* (SHM) [Worden et al., 2007; Farrar and Worden, 2010]. Dans le cadre de la surveillance de l'état des structures, l'état d'une structure est surveillé en permanence, ce qui peut s'avérer utile pour les structures des avions [Molent and Aktepe, 2000; Boller, 2001], des ponts [Li and Ou, 2016] et des pipelines [Arun Sundaram, Kesavan, and Parivallal, 2018].

Ce travail traite principalement de l'application des ondes guidées, mais pas exclusivement, dans le contexte du *contrôle de santé intégré* pour les métaux et les composites. Les métaux et les composites sont avant tout des composants essentiels de la structure. Un examen des ondes guidées pour le *contrôle de santé intégré* est présenté dans [Mitra and Gopalakrishnan, 2016]. Un examen spécifique des applications

dans les matériaux composites figure dans [Ricci et al., 2022]. Afin de surveiller en permanence une structure, des capteurs sont installés en permanence *in situ*, ce qui permet de traiter et d'analyser de grandes quantités de données. Des efforts sont faits pour proposer des stratégies permettant d'extraire des informations de ce type de données. Une base de données ouverte contenant des mesures d'ondes guidées est proposée dans [Marzani et al., 2020]. Le principal défi de l'analyse des données SHM est, étant donné que la structure surveillée est en ligne, de traiter les conditions environnementales et opérationnelles (EOC) [Gorgin, Luo, and Wu, 2020] qui affectent les données, réduisant ainsi la prévisibilité du système.

Conditions Environnementales et Opérationnelles (EOCs). Dans une structure en opération, on peut s'attendre à des variations des conditions environnementales telles que la température, l'humidité de l'air et la vitesse du vent, ainsi que des conditions opérationnelles telles que la charge mécanique et la puissance de fonctionnement. Les effets de la température et de la charge mécanique sont les plus étudiés. Dans ce travail, nous proposons un modèle pour la propagation des ondes ultrasonores dans des conditions de charge opérationnelles et une approche basée sur le modèle pour estimer ces conditions. Cela permet d'estimer les conditions de charge et/ou d'éliminer le biais induit par celles-ci dans les mesures. Plus précisément, nous énonçons les objectifs de cette thèse comme suit :

- **Proposer des outils numériques robustes et génériques pour modéliser comment le chargement mécanique affecte la propagation des ultrasons**
- **Proposer une stratégie d'inversion pour estimer la déformation structurelle à l'aide des ultrasons.**

Les effets de la propagation des ondes dans les solides causés par les contraintes internes – y compris les contraintes induites par la charge – sont appelés effets *acoustoélastiques*. Le cadre de modélisation sur lequel cette thèse est basée est la formulation traditionnelle de l'élastodynamique avec des lois constitutives hyperélastiques. Nous nous référons aux livres [Ogden, 1984; Landau, Lifšic, and Landau, 1986; Ciarlet, 1988b; Holzapfel, 2000; Le Tallec, 2009; Chaves, 2013] pour les méthodologies et les descriptions relatives à la mécanique non linéaire et [Chapelle and Bathe, 2011] pour les structures liées aux coques.

En acoustoélasticité, on étudie à la fois les contraintes résiduelles et les contraintes induites par les charges opérationnelles. Du point de vue des contraintes causées par les conditions de charge opérationnelles, en termes de déformation, les ondes élastiques sont superposées aux ondes structurelles. En mécanique non linéaire, lorsque la loi constitutive est non linéaire et/ou que la description géométrique de la déformation est non linéaire, on s'attend à ce que les déformations superposées ne se comportent pas de la même manière. En élasticité linéaire, aucun effet acoustoélastique ne peut se produire, la réponse du matériau à la déformation est modélisée par la loi de Hooke et seuls deux paramètres sont nécessaires pour la modéliser dans les matériaux isotropes, les paramètres de Lamé λ et μ étant des exemples. Si l'on considère les déformations élastiques du deuxième ordre, les matériaux isotropes nécessitent trois paramètres supplémentaires [Hughes and Kelly, 1953; Shams, Destrade, and Ogden, 2011], les constantes élastiques du troisième ordre (TOEC). Murnaghan a proposé dans [Murnaghan, 1937] les paramètres l, m et n et a développé sa théorie sur la déformation finie dans [Murnaghan, 1951]. Effectivement, dans ce travail, nous traitons l'acoustoélasticité induite par la charge comme une déformation incrémentale dans le contexte de l'élastodynamique non linéaire avec un comportement hyperélastique et nous nous référons à [Ogden, 2007; Shams, Destrade, and Ogden, 2011; Abiza, Destrade, and Ogden, 2012] pour une description plus détaillée. Certains travaux y font référence en tant que *small-on-large*, c'est-à-dire une petite déformation superposée à une grande déformation.

Méthodes numériques dédiées. Dans cette thèse, nous proposons un outil numérique générique et robuste dans le domaine temporel pour la propagation des ondes ultrasonores dans des conditions de charge opérationnelles. Pour ce faire, nous calculons tout d'abord la déformation structurelle causée par les forces externes et, ensuite, nous fournissons la déformation structurelle à un solveur élastodynamique linéarisé efficace. Nous utilisons des éléments finis pour résoudre ces deux problèmes. Nous nous référons au problème quasi-statique comme étant le problème du calcul de la déformation structurelle résultant des conditions de chargement. Dans les applications SHM, les structures minces/coquilles telles que les plaques en polymères renforcés de fibres de carbone (PRFC) sont couramment utilisées et il est connu que, dans de telles structures, les méthodes classiques d'éléments finis souffrent d'un verrouillage numérique [Bathe, 2006; Chapelle and Bathe, 2011]. Pour résoudre ce problème, nous utilisons une formulation de coque 3D [Chapelle and Bathe, 2011] qui, en plus de résoudre le verrouillage numérique, le fait sans hypothèse restrictive sur la loi constitutive hyperélastique et la géométrie, de sorte que la même loi constitutive peut être utilisée pour le problème quasi-statique et le problème de propagation des ondes, qui sera présenté ultérieurement. Nous discutons plus en détail de la formulation des coques dans le Chapter 2. Pour le problème de propagation des ondes, nous utilisons une méthode spectrale d'éléments finis (SFEM) d'ordre élevé [Maday and Patera, 1989; Komatitsch et al., 1999; Cohen, 2002; Joly, 2007], connue pour avoir de bonnes propriétés pour la propagation des ondes [Seriani and Priolo, 1994; Basabe, 2007; Seriani and Oliveira, 2008] dans le domaine temporel.

Estimation des conditions de chargement Ce travail est effectué dans le contexte de l'optimisation des moindres carrés. Notre objectif qui est d'estimer la déformation structurelle peut être interprété comme l'estimation d'un paramètre de modèle. Cette estimation est réalisée en minimisant le carré de la différence entre les données mesurées et les données synthétiques générées par le modèle. Dans notre modèle de propagation des ondes, en modélisant l'opération d'observation, une loi constitutive et des paramètres de matériaux appropriés, nous pouvons supposer que la déformation est la seule inconnue. Ensuite, en ajustant la déformation en minimisant la différence entre les données synthétiques de notre modèle et les mesures, nous espérons estimer la déformation de la structure à partir de laquelle les mesures ont été prises. Le problème peut être décrit en minimisant une fonction à valeur réelle qui quantifie l'écart entre les mesures et les observations synthétiques, et un terme supplémentaire peut être ajouté à la fonction pour régulariser l'inversion.

Comme notre objectif est d'utiliser les ondes ultrasoniques pour effectuer une telle tâche, nous nous concentrons sur les problèmes d'optimisation associés à la dynamique des ondes. La minimisation contrainte par un problème d'évolution temporelle est souvent appelée assimilation de données. Initialement développée pour les prévisions météorologiques, elle est aujourd'hui couramment utilisée dans de nombreux domaines. Les méthodes variationnelles et séquentielles sont deux classes principales de méthodes utilisées dans les problèmes d'assimilation de données et toutes deux peuvent être utilisées dans ce contexte [Tarantola, 1984]. Nous utilisons ici des méthodes séquentielles qui utilisent un flux de données de manière séquentielle pour mettre à jour l'estimation par le biais d'un filtre appliqué à l'inadéquation. Plus spécifiquement, nous avons le filtre de Kalman (KF). En tant que méthode de filtrage, il est couramment utilisé dans un contexte où une estimation en temps réel est souhaitée, comme dans la navigation et le suivi. Il a été conçu à l'origine pour les problèmes linéaires, mais des versions modifiées sont disponibles pour les problèmes non linéaires, comme le filtre de Kalman étendu (EKF) – utilisant la dynamique tangente – et le filtre de Kalman Unscented (UKF), une approche "tangent-free".

Comme indiqué précédemment, ces déformations à estimer peuvent être décrites par un problème quasi statique non linéaire. Cela permet de réduire le type de déformations que nous voulons estimer, ce qui conduit au problème de l'estimation de l'amplitude de quelques modes de vibration de la structure et permet d'utiliser des techniques d'ordre réduit. De plus, notre modèle est basé sur la linéarisation de la formulation élastodynamique non linéaire avec une loi constitutive hyperélastique. Les stratégies inverses qui requièrent la dynamique tangente nécessiteraient la dérivée troisième du potentiel élastique. Pour permettre la généricité, cela impliquerait des calculs peu pratiques. Ainsi, pour notre objectif, les méthodes "tangent-free" sont essentielles et nous choisissons la version d'ordre réduit Unscented du filtre de Kalman, telle que décrite dans cette thèse. Enfin, grâce à une combinaison astucieuse d'approches variationnelles et séquentielles, nous obtenons une version itérative du filtre de Kalman Unscented.

Contents

Introduction

This Ph.D. thesis deals with mechanical and numerical, direct and inverse, modeling of ultrasonic guided wave propagation in structures under operational loading conditions. It has been prepared at CEA List, in the *Laboratoire de Simulation, Modélisation et Analyse* (LSMA) and in the MΞDISIM team, an Inria project team joint with Ecole Polytechnique (*Laboratoire de Mécanique des Solides* (LMS)). It was directed by Dr. Philippe Moireau with Dr. Sébastien Imperiale (Inria, MΞDISIM) and Dr. Alexandre Imperiale (CEA List, LSMA).

The LSMA team works in modeling, simulation and analysis for acoustics and electromagnetics, it is part of the Digital Instrumentation Department (DIN), which has been developing the CIVA platform, specifically dedicated to NDT simulation, on an international scale for many years. The MΞDISIM team research is focused on biomechanical mathematical and numerical models of tissues and organs, with a non-exclusive focus on the cardiovascular system. This research was funded by the following project: "GW4SHM" (gw4shm.eu) project from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie, grant number 860104.

Context

Structural Health Monitoring

Recently, interest in assessing the integrity and remaining useful lifetime of structures is growing with technological advances and aging structures. In addition to safety reasons, effective and frequent assessment avoids inefficient use and replacement of healthy structural components. Among the structures where these types of assessments are critical to avoid accidents are bridges, nuclear power plants, aircraft, pipelines, etc. Structural failure in these cases may cause deaths and enormous environmental impacts. For regulatory reasons, such structures are usually inspected frequently for the detection of any sign of fatigue or potential failure. Some methods may potentially render the structure inoperable permanently, but we will be interested here in nondestructive methods, denominated *Nondestructive Testing* (NDT). Several different methods are employed and validated for such tasks, making use of different physical phenomena, with the main ones being elastic waves (acoustics, ultrasound, structural vibrations) and electromagnetics (visible light, magnetic fields, x-ray).

Among the methods for NDT, low amplitude elastic waves can nondestructively penetrate solids and then be used to retrieve characteristics in their volume. High-frequency elastic waves, also called ultrasounds, are widely used in this context [Workman, Kishoni, and Moore, 2007], where one wants to detect potential defects, stresses, cracks, fatigue or properties of the material. Accessing material characteristics allows one to assess dangers and predict the remaining life of structures. Wave phenomena are known to have their energy density largely decreased due to spreading while propagating. Wave interaction with boundaries may result in Guided Waves. When the propagation is confined between parallel boundaries, interfering and forming propagation modes, we have the Lamb waves, first described in [Lamb, 1917]. Another type of guided wave is the Rayleigh wave [Rayleigh, 1885], which occurs and propagates at a surface, hence it presents low penetration capabilities. As the waves are confined in a region, they can propagate longer distances. Lamb waves are dispersive and multi-modal, *i.e.* they present different modes of propagation per frequency with different dispersion characteristics and wave velocities. The fact that several modes propagating at different speeds are recorded in a signal requires special signal processing techniques and complete distinction of those may not be possible. Although the work done in this thesis can be applied to bulk or guided elastic waves, we focus mainly on Lamb waves.

Employing NDT techniques may render the structure temporarily inoperable, leading to maintenance costs. For instance, an airplane must be moved to a NDT workshop to be inspected, or one of these parts must be removed to go through inspection procedures. To overcome this problem, recent advances in technology such as the development of embedded systems, low energy consumption electronics, signal processing and miniaturized sensors are fostering the development of *in situ* monitoring systems for structures. This field of engineering is called *Structural Health Monitoring* (SHM) [Worden et al., 2007; Farrar and Worden, 2010]. In SHM, the state of a structure is continuously monitored and it is potentially useful for aircraft [Molent and Aktepe, 2000; Boller, 2001], bridge [Li and Ou, 2016] and pipeline [Arun Sundaram, Kesavan, and Parivallal, 2018] structures.

This work deals mainly with the application of guided waves, but not exclusively, in the context of *Structural Health Monitoring* for metals and composites. Metals and composites embody, foremost, essential structure components. A review of guided waves for *Structural Health Monitoring* is found in [Mitra and Gopalakrishnan, 2016]. A specific review for applications in composite materials can be found in [Ricci et al., 2022]. For the purpose of monitoring continuously a structure, sensors are permanently installed *in situ*, eventually resulting in large amounts of data to be processed and analyzed. Efforts are made to propose strategies to extract information for this kind of data. An open database with guided wave measurements is proposed in [Marzani et al., 2020]. The main challenge in analyzing SHM data is, as the monitored structure is online, to deal with Environmental and Operational Conditions (EOCs) [Gorgin, Luo, and Wu, 2020] that affect the data, reducing the system's predictability.

Environmental and Operational Conditions (EOCs)

In an online structure, we can expect variations in environmental conditions such as temperature, air humidity and wind speed and in operational conditions such as mechanical loading and operating power. The effects of temperature and mechanical loading are the most studied ones. For exemplification, pipelines used to transport oil and gas may suffer of high gradient of temperature through its thickness as well as increased internal pressure [Dubuc, Ebrahimkhanlou, and Salamone, 2017a]. An aircraft experiences great changes in temperature and loading conditions from the takeoff to the achievement of cruising speed, with load-induced stresses being particularly relevant for military aircraft [Molent and Aktepe, 2000]. Train rails monitoring suffers from stresses and temperature variations when in operation and is also of interest for SHM systems [Chapuis, 2023]. Experimental data and studies on temperature variations [Wilcox et al., 2007; Song, Huang, and Hu, 2013; Dodson and Inman, 2014; Kijanka et al., 2015] show that its effects are greater than the one caused by stresses [Michaels et al., 2009; Pei and Bond, 2016].

One class of approaches to mitigate the difficulties related to EOCs when analyzing SHM data is data-based [Sohn, 2007]. For instance, one strategy is to perform measurements in a given EOC in the structure in pristine conditions and use them as a baseline to compare with future measurements. By subtracting one recent signal from the baseline, one can isolate the new changes in the structure, hence being capable of identifying potential issues [Croxford et al., 2007; Michaels, Lee, and Michaels, 2011; Tschöke et al., 2017]. Consequently, the baseline method requires measurements for all potential EOCs the structure may meet, or a model that provides baselines through interpolation [Yue and Aliabadi, 2020]. The authors in [Wilcox et al., 2007] show that, even for small imperfections, the baseline subtraction for guided waves may perform poorly, requiring additional post-processing of the signals. Also, by acquiring a large data set, one can perform data analysis to separate changes caused by variations in EOCs from changes caused by other phenomena of interest. Machine learning-based techniques are also proposed [Villares Holguin, Hultmann Ayala, and Kubrusly, 2021; Malekloo et al., 2022].

In this work, we propose a model for ultrasonic wave propagation under operational loading conditions and, following a different path, a model-based approach for estimating those. This enables one to estimate the loading condition and/or to remove the bias induced by it from the measurements. More precisely, we enounce the objectives of this thesis as to

- **Propose robust and generic numerical tools to model how mechanical loading affects ultrasound propagation**

and to

- **Propose an inverse strategy to estimate structural deformation using ultrasound.**

Given these objectives, we discuss below methods that can be used to tackle them, starting with the modeling aspects before discussing the estimation methods.

Modeling the effects of mechanical loading on ultrasound propagation

The effects in wave propagation in solids caused by internal stresses – including load-induced stresses – are denominated *acoustoelastic* effects. The modeling framework on which this thesis is based is the traditional elastodynamics formulation with hyperelastic constitutive laws. We refer to the textbooks [Ogden, 1984; Landau, Lifšic, and Landau, 1986; Ciarlet, 1988b; Holzapfel, 2000; Le Tallec, 2009; Chaves, 2013] for methodologies and descriptions related to nonlinear mechanics and [Chapelle and Bathe, 2011] for shell-related structures. In the following, we do first a review of the development and use of acoustoelasticity. Then, we discuss analytical and numerical methods for (guided) wave propagation in stressed structures. Finally, we present the numerical methods used in this thesis.

Acoustoelasticity

In acoustoelasticity, both residual stresses and stresses induced by operational loads are studied. From the perspective of stresses caused by operational loading conditions, in terms of deformation, the elastic waves are superposed to the structural ones. In nonlinear mechanics, where the constitutive law is nonlinear and/or the geometric description of the deformation is nonlinear, it is expected that superposed deformations do not behave the same. In linear elasticity, no acoustoelastic effect can occur, the material response to strain is modeled by Hooke's law and only two parameters are required to model it in isotropic materials, the Lamé parameters λ and μ being examples. Considering second-order elastic deformations, isotropic materials require three more parameters [Hughes and Kelly, 1953; Shams, Destrade, and Ogden, 2011], the third-order elastic constants (TOECs). Murnaghan proposed in [Murnaghan, 1937] the l, m and n parameters and developed further its theory on finite deformation in [Murnaghan, 1951]. Biot, 1940 studies the effects of initial stresses in wave propagation with a focus on seismic Rayleigh waves. The acoustoelastic theory was further developed in [Hughes and Kelly, 1953], where the authors relate the wave speed to the TOECs, which allows the measurement of TOECs using wave propagation. Toupin and Bernstein, 1961 extended their work to arbitrary symmetry while determining restrictions on the formulation such that materials under homogeneous deformation admit plane waves. Further generalization is made in [Thurston and Brugger, 1964; Brugger, 1965] to obtain TOECs from materials of arbitrary symmetries using wave speed measurements. Most metals present crystalline structure at room temperature, motivating the work in [Hiki, 1981], where the author analyzes the high-order elastic constants (third, fourth, ...) from the perspective of interatomic forces in crystal lattices. The acoustoelastic theory is used for measuring residual stresses using ultrasonic waves in [Pao, 1984]. The same author describes acoustoelasticity for orthotropic media in different and useful coordinate systems in [Pao and Gamer, 1985]. The authors in [Hikata et al., 1956] show that the TOECs are sensitive to plastic deformation. Man and Lu, 1987 treat the formulation for residual stresses while distinguishing these from load-induced stresses. As applied loads may induce anisotropy in wave propagation, the angular dependency of wave speeds can also be used to retrieve the TOECs [Gandhi, Michaels, and Lee, 2012]. Acoustoelastic effects in fiber-reinforced composites are discussed in [Chakrapani, 2017; Rauter and Lammering, 2018].

In this work, we treat load-induced acoustoelasticity as an incremental deformation in the context of nonlinear elastodynamics with hyperelastic behavior and we refer to [Ogden, 2007; Shams, Destrade, and Ogden, 2011; Abiza, Destrade, and Ogden, 2012] for further description. Some works refer to it as *small-on-large*, *i.e.* small deformation superposed to large deformation. Here, constitutive laws will be described by hyperelastic potentials (or strain energy density functions) [Spencer, 1982]. This framework facilitates the stability analysis of hyperelastic laws and incremental behavior. Although, in this work, we assess the potential problem of unstable incremental models, we don't intend to propose solutions for it. Several works have studied the topic of potential stability issues. In [Hayes and Green, 1963], the authors study the uniqueness of solutions in the context of acoustoelasticity. Schröder and Neff, 2003 propose a polyconvex anisotropic framework for hyperelastic laws using deformation invariants. Further discussion on the mathematical properties of such hyperelastic problems can be found in [Itskov and Aksel, 2004; Ndanou, Favrie, and Gavrilyuk, 2014; Clayton and Bliss, 2014].

The literature on experimental acoustoelasticity shows consistently that third-order elastic constants are highly dependent on the fabrication process [Thompson, Lee, and Smith, 1986] and vary significantly from specimen to specimen. In [Muir, Michaels, and Michaels, 2009] the authors compute how usual acquisition uncertainties result in large uncertainties on TOECs when obtaining them using the time of flight of waves. Experimental data shows that measurement errors may reach 100% [Semenov and Beltukov, 2020]. Ultrasonic wave speed measurement is commonly used to retrieve TOECs for solid materials [Hughes and Kelly, 1953; Smith, Stern, and Stephens, 1966; Asay and Guenther, 1967; Prosser and Green, 1990; Dubuget et al., 1996; Santos Jr and Bray, 2002; Muir, Michaels, and Michaels, 2009; Takahashi, 2018]. In [Ponschab, Kiefer, and Rupitsch, 2019], the authors retrieved the TOECs not only by using the time of flight but analyzing dispersion curves obtained through laser vibrometry and space-time Fourier transform.

Analytical and numerical models for acoustoelasticity

Analytical formulations for wave propagation in the context of acoustoelasticity are proposed in [Gandhi, Michaels, and Lee, 2012; Pau and Scalea, 2015; Pei and Bond, 2016; Dubuc, Ebrahimkhanlou, and Salamone, 2017a; Dubuc, Ebrahimkhanlou, and Salamone, 2017b; Mohabuth et al., 2018; Dubuc, Ebrahimkhanlou, and Salamone, 2018], requiring numerical solving for some cases. Although results can be obtained for a relatively low cost, such representations of the phenomena are not capable of treating details in more complex and realistic cases commonly found in the context of *Structural Health Monitoring*.

For complex shapes and more complex loading configurations, semi-analytical methods are often used [Chen and Wilcox, 2007; Loveday, Long, and Wilcox, 2012; Peddeti and Santhanam, 2018; Cheng et al., 2020; Abderahmane, Lhémery, and Daniel, 2021], taking advantage of numerical methods such as finite elements to consider stress inhomogeneity in one or two dimensions. A review of numerical methods for guided waves in SHM can be found in [Willberg et al., 2015]. Semi-analytical methods treat the problem in the frequency domain and efficiently compute dispersion curves of guided waves in specified wave guides, given the stress profile through the thickness.

Analytical and semi-analytical methods present some limitations regarding geometry and applied stresses and have as output the dispersion curves and guided wave mode velocities. In this thesis, we propose a generic and robust numerical tool in the time domain for ultrasonic wave propagation under operational loading conditions. We do it by, first, computing the structural deformation caused by external forces and, secondly, providing the structural deformation to an efficient linearized elastodynamics solver. Both steps are done using robust finite element-based numerical methods [Le Tallec, 1994; Bonet and Wood, 1997; Bathe, 2006], allowing for complex geometry and loading conditions. Additionally, any hyperelastic law can be easily implemented by providing its strain energy density function. In the following, we precise the numerical methods used to solve each of the decomposed problems.

Dedicated numerical methods

As previously mentioned, we use finite elements to solve both presented problems. We refer to the quasi-static problem as the problem of computing the structural deformation resulting from loading conditions. In SHM applications, thin/shell structures such as carbon fiber reinforced polymers (CFRP) plates are commonly used and it is known that, in such structures, classical finite elements methods suffer from numerical locking [Bathe, 2006; Chapelle and Bathe, 2011]. To overcome this issue, we use a 3D shell formulation [Chapelle and Bathe, 2011] that, additionally to mitigating numerical locking, does it without restrictive assumptions on the hyperelastic constitutive law and geometry, thus the exact same constitutive law can be used for the quasi-static problem and the, to be presented, wave propagation problem. We discuss further the shell formulation in Chapter 2.

Guided waves are known to be multi-modal, dispersive and have the capability of propagating long distances. Numerical methods for wave propagation may suffer from numerical dispersion and anisotropy effects that accumulate during the simulation, becoming critical in the long-duration simulations needed for guided waves. Here, we use a high-order spectral finite elements method (SFEM) [Maday and Patera, 1989; Komatitsch et al., 1999; Cohen, 2002; Joly, 2007], known to have good properties regarding these issues [Seriani and Priolo, 1994; Basabe, 2007; Seriani and Oliveira, 2008], in the time domain. Using hexahedral elements, we take advantage of an efficient mass lumping and unassembled stiffness matrix application [Alexandre Imperiale and Demaldent, 2019]. The SFEM implementation used here was validated for SHM applications [Mesnil et al., 2021], including curved structures such as pipes.

Estimating loading conditions

When a structure is mechanically loaded, its internal forces equilibrate in reaction to the applied forces, leading to internal stresses. Due to acoustoelastic effects, these additional stresses, which are not present in the natural state of the structure, cause the medium to behave differently when interacting with other things. In some other cases, in the natural state of the structure, the internal stresses are unevenly balanced, causing some regions to concentrate more stresses than others. These residual stresses originate from different phenomena, for instance, when the material is subjected to plastic deformation or an uneven temperature treatment. Residual stresses are potential causes for structural failure as the forces tend to concentrate in specific regions. To assess residual stresses nondestructively, ultrasonics are one of the employed techniques [Pao, 1984; Santos Jr and Bray, 2002]. Here, we will be interested mainly in load-induced stresses. These types of stresses tend to be distributed through relatively large regions, making guided waves especially sensitive to them as the acoustoelastic effect accumulates through their path.

Many works on measuring these type of stresses using ultrasonics were done [Bergman and Shahbender, 1958; Crecraft, 1967; Abderahmane, Lhémery, and Daniel, 2022; Hsu, 1974; Degtyar and Rokhlin, 1995; Shi, Michaels, and Lee, 2013; Ma et al., 2019; Zhu, Tan, and Song, 2020; Zhao et al., 2022]. All of them rely on obtaining the time of flight of bulk or guided waves, requiring specific experimental setup and post-processing techniques that allow the distinction of wave packets. This leads to potential loss of useful information or limits the type of configurations to be assessed. Our objective is to propose an inverse strategy that can use raw ultrasonic data in any loading scenario in the elastic regime and, coupled with our model for wave propagation, it can estimate complex deformation configurations. We present below a review of optimization methods that can potentially be used for this task, and finally, our strategy of choice.

In this work, we interpret the task of estimating loading conditions as a minimization problem. We refer to the textbooks [Ciarlet, 1988a; Nocedal and Wright, 2006; Luenberger and Ye, 2008] for the traditional optimization techniques used here.

Inverse problems associated with wave propagation

This work is done in the context of least squares optimization. Our objective of estimating the structural deformation can be interpreted as estimating a model parameter. This estimation is done by minimizing the squared misfit between measured data, and synthetic data generated by the model. In our wave propagation model, by modeling the measurement setup with proper constitutive law and material parameters, we can assume the deformation to be the only unknown. Then, adjusting the deformation by minimizing the difference between our model synthetic data and the measurements, we expect to estimate the deformation of the structure from which the measurements were taken. The problem can be described by minimizing a real-valued function that quantifies a discrepancy from the measurements to synthetic observations and an extra term can be added to the functional to regularize the inversion.

As our objective is to use ultrasonic waves to perform such a task, we consequently focus on optimization problems associated with wave dynamics. The minimization constrained by a time evolution problem is often referred to as data assimilation. Initially developed for weather forecasting, it is now commonly used in many fields. Variational and sequential methods are two main classes of methods used in data assimilation problems and both can be in this context [Tarantola, 1984]. Variational methods for minimization problems are usually done iteratively by functional gradient descent. At each iteration, it analyses the whole corpus of measurement data to compute the gradient and/or the Hessian. Sequential methods use a stream of data sequentially to update the estimation through a filter applied to the misfit.

In the class of variational methods, strategies known as *4-Dimensional Variational Data Assimilation* (4D-Var) and *Full-Waveform Inversion* (FWI) [Virieux et al., 2017] are equivalent and were first developed for applications related to seismic, meteorology and oceanography, where huge amounts of data are used simultaneously and real-time estimation is not essential. These denominations refer to the fact that inversion is made using both space and time data. While the denomination 4D-Var is employed more for general space-time data, FWI is specially used in the context of wave propagation. To compute the functional gradient, a specific method called the adjoint method is used [Lions, 1971; Cea, 1986; Plessix, 2006]. Traditional algorithms such as conjugate gradient and quasi-Newton are used to perform the descent. In [Tarantola, 1984; Mora, 1987; Pica, Diet, and Tarantola, 1990; Bunks et al., 1995], the method was first developed and used for nonlinear optimization constrained by a seismic wave propagation problem. Using FWI, they were able to use a seismic propagation model to retrieve sub-surface parameters such as bulk modulus and density. By having a wave propagation model, complex wave interactions such as diffraction and reflections can be taken into account, enhancing the capacity to retrieve information for measurements. A review in FWI for geophysical applications can be found in [Virieux and Operto, 2009]. In [Dimet and Talagrand, 1986; Courtier, Thépaut, and Hollingsworth, 1994; Rabier, Thépaut, and Courtier, 1998], the 4D-Var is applied for meteorology. These methods are usually applied with the misfit computed with an L^2-norm, the squared misfit. Other norms related to the mechanical constitutive relation are also proposed and used in structural mechanics [Feissel and Allix, 2007; Bonnet and Aquino, 2015; Nguyen, Chamoin, and Ha Minh, 2022]. Recent works borrow concepts from optimal transport to overcome the cycle skipping [Engquist, Froese, and Yang, 2016; Métivier et al., 2022]. The nonlinear optimization problems solved in these references are analogous to the one we want to perform.

In the class of sequential methods, we have the Kalman Filter (KF). As a filtering method, it is commonly used in a context where real-time estimation is desired, such as in navigation and tracking. It was originally conceived for linear problems but modified versions are available for nonlinear problems, such as the Extended Kalman Filter (EKF) – using the tangent dynamics – and the Unscented Kalman Filter (UKF), a tangent-free approach. Developed in the context of signal filtering for communications and control, it is also used for data assimilation. In the filtering and control context, we refer to the textbooks [Simon, 2006; Bensoussan et al., 2007]. The Kalman Filter algorithm was first used, for data assimilation purposes, in meteorology [Jones, 1965; Courtier et al., 1993; Houtekamer and Zhang, 2016] and in oceanography [Pham, Verron, and Gourdeau, 1998; Evensen, 2003; Bertino, Evensen, and Wackernagel, 2003; Rozier et al., 2007]. In [Julier, Uhlmann, and Durrant-Whyte, 2000], the authors proposed a tangent-free version of the Kalman Filter to be later denominated Unscented Kalman Filter.

When the parameter to be estimated represents only part of the dynamics state, a reduced order of the filter can be used, as applied to realistic ocean circulation models in [Rozier et al., 2007]. The Kalman Filter is also used in the context of structural mechanics [Mariani and Ghisi, 2007; Marchand, Chamoin, and Rey, 2016], comparing EKF and UKF approaches, and for medical applications [Moireau, Chapelle, and Le Tallec, 2008; Caiazzo et al., 2017]. In the context of SHM, it is used for acoustic source localization and compared to a variational method in [Dehghan-Niri, Farhidzadeh, and Salamone, 2013], among other related applications [Wu, Huang, and Huang, 2004; Zou et al., 2015]. An iterated Kalman Filter is applied for seismic in [Eikrem, Nævdal, and Jakobsen, 2019]. Another iterated use of the Kalman Filter is described in [Furuya and Potthast, 2022]. A review of discrete-time filtering formulation in a data assimilation context can be found in [Moireau, 2022].

Choice of method

Both presented classes of methods are suitable for our problem of estimating the deformation through ultrasound. As previously mentioned, these deformations can be described by a nonlinear quasi-static problem. This helps to reduce the type of deformations we want to estimate, leading to the problem of estimating the amplitude of a few modes of vibration of the structure, thus reduced order techniques can be used. Also, our model is based on the linearization of the nonlinear elastodynamic formulation with hyperelastic constitutive law. Inverse strategies that require the tangent dynamics would require the third derivative of the elastic potential. To allow genericity, this would mean impractical computations. So, for our purpose, tangent-free methods are essential and we choose the Reduced Order Unscented version of the Kalman Filter, as described below.

Structure of this thesis

This thesis is composed of two parts, two chapters each. Chapter 1 and Chapter 3 are review chapters presenting the ingredients used in the direct and inverse modeling. Our original contributions are presented in Chapter 2 and Chapter 4, accompanied in both cases by 3D realistic illustrations. The reader may note that different notations may be used in the chapters related to the original works as they are conceived to be standalone articles.

Part I - Structural Health Monitoring under environmental and operational conditions

Chapter 1. Modeling and numerical methods for linear wave propagation problems

In the context of our first objective, we introduce and motivate our approach to model linear wave propagation. Starting from a weak formulation for nonlinear mechanics, we present the basic assumptions for linearization and how hyperelastic laws can be calibrated in a linearized framework. As we will see in the next chapter, linearization will be the basis for representing ultrasonic waves on a deformed structure. Second, for solving numerically a linearized wave problem, we introduced the class of Galerkin methods with energy conservation and stability conditions in an explicit scheme. The implementation of the spectral finite elements method (SFEM), a Galerkin method, is then detailed featuring mass lumping and unassembled stiffness matrix application. Finally, we show the efficiency of high-order SFEM for solving guided wave propagation problems. This is done by using large-scale problems and comparing high-order with low-order finite elements.

Chapter 2. A time-domain spectral finite element method for acoustoelasticity: modeling the effect of mechanical loading on guided wave propagation

In this chapter, we tackle the first objective of this thesis. Using proper assumptions related to *small-on-large* deformations, two problems emerge, the quasi-static problem and an associated linearized wave propagation problem. After carefully reviewing the modeling ingredients to represent the configurations of interest, we propose an original combination of numerical tools that leads to a computationally efficient algorithm. More specifically, we use 3D shell elements for the quasi-static nonlinear problem and the time-domain spectral finite element method to numerically solve the wave propagation problem. Our approach is generic in the sense that it can represent any type of material constitutive law and geometry. We present realistic numerical results on 3D cases related to the monitoring of both isotropic and anisotropic materials, illustrating the genericity and efficiency of our method. We also validate our approach by comparing it to experimental data from the literature. The chapter takes the form of a pre-print article co-authored by André Dalmora, Alexandre Imperiale, Sébastien Imperiale and Philippe Moireau.

Part II - Loading condition estimation in Structural Health Monitoring

Chapter 3. Introduction to least-squares minimization for solving inverse problems
In this chapter, we review some existing data assimilation methods associated with wave propagation problems. By introducing generic optimization problems of interest, we detail how variational and sequential methods are used to solve them. We implement and discuss these methods in one-dimensional wave propagation inverse problems, where linear and nonlinear dependencies are considered. Finally, we conclude and motivate the use of Reduced-order Unscented Kalman Filter to tackle the second objective of this thesis.

Chapter 4. Kalman-based estimation of loading conditions from ultrasonic guided wave measurements
In this chapter, we tackle the second objective of this thesis. We present a model-based procedure to estimate structural deformation using available guided ultrasonic measurements in SHM systems. The linearized model presented in Chapter 2 is used as a forward model for the inversion algorithm. From a methodological point of view, our approach is original since we have proposed an iterated Reduced-Order Unscented Kalman strategy, which we justify as an alternative to a Levenberg-Marquardt strategy for minimizing the non-quadratic least-squares estimation criteria. Therefore from a data assimilation perspective, we provide a quasi-sequential strategy that can valuably replace more classical variational approaches. Our resulting algorithm proves to be computationally very effective, allowing us to successfully apply our strategy to realistic industrial SHM configurations. The chapter takes the form of a pre-print article co-authored by André Dalmora, Alexandre Imperiale, Sébastien Imperiale and Philippe Moireau.

Henceforth, conclusions and perspectives are presented.

Published articles and pre-prints

- Our direct model was first published as a conference paper in an international conference for NDT. Dalmora, André, Alexandre Imperiale, Sébastien Imperiale, and Philippe Moireau. 2022. "A Generic Numerical Solver for Modeling the Influence of Stress Conditions on Guided Wave Propagation for SHM Applications.", American Society of Mechanical Engineers Digital Collection. `https://doi.org/10.1115/QNDE2022-98682`.
- A detailed description of our direct model is written in a pre-print article format (Chapter 2). Dalmora, André, Alexandre Imperiale, Sébastien Imperiale, and Philippe Moireau. 2023. "A time-domain spectral finite element method for acoustoelasticity: modeling the effect of mechanical loading on guided wave propagation". Pre-print ready for submission.
- Our inversion strategy is detailed in a pre-print article format (Chapter 4). Dalmora, André, Alexandre Imperiale, Sébastien Imperiale, and Philippe Moireau. 2023. "Kalman-based estimation of loading conditions from ultrasonic guided wave measurements". Pre-print ready for submission.

Oral communications

International

- 49th A.R.P. in Quantitative Nondestructive Evaluation – QNDE 2022 – San Diego, CA
  A Generic Numerical Solver for Modeling the Influence of Stress Conditions on Guided Wave Propagation for SHM Applications.
- 20th Anglo-French Physical Acoustics Conference – AFPAC 2023 – Fréjus, France
  Numerical solver for the modelling of guided waves propagation under stress condition in the context of SHM applications.
- 13th European Conference On Non-Destructive Testing - ECNDT 2023 – Lisbon, Portugal
  A Generic Numerical Solver for Modeling the Influence of Stress Conditions on Guided Wave Propagation for SHM Applications.
- Isaac Newton Institute – Cambridge, UK
  Estimating operational loading conditions from wave measurements using Kalman filters.

French

- 2ème colloque du GDR MecaWave, 2021 – Ile d'Oléron, France
  Combining shell elements and transient spectral finite elements for guided wave propagation problems in prestressed thin structures.
- 3ème colloque du GDR MecaWave, 2023 – Porquerolles, France
  Estimating the effect of operational loading condition from ultrasonic guided wave measurements using an iterated unscented Kalman Filter.
- Doctoriales de la COFREND 2023 – Marseille, France
  Solveur numérique générique pour la modélisation de l'influence des contraintes mécaniques sur la propagation des ondes guidées pour les applications SHM.
- 16ème Congrès Français d'Acoustique – CFA 2022 – Marseille, France
  Combinaison d'éléments de coque et d'éléments finis spectraux transitoires pour les problèmes de propagation d'ondes guidées dans des structures déformées.
- Jean Mandel Symposium 2023 – Palaiseau, France
  Estimating the effect of operational loading conditions from ultrasonic guided wave measurements using an iterated Unscented Kalman Filter.

Scientific exchanges

In the context of the "GW4SHM" project, six training weeks were organized online and across Europe (London, Tallinn, Paris and Lisbon), bringing together the doctoral students and the academic and industrial "GW4SHM" participants. Also, as part of this european consortium, I visited for a period of one month each

• Imperial College London: Non-Destructive Evaluation Group – South Kensington, London, UK
Subject: Numerical methods for wave propagation.
• Airbus Defense and Space: Structural Integrity and SHM – Manching, Germany
Subject: Wave propagation in axially loaded stratified composite plates.

Software development

This work was done in a rich environment for the development of numerical applications. At CEA List, I could code the linearized wave propagation model presented in this work using their spectral finite elements core, used in `CIVA`[1] and currently developed by Dr. Alexandre Imperiale, a commercial simulation software specialized in NDT and SHM. Their numerical structure enabled efficient implementation of this work in large realistic cases and to run a considerable amount of those in parallel for the inverse problem. I developed a Python library for the several inverse methods used for

---

[1] `https://www.extende.com`

the illustrations presented in Chapter 3. At Inria, in the MΞDISIM team, their finite elements library `MoReFEM`[2] with the implementation of the 3D shell finite elements method allowed me to implement different hyperelastic laws and invariants to account for transverse isotropy to model stratified composites. Using `MoReFEM`, I was able to validate the tangent stiffness operator implemented in this work. Also in the MΞDISIM team, I have participated in their `AKILLES`[3] Library development that was therefore used for the large-scale illustrations for inversion presented in Chapter 4. I deeply appreciate the help I received from Jérôme Diaz with the `MoReFEM` Library and the implementation of the `AKILLES` Library done by Laurent Steff.

---

[2] `https://gitlab.inria.fr/MoReFEM`

[3] `https://gitlab.inria.fr/AKILLES/AKILLES`

One looks back with appreciation to the brilliant teachers, but with gratitude to those who touched our human feelings. The curriculum is so much necessary raw material, but warmth is the vital element for the growing plant and for the soul of the child.
- Carl Jung

Dawn Treader
William Ackerman

Part I

Structural Health Monitoring under operational conditions

Chapter 1

Modeling and numerical methods for linear wave propagation problems

We model the linear wave propagation using classical linear elastodynamics and an efficient method to solve it numerically using spectral finite elements. Starting from the nonlinear formulation for mechanics, we discuss hyperelastic laws for the isotropic and the transversely isotropic case with a few examples of hyperelastic potentials. Assuming a small strain, the linear elastodynamic problem naturally emerges. From the nonlinear hyperelastic laws, we discuss how they are used within the linear framework. Finally, we present the spectral finite elements method to solve efficiently the wave problem using a second-order Leapfrog scheme with lumped mass matrix and unassembled stiffness matrix applications. A case with realistic dimensions is used to illustrate its performance.

Contents

1.1 Nonlinear mechanics

In this section, we introduce the classical nonlinear mechanics formulation used as a base for our modeling. First, we define the domains that represent the configuration of the structure to be modeled. We consider two configurations:

- **Reference configuration $\hat{\Omega}$**: Also called natural configuration, no external forces are acting in the structure and it is in internal equilibrium. We denote $\hat{\Omega}$ for the domain of the reference configuration and $\hat{x}$ its coordinates system, also called Lagrangian referential.
- **Deformed configuration Ω**: Configuration resulted from external forces acting in the structure in its reference configuration. We denote $\Omega(t)$ for the time-dependent domain of the deformed configuration and x its coordinates system, also called Eulerian referential.

The structural deformation can be represented by a bijective mapping between the configurations, namely

$$\begin{aligned} \phi(t)\ :\ \hat{\Omega} &\longmapsto \Omega(t) \\ \hat{x} &\longrightarrow x(\hat{x},t) = \phi(\hat{x},t). \end{aligned}$$

Associated boundaries are denoted as $\hat{\Gamma}$ and Γ. Figure 1.1 depicts the described notations and relation, with D and N subscripts referring to Dirichlet and Neumann boundary conditions, respectively. With this mapping between coordinates, we can define the displacement field

$$\boldsymbol{u}(\hat{\boldsymbol{x}},t) = \boldsymbol{x}(\hat{\boldsymbol{x}},t) - \hat{\boldsymbol{x}},$$

which characterizes the displacement of a point from the reference configuration to the deformed configuration.

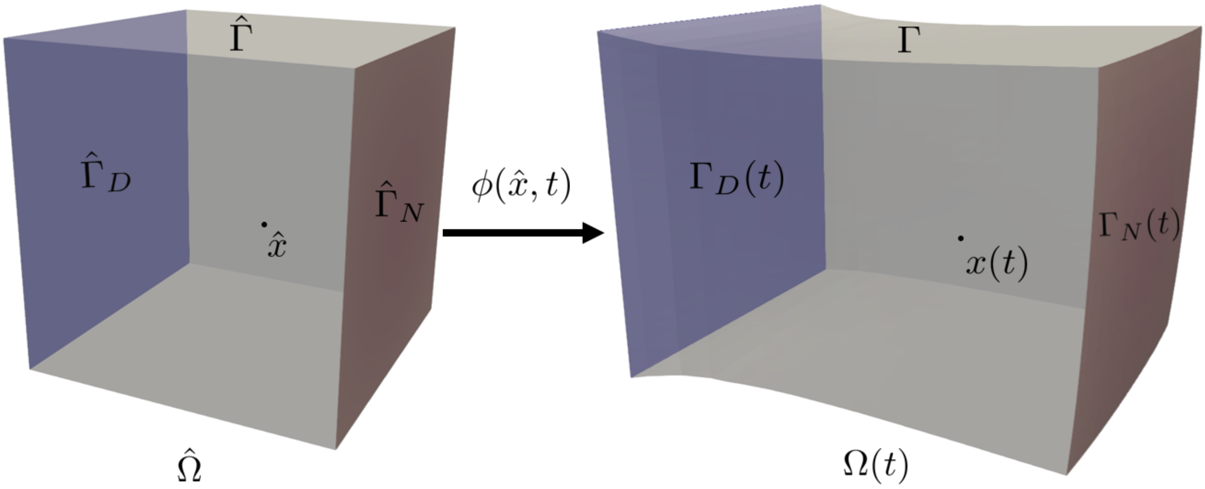


Figure 1.1 – Reference and deformed configurations and their associated notations.

The equation of motion in continuum media with the associated boundary conditions over the domain Ω reads [Ogden, 1984; Temam, Roger and Miranville, Alain, 2005; Chaves, 2013],

$$\begin{cases} \rho(\boldsymbol{x})\boldsymbol{\gamma}(\boldsymbol{x},t) - \boldsymbol{\nabla}_{\boldsymbol{x}}\cdot\boldsymbol{\sigma}(\boldsymbol{x},t) = \rho(\boldsymbol{x})\boldsymbol{f}(x,t) & \text{in } \Omega(t), & (1.1) \\ \boldsymbol{\sigma}(\boldsymbol{x},t)\cdot\boldsymbol{n}(\boldsymbol{x},t) = 0 & \text{in } \Gamma(t), & (1.2) \\ \boldsymbol{\sigma}(\boldsymbol{x},t)\cdot\boldsymbol{n}(\boldsymbol{x},t) = \boldsymbol{f}^s(\boldsymbol{x},t) & \text{in } \Gamma_N(t), & (1.3) \\ \boldsymbol{u}(\boldsymbol{x},t) = 0 & \text{in } \Gamma_D(t), & (1.4) \end{cases}$$

where γ is the acceleration, $\boldsymbol{\sigma}$ is the Cauchy stress tensor that relates strain and stresses through a constitutive law, $\nabla_{\boldsymbol{x}}\cdot$ is the divergence *w.r.t.* the deformed domain coordinates and ρ is the density of the material in the deformed configuration. Additionally, the body and surface forces were introduced as $f(\boldsymbol{x},t)$ and $f^s(\boldsymbol{x},t)$, respectively. The initial conditions are to be defined. The density is included on the right-hand side to model mass-dependent body forces such as gravity.

1.1.1 Weak formulation in the Lagrangian referential

In this section, we rewrite the equations of motion in their weak form and in the Lagrangian referential. We use tensor notation and operations as presented in Section A.1. First, we integrate the equation (1.1) over the domain Ω and multiply it by an arbitrary virtual displacement $w \in \mathcal{V}(\Omega)$, resulting in

$$\int_{\Omega(t)} \rho \boldsymbol{\gamma} \cdot \boldsymbol{w} \, \mathrm{d}\Omega - \int_{\Omega(t)} (\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot \boldsymbol{\sigma}) \cdot \boldsymbol{w} \, \mathrm{d}\Omega = \int_{\Omega(t)} \rho \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\Omega,$$

where $\mathcal{V}(\Omega)$ is the space of admissible displacements to be defined. For brevity, the spatial and time dependencies are omitted. Second, we use Green's formula on the second term and the surface boundary conditions to obtain

$$\int_{\Omega(t)} \rho \boldsymbol{\gamma} \cdot \boldsymbol{w} \, \mathrm{d}\Omega + \int_{\Omega(t)} \boldsymbol{\sigma} : \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{w} \, \mathrm{d}\Omega = \int_{\Omega(t)} \rho \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\Omega + \int_{\Gamma_N(t)} \boldsymbol{f}^s \cdot \boldsymbol{w} \, \mathrm{d}\Gamma. \tag{1.5}$$

The first two terms are related to the inertial and internal virtual work energies, respectively, and the last two terms are related to external virtual work energies. Using the bijective mapping ϕ between domains, we can define the deformation gradient and its Jacobian as

$$\boldsymbol{F}(\hat{\boldsymbol{x}},t) = \boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \phi(\hat{\boldsymbol{x}},t), \quad J = \det F.$$

The following compositions of functions and variables with $\boldsymbol{\phi}(\hat{x},t)$ will be used

$$\boldsymbol{\sigma} \circ \boldsymbol{\phi}, \boldsymbol{f} \circ \boldsymbol{\phi}, \boldsymbol{f}^s \circ \boldsymbol{\phi}, \boldsymbol{w} \circ \boldsymbol{\phi}, \rho \circ \boldsymbol{\phi},$$

without specific notation. Due to mass conservation, we can define the density $\widehat{\rho}(\hat{\boldsymbol{x}}) = J\rho(\phi(\hat{\boldsymbol{x}},t))$. In a change of variables, the infinitesimal volume and surface [Chaves, 2013] from one configuration to another is transformed as

$$\mathrm{d}\Omega = J \mathrm{d}\hat{\Omega}, \quad \mathrm{d}\Gamma = J \| \boldsymbol{F}^{-\top} \cdot \hat{\boldsymbol{n}} \| \mathrm{d}\hat{\Gamma}.$$

Therefore, by a change of variables in (1.5) to pullback to the reference configuration where

$$\boldsymbol{\gamma}(\hat{\boldsymbol{x}},t) = \frac{\partial^2 \boldsymbol{u}}{\partial t^2}(\hat{\boldsymbol{x}},t),$$

and using mass conservation and the fact that $\boldsymbol{\nabla}_{\boldsymbol{x}} \circ = \boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \circ \cdot \boldsymbol{F}^{-1}$, we have

$$\int_{\hat{\Omega}} \widehat{\rho} \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega} + \int_{\hat{\Omega}} \boldsymbol{\sigma} : (\boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \boldsymbol{w} \cdot \boldsymbol{F}^{-1}) \, J \mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^s \cdot \boldsymbol{w} \, J \| \boldsymbol{F}^{-\top} \cdot \hat{\boldsymbol{n}} \| \mathrm{d}\hat{\Gamma} = \int_{\hat{\Omega}} \widehat{\rho} \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega}.$$

Using the first Piola-Kirchhoff stress tensor we rewrite the internal work term

$$\int_{\hat{\Omega}} \widehat{\rho} \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega} + \int_{\hat{\Omega}} \boldsymbol{T} : \boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \boldsymbol{w} \, \mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^s \cdot \boldsymbol{w} \, J \| \boldsymbol{F}^{-\top} \cdot \hat{\boldsymbol{n}} \| \mathrm{d}\hat{\Gamma} = \int_{\hat{\Omega}} \widehat{\rho} \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega}.$$

Moreover, with the second Piola-Kirchhoff (symmetric) stress tensor

$$\boldsymbol{\Sigma} = \boldsymbol{F}^{-1} \cdot \boldsymbol{T} = J\boldsymbol{F}^{-1} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{F}^{-\top}.$$

we have

$$\int_{\hat{\Omega}} \widehat{\rho} \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega} + \int_{\hat{\Omega}} \boldsymbol{\Sigma} : (\boldsymbol{F}^{\top} \cdot \boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \boldsymbol{w}) \, \mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^s \cdot \boldsymbol{w} \, J \| \boldsymbol{F}^{-\top} \cdot \hat{\boldsymbol{n}} \| \mathrm{d}\hat{\Gamma} = \int_{\hat{\Omega}} \widehat{\rho} \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\hat{\Omega}.$$

The Green-Lagrange strain tensor is

$$\boldsymbol{e}(\boldsymbol{u}) = \frac{1}{2}(\boldsymbol{C} - 1) = \frac{1}{2}\Big(\nabla_{\hat{x}}\boldsymbol{u} + (\nabla_{\hat{x}}\boldsymbol{u})^{\top} + (\nabla_{\hat{x}}\boldsymbol{u})^{\top} \cdot \nabla_{\hat{x}}\boldsymbol{u}\Big), \tag{1.6}$$

with $\boldsymbol{C} = \boldsymbol{F}^{\top}\boldsymbol{F}$ being the right Cauchy-Green deformation tensor. Its derivative (Definition A.2.7) is

$$\mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u})\boldsymbol{w} = \frac{1}{2}\Big(\boldsymbol{F}^{\top} \cdot \nabla_{\hat{x}}\boldsymbol{w} + (\nabla_{\hat{x}}\boldsymbol{w})^{\top} \cdot \boldsymbol{F}\Big).$$

As the stress tensor is again symmetric, we can use Property A.1.1 to obtain a weak form of the elastodynamic problem in the reference configuration $\hat{\Omega}$ as presented below.

Nonlinear elastodynamic weak formulation $\forall t > 0$, find $u(t)$ such that

$$\forall w \in \mathcal{V}(\hat{\Omega}), \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}(u, w)_{\mathrm{L}^2(\hat{\Omega})} + \langle A(u), w\rangle = l(w) \tag{1.7}$$

with

$$\begin{aligned}
(u, w)_{\mathrm{L}^2(\hat{\Omega})} &= \int_{\hat{\Omega}} \widehat{\rho}\, \boldsymbol{u} \cdot \boldsymbol{w}\, \mathrm{d}\hat{\Omega}, \\
\langle A(u), w\rangle &= \int_{\hat{\Omega}} \boldsymbol{\Sigma}\big(\boldsymbol{e}(\boldsymbol{u})\big) : \big(\mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u})\boldsymbol{w}\big)\, \mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^{s} \cdot \boldsymbol{w}\, J\|\boldsymbol{F}^{-\top} \cdot \hat{\boldsymbol{n}}\|\mathrm{d}\hat{\Gamma}, \\
l(w) &= \int_{\hat{\Omega}} \widehat{\rho}\boldsymbol{f} \cdot \boldsymbol{w}\, \mathrm{d}\hat{\Omega}.
\end{aligned}$$

Remark. *For sake of brevity, the notation* $\langle\cdot,\cdot\rangle$ *will represents the dual product* $\langle\cdot,\cdot\rangle_{\mathcal{V}'\mathcal{V}}$ *(Definition A.2.6).*

1.1.2 Constitutive laws

The stress-strain relation is characterized by a constitutive law, *i.e.* relation between the second Piola-Kirchhoff stress tensor Σ and the Green-Lagrange strain tensor e, which was considered arbitrary in the presented model. In this work, we will be interested in modeling ultrasonic waves in materials such as metals and carbon fiber composites in elastic regime deformations. In the finite deformations elastic regime presenting nonlinearities, we choose hyperelastic constitutive laws.

Hyperelastic laws. In the case of hyperelastic materials, the stress tensor Σ depends only on space and current state of deformation [Ogden, 1984]. Additionally, they do not dissipate energy when subjected to cyclic homogeneous deformations. In such case, the stress tensor Σ can be derived from an (hyperelastic) potential $\mathscr{W}$, also called strain energy density function

$$\boldsymbol{\Sigma}(\boldsymbol{e}) = \mathrm{D}_{\boldsymbol{e}}\mathscr{W}(\boldsymbol{e}). \tag{1.8}$$

Depending on the symmetries presented by the constitutive behavior of the material, this energy density can expressed using an associated set of invariants of the deformation tensor [Spencer, 1982]. For instance, isotropic materials have their energy density fully described using the first three invariants and transversely isotropic materials make use of the first five invariants. Invariants of the Green-Lagrange strain tensor can be used, but here we use the invariants of the right Cauchy-Green tensor as follows

$$\begin{aligned}
&I_1 = \mathrm{tr}(\boldsymbol{C}), && I_2 = \frac{1}{2}\Big(\mathrm{tr}(\boldsymbol{C})^2 - \mathrm{tr}(\boldsymbol{C}^2)\Big), && I_3 = \det \boldsymbol{C}, \\
&I_4 = \boldsymbol{a} \cdot \boldsymbol{C} \cdot a, && I_5 = \boldsymbol{a} \cdot \boldsymbol{C}^2 \cdot a, &&
\end{aligned}$$

where $\boldsymbol{a}$ is a principal direction defining the plane of transverse isotropy, when applicable. Defining the following generic dependency

$$\mathscr{W} = \mathscr{W}(I_1, I_2, \ldots, I_n),$$

we can express, by using the chain rule, the energy density derivatives *w.r.t.* the strain tensor as

$$\mathrm{D}_{\boldsymbol{e}}\mathscr{W} = \sum_{i=1}^{n} \frac{\partial \mathscr{W}}{\partial I_i} \frac{\partial I_i}{\partial \boldsymbol{C}} : \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{e}} = 2 \sum_{i=1}^{n} \frac{\partial \mathscr{W}}{\partial I_i} \frac{\partial I_i}{\partial \boldsymbol{C}}. \tag{1.9}$$

This form of representing the potential derivatives has a practical use when constructing a generic solver, where the constitutive law can be defined by only giving its potential derivatives as input.

Examples of hyperelastic laws. As mentioned above, the hyperelastic constitutive law can be fully described by a scalar potential. We present here some examples of constitutive laws that can be found in [Ciarlet, 1988b; Chaves, 2013]:

— The **Saint-Venant-Kirchhoff** (SVK):

$$\mathscr{W}^{\mathrm{SVK}} = \frac{(\lambda + 2\mu)}{8} I_1^2 - \frac{(3\lambda + 2\mu)}{4} I_1 - \frac{\mu}{2} I_2 + \frac{(9\lambda + 6\mu)}{8} \tag{1.10}$$

— The **Compressible Neo-Hookean** (CNH):

$$\mathscr{W}^{\mathrm{CNH}} = c_1(\sqrt{I_3} - 1)^2 + c_2\big(I_1 - 3 - \log(I_3)\big). \tag{1.11}$$

— The **Murnaghan**'s (MUR) [Murnaghan, 1951]:

$$\mathscr{W}^{\mathrm{MUR}}(\lambda, \mu, l, m, n) = \frac{\lambda + 2\mu}{8}(I_1 - 3)^2 - \frac{\mu}{2}(3 - 2I_1 + I_2) + \frac{l + 2m}{24}(I_1 - 3)^3 \tag{1.12}$$

$$- \frac{m}{4}(9I_1 - 3I_2 - 2I_1^2 + I_1 I_2 - 9) - \frac{n}{8}(I_1 - I_2 + I_3 - 1). \tag{1.13}$$

The parameters λ and μ are the classic Lamé parameters and l, m and n are the third-order elastic coefficients [Hughes and Kelly, 1953]. When dealing with materials with transversely isotropic symmetries, the invariants I_4 and I_5 are included. This type of symmetry is present, for instance, in a layer of fiber-reinforced polymers where the fibers are oriented in a principal axis. The Carbon Fiber Reinforced Polymer (CFRP) is an example of such material. A Compressible Neo Hookean type energy density is proposed in [Bonet and Wood, 1997]

— Transversely Isotropic law based on CNH (TI):

$$\mathscr{W}^{\mathrm{TI}}(\lambda, \mu, \alpha, \beta, \gamma) = \frac{\mu}{2}(I_1 - 3) - \mu \ln J + \frac{\lambda}{2}(J - 1)^2 + \tag{1.14}$$

$$[\alpha + \beta \ln J + \gamma(I_4 - 1)](I_4 - 1) - \frac{\alpha}{2}(I_5 - 1). \tag{1.15}$$

For this law to be used, additionally to the Lamé parameters, the parameters α, β and γ must be given for the modeled material. These parameters can be calibrated with the linear elasticity tensor as presented in Section 1.2.

1.1.3 Formulation for the quasi-static case

When we consider that the displacement field does not vary in time and the system is in static equilibrium, neglecting the dynamic effects, we call it the quasi-static problem. The quasi-static problem is stated below.

Nonlinear quasi-static weak formulation Find u such that,

$$\forall w \in \mathcal{V}(\hat{\Omega}), \qquad \langle A(u), w\rangle = l(w) \tag{1.16}$$

with

$$\begin{aligned}\langle A(u), w\rangle &= \int_{\hat{\Omega}} \boldsymbol{\Sigma}\big(\boldsymbol{e}(\boldsymbol{u})\big) : \big(\mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u})\boldsymbol{w}\big)\,\mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^s \cdot \boldsymbol{w}\; J\|\boldsymbol{F}^{-\top}\cdot\hat{\boldsymbol{n}}\|\mathrm{d}\hat{\Gamma},\\ l(w) &= \int_{\hat{\Omega}} \widehat{\rho}\boldsymbol{f}\cdot\boldsymbol{w}\,\mathrm{d}\hat{\Omega}.\end{aligned}$$

None of the variables is time-dependent. When solving iteratively the nonlinear problem (1.16), as it will be presented in Chapter 2, the tangent of the stiffness operator A is also needed at each iteration. For hyperelastic laws, this requires the differentiation of $\boldsymbol{\Sigma}$

$$\mathrm{D}_{\boldsymbol{e}}\boldsymbol{\Sigma}(\boldsymbol{e}) = \mathrm{D}^2_{\boldsymbol{e}}\mathscr{W}(\boldsymbol{e}), \tag{1.17}$$

as defined in Definition A.2.8. Expressing this differentiation *w.r.t.* the invariants derivatives we have

$$\mathrm{D}^2_{\boldsymbol{e}}\mathscr{W} = 4\sum_{i=1}^{n}\frac{\partial\mathscr{W}}{\partial I_i}\frac{\partial^2 I_i}{\partial\boldsymbol{C}\partial\boldsymbol{C}} + 4\sum_{i,j=1}^{n}\frac{\partial^2\mathscr{W}}{\partial I_i\partial I_j}\frac{\partial I_i}{\partial\boldsymbol{C}}\otimes\frac{\partial I_j}{\partial\boldsymbol{C}}, \tag{1.18}$$

requiring as input for the implementation in a generic solver, the second derivatives of the invariants and of the hyperelastic potential.

1.2 Modeling linear wave propagation problems

When the source terms are of small amplitude, such as when we are exciting ultrasonic waves in the material with traditional piezoelectric transducers, we expect the displacement field to satisfy $\|u\| \ll 1$ in a suitable norm. This hypothesis motivates the linearization of the presented formulations. In this section, we start from the nonlinear elastodynamic formulation presented above and linearize it around the reference configuration. Assuming $\|u\| \ll 1$ and $\|\nabla_{\hat{x}} u\| \ll 1$ in given norms, the following approximations are done, [Le Tallec, 2009; Chaves, 2013]

$$F \approx \mathbb{1}, \quad J \approx 1, \quad \Omega(t) \approx \hat{\Omega}, \quad \boldsymbol{e}(\boldsymbol{u}) \approx \boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\Big(\nabla_{\hat{x}}\boldsymbol{u} + (\nabla_{\hat{x}}\boldsymbol{u})^\top\Big).$$

Then, we have that

$$\mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{u})\cdot\boldsymbol{w} \approx \boldsymbol{\varepsilon}(\boldsymbol{w}), \quad \forall\boldsymbol{w}\in\mathcal{V}(\hat{\Omega}), \quad \text{and} \quad \boldsymbol{\Sigma}\big(\boldsymbol{e}(\boldsymbol{u})\big) \approx \mathrm{D}_{\boldsymbol{u}}\boldsymbol{\Sigma}(\boldsymbol{0})\cdot\boldsymbol{u}.$$

Using Definition A.1.3, the stress-strain relation becomes

$$\begin{aligned}\mathrm{D}_{\boldsymbol{u}}\boldsymbol{\Sigma}(\boldsymbol{0})\cdot\boldsymbol{u} &= \mathrm{D}^2_{\boldsymbol{e}}\mathscr{W}(\boldsymbol{0}) : \mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{0})\cdot\boldsymbol{u}\\ &= \boldsymbol{C} : \boldsymbol{\varepsilon}(\boldsymbol{u}),\end{aligned}$$

where we defined the fourth-order elasticity tensor $\boldsymbol{C}$. Finally, the stress tensor results in

$$\boldsymbol{\sigma} = J^{-1}\boldsymbol{F}\cdot\boldsymbol{\Sigma}\cdot\boldsymbol{F}^\top \approx \boldsymbol{C} : \boldsymbol{\varepsilon}(\boldsymbol{u}),$$

known as Hooke's law. The strong formulation assumes the form

$$\widehat{\rho}\frac{\partial^2\boldsymbol{u}}{\partial t^2} - \nabla_{\hat{x}}\cdot\Big(\boldsymbol{C}:\boldsymbol{\varepsilon}\Big) = \widehat{\rho}\boldsymbol{f} \quad \text{in } \hat{\Omega}, \tag{1.19}$$

with space and time dependencies omitted. We can define then the linear weak formulation below, for

$$\mathcal{V}(\hat{\Omega}) = \{v \in \mathrm{H}^1(\hat{\Omega})^3 \mid \boldsymbol{v} = 0 \text{ in } \hat{\Gamma}_D\}.$$

Linear elastodynamic weak formulation $\forall t > 0$, find $u(t) \in \mathcal{V}(\hat{\Omega})$ such that

$$\forall w \in \mathcal{V}(\hat{\Omega}), \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}(u, w)_{\mathrm{L}^2(\hat{\Omega})} + \langle Au, w\rangle = l(w) \tag{1.20}$$

with

$$\begin{aligned}
(u, w)_{\mathrm{L}^2(\hat{\Omega})} &= \int_{\hat{\Omega}} \widehat{\rho}\, \boldsymbol{u} \cdot \boldsymbol{w}\, \mathrm{d}\hat{\Omega},\\
\langle Au, w\rangle &= \int_{\hat{\Omega}} \boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{C} : \boldsymbol{\varepsilon}(\boldsymbol{w})\, \mathrm{d}\hat{\Omega}\\
l(w) &= \int_{\hat{\Omega}} \widehat{\rho} \boldsymbol{f} \cdot \boldsymbol{w}\, \mathrm{d}\hat{\Omega} + \int_{\hat{\Gamma}_N} \boldsymbol{f}^s \cdot \boldsymbol{w}\, \mathrm{d}\hat{\Gamma}.
\end{aligned}$$

1.2.1 Constitutive laws in the linear framework

As shown above, in the linear elastodynamic case using hyperelastic laws, the stress-strain relation is given by the elasticity tensor

$$\boldsymbol{C} = \mathrm{D}^2_{\boldsymbol{e}} \mathscr{W}(\boldsymbol{0}).$$

We analyze here the given hyperelastic laws in Section 1.1.2 and how they are represented as the elasticity tensor. Using the engineering (or Voigt) notation [Chaves, 2013], we can represent the tensors as matrix and tensor contractions as scalar products. To assemble the elasticity tensor in the following representation, attention must be given to the differentiation $\mathrm{D}^2_{\boldsymbol{e}}\mathscr{W}$ to ensure coherence between the operator and to what it is applied. Further discussion on compressed matrix representation can be found in [Helnwein, 2001]. The elasticity, strain and stress tensors in Voigt notation are

$$C^v = \begin{pmatrix} C_{1111} & C_{1122} & C_{1133} & C_{1123} & C_{1113} & C_{1112}\\ C_{1122} & C_{2222} & C_{2233} & C_{2223} & C_{2213} & C_{2212}\\ C_{1133} & C_{2233} & C_{3333} & C_{3323} & C_{3313} & C_{3312}\\ C_{1123} & C_{2223} & C_{3323} & C_{2323} & C_{2313} & C_{2312}\\ C_{1113} & C_{2213} & C_{3313} & C_{2313} & C_{1313} & C_{1312}\\ C_{1112} & C_{2212} & C_{3312} & C_{2312} & C_{1312} & C_{1212} \end{pmatrix}, \ \varepsilon^v = \begin{pmatrix} \varepsilon_{11}\\ \varepsilon_{22}\\ \varepsilon_{33}\\ 2\varepsilon_{23}\\ 2\varepsilon_{13}\\ 2\varepsilon_{12}, \end{pmatrix} \text{ and } \sigma^v = C^v\varepsilon^v = \begin{pmatrix} \sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{13}\\ \sigma_{12} \end{pmatrix}.$$

We have the final format in Voigt notation by defining C_{ij}, ε_i and σ_i

$$C^v = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16}\\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26}\\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36}\\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46}\\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56}\\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix}, \ \varepsilon^v = \begin{pmatrix} \varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6 \end{pmatrix} \text{ and } \sigma^v = \begin{pmatrix} \sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6 \end{pmatrix} = C^v\varepsilon^v.$$

Different classes of symmetries for the material will lead to a reduction in the number of independent components. The elasticity tensor related to the Saint-Venant-Kirchhoff constitutive law (1.10) becomes the traditional Hooke's law with the Lamé parameters

$$C^{v,\mathrm{SVK}} = \begin{pmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0\\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0\\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0\\ 0 & 0 & 0 & \mu & 0 & 0\\ 0 & 0 & 0 & 0 & \mu & 0\\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}. \tag{1.21}$$

Using the potential of the hyperelastic constitutive laws presented in Section 1.1.2, we can calibrate its parameters *w.r.t.* the elasticity tensor. For instance, for the Compressible Neo Hookean, writing $\mathrm{D}^2_{\boldsymbol{e}}\mathscr{W}(\mathbf{0})$ in the Voigt notation yields

$$C^{v,\mathrm{CNH}} = \begin{pmatrix} 2c_1+4c_2 & 2c_1 & 2c_1 & 0 & 0 & 0 \\ 2c_1 & 2c_1+4c_2 & 2c_1 & 0 & 0 & 0 \\ 2c_1 & 2c_1 & 2c_1+4c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2c_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2c_2 \end{pmatrix},$$

which, as a result, can be calibrated *w.r.t.* the Lamé parameters yielding

$$c_1 = \frac{\lambda}{2}, \quad c_2 = \frac{\mu}{2}.$$

Another example where the calibration is considered is the case of transversely isotropic laws. In this case, five parameters are needed to describe its behavior. For the principal axis being the first one, the behavior of a transversely isotropic material can be described by the parameters C_{11}, C_{22}, C_{12}, C_{13} and C_{66} in an elasticity tensor of the form

$$C^{v,\mathrm{TI}} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2}(C_{11}-C_{12}) & & \\ 0 & 0 & 0 & 0 & C_{66} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}. \tag{1.22}$$

By developing the elasticity tensor $\mathrm{D}^2_{\boldsymbol{e}}\mathscr{W}(\mathbf{0})$ for the transverse isotropic law (1.14), we have

$$C^{v,\mathrm{TI}} = \begin{pmatrix} -4\alpha+4\beta+8\gamma+\lambda+2\mu & 2\beta+\lambda & 2\beta+\lambda & 0 & 0 & 0 \\ 2\beta+\lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ 2\beta+\lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & -\alpha+\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha+\mu \end{pmatrix},$$

and the following relations are retrieved

$$\lambda = C_{23}, \qquad \mu = \frac{C_{22}-C_{23}}{2},$$

$$\alpha = \frac{C_{22}-C_{23}-2C_{66}}{2}, \qquad \beta = \frac{C_{12}-C_{23}}{2},$$

$$\gamma = \frac{C_{11}-2C_{12}+C_{22}-4C_{66}}{8}.$$

This calibration was also done in [Bonet and Burton, 1998] considering the poisson ratios $\nu_{12} = \nu_{13} = \nu_{23}$, but here we consider that $\nu_{12} = \nu_{13} \neq \nu_{23}$.

1.2.2 Obtaining plane wave velocities from the elasticity tensor

In this section, with the assumption of harmonic plane waves and the Kelvin-Christoffel equations [Musgrave, 1970; Brown, 1989] we obtain the local quasi-longitudinal and quasi-shear velocities for plane waves from the elasticity tensor components [Carcione, 2001]. First, we write (1.19) with no source term in the traditional indicial notation

$$\widehat{\rho}\partial_{tt}u_i - \partial_j\left(C_{ijkl}\varepsilon_{kl}\right) = 0.$$

Using Property A.1.1, yields

$$\widehat{\rho}\partial_{tt}u_i - \partial_j\left(C_{ijkl}\partial_l u_k\right) = 0,$$

and, developing the divergence and considering $\|\nabla_{\hat{x}}\boldsymbol{C}\| \ll 1$ in a suitable norm,

$$\widehat{\rho}\partial_{tt}u_i - C_{ijkl}\partial_j\partial_l u_k = 0.$$

Now, considering a harmonic plane wave with frequency ω_f as a solution,

$$u_i(\hat{\boldsymbol{x}}) = u_i^a e^{\omega_f(s_m\hat{x}_m - t)},$$

where $\boldsymbol{s} = (s_1, s_2, s_3) = v^{-1}\boldsymbol{n}$ is the slowness vector in the direction $\boldsymbol{n}$ and v is the wavespeed, we have

$$\left(C_{ijkl}(s_j s_l) - \widehat{\rho}v^2\delta_{ik}\right)u_k = 0.$$

The eigenproblem can be solved to find $\widehat{\rho}v^2$ and u_k. By solving this eigenproblem, we retrieve up to 3 modes of propagation for bulk waves and their respective velocities, depending on the material symmetries.

1.3 Galerkin approximation and explicit time scheme

In this section, we present the class of (Galerkin) approximations that can be used to solve the linear finite-dimensional problem in the form (1.20). It consists of using a subset of the solution space $\mathcal{V}(\hat{\Omega})$ to solve the weak formulation. It allows us to use a discrete solution space in order to represent the problem as matrix and vector products. Defining the set $\{\boldsymbol{\psi}_{h,I}\}_{I=1}^{N_h}$ of global functions that span the approximation discrete space, dicretized with step h,

$$\mathscr{V}_h^0(\hat{\Omega}) = \mathbf{span}\{\boldsymbol{\psi}_{h,I}\}_{I=1}^{N_h},$$

we consider solving the problem for this subspace with the same bilinear forms and $u_h \in \mathscr{V}_h^0(\hat{\Omega})$.

Leapfrog time scheme. Using a second-order centered time scheme, we approximate the time derivative as

$$\frac{\mathrm{d}^2 u_h}{\mathrm{d}t^2}(n\Delta t) \approx \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, \quad \forall n \in [\![1; N]\!],$$

and obtain the time-discrete problem form

$$\left(\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2}, w_h\right)_{\mathrm{L}^2(\hat{\Omega})} + \langle Au_h^n, w_h\rangle = l(w_h), \quad \forall w_h \in \mathscr{V}_h^0(\hat{\Omega}), \tag{1.23}$$

where Δt is the time step and $u_h^n \approx u_h(n\Delta t)$. With the evaluation of the stiffness operator a at u_h^n, we can retrieve an explicit scheme without the need for inverting the resulting A operator.

Fully discrete form. Using Property A.2.2 and Property A.2.4, we can define $\overrightarrow{u_h}$ and $\overrightarrow{w_h}$ representing the elements and operators in their vectorial and matrix form, yielding

$$(u_h, w_h)_{\mathrm{L}^2(\hat{\Omega})} = \overrightarrow{u_h}^\intercal \mathbb{M}\,\overrightarrow{w_h}, \quad \langle Au_h, w_h\rangle = \overrightarrow{u_h}^\intercal \mathbb{K}\,\overrightarrow{w_h}, \quad l(w_h) = \overrightarrow{f_h}^\intercal \mathbb{M}\,\overrightarrow{w_h},$$

leading to the fully-discrete form

$$\mathbb{M}\left(\frac{\overrightarrow{u_h}^{n+1} - 2\overrightarrow{u_h}^n + \overrightarrow{u_h}^{n-1}}{\Delta t^2}\right) + \mathbb{K}\,\overrightarrow{u_h} = \mathbb{M}\,\overrightarrow{f_h}. \tag{1.24}$$

In order to obtain $\overrightarrow{u_h}$ for a given time step, we can write the time marching operation as

$$\overrightarrow{u_h}^{n+1} = \Delta t^2 \overrightarrow{f_h}^{n} - \mathbb{M}^{-1}\mathbb{K}\,\overrightarrow{u_h}^{n} + 2\overrightarrow{u_h}^{n} - \overrightarrow{u_h}^{n-1}. \tag{1.25}$$

At each time step, the inversion of the mass and the application of the stiffness matrix must be done. As we present next, the use of the spectral finite elements method with mass lumping (Section 1.4) makes this scheme **explicit** and most of the computational cost becomes present in the matrix-vector multiplication $\mathbb{K}\,\overrightarrow{u_h}^{n}$.

Energy conservation and stability. Here, we recall the energy technique for analyzing the stability of the time marching scheme [Cohen, 2002; Joly, 2003]. First, we assume that the source term is zero, *i.e.* no energy input, and a test function in the form of

$$\overrightarrow{w_h} = \frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n+1}}{2\Delta t}$$

being an approximation of $(\mathrm{d}_t\overrightarrow{u_h})(n\Delta t)$. By defining the total energy at $n+\frac{1}{2}$ as

$$\mathscr{E}^{n+\frac{1}{2}} = \frac{1}{2}\left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\}^{\intercal} \mathbb{M} \left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\} + \frac{1}{2}\overrightarrow{u_h}^{n+1}\mathbb{K}\,\overrightarrow{u_h}^{n}, \tag{1.26}$$

one can show that, with the relation (1.25),

$$\frac{\mathscr{E}^{n+\frac{1}{2}} - \mathscr{E}^{n-\frac{1}{2}}}{\Delta t} = 0.$$

This ensures energy conservation. To ensure the stability of the scheme, the energy $\mathscr{E}$ must be positive at all n steps. As $\mathbb{M}$ is positive-definite, the first term in (1.26) is always positive. The second term is further manipulated to obtain

$$\overrightarrow{u_h}^{n+1}\mathbb{K}\,\overrightarrow{u_h}^{n} = \frac{1}{2}\left\{\frac{\overrightarrow{u_h}^{n+1} + \overrightarrow{u_h}^{n}}{2}\right\}^{\intercal} \mathbb{K} \left\{\frac{\overrightarrow{u_h}^{n+1} + \overrightarrow{u_h}^{n}}{2}\right\} - \frac{\Delta t^2}{4}\left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\}^{\intercal} \mathbb{K} \left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\},$$

leading to

$$\mathscr{E}^{n+\frac{1}{2}} = \frac{1}{2}\left\{\frac{\overrightarrow{u_h}^{n+1} + \overrightarrow{u_h}^{n}}{2}\right\}^{\intercal} \mathbb{K} \left\{\frac{\overrightarrow{u_h}^{n+1} + \overrightarrow{u_h}^{n}}{2}\right\} + \frac{\Delta t^2}{4}\left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\}^{\intercal} \left(\mathbb{M} - \frac{\Delta t^2}{4}\mathbb{K}\right) \left\{\frac{\overrightarrow{u_h}^{n+1} - \overrightarrow{u_h}^{n}}{\Delta t}\right\}.$$

The positivity of the first term is ensured by the positivity of $\mathbb{K}$, which is the case for the linear elasticity problem (1.20), but it is not always the case for a linearized problem, as we show in the Chapter 2. The positivity of the second term is ensured by the positivity of $\left(\mathbb{M} - \frac{\Delta t^2}{4}\mathbb{K}\right)$, leading to the Courant-Friedrichs-Lewy condition on the time step

$$\Delta t \leq \frac{2}{\sqrt{r(\mathbb{M}^{-1}\mathbb{K})}}, \tag{1.27}$$

where $r(\cdot)$ represents the spectral radius. This is valid for any Galerkin approximation, including the one presented in the next section.

1.4 Space approximation using the spectral elements method

Here, we precise the methods used for the discretization and the construction of the approximation space $\mathscr{V}_h^0(\hat{\Omega})$. More precisely, we use the Spectral Finite Elements method [Maday and Patera, 1989; Komatitsch and Vilotte, 1998; Joly, 2007]. This method allows efficient use of high-order finite elements, reducing considerably numerical artifacts in wave propagation [Basabe, 2007; Seriani and Oliveira, 2008]. Avoiding numerical artifacts is particularly important for our work as we deal with ultrasonic guided waves that can propagate during long periods and through long distances. Also, the mass lumping and unassembled stiffness matrix application presented in this section reduces the computational cost and the required memory, increasing the number of solvers that can be run in parallel in a given machine. In an inverse problem context such as the one to be presented in Chapter 4, the capacity to launch and run several solvers in parallel is essential. For the sake of simplicity, we will present the spectral elements method for an acoustic case where u is a scalar field. Although, in this work, we deal with the elastodynamics problem (1.20), the techniques presented in this section can be applied to a vectorial field case, being more laborious algebraically but not adding essential difficulty. With the functional space $\mathrm{H}^1(\hat{\Omega})$ defined as in Definition A.2.1, the scalar acoustic problem reads:

> **Linear acoustic weak formulation** $\forall t > 0$, find $u(t) \in \mathrm{H}^1(\hat{\Omega})$ such that
>
> $$\forall w \in \mathrm{H}^1(\hat{\Omega}), \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}(u, w)_{\mathrm{L}^2(\hat{\Omega})} + \langle Au, w\rangle = l(w) \tag{1.28}$$
>
> with
>
> $$\begin{aligned} (u, w)_{\mathrm{L}^2(\hat{\Omega})} &= \int_{\hat{\Omega}} \widehat{\rho}\, u\, w \,\mathrm{d}\hat{\Omega}, \\ \langle Au, w\rangle &= \int_{\hat{\Omega}} c(\hat{x})(\nabla_{\hat{x}} u)^{\intercal} \nabla_{\hat{x}} w \,\mathrm{d}\hat{\Omega}, \\ l(w) &= \int_{\hat{\Omega}} f\, w \,\mathrm{d}\hat{\Omega}. \end{aligned}$$

First, we define a conform tesselation of $\hat{\Omega}$ with elements $K \in \mathcal{T}_h(\hat{\Omega})$ where

$$\hat{\Omega} = \bigcup_{K \in \mathcal{T}_h(\hat{\Omega})} K \quad \text{and} \quad \mathring{K} \cap \mathring{K}' = \emptyset, \ \forall K, K' \in \mathcal{T}_h(\hat{\Omega}) \text{ and } K \neq K',$$

with h being the average element size. Using this definition of the mesh $\mathcal{T}_h(\hat{\Omega})$, we define the discretized solution space as

$$V_h(\hat{\Omega}) = \{v_h \in C^0(\overline{\hat{\Omega}}) \mid \forall K \in \mathcal{T}_h(\hat{\Omega}),\ \exists! \widehat{v} \in Q^k(\widehat{K}),\ v_h|_K = \widehat{v} \circ \Upsilon_K^{-1}\} \subset \mathrm{H}^1(\hat{\Omega}),$$

and

$$V_h^0(\hat{\Omega}) = \{v_h \in V_h(\hat{\Omega}), v_h = 0 \text{ on } \hat{\Gamma}_D\}$$

being a discrete Galerkin approximation, *i.e.* internal approximation of $\mathrm{H}^1(\hat{\Omega})$. This approximates the space with global continuous functions where, for every element K, it has a correspondent local function defined in $\widehat{K}$, the reference element. The functional space of local functions $Q^k(\widehat{K})$ with k-th order polynomials, to be soon precised, is the functional space that represents locally the global functions $v_h \in C^0(\overline{\hat{\Omega}})$. The mapping Υ_K is defined as the bijective map from $\widehat{K}$ to K. We define the set of global functions $\{\psi_{h,I}\}_{I=1}^{N_h}$ that spans the approximation space

$$V_h^0(\hat{\Omega}) = \mathbf{span}\{\psi_{h,I}\}_{I=1}^{N_h}, \qquad \psi_{h,I}(\boldsymbol{\xi}_{h,J}) = \delta_{IJ} \qquad \forall I, J = 0, 1, \ldots, N_h,$$

where $\{\boldsymbol{\xi}_{h,I}\}_{I=1}^{N_h}$ are called degrees of freedom, associated with the functions $\{\psi_{h,I}\}_{I=1}^{N_h}$. Now, we proceed by detailing the use of a local functional basis and the reference element to do operations in the global approximation space $V_h(\hat{\Omega})$, therefore enabling us to solve the weak formulation.

First, in Section 1.4.1, we present how operations can be done with local functions. Using these local operations, we present in Section 1.4.2 how a diagonal mass is obtained through mass lumping and in Section 1.4.3 how parallel and unassembled applications of the stiffness matrix can be done.

1.4.1 Local operations in the reference element

In the definition of the functional space V_h the reference element $\widehat{K}$ is arbitrary. We consider here the reference element as $\widehat{K} = [0,1]^3$. Local coordinates are denoted as $\boldsymbol{\zeta} \in [0,1]^3$. By definition, using the mapping $\boldsymbol{\Upsilon}_K$, every function in $V_h(\hat{\Omega})$ is represented by a local function in $Q^k(\widehat{K})$. Let $Q^k(\widehat{K})$ be the space of polynomials characterized by a set $\{\widehat{\boldsymbol{\xi}}_j\}_{j=1}^{\hat{N}}$ of coordinates representing the local degrees of freedom such that

$$Q^k(\widehat{K}) = \left[\mathcal{P}^k([0,1])\right]^3 = \mathbf{span}\{\widehat{\psi}_i\}_{i=1}^{\hat{N}}, \qquad \widehat{\psi}_i(\widehat{\boldsymbol{\xi}}_j) = \delta_{ij} \qquad \forall i,j \in [\![0;\hat{N}]\!].$$

The one-dimensional polynomials in $\mathcal{P}^k([0,1])$ are characterized by "edge" degrees of freedom $\{\widehat{\xi}_j^{\,\mathrm{1D}}\}_{j=1}^{N_e}$, such that

$$\mathcal{P}^k([0,1]) = \mathbf{span}\{\hat{v}_i\}_{i=1}^{N_e}, \quad \hat{v}_i(\widehat{\xi}_j^{\,\mathrm{1D}}) = \delta_{ij} \quad \text{for} \quad i,j = 1,\dots,N_e.$$

The size of these spaces is related as $\hat{N} = (N_e)^3 = (k+1)^3$. To enable the use of high-order Lagrange polynomials, the one-dimensional degrees of freedom $\{\widehat{\xi}_j^{\,\mathrm{1D}}\}_{j=1}^{N_e}$ are distributed as the Gauss-Lobatto points, avoiding the Runge's phenomenon. We precise then the three-dimensional degrees of freedom coordinates constructed from the one-dimensional ones,

$$\{\widehat{\boldsymbol{\xi}}_i\}_{i=1}^{\hat{N}} = \left\{(\widehat{\xi}_\alpha^{\,\mathrm{1D}}, \widehat{\xi}_\beta^{\,\mathrm{1D}}, \widehat{\xi}_\gamma^{\,\mathrm{1D}}), \quad \forall(\alpha,\beta,\gamma) \in [\![1;N_e]\!]^3\right\}_{i=1}^{N_e^3}.$$

The three-dimensional basis functions are also constructed by using the one-dimensional ones. For $(\hat{v}^{(1)}, \hat{v}^{(2)}, \hat{v}^{(3)}) \in Q^k(\widehat{K})^3$, we have

$$\{\widehat{\psi}_i(\boldsymbol{\zeta})\}_{i=1}^{\hat{N}} = \left\{\hat{v}_\alpha^{(1)}(\zeta_1)\hat{v}_\beta^{(2)}(\zeta_2)\hat{v}_\gamma^{(3)}(\zeta_3), \quad \forall(\alpha,\beta,\gamma) \in [\![1;N_e]\!]^3\right\}_{i=1}^{N_e^3}.$$

In both cases, the relation between the left and right-hand side sets can be defined by an arbitrary one-to-one relation $i \leftrightarrow (\alpha,\beta,\gamma)$, for instance $i = 9(\alpha-1) + 3(\beta-1) + \gamma$. One can note that an anisotropic space Q can be constructed, meaning it is formed by one-dimensional polynomial spaces $\mathcal{P}$ with different orders and degrees of freedom.

Quadrature formula. To integrate functions in $\widehat{K}$ numerically, we must apply a quadrature formula. Attention must be given to ensure that approximation errors do not affect the stability and consistency of the numerical model. Let g be a function to be integrated, a quadrature formula reads

$$\int_{\widehat{K}} g(\boldsymbol{\zeta})\mathrm{d}\widehat{K} \approx \sum_{i=1}^{\hat{N}_q} w_i g(\boldsymbol{\zeta}_i^q), \tag{1.29}$$

with $\hat{N}_q$ being the number of quadrature points, $\{w_i\}_{i=1}^{\hat{N}_q}$ the quadrature weights and $\{\boldsymbol{\zeta}_i^q\}_{i=1}^{\hat{N}_q}$ the quadrature points. For the presented implementation of the spectral finite elements method, the Gauss-Lobatto quadrature formula with $\{\boldsymbol{\zeta}_i^q\}_{i=1}^{\hat{N}_q} = \{\widehat{\boldsymbol{\xi}}_i\}_{i=1}^{\hat{N}}$ is used, leading to the mass lumping and efficient stiffness matrix application, as presented in the following sections. The Gauss-Lobatto quadrature approximation satisfies the necessary properties without losses in convergence [Duruflé, Grob, and Joly, 2009].

Local vector representations. By definition, every global function $v|_K \in C^0(\overline{\hat{\Omega}})$ has a local corresponding function $\widehat{v} \in Q^k(\widehat{K})$. Using property Property A.2.2, a function $\widehat{v} \in Q^k(\widehat{K})$ can be represented in its vectorial form $\overrightarrow{\widehat{v}} = \begin{pmatrix}\widehat{v}_1 & \dots & \widehat{v}_{\hat{N}}\end{pmatrix}^\intercal$ as the local finite elements vector.

1.4.2 Lumping of the mass matrix

To obtain the components of the matrix $\mathbb{M}$ we develop the mass and right-hand side bilinear forms as sums over the tesselation $\mathcal{T}_h$,

$$m(u_h, w_h) = \int_{\hat{\Omega}} \widehat{\rho}\, u_h\, w_h \,\mathrm{d}\hat{\Omega} = \sum_{K\in\mathcal{T}_h} \int_K \widehat{\rho}\, u_h\, w_h \,\mathrm{d}K,$$

$$l(w_h) = \int_{\hat{\Omega}} \widehat{\rho}\, u_h\, w_h \,\mathrm{d}\hat{\Omega} = \sum_{K\in\mathcal{T}_h} \int_K \widehat{\rho}\, u_h\, w_h \,\mathrm{d}K.$$

As we have one bijective map $\boldsymbol{\Upsilon}_K$ from $\widehat{K}$ to every K, we use the local representations $\widehat{u} = u_h \circ \boldsymbol{\Upsilon}_K$, $\widehat{w} = w_h \circ \boldsymbol{\Upsilon}_K$, $\widehat{f} = f_h \circ \boldsymbol{\Upsilon}_K$, $J_K = \det\left(\nabla_{\hat{\boldsymbol{x}}}\boldsymbol{\Upsilon}_K\right)$ to write

$$m(u_h, w_h) = \sum_{K\in\mathcal{T}_h} \int_{\widehat{K}} \widehat{\rho}\,\widehat{u}\,\widehat{w} J_K \,\mathrm{d}\widehat{K} = \sum_{K\in\mathcal{T}_h} \overrightarrow{\widehat{w}}^{\mathsf{T}} \widehat{\mathbb{M}}_K \overrightarrow{\widehat{u}}$$

$$l(w_h) = \sum_{K\in\mathcal{T}_h} \int_{\widehat{K}} \widehat{\rho}\,\widehat{f}\,\widehat{w} J_K \,\mathrm{d}\widehat{K} = \sum_{K\in\mathcal{T}_h} \widehat{\rho}\,\overrightarrow{\widehat{w}}^{\mathsf{T}} \widehat{\mathbb{M}}_K \overrightarrow{\widehat{f}},$$

where $\widehat{\mathbb{M}}_K \in \mathcal{M}(\mathbb{R})_{\hat{N}\times\hat{N}}$ refers to the local mass matrix of the element $K \in \mathcal{T}_h$.

> **Property 1.4.1.** *For $\widehat{u}, \widehat{v} \in Q^k(\widehat{K})$ and a function g, the integral on the reference element can be approximated using* (1.29) *as*
>
> $$\int_{\widehat{K}} g(\boldsymbol{x})\,\widehat{u}\,\widehat{v}\,\mathrm{d}\widehat{K} \approx \overrightarrow{\widehat{u}}^{\mathsf{T}} \widehat{\mathbb{M}}_g \overrightarrow{\widehat{v}} \quad \textit{with} \quad \left(\widehat{\mathbb{M}}_g\right)_{ij} = w_i g(\widehat{\boldsymbol{\xi}}_i)\delta_{ij},$$
>
> *$\overrightarrow{\widehat{u}}$, $\overrightarrow{\widehat{v}}$ being the local finite elements vectors representing $\widehat{u}, \widehat{v} \in Q^k(\widehat{K})$.*

Proof. Using a Gauss-Lobatto quadrature formula with the weights $\{w_i\}_{i=1}^{\hat{N}_q}$ and quadrature points that coincide degrees of freedom $\{\widehat{\boldsymbol{\xi}}_i\}_{i=1}^{\hat{N}}$ of the basis functions, we have

$$\int_{\widehat{K}} g(\boldsymbol{x})\widehat{u}\,\widehat{v}\,\mathrm{d}\widehat{K} \approx \sum_{m=1}^{\hat{N}_q} w_m g(\widehat{\boldsymbol{\xi}}_m)\left(\sum_{j=1}^{\hat{N}} \widehat{u}_j \widehat{\psi}_j(\widehat{\boldsymbol{\xi}}_m)\right)\left(\sum_{k=1}^{\hat{N}} \widehat{v}_k \widehat{\psi}_k(\widehat{\boldsymbol{\xi}}_q)\right).$$

As $\widehat{\psi}_i(\widehat{\boldsymbol{\xi}}_j) = \delta_{ij}$ we obtain

$$\int_{\widehat{K}} g(\boldsymbol{x})\,\widehat{u}\,\widehat{v}\,\mathrm{d}\widehat{K} \approx \sum_{m=1}^{\hat{N}} w_m g(\widehat{\boldsymbol{\xi}}_m)\widehat{u}_m \widehat{v}_m = \overrightarrow{\widehat{u}}^{\mathsf{T}} \widehat{\mathbb{M}}_g \overrightarrow{\widehat{v}}.$$

□

Using Property 1.4.1, the local mass matrix reads

$$\left(\widehat{\mathbb{M}}_K\right)_{ij,\hat{N}\times\hat{N}} = w_i J_K(\widehat{\boldsymbol{\xi}}_i)\rho(\widehat{\boldsymbol{\xi}}_i)\delta_{ij}.$$

From the operations above, one can deduce the assembly of the global mass matrix

$$\mathbb{M} = \mathop{\mathsf{A}}_{K\in\mathcal{T}_h(\hat{\Omega})} \widehat{\mathbb{M}}_K,$$

that will consequently be diagonal as well. The big "A" operator refers to the procedure of assembling the matrix, often used in the finite elements literature. The mass matrix inversion needed in (1.25) is obtained by inexpensively inverting each diagonal component.

1.4.3 Unassembled application of the stiffness matrix

In order to obtain the operations that lead to the elements in the matrix $\mathbb{K}$ we will represent the integration also as the operations by element. For $u_h, w_h \in V_h(\hat{\Omega})$,

$$a(u_h, w_h) = \int_{\hat{\Omega}} c(\hat{x})(\nabla_{\hat{x}} u_h)^\intercal \nabla_{\hat{x}} w_h \, \mathrm{d}\hat{\Omega} = \sum_{K \in \mathcal{T}_h(\hat{\Omega})} \int_K c(\hat{x})(\nabla_{\hat{x}} u_h)^\intercal \nabla_{\hat{x}} w_h \, \mathrm{d}K.$$

Using the bijective map Υ_K, we can write the integral in K, by changing variables, as the local form

$$\begin{aligned}\int_K c(\hat{x})(\nabla_{\hat{x}} u_h)^\intercal \nabla_{\hat{x}} w_h \, \mathrm{d}K &= \int_{\widehat{K}} \hat{c}(\zeta)\Big(\left(\nabla_\zeta \Upsilon_K\right)^{-\intercal} \nabla_\zeta \widehat{u}\Big)^\intercal \Big(\left(\nabla_\zeta \Upsilon_K\right)^{-\intercal} \nabla_\zeta \widehat{w}\Big) J_K \, \mathrm{d}\widehat{K} \\ &= \int_{\widehat{K}} \hat{c}(\zeta)(\nabla_\zeta \widehat{u})^\intercal \left(\nabla_\zeta \Upsilon_K\right)^{-1} \left(\nabla_\zeta \Upsilon_K\right)^{-\intercal} \nabla_\zeta \widehat{w} J_K \, \mathrm{d}\widehat{K}\end{aligned}$$

where

$$\widehat{\boldsymbol{u}} = \boldsymbol{u}_h \circ \Upsilon_K, \quad \widehat{\boldsymbol{w}} = \boldsymbol{w}_h \circ \Upsilon_K, \quad \hat{c} = c \circ \Upsilon_K.$$

The dependency on local coordinates was omitted. Applying a Gauss-Lobatto quadrature formula (1.29), we can approximate this integration as

$$\int_{\widehat{K}} \cdots \, \mathrm{d}\widehat{K} \approx \sum_{m=1}^{\hat{N}} w_m J_K(\widehat{\boldsymbol{\xi}}_m) \hat{c}(\widehat{\boldsymbol{\xi}}_m) \left(\nabla_\zeta \widehat{u}(\widehat{\boldsymbol{\xi}}_m)\right)^\intercal \left(\nabla_\zeta \Upsilon_K(\widehat{\boldsymbol{\xi}}_m)\right)^{-1} \left(\nabla_\zeta \Upsilon_K(\widehat{\boldsymbol{\xi}}_m)\right)^{-\intercal} \nabla_\zeta \widehat{w}(\widehat{\boldsymbol{\xi}}_m).$$

To evaluate this summation we need the gradients ∇_ζ evaluated in the points of quadrature. Finally, we can write

$$a(u_h, w_h) \approx \sum_{K \in \mathcal{T}_h(\hat{\Omega})} \sum_{m=1}^{N_q} w_r J_K \hat{c} \left(\nabla_\zeta \widehat{u}\right)^\intercal \left(\nabla_\zeta \Upsilon_K\right)^{-1} \left(\nabla_\zeta \Upsilon_K\right)^{-\intercal} \nabla_\zeta \widehat{w}$$

with dependency evaluated at $\widehat{\boldsymbol{\xi}}_m$ omitted.

> **Property 1.4.2.** *Using the Gauss-Lobatto quadrature formula presented in this section, the derivative w.r.t. the spatial component j of a local function $\widehat{v} \in Q^k(\widehat{K})$ evaluated at all local degrees of freedom can be expressed as*
>
> $$\overrightarrow{\partial_j \widehat{v}} = \begin{pmatrix} \partial_j \widehat{v}(\widehat{\boldsymbol{\xi}}_1) \\ \vdots \\ \partial_j \widehat{v}(\widehat{\boldsymbol{\xi}}_{\hat{N}}) \end{pmatrix} = \Gamma^{(j)} \overrightarrow{\widehat{v}} \quad \textit{where} \quad \Gamma^{(j)} = \begin{pmatrix} \gamma^{(j)}_{1,1} & \gamma^{(j)}_{1,2} & \cdots & \gamma^{(j)}_{1,\hat{N}} \\ \vdots & & \ddots & \vdots \\ \gamma^{(j)}_{\hat{N},1} & \gamma^{(j)}_{\hat{N},2} & \cdots & \gamma^{(j)}_{\hat{N},\hat{N}} \end{pmatrix}$$
>
> *and $\gamma^{(j)}_{m,n} = \partial_j \widehat{\psi}_n(\widehat{\boldsymbol{\xi}}_m)$ with $\widehat{\psi}_n$ being the n-th local basis function of the space $Q^k(\widehat{K})$ and $\widehat{\boldsymbol{\xi}}_m$ being the m-th local degree of freedom.*

Proof. By definition, $\widehat{v}$ and its derivative *w.r.t.* to the j-th component can be written as

$$\widehat{v}(\zeta) = \sum_{n=1}^{\hat{N}} \widehat{v}_n \widehat{\psi}_n(\zeta) \qquad \text{and} \qquad \partial_j \widehat{v}(\zeta) = \sum_{n=1}^{\hat{N}} \widehat{v}_n \partial_j \widehat{\psi}_n(\zeta),$$

respectively. As $\{\widehat{\psi}_n\}_{n=1}^{\hat{N}}$ are polynomials, its derivatives can also be represented in the same discrete space

$$\partial_j \widehat{\psi}_n(\boldsymbol{\zeta}) = \sum_{m=1}^{\hat{N}} \gamma_{m,n}^{(j)} \widehat{\psi}_m(\boldsymbol{\zeta}), \qquad \text{with} \qquad \gamma_{m,n}^{(j)} = \partial_j \widehat{\psi}_n(\widehat{\boldsymbol{\xi}}_m).$$

Using these representations we have

$$\partial_j \widehat{v}(\boldsymbol{\zeta}) = \sum_{n=1}^{\hat{N}} \widehat{v}_n \sum_{m=1}^{\hat{N}} \gamma_{m,n}^{(j)} \widehat{\psi}_m(\boldsymbol{\zeta}),$$

that, evaluated at a degree of freedom $\widehat{\boldsymbol{\xi}}_i$, yields

$$\partial_j \widehat{v}(\widehat{\boldsymbol{\xi}}_i) = \sum_{n=1}^{\hat{N}} \widehat{v}_n \gamma_{i,n}^{(j)} = \begin{pmatrix} \gamma_{i,1}^{(j)} & \gamma_{i,2}^{(j)} & \cdots & \gamma_{i,\hat{N}}^{(j)} \end{pmatrix} \begin{pmatrix} \widehat{v}_1 \\ \widehat{v}_2 \\ \vdots \\ \widehat{v}_{\hat{N}} \end{pmatrix}.$$

This relation leads to the definition of the matrix $\Gamma^{(j)}$. □

Using Property 1.4.2 to assemble a gradient vector and rearranging the above operations, we can construct an element-wise operation represented by the local stiffness matrix $\{\widehat{\mathbb{K}}_K\}_{K\in\mathcal{T}_h}$, applied to the local finite-element vectors, such as

$$a(u_h, w_h) \approx \sum_{K\in\mathcal{T}_h(\hat{\Omega})} \vec{u}\,\widehat{\mathbb{K}}_K \vec{w}, \quad \therefore \quad \mathbb{K} = \mathop{\mathsf{A}}_{K\in\mathcal{T}_h(\hat{\Omega})} \widehat{\mathbb{K}}_K.$$

Unassembled and parallel computations. When solving (1.25) using the presented spectral elements methods, the inversion of $\mathbb{M}$ becomes computationally cheap as it is diagonal. The most CPU-consuming part becomes the matrix-vector product $\mathbb{K}\,\vec{u_h}^n$ and we present here how it can be done more efficiently.

Property 1.4.3. *In the three dimensional case, the matrix* $\Gamma^{(j)}$ *has a sparsity ratio of* $\frac{k}{k+1}$.

Proof. The basis functions $\{\widehat{\psi}_n\}_{n=1}^{\hat{N}} \in \mathcal{Q}^k(\widehat{K}) = \mathcal{P}^k(\widehat{K})^3$, evaluated at the degrees of freedom, are decomposed by definition as

$$\widehat{\psi}_n(\widehat{\boldsymbol{\xi}}_m) = \prod_{i=1}^{3} \hat{\upsilon}_n^{(i)}(\widehat{\xi}_{m,i}),$$

where $\{\hat{\upsilon}_n^{(i)}\}_{n=1}^{\hat{N}} \in \mathcal{P}^k$ depends only on the i-th coordinate component. The components of $\Gamma^{(j)}$ are

$$\gamma_{m,n}^{(j)} = \partial_j \widehat{\psi}_n(\widehat{\boldsymbol{\xi}}_m) = \partial_j \prod_{i=1}^{d} \hat{\upsilon}_n^{(i)}(\widehat{\xi}_{m,i}),$$

whick leads to the sparsity as by construction, $\widehat{\xi}_{m,i}$ is a degree of freedom of $\hat{\upsilon}_n^{(i)}$. Then we have, for instance, for $j = 1$ in the three-dimensional case

$$\gamma_{m,n}^{(1)} = \hat{\upsilon}_n^{(2)}(\widehat{\xi}_{m,2})\hat{\upsilon}_n^{(3)}(\widehat{\xi}_{m,3}),$$

being non-zero when $\widehat{\boldsymbol{\xi}}_m$ is in the plane orthogonal to the $j = 1$ axis at $\widehat{\xi}_{m,1}$, resulting in a sparsity ratio of $\frac{k}{k+1}$ when Q^k is isotropic. □

In Property 1.4.3, we demonstrate that the gradient operations are sparse, leading to fewer operations when applying $\mathbb{K}$ in comparison to stiffness matrices built with traditional quadrature formulas. As the construction of the stiffness matrix $\mathbb{K}$ is done by summing element-wise contributions and the only need we have for it is its application on a vector, this matrix-vector operation can be done without assembling the stiffness matrix. Additionally, this sum can be done in parallel by treating potential concurrent writing. If $\mathcal{T}_h$ is constructed in a structured manner, concurrent writing can be effectively avoided by grouping the elements $\mathcal{T}_h$ that do not share common degrees of freedom. We denote this aggroupment as $\mathcal{G}$. The matrix-vector product $\mathbb{K}\,\overrightarrow{u_h}^n$ is then described in Algorithm 1. More details can be found in [Alexandre Imperiale and Demaldent, 2019].

Algorithm 1: Unassembled matrix-vector product of the stiffness operator.

Input: Groups E of elements in $\mathcal{G}$, the global finite elements vector $\overrightarrow{u_h}$ and the local-to-global mapping ℓ.

Output: The result of the matrix-vector product $\overrightarrow{V} = \mathbb{K}\,\overrightarrow{u}$.

```
initialize empty global vector $\overrightarrow{V}$
for $G \in \mathcal{G}$ do
    parallel for $e \in \mathrm{E}(G)$ do
        obtain $\overrightarrow{\widetilde{u}^e}$ from $\overrightarrow{u_h}$
        for $i = 1, \ldots, \hat{N}$ do
            $\overrightarrow{V}[\ell(i,e)] \leftarrow \overrightarrow{V}[\ell(i,e)] + \mathbb{K}_e \overrightarrow{\widetilde{u}^e}$
        end
    end
end
```

1.5 Performance illustration on realistic dimensions

The performance of this implementation of the high-order spectral finite elements is illustrated here by modeling two cases. In the first one, ultrasonic waves propagating in an aircraft wing leading edge are modeled using low-order finite elements and high-order finite elements, highlighting the efficiency of using high-order elements for long-range wave propagation. This first illustration was done in the context of the “GW4SHM” project, during the exchange at the Imperial College London. In the second case, an analog of a one-shot aircraft is modeled to assess performance in a computationally demanding configuration. For both configurations, the aluminum material properties are used, namely $\lambda = 54.3\text{GPa}$ and $\mu = 27.17\text{GPa}$.

1.5.1 High-order finite elements for guided wave propagation

We solve a wave propagation problem in an aircraft wing leading edge using two different solvers with the objective of highlighting the particularities of the spectral finite elements method. The leading edge has dimensions of $\sim 160 \times 2300 \times 100\text{mm}$. The model for the wave propagation will be the linear elastodynamics (1.20). The geometry, excitation point and measurement points are depicted in Figure 1.2. Ultrasonic waves are excited with a body force with spatial dependency, for all vectorial components, as

$$e^{-\left(\frac{x-x_0}{\sigma}\right)^2}$$

centered (by setting $\boldsymbol{x}_0$) at the thickness mid-point with $\sigma = 15\text{mm}$. The time-dependent part of the source term is a 200kHz Hanning windowed 5-cycle cosine wave. In the following, we model this setup using two different solvers, one based on low-order finite elements and ours. Our objective is to highlight the importance of using high-order elements for wave propagation over long distances.

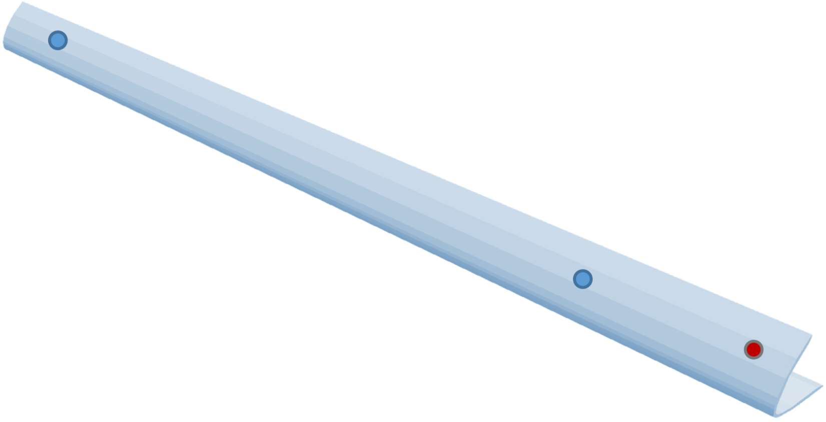
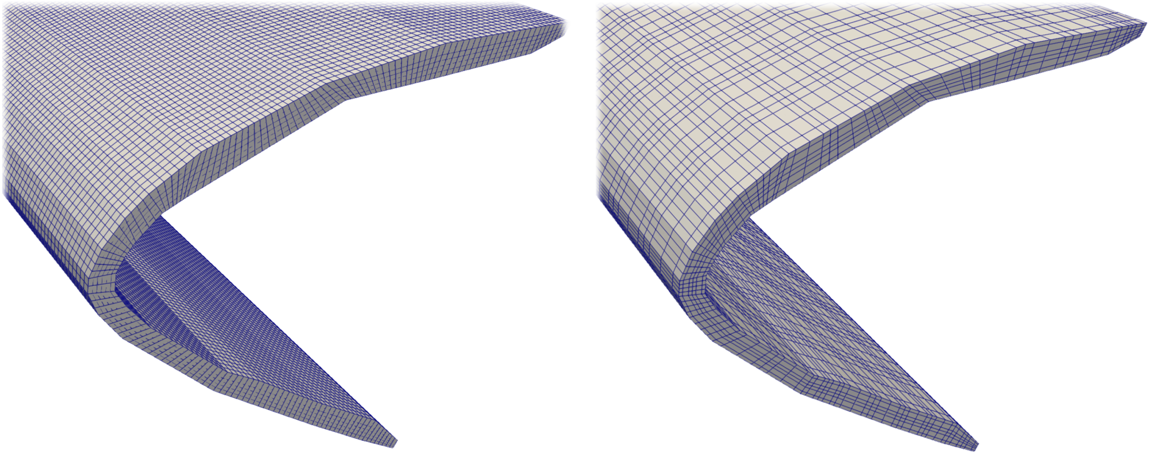


Figure 1.2 – Illustration domain $\hat{\Omega}$, source (red) and measurement points (blue) position.



(a) Order 1 $\mathcal{T}_h(\hat{\Omega})$ discretization.

(b) Order 4 Gauss-Lobatto $\mathcal{T}_h(\hat{\Omega})$ discretization.

Figure 1.3 – Illustration of different types of discretizations.

Using Q^1 finite elements. For the low-order case, the space is discretized with Q^1 finite elements. An example of Q^1 mesh discretization is shown in Figure 1.3a. For this illustration, we use the solver implemented in [Huthwaite, 2014], a highly parallelized approach using GPUs for memory and computation. A thumb rule for good convergence of the solution for this type of element is to set the mesh step h resulting in a minimum of 20 finite elements per wavelength. The maximum time step could be computed using the CFL condition (1.27) but as the computation of $\mathbb{M}^{-1}\mathbb{K}$ is not directly available with the solver, an approximation for low-order finite elements and uniform mesh stepping can be done as

$$\frac{2}{\sqrt{r(\mathbb{M}^{-1}\mathbb{K})}} \approx \frac{h}{c_0}.$$

Therefore, we choose

$$\Delta t = 0.2\frac{h}{c_0},$$

where c_0 is the pressure wave velocity and 0.2 is used as a safety factor to satisfy the CFL condition due to the quality issues of our generated mesh. To test convergence, two different discretizations were set and run: with 20 and 24 nodes per wavelength resulting in 31 million and 52 million degrees of freedom, respectively. We ran the solver up to 600μs. Using 3 GPU GeForce RTX 2080 Ti, the computation took 8 minutes and required 12GB of VRAM for the 20 nodes per wavelength case. For 24 nodes per wavelength, it took 17 minutes and 24GB of VRAM. The Y component is extracted at the closest measurement point, 420mm distant from the source. The extracted signal is plotted in Figure 1.4 for the two discretizations. The results show that the solution is not converged and the numerical error in the signal accumulates from different reflections and paths taken by the wavefront. Increasing the CFL factor did not result in significant improvements.

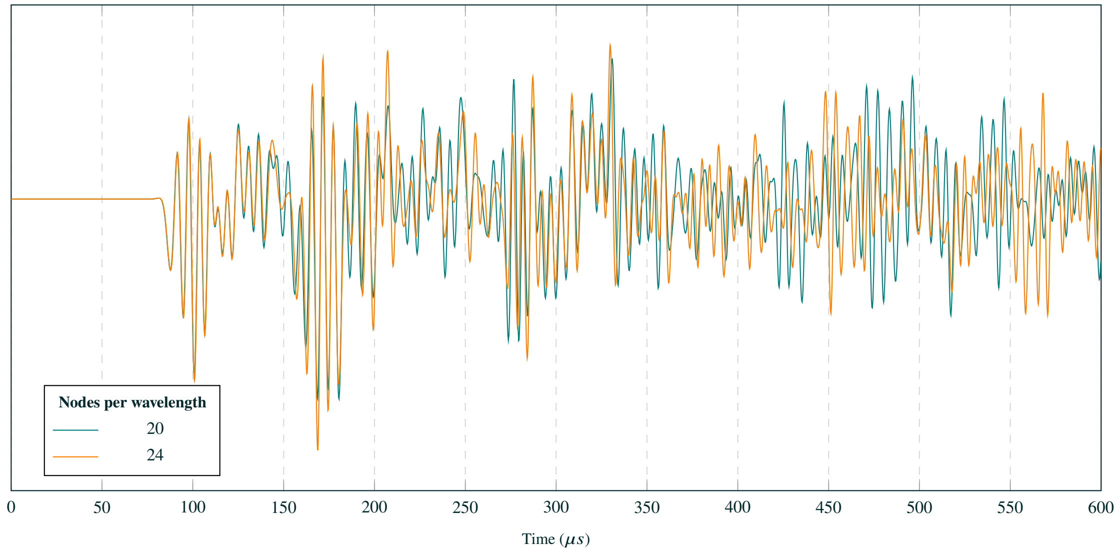


Figure 1.4 – Extracted Y component of the displacement field at 420mm from the source. The simulation is done using the space Q^1.

Using Q^4 finite elements. For the high-order case, we use finite elements constructed by polynomials in Q^4 and a mesh $\mathcal{T}_h$. An example of Q^4 mesh discretization using Gauss-Lobatto points is shown in Figure 1.3b. The mesh step was chosen to ensure a minimum of 8 or 16 nodes per wavelength. Using power iteration, we estimate the greatest eigenvalue of $\mathbb{M}^{-1}\mathbb{K}$, hence being able to compute a time step that satisfies the CFL condition (1.27). For this frequency, a time-step of $\Delta t = 0.02\mu$s was sufficient. Here, we also run the solver up to 600μs using a *Intel i9-9880H* CPU in a laptop workstation. The presented implementation for the spectral finite elements methods required 300MB of RAM and took 40 minutes to run for 8 nodes per wavelength, and 600MB/1h30 minutes for 16 nodes per wavelength. The extracted signals are plotted in Figure 1.5, showing a good convergence already for 8 nodes per wavelength. Snapshots of the computed wavefield are shown in Figure 1.6.

Discussion. In both implementations, the geometry discretization in mesh elements was done by the same basic algorithm, resulting in the same element size quality. Treatment and enhancements on mesh

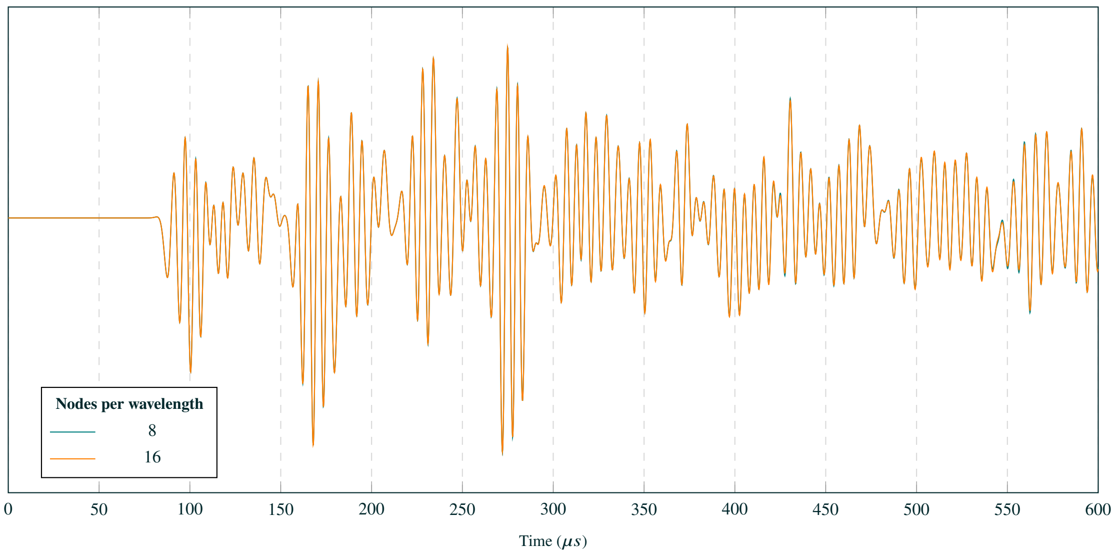
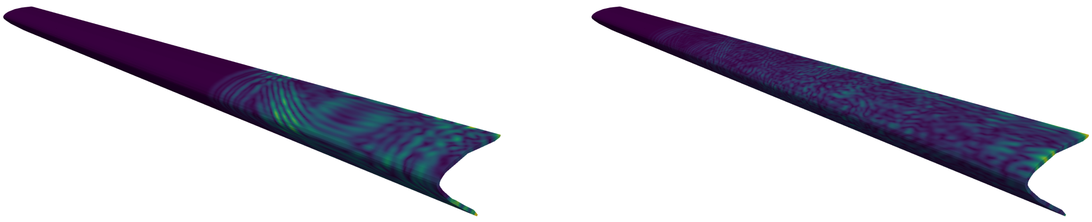


Figure 1.5 – Extracted Y component of the displacement field at 420mm from the source. The simulation was done using the space Q^4.



(a) t=60μs.

(b) t=300μs.

Figure 1.6 – Simulated wavefield generated by the a 200kHz source in a 2.3 meters long leading edge using spectral elements method.

quality could also be done to improve Q^1 discretization results. Treating the mesh quality would decrease computation time for both cases by decreasing the required time step for the CFL condition. Curved elements for the geometry could be used with the implementation of the spectral finite elements, but to ensure fair comparison, linear elements were used for the mesh.

Finally, we compare in Figure 1.7 the first-arrival in the two measurement points for Q^1 and Q^4 cases. We can see that as the wave propagates further, the solution amplitude decreases in Q^1 when compared to the converged Q^4 solution while presenting a delayed time of flight. These effects are due to numerical dissipation and dispersion in low-order finite elements, making it less suitable for medium-long range simulations. The efficient implementation of Q^4 is limited to hexahedral elements while the Q^1 one can be used with tetrahedral elements. This makes the Q^1 implementation more suitable to small (complex) geometries and short-time simulations, than not suitable for guided waves. The high-order spectral finite elements, with low memory consumption and negligible numerical artifacts, allow us to model and simulate long-range wave propagation as required for guided wave simulation.

1.5.2 A case with dimensions of small aircraft

We illustrate the use of the presented numerical method for the linear elastodynamic problem (1.20) with realistic physical dimensions. Using the dimensions of a small aircraft, we model a one-shot wing 6m long, 8mm thick and 1200mm wide. We excite a 50kHz ultrasonic wave with a radial transducer at

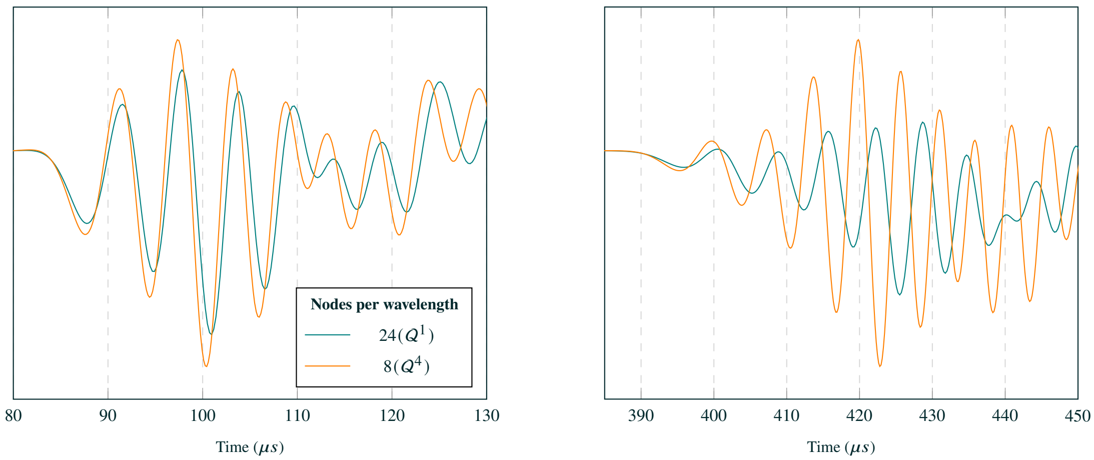
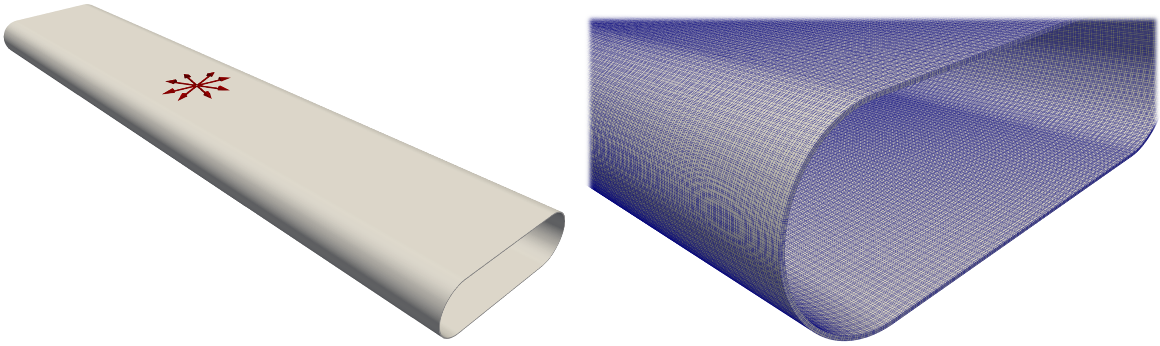


Figure 1.7 – Extracted Y component of the displacement field at 420mm (left) and 2000mm (right) from the source cut at first arrivals. It shows the effects of numerical dissipation and dispersion in Q^1 discretization.

the top and the bottom as depicted in Figure 1.8a. The domain discretization is shown in Figure 1.8b, where we can see the characteristic Gauss-Lobatto distribution of points for every element. The number of mesh points generated is 8.5 million, resulting in 25.5 million degrees of freedom.



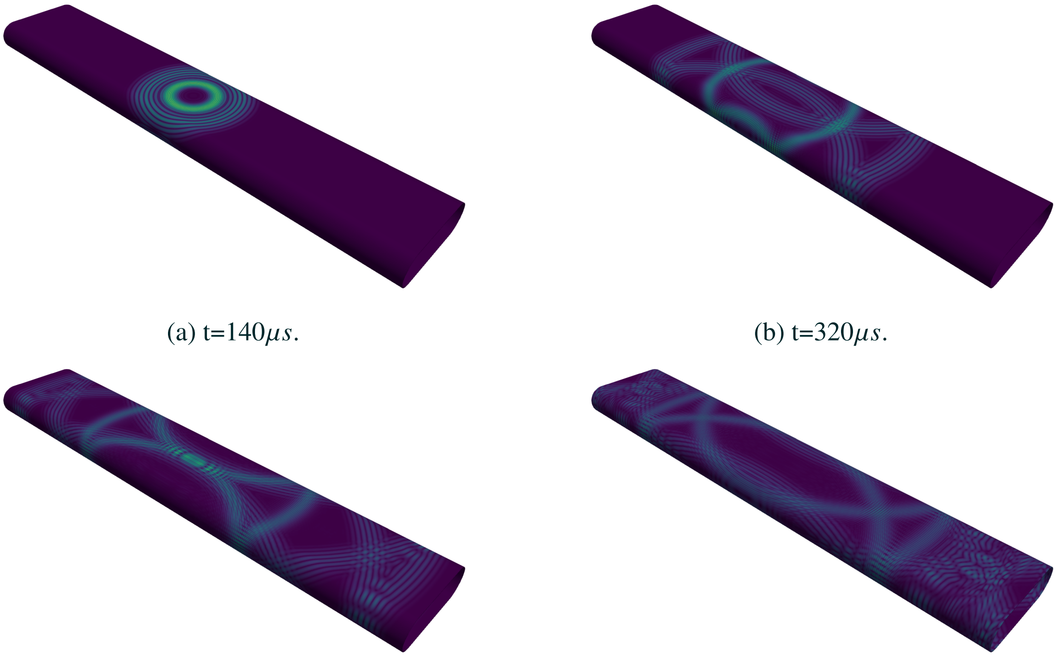
(a) Illustration domain $\hat{\Omega}$ and surface radial source position.

(b) Domain discretization with Gauss-Lobatto points ($\mathcal{T}_h$).

Figure 1.8 – Setup for simulating ultrasonic wave propagation in an aluminum specimen with dimensions of a small aircraft wing.

We consider the elasticity tensor (1.21) with the previously mentioned material parameters for aluminum. The excitation is done in a region of 20mm at the top and the bottom of the specimen. It is a radial source term representing a piezoelectric ultrasonic transducer that expands and contracts when excited with an electrical signal. We satisfy the CFL condition by using $\Delta t = 0.132318\mu$s. We run the simulation up to 1000μs, resulting in 7557 time steps. Due to mass lumping and the efficient application of the stiffness matrix, the computation had a peak memory usage of 920MB RAM and took 45 minutes to complete in a laptop workstation equipped with a *Intel i9-9880H* CPU.

Snapshots of the wavefield generated are shown in Figure 1.9. The ultrasonic excitation being done at the surface excitates waves that, by reflecting on the opposite surfaces, result in a guided wave. Elastic-guided waves can assume different propagation modes and in the case where it is generated by reflections between two parallel surfaces, we have the Lamb wave modes [Lamb, 1917]. The propagation starts at the top and bottom, propagating omnidirectionally and traveling around the specimen. In the snapshots we can recognize the S0 (symmetric) Lamb mode, propagating at a higher velocity, followed by the A0 (asymmetric) Lamb mode. This illustration demonstrates the efficiency of such implementation while outputting the full elastodynamic field, making it suitable for guided wave propagation in complex geometries in realistic scenarios.



(a) t=140μs.

(b) t=320μs.

(c) t=520μs.

(d) t=700μs.

Figure 1.9 – Wavefield simulated in the 6 meter long aluminum specimen. With 25.5 million degrees of freedom and 7557 time steps, the computation required 920MB RAM and took 45 minutes to run up to 1000μs.

1.6 Conclusion

In this chapter, we presented the main concepts of nonlinear mechanics and the hyperelastic constitutive laws. By assuming small displacements and strain, we retrieve a linearized formulation and discuss the constitutive laws in the linear framework. By introducing the Galerkin approximation and using a second-order time scheme we analyze energy conservation and stability properties for a linear elastodynamics problem. The implementation of the spectral finite elements method was detailed. With the chosen time scheme and mass lumping, we have an explicit scheme and by doing an unassembled application of the stiffness matrix, we achieve a low memory usage and efficient parallelization. Finally, we illustrate why high-order finite elements are important for guided-waves propagation and present the performance of the implementation used in this work. The overview of the mechanical and numerical modeling done in this chapter enables us to justify and understand the modeling choices to propose a robust numerical strategy for wave propagation in structures under loading conditions.

Chapter 2

A time-domain spectral finite element method for acoustoelasticity: modeling the effect of mechanical loading on guided wave propagation

---

This chapter aims to present the first part of the original work done in this thesis. The mechanical and numerical modeling of ultrasonic wave propagation in mechanically loaded structures, with a focus on applicability to guided waves, is presented. In the context of acoustoelasticity, we assume the ultrasound to be incremental deformations in addition to a (potentially) large deformation of the structure. We propose an original combination of 3D shell finite elements and transient spectral finite elements to solve the problems that emerge from the *small-on-large* assumption. We illustrate potential stability issues associated with loss of coercivity in different constitutive laws. We present realistic numerical results on 3D cases in both isotropic and anisotropic materials. Finally, the methodology is validated using experimental data from the literature. The chapter takes the form of a pre-print article co-authored by André Dalmora, Alexandre Imperiale, Sébastien Imperiale and Philippe Moireau.

Contents

A time-domain spectral finite element method for acoustoelasticity: modeling the effect of mechanical loading on guided wave propagation

Andre Dalmora[1,2,3], *Alexandre Imperiale*[1], *Sébastien Imperiale*[2,3], *Philippe Moireau*[2,3]

1 - Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France
2 - Inria, Team-MΞDISIM, Inria-Saclay Ile de France 91120 Palaiseau, France
3 - LMS, Ecole Polytechnique, CNRS – Institut Polytechnique de Paris

Abstract. Ultrasonic testing techniques such as guided wave-based structural health monitoring aim to evaluate the integrity of a material with sensors and actuators that operate *in situ*, *i.e.* while the material is in use. Since ultrasonic wave propagation is sensitive to environmental conditions such as pre-deformation of the structure, the design and performance evaluation of monitoring systems in this context is a complicated task that requires quantitative data and the associated modeling effort. In our work, we propose a set of numerical tools to solve the problem of mechanical wave propagation in materials subjected to pre-deformation. This type of configuration is usually treated in the domain of acoustoelasticity. A relevant modeling approach is to consider two different problems: a quasi-static nonlinear problem for the large displacement field of the structure and a linearized time-domain wave propagation problem. After carefully reviewing the modeling ingredients to represent the configurations of interest, we propose an original combination of numerical tools that leads to a computationally efficient algorithm. More specifically, we use 3D shell elements for the quasi-static nonlinear problem and the time-domain spectral finite element method to numerically solve the wave propagation problem. Our approach can represent any type of material constitutive law, geometry or mechanical solicitation. We present realistic numerical results on 3D cases related to the monitoring of both isotropic and anisotropic materials, illustrating the genericity and efficiency of our method. We also validate our approach by comparing it to experimental data from the literature.

2.1 Introduction

In many high-end industries, safety regulations include assessments of the condition of structures and materials used for operations. Typical examples of application areas are nuclear energy, petrochemical industry, transportation, or aeronautics. In fact, any industry where the safety of critical components is of paramount importance. Over time, the need to ensure or assess the health of these components has led to the development of advanced Nondestructive Testing (NDT) techniques such as Structural Health Monitoring (SHM). SHM systems aim to evaluate the structural integrity of the material in question using sensors and processing units *in situ*. In other words, the monitoring systems are attached to the structure once and for all, so that they can continuously generate relevant field data. One means of implementing such systems is to rely on ultrasonic Guided Waves (GW) [Mitra and Gopalakrishnan, 2016; Ricci et al., 2022]. These waves have the advantage of propagating over longer distances – compared to bulk waves, which are arguably more common in the NDT community – and thus probe a larger volume of material. Nonetheless, GW propagation is potentially more complex. Namely, it involves dispersion effects that vary from one propagating mode to another, and in the particular case of SHM systems with *in situ* control processes, it is known to be affected by Environmental and Operational Conditions (EOCs) [Gorgin, Luo, and Wu, 2020]. One of the most general types of EOC that we consider in our work is mechanical solicitations sustained by the structure during its use. These mechanical conditions lead to potential deformations and internal stresses that affect GW propagation. This is the so-called acoustoelastic effect [Shams, Destrade, and Ogden, 2011; Abiza, Destrade, and Ogden, 2012]. To provide meaningful assistance in the development and evaluation of SHM system performance, our work addresses the challenge of incorporating acoustoelastic effects into a wave propagation modeling tool that is both generic – *i.e.* applicable to any geometry, constitutive law or mechanical solicitation – and efficient.

There are numerous previous works in the literature that address this issue but, to our knowledge, they are mainly based on semi-analytical approaches. For instance, in [Gandhi, Michaels, and Lee, 2012; Pau and Scalea, 2015; Pei and Bond, 2016; Dubuc, Ebrahimkhanlou, and Salamone, 2017a; Dubuc, Ebrahimkhanlou, and Salamone, 2017b; Mohabuth et al., 2018; Dubuc, Ebrahimkhanlou, and Salamone, 2018], analytical formulations are proposed and can be solved at low cost. Nevertheless, due to assumptions regarding the geometry or the deformation, such formulations can be limited when precisely modeling complex and realistic cases. For complex shapes and loading configurations, in [Chen and Wilcox, 2007; Loveday, Long, and Wilcox, 2012; Peddeti and Santhanam, 2018; Cheng et al., 2020; Abderahmane, Lhémery, and Daniel, 2021] the authors propose semi-analytical methods that encompass stress inhomogeneity in one or two dimensions. These methods treat the problem in the frequency domain and efficiently compute dispersion curves of guided waves in specified wave guides, given the stress profile through the thickness. In our work, we aim at proposing a time-domain approach for this problem, while considering arbitrary geometries, hyperelastic laws, and inhomogeneous stress profiles caused by mechanical loading.

Traditionally, due to the different characteristic times and amplitudes of the structural strain, one can identify two (unidirectionally coupled) problems: (1) a quasi-static nonlinear mechanical problem for the structural deformation; (2) a time-dependent wave problem resulting from a linearization procedure around the quasi-static deformation. In this work, we aim to extend the reach of acoustoelastic GW propagation modeling by combining a set of generic numerical tools to solve both problems. For the first problem, the main challenge (apart from mastering the inherent nonlinearities) is that classical finite element methods (FEMs) give inadequate results for thin, elongated geometries due to numerical locking [Bathe, 2006; Chapelle and Bathe, 2011]. To overcome this difficulty we use 3D shell elements [Chapelle and Bathe, 2011]. They are based on an asymptotic expansion of the displacement field in the thickness direction and, combined with the so-called Mixed Interpolation of Tensorial Components (MITC) method [Bathe and Dvorkin, 1986; Bucalem and Bathe, 1993; Bathe, Iosilevich, and Chapelle, 2000] they are locking-free. Moreover, they are readily available for any kind of constitutive law. The output of this first computational stage is the (large and quasi-static) displacement field sustained by the structure. This output is then used as an input for the wave propagation problem. For this second problem, we use a different type of discretization, namely the Spectral Finite Element Method (SFEM) in the time domain [Komatitsch et al., 1999; Cohen, 2002; Joly, 2007]. Since the unknown of each problem are fundamentally different in nature, it is natural to resort to different discretization schemes, in particular different meshes. This entails the use of an interpolation operator from the first problem's mesh onto the second one. The SFEM is based on high-order Lagrangian elements defined on Gauss-Lobatto points. This avoids the Runge phenomenon and achieves spectral-like convergence. Using a consistent mass lumping procedure, they lead to a very efficient explicit fully discrete scheme – with a stability condition that depends on the mechanical deformation of the structure. To illustrate the validity and efficiency of our approach, we provide meaningful 3D simulation results associated with SHM configurations, as well as experimental validation results with data from the literature [Gandhi, Michaels, and Lee, 2012].

This article is divided into two main sections. In Section 2.2 we explain the methodology we use in our work. In particular, we first recall the general formalism of nonlinear elastodynamics, followed by the linearization procedure around a quasi-static displacement field. In this way, we can accurately define the wave propagation problem involving the acoustoelastic effects. After comparing our modeling approach with the relevant previous work on acoustoelasticity, we conclude this section with the dedicated time and space numerical schemes applied in order to solve the one-way coupled problems, namely the quasi-static nonlinear problem and the linearized wave propagation problem. The Section 2.3 is devoted to the detailed 3D numerical illustrations of our approach. First, we provide illuminating examples addressing the potential issue of stability of the fully discrete wave propagation model as a function of the type of deformation and the constitutive law. Then we provide illustrations of SHM configurations, namely a tube of isotropic material subjected to a four-point bending test and a stratified anisotropic plate subjected to torsion. These examples illustrate the performances and versatility of our approach, while in the last part of the section we present experimental validations showing the reliability of our modeling toolchain.

2.2 Methodology for modeling the effect of mechanical loading on elastic wave propagation

2.2.1 Nonlinear elastodynamics with quasi-static surface traction & time dependent ultrasonic actuator

This section is dedicated to deriving a model representing the propagation of "*low amplitude – high frequency*" waves, within a material subject to a "*high amplitude – (very) low frequency*" mechanical loading. This model enters the traditional framework of finite deformations of hyperelastic solids. Readers may refer to the reference textbooks [Marsden and Hughes, 1978; Ciarlet, 1988b; Le Tallec, 1994] and references therein for an exhaustive presentation of this generic formalism. We do not seek here to deviate from the very well-understood formalism of hyperelasticity. Rather, we aim at specifying the corresponding modeling elements, in particular the various forces that act on the material of interest. This aspect is of paramount importance through our work, since it leads to a linearization procedure and its subsequent wave propagation model.

Let us denote by $t \geq 0$ the time variable with $T > 0$ the maximal time of interest, and by $\Omega(t) \subset \mathbb{R}^3$ the volume occupied by the material during its deformation, with boundary $\Gamma(t)$. We assume that there exists $\Gamma_D \subset \Gamma(t)$ a clamped part of the boundary. In the following, we consider the functional space $\mathcal{V}(\Omega(t))$ of admissible virtual displacements vanishing on Γ_D.

Quasi-static surface traction. We consider a sub-part of the boundary $\Gamma_0(t) \subset \Gamma$ that is the area of action of a high amplitude and quasi-static force, and such that $\Gamma_0(t) \cap \Gamma_D = \varnothing$. More precisely, let us denote by f_0 a prescribed scalar-valued function representing the amplitude of a surface traction. The associated virtual work amounts to

$$\int_{\Gamma_0(t)} f_0(\boldsymbol{x})\boldsymbol{n}_0(\boldsymbol{x},t) \cdot \boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\Gamma, \qquad \forall \boldsymbol{w} \in \mathcal{V}(\Omega(t)),$$

where $\boldsymbol{n}_0$ is the outgoing unit normal vector field of Γ_0. The quasi-static hypothesis represents the fact that the time variation of this surface traction is assumed to be significantly lower than the ultrasonic source term described subsequently.

Time dependent ultrasonic actuator. We also consider a third part of the boundary $\Gamma_a(t) \subset \Gamma(t)$ where a low amplitude and time-dependent ultrasonic actuator is vibrating, thus generating propagating wave packets. We assume that $\Gamma_a(t) \cap \Gamma_0(t) = \varnothing$ and $\Gamma_a(t) \cap \Gamma_D = \varnothing$. We denote by $\boldsymbol{f}_a^\delta$ a given source term representing this ultrasonic actuator. The superscript δ represents the amplitude ratio of the ultrasonic source term compared to f_0. In the following, we extensively use the assumption that this ratio is small, *i.e.*

$$\delta = \frac{\sup\limits_{(\boldsymbol{x},t)\in\Omega\times[0;T]} \|\boldsymbol{f}_a^\delta(\boldsymbol{x},t)\|_3}{\sup\limits_{\boldsymbol{x}\in\Omega} |f_0(\boldsymbol{x})|} \ll 1. \tag{2.1}$$

In the following, we assume – without loss of generality – that there exists a vector field $\boldsymbol{f}_a$ that corresponds to a rescaled source term, namely $\boldsymbol{f}_a^\delta = \delta \boldsymbol{f}_a$. The associated virtual work becomes

$$\delta \int_{\Gamma_a(t)} \boldsymbol{f}_a(\boldsymbol{x},t) \cdot \boldsymbol{w}(\boldsymbol{x})\,\mathrm{d}\Gamma, \qquad \forall \boldsymbol{w} \in \mathcal{V}(\Omega(t)), \qquad \forall t \in [0;T].$$

Weak formulations of the nonlinear elastodynamic problem. Let us introduce ϱ the mass density of the material. With these notations and the previous definitions of the various boundary conditions, the principle of virtual works states that the velocity field $\boldsymbol{v}(t)$ and the Cauchy stress tensor $\boldsymbol{\sigma}(t)$ satisfy, for any $\boldsymbol{w} \in \mathcal{V}(\Omega(t))$

$$\int_{\Omega(t)} \varrho\big(\partial_t \boldsymbol{v} + (\nabla_x \boldsymbol{v}) \cdot \boldsymbol{v}\big) \cdot \boldsymbol{w} \, \mathrm{d}\Omega + \int_{\Omega(t)} \boldsymbol{\sigma} : \nabla_x \boldsymbol{w} \, \mathrm{d}\Omega = \int_{\Gamma_0(t)} f_0 \boldsymbol{n}_0 \cdot \boldsymbol{w} \, \mathrm{d}\Gamma + \delta \int_{\Gamma_a(t)} \boldsymbol{f}_a \cdot \boldsymbol{w} \, \mathrm{d}\Gamma, \quad (2.2)$$

with given initial conditions $\boldsymbol{v}|_{t=0}$ and $\boldsymbol{\sigma}|_{t=0}$. Albeit being the standard weak form of the equilibrium relation, relation (2.2) deserves some comments. It is not an explicit form because the deformed domain Ω is unknown, nor is it a closed form since it lacks the constitutive law linking the Cauchy stress tensor with the kinematic unknown. It should be noted that this constitutive law is rarely written on a deformed configuration, but rather on a so-called reference configuration. Also, from a purely practical viewpoint, the definition of the source term f_a representing the ultrasonic actuator is more likely to be available on this reference configuration.

From this line of arguments, it is natural to derive a formulation written on a reference coordinate system. Let us denote by $\widehat{\Omega}$ the volume occupied by the material in this reference setting, and by $\boldsymbol{\xi} \in \widehat{\Omega}$ the associated coordinates. For any time $t > 0$, the deformation is represented by the bijective mapping,

$$\begin{aligned} \boldsymbol{\phi}(t) \ : \widehat{\Omega} \ &\longmapsto \Omega(t) \\ \boldsymbol{\xi} \ &\longrightarrow \boldsymbol{x}(\boldsymbol{\xi}, t) = \boldsymbol{\phi}(\boldsymbol{\xi}, t), \end{aligned}$$

and we define the displacement field $\boldsymbol{u}(t)$ as

$$\boldsymbol{u}(\boldsymbol{\xi}, t) = \boldsymbol{x}(\boldsymbol{\xi}, t) - \boldsymbol{\xi}, \qquad \forall \boldsymbol{\xi} \in \widehat{\Omega}.$$

We denote by $\widehat{\Gamma}$ the boundary of $\widehat{\Omega}$, and $\widehat{\Gamma}_D$, $\widehat{\Gamma}_0$, and $\widehat{\Gamma}_a$ the sub-parts of the boundary corresponding to Γ_D, $\Gamma_0(t)$, and $\Gamma_a(t)$ respectively. See Figure 2.1 for an illustration of both coordinate systems and domains.

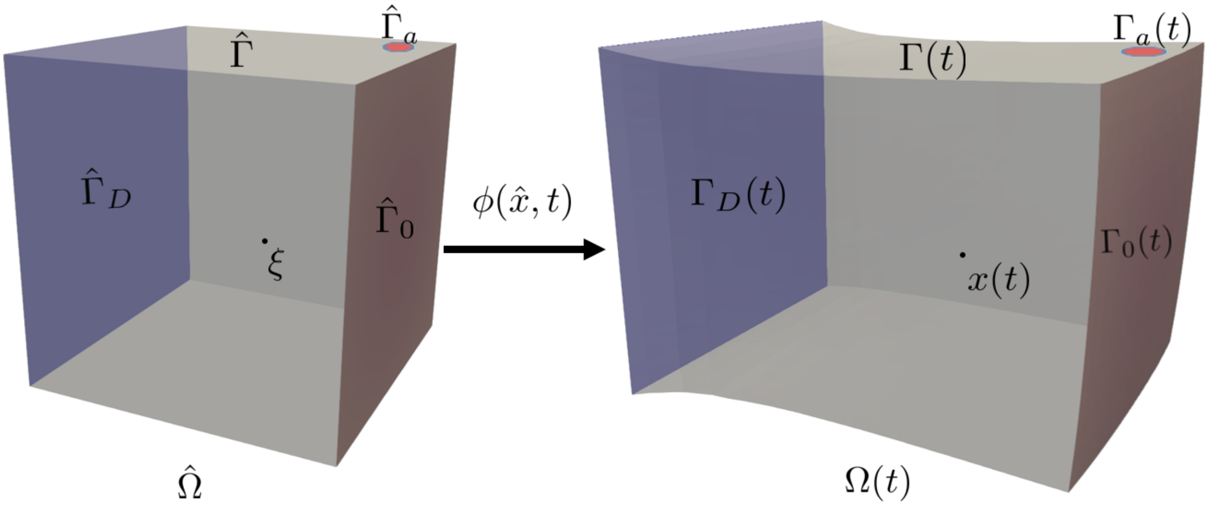


Figure 2.1 – Reference and deformed configurations and their associated notations.

The deformation gradient and its determinant are denoted by

$$\boldsymbol{F}(\boldsymbol{\xi}, t) = \nabla_\xi \boldsymbol{\phi}(\boldsymbol{\xi}, t) = \boldsymbol{I} + \nabla_\xi \boldsymbol{u}(\boldsymbol{\xi}, t) \quad \text{and} \quad J(\boldsymbol{\xi}, t) = \det \boldsymbol{F}(\boldsymbol{\xi}, t), \qquad \forall \boldsymbol{\xi} \in \widehat{\Omega},$$

respectively. With these notations, we derive the equivalent of (2.2) set on the reference configuration, thus explicitly showing the dependency of the displacement field – which is the chosen kinematic unknown in the following. To do so, we apply a standard change of variables from Ω to $\widehat{\Omega}$ – see [Le Tallec, 1994] and references therein for more details. Upon the straightforward relation $\boldsymbol{v} \circ \boldsymbol{\phi} = \partial_t \boldsymbol{u}$, we start with the inertial terms, namely, for all $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$,

$$\int_{\Omega(t)} \varrho \big(\partial_t \boldsymbol{v} + (\nabla_x \boldsymbol{v}) \cdot \boldsymbol{v}\big) \cdot (\boldsymbol{w} \circ \boldsymbol{\phi}^{-1}) \,\mathrm{d}\Omega = \int_{\widehat{\Omega}} J\,(\varrho \circ \boldsymbol{\phi})\, \partial^2_{tt} \boldsymbol{u} \cdot \boldsymbol{w} \,\mathrm{d}\widehat{\Omega} = \int_{\widehat{\Omega}} \widehat{\varrho}\, \partial^2_{tt} \boldsymbol{u} \cdot \boldsymbol{w} \,\mathrm{d}\widehat{\Omega},$$

where $\widehat{\varrho} = (\varrho \circ \boldsymbol{\phi}) J$ is the reference mass density. Concerning the stiffness term, we have, for all $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$,

$$\int_{\Omega(t)} \boldsymbol{\sigma} : \nabla_x (\boldsymbol{w} \circ \boldsymbol{\phi}^{-1}) \,\mathrm{d}\Omega = \int_{\widehat{\Omega}} (\boldsymbol{\sigma} \circ \boldsymbol{\phi}) : (\nabla_\xi \boldsymbol{w} \cdot \boldsymbol{F}^{-1})\, J \mathrm{d}\widehat{\Omega} = \int_{\widehat{\Omega}} \boldsymbol{T} : \nabla_\xi \boldsymbol{w} \,\mathrm{d}\widehat{\Omega},$$

where $\boldsymbol{T} = (\boldsymbol{\sigma} \circ \boldsymbol{\phi}) \cdot \boldsymbol{F}^{-\intercal} J$ is the first Piola-Kirchhoff tensor. Upon the assumption of hyperelasticity [Ciarlet, 1988b; Le Tallec, 1994] there exists a potential function $\mathscr{W}(\boldsymbol{F})$ that depends exclusively on the deformation gradient and such that

$$\boldsymbol{T}(\boldsymbol{F}) = \mathrm{D}_{\boldsymbol{F}} \mathscr{W}(\boldsymbol{F}).$$

Note that prescribing a closed form of this energy potential function actually defines the constitutive relation of the material – a matter addressed in the subsequent section. The first Piola-Kirchhoff tensor is non-symmetric, and in the following, we resort to its symmetric counterpart, which can be introduced as follows,

$$\int_{\widehat{\Omega}} \boldsymbol{T} : \nabla_\xi \boldsymbol{w} \,\mathrm{d}\widehat{\Omega} = \int_{\widehat{\Omega}} (\boldsymbol{F}^{-1} \cdot \boldsymbol{T}) : (\boldsymbol{F}^\intercal \cdot \nabla_\xi \boldsymbol{w}) \,\mathrm{d}\widehat{\Omega} = \int_{\widehat{\Omega}} \boldsymbol{\Sigma} : \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{u}) \boldsymbol{w} \,\mathrm{d}\widehat{\Omega},$$

with $\boldsymbol{\Sigma} = \boldsymbol{F}^{-1} \cdot \boldsymbol{T}$ the (symmetric) second Piola-Kirchhoff tensor, and

$$\boldsymbol{e}(\boldsymbol{u}) = \frac{1}{2}\Big(\nabla_\xi \boldsymbol{u} + \nabla_\xi \boldsymbol{u}^\intercal + \nabla_\xi \boldsymbol{u}^\intercal \cdot \nabla_\xi \boldsymbol{u}\Big), \qquad \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{u}) \boldsymbol{w} = \frac{1}{2}\Big(\boldsymbol{F}^\intercal \cdot \nabla_\xi \boldsymbol{w} + \nabla_\xi \boldsymbol{w}^\intercal \cdot \boldsymbol{F}\Big),$$

the Green-Lagrange tensor and its differential *w.r.t.* the displacement field. For hyperelastic materials, the second Piola-Kirchhoff tensor is usually defined exclusively as a function of the Green-Lagrange tensor, namely $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{u}))$ in the previous expression, and

$$\boldsymbol{\Sigma}(\boldsymbol{e}) = \mathrm{D}_{\boldsymbol{e}} \mathscr{W}(\boldsymbol{e}).$$

For the quasi-static surface traction, one resorts to the formula for transport of normal vector field to derive the following relations

$$\forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}), \qquad \int_{\Gamma_0(t)} f_0 \boldsymbol{n}_0 \cdot (\boldsymbol{w} \circ \boldsymbol{\phi}^{-1}) \,\mathrm{d}\Gamma = \int_{\widehat{\Gamma}_0} \widehat{f}_0\, (\mathrm{co}\boldsymbol{F} \cdot \widehat{\boldsymbol{n}}_0) \cdot \boldsymbol{w} \,\mathrm{d}\widehat{\Gamma},$$

where $\widehat{f}_0 = f_0 \circ \boldsymbol{\phi}$, $\widehat{\boldsymbol{n}}_0$ is the outgoing unit normal vector field of $\widehat{\Gamma}_0$, and $\mathrm{co}\boldsymbol{F} = J \boldsymbol{F}^{-\intercal}$ is the co-factor of the deformation gradient. Finally, concerning the ultrasonic source term, we have

$$\forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}), \qquad \int_{\Gamma_a(t)} \boldsymbol{f}_a \cdot (\boldsymbol{w} \circ \boldsymbol{\phi}^{-1}) \,\mathrm{d}\Gamma = \int_{\widehat{\Gamma}_a} \widehat{\boldsymbol{f}}_a \cdot \boldsymbol{w}\, \widehat{\gamma} \,\mathrm{d}\widehat{\Gamma},$$

where $\widehat{\gamma} = \|\mathrm{co}\boldsymbol{F} \cdot \widehat{\boldsymbol{n}}_a\|_3$ with $\widehat{\boldsymbol{n}}_a$ the unit normal vector field of $\widehat{\Gamma}_a$. This term represents the effect of the deformation on the surface measure. In the previous relation, $\widehat{\boldsymbol{f}}_a$ is the time and space dependency of the actuator on the reference configuration. Its expression depends on the chosen modeling of the actuator. Without being completely exhaustive, we can at least provide three main cases of interest:

- The ultrasonic source is independent of the deformation of the material. For instance, one could imagine configurations where a hydrophone is emitting an ultrasonic beam toward the immersed specimen of interest, subject to the deformation. In that case, we simply have $\widehat{\boldsymbol{f}}_a = \boldsymbol{f}_a \circ \boldsymbol{\phi}$.

- The actuator is considered to be generating a time-dependent source term with polarization in the normal direction of the material. This corresponds to the case where there exists a scalar-valued function f_a such that $\boldsymbol{f}_a = f_a \boldsymbol{n}_a$, where $\boldsymbol{n}_a$ is the outgoing unit normal vector field of Γ_a. In that case, we define
$$\widehat{\boldsymbol{f}}_a = (f_a \circ \boldsymbol{\phi}) \, \frac{\mathrm{co}\boldsymbol{F} \cdot \widehat{\boldsymbol{n}}_a}{\widehat{\gamma}} \, .$$
- The polarization lies within the local tangent plane of the surface Γ_a. More precisely if we denote by $(\boldsymbol{\tau}_{1,a}, \boldsymbol{\tau}_{2,a})$ two vectors field generating this tangent plane, then there exists two scalar-valued functions $f_{1,a}$ and $f_{2,a}$ such that $\boldsymbol{f}_a = f_{1,a}\boldsymbol{\tau}_{1,a} + f_{2,a}\boldsymbol{\tau}_{2,a}$, entailing
$$\widehat{\boldsymbol{f}}_a = \sum_{i=1,2} (f_{i,a} \circ \boldsymbol{\phi}) \, \frac{\boldsymbol{F} \cdot \widehat{\boldsymbol{\tau}}_{i,a}}{\|\boldsymbol{F} \cdot \widehat{\boldsymbol{\tau}}_{i,a}\|_3}, \qquad \boldsymbol{\tau}_{i,a} \circ \boldsymbol{\phi} = \boldsymbol{F} \cdot \widehat{\boldsymbol{\tau}}_{i,a}, \quad \forall i \in \{1, 2\}.$$

We can now give the weak formulation of the problem of interest written in the reference configuration: for any time $t \in [0; T]$, we aim at finding the displacement field $\boldsymbol{u}(t)$ such that, for any test function $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$, we have

$$\int_{\widehat{\Omega}} \widehat{\varrho} \, \partial^2_{tt} \boldsymbol{u} \cdot \boldsymbol{w} \, \mathrm{d}\widehat{\Omega} + \int_{\widehat{\Omega}} \boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{u})) : \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{u}) \boldsymbol{w} \, \mathrm{d}\widehat{\Omega} \\ = \int_{\widehat{\Gamma}_0} \widehat{f}_0 \, (\mathrm{co}\boldsymbol{F} \cdot \widehat{\boldsymbol{n}}_0) \cdot \boldsymbol{w} \, \mathrm{d}\widehat{\Gamma} + \delta \int_{\widehat{\Gamma}_a} \widehat{\boldsymbol{f}}_a \cdot \boldsymbol{w} \, \widehat{\gamma} \, \mathrm{d}\widehat{\Gamma}, \quad (2.3)$$

with given initial conditions $\boldsymbol{u}|_{t=0}$ and $\partial_t \boldsymbol{u}|_{t=0}$.

2.2.2 Linearization around a quasi-static displacement field

Upon the assumption of a quasi-static mechanical loading and a low amplitude time-dependent ultrasonic source term – as depicted in (2.1) – we propose to decompose the solution of (2.3) following the Ansatz

$$\boldsymbol{u}(\boldsymbol{\xi}, t) = \boldsymbol{u}_0(\boldsymbol{\xi}) + \delta \, \boldsymbol{u}_1(\boldsymbol{\xi}, t) + O(\delta^2), \qquad \forall \boldsymbol{\xi} \in \widehat{\Omega}, \qquad \forall t \in [0; T]. \quad (2.4)$$

This (natural) decomposition amounts to seeking a large quasi-static deformation $\boldsymbol{u}_0$ generated by the surface traction term, and a low amplitude time-dependent contribution $\boldsymbol{u}_1$ emanating from the ultrasonic source term. Intuitively, one would expect $\boldsymbol{u}_0$ to satisfy a nonlinear static problem, and $\boldsymbol{u}_1$ a wave equation. In this section, we follow standard linearization arguments to derive these two problems. Moreover, following the existence and uniqueness results given in [Ciarlet, 1988b], we define the space

$$\mathcal{V}(\widehat{\Omega}) = \Big\{ \boldsymbol{w} \in [H^1(\widehat{\Omega})]^3 \mid \boldsymbol{w}|_{\Gamma_D} = 0 \Big\},$$

and the space $\mathcal{W}(\widehat{\Omega})$ of admissible displacements in the reference configuration

$$\mathcal{W}(\widehat{\Omega}) = W^{2,p}(\widehat{\Omega})^3 \cap \mathcal{V}(\widehat{\Omega}), \text{ for some } p > 3. \quad (2.5)$$

We now comment briefly this choice. Imbeddings results (see [Adams and Fournier, 2003; Ciarlet, 1988b]) show that, for any $\boldsymbol{w}^\star \in \mathcal{W}(\widehat{\Omega})$, first, $\nabla_{\boldsymbol{\xi}} \boldsymbol{w}^\star$ belongs to $\mathrm{W}^{1,p}(\widehat{\Omega})^{3\times3}$ that is an algebra (in particular products of functions in this space do belong to the same space), second, $\nabla_{\boldsymbol{\xi}} \boldsymbol{w}^\star$ is continuous and bounded up to the boundary. We deduce that

$$\boldsymbol{e}(\boldsymbol{w}^\star) \in \mathrm{W}^{1,p}(\widehat{\Omega})^{3\times3} \quad \text{and} \quad \forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}), \quad \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{w}^\star)\boldsymbol{w} \in \mathrm{L}^2(\widehat{\Omega})^{3\times3}.$$

Finally, assuming sufficient smoothness of the potential function $\mathscr{W}(\boldsymbol{F})$ it can be inferred that $\boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{w}^\star)) \in \mathrm{W}^{1,p}(\widehat{\Omega})^{3\times3}$ hence, it is continuous and bounded. A similar smoothness property holds for the fourth order tensor $\mathrm{D}_{\boldsymbol{e}}\boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{w}^\star))$. The properties given above are implicitly used to give meanings to the integrals introduced in the section below.

Stiffness and surface traction operators and associated differentials. Let us denote by

$$\begin{aligned} \mathrm{A} \ : \mathcal{W}(\widehat{\Omega}) &\longmapsto \mathcal{V}(\widehat{\Omega})' \\ \boldsymbol{w}^\star &\longrightarrow \mathrm{A}(\boldsymbol{w}^\star), \end{aligned}$$

the nonlinear stiffness operator such that, for any $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$,

$$\langle \mathrm{A}(\boldsymbol{w}^\star), \boldsymbol{w} \rangle = \int_{\widehat{\Omega}} \boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{w}^\star)) : \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{w}^\star)\boldsymbol{w} \, \mathrm{d}\widehat{\Omega},$$

where $\langle \cdot, \cdot \rangle$ denotes the duality product in $\mathcal{V}(\widehat{\Omega})$. The Gâteaux differential of A at $\boldsymbol{w}^\star$ is denoted by $\mathrm{D}_{\boldsymbol{u}} \mathrm{A}(\boldsymbol{w}^\star) \in \mathcal{L}(\mathcal{V}(\widehat{\Omega}), \mathcal{V}(\widehat{\Omega})')$, and reads, for any $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$

$$\langle \mathrm{D}_{\boldsymbol{u}} \mathrm{A}(\boldsymbol{w}^\star)\boldsymbol{v}, \boldsymbol{w} \rangle = \int_{\widehat{\Omega}} \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{w}^\star)\boldsymbol{v} : \mathrm{D}_{\boldsymbol{e}} \boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{w}^\star)) : \mathrm{D}_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{w}^\star)\boldsymbol{w} \, \mathrm{d}\widehat{\Omega} + \int_{\widehat{\Omega}} \boldsymbol{\Sigma}(\boldsymbol{e}(\boldsymbol{w}^\star)) : \mathrm{D}^2_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{v}, \boldsymbol{w}) \, \mathrm{d}\widehat{\Omega}, \tag{2.6}$$

where

$$\mathrm{D}^2_{\boldsymbol{u}} \boldsymbol{e}(\boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2}\Big(\nabla_\xi \boldsymbol{v}^\top \cdot \nabla_\xi \boldsymbol{w} + \nabla_\xi \boldsymbol{w}^\top \cdot \nabla_\xi \boldsymbol{v}\Big).$$

Note that the differential (2.6) is obtained from the differential of a product combined with a chain rule, and the definition of the Green-Lagrange tensor. It actually defines a bilinear form of particular interest in the following. More precisely, for a given $\boldsymbol{w}^\star \in \mathcal{W}(\widehat{\Omega})$ defining

$$\begin{aligned} a(\boldsymbol{w}^\star; \cdot, \cdot) \ : \ \mathcal{V}(\widehat{\Omega}) \times \mathcal{V}(\widehat{\Omega}) &\longmapsto \mathbb{R} \\ (\boldsymbol{v}, \boldsymbol{w}) &\longrightarrow a(\boldsymbol{w}^\star; \boldsymbol{v}, \boldsymbol{w}) = \langle \mathrm{D}_{\boldsymbol{u}} \mathrm{A}(\boldsymbol{w}^\star)\boldsymbol{v}, \boldsymbol{w} \rangle, \end{aligned} \tag{2.7}$$

then one can verify that $a(\boldsymbol{w}^\star; \cdot, \cdot)$ is a symmetric bilinear form. Note that $\mathrm{D}_{\boldsymbol{u}} \mathrm{A}$ is a contribution of two terms: the first one, involving $\mathrm{D}_{\boldsymbol{e}} \boldsymbol{\Sigma}$, is related to constitutive nonlinearities since it involves second order derivatives of the potential $\mathscr{W}$ with respect to the Green-Lagrange tensor $\boldsymbol{e}$; the second one, involving $\boldsymbol{\Sigma}$, is related to geometrical nonlinearities since it comes exclusively from the linearization of the Green-Lagrange tensor. Another more compact form – but less prone to interpretation – of the Gâteaux differential of A can be deduced from the expression

$$\langle \mathrm{A}(\boldsymbol{w}^\star), \boldsymbol{w} \rangle = \int_{\widehat{\Omega}} \boldsymbol{T}(\boldsymbol{F}(\boldsymbol{w}^\star)) : \boldsymbol{\nabla}_\xi \boldsymbol{w} \, \mathrm{d}\widehat{\Omega},$$

where $\boldsymbol{F}(\boldsymbol{w}^\star) = \boldsymbol{I} + \boldsymbol{\nabla}_\xi \boldsymbol{w}^\star$. Computing the Gâteaux differential of A at $\boldsymbol{w}^\star$ gives the alternative expression

$$\begin{aligned} a(\boldsymbol{w}^\star; \boldsymbol{v}, \boldsymbol{w}) = \langle \mathrm{D}_{\boldsymbol{u}} \mathrm{A}(\boldsymbol{w}^\star)\boldsymbol{v}, \boldsymbol{w} \rangle &= \int_{\widehat{\Omega}} \boldsymbol{\nabla}_\xi \boldsymbol{v} : \mathrm{D}_{\boldsymbol{F}} \boldsymbol{T}(\boldsymbol{F}(\boldsymbol{w}^\star)) : \boldsymbol{\nabla}_\xi \boldsymbol{w} \, \mathrm{d}\widehat{\Omega} \\ &= \int_{\widehat{\Omega}} \boldsymbol{\nabla}_\xi \boldsymbol{v} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}\big(\boldsymbol{F}(\boldsymbol{w}^\star)\big) : \boldsymbol{\nabla}_\xi \boldsymbol{w} \, \mathrm{d}\widehat{\Omega}. \end{aligned} \tag{2.8}$$

In the following, we also denote by

$$\begin{aligned} \mathrm{B} \ : \mathcal{W}(\widehat{\Omega}) &\longmapsto \mathcal{V}(\widehat{\Omega})' \\ \boldsymbol{w}^\star &\longrightarrow \mathrm{B}(\boldsymbol{w}^\star), \end{aligned}$$

the nonlinear surface traction operator such that, for any $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$,

$$\langle \mathrm{B}(\boldsymbol{w}^\star), \boldsymbol{w} \rangle = \int_{\widehat{\Gamma}_0} \widehat{f}_0 \, (\mathrm{co}\boldsymbol{F}(\boldsymbol{w}^\star) \cdot \widehat{\boldsymbol{n}}_0) \cdot \boldsymbol{w} \, \mathrm{d}\widehat{\Gamma}.$$

The Gâteaux differential of B at $\boldsymbol{w}^\star$ is denoted by $\mathrm{D}_{\boldsymbol{u}}\mathrm{B}(\boldsymbol{w}^\star) \in \mathcal{L}(\mathcal{V}(\widehat{\Omega}), \mathcal{V}(\widehat{\Omega})')$, and reads, for any $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$

$$\langle \mathrm{D}_{\boldsymbol{u}}\mathrm{B}(\boldsymbol{w}^\star)\boldsymbol{v}, \boldsymbol{w}\rangle = \int_{\widehat{\Gamma}_0} \widehat{f}_0 \left((\mathrm{D}_{\boldsymbol{u}}\mathrm{co}\boldsymbol{F}(\boldsymbol{w}^\star)\boldsymbol{v}) \cdot \widehat{\boldsymbol{n}}_0 \right) \cdot \boldsymbol{w} \, \mathrm{d}\widehat{\Gamma}, \tag{2.9}$$

where

$$\mathrm{D}_{\boldsymbol{u}}\mathrm{co}\boldsymbol{F}(\boldsymbol{w}^\star)\boldsymbol{v} = \Big((\boldsymbol{F}^{-\intercal}(\boldsymbol{w}^\star) : \nabla_{\xi}\boldsymbol{v})\boldsymbol{I} - \boldsymbol{F}^{-\intercal}(\boldsymbol{w}^\star) \cdot \nabla_{\xi}\boldsymbol{v} \Big) \cdot \mathrm{co}\boldsymbol{F}(\boldsymbol{w}^\star).$$

Note that the expression of the differential of the cofactor matrix derives from the following relations

$$\mathrm{D}_{\boldsymbol{u}} J(\boldsymbol{w}^\star)\boldsymbol{v} = \mathrm{co}\boldsymbol{F}(\boldsymbol{w}^\star) : \nabla_{\xi}\boldsymbol{v}, \qquad \mathrm{D}_{\boldsymbol{u}} \left(\boldsymbol{F}^{-\intercal}(\boldsymbol{w}^\star)\right)\boldsymbol{v} = -\boldsymbol{F}^{-\intercal}(\boldsymbol{w}^\star) \cdot \nabla_{\xi}\boldsymbol{v} \cdot \boldsymbol{F}^{-\intercal}(\boldsymbol{w}^\star).$$

Compared to (2.6), no particular property – such as symmetry – can be derived from the bilinear form associated to the differential (2.9). In order to link with the definition of the weak formulation (2.3), we introduce the difference of these two nonlinear operators, namely

$$\begin{aligned} A_b \, : \, \mathcal{W}(\widehat{\Omega}) &\longmapsto \mathcal{V}(\widehat{\Omega})' \\ \boldsymbol{w}^\star &\longrightarrow A_b(\boldsymbol{w}^\star) = \mathrm{A}(\boldsymbol{w}^\star) - \mathrm{B}(\boldsymbol{w}^\star), \end{aligned} \tag{2.10}$$

such that its associated Gâteaux differential, for any $\boldsymbol{w}^\star \in \mathcal{W}(\widehat{\Omega})$, simply reads

$$\mathrm{D}_{\boldsymbol{u}} A_b(\boldsymbol{w}^\star) = \mathrm{D}_{\boldsymbol{u}}\mathrm{A}(\boldsymbol{w}^\star) - \mathrm{D}_{\boldsymbol{u}}\mathrm{B}(\boldsymbol{w}^\star) \;\in\; \mathcal{L}(\mathcal{V}(\widehat{\Omega}), \mathcal{V}(\widehat{\Omega})').$$

Quasi-static and linearized wave problems. Let us denote by $\mathcal{H}(\widehat{\Omega}) = [\mathrm{L}^2(\widehat{\Omega})]^3$ and the associated mass-weighted inner product by

$$(\boldsymbol{v}, \boldsymbol{w})_{\mathcal{H}} = \int_{\widehat{\Omega}} \widehat{\varrho}\, \boldsymbol{v} \cdot \boldsymbol{w} \, \mathrm{d}\widehat{\Omega}, \qquad \forall \boldsymbol{v}, \boldsymbol{w} \in \mathcal{H}.$$

We introduce the nonlinear operator

$$\begin{aligned} L(t,\cdot) \, : \, \mathcal{W}(\widehat{\Omega}) &\longmapsto \mathcal{V}(\widehat{\Omega})' \\ \boldsymbol{w}^\star &\longrightarrow L(t, \boldsymbol{w}^\star), \end{aligned}$$

such that, for any $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$,

$$\langle L(t, \boldsymbol{w}^\star), \boldsymbol{w}\rangle = \int_{\widehat{\Gamma}_a} \widehat{\boldsymbol{f}}_a(\boldsymbol{w}^\star, t) \cdot \boldsymbol{w} \, \widehat{\gamma}(\boldsymbol{w}^\star) \, \mathrm{d}\widehat{\Gamma}. \tag{2.11}$$

From the definition of the operators (2.10) and (2.11), we remark that the weak formulation (2.3) can be written in the following form

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} + \langle A_b(\boldsymbol{u}), \boldsymbol{w}\rangle = \delta\langle L(t, \boldsymbol{u}), \boldsymbol{w}\rangle, \qquad \forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}).$$

Using the Ansatz (2.4), and the time invariance of $\boldsymbol{u}_0$ leads to

$$\langle A_b(\boldsymbol{u}_0), \boldsymbol{w}\rangle + \delta\Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\boldsymbol{u}_1, \boldsymbol{w})_{\mathcal{H}} + \langle \mathrm{D}_{\boldsymbol{u}} A_b(\boldsymbol{u}_0)\boldsymbol{u}_1, \boldsymbol{w}\rangle - \langle L(t, \boldsymbol{u}_0), \boldsymbol{w}\rangle \Big) = O(\delta^2).$$

Hence, satisfying the relation (2.3) up to second order terms *w.r.t* δ amounts to satisfying the following two problems:

i) Find $\boldsymbol{u}_0 \in \mathcal{W}(\widehat{\Omega})$ *the solution, for any* $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$, *of*

$$\langle A_b(\boldsymbol{u}_0), \boldsymbol{w}\rangle = 0.$$

ii) Find $(0,T) \ni t \mapsto \boldsymbol{u}_1(t) \in \mathcal{V}(\widehat{\Omega})$ *satisfying in* $\mathscr{D}'(0,T)$ *– a.k.a in the sense of distribution:*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\boldsymbol{u}_1(\cdot), \boldsymbol{w})_{\mathcal{H}} + \langle \mathrm{D}_{\boldsymbol{u}} A_b(\boldsymbol{u}_0)\boldsymbol{u}_1(\cdot), \boldsymbol{w}\rangle = \langle L(\cdot, \boldsymbol{u}_0), \boldsymbol{w}\rangle, \quad \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}) \tag{2.12}$$

with given initial conditions $\boldsymbol{u}_1(0) \in \mathcal{H}((\widehat{\Omega}))$ *and* $\frac{\mathrm{d}\boldsymbol{u}_1}{\mathrm{d}t}(0) \in \mathcal{V}(\widehat{\Omega})'$.

Approximated linearized wave problem. A major drawback of the formulation (2.12) of the wave problem is that it contains the tangent of the surface traction operator, namely $\mathrm{D}_{\boldsymbol{u}} B(\boldsymbol{u}_0)$, defined in (2.9). It is associated with a non-symmetric bilinear form, with an undefined sign and implying the surface gradient of its first argument. In addition to these unfortunate mathematical properties, one can interpret this term as the effect of the surface traction onto the wave field. In the applications of interest, where the loading area is restricted compared to the complete surface domain, we expect this term to bear little effect. Combining these arguments, we neglect the contribution of the tangent of the surface traction operator in the wave propagation problem, and we consider the following pair of problems:

i) Find $\boldsymbol{u}_0 \in \mathcal{V}(\widehat{\Omega})$ *the solution, for any* $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$*, of*

$$\langle A_b(\boldsymbol{u}_0), \boldsymbol{w} \rangle = 0. \tag{2.13}$$

ii) Find $(0,T) \ni t \mapsto \widetilde{\boldsymbol{u}}(t) \in \mathcal{V}(\widehat{\Omega})$ *satisfying in* $\mathscr{D}'(0,T)$*:*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\widetilde{\boldsymbol{u}}(\cdot), \boldsymbol{w})_{\mathcal{H}} + a(\boldsymbol{u}_0; \widetilde{\boldsymbol{u}}(\cdot), \boldsymbol{w}) = \langle L(\cdot, \boldsymbol{u}_0), \boldsymbol{w} \rangle, \quad \forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}) \tag{2.14}$$

with given initial conditions $\widetilde{\boldsymbol{u}}(0) \in \mathcal{H}(\widehat{\Omega})$ *and* $\frac{\mathrm{d}\widetilde{\boldsymbol{u}}}{\mathrm{d}t}(0) \in \mathcal{V}(\widehat{\Omega})'$*, and* $a(\boldsymbol{u}_0; \cdot, \cdot)$ *defined in* (2.7).

One important remark that can be made at this point is that the two problems previously defined can be addressed successively. From a practical viewpoint this means that one can first solve (2.13), extract the solution $\boldsymbol{u}_0$ to compute the tangent operator and the source term in (2.14) and solve the wave equation to obtain $\widetilde{\boldsymbol{u}}$.

Well-posedness of the wave propagation problem and energy relation. Since the focus of our work is on wave propagation modeling, we do not discuss the questions of the existence and uniqueness of the nonlinear static problem (2.13). Theoretical elements on this matter can be found for instance in [Marsden and Hughes, 1978; Ciarlet, 1988b], they usually entail a so-called "*poly-convexity*" property of $\mathscr{W}$. We assume in the following that there exists at least one solution of (2.13), and rather discuss the well-posedness of the wave propagation problem of interest (2.14). We rely on the standard abstract results that can be found *e.g.* in the reference textbooks [Lions, 1971; Duvaut and Lions, 1976; Dautray and Lions, 1992]. In particular, since the symmetry of the bilinear form is easily verified from its definition given in (2.7), we only need the following assumptions to apply Theorem 4.1 in [Duvaut and Lions, 1976]:

- *The righ-hand side of* (2.14) *satisfies* $L(\cdot, \boldsymbol{u}_0) \in C^1([0,T]; \mathcal{V}(\widehat{\Omega})')$,
- *The bilinear form* $a(\boldsymbol{u}_0; \cdot, \cdot)$ *defined in* (2.7) *is continuous and coercive, i.e. there exists two strictly positive constants* C_0 *and* α_0 *such that*

$$|a(\boldsymbol{u}_0; \boldsymbol{v}, \boldsymbol{w})| \leq C_0 \|\boldsymbol{v}\|_{\mathcal{V}} \|\boldsymbol{w}\|_{\mathcal{V}}, \quad a(\boldsymbol{u}_0; \boldsymbol{v}, \boldsymbol{v}) \geq \alpha_0 \|\boldsymbol{v}\|^2_{\mathcal{V}}, \quad \forall \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}). \tag{2.15}$$

Upon these assumptions the problem (2.14) admits a unique solution such that

$$\widetilde{\boldsymbol{u}} \in \mathrm{L}^2((0,T); \mathcal{V}(\widehat{\Omega})), \quad \frac{\mathrm{d}\widetilde{\boldsymbol{u}}}{\mathrm{d}t} \in \mathrm{L}^2((0,T), \mathcal{H}), \quad \frac{\mathrm{d}^2\widetilde{\boldsymbol{u}}}{\mathrm{d}t^2} \in \mathrm{L}^2((0,T), \mathcal{V}(\widehat{\Omega})')$$

with additionally, from continuous injection results,

$$\widetilde{\boldsymbol{u}} \in C^0([0,T]; \mathcal{H}(\widehat{\Omega})) \cap \mathrm{L}^\infty([0,T]; \mathcal{V}(\widehat{\Omega})), \quad \frac{\mathrm{d}\widetilde{\boldsymbol{u}}}{\mathrm{d}t} \in C^0([0,T]; \mathcal{V}(\widehat{\Omega})') \cap \mathrm{L}^\infty([0,T]; \mathcal{H}(\widehat{\Omega})).$$

Note that the coercivity assumption could be replaced by a weaker hypothesis, assuming the coercivity up to a term proportional to $\|\boldsymbol{v}\|^2_{\mathcal{H}}$. It is clear from the expression of the bilinear form that the coercivity constant depends intricately on the choice of the potential function $\mathscr{W}$ and on the linearization point $\boldsymbol{u}_0$. In particular, in Section 2.3 we provide illuminating examples of numerical solutions where, depending on the solution of the static problem, we observe a loss of coercivity leading to an ill-posed wave propagation problem. In addition to the previous well-posedness result, we recall the standard energy relation which is an important framework to devise a stable time discretization procedure for the problem (2.14). Let us denote the kinematic, potential and total energies associated with the unique solution of (2.14) by

$$\mathscr{K}(t) = \frac{1}{2}\Big\|\frac{\mathrm{d}\widetilde{\boldsymbol{u}}}{\mathrm{d}t}\Big\|^2_{\mathcal{H}}, \qquad \mathscr{P}(t) = \frac{1}{2}a(\boldsymbol{u}_0;\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}), \qquad \mathscr{E} = \mathscr{K} + \mathscr{P}.$$

Assuming a null source term in (2.14), namely $L(t;\boldsymbol{u}_0) = 0$, then, by using formally the time derivative of the solution as a test function in (2.14), we recover the following energy conservation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}(t) = 0. \tag{2.16}$$

Energy potentials. Our numerical investigations mainly concern frame-indifferent hyperelastic material with no residual stress. The material being frame-indifferent the hyperelastic potential – assumed smooth – is only dependent on the Green-Lagrange tensor $\boldsymbol{e}$

$$\mathscr{W}(\boldsymbol{F}) \equiv \mathscr{W}(\boldsymbol{e}).$$

First, we consider isotropic materials. Since the material has no residual stress, we necessarily have (see [Ciarlet, 1988b]),

$$\mathscr{W}(\boldsymbol{e}) = \frac{\lambda}{2}(\operatorname{tr}\boldsymbol{e})^2 + \mu\operatorname{tr}\boldsymbol{e}^2 + \alpha_1(\operatorname{tr}\boldsymbol{e})^3 + \alpha_2(\operatorname{tr}\boldsymbol{e})\operatorname{tr}\boldsymbol{e}^2 + \alpha_3\operatorname{tr}\boldsymbol{e}^3 + o(|\boldsymbol{e}|^3). \tag{2.17}$$

Note that in the expression above, if the material is heterogeneous, the parameter λ, μ and the $\{\alpha_i\}_{i=1}^3$ may depend on $\boldsymbol{\xi}$. The coefficients associated with third-order terms, namely $\{\alpha_i\}_{i=1}^3$, are usually referred to as the Third Order Elastic Constants (TOECs). The two first terms of this expansion correspond to the so-called Saint-Venant-Kirchhoff (SVK) potential

$$\mathscr{W}^{\mathrm{SVK}}(\boldsymbol{e}) = \frac{\lambda}{2}(\operatorname{tr}\boldsymbol{e})^2 + \mu\operatorname{tr}\boldsymbol{e}^2,$$

while, neglecting higher-order terms, yields the potential

$$\mathscr{W}^{\mathrm{3rd}}(\boldsymbol{e}) = \mathscr{W}^{\mathrm{SVK}}(\boldsymbol{e}) + \alpha_1(\operatorname{tr}\boldsymbol{e})^3 + \alpha_2(\operatorname{tr}\boldsymbol{e})\operatorname{tr}\boldsymbol{e}^2 + \alpha_3\operatorname{tr}\boldsymbol{e}^3, \tag{2.18}$$

such potential can be written using Landau's coefficients [Landau, Lifšic, and Landau, 1986]

$$A = 3\alpha_3, \quad B = \alpha_2 \quad \text{and} \quad C = 3\alpha_1.$$

Note that these two potentials are not poly-convex in general, hence stability of the nonlinear mechanical problem – and therefore of its linearized version – may be an issue. The most simple elastic potential mitigating this problem is the Compressible Neo-Hookean (CNH) potential,

$$\mathscr{W}^{\mathrm{CNH}}(\boldsymbol{e}) = \frac{\lambda}{2}\big(\sqrt{\det(\boldsymbol{I}+2\boldsymbol{e})} - 1\big)^2 + \mu\operatorname{tr}\boldsymbol{e} - \frac{\mu}{2}\log\big(\det(\boldsymbol{I}+2\boldsymbol{e})\big).$$

The potential $\mathscr{W}^{\mathrm{CNH}}$ satisfies the following expansion,

$$\mathscr{W}^{\mathrm{CNH}}(\boldsymbol{e}) = \frac{\lambda}{2}(\operatorname{tr}\boldsymbol{e})^2 + \mu\operatorname{tr}\boldsymbol{e}^2 + \frac{\lambda}{2}(\operatorname{tr}\boldsymbol{e})^3 - \lambda(\operatorname{tr}\boldsymbol{e})\operatorname{tr}\boldsymbol{e}^2 - \frac{4\mu}{3}\operatorname{tr}\boldsymbol{e}^3 + o(|\boldsymbol{e}|^3),$$

showing that, up to third order term in $|\boldsymbol{e}|^3$ the CNH constitutive law corresponds to (2.18) with $\alpha_1 = \lambda/2$, $\alpha_2 = -\lambda$ and $\alpha_3 = -4\mu/3$. The hyperelastic potential of frame-indifferent isotropic material can be re-written in terms of the first three invariants of the Cauchy-Green deformation tensor $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$,

$$I_1 = \operatorname{tr}(\boldsymbol{C}), \qquad I_2 = \frac{1}{2}\Big(\operatorname{tr}(\boldsymbol{C})^2 - \operatorname{tr}(\boldsymbol{C}^2)\Big), \qquad I_3 = \det\boldsymbol{C}.$$

It can be verified that

$$\mathscr{W}^{\mathrm{SVK}} = \frac{(\lambda+2\mu)}{8} I_1^2 - \frac{(3\lambda+2\mu)}{4} I_1 - \frac{\mu}{2} I_2 + \frac{(9\lambda+6\mu)}{8} \tag{2.19}$$

and

$$\mathscr{W}^{\mathrm{CNH}} = \frac{\lambda}{2}(\sqrt{I_3}-1)^2 + \frac{\mu}{2}(I_1 - 3 - \log(I_3)). \tag{2.20}$$

Murnaghan [Murnaghan, 1951] proposed different TOECs to characterize the second-order elastic deformations, namely l, m and n,

$$\mathscr{W}^{3\mathrm{rd}} = \mathscr{W}^{\mathrm{SVK}} + \frac{l}{24}(I_1-3)^3 + \frac{m}{12}(I_1-3)(I_1^2-3I_2) + \frac{n}{8}(I_1 - I_2 + I_3 - 1). \tag{2.21}$$

In the context of acoustoelasticity, Hughes and Kelly related these TOECs with changes in bulk wave velocities for prestressed isotropic materials and experimentally assessed them in [Hughes and Kelly, 1953]. The relations between (l, m, n) and $(\alpha_1, \alpha_2, \alpha_3)$ appearing in the expression (2.18) are given by

$$\alpha_1 = l - \frac{m}{3} + \frac{n}{6}, \quad \alpha_2 = m - \frac{n}{2}, \quad \alpha_3 = \frac{n}{3}.$$

We also investigate an orthotropic configuration where one plane of symmetry is lost, orthogonal to a principal axis a, in that case, the material is said to be transversely isotropic. To fully describe its constitutive behavior we need to add two more invariants

$$I_4 = a \cdot \boldsymbol{C} \cdot a, \qquad\qquad I_5 = a \cdot \boldsymbol{C}^2 \cdot a.$$

One example of the hyperelastic potential for transversely isotropic materials is the one from [Bonet and Burton, 1998],

$$\mathscr{W}^{\mathrm{CNH\text{-}TI}} = \mathscr{W}^{\mathrm{CNH}} + [\alpha + \beta \ln J + \gamma(I_4 - 1)](I_4 - 1) - \frac{\alpha}{2}(I_5 - 1), \tag{2.22}$$

where α, β and γ are additional parameters that can be calibrated to match observed transversely isotropic behavior.

Acoustoelastic effects. The change in the behavior of elastic waves when propagating in a deformed or prestressed medium is known as acoustoelastic effect. This phenomenon is extensively studied, see *e.g.* [Biot, 1940; Murnaghan, 1951; Hughes and Kelly, 1953]. In [Ogden, 1984], the author analyses the structure and properties of incremental elastic deformations models, including the unicity and stability aspects with respect to the constitutive law. As far as isotropic materials are considered, the behavior of elastic waves – bulk wave velocities, in particular – has been mathematically studied by the authors in [Shams, Destrade, and Ogden, 2011]. We can relate their considerations to stability properties and positivity of the bilinear form $a(\boldsymbol{u}_0; \cdot, \cdot)$. Indeed, from the expression (2.8), assuming that the solution behaves locally as a plane wave, with polarization $\boldsymbol{m}$ propagating in the direction $\boldsymbol{n}$, and assuming smoothness of the parameters and of $\boldsymbol{u}_0$, we arrive at the necessary condition for the positivity of $a(\boldsymbol{u}_0; \cdot, \cdot)$,

$$\forall\, (\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}^3 \times \mathbb{R}^3, \quad \boldsymbol{m} \otimes \boldsymbol{n} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}\big(\boldsymbol{F}(\boldsymbol{u}_0)\big) : \boldsymbol{m} \otimes \boldsymbol{n} \geq 0. \tag{2.23}$$

Note that the inequality above should be verified for each $\boldsymbol{\xi} \in \widehat{\Omega}$ since $\boldsymbol{u}_0$ is a function of $\boldsymbol{\xi}$ even when the material is homogeneous. Now, following [Shams, Destrade, and Ogden, 2011], this amounts to checking the positivity of the so-called (symmetric) Christoffel tensor $\boldsymbol{Q}(\boldsymbol{u}_0, \boldsymbol{n}) \in \mathbb{R}^{3\times 3}$, indeed,

$$\forall\, (\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}^3 \times \mathbb{R}^3, \quad \boldsymbol{m} \otimes \boldsymbol{n} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}\big(\boldsymbol{F}(\boldsymbol{u}_0)\big) : \boldsymbol{m} \otimes \boldsymbol{n} = \boldsymbol{m} \cdot \boldsymbol{Q}(\boldsymbol{u}_0, \boldsymbol{n}) \cdot \boldsymbol{m}.$$

To further analyse it, we use the Christoffel tensor $\boldsymbol{Q}(\boldsymbol{u}_0, \boldsymbol{n})$ given in [Shams, Destrade, and Ogden, 2011] in the isotropic case and for the reference configuration. It depends on $\mathscr{W}_i$ and $\mathscr{W}_{ij}$, the first and second-order derivatives of the potential with respect to the invariant I_i and I_j respectively. Note that the dependence in $\boldsymbol{u}_0$ of $\mathscr{W}_i$ and $\mathscr{W}_{ij}$ is temporarily omitted for the sake of clarity. In the case where

$$\mathscr{W}_{12} = \mathscr{W}_{22} = \mathscr{W}_{13} = \mathscr{W}_{23} = 0,$$

which is verified for the Saint-Venant-Kirchhoff and Compressible Neo-Hookean constitutive laws, the Christoffel tensor is given by

$$\begin{aligned}\boldsymbol{Q}(\boldsymbol{u}_0, \boldsymbol{n}) = {} & (4I_3^2\mathscr{W}_{33} + 2I_3\mathscr{W}_3)\hat{\boldsymbol{n}} \otimes \hat{\boldsymbol{n}} + 2(\mathscr{W}_1 + I_1\mathscr{W}_2)\boldsymbol{I} \\ & + \big(4\mathscr{W}_{11} + 2\mathscr{W}_2\big)\big(\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\big) \otimes \big(\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\big) - 2\mathscr{W}_2\Big(\|\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\|_3^2\boldsymbol{I} + \boldsymbol{B}(\boldsymbol{u}_0)\Big), \quad (2.24)\end{aligned}$$

where $\boldsymbol{B}(\boldsymbol{u}_0) = \boldsymbol{F}(\boldsymbol{u}_0)\boldsymbol{F}^{\top}(\boldsymbol{u}_0)$ and $\hat{\boldsymbol{n}} = \boldsymbol{F}^{-\top}(\boldsymbol{u}_0)\cdot\boldsymbol{n}$. The expression above takes a simple form when considering the CNH constitutive law

$$\boldsymbol{Q}^{\mathrm{CNH}}(\boldsymbol{u}_0, \boldsymbol{n}) = \big(\lambda I_3(\boldsymbol{u}_0) + \mu\big)\hat{\boldsymbol{n}} \otimes \hat{\boldsymbol{n}} + \mu\boldsymbol{I}.$$

When considering the SVK constitutive law, we get

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{SVK}}(\boldsymbol{u}_0, \boldsymbol{n}) = {} & \frac{1}{2}\big(I_1(\boldsymbol{u}_0)\lambda - 3\lambda - 2\mu + 2\mu\|\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\|^2\big)\boldsymbol{I} \\ & + \big(\lambda + \mu\big)\big(\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\big) \otimes \big(\boldsymbol{F}(\boldsymbol{u}_0)\cdot\boldsymbol{n}\big) + \mu\boldsymbol{B}(\boldsymbol{u}_0). \quad (2.25)\end{aligned}$$

Considering pure dilatation deformation, *i.e.*, $\boldsymbol{u}_0 = \boldsymbol{\xi}(d_0 - 1)$ and $\boldsymbol{F}(\boldsymbol{u}_0) = d_0\boldsymbol{I}$ with $d_0 \geq 1$ we have,

$$\boldsymbol{Q}^{\mathrm{SVK}}(\boldsymbol{\xi}(d_0 - 1), \boldsymbol{n}) = \frac{(d_0^2 - 1)}{2}(3\lambda + 2\mu)\boldsymbol{I} + d_0^2\Big[\big(\lambda + \mu\big)\boldsymbol{n} \otimes \boldsymbol{n} + \mu\boldsymbol{I}\Big]$$

and

$$\boldsymbol{Q}^{\mathrm{CNH}}(\boldsymbol{\xi}(d_0 - 1), \boldsymbol{n}) = \big(\lambda d_0^4 + \mu d_0^{-2}\big)\boldsymbol{n} \otimes \boldsymbol{n} + \mu\boldsymbol{I},$$

One can see that the eigenvalues of $\boldsymbol{Q}^{\mathrm{SVK}}(\boldsymbol{\xi}(d_0 - 1), \boldsymbol{n})$, are increasing functions of d_0, meaning that waves tend to propagate faster – at least locally. The CNH constitutive law behaves similarly when $2\lambda \geq \mu$. This behavior contradicts observed experiments [Hughes and Kelly, 1953] where velocities of the shear and pressure waves decrease for some materials – aluminum typically – that undergoes an increase in volume. In this case, proper behavior can be modeled using high-order elasticity models such as the Murnaghan's, despite some expected and observed stability issues (see the numerical experiments after). In an hypothetical case of pure stretching in one direction, we have locally

$$\boldsymbol{F}(\boldsymbol{u}_0) = \begin{pmatrix} d_0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

that is to say $\boldsymbol{u}_0 = \boldsymbol{e}_1(\boldsymbol{e}_1 \cdot \boldsymbol{\xi})(d_0 - 1)$, with $d_0 \geq 1$, we have

$$\boldsymbol{Q}^{\mathrm{SVK}}(\boldsymbol{u}_0, \boldsymbol{e}_1) = \Big(\frac{d_0^2 - 1}{2}(\lambda + 2\mu) + \mu\Big)\boldsymbol{I} + \big(d_0^2(\lambda + 2\mu) - \mu\big)\boldsymbol{e}_1 \otimes \boldsymbol{e}_1,$$

and

$$\boldsymbol{Q}^{\mathrm{CNH}}(\boldsymbol{u}_0, \boldsymbol{e}_1) = \big(\lambda + d_0^{-2}\mu\big)\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \mu\boldsymbol{I}.$$

In this particular case, the behavior of SVK and CNH constitutive laws differs. Indeed one can observe that for increasing values of d_0 the eigenvalues of $\boldsymbol{Q}^{\mathrm{SVK}}$ are all increasing while the eigenvalues of $\boldsymbol{Q}^{\mathrm{CNH}}$ associated to the eigenvector $\boldsymbol{e}_1$ is given by $\lambda + (d_0^{-2} + 1)\mu$, hence it decreases with d_0. This behavior is observed in our numerical results.

In the acoustoelasticity literature, a more general formulation comprising anisotropic behavior is usually presented. The expansion of the potential up to third-order terms are written in the indicial notation as

$$\mathscr{W}(e) \approx C_{ijkl} e_{ij} e_{kl} + C_{ijklmn} e_{ij} e_{kl} e_{mn}, \tag{2.26}$$

which is a generalization of the expansion (2.17). The fourth-order tensor C_{ijkl} is the traditional elasticity tensor, with components related to the *second order elastic constants* from Hooke's law, and the sixth-order tensor C_{ijklmn} has its components related to the *third-order elastic constants*. One may note that this description of the constitutive behavior complexifies significantly the study of poly-convexity.

Obtaining material parameters for suitable constitutive models is an essential and difficult task. Many works concern the experimental assessment of these parameters. For example, the seminal work presented in [Hughes and Kelly, 1953] gives the TOECs for polystyrene, iron, and Pyrex glass. In [Hikata et al., 1956; Bergman and Shahbender, 1958; Smith, Stern, and Stephens, 1966; Crecraft, 1967; Asay and Guenther, 1967; Hsu, 1974; Santos Jr and Bray, 2002; Muir, Michaels, and Michaels, 2009; Gandhi, Michaels, and Lee, 2012; Shi, Michaels, and Lee, 2013; Song, Huang, and Hu, 2013], the TOECs are obtained for aluminum and other metals. Transversely isotropic materials are specifically treated in [Prosser and Green, 1990; Bulletti and Capineri, 2015; Kubrusly et al., 2016; Dubuc, Ebrahimkhanlou, and Salamone, 2017b]. Most of the presented methods rely on pitch-catch time of flight measurements, which is potentially subject to poor robustness [Muir, Michaels, and Michaels, 2009]. Alternatively, more sophisticated methods based on laser vibrometer aided measurements and space-time Fourier Transform can be applied to enhance the quality of the parameteres reconstruction [Ponschab, Kiefer, and Rupitsch, 2019]. Note that, in practice, it appears that experimental set up uncertainties can have a large influence on the estimation of the parameters, as studied in [Muir, Michaels, and Michaels, 2009]. Additionally, fabrication processes may significantly modify the acoustoelastic response of the material [Hsu, 1974; Thompson, Lee, and Smith, 1986].

2.2.3 Dedicated time & space numerical schemes

We now move on to proposing a discretization strategy that addresses the problem of solving the one-way coupled problems "*i*) ⇒ *ii*)" defined in (2.13) and (2.14). Each problem, namely the nonlinear static problem and the wave propagation problem, presents specific numerical challenges – that have been tackled independently in the literature – and we propose here to aggregate the suitable numerical methods to derive a complete and efficient fully discrete algorithm.

Solving the quasi-static problem using 3D shell elements. The nonlinearities associated to problem (2.13) are handled using Newton-Raphson iterations: for a given $\boldsymbol{u}_0^k$, find $\boldsymbol{u}_0^{k+1} \in \mathcal{W}(\widehat{\Omega})$ such that

$$\langle \mathrm{D}_{\boldsymbol{u}} A_b(\boldsymbol{u}_0^k)(\boldsymbol{u}_0^{k+1} - \boldsymbol{u}_0^k), \boldsymbol{w} \rangle = \langle A_b(\boldsymbol{u}_0^k), \boldsymbol{w} \rangle, \qquad \forall \boldsymbol{w} \in \mathcal{V}(\widehat{\Omega}),$$

until convergence. In some cases, typically when the maximal amplitude of the surface traction is large, this iterative process may not converge. To increase robustness in these cases, we apply a continuation algorithm – *i.e.* successive reductions of the loading amplitude, see [Léger, Deteix, and Fortin, 2015] and references therein for more details. Writing the increment $\boldsymbol{u} = \boldsymbol{u}_0^{k+1} - \boldsymbol{u}_0^k$, we need to solve at each iteration a linearized elasticity problem in the form of

Find $\boldsymbol{u} \in \mathcal{W}(\widehat{\Omega})$ the solution, for any $\boldsymbol{w} \in \mathcal{V}(\widehat{\Omega})$ of

$$a_k(\boldsymbol{u}, \boldsymbol{w}) - b_k(\boldsymbol{u}, \boldsymbol{w}) = \ell_k(\boldsymbol{w}), \tag{2.27}$$

where $a_k(\cdot,\cdot) = a(\boldsymbol{u}_0^k; \cdot, \cdot)$ defined in (2.7), and

$$b_k(\boldsymbol{u}, \boldsymbol{w}) = \langle \mathrm{D}_{\boldsymbol{u}} \mathrm{B}(\boldsymbol{u}_0^k)\boldsymbol{u}, \boldsymbol{w} \rangle, \quad \ell_k(\boldsymbol{w}) = \langle A_b(\boldsymbol{u}_0^k), \boldsymbol{w} \rangle.$$

It is well-known that for configurations where $\widehat{\Omega}$ is a thin elongated geometry, standard finite element procedures have very poor performances, or can even fail to provide meaningful results due to locking phenomena [Bathe, 2006; Chapelle and Bathe, 2011]. To circumvent this difficulty we use 3D shell elements [Chapelle and Bathe, 2011] that enter the framework of general shell elements [Bathe and Dvorkin, 1986; Bucalem and Bathe, 1993]. These elements possess desirable properties in the scope of our work: (1) they are locking-free; (2) they rely on the full 3D variational formulations of the type of (2.27), hence, they can be applied to any type of constitutive law; (3) they rely on a quadratic expansion of the displacement unknown, involving nodes on the outer surface. This makes them readily suitable for coupling with other elements or with themselves, *i.e.* stacking of shell elements. This aspect is particularly useful to represent stratified materials.

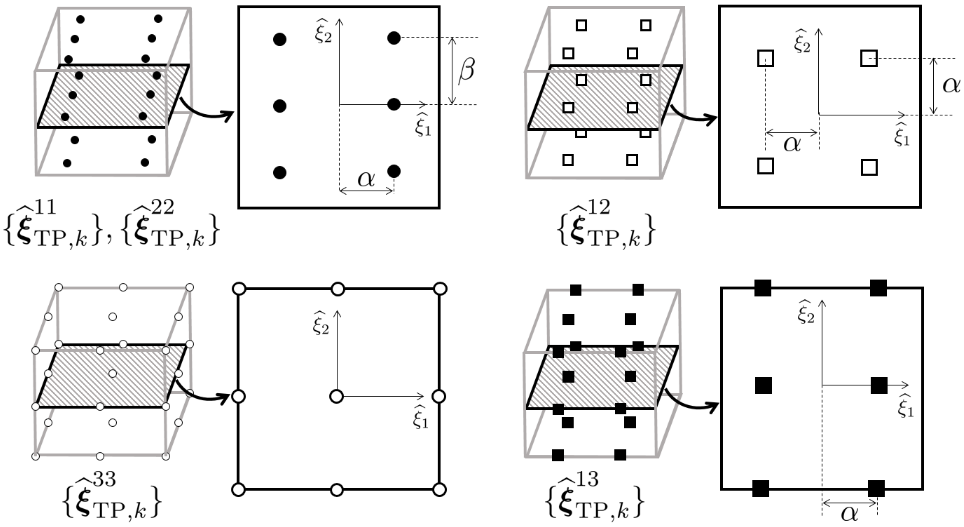


Figure 2.2 – Definition of the MITC9 tying points on the reference hexahedron – extracted from [Bathe, Iosilevich, and Chapelle, 2000] – with $\alpha = 1/\sqrt{3}$ and $\beta = \sqrt{3}/\sqrt{5}$. The tying points for the remaining tensorial components are deduced by symmetry.

Let us assume for illustration purposes that the material at hand is homogeneous and that the domain of interest can be expressed as $\widehat{\Omega} = \widehat{\omega} \times]-\eta;\eta[$, where $\widehat{\omega} \subset \mathbb{R}^2$ is the mid-surface, and $\eta > 0$ is the half-thickness of the domain. Since 3D shell elements rely on a quadratic expansion of the displacement unknown, after defining quadrangular quadratic finite elements on the mid-surface, they amount to considering standard Q^2 Lagrange elements on the full 3D domain [Chapelle and Bathe, 2011]. More precisely, let us define $\mathscr{T}_H$, a tesselation of $\widehat{\Omega}$ with mesh step H and composed of hexahedra with one element in the thickness. Denoting by $\widehat{K} = [-1;1]^3$ the reference element, any hexahedron $K \in \mathscr{T}_H$ is the deformation of the reference element through a C^1-diffeomorphism denoted by $\mathcal{F}_K$. We consider the following approximation space

$$V_H(\widehat{\Omega}) = \left\{ w_H \in C^0(\overline{\widehat{\Omega}}) \;\middle|\; \forall K \in \mathscr{T}_H,\ \exists \widehat{w} \in Q^2(\widehat{K}),\ w_H|_K = \widehat{w} \circ \mathcal{F}_K^{-1} \right\},$$

and its vectorial counterpart including the relevant boundary conditions

$$\mathcal{V}_H(\widehat{\Omega}) = \left\{ \boldsymbol{w}_H \in [V_H(\widehat{\Omega})]^3,\ \boldsymbol{w}_H|_{\Gamma_D} = 0 \right\} \subset \mathcal{V}(\widehat{\Omega}).$$

Considering a direct application of an approximation of problem (2.27) using $\mathcal{V}_H(\widehat{\Omega})$ would not alleviate the locking phenomena. To do so, one can resort to the so-called Mixed Interpolation of Tensorial Components (MITC) method [Bathe and Dvorkin, 1986; Bucalem and Bathe, 1993; Bathe, Iosilevich, and Chapelle, 2000]. In essence, this approach consists of interpolating the (discrete) Green-Lagrange tensor on so-called tying points. The location of these points is different from the Lagrange interpolation points defining the discrete displacement unknown and is also different from one component of the interpolated tensor to another. In our work, we use the MITC9 point distribution since we consider quadratic elements – see Figure 2.2 for the positioning of these points. Let $i, j \in \{1, 2, 3\}$ be indexes of the components of the Green-Lagrange tensor, and n_{TP}^{ij} the total number of the tying points denoted by

$$\{\widehat{\boldsymbol{\xi}}_{\mathrm{TP},k}^{ij}\}_{k=1}^{n_{\mathrm{TP}}^{ij}} \subset \widehat{K},$$

and defined on the reference hexahedron. The MITC process applied to a component e_{ij} of the Green-Lagrange tensor is denoted by $\mathscr{I}_{\mathrm{TP}}(e_{ij})$ and reads

$$\forall K \in \mathscr{T}_H, \quad \forall \boldsymbol{\xi} \in K, \quad \mathscr{I}_{\mathrm{TP}}(e_{ij})|_K(\boldsymbol{\xi}) = \sum_{k=1}^{n_{\mathrm{TP}}^{ij}} \widehat{\lambda}_{\mathrm{TP},k}\big(\mathcal{F}_K^{-1}(\boldsymbol{\xi})\big)\, e_{ij}(\mathcal{F}_K(\widehat{\boldsymbol{\xi}}_{\mathrm{TP},k}^{ij})),$$

where $\{\widehat{\lambda}_{\mathrm{TP},i}\}_{i=1}^{n_{\mathrm{TP}}}$ are the Lagrange shape functions associated with the tying points. In the following, we denote by $\overline{\boldsymbol{e}} = \mathscr{I}_{\mathrm{TP}}(\boldsymbol{e})$ the MITC process applied to tensors, in the sense that each component is interpolated on the tying points, namely

$$\forall i, j \in \{1, 2, 3\}, \quad \overline{e}_{ij} = \mathscr{I}_{\mathrm{TP}}(e)_{ij} = \mathscr{I}_{\mathrm{TP}}(e_{ij}).$$

In the same fashion, considering a discrete linearization point $\boldsymbol{w}_H^\star \in \mathcal{V}_H(\widehat{\Omega})$, we introduce, for any $\boldsymbol{v}_H, \boldsymbol{w}_H \in \mathcal{V}_H(\widehat{\Omega})$,

$$\mathrm{D}_{\boldsymbol{u}}\overline{\boldsymbol{e}}(\boldsymbol{w}_H^\star) \cdot \boldsymbol{w}_H = \mathscr{I}_{\mathrm{TP}}(\mathrm{D}_{\boldsymbol{u}}\boldsymbol{e}(\boldsymbol{w}_H^\star) \cdot \boldsymbol{w}_H), \quad \mathrm{D}_{\boldsymbol{u}}^2\overline{\boldsymbol{e}}(\boldsymbol{v}_H, \boldsymbol{w}_H) = \mathscr{I}_{\mathrm{TP}}(\mathrm{D}_{\boldsymbol{u}}^2\boldsymbol{e}(\boldsymbol{v}_H, \boldsymbol{w}_H)),$$

the interpolation on the tying points of the first and second differential of the Green-Lagrange tensor. With these notations, the discrete problem associated with (2.27) reads:

Find $\boldsymbol{u}_H \in \mathcal{V}_H(\widehat{\Omega})$ *the solution, for any* $\boldsymbol{w}_H \in \mathcal{V}_H(\widehat{\Omega})$*, of*

$$\overline{a}_k(\boldsymbol{u}_H, \boldsymbol{w}_H) - b_k(\boldsymbol{u}_H, \boldsymbol{w}_H) = \ell_k(\boldsymbol{w}_H),$$

with

$$\overline{a}_k(\boldsymbol{u}_H, \boldsymbol{w}_H) = \int_{\widehat{\Omega}} \mathrm{D}_{\boldsymbol{u}}\overline{\boldsymbol{e}}(\boldsymbol{u}_{0,H}^k)\boldsymbol{u}_H : \mathrm{D}_{\boldsymbol{e}}\boldsymbol{\Sigma}(\overline{\boldsymbol{e}}(\boldsymbol{u}_{0,H}^k)) : \mathrm{D}_{\boldsymbol{u}}\overline{\boldsymbol{e}}(\boldsymbol{u}_{0,H}^k)\boldsymbol{w}_H \,\mathrm{d}\widehat{\Omega} \\ + \int_{\widehat{\Omega}} \boldsymbol{\Sigma}(\overline{\boldsymbol{e}}(\boldsymbol{u}_{0,H}^k)) : \mathrm{D}_{\boldsymbol{u}}^2\overline{\boldsymbol{e}}(\boldsymbol{u}_H, \boldsymbol{w}_H) \,\mathrm{d}\widehat{\Omega}.$$

In the previous expression of the discrete problem, $\boldsymbol{u}_{0,H}^k \in \mathcal{V}_H(\widehat{\Omega})$ represents the discrete displacement field obtained from the prior Newton-Raphson iteration, and is also used as a linearization point for the bilinear and linear forms $b_k(\cdot, \cdot)$ and $\ell_k(\cdot)$.

Spectral elements and explicit leap-frog scheme for the wave propagation problem. We now provide key elements for efficient space and time discretization of the wave problem (2.14) relying on the combination of spectral finite elements, the mass-lumping technique and the explicit leapfrog scheme. Readers may refer to [Komatitsch et al., 1999; Cohen, 2002; Joly, 2007] for a more detailed description of these subjects.

To start with, let us denote by $\boldsymbol{p} = (p_1, p_2, p_3)^\intercal \in \mathbb{N}^{*3}$ an integer triplet representing orders of polynomial on every direction of the reference hexahedron. We define the space of polynomial with "*anisotropic orders*" as

$$Q^{\boldsymbol{p}}(\widehat{K}) = \mathrm{span}\left\{ \widehat{s}_1^i \widehat{s}_2^j \widehat{s}_3^k, \quad \forall (i, j, k) \in \mathbb{N} \;:\; i \leq p_1,\; j \leq p_2,\; k \leq p_3 \right\},$$

where $\widehat{\boldsymbol{s}} = (\widehat{s}_1, \widehat{s}_2, \widehat{s}_3)^\intercal$ are the local coordinates on the reference element. From this local space, we can build the global conforming approximation space that supports the semi-discretization procedure of the wave problem (2.14). To do so, we consider $\mathscr{T}_h$ a hexahedral mesh of $\widehat{\Omega}$ – different from $\mathscr{T}_H$ – with mesh step h and such that any hexahedron $K \in \mathscr{T}_h$ is the deformation of the reference element through a C^1-diffeomorphism denoted by $\mathcal{G}_K$. The scalar-valued approximation space reads

$$V_h(\widehat{\Omega}) = \left\{ w_h \in C^0(\overline{\widehat{\Omega}}) \,\middle|\, \forall K \in \mathscr{T}_h,\ \exists \widehat{w} \in Q^{\boldsymbol{p}}(\widehat{K}),\ w_h|_K = \widehat{w} \circ \mathcal{G}_K^{-1} \right\},$$

and, as previously, its vectorial counterpart including relevant boundary conditions is

$$\mathcal{V}_h(\widehat{\Omega}) = \left\{ \boldsymbol{w}_h \in [V_h(\widehat{\Omega})]^3,\ \boldsymbol{w}_h|_{\Gamma_D} = 0 \right\} \subset \mathcal{V}(\widehat{\Omega}).$$

The local Lagrange (polynomial) basis functions are defined on the reference element from Gauss-Lobatto points. This is a key point in order to avoid Runge's phenomenon, and to render spectral-like convergence as the order of the local polynomials increases. In practice, these points are set on the reference edge $\widehat{e} = [-1; 1]$ for any order of approximation (see *e.g.* [Joly, 2007] for an explicit definition and related examples of these points), and the location of points are deduced in the reference cube by a simple extrusion process.

In addition to this specific choice of approximation space, we use a quadrature formula to numerically compute the integrals involved in the bilinear forms $(\cdot, \cdot)_{\mathcal{H}}$ and $a(\boldsymbol{u}_0; \cdot, \cdot)$ appearing in (2.14). The formula is defined on the reference hexahedron. We denote the $\widehat{n}_q$ associated weights and points respectively by

$$\left(\{\widehat{\omega}_i\}_{i=1}^{\widehat{n}_q}, \{\widehat{\boldsymbol{s}}_i\}_{i=1}^{\widehat{n}_q} \right).$$

For any function f at least continuous over each element in $\mathscr{T}_h$, the numerical integration process reads

$$\int_\Omega f \, \mathrm{d}\Omega \;\approx\; \sum_{K \in \mathscr{T}_h} \sum_{k=1}^{\widehat{n}_q} \widehat{\omega}_k \, f(\mathcal{G}_K(\widehat{\boldsymbol{s}}_k)) \left| \det \boldsymbol{\nabla}_{\widehat{\boldsymbol{s}}} \mathcal{G}_K(\widehat{\boldsymbol{s}}_k) \right|.$$

The quadrature points are identical to the Gauss-Lobatto points, and the quadrature weights are the strictly positive Gauss-Lobatto weights. Matching the quadrature points with the points associated with the local Lagrange basis functions is referred to as "mass lumping" [Komatitsch et al., 1999; Cohen, 2002; Joly, 2007]. It leads to a diagonal approximated mass matrix. It has been the subject of particular interest in the context of finite element methods applied to time domain wave propagation problems since it enables fully explicit time discretization schemes. To conserve stability and accuracy, specific conditions must be enforced on the quadrature points and weights. For quadrangles and hexahedra, these conditions are met at any order of approximations by the Gauss-Lobatto rules [Joly, 2007; Duruflé, Grob, and Joly, 2009]. Following this numerical integration procedure, we denote by

$$(\cdot, \cdot)_{\mathcal{H}} \approx (\cdot, \cdot)_h, \qquad a(\boldsymbol{u}_0; \cdot, \cdot) \approx a_h(\boldsymbol{u}_0; \cdot, \cdot), \tag{2.28}$$

the approximated bilinear forms so that the semi-discrete formulation of the wave problem (2.14) reads

For any time $t \in [0; T]$, find $\widetilde{\boldsymbol{u}}_h(t) \in \mathcal{V}_h(\widehat{\Omega})$ the solution, for any $\boldsymbol{w}_h \in \mathcal{V}_h(\widehat{\Omega})$, of

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} (\widetilde{\boldsymbol{u}}_h, \boldsymbol{w}_h)_h + a_h(\boldsymbol{u}_{0,H}; \widetilde{\boldsymbol{u}}_h, \boldsymbol{w}_h) = \langle L(t, \boldsymbol{u}_{0,H}), \boldsymbol{w}_h \rangle. \tag{2.29}$$

with given initial conditions $\widetilde{\boldsymbol{u}}_h|_{t=0}$ and $\partial_t \widetilde{\boldsymbol{u}}_h|_{t=0}$.

Note that in (2.29) $\boldsymbol{u}_{0,H}$ is the discrete solution – after convergence of the Newton-Raphson iterations – of problem (2.13). Moving on to the time discretization of the wave propagation problem, we define $\Delta t > 0$ as the time step. We denote by $t^n = n\Delta t$, and we assume that there exists $N \in \mathbb{N}^*$ such that $T = N\Delta t$. To obtain the fully discrete scheme associated with (2.29), we apply an explicit second-order time scheme centered at t^n, *a.k.a.* the leapfrog scheme [Cohen, 2002; Joly, 2003]. Namely, denoting by $\widetilde{u}_h^n$ the approximation of $\widetilde{u}_h(t^n)$ solution of (2.29), we define the following fully discrete scheme:

For any $n \in [\![1;N]\!]$, *find* $\widetilde{\boldsymbol{u}}_h^{n+1} \in \mathcal{V}_h(\widehat{\Omega})$ *the solution, for any* $\boldsymbol{w}_h \in \mathcal{V}_h(\widehat{\Omega})$, *of*

$$\left(\frac{\widetilde{\boldsymbol{u}}_h^{n+1} - 2\widetilde{\boldsymbol{u}}_h^{n} + \widetilde{\boldsymbol{u}}_h^{n-1}}{\Delta t^2}, \boldsymbol{w}_h\right)_h + a_h(\boldsymbol{u}_{0,H}; \widetilde{\boldsymbol{u}}_h^{n}, \boldsymbol{w}_h) = \langle L(t^n, \boldsymbol{u}_{0,H}), \boldsymbol{w}_h \rangle. \tag{2.30}$$

with given initial steps $\widetilde{\boldsymbol{u}}_h^0$ *and* $\widetilde{\boldsymbol{u}}_h^1$.

From this expression of the fully discrete problem, it is possible to recover a discrete energy conservation property, which is the exact translation of (2.16). Denoting by $\|\cdot\|_h$ the norm associated with the discrete inner-product $(\cdot,\cdot)_h$, we define the kinematic and potential discrete energies associated with the unique solution of (2.30) as

$$\mathscr{K}_h^{n+1/2} = \frac{1}{2}\left\|\frac{\widetilde{\boldsymbol{u}}_h^{n+1} - \widetilde{\boldsymbol{u}}_h^{n}}{\Delta t}\right\|_h^2, \qquad \mathscr{P}_h^{n+1/2} = \frac{1}{2} a_h(\boldsymbol{u}_{0,H}; \widetilde{\boldsymbol{u}}_h^{n+1}, \widetilde{\boldsymbol{u}}_h^{n}),$$

and the total energy as $\mathscr{E}_h^{n+1/2} = \mathscr{K}_h^{n+1/2} + \mathscr{P}_h^{n+1/2}$. Assuming a null source term in (2.30), namely $L(t^n, \boldsymbol{u}_{0,H}) = 0$, then, taking $\boldsymbol{w}_h = \frac{1}{2\Delta t}(\widetilde{\boldsymbol{u}}_h^{n+1} - \widetilde{\boldsymbol{u}}_h^{n-1})$ as test function and using the symmetry of the stiffness bilinear form, we have the following energy conservation

$$\frac{\mathscr{E}_h^{n+1/2} - \mathscr{E}_h^{n-1/2}}{\Delta t} = 0. \tag{2.31}$$

This conservation property is critical since it leads to the conditions – on the stiffness bilinear form and on the time step – for the fully discrete (2.30) to be stable. Indeed, note that the total energy can be re-written in the following form

$$\begin{aligned}\mathscr{E}_h^{n+1/2} = \frac{1}{2}\left\|\frac{\widetilde{\boldsymbol{u}}_h^{n+1} - \widetilde{\boldsymbol{u}}_h^{n}}{\Delta t}\right\|_h^2 - \frac{\Delta t^2}{8} a_h\left(\boldsymbol{u}_{0,H}; \frac{\widetilde{\boldsymbol{u}}_h^{n+1} - \widetilde{\boldsymbol{u}}_h^{n}}{\Delta t}, \frac{\widetilde{\boldsymbol{u}}_h^{n+1} - \widetilde{\boldsymbol{u}}_h^{n}}{\Delta t}\right) \\ + a_h\left(\boldsymbol{u}_{0,H}; \frac{\widetilde{\boldsymbol{u}}_h^{n+1} + \widetilde{\boldsymbol{u}}_h^{n}}{2}, \frac{\widetilde{\boldsymbol{u}}_h^{n+1} + \widetilde{\boldsymbol{u}}_h^{n}}{2}\right).\end{aligned}$$

Hence, upon the assumption that

- *The stiffness bilinear form in* (2.30) *satisfies the coercivity property* (2.15),
- *The time step satisfies*

$$\Delta t \leq 2\left(\sup_{\boldsymbol{w}_h \in \mathcal{V}_h(\widehat{\Omega})} \frac{a_h(\boldsymbol{u}_{0,H}; \boldsymbol{w}_h, \boldsymbol{w}_h)}{\|\boldsymbol{w}_h\|_h^2}\right)^{-\frac{1}{2}}, \tag{2.32}$$

we can ensure that the total energy is a positive functional and that it actually corresponds to an equivalent norm on the discrete solution. Following arguments provided *e.g.* in [Joly, 2003; Chabassier and Imperiale, 2013; Chabassier and Imperiale, 2021] one can obtain the following stability result: there exists a constant $C > 0$ that depends on the initial steps $\widetilde{\boldsymbol{u}}_h^0$ and $\widetilde{\boldsymbol{u}}_h^1$, the source term $L(t^n; \boldsymbol{u}_{0,H})$ and the maximal time T, such that the solution of (2.30) satisfies

$$\|\widetilde{\boldsymbol{u}}_h^{n+1}\|_h \leq C(T, L, \widetilde{\boldsymbol{u}}_h^0, \widetilde{\boldsymbol{u}}_h^1), \qquad \forall n \in [\![1;N]\!].$$

Remarks on practical considerations. Prior to giving substantiated numerical results proving the soundness of the presented modeling strategy, let us gather some important remarks.

First, in the definition of the discrete space $\mathcal{V}_h(\widehat{\Omega})$ we consider anisotropic orders of approximation. This aspect appears to be particularly convenient for thin and potentially stratified materials. Indeed, it allows for high-order elements in the mid-surface of the domain – with low numerical dispersion or numerical anisotropy [Cohen, 2002; Basabe, 2007; Seriani and Oliveira, 2008] – while maintaining a reasonable number of Degrees of Freedom (DoF) by using lower order in the thickness of the material.

Material	Density($\mathrm{g/cm^3}$)	λ(GPa)	μ(GPa)	l(GPa)	m(GPa)	n(GPa)
Aluminum	2.704	54.308	27.174	-181.0	-289.0	-336.0
Steel	7.850	111	82.1	-459.0	-461.0	-358.0

Table 2.1 – Material parameters for the metals used in this work.

Second, the use of approximated numerical integration as denoted in (2.28) implies in particular a lumped, *i.e.* diagonal, mass matrix. This is an important aspect in the context of high-frequency wave propagation since the leapfrog scheme (2.30) thus becomes fully explicit. At each time step, only a diagonal matrix needs to be inverted. Additionally, using numerical integration with points matching the ones defining the local Lagrange polynomial leads to very sparse discrete gradient computations. Hence, the approximation $a(\boldsymbol{u}_0;\cdot,\cdot) \approx a_h(\boldsymbol{u}_0;\cdot,\cdot)$ also implies significant performance gains – see *e.g.* [Alexandre Imperiale and Demaldent, 2019] for more details on these practical aspects.

Moreover, in practice, the mesh step for the locking-free shell elements depends mostly on the geometrical characteristics of the domain and the material variations – typically one element per material layer in the case of stratified materials. For the wave problem, to these constraints is added the constraint of having enough discretization points per wavelength. With these qualitative meshing directives in mind, in the context of high-frequency wave propagation, one can expect that the $\mathscr{T}_h$ is significantly more refined than $\mathscr{T}_H$, *i.e.* $h \ll H$.

In light of the previous remark, in order to solve the fully discrete one-way coupled problems, one needs to resort to a mean of interpolating $\boldsymbol{u}_{0,H} \in \mathcal{V}_H(\widehat{\Omega})$ onto $\mathcal{V}_h(\widehat{\Omega})$. Let us denote by $\mathscr{I}_h(\cdot)$ the nodal interpolator from $\mathcal{V}_H(\widehat{\Omega})$ to $\mathcal{V}_h(\widehat{\Omega})$, without entering into the details of the exact expression of this interpolator, let us simply remark that the stiffness bilinear form actually reads

$$a_h(\boldsymbol{u}_{0,H};\cdot,\cdot) = a_h(\mathscr{I}_h(\boldsymbol{u}_{0,H});\cdot,\cdot).$$

Finally, the condition (2.32) on the time step is often referred to as the Courant – Friedrichs – Lewy (CFL) condition of the fully discrete scheme. Note that in the particular cases of interest in our work, this CFL condition naturally depends on the linearization point, *i.e.* the displacement field $\boldsymbol{u}_{0,H}$. Furthermore, in terms of stability, in addition to this CFL condition, the coercivity assumption is capital – a condition that is also required for the well-posedness of the continuous problem.

2.3 Numerical results and experimental validation

In this section, we apply our model and numerical methods for different configurations. First, we assess the stability issue caused by a potential failure to ensure coercivity. Second, we apply our methods to realistic cases in the SHM context. At last, we present how we validate our model with experimental data available in the literature. The computations were done using a laptop workstation equipped with a *Intel i9-9880H* CPU and 32GB of RAM. The wave propagation simulations are performed within the SFEM kernel of the CIVA platform [1] developed at CEA.

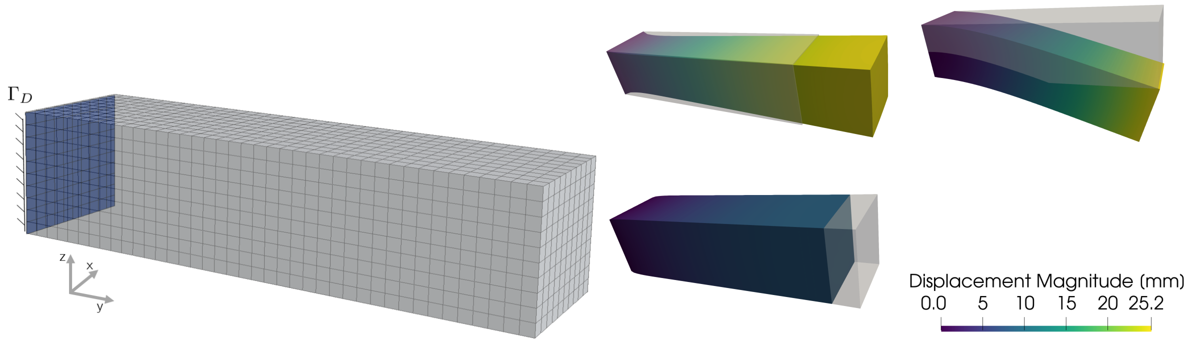
2.3.1 Illustration of potential stability issues of the fully discrete scheme

As discussed previously, the tangent stiffness operator in (2.14) must satisfy the coercivity assumption (2.15), otherwise the well-posedness of the problem cannot be guaranteed. This assumption cannot be satisfied for any deformation or constitutive law. In this section, we aim to illustrate this problem in a test model. At a discrete level, coercivity is equivalent to positiveness of the operator, which can be assessed by computing its eigenvalues. In the zero-deformation case, the presented model equates to linear elasticity (Hooke's law), and $a(0;\cdot,\cdot)$ is known to be coercive [Ciarlet, 1988b]. Here, we assess

---

[1] `https://www.extende.com`

stability for deformed cases by analyzing the evolution of the system energy and the eigenvalues of the local stiffness matrices.



(a) Mesh and boundary conditions for the test case for model stability.

(b) Deformations caused by body forces in three different directions.

Figure 2.3 – Configuration and deformations for illustrating the potential stability issues.

If instability would appear by the deformation, it could be interpreted as a shift of tangent stiffness operator eigenvalues towards negative values. In this case, the negative eigenvalues that are related to the instability of the scheme are expected to be among those of the smallest amplitude, *i.e.* low spatial frequency. Computing the lowest eigenvalues for the stiffness operator may have prohibitive computational cost due to the size of the problems at hand. To circumvent this computational difficulty, we analyze instead the eigenvalues of the local stiffness matrix related to each finite element. Although negative eigenvalues for the local stiffness matrices do not necessarily mean negative eigenvalues for the global matrix, it does give an insight into the problematic regions and related severity. In addition to the study of the lowest eigenvalue, we analyze the evolution in time of the total system's kinetic and potential energy for different initial conditions, showing the nature of such stability issues.

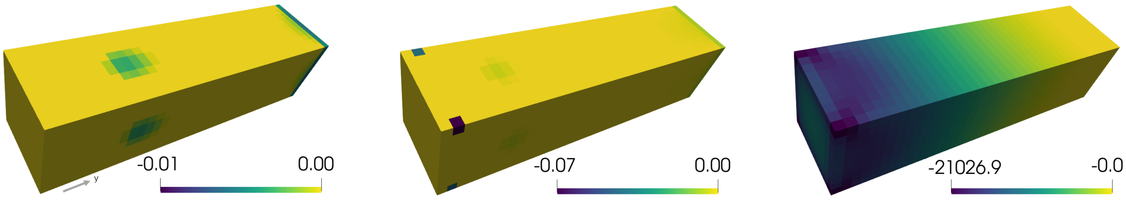
The model configuration used is an aluminum cuboid of 25mm × 25mm × 100mm depicted in Figure 2.3a. Dirichlet boundary condition is considered at x=0. The material parameters are present in Table 2.1. We analyze different constitutive laws and types of mechanical solicitations. The analysis is done using the aforementioned SVK, CNH, and Murnaghan's constitutive laws. For each case, three different types of mechanical solicitations as body forces will be considered:

- **Stretching**: $300\text{N}/\text{mm}^3$ (Y direction),
- **Bending**: $7.5\text{N}/\text{mm}^3$ (-Z direction),
- **Compresing**: $135\text{N}/\text{mm}^3$ (-Y direction).

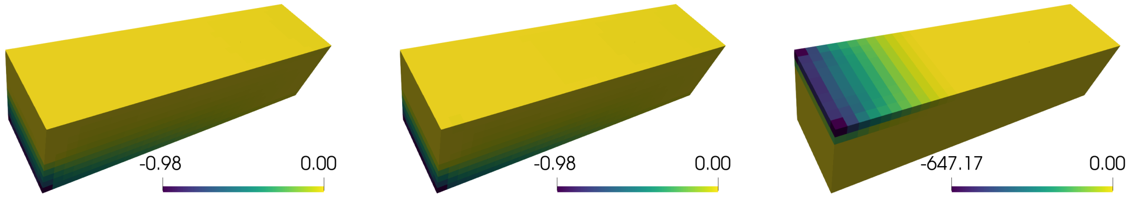
The associated deformations are illustrated in Figure 2.3b for the CNH case. The SVK law has slightly different deformation with the same nature and amplitudes. Computing the quasi-static deformation using Murnaghan's constitutive law did not converge due to its particularly ill-posed tangent operator.

After computing the deformations caused by these forces for each constitutive law, the displacement field is given as input to our wave propagation solver. The same constitutive law is used for the quasi-static problem and the wave propagation problem, except in Murnaghan case. For the Murnaghan case, we compute the deformation with CNH instead. The deformation field is interpolated from the quasi-static problem mesh to the wave propagation problem mesh. After discretizing the functional space using spectral finite elements, we assemble the local tangent stiffness matrices and compute their eigenvalues. The minimum eigenvalues, for each element, are shown in Figure 2.4 for different types of deformations and constitutive laws.

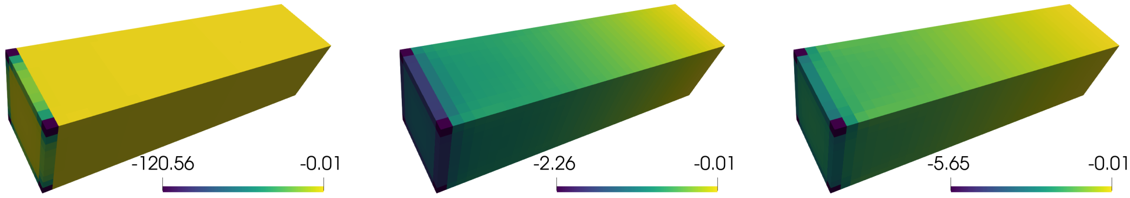
The CNH constitutive law presents less severity, *i.e.* lowest amplitude negative eigenvalues. The SVK and CNH cases are more unstable for compressed cases – cases associated with buckling phenomena. The Murnaghan's law presents the highest negative eigenvalues with critical behavior in the stretched case. The stability issue can be also assessed by observing the evolution of the system's potential and



(a) Stretched SVK. (b) Stretched CNH. (c) Stretched MUR.



(d) Bended SVK. (e) Bended CNH. (f) Bended MUR.



(g) Compressed SVK. (h) Compressed CNH. (i) Compressed MUR.

Figure 2.4 – Minimum (local) tangent stiffness matrix eigenvalues for different types of deformation and constitutive laws.

kinetic energy. We run the wave propagation problem, without source-term, from initial conditions with specific spatial frequency signature, namely for $n = 1, 2, 3$

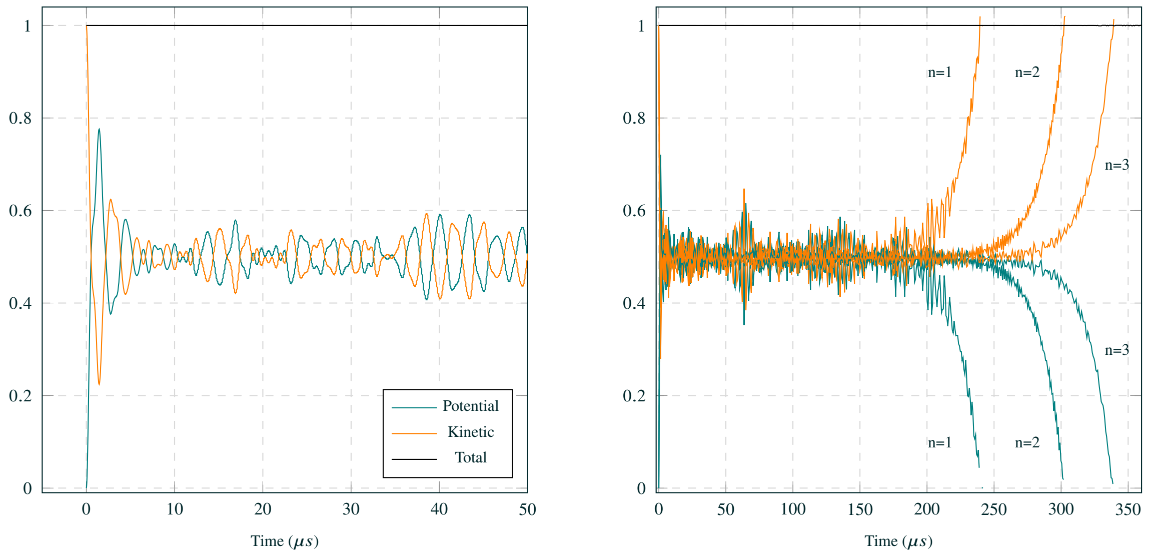
$$\begin{aligned}\widetilde{\boldsymbol{u}}_h^0 &= -\left[\sin\left(\frac{25-x}{25}2\pi n\right)\sin\left(\frac{25-y}{25}2\pi n\right)\sin\left(\frac{25-z}{25}2\pi n\right)\right](\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z),\\ \widetilde{\boldsymbol{u}}_h^1 &= -\widetilde{\boldsymbol{u}}_h^0.\end{aligned}$$

It allows us to solicit the low-frequency modes and assess the instability, relating the frequency content to the issue. Running simulations with these initial conditions, we monitor the total, kinetic and potential energies as shown in Figure 2.5.

As the total energy is conserved, if both kinetic and potential energies are positive, the system's energy is bounded. However, if any of the energies become negative, the solution may diverge. For the stable cases, the energy transfer between kinetic and potential is kept throughout the simulation. For the unstable cases, as soon as the potential energy presents negative values, the system starts to diverge. The excitation of lower frequencies presented instability in fewer time steps as they excite lower frequency modes, *i.e.* modes related to lower eigenvalues. For the presented forces, the CNH case is unstable only under compression, SVK is unstable for bending and compression and Murnaghan's is unstable for all presented deformations. In general, the stability issues are not only dependent of buckling phenomena but also on the choice of constitutive law.

2.3.2 Illustrations related to *Structural Health Monitoring*

We now illustrate the generic aspect of our work by applying it to different cases. The first case is a steel pipe subjected to the 4-point bending test, inspired by an experiment done in [Tschöke et al., 2017]. We simulate ultrasonic propagation with different deformation intensities and compare the acoustoelastic effects in the signal with ones caused by a crack through the thickness of the pipe. In the second case, we



(a) Stable behavior for a stable case (Stretched; CNH).

(b) Unstable case (Bended; CNH) for different initial conditions.

Figure 2.5 – Evolution of total energies ($a.u.$) for the wave propagation problem.

model ultrasonic wave propagation in a 16-layer stratified composite plate under torsion forces. For such modeling, we use the transversely isotropic law (2.22) for each of the 16 layers associated with a principal fiber direction. These examples illustrate the application of our method to realistic and complex loading configurations, and the simulation results could be used to assess the efficiency of a SHM system.

Illustration on a pipe subject to a 4-point bending test. We present here a case of wave propagation in a steel pipe subjected to a 4-point bending test. This configuration is usually used to perform fatigue tests in welded pipes. The experimental setup is inspired by the experiment done in [Tschöke et al., 2017]. We model and simulate the wave propagation in the pipe subjected to such bending forces with different intensity levels. We use Murnaghan's constitutive law with material parameters for steel (Table 2.1). For these levels of deformation and material parameters the wave propagation problem does not present stability issues. The surface force is applied at two regions close to the pipe extremities. The configuration is depicted in Figure 2.6, with a 2.94m long, 8mm thick pipe and an inner radius of 102.65mm. The resulting deformation for a total force of 220kN is shown in Figure 2.7.

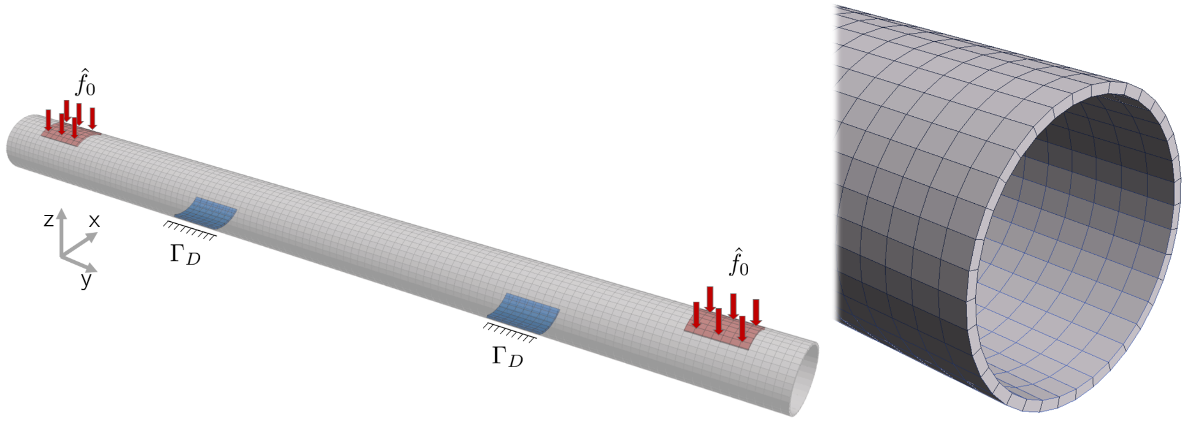


Figure 2.6 – Numerical setup and mesh for the quasi-static problem.

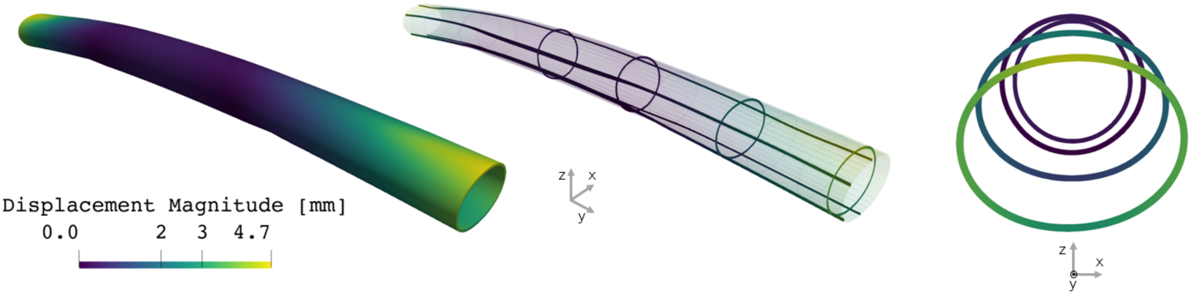


Figure 2.7 – Deformation of the pipe for a 220kN loading. (Deformation scaled 20x for visualization).

Given the different deformations as input, we interpolate them in the wave propagation problem mesh and compute the wavefield generated by a unique source term. The wave problem mesh, source location and measurement point are depicted in Figure 2.8, as well as the position of a potential crack. The ultrasonic waves are excited close to one edge of the pipe and measured at the other edge. For a case with no cracks, we simulate the wavefield and extract the signal from the modeled transducer at the measurement point and plot in Figure 2.9 for different levels of mechanical loadings. Snapshots of the wavefield are shown in Figure 2.10.

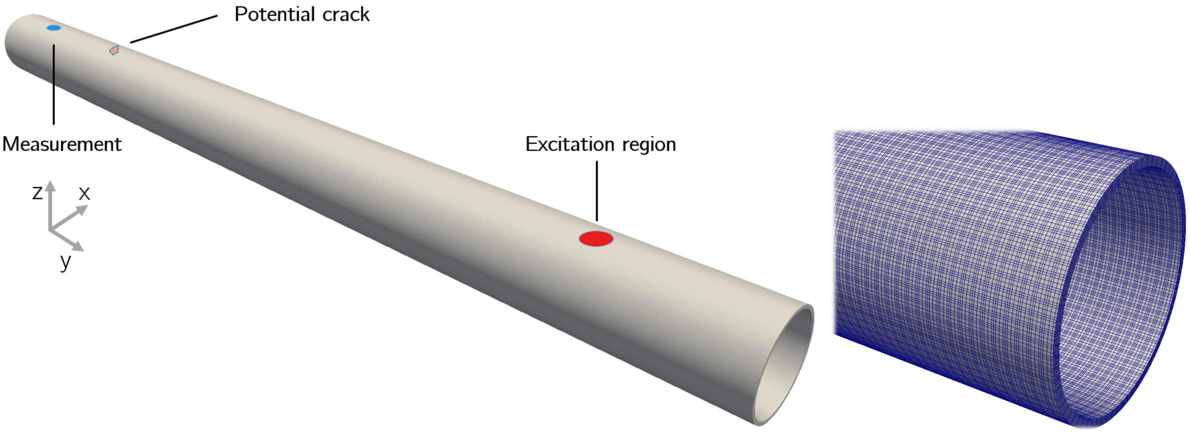


Figure 2.8 – Transducer, potential crack and measurement point positions for the wave propagation problem. On the right-hand side, the mesh for the wave propagation problem characterized by the Gauss-Lobatto points.

In Figure 2.9, we can see that the effect of bending forces results in a change in time of flight, decreasing the wave velocity, and also a change in amplitude of the signals with higher forces.

To evaluate how the effect of mechanical loading on the signal compares with the one caused by a crack, we set up and simulate three numerical configurations:

- S: non deformed pipe (considered as baseline; no defects)
  Simulation is done considering no deformation with a pristine pipe,
- S_p: deformed pipe (without defect)
  Simulation is done considering the pipe under bending forces of 220kN,
- S_d: pipe with defect (without deformation)
  Simulation is done considering the undeformed pipe with a crack of 19mm wide through half the thickness.

From each case, we extract the signal of a modeled transducer at the measurement point. The defect through half the thickness was modeled by disconnecting elements. By subtracting the signals from the baseline, we can isolate the individual contribution to the signal of the mechanical loading and the crack.

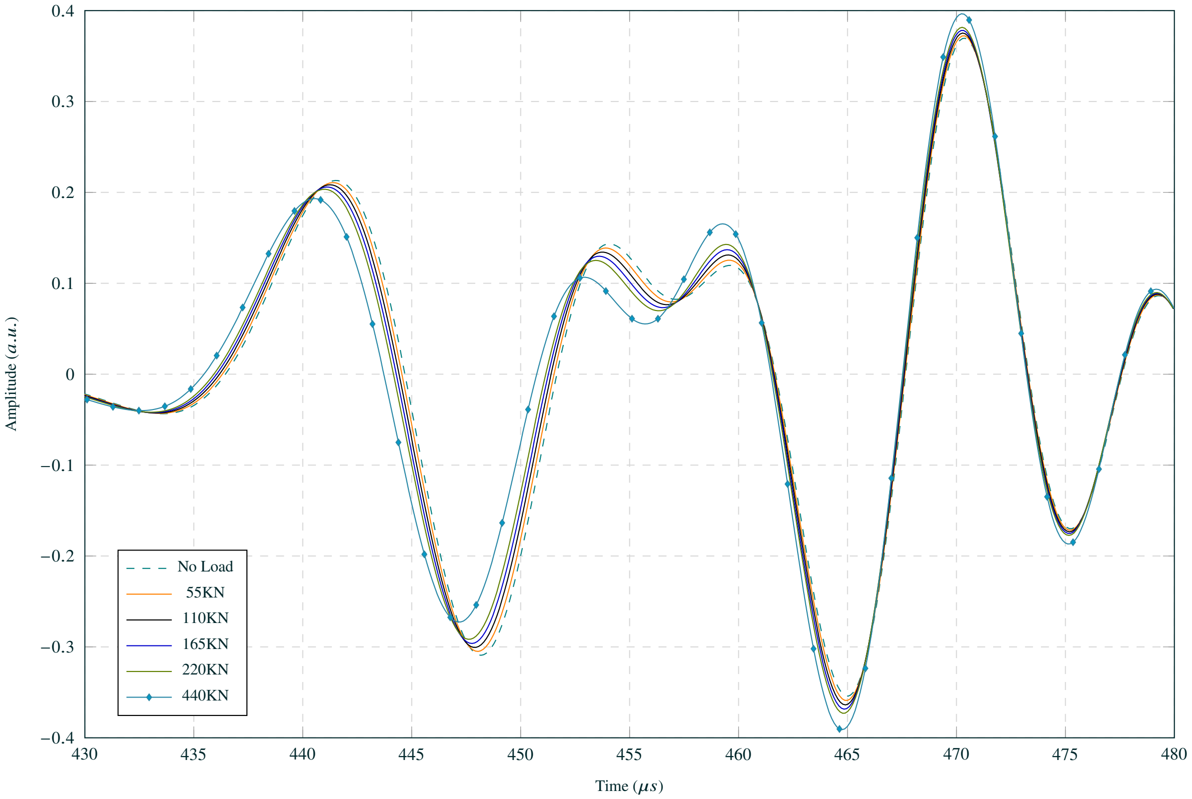


Figure 2.9 – Transducer signal at the measurement point for different bending forces.

Material	Density(g/cm^3)	C_{11}	C_{22}	C_{12}	C_{23}	C_{66}
CFRP	1.6	155.17	13.22	5.34	6.98	5.77

Table 2.2 – Material parameters for CFRP. [Lonné, 2003]

The baseline signal S is subtracted from S_p and S_d and the results are plotted in Figure 2.11.

Regarding performances, the quasi-static problem is 148k DoFs and takes 3 minutes to run 5 Newton-Raphson iterations while using 2.2GB of RAM. The wave propagation problem requires around 3.86M mesh points to ensure at least two elements of fourth-order per wavelength, resulting in 11.58M DoFs. The timestep is $\Delta t = 0.11\mu$s to comply with the CFL condition (2.32). Using an efficient implementation of the spectral finite elements, the numerical solver uses 2.3GB of RAM, 73 seconds for initialization (starting the model, initial outputs and interpolating the structure displacement field) and 40 minutes to compute approximately 4336 timesteps (including wavefield outputs).

The effect on the signal caused by the deformation alone is comparable to the effect caused by a crack for such configuration and type of measurement. Thus, not considering these effects when analysing wave propagation data for this level of mechanical loading results in less reliable detection and evaluation of such type of damages.

Illustration on a stratified anisotropic plate. Carbon fiber-reinforced polymers are largely employed in aerospace industries, where ensuring safety is essential. They are made of a polymer matrix reinforced with stacked layers of fibers. Usually, each fiber layer has a principal direction, presenting a transversely isotropic response to mechanical solicitations. Here, we model the wave propagation in a 16-layer Carbon Fiber Reinforced Polymer (CFRP) plate with dimensions of $900 \times 300 \times 2\,(16 \times 0.125)$ mm^3 under torsion as illustrated in Figure 2.6.

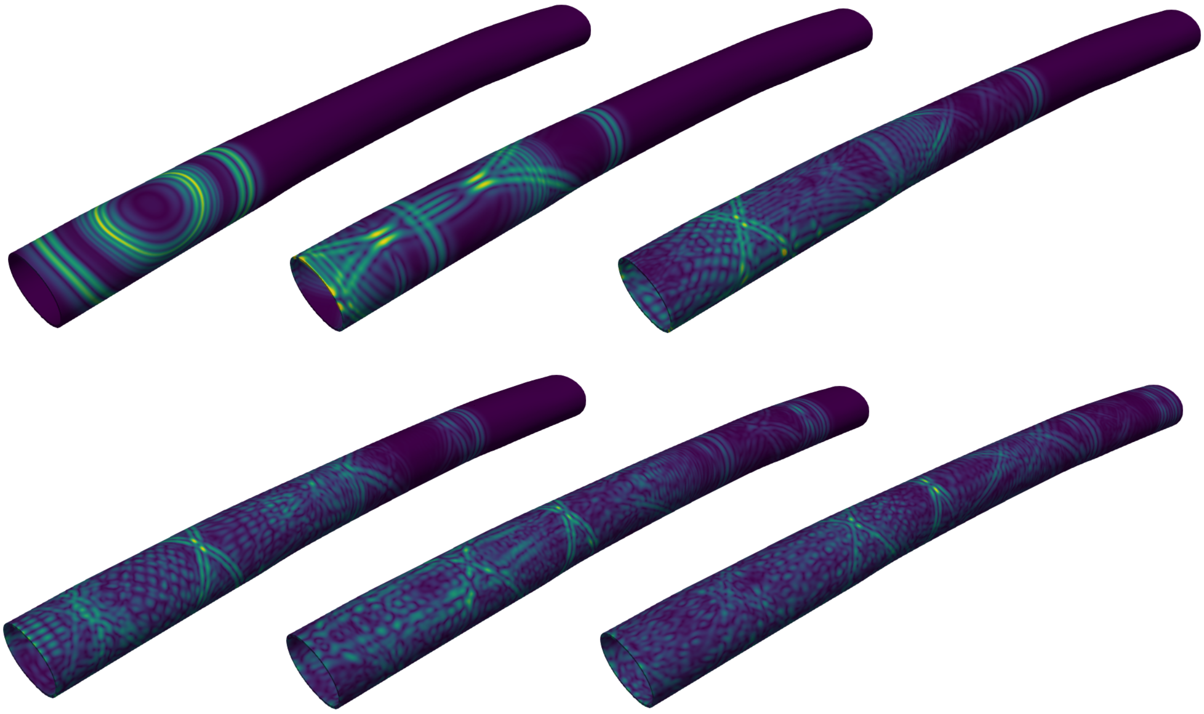


Figure 2.10 – Snapshots of the wavefield computed in the steel pipe subjected to the 4-point bending test at t = 70, 140, 230, 300, 390 and 480μs (left to right, top to bottom).

Each layer is modeled using the hyperelastic potential (2.22) where the direction related to the invariants I_4 and I_5 were set accordingly to the fiber directions in Figure 2.12. The material parameters are shown in Table 2.2. The deformation was caused by setting the Dirichlet boundary condition at $x = 0$ and applying a body force with the following form

$$f = \frac{y - 150}{300}\ 10^{-3}\mathrm{N/mm}^3, \quad \forall y \in [0, 300].$$

The resulting deformation by solving the quasi-static problem is shown in Figure 2.13. Snapshots of the wavefield computed solving the wave propagation problem can be seen in Figure 2.14.

The quasi-static problem has 270k DoFs and takes 50 minutes to run 10 Newton-Raphson iterations while using 12GB of RAM. The wave propagation problem requires around 8.3M mesh points, having at least two elements of fourth-order per wavelength, resulting in 25M DoFs. Satisfying the CFL condition (2.32), the timestep is $\Delta t = 0.0159\mu$s. The wave propagation solver requires 4GB of RAM, 9 minutes for initialization (starting the model, initial outputs and interpolating the structure displacement field) and 2 hours and 54 minutes to compute approximately 18867 timesteps (including wavefield outputs). This computation is particularly expensive as we need to satisfy the CFL condition for the mesh step related to the thickness discretization (16 elements for 2mm). This problem could be mitigated by using implicit-explicit schemes as in [Methenni, 2021].

This illustration shows the capability of our simulation tools to model thin structures with the possibility of choosing a hyperelastic law that best suits the modeling objectives. As discussed in the paragraph about acoustoelasticity, the acoustoelastic effect is controlled by the TOECs. Here, we choose to use a hyperelastic law that matches only second-order elastic constants although the implementation of laws that include the TOECs can also be considered.

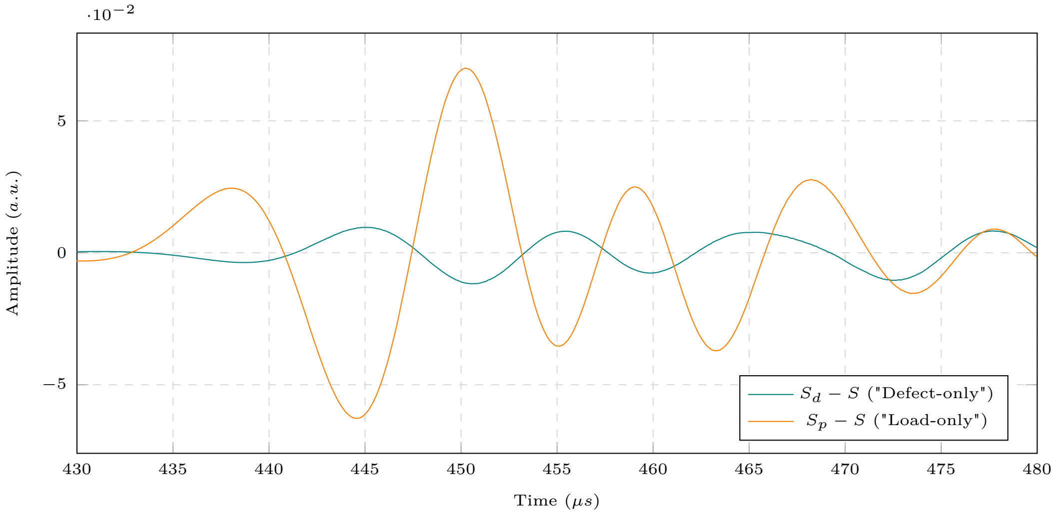


Figure 2.11 – Transducer signal at measure point subtracted from the baseline for the case with defect-only and load-only.

Material	Density (g/cm^3)	λ (GPa)	μ	l	m	n
Aluminum (fitted)	2.704	54.308	27.174	-181.0	-289.0	-336.0

Table 2.3 – Material parameters for aluminum, fitted to experimental data.

2.3.3 Experimental validation on an isotropic plate

Finally, in this section, we model and simulate an experimental setup to validate our model in the isotropic case. The experiment is presented in [Gandhi, Michaels, and Lee, 2012], where the authors apply axial forces to an aluminum plate and measure the change in wave speed caused by the load-induced deformation, for different directions and guided wave modes. Using Murnaghan's constitutive law, the authors propose an eigenproblem to obtain guided wave modes and their wave speeds in biaxially stressed isotropic materials. By applying canonical deformations and measuring the change in wave speed for different angles, the authors calibrated the TOECs of Murnaghan's constitutive law using experimental data. The obtained parameters are presented in Table 2.3. Here, we aim at modeling and simulating this experiment using the hyperelastic potential (2.21) with these parameters. We compare the simulated results with their experimental data to validate our model and numerical tools. The experiment is depicted in Figure 2.15.

In the experimental setup, eleven levels of axial loading are applied from 0MPa to 57.5MPa and the ultrasonic wave speeds are measured for 9 different angles of propagation with transducers positioned at 109mm from the center in the undeformed configuration [Gandhi, M., 2010]. For the validation procedures, we use their measured changes in wave speed for the S0 mode excited with a 250kHz 5-cycle Hanning windowed cosine. The wave speed is obtained by retrieving the ToF between transducer pairs for the S0 mode at a given load. To avoid effects of dispersion on the ToF measurement, we identify the 4th zero-crossing of the signal as the least affected by dispersion. As our solution is computed in the reference configuration, there is no need for correcting the distance change between transducers caused by the deformation. The authors in [Gandhi, Michaels, and Lee, 2012] fitted their parameters while neglecting this distance difference when computing the wave speed using the ToF. The computed displacement between the axial pair (#1,#6) has a maximum of 0.173mm at 57.5MPa. It represents a change in the total pair distance of around 0.08%, or, in terms of wave speed, around 4.2m/s when considering the mean wave speed of 5300m/s.

We simulate and obtain, for the different load levels, the changes in time of flight (ToF) for the angles

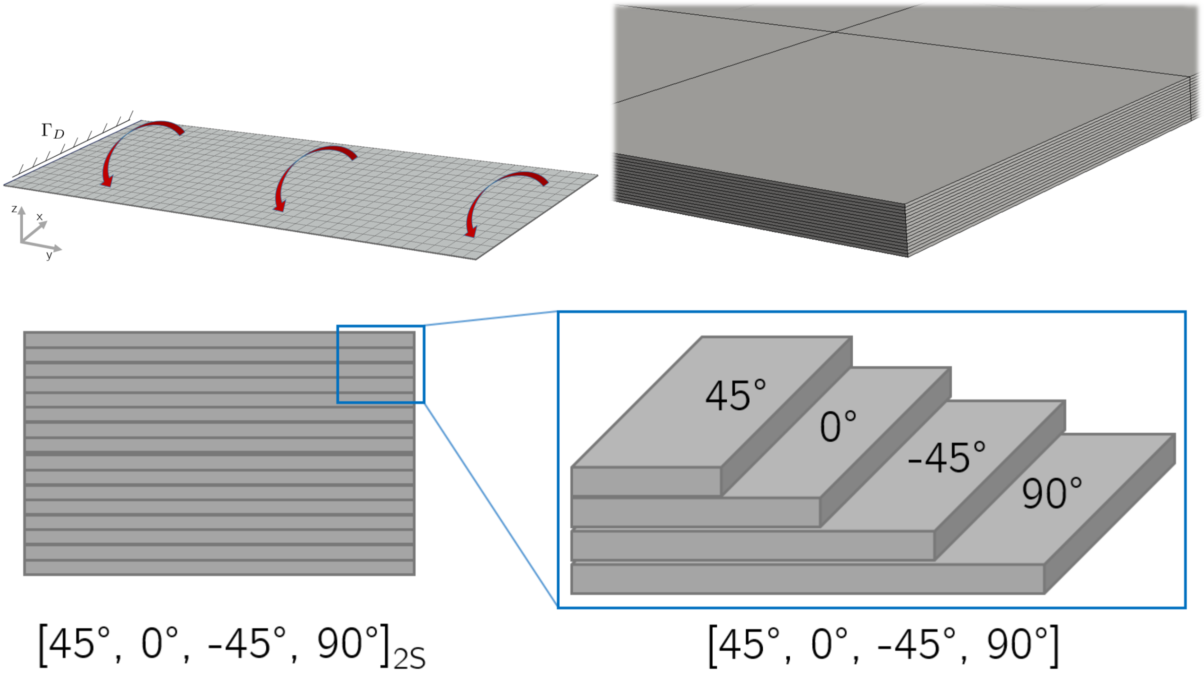


Figure 2.12 – Configuration for the quasi-static problem (upper left). Zoom showing the stratified aspect (upper right). Ply configuration (bottom).

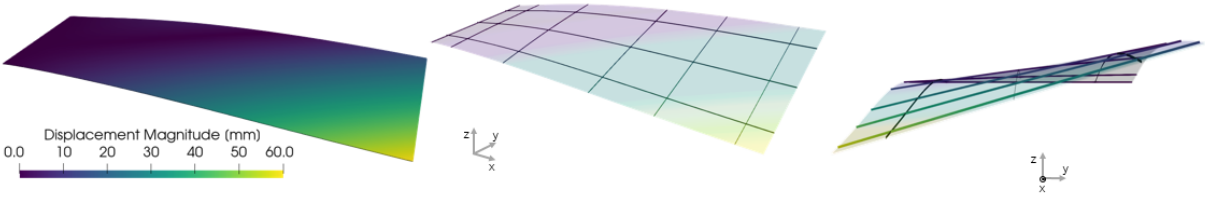


Figure 2.13 – Deformation of the CFRP plate under torsion. (visualization scaled 2x).

90º, 45º, 0º, -22.5º, -67.5º, corresponding to the transducer pairs (#1,#6), (#5,#10), (#4,#9), (#2,#7) and (#3,#8). The Y axis is considered to be at an angle of 90º. The changes in wave speed in the loading axis with respect to the load intensity are plotted in Figure 2.16a. Here, simulated results with different constitutive laws (SVK, CNH and Murnaghan's) are compared to experimental data and the analytical reference solution proposed in [Gandhi, Michaels, and Lee, 2012]. The behavior of SVK and CNH constitutive laws reflect the analysis done in the paragraph about acoustoelastic effects. With the SVK law, increasing axial load leads to increasing incremental wave speeds, what is opposite to the observed effect. Although the CNH models the right tendency of decreasing the wave speed, it underestimate its effects. Making use of the TOECs, the Murnaghan's law is able to properly model the experimental data. For the maximum load of 57.5MPa, the changes in wave speed with respect to the angle of propagation are plotted in Figure 2.16b. Here, our method implemented with the Murnaghan's law is also compared with the experimental data and the analytical proposition when obtaining the angular-dependency of the changes in wave speed. With our simulation, we retrieve the expected bell-shaped curve, effect of the load-induced anisotropy also obtained from the ToF measurements from experimental and simulated data.

Summarizing our validation achievement, the experiment is modeled and simulated starting from the mechanical loading that results in the deformation of the aluminum plate. This deformation is used as input to our linearized wave propagation model, and, using the proposed constitutive law and material parameters we are able to extract the ultrasonic signals of the same nature as in the experiment. Our

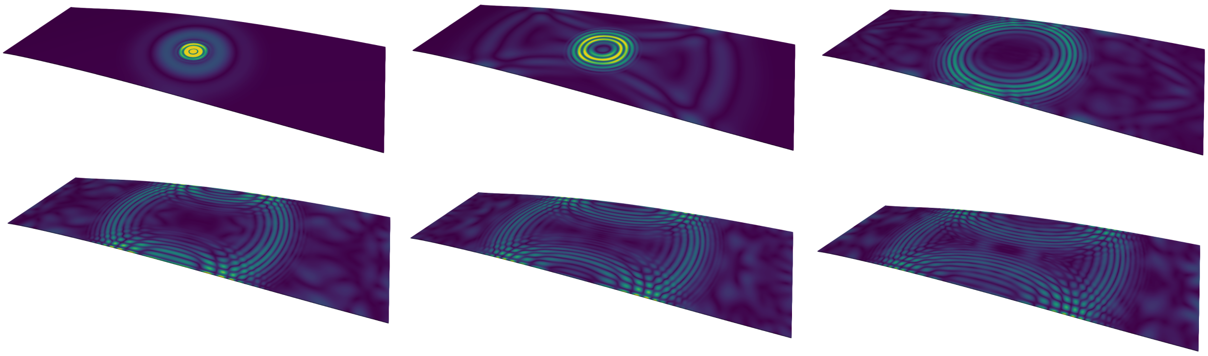


Figure 2.14 – Wavefield computed in a 16 layers CFRP plate under torsion at t = 45, 90, 180, 240, 270 and 300μs.

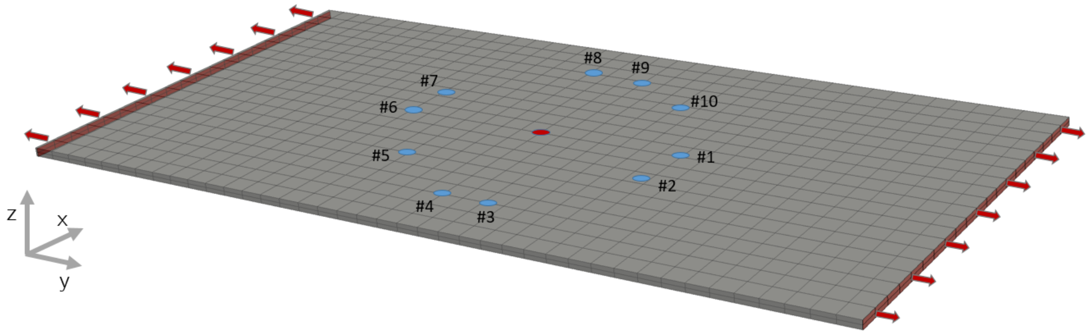


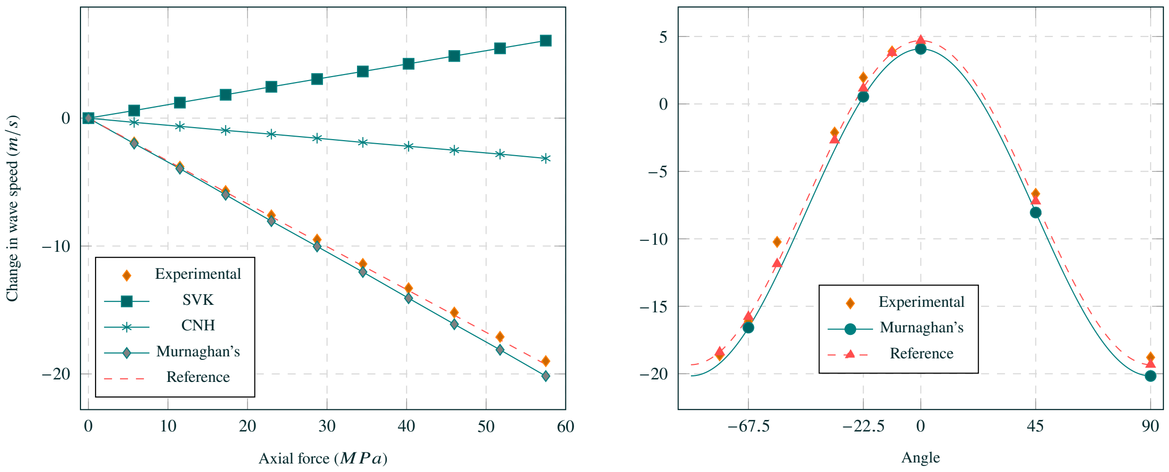
Figure 2.15 – Experimental setup for calibrating the TOECs using the acoustoelastic effect.

simulation reliably matches the experimental results while computing the whole wave field. We also implement different constitutive laws, obtaining their expected behavior. This further validates each modeling component and solvers used.

2.4 Conclusion & perspectives

In our work, we have proposed a set of numerical tools addressing the problem of ultrasonic guided-wave propagation in materials subject to pre-deformation. Modeling of such configurations includes two different problems: a quasi-static (time-independent) nonlinear problem satisfied by the large displacement field, and a linearized time-domain wave propagation problem. Our approach is generic in the sense that it can represent any type of constitutive law and geometries. For each problem, we have carefully chosen a dedicated numerical method – namely the 3D shell elements for the first problem and the time-domain SFEM for the second one – leading to computationally efficient algorithms. We have presented numerical results on 3D cases related to SHM configurations, thus illustrating the performances of our method for both isotropic and anisotropic materials. We also have compared our results against experimental data extracted from the literature, thus validating the overall modeling and numerical toolchain.

Thanks to the presented approach, we were able to illustrate how the effects of a pre-deformation can alter the ultrasonic signals used to monitor a specimen under inspection. It makes this work in a good position to get closer to real applications in industrial contexts, and to provide meaningful modeling tools for potential industrial partners aiming at designing SHM systems. In order to enhance the attractiveness of these modeling tools, it would be interesting to include a mean of automatically removing the bias induced by the pre-deformation or the EOCs in general. This could typically be done by considering an inverse problem where we would aim at reconstructing the pre-deformation sustained by the structure from the guided-wave measurements.



(a) Axial load-induced wavespeed changes.

(b) Load-induced wavespeed angular changes (57.5MPa).

Figure 2.16 – Comparison between simulated data (for the S0 mode at 250kHz and different constitutive laws) and experimental data, validating our proposed model.

Minha filha, desce das nuvens

Hoje, sentada ao lado
Da hortinha da minha mãe
Pensei Meu Deus, como a vida é besta.
Tanta conta pra pagar
E eu aqui sem um puto.
Minha filha, desce das nuvens,
Me diz meu terapeuta,
Uma queda dessas, se não mata
Pode lhe quebrar as pernas.
Mas hoje, eu aqui, ao som dos insetinhos
Penso
Meu Deus como a vida é besta
As nuvens refletidas na água da piscininha
E minha avó
Matando as borboletas chicoteando um pano de prato.
Criaturas horríveis, ela diz
Põe as larvas nas folhas
E estragam meu trabalho.
E eu que nem nunca tinha pensado nisso.
Se eu tivesse coragem vó
Eu saia das águas, das nuvens
E te ajudava
A matar borboletas
Mas é que elas são tão bonitas.
- Renata Dalmora

For the Asking
William Ackerman

Part II

Loading condition estimation in Structural Health Monitoring

Chapter 3

Introduction to least-squares minimization for solving inverse problems

In this chapter, we have the objective of selecting a suitable method for reconstructing structural displacement using ultrasonic waves. We do an introduction to least squares minimization in inverse problems related to wave propagation. The inverse problem is interpreted as a minimization problem where the misfit between observed data and a model is minimized by adjusting its parameters. We first introduce the formalism and the problems of interest, then we recall classical variational and sequential methods. Finally, we apply those methods to the problems of interest in one dimension for illustration and clarity purposes. By analyzing the cost and requirements of the presented methods, the Reduced-Order Unscented Kalman Filter was chosen as the most adequate method for our problem of estimating structural deformation.

Contents

Low amplitude wave phenomenon is intensively used to access information that is not in reach nor accessible without deconfiguring or destroying the medium of interest. In the specific case of Nondestructive Testing and Structural Health Monitoring, ultrasonic waves are one of the most employed, and more recently, in its guided configuration [Croxford et al., 2007; Mitra and Gopalakrishnan, 2016]. While particular characteristics of a propagating wave such as time of flight, dispersion, amplitude and phase can be used as a form of compressing and filtering the information transported by the wave, here we present inverse strategies that try to maximize the retrieved information by using raw ultrasonic data from transducers. The use of raw data requires proper dealing with “noise”, otherwise the method may present instability or a slow rate of convergence. Here, we aim to give an introduction to some traditional methods that can be used for estimation using wave propagation.

Given a specific wave propagation problem as the one presented in the first part, our objective is to estimate parameters of the wave model that bear uncertainties in the case where the information available in the wavefield is limited. This is usually the case in practical systems where just a few sensors are placed to measure it while most of it is left unmeasured. By having a model of the system with *a priori* initial conditions and parameters, the whole wavefield can be simulated. We call by *unknown* the inputs of the model that bear uncertainties. The misfit between measured and simulated data is used to update the unknown towards a better estimation, *i.e.* a value that results in simulated data closer to the measured one. The inverse problem of estimating this unknown can be treated in the context of least-squares optimization where the estimation is done by minimizing the squared misfit in a given norm.

The objective of this chapter is to introduce inverse problems related to wave propagation and associated existing inverse methods. By dealing with nonlinear inverse problems, we can consider the task of estimating structural deformation using limited ultrasonic measurements with the help of the “direct” model presented in Chapter 2. The methods are presented in their time-continuous form for conciseness and clarity, and in their time-discrete form for further detailing their practical aspects [Moireau, 2022]. In Section 3.1 we introduce the formalism and enunciate the inverse problems of interest interpreted as least squares minimization problems. In Section 3.2 the class of variational methods for minimization are presented and in Section 3.3, the sequential methods. In Section 3.4, we apply such methods in a one-dimensional framework for illustration and comparison. Finally, in Section 3.5 we briefly discuss the presented results and the arguments for choosing the Reduced-Order Unscented Kalman Filter for our problem of estimating the structural deformation. As an abuse of notation, in this chapter, we make no distinction when writing the function and its corresponding vector of components $\mathbb{R}^N$, in a given basis.

3.1 Introduction

In this section, we aim to introduce in a compact form the problems of interest and the associated formalism to treat them. As mentioned, the objective is to estimate a model unknown utilizing limited observed data. The difficulty of the unknown estimation task is highly dependent on how the unknown acts on the system. Therefore, we present two main classes of problems to be treated:

- Linear-quadratic problems: The dynamics depend linearly on the parameter. One example, that will be detailed later, is when the unknown appears in the source term of the wave problem.
- Nonlinear problems: The functional depends nonlinearly on the parameter to be reconstructed, such as in the state-parameter joint-estimation or in the first part of this thesis where the constitutive operator is parametrized by a quasi-static deformation.

More specifically, for better illustration, we exemplify two types of problems:

- Initial state reconstruction (nonlinear): The objective is to estimate the initial state of nonlinear dynamics. We present the Problem A.1, for its time-continuous version, and Problem I.1, for its time-discrete version.
- Parameter estimation (nonlinear and linear-quadratic): A time-independent parameter is estimated. The dependency can be a constitutive or a source term dependency. We present Problem B.1 and Problem II.1 for a nonlinear parametric dependency for the time-continuous and

time-discrete cases, respectively. For a linear case, we present Problem B.2 and Problem II.2. We show that the problem of estimating a time-independent parameter can be interpreted as an initial state estimation and, for this reason, the initial state estimation problem will be considered as the generic case. Therefore, we will focus on the presentation of estimation methods for estimating an initial state, although some links between the different cases will be given. The inverse problem of reconstructing the unknown is solved by minimizing a functional containing a term representing the misfit between measurements and simulated data and a term for regularizing the inverse problem. A model (dynamics) of the system is also given with an associated observation operator that produces the simulated data.

In Section 3.1.1 we present the formalism and generally an inverse problem in a time-independent framework as an introduction, in Section 3.1.2 we introduce the overall format of the dynamical systems of interest that will be related to wave propagation problems and to what was presented in Section 3.1.1. Finally, in Section 3.1.3 and Section 3.1.4 we formally present the inverse problems of interest in its time-continuous and time-discrete formulations.

3.1.1 General formulation

In this section we present a simple and comprehensive formulation for the inverse problems of interest, introducing the main elements and notations. As our objective is to apply such methods in a discrete context, we simplify the presentation by considering only finite-dimensional Hilbert spaces if not explicitly specified otherwise. To introduce the notations related to the inverse problems, we start by defining a nonlinear “input-output“ operator $\Psi : \Theta \mapsto \mathcal{Y}$ that generates the observations $y \in \mathcal{Y}$ from a parameter $\theta \in \Theta$, where Θ is the parametric space and $\mathcal{Y}$ is the space of observations. We also introduce the measurement y

$$y = \Psi(\theta) + \chi,$$

which represents the measured data and indirectly contains information about the unknown to be estimated. The function χ represents a disturbance modeling potential measurement noises. Generally, the inverse problem will be stated as: for given observations y, find the parameter θ. This could also be stated as

$$\text{“}\theta = \Psi^{-1}(y)\text{”}.$$

Additionally to the difficulty posed by the observations disturbance χ, the inverse Ψ^{-1} is usually not defined or costly to compute, making its direct inversion not feasible. Instead, the inverse problem can be interpreted as finding $\overline{\theta}$ such as [Nocedal and Wright, 2006; Luenberger and Ye, 2008]

$$\overline{\theta} = \operatorname*{argmin}_{\theta \in \Theta} \mathcal{J}(\theta) = \operatorname*{argmin}_{\theta \in \Theta} \left\{ r_\Theta(\theta_0 - \theta) + \frac{1}{2} \|y - \Psi(\theta)\|^2_{\mathcal{Y}} \right\}, \tag{3.1}$$

where r_Θ is an operator on Θ and $\mathcal{J} : \Theta \mapsto \mathbb{R}$ is the functional. The operator r_Θ acts on the functional to regularize the associated inverse problem. In most cases presented here, this regularization term will have a quadratic form as

$$r_\Theta(\theta - \theta_0) = \frac{1}{2}\big(R_\Theta(\theta - \theta_0), (\theta - \theta_0)\big)_\Theta,$$

where $\theta_0 \in \Theta$ represents an initial guess of the parameter to be estimated and $R_\Theta : \Theta \mapsto \Theta$ a self-adjoint positive regularizing operator.

Linear case. In cases where $\Psi \in \mathcal{L}(\Theta, \mathcal{Y})$, $\Psi(\theta)$ becomes $\Psi\theta$ and the minimization (3.1) becomes

$$\overline{\theta} = \operatorname*{argmin}_{\theta \in \Theta} \mathcal{J}(\theta) = \operatorname*{argmin}_{\theta \in \Theta} \left\{ r_\Theta(\theta) + \frac{1}{2} \|y - \Psi\theta\|^2_{\mathcal{Y}} \right\}. \tag{3.2}$$

This linear minimization problem has its minimum in a closed form

$$\overline{\theta} = \Big(R_\Theta + \Psi^*\Psi\Big)^{-1}\Big(R_\Theta\theta_0 + \Psi y\Big). \tag{3.3}$$

3.1.2 Time-dependent systems

For the inverse problems of interest, the unknowns are not directly observable. Instead, the observations can only be done on the state of dynamics controlled by the unknown. For instance, the unknown can present in the constitutive part of a wave dynamics or in its source term, and the observations be done in its displacement field. Therefore, we present here the dynamic systems we deal with and how to interpret an inverse problem involving such dynamics. As we will see, information about the unknown can be retrieved through a model that relates the unknown and the observed state. We briefly introduce in this section the generic form of such dynamics and its operators. We first present the time-continuous case and then the time-discrete case.

3.1.2.1 Time-continuous dynamics

Here, we present the time-continuous dynamics of interest in a sufficient generic form. We introduce a state variable, the potentially nonlinear dynamics operator A and the observation operator C as

$$z \in \mathcal{Z}_T = C^1([0,T];\mathcal{Z}), \quad A : \mathcal{Z} \times [0,T] \mapsto \mathcal{Z} \quad \text{and} \quad C : \mathcal{Z} \mapsto \mathcal{Y},$$

for a maximum time T. We assume that the evolution of state through time can be described using the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = A(z(t),t) + b(t), \quad \forall t \in [0,T], \tag{3.4}$$

where $z(0)$ is given. We assume that observations $y \in \mathcal{Y}_T = \mathrm{L}^2(0,T;\mathcal{Y})$ are generated from the state by the observation operator $C\big(z(t)\big)$, as

$$y(t) = C\big(z(t)\big) + \chi(t), \quad \forall t \in [0,T]. \tag{3.5}$$

The observations $y(t)$, or measurements, will represent the available data retrieved from the system we are modeling. A disturbance $\chi(t)$ is added to represent potential unmodeled effects on the process of measurement such as electronic or ambient noise. Here, we have presented the dynamics in the format that will be used throughout the following sections. The dynamics operators and how the unknowns act in them will be defined differently for each illustration or case of application.

Linear case. If the operator A is linear *w.r.t.* the state $z(t)$, namely $A(z(t),t) = A(t)z(t)$, we have the linear dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = A(t)z(t) + b(t), \quad \forall t \in [0,T], \tag{3.6}$$

where $z(0)$ is given. If the observation operator is also linear, we have

$$y(t) = Cz(t) + \chi(t), \quad \forall t \in [0,T]. \tag{3.7}$$

Formulation for wave propagation problems. We show here that the dynamics (3.6) written in the first-order (time-derivative) form, comprises wave propagation problems. Such demonstration is particularly interesting as the strategies to be presented in the next sections consider these first-order dynamics but our objective is to apply those strategies to wave propagation inverse problems. To do so, starting from the finite-dimensional second-order wave propagation dynamics, we rewrite it as a first-order dynamics. For a displacement field $u \in C^2([0,T],V)$ and V a finite-dimensional Hilbert space, we introduce a discretized wave propagation problem in the form

$$\begin{cases} \dfrac{\mathrm{d}^2}{\mathrm{d}t^2} u(t) + A_u u(t) = f(t), \\[2ex] u(0) = u^0, \quad \dfrac{\mathrm{d}u}{\mathrm{d}t}(0) = v^0, \end{cases} \tag{3.8}$$

where $A_u \in \mathcal{L}(V)$ is a self-adjoint positive operator, $f \in C^0([0,T],V)$ is the source term. We complete the description of such a problem by defining the observations $y_u \in \mathrm{L}^2(0,T;\mathcal{Y}_u)$ with

$$y_u(t) = C_u\big(u(t)\big) + \chi_u(t),$$

where $C_u : V \mapsto \mathcal{Y}_u$ and $\chi_u \in \mathrm{L}^2(0,T;\mathcal{Y}_u)$ is a disturbance. With the objective of writing the wave propagation dynamics in the form of (3.6) we define the velocity, state variables and disturbance as

$$v = \frac{\mathrm{d}u}{\mathrm{d}t}, \quad z = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \chi(t) = \begin{pmatrix} 0 \\ \chi_u(t) \end{pmatrix}.$$

Finally, writing the operator and source term of (3.4) as

$$A\big(z(t)\big) = \begin{pmatrix} 0 & I \\ -A_u & 0 \end{pmatrix} z(t), \quad b(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix},$$

with I being the identity operator, we can then write the wave propagation problem (3.8) in the form of the first-order dynamics (3.6)

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} 0 & I \\ -A_u & 0 \end{pmatrix}\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix} \\ z(0) = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \end{cases} \tag{3.9}$$

The observations are analogously obtained by defining the operator in (3.5) as

$$C\big(z(t)\big) = C_u\big(u(t)\big),$$

that, if it is linear, analogously to (3.7) we have

$$Cz(t) = \begin{pmatrix} C_u & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}(t).$$

Therefore, the methods discussed in the following sections for solving inverse problems with an associated first-order differential equation (*i.e.* (3.6)) can also be used for wave propagation problems. In practice, when finite elements procedures are applied, we have naturally the inner product for the space V as $(u,u)_V = (\mathbb{M}u,u)_2$ (Euclidian norm) and the operator $A_u := \mathbb{M}^{-1}\mathbb{K} \in \mathcal{L}(\mathbb{R}^{N_u})$, where the matrices $\mathbb{M}$ and $\mathbb{K}$ are the traditional mass and stiffness matrices. In some cases, as presented in the next sections, the operator A_u may depend on the state, hence becoming a nonlinear problem as (3.4).

3.1.2.2 Time-discrete dynamics

Given a time step Δt, the N-steps time discretization is done as $z^n \approx z(n\Delta t)$ with $n \in [\![0;N]\!]$, where Δt usually is fixed to satisfy stability conditions. The approximation error is expected to decrease as Δt decreases. We would like to write the dynamics (3.4) with a time-stepping procedure, with $b^n = b(n\Delta t)$, such as

$$z^{n+1} = \Phi_{n+1|n}(z^n) + b^n, \quad \forall n \in [\![0;N-1]\!], \tag{3.10}$$

where we define $\Phi_{n+1|n} : \mathcal{Z} \mapsto \mathcal{Z}$ the discretized version of the flow defined by the A operator, transitioning a state from the step n to $n+1$. When the operator acts linearly in the state we represent it as

$$\Phi_{n+1|n}(z^n) = \Phi_{n+1|n}z^n,$$

so the dynamics become

$$z^{n+1} = \Phi_{n+1|n}z^n + b^n, \quad \forall n \in [\![0;N-1]\!], \tag{3.11}$$

where z^0 is given. We assume that observations $\{y^n\}_{n=0}^N$ are generated from the state from the same observation operator C, having

$$y^n = C(z^n) + \chi^n, \tag{3.12}$$

with $\chi^n = \chi(n\Delta t)$.

Formulation for wave propagation problems. As done for the time-continuous case, we can write a discrete wave propagation problem in the form of (3.10) for the nonlinear case and in the form of (3.11) for the linear case. We start by discretizing the wave propagation problem in (3.8), approximating the time derivative with a second order leapfrog scheme with time step Δt such that $u^n \approx u(n\Delta t)$ and using $f^n = f(n\Delta t)$

$$\mathbb{M}\,\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} + \mathbb{K}\,u^n = f^n, \quad u^1 = u^0 = 0, \quad \forall n \in [\![1\,;N-1]\!], \tag{3.13}$$

where $u \in \mathbb{R}^{N_u}$ is the discretized displacement field, for instance, a finite element vector. The matrices $\mathbb{M}$ and $\mathbb{K}$ represent the mass and stiffness matrices obtained also from finite elements procedures. To use the estimation methods described in this section, we write the wave propagation problem as a discrete first-order dynamics by introducing the velocity unknown v

$$\begin{cases} \mathbb{M}\,\dfrac{u^{n+1} - u^n}{\Delta t} - \mathbb{M}\,v^{n+1} = 0, \\ \mathbb{M}\,\dfrac{v^{n+1} - v^n}{\Delta t} + \mathbb{K}\,u^n = f^n. \end{cases}$$

Introducing the state variable as $z^n = \begin{pmatrix} u^n & v^n \end{pmatrix}^\intercal \in \mathcal{Z}$, we can write this system as

$$\Phi_1 z^{n+1} = \Phi_0 z^n + f_z^n, \quad z^0 = 0, \quad \forall n \in [\![0\,;N-1]\!],$$

with

$$\Phi_1 = \begin{pmatrix} \mathbb{M} & -\Delta t\mathbb{M} \\ 0 & \mathbb{M} \end{pmatrix}, \quad \Phi_0 = \begin{pmatrix} \mathbb{M} & 0 \\ -\Delta t\mathbb{K} & \mathbb{M} \end{pmatrix}, \quad f_z^n = \begin{pmatrix} 0 \\ \Delta t f^n \end{pmatrix}, \tag{3.14}$$

analogously to (3.11), yielding

$$z^{n+1} = \Phi_{n+1|n} z^n + b^n, \quad \forall n \in [\![0\,;N-1]\!] \tag{3.15}$$

with

$$\Phi_{n+1|n} = \Phi_1^{-1}\Phi_0, \quad \text{and} \quad b^n = \Phi_1^{-1} f_z^n.$$

Using this notation for Φ, one can write

$$z^n = \Phi_{n|0} z^0 + \sum_{i=0}^{n-1} \Phi_{n|i+1} b^i,$$

where the operator subscript “a|b” means taking the step b as input and forwarding it to the step a.

3.1.3 Time-continuous inverse problems of interest

In this section, we present the problems of interest in their time-continuous form. The first is the reconstruction of the initial state of a dynamical system, which is nonlinear, *i.e.* the functional to be minimized depends nonlinearly on the unknown. The second is to estimate a parameter of the dynamical system. The second problem is divided into two sub-problems: linear-quadratic and nonlinear. In the linear-quadratic case, the dependency *w.r.t.* the unknown is characterized linearly on the dynamics but quadratically in the least-squares misfit term. We also show that the parametric estimation, where the parameter is time-independent, can be written in the form of initial state reconstruction, making the initial state reconstruction case the most generic among the presented problems. For the sake of conciseness, the problems are described by defining the:

1. *Functional/Cost function*: A functional with the unknown as argument. Its minimization *w.r.t.* the unknown is interpreted as the inverse problem.

2. *Model (Dynamics)*: How the state evolves in time defining the operator and source term in (3.4) and how the state depends on the unknown. Used to generate the simulated data.
3. *Available data*: The available observations, how they are obtained and *a priori* guess about the unknown.

The *functional*, or *cost function*, quantifies the misfit between the simulated data (generated with an estimation of the unknown) with the given observations. This misfit is computed as the square, in a given norm, of the difference between the observed state of the model and the given observations. As disturbances in the observations are often present, a regularization term is added to the functional that aims at controlling effectively the way deviations from the *a priori* unknown are penalized. The *dynamics* represents the model of the target system by an equation that describes the time evolution of the system state from its initial conditions. We call as target system the system from which the observations were obtained, the denomination "target" is due to the fact that we aim at estimating its parameters that we consider unknown or bearing uncertainties. The *available data* comprises, additionally to the model, actual measurements on the target system state and an initial guess of the unknown. These measurements are usually done by physical sensors (*i.e.* accelerometers, gauges, temperature sensors, piezoelectric transducers, etc.), which are modeled through observation operators. The link between this form of interpreting an inverse problem and the form with an "input-output" operator will be established to elucidate the relation between the unconstrained inverse problem (3.1) and one constrained by dynamics.

Given this format of presenting an inverse problem, we describe in the following sections the problem of reconstructing the initial state of time-continuous dynamics and the problem of estimating a parameter acting in its constitutive operator or in the source term.

3.1.3.1 Initial state reconstruction

By using the proposed structure for presenting the inverse problems of interest, we describe the first one that aims at reconstructing an error ζ in the initial state of a dynamical system by means of observations on the state. The problem can be described as follows. The functional is regularized by an operator $r(\zeta) : \mathcal{Z} \mapsto \mathbb{R}$ defined as

$$r(\zeta) = \frac{1}{2}(R\zeta, \zeta)_{\mathcal{Z}}, \quad \forall \zeta \in \mathcal{Z},$$

where $R : \mathcal{Z} \mapsto \mathcal{Z}$ is a self-adjoint positive operator.

Nonlinear problem A.1

1. **Functional**:

$$\mathfrak{J}_T(\zeta) = \left\{ r(\zeta) + \frac{1}{2}\int_0^T \|y(t) - C\big(z_\zeta(t)\big)\|^2_{\mathcal{Y}}\, \mathrm{d}t \right\} \tag{3.16}$$

2. **Dynamics**: (3.4) with initial condition bearing uncertainty, namely,

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} z_\zeta(t) = A(z_\zeta(t), t) + b(t), & \forall t \in [0, T], \\ \\ z_\zeta(0) = z_0 + \zeta. \end{cases} \tag{3.17}$$

3. **Available data**

   An initial guess z_0 and noisy observations $y(t)$ from (3.5) with $z(t)$ generated from (3.4) with $z(0) = z_0 + \zeta^*$.

The objective is to estimate the error ζ^* in the initial state through the minimization of (3.16).

This problem is so-called "nonlinear" because its functional depends nonlinearly on the unknown. Differing from the functional in (3.1), in the case where the minimization is constrained to a time-dependent dynamics, one can notice the dependency of the optimal value *w.r.t.* to T. This dependency means that, as intuitively expected, the optimal value depends on the amount of information available through observations.

Initial state reconstruction for wave propagation problems. Using the wave dynamics in its first-order form (3.9), one can propose an inverse problem to reconstruct the initial displacement field u_0 using noisy measurements up to a given time T. As an example, in practice, the observation operator C could represent piezoelectric transducers. Such transducers are commonly used for measuring ultrasonic wave propagation. They are bonded to the structure surface, generating electrical signals from the surface displacement field. The initial state of the system can then be reconstructed using the noisy information acquired during a finite period.

Unconstrained formulation. We give here an example of how problems such as Problem A.1 can be related to the time-independent inverse problem presented in Section 3.1.1. The idea is that even with the unknown not being directly observable in the physical system, an operator that outputs the observations from the unknown can be constructed by using the given model (dynamics). To do so, we develop the idea for a linear dynamics case where $A\big(z_\zeta(t)\big)$ becomes $Az_\zeta(t)$ with functional as (3.1). The minimization in the form of (3.16) can be represented in an alternative form of (3.1)

$$\min_{\zeta\in\mathcal{Z}} \mathcal{J}(\theta) = \min_{\zeta\in\mathcal{Z}} \left\{ r(\zeta) + \frac{1}{2}\,\|y-\Psi(\zeta)\|^2_{\mathcal{Y}_T} \right\},$$

where the norm $\|\cdot\|^2_{\mathcal{Y}_T}$ is defined as

$$\|y-\Psi(\zeta)\|^2_{\mathcal{Y}_T} = \int_0^T \|y(t)-\Psi(\zeta)(t)\|^2_{\mathcal{Y}}\;\mathrm{d}t.$$

Refering to (3.17), $\Psi : \mathcal{Z} \mapsto \mathcal{Y}_T$ is defined as

$$\Psi(\zeta)(t) = C\big(z_\zeta(t)\big), \quad \forall t\in[0,T],$$

where, using the Duhamel's formula [Bhatia, 2015], we have

$$z_\zeta(t) = e^{tA}(z_0+\zeta) + \int_0^t e^{(t-s)A} b(s)\;\mathrm{d}s, \quad \forall t\in[0,T].$$

This formulation makes explicit how the observations are related to the unknown, the complexity and the ill-posedness of the problem. Additionally, the observability of the inverse problem can be analyzed through this relation, namely how much information can be retrieved given the observations and the model. We will show in the time-discrete section that when the problem is time-discrete, the time-dependent problem can be written explicitly in the form of (3.1).

3.1.3.2 Parameter reconstruction

We treat here the case where the operator A depends on a parameter bearing uncertainties and the initial state is considered to be fully known. The parameter estimation problem will be divided into two types: linear-quadratic and nonlinear problems. The state dynamics is linear for the linear-quadratic case or nonlinear, otherwise, *w.r.t.* the parameter and the functional to be minimized is quadratic *w.r.t.* to the observations. To state such problems, we redefine the operator A in (3.4) with a nonlinear dependency *w.r.t.* the parameter and state, namely

$$A(z(t),t) \text{ redefined as } A(z(t),\theta,t)$$

then the nonlinear dynamics are written

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = A(z(t), \theta, t) + b(t), \quad \forall t \in [0, T], \tag{3.18}$$

where $z(0)$ is given. The objective will be then to estimate an associated parameter θ using observations. The functionals representing the following inverse problems are regularized by an operator $r_\Theta(\zeta) : \Theta \mapsto \mathbb{R}$ defined as

$$r_\Theta(\theta) = \frac{1}{2}(R_\Theta \theta, \theta)_\Theta, \quad \forall \theta \in \Theta, \tag{3.19}$$

and $R_\Theta : \Theta \mapsto \Theta$ is a self-adjoint positive operator. The observations are considered to be generated from (3.18) with θ^* with added disturbance χ.

Nonlinear problem. Here, we consider that the dynamics has a nonlinear dependency on the parameter and applies linearly to the state, namely

$$A(z(t), \theta, t) = A(\theta, t) z(t), \quad A : \Theta \times [0, T] \mapsto \mathcal{L}(\mathcal{Z}).$$

First, we describe the problem in a more natural description below.

> 1. **Functional**:
>
> $$\mathcal{J}_T(\theta) = \left\{ r_\Theta(\theta - \theta_0) + \frac{1}{2} \int_0^T \| y(t) - C\big(z_\theta(t)\big) \|_{\mathcal{Y}}^2 \, \mathrm{d}t \right\}, \tag{3.20}$$
>
> 2. **Dynamics**: (3.18) with parameter bearing partial uncertainty and known initial condition, namely,
>
> $$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} z_\theta(t) = A(\theta, t) z_\theta(t) + b(t), & \forall t \in [0, T], \\ z_\theta(0) = z(0). \end{cases} \tag{3.21}$$
>
> 3. **Available data**
>
>   An initial guess θ_0 and noisy observations $y(t)$ from (3.5) with $z(t)$ generated from (3.18) with $\theta = \theta^*$.
>
> The objective is to estimate θ^* through the minimization of (3.20).

In fact, all problems to be solved can be treated as an initial state reconstruction problem. We are then interested in presenting the parameter estimation as the initial state reconstruction problem. In the following, we show that the parameter reconstruction problem can be written as Problem A.1 by augmenting the state variable and the dynamics with

$$^\sharp z(t) \in {}^\sharp\mathcal{Z} = \mathcal{Z} \times \Theta, \quad {}^\sharp A : \Theta \mapsto \mathcal{L}({}^\sharp\mathcal{Z}) \quad \text{and} \quad {}^\sharp b(t) \in \mathcal{Z} \times \Theta, \quad \forall t \in [0, T],$$

giving

$$^\sharp z(t) = \begin{pmatrix} z(t) \\ \theta \end{pmatrix}, \quad {}^\sharp A(\theta, t) = \begin{pmatrix} A(\theta, t) & 0 \\ 0 & 0 \end{pmatrix}, \quad {}^\sharp b(t) = \begin{pmatrix} b(t) \\ 0 \end{pmatrix}.$$

The dynamics then becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} {}^\sharp z(t) = {}^\sharp A(\theta, t) \, {}^\sharp z(t) + {}^\sharp b(t), \quad \forall t \in [0, T], \tag{3.22}$$

where $^\sharp z(0)$ is given. The observations are obtained using an augmented observation operator $^\sharp C : {}^\sharp\mathcal{Z} \mapsto \mathcal{Y}$ as

$$y(t) = {}^\sharp C\big({}^\sharp z(t)\big) + \chi(t) = C\big(z(t)\big) + \chi(t), \quad \forall t \in [0, T]. \tag{3.23}$$

As denoted, the parameter is expected to be time-independent. Hence, we interpret the problem of estimating the parameter as an initial state estimation as below, where we reconstruct the parameter-only initial condition error, denoted as ζ_θ.

Nonlinear problem B.1

1. **Functional**:

$$\mathcal{J}_T(\zeta_\theta) = \left\{ r_\Theta(\zeta_\theta) + \frac{1}{2}\int_0^T \|y(t) - {}^\sharp C\big({}^\sharp z_\theta(t)\big)\|^2_{\mathcal{Y}} \,\mathrm{d}t \right\} \tag{3.24}$$

2. **Dynamics**: (3.22) with initial condition bearing partial uncertainty, namely,

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^\sharp z_\theta(t) = {}^\sharp A(\theta, t)\,{}^\sharp z_\theta(t) + {}^\sharp b(t), \qquad \forall t \in [0, T], \\ {}^\sharp z_\theta(0) = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta \end{pmatrix}. \end{cases} \tag{3.25}$$

3. **Available data**

   An initial guess θ_0 and noisy observations $y(t)$ from (3.23) with ${}^\sharp z(t)$ generated from (3.22) with

$$^\sharp z(0) = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta{}^* \end{pmatrix}.$$

The objective is to estimate the parameter-only error $\zeta_\theta{}^*$ through the minimization of (3.24).

This type of problem is illustrated in Section 3.4.2 and Section 3.4.3, where perturbations in the medium wave velocity are constructed using limited observed wavefield data.

Linear-quadratic problem. In the inverse problem case where the dynamics have linear dependency *w.r.t.* the parameter, namely Linear-Quadratic (LQ) problem, we can write the parameter-estimation problem like Problem B.1 by defining the operator

$$A^L \in \mathcal{L}({}^\sharp\mathcal{Z}), \quad A^L(t) = \begin{pmatrix} A & B(t) \\ 0 & 0 \end{pmatrix}, \quad \forall t \in [0, T],$$

where $B(t) \in \mathcal{L}(\Theta, \mathcal{Z})$ provides a source term from the parameter. Then, $b(t) = 0$ as no extra source terms are considered. The modified dynamics becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\,{}^\sharp z(t) = A^L(t)\,{}^\sharp z(t), \quad \forall t \in [0, T], \tag{3.26}$$

where ${}^\sharp z$ is the augmented state (state-parameter) and ${}^\sharp z(0)$ is given. Then the problem can be stated in the form of an initial state reconstruction as below, which is a lightly modified version of Problem B.1.

Linear-quadratic problem B.2

1. **Functional**:

$$\mathfrak{J}_T(\zeta_\theta) = \left\{ r_\Theta(\zeta_\theta) + \frac{1}{2}\int_0^T \|y(t) - {}^\sharp C\left({}^\sharp z_\theta(t)\right)\|^2_{\mathcal{Y}}\, \mathrm{d}t \right\} \tag{3.27}$$

2. **Dynamics**: (3.26) with initial condition bearing partial uncertainty, namely,

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^\sharp z_\theta(t) = A^L(t)\,{}^\sharp z_\theta(t), \qquad \forall t \in [0,T], \\ {}^\sharp z_\theta(0) = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta \end{pmatrix}. \end{cases} \tag{3.28}$$

3. **Available data**

   An initial guess θ_0 and noisy observations $y(t)$ from (3.23) with ${}^\sharp z(t)$ generated from (3.26) with

$${}^\sharp z(0) = \begin{pmatrix} z_0 \\ \theta_0 + {\zeta_\theta}^* \end{pmatrix}.$$

The objective is to estimate the parameter-only error ${\zeta_\theta}^*$ through the minimization of (3.27).

This type of problem is illustrated in Section 3.4.1, where a source term spatial dependency is constructed by means of limited observed wavefield data.

3.1.4 Time-discrete inverse problems of interest

Although it will be simpler to present some optimization methods for time-continuous inverse problems, it is important to write and analyze them in their discrete form. One reasonable path would be to propose optimization methods in the time-continuous formulation and then discretize to solve it numerically. However, if the time-continuous minimization strategy is discretized, it may not satisfy optimality for the discrete version of the dynamics, we then do the approach "discretization-and-then-control" [Zuazua, 2005; Moireau, 2022]. For that reason, we write the problems directly in their discrete form to then propose respective minimization methods. We write here the Problem A.1, Problem B.1 and Problem B.2 in their time-discrete forms.

3.1.4.1 Initial state reconstruction

Based on (3.10), we write Problem A.1 in its discrete form.

Nonlinear discrete problem I.1

1. **Functional**:

$$\mathfrak{J}_N(\zeta) = \left\{ r(\zeta) + \frac{\Delta t}{2}\sum_{n=0}^{N} \|y^n - C(z_\zeta^n)\|^2_{\mathcal{Y}} \right\} \tag{3.29}$$

2. **Dynamics**: (3.10) with initial conditions bearing uncertainty, namely,

$$\begin{cases} z_\zeta^{n+1} = \Phi_{n+1|n}(z_\zeta^n) + b^n, \quad \forall n \in [\![0; N-1]\!], \\ z_\zeta^0 = z_0 + \zeta. \end{cases} \tag{3.30}$$

3. **Available data**

> An initial guess $z_\circ$ and noisy observations $\{y^n\}_{n=0}^N$ from (3.12) with $\{z^n\}_{n=0}^N$ generated from (3.10) with $z^0 = z_\circ + \zeta^*$.
>
> The objective is to estimate the error ζ^* in the initial state through the minimization of (3.29).

Unconstrained formulation. As done in the time-continuous case, a link between the minimization in Problem I.1 and the time-independent inverse problem (3.1) can be established. For the discrete case, an explicit relation can be obtained. First, we define the Hilbert space $\mathcal{Y}_N$ and stack the vectors by defining $y_N \in \mathcal{Y}_N$ and $\Psi_N(\zeta) \in \mathcal{Y}_N$

$$y_N = \begin{pmatrix} y^0 \\ \vdots \\ y^N \end{pmatrix}, \quad \Psi_N(\zeta) = \begin{pmatrix} C(z_\zeta^0) \\ \vdots \\ C(z_\zeta^N) \end{pmatrix}$$

where the modified space of observations $\mathcal{Y}_N$ has the inner product defined as

$$(u_N, v_N)_{\mathcal{Y}_N} = \Delta t \sum_{n=0}^{N} (u^n, v^n)_{\mathcal{Y}}, \quad \forall u_N, v_N \in \mathcal{Y}_N.$$

The dynamics solution $\{z_\zeta^n\}_{n=0}^N$ present in the $\Psi_N(\zeta)$ vector can then be written in its closed form

$$z_\zeta^n = \Phi_{n|n-1}\Big(\Phi_{n-1|n-2}\big(\cdots \Phi_{1|0}(z_\circ + \zeta) + b^0 \cdots\big) + b^{n-2}\Big) + b^{n-1}.$$

With these definitions, the functional in (3.29) becomes

$$\mathcal{J}_N(\zeta) = \left\{ r(\zeta) + \frac{1}{2} \|y_N - \Psi_N(\zeta)\|_{\mathcal{Y}_N}^2 \right\}.$$

3.1.4.2 Parameter reconstruction

As done for the time-continuous case, we write the discrete version of the parameter estimation problem as an initial state reconstruction.

Nonlinear case. We start by introducing the augmented discrete operator by redefining

$$\Phi_{n+1|n}(z^n) \text{ as } {}^\sharp\Phi_{n+1|n}(\theta^n) z^n.$$

We specify the augmented state and operators as

$$\begin{aligned} {}^\sharp z^n &\in {}^\sharp\mathcal{Z} = \mathcal{Z} \times \Theta, && \forall n \in [\![0; N]\!]. \\ {}^\sharp\Phi_{n+1|n} &: \Theta \mapsto \mathcal{L}({}^\sharp\mathcal{Z}) \quad \text{and} \quad {}^\sharp b^n \in {}^\sharp\mathcal{Z}, && \forall n \in [\![0; N-1]\!]. \end{aligned}$$

Considering the problem Problem I.1, we precise these augmented variables and operator as

$${}^\sharp z^n = \begin{pmatrix} z^n \\ \theta^n \end{pmatrix}, \quad {}^\sharp\Phi_{n+1|n}(\theta^n) = \begin{pmatrix} \Phi_{n+1|n}(\theta^n) & 0 \\ 0 & 0 \end{pmatrix}, \quad {}^\sharp b^n = \begin{pmatrix} b^n \\ 0 \end{pmatrix}.$$

such that (3.10) becomes

$${}^\sharp z^{n+1} = {}^\sharp\Phi_{n+1|n}(\theta^n)\, {}^\sharp z^n + {}^\sharp b^n, \quad \forall n \in [\![0; N-1]\!], \tag{3.31}$$

where ${}^\sharp z^0$ is given. The observations are obtained using the augmented observation operator ${}^\sharp C$ as

$$y^n = {}^\sharp C({}^\sharp z^n) + \chi^n = C(z^n) + \chi^n, \quad \forall n \in [\![0; N]\!]. \tag{3.32}$$

With the augmented discrete state and operator, we can write the parameter reconstruction problem as below.

Nonlinear discrete problem II.1

1. **Functional**:

$$\mathfrak{J}_N(\zeta_\theta) = \left\{ r_\Theta(\zeta_\theta) + \frac{\Delta t}{2} \sum_{n=0}^{N} \| y^n - {}^\sharp C({}^\sharp z_\theta^n) \|_{\mathcal{Y}}^2 \right\} \tag{3.33}$$

2. **Dynamics**: (3.31) with initial conditions bearing uncertainty, namely,

$$\left\{ \begin{aligned} & {}^\sharp z_\theta^{n+1} = {}^\sharp\Phi_{n+1|n}({}^\sharp z_\theta^n) + {}^\sharp b^n, \qquad n \in [\![0; N-1]\!], \\ & {}^\sharp z_\theta^0 = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta \end{pmatrix}. \end{aligned} \right. \tag{3.34}$$

3. **Available data**

   An initial guess θ_0 and noisy observations $\{y^n\}_{n=0}^N$ from (3.32) with ${}^\sharp z(t)$ generated from (3.31) with

$$ {}^\sharp z^0 = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta{}^* \end{pmatrix}.$$

The objective is to estimate the parameter-only error $\zeta_\theta{}^*$ through the minimization of (3.33).

Linear case. We can do the same for the linear case Problem B.2 by using the operator

$$\Phi^L{}_{n+1|n} \in \mathcal{L}({}^\sharp\mathcal{Z}), \quad \Phi^L{}_{n+1|n} = \begin{pmatrix} \Phi_{n+1|n} & B^n \\ 0 & I \end{pmatrix}, \quad \forall n \in [\![0; N-1]\!].$$

Then the linear version can be written as below in analogy to Problem II.1 by changing the operator and considering no other source term, ${}^\sharp b^n = 0$, yielding

$${}^\sharp z^{n+1} = \Phi^L{}_{n+1|n}(\theta^n)\, {}^\sharp z^n, \quad \forall n \in [\![0; N-1]\!]. \tag{3.35}$$

Linear-quadratic discrete problem II.2

1. **Functional**:

$$\mathfrak{J}_N(\zeta_\theta) = \left\{ r_\Theta(\zeta_\theta) + \frac{\Delta t}{2} \sum_{n=0}^{N} \| y^n - {}^\sharp C({}^\sharp z_\theta^n) \|_{\mathcal{Y}}^2 \right\} \tag{3.36}$$

2. **Dynamics**: (3.31) with initial conditions bearing uncertainty, namely,

$$\left\{ \begin{aligned} & {}^\sharp z_\theta^{n+1} = \Phi^L{}_{n+1|n}\, {}^\sharp z_\theta^n, \qquad \forall n \in [\![0; N-1]\!], \\ & {}^\sharp z_\theta^0 = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta \end{pmatrix}. \end{aligned} \right. \tag{3.37}$$

3. **Available data**

   An initial guess θ_0 and noisy observations $\{y^n\}_{n=0}^N$ from (3.32) with ${}^\sharp z^n$ generated from (3.35) with

$$ {}^\sharp z^0 = \begin{pmatrix} z_0 \\ \theta_0 + \zeta_\theta{}^* \end{pmatrix}.$$

The objective is to estimate the parameter-only error $\zeta_\theta{}^*$ through the minimization of (3.36).

3.2 Variational methods

In the previous section, we presented the least-squares minimization formalism related to the inverse problems in their time-continuous and time-discrete forms. In this section, we present the variational methods for solving such minimization problems [Tarantola, 1984]. The variational methods are a class of methods based on finding the minima iteratively by using the functional gradient. Variational methods iteratively update the unknown using all the system's available observations, as computing the gradient requires all data at once. The gradient is used for updating the unknown towards a better estimation.

Variational methods are well known for solving large-scale ill-posed non-linear optimization problems. Among the usages of such methods, it is used in geophysics [Mora, 1987; Pica, Diet, and Tarantola, 1990; Tarantola, 1984; Virieux and Operto, 2009; Aghamiry et al., 2022], meteorology [Dimet and Talagrand, 1986; Courtier, Thépaut, and Hollingsworth, 1994; Rabier, Thépaut, and Courtier, 1998], structural mechanics [Feissel and Allix, 2007; Bonnet and Aquino, 2015; Nguyen, Chamoin, and Ha Minh, 2022], nondestructive testing [Rao, Ratassepp, and Fan, 2016; Xu et al., 2021; Ratassepp et al., 2021], medical imaging [Agudo et al., 2018; Guasch et al., 2020; Lucka et al., 2021], among others. Although we focus here in the L^2-norm, the misfit term in the function can be represented by different norms, for instance the L^1-norm [Ma et al., 2020] or the Wasserstein norm [Engquist, Froese, and Yang, 2016], from optimal transport [Métivier et al., 2022]. To compute the functional gradient, the adjoint method is mostly used [Cea, 1986; Lions, 1971; Plessix, 2006].

In Section 3.2.1 we introduce the traditional descent methods while introducing the notation and formalism. The textbooks [Nocedal and Wright, 2006; Luenberger and Ye, 2008] are used as the main references throughout this section. In Section 3.2.2 we show how the functional gradient constrained to a given dynamics can be obtained through the adjoint method and apply it with descent methods to the problems of interest.

3.2.1 Descent methods

In the context of optimization by least-squares minimization, the Descent Methods are a well-known class of methods to minimize functionals using the functional's gradient. They consist of iteratively updating the parameter in a specific direction that decreases the functional. Here, we present such methods and some of the most used versions and ways of computing this specific direction. We recall the unconstrained inverse problem as the minimization of the functional (3.1),

$$\mathcal{J}(\theta) = \left\{ r_\Theta(\theta) + \frac{1}{2} \left\| y - \Psi(\theta) \right\|^2_{\mathcal{Y}} \right\},$$

with $\theta \in \Theta$, $\Psi : \Theta \mapsto \mathcal{Y}$ and $y \in \mathcal{Y}$. The operator Ψ is considered to be twice-differentiable in terms of the *Fréchet* differentiation. In this section, r_Θ will have the form

$$r_\Theta(\theta) = \frac{1}{2}(R_\Theta \theta, \theta)_\Theta,$$

with R_Θ being a self-adjoint positive linear operator. The descent methods are characterized by updating the parameter as [Nocedal and Wright, 2006]

$$\theta^{k+1} = \theta^k + \alpha^k d^k, \tag{3.38}$$

where $\theta^k \in \Theta$ is the parameter at the k-th step, $\alpha^k > 0$ is the descent step and $d^k \in \Theta$ is the descent direction that decreases the functional $\mathcal{J}$. The direction d^k can be obtained by different methods that will be presented next. The descent step can be computed from a line search problem as

$$\alpha^k = \operatorname*{argmin}_{\alpha > 0} \mathcal{J}(\theta^k + \alpha d^k)$$

meaning that we look for a descent step that minimizes the functional when updating the parameter in the direction of d^k. If the minimum is attained, we call it an exact line search while if it is not attained we call it an inexact line search. Using Definition A.2.7, we can develop the gradient of the functional as

$$\mathrm{D}\mathcal{J}(\theta)\widetilde{\theta} = (R_\Theta\theta, \widetilde{\theta})_\Theta + \big(\mathrm{D}\Psi(\theta)\widetilde{\theta}, \Psi(\theta) - y\big)_{\mathcal{Y}}, \quad \forall\widetilde{\theta} \in \Theta, \tag{3.39}$$

where $\mathrm{D}\mathcal{J}(\theta) \in \mathcal{L}(\Theta, \mathbb{R})$ and $\mathrm{D}\Psi(\theta) \in \mathcal{L}(\Theta, \mathcal{Y})$ are the functional and the "input-output" operator differentiations, respectively. By using Property A.2.1, we denote as $\nabla\mathcal{J}$ the Riesz representation of $\mathrm{D}\mathcal{J}$, then

$$\langle \mathrm{D}\mathcal{J}(\theta), \widetilde{\theta}\rangle_{\Theta'\Theta} = (\nabla\mathcal{J}, \widetilde{\theta})_\Theta, \quad \forall\widetilde{\theta} \in \Theta.$$

In the following, we may use $\nabla\mathcal{J}$ without writing the norm in which the scalar product should be taken. It means that, we assume a discrete norm $(\theta, \theta')_\Theta = \theta^\intercal\theta'$, $\forall\theta, \theta' \in \Theta$. When it is not the case, for instance, where Θ is a finite element space and a L^2 norm is associated with it, the norm should be present

$$(\nabla\mathcal{J}, \widetilde{\theta})_\Theta = \widetilde{\theta}^\intercal \mathbb{M}\, \nabla\mathcal{J}, \quad \forall\widetilde{\theta} \in \Theta,$$

where the application of the mass matrix $\mathbb{M}$, as seen in Chapter 1, results in the L^2 norm.

3.2.1.1 Steepest descent

The steepest descent (SD) method is a descent method where we choose the direction d^k in which the iteration (3.38) directly satisfies the descent condition

$$\mathcal{J}(\theta^{k+1}) \leq \mathcal{J}(\theta^k)$$

when considering a first-order expansion. By doing the Taylor expansion in $\mathcal{J}$ up to the first-order term, where $\mathrm{D}\mathcal{J} \in \mathcal{L}(\Theta, \mathbb{R})$, we have

$$\mathcal{J}(\theta^{k+1}) = \mathcal{J}(\theta^k + \alpha^k d^k) \approx \mathcal{J}(\theta^k) + \alpha^k \mathrm{D}\mathcal{J}(\theta^k) d^k,$$

then, to satisfy the descent condition $\forall\alpha^k > 0$,

$$\begin{aligned} \mathcal{J}(\theta^k) + \alpha^k \mathrm{D}\mathcal{J}(\theta^k) d^k &\leq \mathcal{J}(\theta^k) \\ \mathrm{D}\mathcal{J}(\theta^k) d^k &\leq 0, \end{aligned}$$

one can choose the direction

$$d^k = -\nabla\mathcal{J}(\theta^k) = -\nabla\mathcal{J}^{(k)}, \tag{3.40}$$

where we introduce the notation $\nabla\mathcal{J}(\theta^k) = \nabla\mathcal{J}^{(k)}$. To exemplify the effect of a chosen norm in Θ, in the aforementioned case of a L^2 norm, we have

$$d^k = -\mathbb{M}\, \nabla\mathcal{J}^{(k)}.$$

The expected order of convergence of the steepest descent method being of the first-order [Nocedal and Wright, 2006; Luenberger and Ye, 2008], we present the following methods with an expected better order of convergence.

3.2.1.2 Conjugate gradient

In conjugate gradient (CG) methods, the direction d^k is defined as

$$d^k = \begin{cases} -\nabla\mathcal{J}^{(k)} & \text{for } k = 0 \\ -\nabla\mathcal{J}^{(k)} + \beta^k d^{k-1} & \text{for } k > 0 \end{cases}. \tag{3.41}$$

where β^k is to be choosen. One example of choice of coefficient is the Fletcher-Reeves β^k coefficient [Luenberger and Ye, 2008]

$$\beta^k = \frac{\left\|\nabla\mathcal{J}^{(k+1)}\right\|_\Theta^2}{\left\|\nabla\mathcal{J}^{(k)}\right\|_\Theta^2}, \quad \forall k > 0.$$

Other choices of β^k are possible [Nocedal and Wright, 2006; Nakayama, 2019]. The descent step α^k is obtained through a line search.

3.2.1.3 Newton-Raphson

The Newton-Raphson (NR) method provides a better convergence than the aforementioned methods, where the d^k is defined as

$$d^k = (H^{(k)})^{-1}\nabla\mathcal{J}^{(k)}, \tag{3.42}$$

where the Hessian operator $H^{(k)}$ emerges from the second (*Fréchet*) differentiation of the functional at θ^k as defined in Definition A.2.8. We assume the Hessian to be invertible. Differentiating twice the functional yields, $\forall\widetilde{\theta},\widetilde{\theta}' \in \Theta$,

$$\begin{aligned} \mathrm{D}^2\mathcal{J}(\theta^k)(\widetilde{\theta},\widetilde{\theta}') &= (R_\Theta\widetilde{\theta}',\widetilde{\theta})_\Theta + \left(\mathrm{D}\Psi^{(k)}\widetilde{\theta},\mathrm{D}\Psi^{(k)}\widetilde{\theta}'\right)_{\mathcal{Y}} + \left(\left[\mathrm{D}^2\Psi(\theta^k)\widetilde{\theta}\right]\widetilde{\theta}',\Psi^{(k)} - y\right)_{\mathcal{Y}} \\ &= [H^{(k)}\widetilde{\theta}]\widetilde{\theta}', \end{aligned} \tag{3.43}$$

where we make explicit the Hessian and introduce the compact notations: $\Psi^{(k)} = \Psi(\theta^k)$, $\mathrm{D}\Psi^{(k)} = \mathrm{D}\Psi(\theta^k)$ and $\mathrm{D}\Psi^{(k)*} = \mathrm{D}\Psi^*(\theta^k)$. It is possible to prove that the Newton-Raphson method, in a particular case where $\mathcal{J}$ is convex, benefits from an order-two convergence towards the global minimum, see [Luenberger and Ye, 2008] for more details. Although it has the advantage of converging at a better rate than steepest-descent and conjugate-gradient, differentiating twice the cost function means differentiating twice the operator A *w.r.t.* the parameter. Usually, this differentiation is not straightforward to compute and, for some large cases, the amount of required memory may be a limiting factor as it, *a priori*, does not result in a sparse matrix.

3.2.1.4 Quasi-Newton methods

Considering limiting factors of computing the Hessian, approximated Hessians are used in the so-called quasi-Newton methods. For presenting the methods, we will denote the approximation of the Hessian $\widetilde{H}$ and its inverse as $\widetilde{B}$.

Gauss-Newton. In the Gauss-Newton method we suppose that the second-derivative term in (3.43) is small or negligible, so the Hessian is approximated as

$$\begin{aligned} [H^{(k)}\widetilde{\theta}]\widetilde{\theta}' \approx [\widetilde{H}^{(k)}\widetilde{\theta}]\widetilde{\theta}' &= (R_\Theta\widetilde{\theta}',\widetilde{\theta})_\Theta + \left(\mathrm{D}\Psi^{(k)}\widetilde{\theta},\mathrm{D}\Psi^{(k)}\widetilde{\theta}'\right)_{\mathcal{Y}} \\ &= (R_\Theta\widetilde{\theta}',\widetilde{\theta})_\Theta + \left(\mathrm{D}\Psi^{(k)*}\mathrm{D}\Psi^{(k)}\widetilde{\theta}',\widetilde{\theta}\right)_\Theta \\ &= \left((R_\Theta + \mathrm{D}\Psi^{(k)*}\mathrm{D}\Psi^{(k)})\widetilde{\theta}',\widetilde{\theta}\right)_\Theta. \end{aligned} \tag{3.44}$$

Iterated Gauss-Newton. Using this approximation, we can show that each descent iteration (3.38) can be interpreted as a linearized minimization problem. To do so, we first develop the iterations for the Newton-Raphson Method using the Gauss-Newton approximation and (3.39),

$$\begin{aligned} \theta^{k+1} &= \theta^k - \alpha_k(\widetilde{H}^{(k)})^{-1}\nabla\mathcal{J}^{(k)}. \\ &= \theta^k - \alpha_k\left(R_\Theta + \mathrm{D}\Psi^{(k)*}\mathrm{D}\Psi^{(k)}\right)^{-1}\left(R_\Theta\theta^k + \mathrm{D}\Psi^{(k)}(\Psi^{(k)} - y)\right). \end{aligned}$$

By defining $\delta y = y - \Psi^{(k)}$ and $\delta\theta = (\theta^{k+1} - \theta^k)/\alpha_k$, we have

$$\delta\theta = \left(R_\Theta + \mathrm{D}\Psi^{(k)*}\mathrm{D}\Psi^{(k)}\right)^{-1}\left(R_\Theta\theta^k + \mathrm{D}\Psi^{(k)}\delta y\right)$$

that is the closed-form minimum of the a new functional $\mathfrak{J}^k$ with the linearized operator

$$\min_{\delta\theta\in\Theta} \mathfrak{J}^k(\delta\theta) = \min_{\delta\theta\in\Theta}\left\{r_\Theta(\theta^k) + \frac{1}{2}\|\delta y - \mathrm{D}\Psi^{(k)}\delta\theta\|^2_{\mathcal{Y}}\right\}.$$

The iterated version of the Gauss-Newton methods is useful for wave propagation inverse problems as the assembly of the "input-output" tangent operator is prohibitive and is represented by a wave propagation problem. Then, using tangent dynamics at each sub-iteration, one can compute the gradient to minimize the "linearized" functional.

BFGS (Broyden-Fletcher-Goldfarb-Shanno). Using the fact that the Hessian evolution through descent iterations is associated with the evolution of the functional gradient, the BFGS method consists in generating an approximation of the Hessian at each descent step using only the functional gradient and the parameter itself [Luenberger and Ye, 2008]. Starting from an *a priori* Hessian, it is updated at each descent step by using the gradient and parameter innovations. As the direction (3.42) is used, we are more interested in approximating the Hessian inverse directly than approximating the Hessian and then inversing it. We present here the BFGS version in which the inverse of the Hessian is estimated, avoiding solving a linear system at each iteration. We present the BFGS algorithm to compute the approximated Hessian inverse $\widetilde{B}$.

$$\widetilde{B}^{k+1} = V^{(k)}\widetilde{B}^k V^{(k)*} + \rho^{(k)} y^{(k)*} y^{(k)}, \tag{3.45}$$

where

$$\rho^{(k)} = \frac{1}{y^{(k)} s^{(k)}}, \qquad V^{(k)} = I - \rho^{(k)} s^{(k)} y^{(k)},$$

and

$$s^{(k)} = \theta^{k+1} - \theta^k, \qquad y^{(k)} = \mathrm{D}\mathfrak{J}^{(k+1)} - \mathrm{D}\mathfrak{J}^{(k)},$$

where

$$\widetilde{B}^k \in \mathcal{L}(\Theta), \qquad s^{(k)} \in \Theta, \qquad y^{(k)} \in \mathcal{L}(\Theta, \mathbb{R}) \quad \text{and} \quad V^{(k)} \in \mathcal{L}(\Theta).$$

BFGS: discrete case. Considering the parametric space $\Theta \in \mathbb{R}^{N_\theta}$ with norm $\|\theta\|^2_\Theta = \theta^\top\theta, \forall\theta \in \Theta$, the update algorithm is written as

$$\widetilde{\mathbb{B}}^{(k+1)} = \mathbb{V}^{(k)}\widetilde{\mathbb{B}}^{(k)}\mathbb{V}^{(k)\top} + \rho^{(k)} y^{(k)} y^{(k)\top}, \tag{3.46}$$

where

$$\rho^{(k)} = \frac{1}{y^{(k)\top} s^{(k)}}, \qquad \mathbb{V}^{(k)} = (\mathbb{I} - \rho^{(k)} s^{(k)} y^{(k)\top}),$$

and

$$s^{(k)} = \theta^{k+1} - \theta^k, \qquad y^{(k)} = \nabla_\theta\mathfrak{J}^{(k+1)} - \nabla_\theta\mathfrak{J}^{(k)}.$$

The *a priori* Hessian inverse, namely $\widetilde{\mathbb{B}}^{(0)}$, must be given. Usually we start with $\widetilde{\mathbb{B}}^{(0)} = \mathbb{R}_\Theta^{-1}$, $s^{(0)} = -\nabla_\theta\mathfrak{J}^{(0)}$, then $\nabla_\theta\mathfrak{J}^{(1)} = \nabla_\theta\mathfrak{J}(\theta^0 + s^{(0)})$.

L-BFGS: discrete case. In cases where the size of parameter space makes storing the Hessian matrix a limiting factor, the Limited-Memory BFGS (L-BFGS) proposes to store vectors used to update the Hessian instead of the matrices themselves. As this matrix is updated iteratively as in (3.45), it can be rewritten as

$$\begin{aligned}\widetilde{\mathbb{B}}^{(k)} = {} & (\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{(0)\intercal})\widetilde{\mathbb{B}}^{(0)}(\mathbb{V}^{(0)} \dots \mathbb{V}^{(k-1)}) \\ & + \rho^{(0)}(\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{(1)\intercal})s^{(0)}s^{(0)\intercal}(\mathbb{V}^{(1)} \dots \mathbb{V}^{(k-1)}) \\ & + \rho^{(1)}(\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{2\intercal})s^{(1)}s^{(1)\intercal}(\mathbb{V}^{2} \dots \mathbb{V}^{(k-1)}) \\ & + \dots + \\ & + \rho^{(k-1)}s^{(k-1)}s^{(k-1)\intercal}.\end{aligned}$$

Using this formulation, instead of storing the approximation of the Hessian matrix and the auxiliary matrices $\mathbb{V}$, we can store only a set of vector pairs$\{s^{(i)}, y^{(i)}\}_{i=k-m}^{k-1}$ where m is the number of stored pairs. For instance, if $m = k$ we always store all vectors, if $m = 3$ we store only the last three vectors. The limited-memory approximation becomes

$$\begin{aligned}\widetilde{\mathbb{B}}^{(k)} = {} & (\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{(k-m)\intercal})\widetilde{\mathbb{B}}^{0(k)}(\mathbb{V}^{(k-m)} \dots \mathbb{V}^{(k-1)}) \\ & + \rho^{(k-m)}(\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{(k-m+1)\intercal})s^{(k-m)}s^{(k-m)\intercal}(\mathbb{V}^{(k-m+1)} \dots \mathbb{V}^{(k-1)}) \\ & + \rho^{(k-m+1)}(\mathbb{V}^{(k-1)\intercal} \dots \mathbb{V}^{(k-m+2)\intercal})s^{(k-m+1)}s^{(k-m+1)\intercal}(\mathbb{V}^{(k-m+2)} \dots \mathbb{V}^{(k-1)}) \\ & + \dots + \\ & + \rho^{(k-1)}s^{(k-1)}s^{(k-1)\intercal}.\end{aligned}$$

The identity matrix or another set of previously stored vectors can be used as the *a priori* $\widetilde{\mathbb{B}}^{0(k)}$.

Memoryless quasi-Newton. In the case where $\widetilde{\mathbb{B}}^{0(k)} = \mathbb{I}$ and $m = 1$, no extra vectors or matrices are kept in memory and the iteration (3.38) is done directly using $s^{(k)}$ and $y^{(k)}$. This method has a similar memory usage and computational cost as the Conjugate Gradient while performing slightly better, as shown in [Liu et al., 2015].

3.2.2 Functional gradient for time-dependent systems

We presented the above descent methods to minimize a functional using its gradient and/or Hessian. To complete its implementation for the problems of interest, we discuss here how to obtain the functional gradient and Hessian for time-dependent systems such as the ones presented in Section 3.1. By redefining the "input-output" operator Ψ, including a time dependency, the cost function takes the form of $\mathfrak{J}_T$

$$\mathfrak{J}_T(\theta) = \left\{ r_\Theta(\theta) + \frac{1}{2}\int_0^T \|y(t) - \Psi(\theta)(t)\|^2_{\mathcal{Y}} \, \mathrm{d}t \right\}.$$

First, in Section 3.2.2.1 we present a direct and costly method for computing both gradient and Hessian using the sensitivity and cross-sensitivity solutions. We present them for the initial state reconstruction problem. In Section 3.2.2.2 a more efficient method to compute the gradient is presented, the adjoint method, for the time-continuous case. In Section 3.2.2.3, the adjoint method is presented for the fully discrete case.

3.2.2.1 Sensitivity and cross-sensitivity approach for initial state reconstruction

We focus here on computing the gradient and Hessian of the functional representing the Problem A.1 by using sensitivity and cross-sensitivity, respectively. The sensitivity can be used to compute the functional gradient. The gradient of the functional in (3.16) can be written as

$$\mathrm{D}\mathcal{J}_T(\zeta)\tilde{\zeta} = (R\zeta,\tilde{\zeta})_{\mathcal{Z}} + \int_0^T \Big(\mathrm{D}C S_\zeta(t)\tilde{\zeta}, C\big(z_\zeta(t)\big) - y(t)\Big)_{\mathcal{Y}}\,\mathrm{d}t \tag{3.47}$$

where the sensitivity trajectory $S_\zeta(t) = \mathrm{D}_\zeta z_\zeta(t)$, $S_\zeta(t) \in \mathcal{L}(\mathcal{Z})$ is introduced below and $\mathrm{D}C = \mathrm{D}C(z_\zeta(t)) \in \mathcal{L}(\mathcal{Z},\mathcal{Y})$. The sensitivity is obtained differentiating the dynamics (3.17):

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}S_\zeta(t) = \mathrm{D}A(z)S_\zeta(t), \quad \forall t \in [0,T], \\[2ex] S_\zeta(0) = I, \end{cases} \tag{3.48}$$

for $\mathrm{D}A(z) \in \mathcal{L}(\mathcal{Z})$. Differentiating twice, the Hessian of the functional is

$$\begin{aligned} \mathrm{D}^2_\zeta\mathcal{J}_T(\zeta)(\tilde{\zeta},\tilde{\zeta}') = (R\tilde{\zeta}',\tilde{\zeta})_{\mathcal{Z}} + \int_0^T &(\mathrm{D}C S_\zeta(t)\tilde{\zeta}, \mathrm{D}C S_\zeta(t)\tilde{\zeta}')_{\mathcal{Y}} \\ &+ \Big(\mathrm{D}C[W_\zeta(t)\tilde{\zeta}]\tilde{\zeta}', C\big(z_\zeta(t)\big) - y(t)\Big)_{\mathcal{Y}}\,\mathrm{d}t. \end{aligned}$$

where we define the cross-sensitivity trajectory $W_\zeta(t) = \mathrm{D}^2_\zeta z_\zeta(t) \in \mathcal{L}(\mathcal{Z},\mathcal{L}(\mathcal{Z}))$. We obtain the cross-sensitivity by differentiating twice (3.17):

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}[W_\zeta(t)\tilde{\zeta}]\tilde{\zeta}' = \big[\mathrm{D}^2A\big(z_\zeta(t)\big)(S_\zeta(t)\tilde{\zeta})\big](S_\zeta(t)\tilde{\zeta}') + \mathrm{D}A\big(z_\zeta(t)\big)\big[W_\zeta(t)\tilde{\zeta}\big]\tilde{\zeta}', \qquad \forall t \in [0,T], \\[2ex] [W_\zeta(0)\tilde{\zeta}]\tilde{\zeta}' = 0. \end{cases}$$

for the second-order tangent operator $\mathrm{D}^2A\big(z_\zeta(t)\big) \in \mathcal{L}(\mathcal{Z},\mathcal{L}(\mathcal{Z}))$.

Sensitivity and cross-sensitivity in the finite-dimensional cases. The computation of the gradient using sensitivity means that, on top of the original forward dynamics to be solved, other N_z forward problems need to be solved when $\zeta \in \mathbb{R}^{N_z}$. As $(W_\zeta)_{ij}$ is symmetric in i, j, the number of extra problems to be solved to compute the Hessian are $\frac{N_z(N_z+1)}{2}$ for the cross-sensitivity. The number of problems to be solved is then proportional to the square of the size of the unknown. This high computational cost motivates the use of the adjoint method presented in the next subsection.

3.2.2.2 Time-continuous adjoint method

In practice, computing the gradient by using the sensitivity requires several direct problems to be launched proportional to the size of the parametric space. Here, we present a better method for computing the functional gradient by adding only one extra problem to be solved: the adjoint dynamics. We develop it for the time-continuous case, Problem A.1. The gradient (3.47) can be computed as below.

Property 3.2.1. *Considering a adjoint trajectory $p(t) \in \mathcal{Z}$ as*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}p(t) + \mathrm{D}A\big(z_\zeta(t)\big)^* p(t) = \mathrm{D}C\big(z_\zeta(t)\big)^*\Big(C\big(z_\zeta(t)\big) - y(t)\Big), \quad \forall t \in [0;T], \\[2ex] p(T) = 0, \end{cases} \tag{3.49}$$

the functional gradient of Problem A.1 can be expressed as

$$\mathrm{D}\mathcal{J}_T(\zeta) = R\zeta - p(0).$$

Proof. Replacing the adjoint trajectory in (3.47) we have

$$\begin{aligned}
\mathrm{D}\mathcal{J}_T(\zeta)\widetilde{\zeta} &= (R\zeta,\widetilde{\zeta})_{\mathcal{Z}} + \int_0^T \Big(\mathrm{D}_\zeta z_\zeta(t)\widetilde{\zeta}, \mathrm{D}C\big(z_\zeta(t)\big)^*\big(C(z_\zeta(t)) - y(t)\big)\Big)_{\mathcal{Z}}\,\mathrm{d}t \\
&= (R\zeta,\widetilde{\zeta})_{\mathcal{Z}} + \int_0^T \Big(\mathrm{D}_\zeta z_\zeta(t)\widetilde{\zeta}, \frac{\mathrm{d}}{\mathrm{d}t}p(t) + \mathrm{D}A\big(z_\zeta(t)\big)^*\,p(t)\Big)_{\mathcal{Z}}\,\mathrm{d}t \\
&= (R\zeta,\widetilde{\zeta})_{\mathcal{Z}} + \int_0^T \Big(\mathrm{D}_\zeta z_\zeta(t)^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}p(t) + \mathrm{D}A\big(z_\zeta(t)\big)^*\,p(t)\Big), \widetilde{\zeta}\Big)_{\mathcal{Z}}\,\mathrm{d}t,
\end{aligned}$$

and, as it is valid for any $\widetilde{\zeta}$, we have

$$\nabla\mathcal{J}_T(\zeta) = R\zeta + \int_0^T \mathrm{D}_\zeta z_\zeta(t)^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}p(t) + \mathrm{D}A\big(z_\zeta(t)\big)^*\,p(t)\Big)\,\mathrm{d}t$$

Using integration by parts, the definition of the sensitivity (3.48) and its initial condition, the gradient becomes

$$\begin{aligned}
\nabla\mathcal{J}_T(\zeta) &= R\zeta + \int_0^T \mathrm{D}_\zeta z_\zeta(t)^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}p(t)\Big) + \mathrm{D}_\zeta z_\zeta(t)^*\mathrm{D}A\big(z_\zeta(t)\big)^*p(t)\,\mathrm{d}t \\
&= R\zeta + \int_0^T \mathrm{D}_\zeta z_\zeta(t)^*\Big(\frac{\mathrm{d}}{\mathrm{d}t}p(t)\Big) + \Big(\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{D}_\zeta z_\zeta(t)\big)^*\Big)p(t)\,\mathrm{d}t. \\
&= R\zeta + \Big\{\big[\big(\mathrm{D}_\zeta z_\zeta(t)\big)^*p(t)\big]_0^T - \int_0^T \Big(\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{D}_\zeta z_\zeta(t)\big)^*\Big)p(t)\,\mathrm{d}t + \int_0^T \Big(\frac{\mathrm{d}}{\mathrm{d}t}\big(\mathrm{D}_\zeta z_\zeta(t)\big)^*\Big)p(t)\,\mathrm{d}t\Big\} \\
&= R\zeta - p(0).
\end{aligned}$$

□

Two-end problem By seeking $\nabla\mathcal{J}_T(\zeta) = 0$ to satisfy a first-order optimality condition, the state dynamics is coupled with the adjoint dynamics, yielding a two-end problem [Lions, 1971; Bensoussan et al., 2007]

$$\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{d}t}\overline{z}(t) = A(\overline{z}(t),t) + b(t), & \forall t \in [0;T], \\[2ex]
\dfrac{\mathrm{d}}{\mathrm{d}t}\overline{p}(t) + \mathrm{D}A(\overline{z}(t),t)^*\,\overline{p}(t) = \mathrm{D}C\big(\overline{z}(t)\big)^*\Big(C\big(\overline{z}(t)\big) - y(t)\Big), & \forall t \in [0;T], \\[2ex]
\overline{p}(T) = 0, \quad \overline{z}(0) = z_0 + R^{-1}\,\overline{p}(0).
\end{cases} \tag{3.50}$$

Through the two-end problem, we obtain the optimally estimated initial condition $\overline{z}(0)$. The name two-end comes from the fact that we have initial conditions for one trajectory and final conditions for the other one, while they are coupled. It couples the forward (state) dynamics and the backward (adjoint) dynamics. Although numerical methods are proposed to solve this class of problems, such as iterative strategies or space-time finite elements [Cîndea and Münch, 2015], their cost is prohibitive for the large wave propagation problems we aim for in this thesis. The difficulty posed to solve this type of coupled forward-backward problem makes the adjoint dynamics more useful in computing the gradient, as in Property 3.2.1, for iterative methods.

Application to parameter reconstruction. Extending the system as done in Problem B.1, we can obtain the augmented adjoint dynamics for computing its functional gradient as well as for Problem B.2.

In Section 3.4.1 we exemplify the use of the adjoint dynamics for solving a problem of the same type as Problem II.2.

3.2.2.3 Time-discrete adjoint method

We will now look at the discrete version of adjoint dynamics for computing the functional gradient in initial state reconstruction Problem I.1. We do it by first getting the optimality condition through a two-end problem defined below [Moireau, 2022].

> **Property 3.2.2.** *Introducing the discrete adjoint dynamics, the discrete version of the two-end problem is*
>
> $$\begin{cases} \overline{z}^{n+1} = \Phi_{n+1|n}(\overline{z}^n) + b^n, & \forall n \in [\![0; N-1]\!], \\ \overline{p}^n - \mathrm{D}\Phi_{n+1|n}(\overline{z}^n)^* \, \overline{p}^{n+1} = \Delta t \mathrm{D}C(\overline{z}^n)^*(y^n - C(\overline{z}^n)), & \forall n \in [\![N; 0]\!], \\ \overline{p}^{N+1} = 0, \quad \overline{z}^0 = z_0 + R^{-1}\overline{p}^0. \end{cases} \tag{3.51}$$

Proof. Using the functional of Problem I.1, the discrete adjoint dynamics can be obtained by defining the Lagrangian

$$\mathcal{L}_N\left(\{z_\zeta^n\}_{n=0}^{N-1}, \{p_\zeta^n\}_{n=1}^{N}\right) = \mathcal{J}_N(\zeta) + \sum_{n=0}^{N-1} \left(z_\zeta^{n+1} - \Phi_{n+1|n}(z_\zeta^n) - b^n, p_\zeta^{n+1}\right)_{\mathcal{Z}},$$

where $\{p_\zeta^n\}_{n=1}^N$ is the time-discrete adjoint variable. Satisfying the optimality conditions requires differentiating the Lagrangian by the adjoint and state variables. To do so, we separate the differentiation by

Differentiating w.r.t. adjoint variable at $n \in [\![1; N]\!]$:

$$\mathrm{D}_{p_\zeta^n}\mathcal{L}_N \widetilde{p}^n = (z_\zeta^n - \Phi_{n|n-1}(z_\zeta^{n-1}) - b^{n-1}, \widetilde{p}^n)_{\mathcal{Z}},$$

hence, satisfying the optimality conditions leads to the dynamics

$$\overline{z}^{n+1} = \Phi_{n+1|n}(\overline{z}^n) + b^n, \quad \forall n \in [\![0; N-1]\!],$$

where $\overline{z}$ is introduced as the optimal trajectory with initial conditions yet to be obtained.

Differentiating w.r.t. state variable at $n \in [\![1; N-1]\!]$:

$$\begin{aligned} \mathrm{D}_{z^n}\mathcal{L}_N \widetilde{z}^n &= -\Delta t\left(y^n - C(z_\zeta^n), \mathrm{D}C^{(n)}\widetilde{z}^n\right)_{\mathcal{Y}} + \left(p_\zeta^n, \widetilde{z}^n\right)_{\mathcal{Z}} - \left(p_\zeta^{n+1}, \mathrm{D}\Phi_{n+1|n}^{(n)}\widetilde{z}^n\right)_{\mathcal{Z}} \\ &= -\left(\Delta t(\mathrm{D}C^{(n)})^*\left(y^n - C(z_\zeta^n)\right), \widetilde{z}^n\right)_{\mathcal{Z}} + \left(p_\zeta^n - (\mathrm{D}\Phi_{n+1|n}^{(n)})^* p_\zeta^{n+1}, \widetilde{z}^n\right)_{\mathcal{Z}} \end{aligned}$$

where $\mathrm{D}C^{(n)} = \mathrm{D}C(z_\zeta^n)$ and $\mathrm{D}\Phi_{n+1|n}^{(n)} = \mathrm{D}\Phi_{n+1|n}(z_\zeta^n)$, leading to the adjoint optimal discrete trajectory

$$\overline{p}^n - (\mathrm{D}\Phi_{n+1|n}^{(n)})^* \, \overline{p}^{n+1} = \Delta t(\mathrm{D}C^{(n)})^*\left(y^n - C(z_\zeta^n)\right). \quad \forall n \in [\![1; N-1]\!].$$

Differentiating w.r.t. state variable at $n = N$:

$$\begin{aligned}\mathrm{D}_{z_\zeta^n}\mathcal{L}_N\widetilde{z}^N &= -\Delta t\big(y^N - C(z_\zeta^N), \mathrm{D}C^{(N)}\widetilde{z}^N\big)_{\mathcal{Y}} + \big(p_\zeta^N, \widetilde{z}^N\big)_{\mathcal{Z}}\\ &= -\Delta t\Big((\mathrm{D}C^{(n)})^*\big(y^N - C(z_\zeta^N)\big), \widetilde{z}^N\Big)_{\mathcal{Z}} + \big(p_\zeta^N, \widetilde{z}^N\big)_{\mathcal{Z}}\end{aligned}$$

This yields $\overline{p}^N = \Delta t(\mathrm{D}C^{(n)})^*(y^N - C(z_\zeta^N))$ and, to satisfy the equation of the adjoint optimal discrete trajectory for $n = N$, we impose

$$\overline{p}^{N+1} = 0.$$

Differentiating w.r.t. state variable at $n = 0$: Recalling that that $\zeta = z_\zeta^0 - z_0$,

$$\begin{aligned}\mathrm{D}_{z_\zeta^0}\mathcal{L}_N\widetilde{z}^0 &= \big(R(z_\zeta^0 - z_0), \widetilde{z}^0\big)_{\mathcal{Z}} - \Delta t\big(y^0 - C(z_\zeta^0), \mathrm{D}C^{(0)}\widetilde{z}^0\big)_{\mathcal{Y}} - \big(p_\zeta^1, \mathrm{D}\Phi_{1|0}^{(0)}\widetilde{z}^0\big)_{\mathcal{Z}}\\ &= \big(R(z_\zeta^0 - z_0), \widetilde{z}^0\big)_{\mathcal{Z}} - \Delta t\Big((\mathrm{D}C^{(0)})^*\big(y^0 - C(z_\zeta^0)\big), \widetilde{z}^0\Big)_{\mathcal{Z}} - \big((\mathrm{D}\Phi_{1|0}^{(0)})^* p_\zeta^1, \widetilde{z}^0\big)_{\mathcal{Z}}.\end{aligned}$$

Naturally introducing the final step of the adjoint dynamics $\overline{p}^0$ such that

$$\overline{p}^0 - (\mathrm{D}\Phi_{1|0}^{(0)})^*\,\overline{p}^1 = \Delta t(\mathrm{D}C^{(0)})^*\big(y^0 - C(z_\zeta^0)\big),$$

the optimality condition leads to

$$R^{-1}\overline{p}^0 = \widetilde{z}^0 - z_0.$$

The two-end problem can be then written. □

As in the time-continuous case, the time-discrete two-end problem poses the same difficulties for solving it. The adjoint dynamics is useful to compute the functional gradient as shown below, in Property 3.2.3.

Property 3.2.3. *The functional gradient can be written as*

$$\mathrm{D}\mathcal{J}_N(\zeta) = R\zeta - p_\zeta^0 \tag{3.52}$$

with p_ζ obtained from the adjoint dynamics

$$\begin{cases} p_\zeta^n - (\mathrm{D}\Phi_{n+1|n}^{(n)})^* p_\zeta^{n+1} = \Delta t(\mathrm{D}C^{(n)})^*(y^n - C\big(z_\zeta^n)\big), & \forall n \in [\![N; 0]\!],\\ p_\zeta^{N+1} = 0 \end{cases} \tag{3.53}$$

and z_ζ from Problem I.1.

Proof. From the definition of the adjoint variable dynamics, it is possible to derive a direct expression of the cost function gradient. We have, by differentiating (3.30),

$$\begin{cases} \mathrm{D}_\zeta z_\zeta^{n+1} = \mathrm{D}\Phi_{n+1|n}^{(n)} \mathrm{D}_\zeta z_\zeta^n, & \forall n \in [\![0; N-1]\!],\\ \mathrm{D}_\zeta z_\zeta^0 = I. \end{cases}$$

Differentiating the functional in (3.29) we obtain, $\forall \widetilde{\zeta} \in \mathcal{Z}$,

$$\mathrm{D}\mathcal{J}_N(\zeta)\widetilde{\zeta} = (R\zeta, \widetilde{\zeta})_{\mathcal{Z}} - \Delta t \sum_{n=0}^{N} \Big((\mathrm{D}C^{(n)})^* (y^n - C(z_\zeta^n)), \mathrm{D}_\zeta z_\zeta^n \widetilde{\zeta} \Big)_{\mathcal{Z}}.$$

The adjoint dynamics p_ζ leads to

$$\begin{aligned} \mathrm{D}\mathcal{J}_N(\zeta)\widetilde{\zeta} &= (R\zeta, \widetilde{\zeta})_{\mathcal{Z}} - \sum_{n=0}^{N} \Big(p_\zeta^n - (\mathrm{D}\Phi_{n+1|n}^{(n)})^* p_\zeta^{n+1}, \mathrm{D}_\zeta z_\zeta^n \widetilde{\zeta} \Big)_{\mathcal{Z}} \\ &= (R\zeta, \widetilde{\zeta})_{\mathcal{Z}} - \sum_{n=0}^{N} (p_\zeta^n, \mathrm{D}_\zeta z_\zeta^n \widetilde{\zeta})_{\mathcal{Z}} + \sum_{n=0}^{N} (p_\zeta^{n+1}, \mathrm{D}_\zeta z_\zeta^{n+1} \widetilde{\zeta})_{\mathcal{Z}}, \\ &= (R\zeta - p_\zeta^0, \widetilde{\zeta})_{\mathcal{Z}}. \end{aligned}$$

Finally, we have the expression for the functional gradient. □

Using the property defined above, the adjoint dynamics can then be used to obtain the functional gradient with the cost of computing one extra dynamics, hence it allows to use effectively the methods described in Section 3.2.1.

3.3 Sequential methods

In the previous section, the optimality condition for the inverse problems of interest is satisfied by solving a two-end problem as (3.50). The associated adjoint dynamics can be used to compute the gradient to be used in descent methods. In this section, we follow another approach for estimation, the sequential method. Sequential (or filtering) methods have the particularity of updating the estimation of the unknown at each data entry (or time step, for the case of dynamics), contrary to variational methods where all the data is used at each descent step. One of the advantages of sequential methods is that it allows the estimation to be continued when new data is available without recomputing the gradients using the whole data. Although the methodologies differ, we can relate sequential methods to variational methods achieving optimality in a linear-quadratic case, using the so-called Kalman Filter (KF). This motivates, in this work, the use of Kalman filters in the context of least-squares optimization.

As well as variational methods, the use of Kalman Filter in data assimilation context has been of increasing relevance. Although first conceived for real-time signal filtering and communications, its also currently used in meteorology [Jones, 1965; Courtier et al., 1993; Houtekamer and Zhang, 2016], oceanography [Pham, Verron, and Gourdeau, 1998; Evensen, 2003; Bertino, Evensen, and Wackernagel, 2003; Rozier et al., 2007], geophysics [Huang and Zhu, 2019; Eikrem, Nævdal, and Jakobsen, 2019], medical applications [Moireau, Chapelle, and Le Tallec, 2008; Chabiniok et al., 2012; Caiazzo et al., 2017], structural mechanics [Mariani and Ghisi, 2007; Marchand, Chamoin, and Rey, 2016] and NDT/SHM [Wu, Huang, and Huang, 2004; Dehghan-Niri, Farhidzadeh, and Salamone, 2013; Zou et al., 2015], to provide some examples. Here, we treat the Kalman Filter from a deterministic point of view while also potentially using terminology from a stochastic framework such as “covariance” for the Riccati operator. We refer to [Särkkä, 2013] for the Kalman Filter presentation in a stochastic framework.

In the first subsection, the main concepts and algorithm for filtering are presented. From this framework, the Kalman Filter is then presented as a link between filtering and optimality. We then present the Kalman Filter in different versions for linear and nonlinear inverse problems, such as those problems introduced in Section 3.1. Our introduction to the methods presented here are mainly based in the works done in [Simon, 2006; Bensoussan et al., 2007; Moireau, 2022].

3.3.1 Filtering methods

To introduce the filtering methods, we present a case with linear dynamics where error estimation is done by adding feedback to the dynamics proportional to the misfit between the model and observations. Let a reference linear dynamics be $z(t) \in \mathcal{Z}$

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} z(t) = Az(t) + b(t), \qquad \forall t \in [0,T], \\[2ex] z(0) = z_0, \end{cases} \tag{3.54}$$

with $A \in \mathcal{L}(\mathcal{Z})$ and $b(t) \in \mathcal{Z}$. We then define a dynamics, $\forall t \in [0,T]$ and $\widehat{z}(t) \in \mathcal{Z}$, as

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \widehat{z}(t) = A\widehat{z}(t) + b(t) + G(t)\big(y(t) - C\widehat{z}(t)\big), \qquad \forall t \in [0,T], \\[2ex] \widehat{z}(0) = z_0 + \zeta, \end{cases} \tag{3.55}$$

with $G(t) \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$. This dynamics is known as an **observer** of the original system [Luenberger, 1971], also referred to as "sequential estimator" [Särkkä, 2013] in a stochastic framework. This observer has the objective of representing the original trajectory ($z(t)$ without error in the initial condition) while removing the bias from the added error ζ by using the misfit between measurements y and itself. In the observer dynamics, we note the presence of the feedback operator and an error in the initial condition. For exemplification, the original and observer trajectories are observed through a linear observation operator $C \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$. The disturbed measurements are defined as

$$y(t) = Cz(t) + \chi(t).$$

We introduce a dynamics representing the residual system $\tilde{z}(t) = z(t) - \widehat{z}(t)$ that will satisfy the dynamics

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \tilde{z}(t) = (A - G(t)C)\tilde{z}(t) + G(t)\chi(t), \qquad \forall t \in [0,T], \\[2ex] \tilde{z}(0) = \zeta. \end{cases}$$

By choosing G in a way that the residual dynamics are dissipative, we ensure that the error will be reduced through the observer dynamics. At the same time, attention must be given to the noise gain in $G(t)\chi(t)$.

3.3.2 Time-continuous Kalman Filter

Although the error will decrease when choosing a proper G, no relation between the optimal condition and the final state $\widehat{z}(T)$ is ensured in the context of least-squares minimization. We show in this section that we can attain optimality, for linear-quadratic optimization problems, by decomposing the two-end problem (3.50) in a Cauchy problem, *i.e.* a problem with exclusively initial or final conditions. This decomposition leads to the definition of the Kalman Filter [Simon, 2006].

> **Property 3.3.1.** *Introducing an observer $\widehat{z} \in C^1([0,T];\mathcal{Z})$, the following coupled dynamics satisfy the optimality condition $\widehat{z}(T) = \bar{z}(T)$ from (3.50), where the dynamics and observation operator are defined as in Section 3.3.1, i.e. linear dynamics and linear observation operator.*

Denoting by $P \in C^1([0;T];\mathcal{L}(\mathcal{Z}))$, a self-adjoint operator, the trajectories satisfies the dynamics

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}P(t) - P(t)A^* - AP(t) + P(t)C^*CP(t) = 0, & \forall t \in [0;T], \\[2ex] \dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{z}(t) = A\widehat{z}(t) + b(t) + G(t)\Big(y(t) - C\widehat{z}(t)\Big), & \forall t \in [0;T], \\[2ex] \widehat{z}(0) = z_0, \quad P(0) = R^{-1}, \quad G(t) = P(t)C^*. \end{cases} \tag{3.56}$$

The $P(t)$ dynamics will be referred to as the Riccati dynamics. The existence and unicity of the solution in P is ensured by the Cauchy-Lipschitz theorem [Demailly, 2006].

Proof. Defining

$$\eta(t) = \overline{z}(t) - \widehat{z}(t) - P(t)\overline{p}(t), \quad \forall t \in [0;T], \tag{3.57}$$

where $\overline{p}(t)$ is the adjoint dynamics from the two-end problem (3.50). We show that $\eta(t)$ is necessarily null for all time $t \in [0;T]$, yielding the desired result. Therefore, we develop a dynamics on η by computing the derivative it *w.r.t.* time,

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta(t) = \frac{\mathrm{d}}{\mathrm{d}t}\overline{z} - \frac{\mathrm{d}}{\mathrm{d}t}\widehat{z} - \frac{\mathrm{d}}{\mathrm{d}t}(P(t)\overline{p}).$$

Further developing the $\eta(t)$ dynamics by using the dynamics of $\overline{z}$, $\widehat{z}$, $P(t)$ and $\overline{p}$, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\eta(t) &= \big(A\overline{z} + b(t)\big) - \big(A\widehat{z} + b(t) + P(t)C^*CP(t)\big) \\ &\quad - \Big(\big(P(t)A^* + AP(t) - P(t)C^*CP(t)\big)\overline{p}(t) + P(t)\big(A\overline{p}(t) + C^*(y(t) - C\overline{z})\big)\Big), \\ &= A\big(\overline{z} - \widehat{z} - P(t)\overline{p}(t)\big) + P(t)C^*C\big(\widehat{z} - \overline{z} + P(t)\overline{p}(t)\big) \\ &= A\eta - P(t)C^*C\eta. \end{aligned}$$

As $\eta(0) = 0$ without a source term in its dynamics, it is necessarily null for $t \in [0,T]$, by unicity. Then, as

$$\eta(T) = 0 = \overline{z}(T) - \widehat{z}(T) - P(T)\overline{p}(T)$$

and $\overline{p}(T) = 0$, we have that

$$\overline{z}(T) = \widehat{z}(T).$$

□

We've shown that by choosing $G(t) = P(t)C^*$ in (3.55), we satisfy the optimality condition at the end of the observer trajectory in the linear-quadratic case. An important aspect of this result is that only the state at time T satisfies optimality, hence its utility to estimate initial state conditions and time-independent parameters. Although an error in the model A can be included in the filtering method, in this section we consider that such an error is not present. The Riccati dynamics followed by the operator P – called covariance operator – propagates the covariance of $\widehat{z}$, being useful for sensitivity and uncertainty quantification analysis. As it is an operator $\mathcal{L}(\mathcal{Z})$, in practice it can be computationally expensive to store and propagate such an operator as we will see in the time-discrete Kalman Filter.

3.3.2.1 Extended Kalman Filter

In the previous subsection, we presented the Kalman Filter for the linear-quadratic case. The Extended Kalman Filter (EKF) [Simon, 2006] proposes an extension of the Kalman Filter to nonlinear dynamics such as (3.4), although no optimality condition can be ensured. The idea is to use the tangent of the nonlinear operator evaluated at the current state, namely $\mathrm{D}A(\widehat{z}(t)) \in \mathcal{L}(\mathcal{Z})$ and $\mathrm{D}C(\widehat{z}(t)) \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$. So, from (3.56), we have

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}P(t) - P(t)(\mathrm{D}A)^* - \mathrm{D}AP(t) + P(t)\mathrm{D}C^*\mathrm{D}CP(t) = 0, & \forall t \in [0;T], \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{z}(t) = \mathrm{D}A\widehat{z}(t) + b(t) + G(t)\Big(y(t) - C\big(\widehat{z}(t)\big)\Big), & \forall t \in [0;T], \\ \widehat{z}(0) = z_0, \quad P(0) = R, \quad G(t) = P(t)\mathrm{D}C^*, \end{cases} \tag{3.58}$$

with omitted dependencies on $\mathrm{D}A(\widehat{z}(t))$ and $\mathrm{D}C(\widehat{z}(t))$.

3.3.2.2 Reduced-Order Kalman Filter

In the cases where we have only partial uncertainty on the initial state, the reduction of the unknown space leads to the Reduced-Order version of the Kalman filter (ROKF). In this section, we suppose the uncertainty to be only present partially in the initial state of the dynamics, as in Problem B.2. Linear dynamics will be considered. We recall here the augmented state,

$$^{\sharp}z(t) = \begin{pmatrix} z(t) \\ \theta(t) \end{pmatrix} \in {}^{\sharp}\mathcal{Z}, \quad z(t) \in \mathcal{Z}, \quad \theta(t) \in \Theta, \quad \forall t \in [0;T].$$

The dynamics for the linear-quadratic case (3.26) will be considered with a linear observation operator,

$$^{\sharp}C\big({}^{\sharp}z(t)\big) = {}^{\sharp}C\,{}^{\sharp}z(t) = \begin{pmatrix} C & 0 \end{pmatrix} {}^{\sharp}z(t) = Cz(t).$$

As done for the Kalman Filter, we start from a two-end problem (3.56) and use the fact that we have parameter-only uncertainty to obtain a reduced-order version of the Kalman Filter.

Property 3.3.2. *Assuming that the Riccati equation in (3.56) is well-posed, in particular, that there exists a unique solution, and we have the state defined as above with uncertainty present only in the parameter θ, the optimal Reduced-Order Kalman filter is given by*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{z}(t) = A\widehat{z}(t) + B(t)\widehat{\theta}(t) + b(t) + L_z(t)\dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\theta}(t), & \forall t \in [0;T], \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\theta}(t) = U^{-1}(t)L_z(t)^*C^*\Big(y(t) - C\widehat{z}(t)\Big), & \forall t \in [0;T], \\ \dfrac{\mathrm{d}}{\mathrm{d}t}L_z(t) = AL_z(t) + B(t), & \forall t \in [0;T] \\ \dfrac{\mathrm{d}}{\mathrm{d}t}U(t) = L_z(t)^*C^*CL_z(t), & \forall t \in [0;T] \\ \widehat{z}(0) = z_0, \quad \widehat{\theta}(0) = \theta_0, \quad L_z(0) = 0, \quad U(0) = R_\Theta. \end{cases} \tag{3.59}$$

Proof. When augmenting the Riccati dynamics in (3.56) we have

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{\sharp}P(t) - {}^{\sharp}P(t)A^{L*} - A^{L}\,{}^{\sharp}P(t) + {}^{\sharp}P(t)\,{}^{\sharp}C^{*}\,{}^{\sharp}C\,{}^{\sharp}P(t) = 0, & \forall t \in [0;T], \\ {}^{\sharp}P(0) = {}^{\sharp}R^{-1}, \end{cases}$$

where ${}^{\sharp}R$ is decomposed as

$${}^{\sharp}R = {}^{\sharp}L_0 U_0^{-1}\,{}^{\sharp}L_0^{*} \in \mathcal{L}({}^{\sharp}\mathcal{Z}), \qquad \text{with} \quad {}^{\sharp}L_0 = \begin{pmatrix} 0 \\ I \end{pmatrix} \in \mathcal{L}(\Theta, {}^{\sharp}\mathcal{Z}),$$

with $U_0 = R_\Theta$, the operator defined in (3.19). Now, one would like to decompose ${}^{\sharp}P$ in a reduced form, it only needs to propagate information that relates the parameter space and the observed state space. Denoting by $\{U(t)\}_{t\in[0;T]} \subset \mathcal{L}(\Theta)$ and $\{{}^{\sharp}L(t)\}_{t\in[0;T]} \subset \mathcal{L}(\Theta, {}^{\sharp}\mathcal{Z})$ the trajectories satisfying the following dynamics

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} U(t) = {}^{\sharp}L(t)^{*}\,{}^{\sharp}C^{*}\,{}^{\sharp}C\,{}^{\sharp}L(t), \\ U(0) = U_0, \end{cases} \qquad \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\,{}^{\sharp}L(t) = A^{L}(t)\,{}^{\sharp}L(t), \\ L(0) = L_0, \end{cases}$$

then we have that $\{Q(t)\}_{t\in[0;T]} \subset \mathcal{L}({}^{\sharp}\mathcal{Z})$ defined by

$$Q(t) = {}^{\sharp}L(t)U(t)^{-1}\,{}^{\sharp}L(t)^{*}, \quad \forall t \in [0;T], \tag{3.60}$$

satisfies the Riccati equation (3.56) as shown in the following. Differentiating (3.60) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = \left(\frac{\mathrm{d}}{\mathrm{d}t}\,{}^{\sharp}L\right) U^{-1}\,{}^{\sharp}L^{*} + {}^{\sharp}LU^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\,{}^{\sharp}L\right)^{*} + {}^{\sharp}L\left(\frac{\mathrm{d}}{\mathrm{d}t}U^{-1}\right){}^{\sharp}L^{*},$$

since

$$\frac{\mathrm{d}}{\mathrm{d}t}U^{-1} = -U^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}U\right)U^{-1} = -U^{-1}\,{}^{\sharp}L^{*}\,{}^{\sharp}C^{*}\,{}^{\sharp}C\,{}^{\sharp}LU^{-1},$$

with time-dependencies in $U(t)$ and ${}^{\sharp}L(t)$ omitted. With that we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = \left(A^{L}\,{}^{\sharp}L\right)U^{-1}\,{}^{\sharp}L^{*} + {}^{\sharp}LU^{-1}\left(A^{L}\,{}^{\sharp}L\right)^{*} - {}^{\sharp}L\left(U^{-1}\,{}^{\sharp}L^{*}\,{}^{\sharp}C^{*}\,{}^{\sharp}C\,{}^{\sharp}LU^{-1}\right){}^{\sharp}L^{*},$$

that, rearranging, gives

$$\frac{\mathrm{d}}{\mathrm{d}t}Q = A^{L}\left({}^{\sharp}LU^{-1}\,{}^{\sharp}L^{*}\right) + \left({}^{\sharp}LU^{-1}\,{}^{\sharp}L^{*}\right)A^{L*} - {}^{\sharp}LU^{-1}\,{}^{\sharp}L^{*}\,{}^{\sharp}C^{*}\,{}^{\sharp}C\left({}^{\sharp}LU^{-1}\,{}^{\sharp}L^{*}\right),$$

showing that the definition (3.60) satisfies the Riccati equation in (3.56). Now, we decompose the dynamics of ${}^{\sharp}L$ introducing the block

$${}^{\sharp}L(t) = \begin{pmatrix} L_z(t) \\ L_\theta(t) \end{pmatrix}, \quad \text{with} \quad L_z(t) \in \mathcal{L}(\Theta, \mathcal{Z}), \quad L_\theta(t) \in \mathcal{L}(\Theta), \quad \forall t \in [0,T].$$

Using the definition of A^{L}, from (3.26), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}L_\theta(t) = 0, \quad \forall t \in [0,T],$$

then $L_\theta(t) = L_\theta(0) = I$, so the ${}^{\sharp}L$ dynamics is reduced to the parameter-state sensitivity

$$\frac{\mathrm{d}}{\mathrm{d}t}L_z(t) = AL_z(t) + B(t), \forall t \in [0;T].$$

Taking the dynamics for $\widehat{z}$ in (3.56) for its augmented version, considering the state-only observation operator and inserting Q as the Riccati solution, we have the state observer

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{z}(t) = A\widehat{z}(t) + b(t) + {}^{\sharp}L(t)U(t)^{-1}\,{}^{\sharp}L(t)^{*}C^{*}\Big(y(t) - C\widehat{z}(t)\Big), \quad \forall t \in [0;T],$$

leading to the desired observer dynamics and to the parameter observer dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\theta}(t) = U^{-1}(t)L_z(t)^{*}C^{*}\Big(y(t) - C\widehat{z}(t)\Big), \quad \forall t \in [0;T].$$

Noting that the time-derivative of the parameter observer appears in the state observer, we arrive at the desired formulation. □

Solving the dynamics of Property 3.3.1 for the Problem B.2, one can show that the optimal solution $\overline{\zeta_\theta} \in \Theta$ of the minimization problem (3.27) is obtained at the end of the estimation procedure, namely

$$\overline{\zeta_\theta} = \widehat{\theta}(T) - \theta_0.$$

As mentioned, propagating the operator P when estimating using Property 3.3.1 may be expensive in terms of memory and computation. Using the Reduced-Order Kalman Filtering, instead of propagating $P \in \mathcal{L}(\mathcal{Z})$, one needs to propagate the sensitivity $L_z(t) \in \mathcal{L}(\Theta, \mathcal{Z})$ which size is expected to be smaller.

3.3.3 Time-discrete Kalman Filter

We present here the discrete version of the Kalman Filter [Moireau, 2022] by recalling the discrete dynamics (3.10)

$$z^{n+1} = \Phi_{n+1|n}(z^n) + b^n, \quad \forall n \in [\![0; N-1]\!],$$

where z^0 is given, and the observations obtained by a potentially nonlinear observation operator

$$y^n = C(z^n) + \chi^n, \quad \forall n \in [\![0; N]\!],$$

with disturbance χ^n. In analogy with the time-continuous case, the discrete observer dynamics is

$$\widehat{z}^{n+1} = \Phi_{n+1|n}(\widehat{z}^n) + b^n + G^n\big(y^n - C(\widehat{z}^n)\big), \quad \forall n \in [\![0; N-1]\!],$$

where $\widehat{z}^0$ is given.

3.3.3.1 Optimal derivation for linear-quadratic systems

To show the optimality properties of the discrete Kalman Filter we present the observer for an inverse problem in the linear form

$$\widehat{z}^{n+1} = \Phi_{n+1|n}\widehat{z}^n + b^n + G^n(y^n - C\widehat{z}^n), \quad \forall n \in [\![0; N-1]\!],$$

where $\widehat{z}^0$ is given. The approach to be presented is to define G^n in a way that $\widehat{z}^N = \overline{z}^N$, where $\overline{z}$ is the optimal state trajectory in (3.51) for the linear case. This demonstration differs from the one done for the time-continuous Kalman Filter, where we deducted the optimal condition from the two-end problem. We state our objective as: given the observations $\{y^n\}_0^N$, we want to minimize

$$\mathcal{J}_N(z_\zeta^0) = \left\{ r(z_\zeta^0) + \frac{\Delta t}{2}\sum_{n=0}^{N} \|y^n - Cz_\zeta^n\|_{\mathcal{Y}}^2 \right\}$$

where $\{z^n_\zeta\}_0^N$ is represented by a dynamics with uncertain initial conditions

$$\begin{cases} z^{n+1}_\zeta = \Phi_{n+1|n} z^n_\zeta + b^n, \quad \forall n \in [\![0; N-1]\!], \\ \\ z^0_\zeta = z_0 + \zeta. \end{cases}$$

One can note that minimizing this problem *w.r.t.* z^0_ζ is equivalent to minimizing *w.r.t.* ζ. We write these dynamics in a non-recursive form

$$\begin{aligned} z^{n+1}_\zeta &= \Phi_{n+1|n} \cdots \Phi_{1|0} z^0_\zeta + b^n \\ &\quad + \Phi_{n+1|n} b^{n-1} + \Phi_{n+1|n}\Phi_{n|n-1} b^{n-2} + \cdots + \Phi_{n+1|n} \cdots \Phi_{2|1} b^0, \end{aligned}$$

leading to

$$z^n_\zeta = \Phi_{n|0} z^0_\zeta + \sum_{k=0}^{n-1} \Phi_{n|k+1} b^k, \tag{3.61}$$

where $\Phi_{n|0}$ takes the initial state z^0_ζ directly to the step n. Having a closed form for z^n_ζ, we can write an observation operator that, applied to the initial state, outputs observations up to n as

$$\widetilde{C}_{|n} = \begin{pmatrix} C\Phi_{0|0} \\ \vdots \\ C\Phi_{n|0} \end{pmatrix}.$$

Additionally, introducing a modified observations vector

$$\widetilde{y}_{|n} = \begin{pmatrix} \widetilde{y}^0 \\ \vdots \\ \widetilde{y}^n \end{pmatrix} = \begin{pmatrix} y^0 \\ \vdots \\ y^n - C\sum_{k=0}^{n-1} \Phi_{n|k+1} b^k \end{pmatrix} \quad \text{and} \quad \|\widetilde{y}_{|n}\|^2_{\mathcal{Y}_{|n}} = \Delta t \sum_{k=0}^{n} (\widetilde{y}^k, \widetilde{y}^k)_{\mathcal{Y}},$$

the optimization can be rewritten as

$$\mathcal{J}_N(z^0_\zeta) = \frac{1}{2}\left\{ (R z^0_\zeta, z^0_\zeta)_{\mathcal{Z}} + \left\| \widetilde{y}_{|N} - \widetilde{C}_{|N} z^0_\zeta \right\|^2_{\mathcal{Y}_{|N}} \right\}.$$

The subscript "$|n$" means "knowing n samples". The minimization of this functional is interpreted as finding the initial condition $z^0_\zeta = z_0 + \zeta$ while including a regularization criteria represented by R. We recognize this functional in analogy with the unconstrained case (3.2) which minimum has the closed form

$$\begin{aligned} \overline{z^0_\zeta}_{|n} &= \left(R + \Delta t \widetilde{C}^*_{|n} \widetilde{C}_{|n}\right)^{-1} \left(R z_0 + \Delta t \widetilde{C}^*_{|n} \widetilde{y}_{|n}\right) \\ &= \left(R + \Delta t \sum_{k=0}^{n} (C\Phi_{k|0})^* (C\Phi_{k|0})\right)^{-1} \left(R z_0 + \Delta t \sum_{k=0}^{n} (C\Phi_{k|0})^* \widetilde{y}^n \right), \end{aligned} \tag{3.62}$$

so $\overline{z^0_\zeta}_{|n}$ means the optimal initial state obtained through minimization when n samples of observed data is available. Now, we define

$$\left| \begin{aligned} (P_{|n})^{-1} &= R + \Delta t \sum_{k=0}^{n} (C\Phi_{k|0})^* (C\Phi_{k|0}), \quad \forall n \in 0, \ldots, N, \\ \\ &= (P_{|n-1})^{-1} + \Delta t (C\Phi_{n|0})^* C\Phi_{n|0}, \quad \text{with} \quad (P_{|0})^{-1} = R. \end{aligned} \right. \tag{3.63}$$

Using this definition in (3.62), the optimal initial state satisfy the recursivity

$$\overline{z^0_\zeta}_{|n} = P_{|n} \left(R z_0 + \Delta t \sum_{k=0}^{n} \Phi^*_{n|0} C^* \widetilde{y}^k \right).$$

Then, the optimal initial state $\overline{z^0_\zeta}_{|n}$ can be written in a recursive form as

$$\begin{aligned}\overline{z^0_\zeta}_{|n} &= P_{|n}\left((P_{|n-1})^{-1}\overline{z^0_\zeta}_{n-1} + \Delta t\Phi^*_{n|0}C^*\widetilde{y}^n\right)\\ &= P_{|n}(P_{|n-1})^{-1}\overline{z^0_\zeta}_{n-1} + \Delta t P_{|n}\Phi^*_{n|0}C^*\widetilde{y}^n.\end{aligned} \tag{3.64}$$

Using the inversion lemma, Property A.2.5, we develop the inverse $(P_{|n})^{-1}$

$$P_{|n} = P_{|n-1} - \Delta t P_{|n-1}\Phi^*_{n|0}C^*(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)^{-1}C\Phi_{n|0}P_{|n-1}.$$

Replacing $P_{|n}$ in (3.64) we have

$$\overline{z^0_\zeta}_{|n} = \overline{z^0_\zeta}_{n-1} - G^n_0 C\Phi_{n|0}\overline{z^0_\zeta}_{n-1} + \Delta t P_{|n}\Phi^*_{n|0}C^*\widetilde{y}^n,$$

with

$$\begin{aligned}G^n_0 &= \Delta t P_{|n-1}\Phi^*_{n|0}C^*(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)^{-1}\\ &= \Delta t P_{|n}(P_{|n})^{-1}P_{|n-1}\Phi^*_{n|0}C^*(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)^{-1}\\ &= \Delta t P_{|n}((P_{|n-1})^{-1} + \Delta t\Phi^*_{n|0}C^*C\Phi_{n|0})P_{|n-1}\Phi^*_{n|0}C^*(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)^{-1}\\ &= \Delta t P_{|n}\Phi^*_{n|0}C^*(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)(I + \Delta t C\Phi_{n|0}P_{|n-1}\Phi^*_{n|0}C^*)^{-1}\\ &= \Delta t P_{|n}\Phi^*_{n|0}C^*.\end{aligned}$$

G^n_0 is a “correction” operator for the estimated initial condition by using n samples of the observed data. Using the above, the recursive optimal trajectory satisfies

$$\overline{z^0_\zeta}_{|n} = \overline{z^0_\zeta}_{|n-1} + G^n_0(\widetilde{y}^n - C\Phi_{n|0}\overline{z^0_\zeta}_{|n-1}). \tag{3.65}$$

The filter will follow from this expression where we now aim at writing it in a recursive form without the need for $\Phi_{n|0}$. Let $\widehat{z}^n_+$ be the “corrected” state, the evolution of the dynamics up to the n-th step from the optimal initial state $\overline{z^0_\zeta}_{|n}$. We start by propagating the optimal initial state in the non-recursive form as in (3.61)

$$\widehat{z}^n_+ = \Phi_{n|0}\overline{z^0_\zeta}_{|n} + \sum_{k=0}^{n-1}\Phi_{n|k+1}b^k. \tag{3.66}$$

Using this formula, we have for step $n-1$ that

$$\begin{aligned}\overline{z^0_\zeta}_{|n-1} &= \Phi^{-1}_{n-1|0}\widehat{z}^{n-1}_+ - \Phi^{-1}_{n-1|0}\sum_{k=0}^{n-2}\Phi_{n-1|k+1}b^k\\ &= \Phi^{-1}_{n-1|0}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi^{-1}_{k+1|0}b^k.\end{aligned}$$

Replacing (3.65) in (3.66), using the $\overline{z^0_\zeta}_{|n-1}$ above and the definition of $\widetilde{y}^n$, we have

$$
\begin{aligned}
\widehat{z}^n_+ &= \Phi_{n|0}\Big(\overline{z^0_\zeta}_{|n-1} + G^n_0(\widetilde{y}^n - C\Phi_{n|0}\overline{z^0_\zeta}_{|n-1})\Big) + \sum_{k=0}^{n-1}\Phi_{n|k+1}b^k \\
&= \Phi_{n|0}\Big(\Phi^{-1}_{n-1|0}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi^{-1}_{k+1|0}b^k + G^n_0\big(\widetilde{y}^n - C\Phi_{n|0}(\Phi^{-1}_{n-1|0}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi^{-1}_{k+1|0}b^k)\big)\Big) + \sum_{k=0}^{n-1}\Phi_{n|k+1}b^k \\
&= \Big(\Phi_{n|n-1}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi_{n|k+1}b^k + \Phi_{n|0}G^n_0\big(\widetilde{y}^n - C(\Phi_{n|n-1}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi_{n|k+1}b^k)\big)\Big) + \sum_{k=0}^{n-1}\Phi_{n|k+1}b^k \\
&= \Phi_{n|n-1}\widehat{z}^{n-1}_+ + b^{n-1} + \Phi_{n|0}G^n_0\Big(\widetilde{y}^n - C(\Phi_{n|n-1}\widehat{z}^{n-1}_+ - \sum_{k=0}^{n-2}\Phi_{n|k+1}b^k)\Big) \\
&= \Phi_{n|n-1}\widehat{z}^{n-1}_+ + b^{n-1} + \Phi_{n|0}G^n_0\Big(y^n - C\sum_{k=0}^{n-1}\Phi_{n|k+1}b^k - C\Phi_{n|n-1}\widehat{z}^{n-1}_+ + C\sum_{k=0}^{n-2}\Phi_{n|k+1}b^k\Big) \\
&= \Phi_{n|n-1}\widehat{z}^{n-1}_+ + b^{n-1} + \Phi_{n|0}G^n_0\Big(y^n - C(\Phi_{n|n-1}\widehat{z}^{n-1}_+ + b^{n-1})\Big).
\end{aligned}
$$

By defining a "predicted" state as

$$\widehat{z}^n_- = \Phi_{n|n-1}\widehat{z}^{n-1}_+ + b^{n-1},$$

$\widehat{z}^n_+$ can be written as

$$\widehat{z}^n_+ = \widehat{z}^n_- + \Phi_{n|0}G^n_0(y^n - C\widehat{z}^n_-).$$

By the nature of these operations, $\widehat{z}_-$ outcomes from a prediction step and $\widehat{z}_+$ from a correction step. Finally, we would like to have a recursive form of the operation $\Phi_{n|0}G^n_0$. To do so, we similarly consider the propagated covariance from its initial optimal condition, namely the "corrected" covariance

$$P^n_+ = \Phi_{n|0}P_{|n}\Phi^*_{n|0}.$$

We define the "predicted" covariance in a recursive form as

$$P^n_- = \Phi_{n|n-1}P^{n-1}_+\Phi^*_{n|n-1}. \tag{3.67}$$

Using (3.63), we define a "corrected" covariance as

$$
\begin{aligned}
(P^n_+)^{-1} &= {\Phi^*_{n|0}}^{-1}(P_{|n})^{-1}\Phi^{-1}_{n|0} \\
&= {\Phi^*_{n|0}}^{-1}\Big((P_{|n-1})^{-1} + \Delta t\Phi^*_{n|0}C^*C\Phi_{n|0}\Big)\Phi^{-1}_{n|0} \\
&= (P^n_-)^{-1} + \Delta t C^*C.
\end{aligned}
$$

Finally, using the inversion lemma,

$$P^n_+ = P^n_- - \Delta t P^n_- C^*\Big(I + \Delta t C P^n_- C^*\Big)^{-1} C P^n_-,$$

we define the discrete gain operator

$$G^n = \Phi_{n|0}G^n_0 = \Delta t\Phi_{n|0}P_{|n}\Phi^*_{n|0}C^* = \Delta t P^n_+ C^*.$$

With these developments, we started from the propagation of the optimal initial state (where the initial state is estimated considering n available measurements) and obtained a two-stage recursive algorithm (prediction and correction) that results in the same optimal observer. The algorithm then reads:

$$
\begin{aligned}
&(\text{Initialization}) \left\{ \begin{array}{ll}
P_-^0 & = R^{-1} \\
P_+^0 & = P_-^0 - \Delta t P_-^0 C^* \left(I + \Delta t C P_-^0 C^* \right)^{-1} C P_-^0, \\
G^0 & = \Delta t P_+^0 C^*, \\
\widehat{z}_+^0 & = z_0 + G^0 (\widetilde{y}^0 - C z_0)
\end{array} \right. \\
&(\text{Prediction}) \left\{ \begin{array}{ll}
\widehat{z}_-^{n+1} & = \Phi_{n+1|n} \widehat{z}_+^n + b^n \\
P_-^{n+1} & = \Phi_{n+1|n} P_+^n \Phi_{n+1|n}^*
\end{array} \right. \\
&(\text{Correction}) \left\{ \begin{array}{ll}
P_+^{n+1} & = P_-^{n+1} - \Delta t P_-^{n+1} C^* \left(I + \Delta t C P_-^{n+1} C^* \right)^{-1} C P_-^{n+1}, \\
G^{n+1} & = \Delta t P_+^{n+1} C^*, \\
\widehat{z}_+^{n+1} & = \widehat{z}_-^{n+1} + G^{n+1} (\widetilde{y}^{n+1} - C \widehat{z}_-^{n+1})
\end{array} \right.
\end{aligned}
\tag{3.68}
$$

3.3.3.2 Extended Kalman Filter

As done for the time-continuous case, we present here the time-discrete Extended Kalman Filter in which the nonlinear operators are linearized around the current state. This formulation is useful for the nonlinear case (3.10). The algorithm reads

$$
\begin{aligned}
&(\text{Initialization}) \left\{ \begin{array}{ll}
P_-^0 & = R^{-1} \\
P_+^0 & = P_-^0 - \Delta t P_-^0 (\mathrm{D}C^{(0)})^* \left(I + \Delta t \mathrm{D}C^{(0)} P_-^0 (\mathrm{D}C^{(0)})^* \right)^{-1} \mathrm{D}C^{(0)} P_-^0, \\
G^0 & = \Delta t P_+^0 (\mathrm{D}C^{(n)})^*, \\
\widehat{z}_+^0 & = z_0 + G^0 (\widetilde{y}^0 - C(z_0))
\end{array} \right. \\
&(\text{Prediction}) \left\{ \begin{array}{ll}
\widehat{z}_-^{n+1} & = \Phi_{n+1|n} (\widehat{z}_+^n) + b^n \\
P_-^{n+1} & = \mathrm{D}\Phi_{n+1|n} P_+^n (\mathrm{D}\Phi_{n+1|n})^*
\end{array} \right. \\
&(\text{Correction}) \left\{ \begin{array}{ll}
P_+^{n+1} & = P_-^{n+1} - \Delta t P_-^{n+1} (\mathrm{D}C^{(n+1)})^* \left(I + \Delta t \mathrm{D}C^{(n+1)} P_-^{n+1} (\mathrm{D}C^{(n+1)})^* \right)^{-1} \mathrm{D}C^{(n+1)} P_-^{n+1}, \\
G^{n+1} & = P_+^{n+1} (\mathrm{D}C^{(n+1)})^*, \\
\widehat{z}_+^{n+1} & = \widehat{z}_-^{n+1} + G^{n+1} \left(\widetilde{y}^{n+1} - C(\widehat{z}_-^{n+1}) \right)
\end{array} \right.
\end{aligned}
\tag{3.69}
$$

where the dependency in $\mathrm{D}\Phi_{n+1|n}(\widehat{z}_-^n)$ was omitted and $\mathrm{D}C^{(n)} = \mathrm{D}C(\widehat{z}_-^n)$.

3.3.3.3 Unscented Kalman Filter

In the context of estimation in a nonlinear framework, another alternative is the Unscented Kalman Filter (UKF) [Julier, Uhlmann, and Durrant-Whyte, 2000; Simon, 2006; Moireau and Chapelle, 2011]. As shown, the Extended KF alternative requires the computation of the tangent operators, namely $\mathrm{D}\Phi$ and $\mathrm{D}C$, for propagating the covariance which can be prohibitively complex. Instead of computing directly the tangents, the propagation of the covariance can be done empirically through a particular set of particles propagated through the nonlinear operator.

Definition 3.3.1. ***Empiric mean.*** *Given $\mathcal{V}$ a Hilbert Space. Let $\{\alpha_i\}_{i=1}^r \in \mathbb{R}_+^r$ such that $\sum_{i=1}^r \alpha_i =$ 1. We define the empirical mean operator $E_\alpha : \mathcal{V}^r \mapsto \mathcal{V}$ such that for any $\{v^{(i)}\}_{i=1}^r \in \mathcal{V}^r$,*

$$
E_\alpha(\{v^{(i)}\}_{i=1}^r) = \sum_{i=1}^r \alpha_i v^{(i)}.
$$

We may use the shortened notation $E_\alpha(v^{(*)}) := E_\alpha(\{v^{(i)}\}_{i=1}^r)$.

Definition 3.3.2. ***Empiric covariance.*** *Given $\mathcal{V}$ and $\mathcal{W}$ two Hilbert Spaces. Let $\{\alpha_i\}_{i=1}^r \in \mathbb{R}_+^r$ such that $\sum_{i=1}^r \alpha_i = 1$. We define the empirical covariance* $\mathrm{Cov}_\alpha : \mathcal{V}^r \times \mathcal{W}^r \mapsto \mathcal{L}(\mathcal{W}, \mathcal{V})$ *such that for two sets $\{v^{(i)}\}_{i=1}^r \in \mathcal{V}^r$ and $\{w^{(i)}\}_{i=1}^r \in \mathcal{W}^r$*

$$\mathrm{Cov}_\alpha(\{v^{(i)}\}_{i=1}^r, \{w^{(i)}\}_{i=1}^r) : \mathcal{W} \ni u \mapsto \sum_{i=1}^r \alpha_i \Big(w^{(i)} - E_\alpha(w^{(*)}), u\Big)_{\mathcal{W}} (v^{(i)} - E_\alpha(v^{(*)})) \in \mathcal{V}.$$

We may use the shortened notations $\mathrm{Cov}_\alpha(v^{(*)}, w^{(*)}) := \mathrm{Cov}_\alpha(\{v^{(i)}\}_{i=1}^r, \{w^{(i)}\}_{i=1}^r)$ *and* $\mathrm{Cov}_\alpha(v^{(*)}) := \mathrm{Cov}_\alpha(v^{(*)}, v^{(*)})$.

Definition 3.3.3. ***Sigma-points.*** *Given $\mathcal{V}$ a Hilbert Space. Let $\{v^{(i)}\}_{i=1}^r \in \mathcal{V}^r$ be a set of elements. Its mean and covariance as $E(v^{(*)})$ and* $\mathrm{Cov}(v^{(*)})$*, respectively. We defined another set $\{\widetilde{v}^{(i)}\}_{i=1}^r \in \mathcal{V}^r$ such that the elements $\{v^{(i)}\}_{i=1}^r$ are constructed as*

$$v^{(i)} = E(v^{(*)}) + \widetilde{v}^{(i)}, \quad 1 \le i \le r.$$

We struct the set $\{\widetilde{v}^{(i)}\}_{i=1}^r$ with the following conditions, associating each element to a coefficient in $\{\alpha_i\}_{i=1}^r \in \mathbb{R}_+^r$,

$$\begin{cases} \sum_{i=1}^r \alpha_i & = 1 \\ E_\alpha(v^{(*)}) & = E(v^{(*)}) \\ \mathrm{Cov}_\alpha(v^{(*)}) & = \mathrm{Cov}(v^{(*)}). \end{cases}$$

These conditions implies for the set $\widetilde{v}^{(*)}$

$$\begin{cases} E_\alpha(\widetilde{v}^{(*)}) & = 0 \\ \mathrm{Cov}_\alpha(\widetilde{v}^{(*)}) & = \mathrm{Cov}(v^{(*)}). \end{cases}$$

The set of points $\{v^{(i)}\}_{i=1}^r$ satisfying these conditions are called sigma-points.

The Unscented Kalman Filter algorithm is presented below. We introduce the zero-mean and unit covariance set of *sigma-points* $\{\widetilde{I}^{(i)}\}_{i=1}^r \in \mathcal{Z}^r$ using Definition 3.3.1, Definition 3.3.2 and Definition 3.3.3. Here, instead of propagating only one observer $\widehat{z}$, we have r observers $\{\widehat{z}^{(i)}\}_{i=1}^r$ propagated through the dynamics and exploring the state space, retrieving information about how the uncertainties propagate through the dynamics. This exploration is controlled by the estimated covariance and the *sigma-points*. Additionally, we define the square root of a self-adjoint operator P in the decomposition

$$P = \sqrt{P}\sqrt{P}.$$

$$
\begin{aligned}
&\text{(Initialization)} \begin{cases} P_+^0 &= R^{-1}, \\ \widehat{z}_+^0 &= z_0. \end{cases} \\
&\text{(Sampling)} \quad \widehat{z}_+^{n(i)} = \widehat{z}_+^n + \sqrt{P_+^n}\widetilde{I}^{(i)}. \\
&\text{(Prediction)} \begin{cases} \widehat{z}_-^{n+1(i)} &= \Phi_{n+1|n}(\widehat{z}_+^{n(i)}) + b^n, \\ \widehat{z}_-^{n+1} &= E_\alpha(\widehat{z}_-^{n+1(*)}), \\ P_-^{n+1} &= \mathrm{Cov}_\alpha(\widehat{z}_-^{n+1(*)}). \end{cases} \\
&\text{(Resampling)} \begin{cases} \widehat{z}_-^{n+1(i)} &= \widehat{z}_-^{n+1} + \sqrt{P_-^{n+1}}\widetilde{I}^{(i)}, \\ \widehat{y}^{n+1(i)} &= C(\widehat{z}_-^{n+1(i)}). \end{cases} \\
&\text{(Correction)} \begin{cases} P_{z\widehat{y}} &= \mathrm{Cov}_\alpha(\widehat{z}_-^{n+1(*)}, \widehat{y}^{n+1(*)}), \\ P_{\widehat{y}\widehat{y}} &= I + \Delta t \mathrm{Cov}_\alpha(\widehat{y}^{n+1(*)}), \\ G^{n+1} &= \Delta t P_{z\widehat{y}}(P_{\widehat{y}\widehat{y}})^{-1}, \\ \widehat{z}_+^{n+1} &= \widehat{z}_-^{n+1} + G^{n+1}\big(y^n - E_\alpha(\widehat{y}^{n+1(*)})\big), \\ P_+^{n+1} &= P_-^{n+1} - G^{n+1}P_{\widehat{y}\widehat{y}}(G^{n+1})^*. \end{cases}
\end{aligned}
\tag{3.70}
$$

To better understand the particularities of this filter, one can first focus on the steps of the algorithm that is analogous to the extended Kalman Filter (3.69). In both cases, the prediction step uses the model to forward the state and to obtain a covariance from the model. In the correction phase, the state is corrected by using the propagated covariance and the observation misfit. Differently from the Extended version where the covariance is propagated using the tangent (linear) dynamics, the Unscented version retrieves the covariance by exploring the state space through the propagation of particles $\{\widehat{z}^{(i)}\}_{i=1}^r$ and gathering the outcome. This means that a trade-off for not requiring the tangent dynamics is the parallel propagation of r particles. As we will see in Section 3.4 this trade-off does not add much computational cost as the propagations can be done in parallel and the total memory required is proportional to the size of the state space squared. For instance, using the Extended version for the wave propagation problem we need to propagate the sensitivities, further details can be found in Section 3.4.1. Moreover, the Unscented version enables the choice of different stencil schemes for the sigma points. One can increase the number and use different distribution of sigma points, what may increase the computational cost but also increase the stability and convergence. A few examples of potential stencil schemes are presented in Chapter 4.

3.3.3.4 Reduced-Order Kalman Filter

Here we present the time-discrete version of the Reduced-Order Kalman Filter, as its demonstration differs slightly from the time-continuous case. We take as reference the linear-quadratic Problem II.2 with the associated definitions. The augmented dynamics of the linear-quadratic problem is expanded as

$$
\begin{cases}
z_\theta^{n+1} = \Phi_{n+1|n} z_\theta^n + B^n \theta^n, & \forall n \in [\![0; N-1]\!], \\
\theta^{n+1} = \theta^n, & \forall n \in [\![0; N-1]\!], \\
z_\theta^0 = z_0, \\
\theta^0 = \theta_0 + \zeta_\theta.
\end{cases}
$$

An important aspect of the ROKF is that only part of the augmented state ${}^\sharp z$ bears uncertainty. That said, we would like to propagate only the covariance related to this part. By rewriting (3.68) with the augmented operators and variables from Problem II.2, we have

$$
\begin{aligned}
&\text{(Initialization)}\begin{cases}
{}^{\sharp}P^0_- &= {}^{\sharp}R^{-1}\\
{}^{\sharp}P^0_+ &= {}^{\sharp}P^0_- - \Delta t\,{}^{\sharp}P^0_-\,{}^{\sharp}C^*\left(I + \Delta t\,{}^{\sharp}C\,{}^{\sharp}P^0_-\,{}^{\sharp}C^*\right)^{-1}{}^{\sharp}C\,{}^{\sharp}P^0_-,\\
{}^{\sharp}G^0 &= \Delta t\,{}^{\sharp}P^0_+\,{}^{\sharp}C^*,\\
{}^{\sharp}\widehat{z}^0_+ &= z_0 + {}^{\sharp}G^0(\widetilde{y}^0 - {}^{\sharp}C z_0).
\end{cases}\\
&\text{(Prediction)}\begin{cases}
{}^{\sharp}\widehat{z}^{n+1}_- &= \Phi^L{}_{n+1|n}\,{}^{\sharp}\widehat{z}^n_+ + b^n,\\
{}^{\sharp}P^{n+1}_- &= \Phi^L{}_{n+1|n}\,{}^{\sharp}P^n_+\Phi^{L*}{}_{n+1|n}.
\end{cases}\\
&\text{(Correction)}\begin{cases}
{}^{\sharp}P^{n+1}_+ &= {}^{\sharp}P^{n+1}_- - \Delta t\,{}^{\sharp}P^{n+1}_-\,{}^{\sharp}C^*\left(I + \Delta t\,{}^{\sharp}C\,{}^{\sharp}P^{n+1}_-\,{}^{\sharp}C^*\right)^{-1}{}^{\sharp}C\,{}^{\sharp}P^{n+1}_-,\\
{}^{\sharp}G^{n+1} &= \Delta t\,{}^{\sharp}P^{n+1}_+\,{}^{\sharp}C^*,\\
{}^{\sharp}\widehat{z}^{n+1}_+ &= {}^{\sharp}\widehat{z}^{n+1}_- + {}^{\sharp}G^{n+1}(\widetilde{y}^{n+1} - {}^{\sharp}C\,{}^{\sharp}\widehat{z}^{n+1}_-).
\end{cases}
\end{aligned}
\tag{3.71}
$$

We start by showing a possible decomposition of the augmented ${}^{\sharp}P$ trajectory.

> **Property 3.3.3.** *By decomposing*
>
> $${}^{\sharp}P^n_- = {}^{\sharp}L^n(U^{n-1})^{-1}({}^{\sharp}L^n)^*$$
>
> *with*
>
> $$\begin{cases} {}^{\sharp}L^{n+1} = \Phi^L{}_{n+1|n}\,{}^{\sharp}L^n\\ U^{n+1} = U^n + \Delta t({}^{\sharp}L^{n+1})^*({}^{\sharp}C)^*\,{}^{\sharp}C\,{}^{\sharp}L^{n+1}\end{cases}$$
>
> *we have that*
>
> $${}^{\sharp}P^n_+ = {}^{\sharp}L^n(U^n)^{-1}({}^{\sharp}L^n)^*.$$

Proof. We start from the prediction step of ${}^{\sharp}P$:

$${}^{\sharp}P^{n+1}_- = \Phi^L{}_{n+1|n}\,{}^{\sharp}P^n_+\Phi^{L*}{}_{n+1|n}$$

and replace the step ${}^{\sharp}P^n_+$

$${}^{\sharp}P^n_+ = {}^{\sharp}P^n_- - \Delta t\,{}^{\sharp}P^n_-\,{}^{\sharp}C^*\left(I + \Delta t\,{}^{\sharp}C\,{}^{\sharp}P^n_-\,{}^{\sharp}C^*\right)^{-1}{}^{\sharp}C\,{}^{\sharp}P^n_-.$$

Therefore, using the decomposition, we have

$$
\begin{aligned}
{}^{\sharp}P^{n+1}_+ &= \Phi^L{}_{n+1|n}\left({}^{\sharp}P^n_- - \Delta t\,{}^{\sharp}P^n_-\,{}^{\sharp}C^*\left(I + \Delta t\,{}^{\sharp}C\,{}^{\sharp}P^n_-\,{}^{\sharp}C^*\right)^{-1}{}^{\sharp}C\,{}^{\sharp}P^n_-\right)\Phi^{L*}{}_{n+1|n}\\
&= \Phi^L{}_{n+1|n}\,{}^{\sharp}L^n\\
&\quad\left((U^{n-1})^{-1} - \Delta t(U^{n-1})^{-1}({}^{\sharp}L^n)^*\,{}^{\sharp}C^*\left(I + \Delta t\,{}^{\sharp}C\,{}^{\sharp}L^n(U^{n-1})^{-1}({}^{\sharp}L^n)^*\,{}^{\sharp}C^*\right)^{-1}{}^{\sharp}C\,{}^{\sharp}L^n(U^{n-1})^{-1}\right)\\
&\quad({}^{\sharp}L^n)^*\Phi^{L*}{}_{n+1|n}.
\end{aligned}
$$

Finally, using the inversion lemma (Property A.2.5), we recognize

$${}^{\sharp}P^{n+1}_+ = {}^{\sharp}L^{n+1}(U^n)^{-1}({}^{\sharp}L^{n+1})^*.$$

□

The initialization of P^0_- is decomposed as

$$P^0_- = {}^{\sharp}L^0(U^0)^{-1}({}^{\sharp}L^0)^*,\quad {}^{\sharp}L^0 = \begin{pmatrix} L^0_z\\ L^0_\theta\end{pmatrix} = \begin{pmatrix} 0\\ I\end{pmatrix},\quad U^0 = R_\Theta$$

such that

$$P^0_- = \begin{pmatrix} 0 & 0 \\ 0 & R_\Theta^{-1} \end{pmatrix}.$$

By introducing the augmented state and the decomposition of the propagated covariance (Property 3.3.3) in the augmented Kalman Filter (3.71), we have

$$\begin{aligned}
&\text{(Initialization)} \left\{ \begin{aligned} L_z^0 &= 0 \\ U^0 &= R_\Theta \\ \widehat{z}_+^0 &= z_0 + \Delta t (U^0)^{-1} (L_z^0)^* C^* (\widehat{y}^0 - C z_0) \\ \widehat{\theta}_+^0 &= \theta_0 \end{aligned} \right. \\
&\text{(Prediction)} \left\{ \begin{aligned} \widehat{z}_-^{n+1} &= \Phi_{n+1|n} \widehat{z}_+^n + B^n \widehat{\theta}_+^n \\ \widehat{\theta}_-^{n+1} &= \widehat{\theta}_+^n, \\ L_z^{n+1} &= \Phi_{n+1|n} L_z^n + B^n \end{aligned} \right. \\
&\text{(Correction)} \left\{ \begin{aligned} U^{n+1} &= U^n + \Delta t (L_z^{n+1})^* C^* C L_z^{n+1} \\ G_\theta &= \Delta t (U^{n+1})^{-1} (L_z^{n+1})^* C^*, \\ \widehat{\theta}_+^{n+1} &= \widehat{\theta}_-^{n+1} + G_\theta (y^{n+1} - C \widehat{z}_-^{n+1}) \\ \widehat{z}_+^{n+1} &= \widehat{z}_-^{n+1} + \Delta t L_z^{n+1} (\widehat{\theta}_+^{n+1} - \widehat{\theta}_-^{n+1}) \end{aligned} \right.
\end{aligned} \tag{3.72}$$

As the uncertainty is present only in θ, we propagate only the covariance related to the parameter by propagating the sensitivity of the state *w.r.t.* the parameter L_z and the inverse of the parameter covariance U. This reduces the cost in memory and computation when compared with the nonreduced version.

3.3.3.5 Reduced-Order Extended Kalman Filter

In the case, we have nonlinear operators $\Phi_{n+1|n}$ and C *w.r.t.* θ as in Problem II.1, the extended form of the ROKF can be written by using their tangent as below. The Reduced-Order Extended Kalman Filter (ROEKF) reads

$$\begin{aligned}
&\text{(Initialization)} \left\{ \begin{aligned} L_z^0 &= 0, \\ U^0 &= R_\Theta, \\ \widehat{z}_+^0 &= z_0 + \Delta t (U^0)^{-1} (L_z^0)^* \big(\mathrm{D}C(z_0)\big)^* \big(\widehat{y}^0 - C(z_0)\big), \\ \widehat{\theta}_+^0 &= \theta_0. \end{aligned} \right. \\
&\text{(Prediction)} \left\{ \begin{aligned} \widehat{z}_-^{n+1} &= \Phi_{n+1|n}(\widehat{\theta}_+^n) \widehat{z}_+^n + b^n \\ L_z^{n+1} &= \big(\mathrm{D}\Phi_{n+1|n}(\widehat{\theta}_+^n) \circ \big) L_z^n, \\ \widehat{\theta}_-^{n+1} &= \widehat{\theta}_+^n. \end{aligned} \right. \\
&\text{(Correction)} \left\{ \begin{aligned} U^{n+1} &= U^n + \Delta t (L_z^{n+1})^* \big(\mathrm{D}C(\widehat{z}_-^{n+1})\big)^* \mathrm{D}C(\widehat{z}_-^{n+1}) L_z^{n+1} \\ G_\theta &= \Delta t (U^{n+1})^{-1} (L_z^{n+1})^* \big(\mathrm{D}C(\widehat{z}_-^{n+1})\big)^*, \\ \widehat{\theta}_+^{n+1} &= \widehat{\theta}_-^{n+1} + G_\theta \big(y^{n+1} - C(\widehat{z}_-^{n+1})\big), \\ \widehat{z}_+^{n+1} &= \widehat{z}_-^{n+1} + \Delta t L_z^{n+1} (\widehat{\theta}_+^{n+1} - \widehat{\theta}_-^{n+1}), \end{aligned} \right.
\end{aligned} \tag{3.73}$$

where $\big(\mathrm{D}\Phi_{n+1|n}(\widehat{\theta}_+^n) \circ \big) \in \mathcal{L}(\Theta, \mathcal{L}(\mathcal{Z}, \mathcal{Z}))$.

3.3.3.6 Reduced-Order Unscented Kalman Filter

When the uncertainty is reduced to only part of the state, the Unscented Kalman Filter can also be written in a decomposed form. In this case, the propagated covariance will comprise only the part related to the uncertainty, reducing the computational cost. Here, we aim to propose a sequential method for the nonlinear Problem II.1 without the need for tangent dynamics. For that, we do the same decomposition of the covariance as in Section 3.3.3.4 for the ROKF. Let $\{\widetilde{I}^{(i)}\}_{i=1}^r \in \Theta^r$ be a set of sigma-points satisfying the conditions presented in Definition 3.3.3, the Reduced-Order Unscented Kalman Filter (ROUKF) is presented below with the parameter-state sensitivity and parameter sensitivity operators; and inverse parameter covariance operator

$$L_{z\theta} \in \mathcal{L}(\Theta, \mathcal{Z}), \quad L_\theta^n \in \mathcal{L}(\Theta) \quad \text{and} \quad U^n \in \mathcal{L}(\Theta),$$

respectively. To obtain the ROUKF algorithm, we departed from (3.70) and introduced the augmented state variables and operators from Problem B.1 and the definition of the augmented covariance ${}^\sharp P$ presented in Section 3.3.3.4.

$$\begin{aligned}
&\text{(Initialization)} \left\{ \begin{array}{ll} U^0 & = R_\Theta, \\ \widehat{z}_+^0 & = z_0. \end{array} \right. \\
&\text{(Sampling)} \left\{ \begin{array}{ll} \widehat{z}_+^{n(i)} & = \widehat{z}_+^n, \\ \widehat{\theta}_+^{n(i)} & = \widehat{\theta}_+^n + L_\theta^n \sqrt{(U^n)^{-1}} \widetilde{I}^{(i)}. \end{array} \right. \\
&\text{(Prediction)} \left\{ \begin{array}{ll} \widehat{z}_-^{n+1(i)} & = \Phi_{n+1|n}(\widehat{\theta}_+^n) \widehat{z}_+^{n(i)} + b^n, \\ \widehat{\theta}_-^{n+1(i)} & = \widehat{\theta}_+^{n(i)}, \\ \widehat{y}^{n+1(i)} & = C(\widehat{z}_-^{n+1(i)}), \\ \widehat{z}_-^{n+1} & = E_\alpha(\widehat{z}_-^{n+1(*)}). \end{array} \right. \\
&\text{(Correction)} \left\{ \begin{array}{ll} L_{z\theta} & = \mathrm{Cov}_\alpha(\widehat{z}_-^{n+1(*)}, \widetilde{I}^{(*)}), \\ L_\theta & = \mathrm{Cov}_\alpha(\widehat{\theta}_-^{n+1(*)}, \widetilde{I}^{(*)}), \\ L_{y\theta} & = \mathrm{Cov}_\alpha(\widehat{y}^{n+1(*)}, \widetilde{I}^{(*)}), \\ U^{n+1} & = \mathrm{Cov}_\alpha(\widetilde{I}^{(*)}) + \Delta t (L_{y\theta})^* L_{y\theta}, \\ \widehat{z}_+^{n+1} & = \widehat{z}_-^{n+1} + \Delta t L_{z\theta} (U^{n+1})^{-1} (L_{y\theta})^* \big(y^n - E_\alpha(\widehat{y}^{n+1(*)})\big), \\ \widehat{\theta}_+^{n+1} & = \widehat{\theta}_-^{n+1} + \Delta t L_\theta (U^{n+1})^{-1} (L_{y\theta})^* \big(y^n - E_\alpha(\widehat{y}^{n+1(*)})\big). \end{array} \right.
\end{aligned} \tag{3.74}$$

As in the nonreduced form, the Unscented version propagates the covariance using particles distributed using *sigma-points*. In this case, as the uncertainty is present only in the parameter, the *sigma-points* are related to the parametric space Θ as well as the covariance that is propagated through the parameter-state sensitivity $L_{z\theta}$ and the parameter-parameter sensitivity L_θ^{n+1}. The parameter covariance at a given time step can be retrieved as

$$P_{\theta\theta}^n = L_\theta^n (U^n)^{-1} (L_\theta^n)^*.$$

Remarks on implementation. When using this algorithm with discretized spaces in $\mathbb{R}$, the *sigma-points* as in Definition 3.3.3 can be precisely chosen. One possible set of points is the simplex *sigma-points*, which has the smallest number of necessary points satisfying the conditions. This set $\{\widetilde{I}^{(i)}\}_{i=1}^p$ is generated as the columns of the matrix $[\widetilde{I}_p^{(*)}]$ that satisfies the following recursion

$$\left\{ \begin{array}{ll} [\widetilde{I}_1^{(*)}] & = \sqrt{p} \left(-\frac{1}{\sqrt{2\alpha}}, \frac{1}{\sqrt{2\alpha}} \right), \alpha = \frac{p}{p+1} \\ & \\ [\widetilde{I}_d^{(*)}] & = \sqrt{p} \left(\begin{array}{cccc} & & & 0 \\ & [\widetilde{I}_{d-1}^{(*)}] & & \vdots \\ & & & 0 \\ \frac{1}{\sqrt{\alpha d(d+1)}} & \vdots & \frac{1}{\sqrt{\alpha d(d+1)}} & \frac{-d}{\sqrt{\alpha d(d+1)}} \end{array} \right), \quad 2 \leq d \leq p, \end{array} \right. \tag{3.75}$$

where, for the reduced-order case, $p = N_\theta + 1$.

Another important aspect when constructing the discrete algorithm is to take into account the norm of the space where the operations are done while discretizing. For instance, the operator $C \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ has its matrix form obtained as in Property A.2.4, so

$$\forall z \in \mathcal{Z} = \mathbb{R}^{N_z}, \forall y \in \mathcal{Y} = \mathbb{R}^{N_\theta}, \quad (Cz, y)_{\mathcal{Y}} = z^\intercal \mathbb{C}^\intercal \mathbb{Y} y.$$

This is further illustrated below.

ROUKF implementation. Due to its particularity concerning the construction of sigma points and the computation of the covariance, it is interesting to detail the discrete implementation of this algorithm as done in [Moireau and Chapelle, 2011]. Considering that we have a set $\{\widetilde{I}^{(i)}\}_{i=1}^{p} \in \Theta^{p}$ of sigma-points being the columns of the matrix $[\widetilde{I}^{(*)}]$ generated as in (3.75). By defining the matrix

$$\mathbb{D}_{\alpha} = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \alpha \end{pmatrix}$$

and a set of particles $\{x^{(i)}\}_{i=1}^{p}$ analogously represented in its column-matrix form $[x^{(*)}]$, one can show that

$$\mathrm{Cov}_{\alpha}(x^{(*)}, \widetilde{I}^{(*)}) = [x^{(*)}] D_{\alpha} [\widetilde{I}^{(*)}]^{\intercal},$$

leading to the fully discrete algorithm below.

$$\begin{aligned}
&\text{(Initialization)} \begin{cases} \mathbb{U}^{0} & = \mathbb{R}_{\Theta}, \\ \mathbb{L}_{z\theta}^{0} & = 0, \quad \mathbb{L}_{\theta}^{0} = \mathbb{I}_{\Theta}, \\ \widehat{z}_{+}^{0} & = z_{0}. \end{cases} \\
&\text{(Sampling)} \begin{cases} \widehat{z}_{+}^{n(i)} & = \widehat{z}_{+}^{n}, \\ \widehat{\theta}_{+}^{n(i)} & = \widehat{\theta}_{+}^{n} + \mathbb{L}_{\theta}^{n} \sqrt{(\mathbb{U}^{n})^{-1}} \widetilde{I}^{(i)}. \end{cases} \\
&\text{(Prediction)} \begin{cases} \widehat{z}_{-}^{n+1(i)} & = \Phi_{n+1|n}(\widehat{\theta}_{+}^{n}) \widehat{z}_{+}^{n(i)} + b^{n}, \\ \widehat{\theta}_{-}^{n+1(i)} & = \widehat{\theta}_{+}^{n(i)}, \\ \widehat{y}^{n+1(i)} & = C(\widehat{z}_{-}^{n+1(i)}), \\ \widehat{z}_{-}^{n+1} & = E_{\alpha}(\widehat{z}_{-}^{n+1(*)}). \end{cases} \\
&\text{(Correction)} \begin{cases} \mathbb{L}_{z\theta}^{n+1} & = [\widehat{z}_{-}^{n+1(*)}] \mathbb{D}_{\alpha} [\widetilde{I}^{(*)}], \\ \mathbb{L}_{\theta}^{n+1} & = [\widehat{\theta}_{-}^{n+1(*)}] \mathbb{D}_{\alpha} [\widetilde{I}^{(*)}], \\ \mathbb{L}_{y\theta}^{n+1} & = [\widehat{y}^{n+1(*)}] \mathbb{D}_{\alpha} [\widetilde{I}^{(*)}], \\ \mathbb{U}^{n+1} & = \mathbb{I} + \Delta t (\mathbb{L}_{y\theta}^{n+1})^{\intercal} \mathbb{Y} \mathbb{L}_{y\theta}^{n+1}, \\ \widehat{z}_{+}^{n+1} & = \widehat{z}_{-}^{n+1} + \Delta t \mathbb{L}_{z\theta}^{n+1} (\mathbb{U}^{n+1})^{-1} (\mathbb{L}_{y\theta}^{n+1})^{\intercal} \mathbb{Y} \big(y^{n} - E_{\alpha}(\widehat{y}^{n+1(*)})\big), \\ \widehat{\theta}_{+}^{n+1} & = \widehat{\theta}_{-}^{n+1} + \Delta t \mathbb{L}_{\theta}^{n+1} (\mathbb{U}^{n+1})^{-1} (\mathbb{L}_{y\theta}^{n+1})^{\intercal} \mathbb{Y} \big(y^{n} - E_{\alpha}(\widehat{y}^{n+1(*)})\big). \end{cases}
\end{aligned} \tag{3.76}$$

In this case, the particles related to the sigma points can be run in parallel.

3.4 Application to one-dimensional wave propagation inverse problems

In this section, we apply the presented methods in the previous sections to one-dimensional wave propagation problems, while alluding to potential industrial applications. More specifically, we focus on parameter reconstruction by presenting two problems: a **linear-quadratic** problem (Section 3.4.1) as in Problem B.2 and a **nonlinear** problem (Section 3.4.2 and Section 3.4.3) as in Problem B.1. In both cases, we have a one-dimensional domain Ω and we will aim at constructing a parameter $\theta \in \Theta(\Omega)$ that varies in space. The parametric space is constructed by an arbitrary finite-dimensional basis $\{\psi_{i}^{\Theta}\}_{i=1}^{N_{\Theta}}$

$$\forall \theta \in \Theta(\Omega), \quad \exists! \{\theta_{i}\}_{i=1}^{N_{\Theta}} \in \mathbb{R}^{N_{\Theta}}, \quad \theta(x) = \sum_{i=1}^{N_{\Theta}} \theta_{i} \psi_{i}^{\Theta}(x), \quad \forall x \in \Omega. \tag{3.77}$$

We recall that no distinction will be made between the function $\theta \in \Theta(\Omega)$ and its corresponding vector of components $\theta \in \mathbb{R}^{N_{\Theta}}$. Both problems will be related to a scalar displacement field u_{θ}, associated with a parameter θ, being the solution of wave propagation dynamics. The wave propagation problem will be written as a second-order time-differential equation. For the first problem treated in Section 3.4.1, this parameter will represent the spatial dependency of a source term. In the second problem, the parameter will represent the temperature of the material that is related to the wave velocity in the medium. It is treated in Section 3.4.2 using the sensitivity for computing the functional gradient with the objective of detailing the development of the inverse strategy and discussing some particularities of wave propagation-related inverse problems. The second problem is also treated in Section 3.4.3 with sequential methods.

3.4.1 Linear-quadratic problem: Estimating the spatial dependency of a source term

Here, we use the methods presented in previous sections to treat a linear-quadratic inverse problem analog to Problem B.2. Let the unknown $\theta \in \Theta$ be a function of space representing the spatial dependency of the source term of a wave propagation problem. Let $u_\theta(t) \in \mathrm{H}^1(\Omega)$ be the displacement field associated with θ, the wave propagation problem reads

$$\begin{cases} \partial^2_{tt} u_\theta(t)(x) - c^2 \partial_{xx} u_\theta(t)(x) = \theta(x) f(t), & \text{in } \Omega \times]0;T[, \\ \partial_x u_\theta(t)(0) = \partial_x u_\theta(t)(L) = 0, \\ u_\theta(0)(x) = \partial_t u(0)(x) = 0. \end{cases} \tag{3.78}$$

The displacement field is observed through a linear operator $C_u \in \mathcal{L}(\mathrm{H}^1(\Omega), \mathcal{Y})$ and the observed data is

$$y(t) = C_u u^*(t),$$

where u^* is the wavefield associated with θ^*, the parameter of the target system. For this illustration case, no disturbance is present in the observations. Considering the parameter $\theta = \theta_0 + \zeta_\theta$, where θ_0 is an initial guess, we interpret the inverse problem of estimating the error ζ_θ as minimizing

$$\mathcal{J}_T(\zeta_\theta) = \underset{\zeta_\theta \in \Theta}{\operatorname{argmin}} \left\{ r_\Theta(\zeta_\theta) + \frac{1}{2} \int_0^T \| y(t) - C_u u_\theta(t) \|^2_{\mathcal{Y}} \, \mathrm{d}t \right\}$$

where we define the regularization analogue to the $\mathrm{H}^1(\Omega)$ norm as

$$r_\Theta(\zeta_\theta) = \beta \int_\Omega (\zeta_\theta)^2 \mathrm{d}\Omega + \gamma \int_\Omega (\partial_x \zeta_\theta)^2 \mathrm{d}\Omega,$$

with β and γ allowing to adjust the regularization.

3.4.1.1 Discretized problem

We present here the discrete version of the estimation problem and brief details of the discretization procedure. Approximating the functional space using the finite elements method as presented in Section 1.3, we have

$$V_h(\Omega) = \{ v_h \in C^0(\overline{\Omega}), \quad \forall K \in \mathcal{T}_h, \quad v_h\Big|_K \in P^k(K) \},$$

representing an approximation of the functional space $\mathrm{H}^1(\Omega)$ where we look for the solution of the wave propagation problems. We define $\mathcal{T}_h$ as the one-dimensional mesh, related to the mesh step h, and P^k the space of Lagrange polynomials of order k. The approximation space is generated by the set of Lagrange basis functions

$$V_h(\Omega) = \operatorname{span}\{\psi_I\}_{I=1}^{N_h},$$

where we associate $\{\psi_I\}_{I=1}^{N_h}$ to the interpolation points

$$\{\xi_I\}_{I=1}^{N_h} \subset \overline{\Omega} \quad \text{with} \quad \psi_I(\xi_J) = \delta_{IJ}, \quad \forall I, J \in [\![1; N_h]\!].$$

Using Property A.2.2, we represent the displacement field $u_\theta \in V_h(\Omega)$ by the vector $\begin{pmatrix} u_1 & \cdots u_{N_h} \end{pmatrix}$. Also, by finite elements procedures, we have the associated mass and stiffness matrices $\mathbb{M}$ and $\mathbb{K}$, respectively. We further detail the assembly of such matrices in the next section, when illustrating the nonlinear case. In the following, we denote this finite element vector as $u_\theta \in \mathbb{R}^{N_h}$. Using a second-order time scheme and discretizing the operators we can write the discrete dynamics

$$\mathbb{M}\,\frac{u_\theta^{n+1} - 2u_\theta^n + u_\theta^{n-1}}{\Delta t^2} + \mathbb{K}\,u_\theta^n = f^n \mathbb{P}\theta, \quad \forall n \in \llbracket 1; N-1 \rrbracket,$$

associated with zero initial conditions and where $f^n = f(n\Delta t)$ and $\mathbb{P}_{ij} = \psi_j^\Theta(\xi_i)$ is the matrix that projects $\Theta(\Omega)$ into $V_h(\Omega)$. We rewrite the discrete wave equation as done in (3.15)

$$\Phi_1 z^{n+1} = \Phi_0 z^n + \mathbb{B}_0^n \theta, \quad z^0 = 0, \quad \forall n \in \llbracket 0; N-1 \rrbracket,$$

with

$$\Phi_1 = \begin{pmatrix} \mathbb{M} & -\Delta t \mathbb{M} \\ 0 & \mathbb{M} \end{pmatrix}, \quad \Phi_0 = \begin{pmatrix} \mathbb{M} & 0 \\ -\Delta t \mathbb{K} & \mathbb{M} \end{pmatrix}, \quad \mathbb{B}_0^n = \begin{pmatrix} 0 \\ \Delta t f^n \mathbb{P} \end{pmatrix}.$$

We proceed by writing the dynamics in the same form as in Problem II.2, with an augmented state-parameter variable ${}^\sharp z_\theta^n = \begin{pmatrix} z^n & \theta \end{pmatrix}^\intercal$ leading to the dynamics

$${}^\sharp z_\theta^{n+1} = {}^\sharp\Phi_{n+1|n}\, {}^\sharp z_\theta^n, \quad \forall n \in \llbracket 0; N-1 \rrbracket,$$

with

$${}^\sharp\Phi_{n+1|n} = {}^\sharp\Phi_1^{-1}\, {}^\sharp\Phi_0^n, \quad {}^\sharp\Phi_1 = \begin{pmatrix} \Phi_1 & 0 \\ 0 & I_\Theta \end{pmatrix}, \quad {}^\sharp\Phi_0^n = \begin{pmatrix} \Phi_0 & \mathbb{B}_0^n \\ 0 & I_\Theta \end{pmatrix}.$$

Then, the inverse problem in its discrete form reads

$$\overline{\zeta_\theta} = \operatorname*{argmin}_{\zeta_\theta \in \mathbb{R}^{N_\theta}} \mathcal{J}_N(\zeta_\theta) = \operatorname*{argmin}_{\zeta_\theta \in \mathbb{R}^{N_\theta}} \left\{ r_\Theta(\zeta_\theta) + \frac{\Delta t}{2} \sum_{n=0}^{N} \|y^n - {}^\sharp\mathbb{C}\, {}^\sharp z_\theta{}^n\|_{\mathcal{Y}}^2 \right\}, \tag{3.79}$$

where the discrete augmented observation operator is

$${}^\sharp\mathbb{C} = \begin{pmatrix} \mathbb{C} & 0 \end{pmatrix} \text{ and } \mathbb{C} = \begin{pmatrix} \mathbb{C}_u & \mathbb{C}_v \end{pmatrix}.$$

Except for the initial detailing of the adjoint method, we will consider for all cases $\mathbb{C}_v = 0$. We consider that the observation operator $\mathbb{C}_u$ retrieves the displacement field in a restricted domain ω. The discrete space of observations is defined as $\mathcal{Y} = \mathbb{R}^{N_y}$ with associated norm

$$\|y^n\|_{\mathcal{Y}}^2 = y^{n\intercal}\mathbb{Y}y^n, \quad \forall y^n \in \mathcal{Y}, \quad \forall n \in \llbracket 0; N \rrbracket,$$

where $\mathbb{Y}$ is a positive-definite symmetric matrix. The regularization is represented by the positive-definite symmetric matrix $\mathbb{R}_\Theta$,

$$r_\Theta(\zeta_\theta) = \frac{1}{2}(R_\Theta \zeta_\theta, \zeta_\theta)_\Theta = \frac{1}{2}\zeta_\theta{}^\intercal \mathbb{R}_\Theta \zeta_\theta,$$

being a combination of the mass and stifness matrices, as precised later in this section. In the following, to implement the different methods to minimize (3.79), we first compute the functional gradient to apply variational methods. We present how to compute the gradient using sensitivity and then the adjoint method. We show that by using the sensitivity dynamics when the problem is sufficiently regularized, we can compute directly the functional minimizer and use it as the reference to assess the other strategies. Using the gradient obtained by the adjoint method, we apply the gradient descent class of methods: steepest descent with optimal fixed step, the conjugate gradient with Fletcher-Reeves coefficient, the Gauss-Newton method and the BFGS. For the sequential methods, we apply the ROKF and ROUKF methods.

Functional gradient using sensitivity. First, we present the computation of the gradient using the sensitivity of the dynamics *w.r.t.* the parameter. As the system depends linearly on the parameter, the sensitivity $\mathrm{D}_{\zeta_\theta} z_\theta \in \mathcal{L}(\Theta, \mathcal{Z})$ in its discrete form $\mathbb{S}_{\zeta_\theta} \in \mathcal{M}(\mathbb{R})_{N_z \times N_\theta}$ is computed as

$$\begin{cases} \mathbb{S}_{\zeta_\theta}^{n+1} = \Phi_1^{-1}\Phi_0^n \mathbb{S}_{\zeta_\theta}^n + \Phi_1^{-1}\mathbb{B}_0^n, \\ \mathbb{S}_{\zeta_\theta}^0 = 0, \end{cases}$$

where

$$\mathbb{S}_{\zeta_\theta}^n = \begin{pmatrix} \mathbb{S}_u^n \\ \mathbb{S}_v^n \end{pmatrix}$$

Denoting the i-th column of $\mathbb{S}_u$ as s_i, we can compute the displacement sensitivity matrix column by column with a set of wave propagation problems, namely for $i \in [\![1; N_\theta]\!]$,

$$\mathbb{M}\,\frac{s_i^{n+1} - 2s_i^n + s_i^{n-1}}{\Delta t^2} + \mathbb{K}\, s_i^n = f^n \mathbb{P}_i, \quad s_i^1 = s_i^0 = 0, \quad \forall n \in [\![1; N-1]\!]$$

where $\mathbb{P}_i$ is the i-th column of $\mathbb{P}$. One can show that $u_\theta^n = \mathbb{S}_u^n \theta$. Then, the gradient of the functional (3.79) can be computed using the sensitivity matrix as

$$\nabla \mathcal{J}_N(\zeta_\theta) = \mathbb{R}_\Theta \zeta_\theta - \Delta t \sum_{n=0}^{N} (\mathbb{S}_u^n)^\intercal \mathbb{C}_u^\intercal \mathbb{Y}(y^n - \mathbb{C}_u u_\theta^n).$$

Optimal solution. Using the sensitivity, we can compute directly the minimizer of (3.79) and use it as the reference. As $y^n = \mathbb{C}_u z_{\theta^*} = \mathbb{C}_u \mathbb{S}_u^n \theta^*$ and $u_\theta^n = \mathbb{S}_u^n(\theta_0 + \zeta_\theta)$, the observed target system, we have from the functional gradient above,

$$\begin{aligned} \nabla \mathcal{J}_N(\zeta_\theta) &= \mathbb{R}_\Theta \zeta_\theta - \Delta t \sum_{n=0}^{N} (\mathbb{S}_u^n)^\intercal \mathbb{C}_u^\intercal \mathbb{Y}\big(\mathbb{C}_u \mathbb{S}_u^n \theta^* - \mathbb{C}_u \mathbb{S}_u^n(\theta_0 + \zeta_\theta)\big) \\ &= \mathbb{R}_\Theta \zeta_\theta - \left(\Delta t \sum_{n=0}^{N} (\mathbb{S}_u^n)^\intercal \mathbb{C}_u^\intercal \mathbb{Y} \mathbb{C}_u \mathbb{S}_u^n\right)(\theta^* - \theta_0 - \zeta_\theta), \end{aligned}$$

from where we define the Gramian

$$\mathbb{G}_r = \Delta t \sum_{n=0}^{N} (\mathbb{S}_u^n)^\intercal \mathbb{C}_u^\intercal \mathbb{Y} \mathbb{C}_u \mathbb{S}_u^n.$$

Using the Gramian, we have

$$\nabla \mathcal{J}_N(\zeta_\theta) = \mathbb{R}_\Theta \zeta_\theta - \mathbb{G}_r(\theta^* - \theta_0 - \zeta_\theta),$$

so the minimizer $\overline{\zeta_\theta}$ where $\nabla \mathcal{J}_N(\overline{\zeta_\theta}) = 0$ is

$$\overline{\zeta_\theta} = \big(\mathbb{R}_\Theta + \mathbb{G}_r\big)^{-1} \mathbb{G}_r(\theta^* - \theta_0).$$

We can use this optimal solution as a reference to compare and validate the minimization methods. Due to regularization the target θ^* differs from the optimal $\theta_0 + \overline{\zeta_\theta}$, even though no disturbance is present in the observations y. One can expect that, as $\mathbb{C}_u$ is usually a low-rank matrix, the well-posedness of the inversion of $\mathbb{R}_\Theta + \mathbb{G}_r$ is determined by the presence and quality of $\mathbb{R}_\Theta$. Therefore, the conditioning of the Gramian carries a measurement of observability of the unknown for the given observation method.

Computing the gradient using the adjoint method. As previously discussed, the sensitivity approach to compute the gradient requires, for each evaluation, solving several wave propagation problems. The adjoint method is detailed here for this illustration case. In the case of a wave equation, the transition operator in the augmented version of the adjoint dynamics (3.53) reads

$$^{\sharp}\Phi^{\intercal}_{n+1|n} = {}^{\sharp}\Phi_0^{n\,\intercal}\,{}^{\sharp}\Phi_1^{-\intercal}.$$

Referring to (3.53) and introducing the variable ${}^{\sharp}q^n = {}^{\sharp}\Phi_1^{-\intercal} p_{\zeta}^n$, we can write the adjoint dynamics as a backward propagation problem

$$^{\sharp}\Phi_1^{\intercal}\,{}^{\sharp}q^n - {}^{\sharp}\Phi_0^{n\,\intercal}\,{}^{\sharp}q^{n+1} = \Delta t\,{}^{\sharp}\mathbb{C}^{\intercal}\mathbb{Y}(y^n - {}^{\sharp}\mathbb{C}\,{}^{\sharp}z^n).$$

Separating the state and parameter part of this augmented adjoint variable as ${}^{\sharp}q^{n\,\intercal} = \left(q_u^{n\,\intercal}\ q_v^{n\,\intercal}\ q_\theta^{n\,\intercal}\right)$, we obtain

$$\begin{cases}
\mathbb{M}\,\dfrac{q_u^n - q_u^{n+1}}{\Delta t} + \mathbb{K}\,q_v^{n+1} = \Delta t \mathbb{C}_u^{\intercal}\mathbb{Y}(y^n - \mathbb{C}_u u^n - \mathbb{C}_v v^n), \quad \forall n \in [\![N;0]\!],\\[2ex]
\mathbb{M}\,\dfrac{q_v^n - q_v^{n+1}}{\Delta t} - \mathbb{M}\,q_u^n = \Delta t \mathbb{C}_v^{\intercal}\mathbb{Y}(y^n - \mathbb{C}_u u^n - \mathbb{C}_v v^n), \quad \forall n \in [\![N;0]\!],\\[2ex]
\dfrac{q_\theta^n - q_\theta^{n+1}}{\Delta t} = \mathbb{B}_0^{n\,\intercal} q_v^{n+1}, \quad \forall n \in [\![N;0]\!],\\[2ex]
q_u^{N+1} = q_v^{N+1} = 0, \quad q_\theta^{N+1} = 0.
\end{cases}$$

Assuming pure displacement observations, i.e. $\mathbb{C}_v = 0$, this system reduces to

$$\begin{cases}
\mathbb{M}\,\dfrac{q_v^{n-1} - 2q_v^n + q_v^{n+1}}{\Delta t^2} + \mathbb{K}\,q_v^n = \Delta t \mathbb{C}_u^{\intercal}\mathbb{Y}(y^{n-1} - \mathbb{C}_u u^{n-1}), \qquad \forall n \in [\![N;1]\!],\\[2ex]
\dfrac{q_\theta^n - q_\theta^{n+1}}{\Delta t} = \mathbb{B}_0^{n\,\intercal} q_v^{n+1}, \qquad \forall n \in [\![N;0]\!],
\end{cases}$$

with the following final conditions

$$q_v^N = (\Delta t^2)\Delta t \mathbb{M}^{-1}\mathbb{C}_u^{\intercal}\mathbb{Y}(y^N - \mathbb{C}_u u^N), \quad q_v^{N+1} = 0, \quad q_\theta^{N+1} = 0.$$

Remarking that

$$^{\sharp}\Phi_1^{\intercal} q^n = \begin{pmatrix} \Phi_1^{\intercal} & 0 \\ 0 & \mathbb{I}_\Theta \end{pmatrix} \begin{pmatrix} q_z^n \\ q_\theta^n \end{pmatrix} = \begin{pmatrix} p_z^n \\ p_\theta^n \end{pmatrix},$$

which, in particular, entails that $q_\theta^n = p_\theta^n$, we write the gradient from (3.52) as

$$\nabla \mathcal{J}_N(\zeta_\theta) = \mathbb{R}_\Theta \zeta_\theta - q_\theta^0.$$

To compute it, we develop the dynamic of q_θ we have

$$q_\theta^n = q_\theta^{n+1} + \Delta t \mathbb{B}_0^{n\,\intercal} q_v^{n+1}, \quad \forall n \in [\![N;0]\!],$$

and denesting the recurrency of q_θ in terms of q_v,

$$q_\theta^0 = \Delta t \left(\mathbb{B}_0^{N-1\,\intercal} q_v^N + \mathbb{B}_0^{N-2\,\intercal} q_v^{N-1} + \cdots + \mathbb{B}_0^{0\,\intercal} q_v^1\right), \quad \forall n \in [\![N;0]\!],$$

we can write the gradient as:

$$\nabla\mathfrak{J}_N(\zeta_\theta) = \mathbb{R}_\Theta\zeta_\theta - \Delta t\sum_{n=0}^{N-1}\mathbb{B}_0^{n\intercal}q_v^{n+1}. \tag{3.80}$$

The associated adjoint system for computing this gradient, for u_θ, is

$$\begin{cases}\mathbb{M}\,\dfrac{u_\theta^{n+1}-2u_\theta^n+u_\theta^{n-1}}{\Delta t^2}+\mathbb{K}\,u_\theta^n = \mathbb{B}_0^n(\theta_0+\zeta_\theta), \quad \forall n\in[\![0;N-1]\!],\\[2em] \mathbb{M}\,\dfrac{q_v^{n-1}-2q_v^n+q_v^{n+1}}{\Delta t^2}+\mathbb{K}\,q_v^n = \Delta t\mathbb{C}_u^\intercal\mathbb{Y}(y^{n-1}-\mathbb{C}_u u_\theta^{n-1}), \quad \forall n\in[\![N+1;2]\!],\\[2em] u_\theta^{-1}=u_\theta^0=q_v^{N+2}=q_v^{N+1}=0.\end{cases}$$

This method is used in the following to compute the functional gradient.

Steepest descent and optimal descent step. By choosing the steepest descent direction (3.40) and using the gradient computed using the expression (3.80) we perform a gradient descent with fixed descent step. We choose a fixed optimal descent step α as in [Ciarlet, 1988a] by computing the set $\{\lambda_i\}$ of eigenvalues of the following eigenproblem

$$\mathbb{G}_r v = \lambda\mathbb{R}_\Theta v$$

and taking

$$\alpha = \max\left\{\frac{\min\{\lambda_i\}}{\max\{\lambda_i\}^2};\frac{2}{\min\{\lambda_i\}+\max\{\lambda_i\}}\right\}.$$

This gives the larger stable step α. We then update the estimated error as

$${\zeta_\theta}^{k+1} = {\zeta_\theta}^k - \alpha\nabla\mathfrak{J}_N({\zeta_\theta}^k).$$

Conjugate gradient. For the conjugate gradient implementation, the error is updated as

$${\zeta_\theta}^{k+1} = {\zeta_\theta}^k - d^k$$

with d^k = choosen as in (3.41) and for β^k we use the Fletcher-Reeves coefficient

$$\beta^k = \frac{(\nabla\mathfrak{J}_N({\zeta_\theta}^{k+1}))^\intercal\nabla\mathfrak{J}_N({\zeta_\theta}^{k+1})}{(\nabla\mathfrak{J}_N({\zeta_\theta}^{k}))^\intercal\nabla\mathfrak{J}_N({\zeta_\theta}^{k})}.$$

As the functional is quadratic *w.r.t.* the parameter, we perform an exact line search by evaluating the functional at three different points.

Gauss-Newton method. Using the gradient and the sensitivity we can perform the descent iterations (3.42) using the Gauss-Newton approximation (3.44) of the Hessian. The Gauss-Newton approximation of the Hessian using the sensitivity is equivalent to the Gramian

$$\widetilde{H}^{(k)} = \Delta t\sum_{n=0}^{N}(\mathbb{S}_u^n)^\intercal\mathbb{C}_u^\intercal\mathbb{Y}\mathbb{C}_u\mathbb{S}_u^n = \mathbb{G}_r.$$

One can show that for the linear-quadratic case, this is exactly the Hessian, hence only one iteration is needed to attain the minimizer.

BFGS. We use the BFGS approximation of the Hessian inverse (3.46) by initializing it with

$$\widetilde{\mathbb{B}}^{(0)} = \mathbb{R}_{\Theta}^{-1}, \quad s^{(0)} = -\nabla \mathcal{J}_N({\zeta_\theta}^0)$$

then

$$\nabla \mathcal{J}_N({\zeta_\theta}^1) = \nabla \mathcal{J}_N(\theta^0 + s^{(0)}).$$

Then, the descent is performed as in (3.42) with an optimal descent step α_k as done for the conjugate gradient case.

Reduced-Order Kalman Filter. The first sequential method implemented for this illustration is the ROKF. We rewrite here the algorithm (3.72) using the case-specific discrete definitions and norms

$$\text{(Initialization)} \begin{cases} \mathbb{L}_z^0 & = 0 \\ \mathbb{U}^0 & = \mathbb{R}_\Theta \\ \widehat{z}_+^0 & = z_0 + \mathbb{G}^0(\widetilde{y}^0 - \mathbb{C} z_0) \\ \widehat{\zeta}_{\theta +}^{\,0} & = 0 \end{cases}$$

$$\text{(Prediction)} \begin{cases} \widehat{z}_-^{n+1} & = \Phi_1^{-1}\Phi_0^n \widehat{z}_+^n + \Phi_1^{-1}\mathbb{B}_0^n(\theta_0 + \widehat{\zeta}_{\theta +}^{\,n}) \\ \widehat{\zeta}_{\theta -}^{\,n+1} & = \widehat{\zeta}_{\theta +}^{\,n} \end{cases}$$

$$\text{(Correction)} \begin{cases} \mathbb{L}_z^{n+1} & = \Phi_1^{-1}\Phi_0^n \mathbb{L}_z^n + \Phi_1^{-1}\mathbb{B}_0^n \\ \mathbb{U}^{n+1} & = \mathbb{U}^n + \Delta t (\mathbb{L}_z^{n+1})^\intercal \mathbb{C}^\intercal \mathbb{C} \mathbb{L}_z^{n+1} \\ \mathbb{G}_{\zeta_\theta} & = \Delta t (\mathbb{U}^{n+1})^{-1} (\mathbb{L}_z^{n+1})^\intercal \mathbb{Y} \mathbb{C}^\intercal \mathbb{Y}, \\ \widehat{\zeta}_{\theta +}^{\,n+1} & = \widehat{\zeta}_{\theta -}^{\,n+1} + \mathbb{G}_{\zeta_\theta}(y^{n+1} - \mathbb{C}\widehat{z}_-^{n+1}) \\ \widehat{z}_+^{n+1} & = \widehat{z}_-^{n+1} + \Delta t \mathbb{L}_z^{n+1}(\widehat{\zeta}_{\theta +}^{\,n+1} - \widehat{\zeta}_{\theta -}^{\,n+1}) \end{cases}$$

We recognize that $\mathbb{L}_z^n = \mathbb{S}_{\zeta_\theta}^n$, the sensitivity of the state *w.r.t.* the parameter. Using the definitions of Φ_1 and Φ_0, replacing it in the discrete algorithm above, we have

$$\text{(Initialization)} \begin{cases} \mathbb{S}_u^0 & = \mathbb{S}_u^1 = 0 \\ \widehat{u}_+^0 & = \widehat{u}_+^1 = 0 \\ \mathbb{U}^0 & = \mathbb{R}_\Theta \\ \widehat{\zeta}_\theta^{\,1} & = 0 \end{cases}$$

$$\text{(Prediction)} \begin{cases} \mathbb{M}\dfrac{\widehat{u}_-^{n+1} - 2\widehat{u}_+^n + \widehat{u}_+^{n-1}}{\Delta t^2} + \mathbb{K}\,\widehat{u}_+^n = f^n \mathbb{P}(\theta_0 + \widehat{\zeta}_\theta^{\,n}), \\ \mathbb{M}\dfrac{\mathbb{S}_u^{n+1} - 2\mathbb{S}_u^n + \mathbb{S}_u^{n-1}}{\Delta t^2} + \mathbb{K}\,\mathbb{S}_u^n = f^n \mathbb{P} \end{cases}$$

$$\text{(Correction)} \begin{cases} \mathbb{U}^{n+1} & = \mathbb{U}^n + \Delta t (\mathbb{S}_u^{n+1})^\intercal \mathbb{C}_u^\intercal \mathbb{Y} \mathbb{C}_u \mathbb{S}_u^{n+1} \\ \mathbb{G}_\theta & = \Delta t (\mathbb{U}^{n+1})^{-1} (\mathbb{S}_u^{n+1})^\intercal \mathbb{C}_u^\intercal \mathbb{Y} \\ \widehat{\zeta}_\theta^{\,n+1} & = \widehat{\zeta}_\theta^{\,n} + \mathbb{G}_\theta(y^{n+1} - \widehat{u}_-^{n+1}) \\ \widehat{u}_+^{n+1} & = \widehat{u}_-^{n+1} + \mathbb{S}_u^{n+1}\mathbb{G}_\theta(y^{n+1} - \widehat{u}_-^{n+1}) \\ \widehat{u}_+^n & = \widehat{u}_-^n + \mathbb{S}_u^n \mathbb{G}_\theta(y^{n+1} - \widehat{u}_-^{n+1}) \end{cases}$$

Reduced-Order Unscented Kalman Filter. The ROUKF method is implemented as described in (3.76), using the simplex *sigma-points*.

Illustration specifications. We specify here the remaining configuration for the one-dimensional wave propagation inverse problem. We define a domain $\Omega =]0, 600[$ mm and a duration of $T = 400\mu$s for the dynamics. The observation operator retrieves the displacement field at the observation region

$\omega =]0, 100[$ mm. The wave-speed c is defined as $3\text{mm} \cdot \mu\text{s}^{-1}$. The time dependency f of the source is a gated cosine

$$f(t) = \mathrm{e}^{-\frac{(t-t_0)^2}{\sigma^2}} \cos\left(2\pi\omega_f(t - t_0)\right)$$

with $t_0 = 50\mu$s, $\sigma = 12$mm and $\omega_f = 100$kHz. We choose the same basis for the parametric space as the displacement field

$$\{\psi_i^{\Theta}\}_{i=1}^{N_\Theta} = \{\psi_i\}_{i=1}^{N_h},$$

meaning that we reconstruct the finite element vector for θ. With this choice, we can write the equivalent discrete version of the regularization as $\mathbb{R}_\Theta = \beta\mathbb{M} + \gamma\mathbb{K}$, so

$$r_\Theta(\zeta_\theta) = \frac{1}{2}\zeta_\theta{}^\intercal\left(\beta\mathbb{M} + \gamma\mathbb{K}\right)\zeta_\theta.$$

The space $V_h(\Omega)$ is constructed using Lagrange polynomials of fourth-order as local basis functions. We construct the mesh $\mathcal{T}_h$ to ensure at least 2 elements per wavelength. The wavelength is computed using the wave speed c and frequency ω_f. The time step satisfies the CFL condition (more details in Chapter 1)

$$\Delta t \leq \frac{2}{\sqrt{r(\mathbb{M}^{-1}\mathbb{K})}},$$

where r is the spectral radius. In this illustration, we reduce the time step by a factor of 0.9 as the spectral radius is computed approximately. The maximum number of iterations for the gradient descent methods is fixed at 1300. Denoting the mass matrix in $V_h(\omega)$ as $\mathbb{M}_\omega$ we consider $\mathbb{Y} = \mathbb{M}_\omega$, since we want to use the L^2 norm for the observation space. The objective will be to reconstruct the spatial dependency of the source term $\theta^*(x)$ from a first guess $\theta_0(x)$ and regularization defined for β and γ. These parameters are generated by projecting the Gaussian

$$A\frac{e^{-(\frac{x-x_0}{2\sigma})^2}}{\sigma\sqrt{2\pi}}$$

into Θ, with $\sigma = 10$mm. For the guess, $A = 32$ and $x_0 = 400$mm and for the target $A = 40$ and $x_0 = 500$mm. The guess and target parameters are shown in Figure 3.1. As the basis for the parametric space is the finite element space, arbitrary forms of spatial dependency could be chosen. Snapshots of the u_{θ^*} are shown in Figure 3.2.

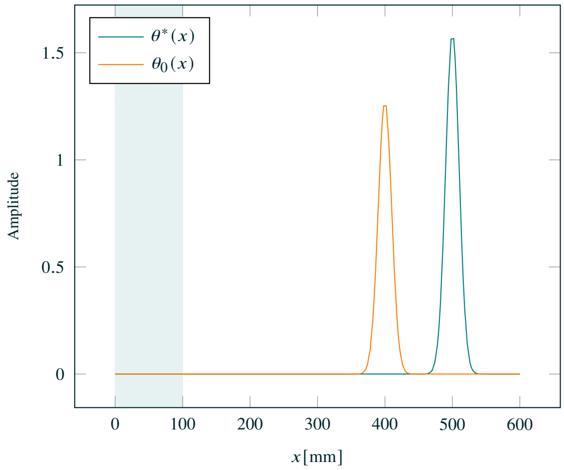


Figure 3.1 – Target and guess spatial dependency for the linear-quadratic inverse problem illustration.

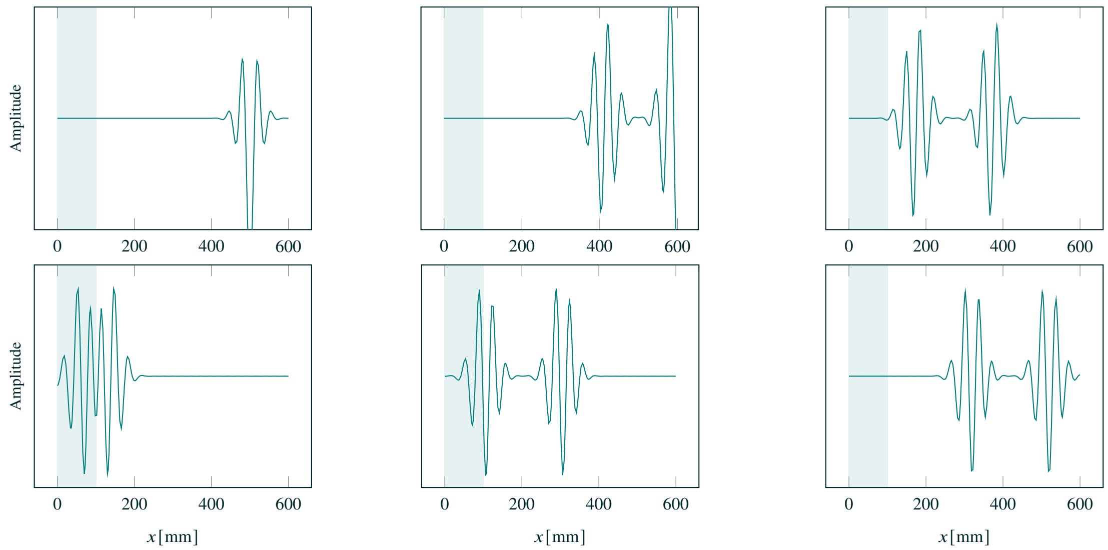


Figure 3.2 – Snapshots of the target displacement field u_{θ^*} at $t = 45, 80, 150, 230, 300$ and 380μs (left-right, top-bottom). The observed region is highlighted.

Convergence rate results. Using the optimal solution as the reference, we compute the relative squared error *w.r.t.* the optimal solution at each iteration step, for variational methods, and at each time step, for sequential methods using

$$\frac{\left\|\theta^i - \theta^*\right\|_\Theta^2}{\|\theta^*\|_\Theta^2} = \frac{(\theta^i - \theta^*)^\top \mathbb{M}\,(\theta^i - \theta^*)}{(\theta^*)^\top \mathbb{M}\,(\theta^*)},$$

where θ^i is the i-th estimated parameter. For these computations, the regularization was adjusted with $\beta = 10^{-10}$ and $\gamma = 10^{-5}$. As shown in Section 3.2 and Section 3.3, variational and sequential methods are equivalent for linear-quadratic problems, *i.e.* both converge towards the minimizer. The error is plotted for both classes of methods in Figure 3.3. For the variational methods, we computed 1300 iterations. The Gauss-Newton method took one iteration to achieve a relative squared error of 10^{-13}. Even with an optimal descent step, the Steepest Descent case did not converge. The Conjugate Gradient method performed slightly better but still presented a large error. The BFGS method presented the best results among the iterative variational methods when discarding the Gauss-Newton method. For the sequential methods, both ROKF and ROUKF achieved relative squared errors below 10^{-12} in one run/use of the observed data.

Cost analysis and discussion. We analyze here the memory and computational cost for these methods, summarized in Table 3.1 and Table 3.2. We denote by N_{RAM} the amount of memory necessary for running one wave propagation problem, also called forward problem. For the steepest descent, although an optimal descent step is used, it presents very low performance. If the fixed descent step is chosen using other techniques, it would require propagating only two forward problems for each step. Other techniques for adapting the descent step can be sought [Nocedal and Wright, 2006]. The Conjugate Gradient presented an improved convergence with the additional cost of performing a line search but keeping relatively low memory requirements. The Gauss-Newton method presents a one-shot convergence for the linear-quadratic case as the functional is convex but requires N_θ forward problems to be solved to compute the Gramian and a $N_\theta \times N_\theta$ matrix inversion.

In the case of the sequential methods ROKF and ROUKF, instead of computing the gradient at each estimation, the estimation is corrected using the observed data step by step, so the computational cost is estimated step by step. Both sequential methods present a one-shot convergence for the linear-quadratic case, although the ROUKF propagates the covariance empirically. At each time step, the inversion of the

Method	**Memory Cost**	**Computational Cost**	**Conclusion**
Steepest Descent w. optimal step	• N_θ (Gradient) • $N_{RAM} \times N_\theta$ (Sensitivity) • $2N_{RAM}$ (Adjoint)	• N_θ forwards (Once) • 2 forwards (Adjoint)	– Low performance, – Sensitivity comput..
Conjugate Gradient	• $2N_\theta$ (Gradients) • $2N_{RAM}$ (Adjoint)	• 2 forwards (Adjoint) • $\sim 3\mathcal{J}_T$ eval. (line search)	↓ Performance.
BFGS	• $2N_\theta$ (Gradients) • N_θ^2 (Hessian) • $2N_{RAM}$ (Adjoint)	• 2 forwards (Adjoint) • $\sim 3\mathcal{J}_T$ eval. (line search)	+ Low cost, + Hessian approximation, · Medium performance.
Gauss-Newton	• N_θ (Gradient) • N_θ^2 (Gramian) • $N_{RAM} \times N_\theta$ (Sensitivity) • $2N_{RAM}$ (Adjoint)	• N_θ forwards (Once) • 2 forwards (Adjoint) • N_θ^2 matrix inversion	+ Good convergence, – N_θ^2 inversion, – Need for sensitivity.

Table 3.1 – Analysis of costs for the different implemented variational methods.

Method	**Memory Cost**	**Step cost**	**Conclusion**
ROKF	• N_{RAM} (Observer) • N_θ (Observer) • $N_{RAM} \times N_\theta$ (Sensitivity) • N_θ^2 (Covariance)	• 1 forward step (Observer) • N_θ forward steps (Sensitivity) • N_θ^2 matrix inversion	+ Good convergence, + Hessian approximation, + Parallelization, – N_θ^2 inversion.
ROUKF (simplex)	• N_{RAM} (Observer) • N_θ (Observer) • $N_{RAM} \times N_\theta$ (Particles) • N_θ^2 (Covariance)	• $N_\theta + 1$ forward steps (Particles) • N_θ^2 matrix inversion • N_θ^2 Cholesky decomposition	+ Good convergence, + Hessian approximation, + Parallelization, + Tangent-free, – N_θ^2 inversion.

Table 3.2 – Analysis of costs for the different implemented sequential methods.

inverse covariance matrix U must be done and for the ROUKF this covariance must be decomposed using Cholesky decomposition. In both cases, $N_\theta + 1$ wave propagation problems must be forward step by step, an operation that can be done in parallel. The storage of $N_\theta \times N_\theta$ and associated operations may result in high computational costs, in such cases the Conjugate Gradient and L-BFGS can be used to mitigate this issue. We will consider for our cases that this storage and operations will not be prohibitive.

Finally, we compute the total computational cost for this illustration: the BFGS computed 2600 forward problems and 3900 functional evaluations sequentially to achieve 10^{-8} relative squared error; the ROUKF computed $N_\theta + 1$ forward problems in parallel, 712 $N_\theta \times N_\theta$ matrix inversions and Cholesky decompositions. The efficiency of each method will depend on the cost of the forward problem and the size of the parametric space N_θ. For an expensive forward problem and a reduced parametric space, the ROUKF presents the best efficiency.

Although we ran these implementations for a linear case, some remarks can be made for the nonlinear case. In the linear case, the sensitivity used for the variational methods only had to be computed once, but in the nonlinear case, it would have to be done at each iteration. This is less critical in sequential methods as the sensitivity is computed and corrected during one estimation. In the linear case, not much difference is expected between the ROKF and ROUKF versions. In the nonlinear case, the Extended version (ROEKF) would have to be used instead of ROKF with the need for the tangent dynamics, while the Unscented version (ROUKF) only needs the forward operator. All presented cases, except the ROUKF, require the tangent dynamics for a nonlinear problem. A potential computational burden for the sequential method in a nonlinear case is to recompute the forward operator at each time step, as the parameter changes.

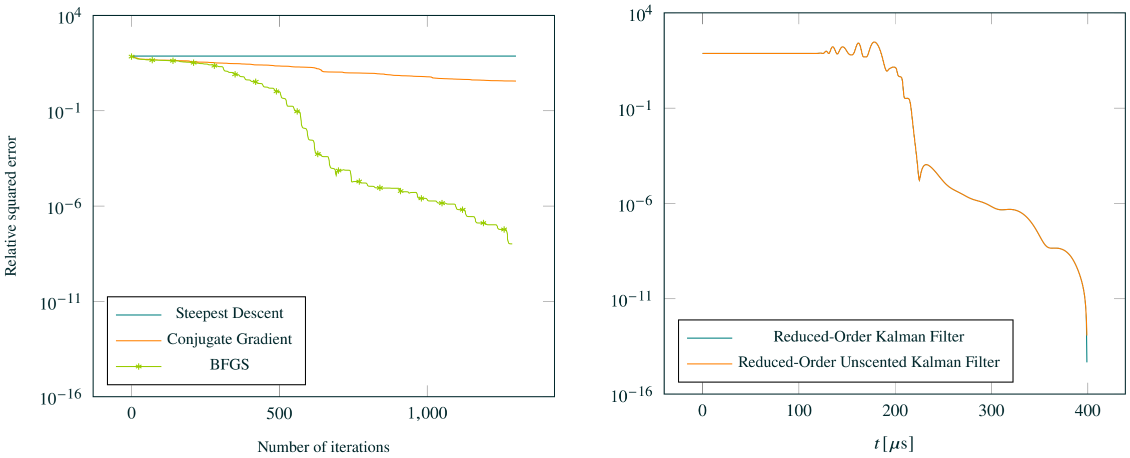


Figure 3.3 – Convergence of the minimization methods using the optimal solution as the reference. Variational Methods (left), Sequential Methods (right).

3.4.2 Nonlinear problem: A detailed sensitivity approach and inverse problem analysis

In the previous section, we implement variational and sequential methods for a linear-quadratic problem. In this section we will apply the sensitivity approach to a nonlinear problem, giving a better insight into how the discrete form of the gradient is computed component by component. Also, with a spatial dependent unknown that affects the wave speed, we reproduce and analyze the phenomenon of cycle skipping in wave propagation-related inverse problems. Let θ be the unknown temperature of a one-dimensional material occupying the domain $\Omega =]0; L[$, where $L > 0$ is the given total length of the domain. The temperature varies in space and, as previously mentioned, Θ is spaned by $\{\psi_i^{\Theta}\}_{i=1}^{N_\theta}$. We assume that $u_\theta(t) \in \mathrm{H}^1(\Omega)$ is the solution to the following wave propagation problem with free-surface boundary conditions and zero initial conditions, where f is a smooth source term to be defined,

$$\begin{cases} \partial_{tt}^2 u_\theta(x,t) - \partial_x\Big(c^2(\theta)\partial_x u_\theta(x,t)\Big) = f(x,t), & \text{in } \Omega\times]0;T[, \\ \partial_x u_\theta(0,t) = \partial_x u_\theta(L,t) = 0, \\ u_\theta(x,0) = \partial_t u_\theta(x,0) = 0 \end{cases} \tag{3.81}$$

The wave velocity depends on the temperature and we assume that the relation temperature-velocity is given by

$$c\big(\theta(x)\big) = k_s\theta(x) + c_0, \quad \forall x \in \Omega, \tag{3.82}$$

where $(k_s, c_0) \in \mathbb{R}^2$ are material-dependent constants. Considering $\omega \subset \Omega$ the “observation“ subdomain and the observation operator $C_u \in \mathcal{L}(\mathrm{H}^1(\Omega), \mathrm{H}^1(\omega))$ as

$$\begin{aligned} C_u \ : \ \mathrm{H}^1(\Omega) \ &\longmapsto \ \mathrm{H}^1(\omega) \\ u \ &\longrightarrow \ u\Big|_{\omega}, \end{aligned}$$

as in the previous illustration, C_u extracts the displacement field u in ω. We suppose that we have observations $y \in \mathcal{Y}$ from a target system

$$y(t) = C_u u^*(t) = u^*\Big|_{\omega}(t), \tag{3.83}$$

where u^* is the displacement field generated in the target system with θ^* as temperature field. Time and space dependencies were omitted. We will interpret the inverse problem as the following minimization problem

$$\overline{\theta} = \underset{\theta\in\Theta}{\operatorname{argmin}}\ \mathcal{J}_T(\theta) = \underset{\theta\in\Theta}{\operatorname{argmin}}\ \left\{\frac{1}{2}\int_0^T \|y(t) - C_u u_\theta(t)\|^2_{\mathrm{H}^1(\omega)}\,\mathrm{d}t\right\}, \tag{3.84}$$

where $\mathfrak{J}_T(\theta)$ is an L^2-based, in time, discrepancy measure or "cost function". Here, we consider a case without regularization and no disturbances in the observations, so we expect to find $\overline{\theta} = \theta^*$. In addition to the observations (3.83), we assume to have a first guess $\theta_0 \in \Theta$ on the sought temperature. In the following, we apply the method presented in Section 3.2.2.1 to compute the gradient of the functional $\mathfrak{J}_T$ and then perform the gradient descent to iteratively estimate $\overline{\theta}$. For a better illustration of such implementation, we present the method by making explicit the derivatives *w.r.t.* to the parameter components.

Steepest descent. First, we develop for this illustration case the steepest descent with a fixed time step using the gradient computed by the sensitivity approach. As seen in Section 3.2, the functional gradient can be used to perform the steepest descent and update iteratively an estimation of the parameter, where each parameter component is updated as

$$\theta_i^{k+1} = \theta_i^k - \alpha^k \frac{\partial \mathfrak{J}_T}{\partial \theta_i}(\theta^k)$$

with a descent step α^k. All components are updated at once.

Functional gradient using sensitivity Based in the method presented in Section 3.2.2.1, the computation of the gradient requires the sensitivity $s_\theta(t) \in \mathcal{L}(\Theta, \mathcal{Z})$

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} s_\theta(t) = A(\theta) s_\theta(t) + \mathrm{D}A(\theta) z_\theta(t), & \forall t \in [0, T], \\ \\ s_\theta(0) = 0, \end{cases}$$

that is equivalent, for this case, to (3.48). The functional gradient applied to an increment is written as

$$\nabla \mathfrak{J}_T(\theta)\widetilde{\theta} = \int_0^T \Big(C s_\theta(t)\widetilde{\theta}, C z_\theta(t) - y(t) \Big)_{\mathrm{H}^1(\omega)} \,\mathrm{d}t, \quad \forall \widetilde{\theta} \in \Theta.$$

In the following, we use the second-order form of the wave equation to express the functional gradient computations. The functional is written as

$$\begin{aligned} \mathfrak{J}_T(\theta) &= \frac{1}{2} \int_0^T \|y - C_u u\|^2_{\mathrm{H}^1(\omega)} \,\mathrm{d}t \\ &= \frac{1}{2} \int_0^T (y - C_u u, y - C_u u)_{\mathrm{H}^1(\omega)} \,\mathrm{d}t, \end{aligned} \tag{3.85}$$

We obtain the functional gradient by differentiating $\mathfrak{J}_T$ *w.r.t.* each component of θ, namely θ_i, $\forall i = 1, \ldots, N^\Theta$, yielding

$$\begin{aligned} \frac{\partial \mathfrak{J}_T}{\partial \theta_i}(\theta) &= \frac{1}{2} \int_0^T \frac{\partial}{\partial \theta_i} (y - C_u u, y - C_u u)_{\mathrm{H}^1(\omega)} \,\mathrm{d}t \\ &= -\int_0^T (C_u \frac{\partial u}{\partial \theta_i}, y - C_u u)_{\mathrm{H}^1(\omega)} \,\mathrm{d}t, \end{aligned}$$

leading to

$$\frac{\partial \mathfrak{J}_T}{\partial \theta_i}(\theta) = -\int_0^T (C_u s_i, y - C_u u)_{\mathrm{H}^1(\omega)} \,\mathrm{d}t, \tag{3.86}$$

where we denote the components of the sensitivity $s_i \in \mathcal{L}(\Theta, \mathrm{H}^1(\Omega))$, for brevity but with the risk of being mistaken with $s_\theta(t) \in \mathcal{L}(\Theta, \mathcal{Z})$,

$$s_i(t) = \frac{\partial u_\theta}{\partial \theta_i}(t) = \partial_{\theta_i} u_\theta(t), \quad \forall i \in [\![1; N^\Theta]\!],$$

introducing the short notation ∂_{θ_i} for partial differentiation. To obtain the dynamics for the sensitivity explicitly, we proceed by differentiating (3.81) *w.r.t.* the components of θ

$$\partial^2_{tt}\partial_{\theta_i}u_\theta - \partial_x\Big(\partial_{\theta_i}[c^2(\theta)]\partial_x u_\theta + c^2(\theta)\partial_x\partial_{\theta_i}u_\theta\Big) = \partial_{\theta_i} f.$$

By introducing the notation $\beta_i(\theta) = \partial_{\theta_i}[c^2(\theta)] = 2k_s\psi_i^\Theta c(\theta)$, with omitted x dependency, it yields

$$\partial^2_{tt}s_i - \partial_x\Big(\beta_i(\theta)\partial_x u + c^2(\theta)\partial_x s_i\Big) = 0.$$

To apply finite element procedures, we write the previous expression in its weak form and obtain the operators in their bilinear forms. By multiplying it by a test function $w \subset \mathrm{H}^1(\Omega)$ and integrating over the domain, we have

$$\partial^2_{tt}\int_\Omega s_i w\,\mathrm{d}\Omega - \int_\Omega \partial_x\Big(\beta_i(\theta)\partial_x u\Big)w\,\mathrm{d}\Omega - \int_\Omega \partial_x\Big(c^2(\theta)\partial_x s_i\Big)w\,\mathrm{d}\Omega = 0.$$

Using integration by parts and using the free boundary conditions, we get

$$\partial^2_{tt}\int_\Omega s_i w\,\mathrm{d}\Omega + \int_\Omega c^2(\theta)\partial_x s_i\partial_x w\,\mathrm{d}\Omega = -\int_\Omega \beta_i(\theta)\partial_x u\partial_x w\,\mathrm{d}\Omega.$$

The equation above written using bilinear forms is

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(s_i, w)_{\mathrm{L}^2(\Omega)} + k(c^2(\theta); s_i, w) = -k(\beta_i(\theta); u, w), \quad \forall i \in \llbracket 1; N^\Theta \rrbracket, \tag{3.87}$$

where, for any sufficiently regular function f,

$$k(f; v, w) = \int_\Omega f\partial_x v\,\partial_x w\,\mathrm{d}x, \quad \forall v, w \in \mathrm{H}^1(\Omega).$$

These propagation problems are completed with zero initial conditions. The sensitivity dynamics is no other than the displacement dynamics with a source term coming from the linearization of the k bilinear form. We deduce from (3.86) and (3.87) that using the sensitivity method to compute the functional gradient implies solving $N^\Theta + 1$ wave propagation problems.

Newton-Raphson descent method. As presented in Section 3.2, another method of descent is the Newton-Raphson method with an expected better order of convergence when compared with the Steepest Descent. With the object of using the descent in the form of (3.43), we compute the Hessian (3.42), for the presented case the second differentiation of the functional, the Hessian

$$\big(H^{(k)}\big)_{i,j} = \frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j}(\theta^k), \quad \forall i, j \in \llbracket 1; N^\Theta \rrbracket.$$

Obtaining the Hessian by cross-sensitivity problems. Differentiating twice $\mathfrak{J}_T$ *w.r.t.* θ for each component and using the previous results, we have

$$\frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j} = \int_0^T \frac{\partial}{\partial\theta_j}\Big(-(C_u s_i, y - C_u u)_{\mathrm{H}^1(\omega)}\Big)\,\mathrm{d}t =$$
$$\int_0^T (C_u s_i, C_u s_j)_{\mathrm{H}^1(\omega)} - (C_u h_{ij}, y - C_u u)_{\mathrm{H}^1(\omega)}\,\mathrm{d}t,$$

where we introduce the cross-sensitivity h as

$$h_{ij}(\theta, x, t) = \frac{\partial^2 u}{\partial\theta_i\partial\theta_j}(\theta, x, t), \quad \forall i, j \in \llbracket 1; N^\Theta \rrbracket.$$

Then, the Hessian is computed as

$$\begin{aligned}\left(H^{(k)}\right)_{i,j} = \left(\mathrm{D}^2\mathfrak{J}_T(\theta)\right)_{i,j} &= \frac{\partial^2 \mathfrak{J}_T}{\partial\theta_i \partial\theta_j} \\ &= \int_0^T (C_u s_i, C_u s_j)_{\mathrm{H}^1(\omega)} - (C_u h_{ij}, y - C_u u)_{\mathrm{H}^1(\omega)} \, \mathrm{d}t. \quad \forall i, j \in [\![1; N^\Theta]\!].\end{aligned}$$

Analog to the sensitivity case, we can obtain dynamics for the cross-sensitivity by differentiating (3.87)

$$\partial_{\theta_j}\left[\partial_{tt}^2 \int_\Omega s_i w \,\mathrm{d}\Omega + \int_\Omega c^2 \partial_x s_i \partial_x w \,\mathrm{d}\Omega\right] = -\partial_{\theta_j}\left[\int_\Omega \beta_i \partial_x u \partial_x w \,\mathrm{d}\Omega\right].$$

Permutating the derivatives and using the definition of h_{ij} we obtain

$$\partial_{tt}^2 \int_\Omega h_{ij} w \,\mathrm{d}\Omega + \int_\Omega \partial_{\theta_j}\left(c^2 \partial_x s_i \partial_x w\right) \mathrm{d}\Omega = -\int_\Omega \partial_{\theta_j}\left(\beta_i \partial_x u \partial_x w\right) \mathrm{d}\Omega.$$

Applying the product-rule and using the definition of h_{ij} and s_j, yields

$$\partial_{tt}^2 \int_\Omega h_{ij} w \,\mathrm{d}\Omega + \int_\Omega \left(\partial_{\theta_j} c^2 \partial_x s_i + c^2 \partial_x h_{ij}\right) \partial_x w \,\mathrm{d}\Omega = -\int_\Omega \left(\partial_{\theta_j} \beta_i \partial_x u + \beta_i \partial_x s_j\right) \partial_x w \,\mathrm{d}\Omega.$$

Finally rearranging the right and left-hand side

$$\partial_{tt}^2 \int_\Omega h_{ij} w \,\mathrm{d}\Omega + \int_\Omega c^2 \partial_x h_{ij} \partial_x w \,\mathrm{d}\Omega = -\int_\Omega (\beta_j \partial_x s_i + \beta_i \partial_x s_j) \partial_x w \,\mathrm{d}\Omega - \int_\Omega \partial_{\theta_j} \beta_i \partial_x u \partial_x w \,\mathrm{d}\Omega,$$

we may write it using the defined bilinear forms

$$\partial_{tt}^2 (h_{ij}, w)_{\mathrm{L}^2(\Omega)} + k(c^2; h_{ij}, w) = -k(\beta_j; s_i, w) - k(\beta_i; s_j, w) - k(\gamma_{ij}; u, w) \quad \forall i, j \in [\![1; N^\Theta]\!], \tag{3.88}$$

introducing $\gamma_{ij} = 2k_s^2 \psi_i^\Theta \psi_j^\Theta$, with omitted x dependency. Note that to use the Newton-Raphson method, the following numbers of propagation problems need to be solved for each descent step: one for u; N^Θ for s_i; $\frac{1}{2}N^\Theta(N^\Theta + 1)$ for the (symmetric) Hessian, having a total of

$$\frac{N^\Theta(N^\Theta + 1)}{2} + N^\Theta + 1$$

dynamics to solve.

Gauss-Newton method. One of the quasi-Newton methods with straightforward implementation is the Gauss-Newton, (3.44). Here, this approximation is equivalent to

$$\frac{\partial^2 \mathfrak{J}_T}{\partial\theta_i \partial\theta_j} \approx \int_0^T (C_u s_i, C_u s_j)_{\mathrm{H}^1(\omega)} \, \mathrm{d}t,$$

eliminating the need to propagate the cross-sensitivity dynamics.

3.4.2.1 Discretized problem

Using a Galerkin method as described in Chapter 1, we can represent the wave equations from their weak formulations in a matrix formulation. We define the discrete finite element space

$$V_h(\Omega) = \{v_h \in C^0(\overline{\Omega}), \quad \forall K \in \mathcal{T}_h, \quad v_h\big|_K \in P^k(K)\},$$

representing an approximation of the functional space $\mathrm{H}^1(\Omega)$ where we look for the solution of the wave propagation problems. We define $\mathcal{T}_h$ as the one-dimensional mesh, related to the mesh step h, and P^k

the space of polynomials of order k. Hence, $\mathcal{T}_h(\Omega)$ and $\mathcal{T}_h(\omega)$, are the mesh for the whole domain and its restriction to the observation region, respectively. The approximation space is generated by the set of Lagrange basis functions

$$V_h(\Omega) = \operatorname{span}\{\psi_I\}_{I=1}^{N_h} \quad \text{and,}$$

where we associate $\{\psi_I\}_{I=1}^{N_h}$ to the interpolation points

$$\{\xi_I\}_{I=1}^{N_h} \subset \overline{\Omega} \quad \text{with} \quad \psi_I(\xi_J) = \delta_{IJ}, \quad \forall I, J \in [\![1; N_h]\!].$$

We then represent the bilinear operators used in (3.87) and (3.88) in the finite element space, as the matrices

$$\mathbb{M}_{IJ} = (\psi_I, \psi_J)_{\mathrm{L}^2(\Omega)}, \quad \mathbb{K}_{g,IJ} = k(g; \psi_I, \psi_J),$$

for any function g. This matrices can be used, in addition to a discretized second-order time scheme, to rewrite the wave propagation problems (3.81), (3.87) and (3.88) in their discrete forms

- **Model**

$$\mathbb{M}\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} + \mathbb{K}_{c^2}u^n = f^n, \tag{3.89}$$

- **Sensitivity**

$$\mathbb{M}\frac{s_i^{n+1} - 2s_i^n + s_i^{n-1}}{\Delta t^2} + \mathbb{K}_{c^2}s_i^n = -\mathbb{K}_{\beta_i}u^n, \quad \forall i \in [\![1; N^\Theta]\!], \tag{3.90}$$

- **Cross-sensitivity**

$$\mathbb{M}\frac{h_{ij}^{n+1} - 2h_{ij}^n + h_{ij}^{n-1}}{\Delta t^2} + \mathbb{K}_{c^2}h_{ij}^n = -\mathbb{K}_{\beta_j}s_i^n - \mathbb{K}_{\beta_i}s_j^n - \mathbb{K}_{\gamma_{ij}}u^n, \quad \forall i, j \in [\![1; N^\Theta]\!], \tag{3.91}$$

where $f^n = f(n\Delta t)$, with f being the discretized right-hand side in $V_h(\Omega)$ and the dependencies on θ are omitted. To ensure the stability of both time schemes, the time step must satisfy the CFL condition, as in (1.27),

$$\Delta t \leq \frac{2}{\sqrt{r(\mathbb{M}^{-1}\mathbb{K}_{c^2})}},$$

that should be computed at each descent iteration as $c^2(\theta)$ changes. Satisfying the CFL condition for these wave propagation problems enabled us to set a unique time step as the matrices $\mathbb{M}$ and $\mathbb{K}_{c^2}$ do not change during the descent iteration. But as the observations may be in a different time sampling and discrete operations are made between the model and the observations, a linear interpolation in time must done keeping the smallest time step among them. This choice of interpolation is crucial as the method will not converge otherwise. The observation operator is represented by the matrix $\mathbb{C}_u$ that, given a solution, returns its content in the ω region.

The functional in its discrete form is

$$\mathcal{J}_N(\theta) = \frac{\Delta t}{2}\sum_{n=1}^{N}(y^n - \mathbb{C}_u u^n)^\intercal(\mathbb{M}_\omega + \mathbb{K}_\omega)(y^n - \mathbb{C}_u u^n), \tag{3.92}$$

where $\mathbb{M}_\omega$ and $\mathbb{K}_\omega$ are the equivalent mass and stiffness matrices for $V_h(\omega)$. After setting up the numerical scheme described here, we first specify the material parameters, source, domain and discretization. We apply the Steepest Descent, Newton-Raphson and Gauss-Newton methods for reconstructing the temperature field in two cases: when the temperature field is locally perturbed and when the perturbation is uniform in the whole domain Ω.

Wave equation as a first-order ordinary differential equation. The discrete versions of the model (3.89), the sensitivity (3.90) and cross-sensitivity (3.91) can be written in the form of a first-order discrete equation as done in (3.15). For each of these propagators, only the right-hand side will assume different forms.

Illustration specifications. The configuration used to illustrate in this section is described here. The temperature-dependent wave speed c will be defined as in (3.82) with

$$k_s = -0.752 \cdot 10^{-3}\,\mathrm{mm} \cdot \mu\mathrm{s}^{-1} \quad \text{and} \quad c_0 = 3.103\,\mathrm{mm} \cdot \mu\mathrm{s}^{-1},$$

where k_s and c_0 are material dependent and we use the values for aluminum [Croxford et al., 2007]. For the following illustrations, we will consider a reference temperature of 25°. The domain is defined as $\Omega =]0, 3000[$mm. The source function $f(x,t)$ is decomposed in a time-dependent part multiplied by a spatial-dependent part. The time-dependent part is a gated cosine with frequency 50kHz and its spatial-dependency is a rectangular function with amplitude one over $x \in [700, 800]$. We solved numerically (3.81) for N_t time steps, enough to propagate the wave through the whole domain, precisely 500μs for the case of the local variation and 590μs for the case of the global variation. The domain was meshed while ensuring three elements per wavelength elements of order 5 were used for discretization, *i.e.* $k = 5$.

Local variations of the temperature field. Here, we run a case where the temperature field is perturbed locally. The locally varying temperature field will have its "perturbation" domain with the length of the same order as the wavelength, leading to wave reflections. The perturbation is modeled as a rectangular function starting at point x_p^0 and ending at x_p^1, more precisely $\psi_1^\Theta(x)$ is this rectangular function and the temperature is described with an offset T_0, the reference temperature,

$$\theta(x) = T_0 + \theta_1 \psi_1^\Theta(x), \quad \forall x \in \Omega,$$

with the amplitude as the unknown parameter. A constant is included to take into account a reference temperature $T_0 = 25°$. As the methods described previously try to minimize cost function $\mathcal{J}_T(\theta)$, it is interesting to know the behavior of $\mathcal{J}_T$ for the considered case and region of interest. Therefore, we do the numerical evaluation of the cost function over a range of θ with a local varying temperature field. The temperature field will be T_0 with a local temperature perturbation as defined above with:

$$\omega =]700, 1400[\quad N_\Theta = 1, \quad x_p^0 = 1475 \quad x_p^1 = 1525 \quad \text{and} \quad {\theta^*}_1 = 250°,$$

so the observation domain and the perturbation do not intersect. Using the sensitivity approach we compute $\mathcal{J}_T$ (Figure 3.4) for a range of θ that comprises θ^* and the operational temperatures of an aircraft [Castanie, Bouvet, and Ginot, 2020]. The computed $\mathcal{J}_T$ is convex in θ and has no local minima, satisfying good convergence for the descent method. The local disturbances in the wave speed imply reflections and the reflected amplitude is related to the perturbation amplitude, not changing the overall time-of-flight of the signal. A difference norm L^2 will then account for it effectively, explaining a convex cost function. This can be verified in the snapshots shown in Figure 3.5 for $\theta = 119°$ and $\theta = 257°$.

Global variations of the temperature field. For the case of globally varying temperature in Ω, we will consider that the perturbation is over the whole domain. The target constant temperature will be $\theta^* = 250$ over the whole domain, keeping the same other configuration parameters except the observation domain $\omega =]200, 1200[$. The cost function $\mathcal{J}_T$, as in the local perturbation case, is also evaluated over the range of θ (Figure 3.7 (left)). As the change in wave speed is over the whole domain, the differences between observations will be mostly in the time-of-flight and not in the amplitude, causing the L^2 norm to be a less efficient functional, presenting local minima. Depending on the chosen θ to start the gradient descent, the resulting estimation may not be the global minimum. This must be considered as there could be local minima inside the operational range, for instance the case illustrated here. The phenomenon that caused these local minima is called cycle-skipping [Virieux et al., 2017] and is illustrated in the snapshots (Figure 3.8). The cycle-skipping happens due to the similarity of the signal, and its correlation when dephased by one or multiple cycles. For instance, if we use a signal with fewer oscillations, we expect to have fewer local minima. To illustrate it, we replace the gated cosine in time and use the Ricker wavelet. Both functions are presented in Figure 3.6. With this new time-dependency, the cycle-skipping effect is totally removed, as seen in the recomputed functional in Figure 3.7 (right). It has smoothed out the two closest local minima.

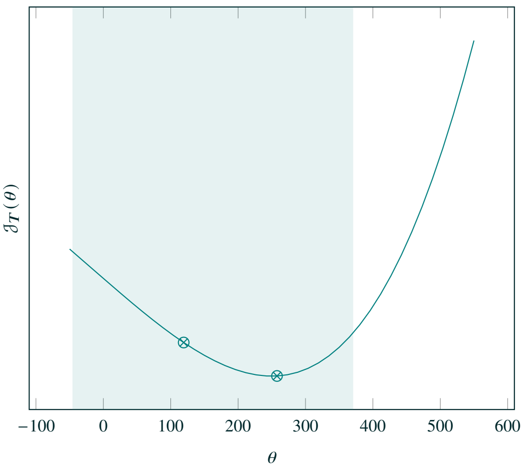


Figure 3.4 – Cost-function evaluated over a θ range. Snapshots were taken at the marked values (Figure 3.5). A supposed region of operation is highlighted.

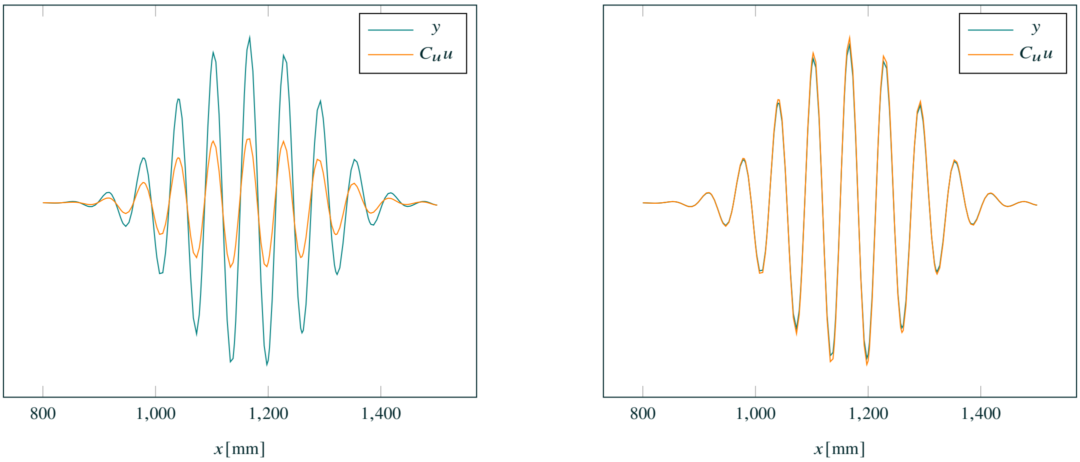


Figure 3.5 – Snapshots of the observed data and the observation operator applied to the model are shown for $\theta = 119°$ and $\theta = 257°$ at time 460 μs.

Gradient descent pseudo-algorithm. The pseudo-algorithm implementing the gradient computation is presented in Algorithm 2, where $c_k^2 = c^2(\theta_k)$ and $\beta_{i,k} = \beta_i(\theta_k)$. The observed data y, representing the measurements, is synthetically generated using the same model for wave propagation but using the target parameter θ^*. One may note that the loops of the wave propagation, sensitivity and cross-sensitivity can be shared and not necessarily done sequentially.

Testing different descent steps and first guesses. We illustrate here some of the main possible outcomes depending on the inversion parameters by presenting the evolution of the estimated parameter through iterations for different initial guesses θ_0 and fixed descent steps α. The stopping criteria for the iterations is

$$\frac{|\theta_k - \theta_{k-1}|}{|\theta_k|} \geq r \quad \text{or} \quad k \geq k_{max}.$$

This evolution is plotted in Figure 3.9 for local variations and Figure 3.10 for a global variation.

For both cases, a misadjusted descent step can lead to oscillations around the minimum and not converge at all, for instance in Figure 3.10 (right). For the local perturbation case, whose cost function is convex, there is only one minimum to which the descent method tends and, depending on the step size,

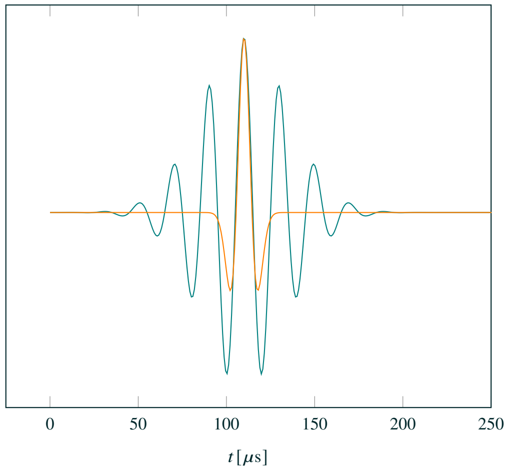


Figure 3.6 – Gated Cosine and Ricker type of signals, illustrating the difference in their oscillatory aspect.

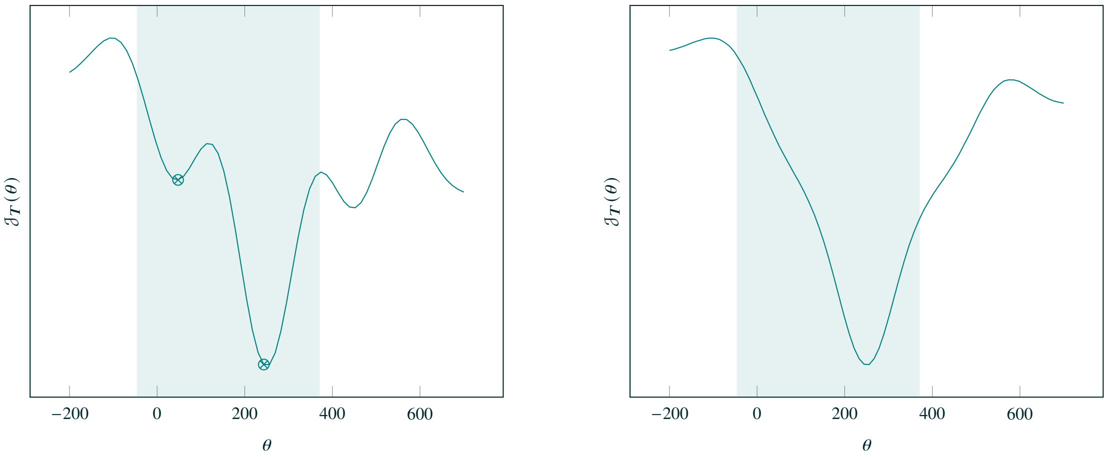


Figure 3.7 – Cost-function evaluated over a θ range for a gated-cosine (left) and a Ricker wavelet (right) time signature. Snapshots were taken at the marked values (Figure 3.8). A supposed region of operation is highlighted.

the convergence can be achieved in fewer iterations. The method can diverge if the step size is too big (for instance the $\alpha^k = 300$ case) or cannot attain convergence for small k_{max}. For the case of global perturbation, as its cost function is not convex, the descent method may converge to a local minimum. This wrong convergence can be triggered through slight changes in the initial guess around a (local) maximum. When using the Ricker wavelet the local minima were smoothed out and a good convergence is achieved for a wider range of θ as shown in Figure 3.11. The cycle skipping is associated with the choice of the L^2 norm for quantifying the misfit. Changing the norm to the Wasserstein metric [Mainini, 2012; Engquist, Ren, and Yang, 2020] is a form of mitigating the problem, but requires adjustments as it is fitted for nonnegative data.

Newton-Raphson method. As the Newton-Raphson method developed here looks for $\nabla\mathcal{J}_T = 0$, before implementing it, we plot the functional gradient with respect to the parameter to better understand and anticipate the behavior of the method. The functional gradient is plotted in Figure 3.12 for the local and global varying temperature cases. For the global case, the convergence to the global maximum will be only achieved for a narrower range of first guesses, when compared with the steepest descent method, between $\theta = 200$ and $\theta = 300$. The reason for this is that the derivative of the cost function crosses zero

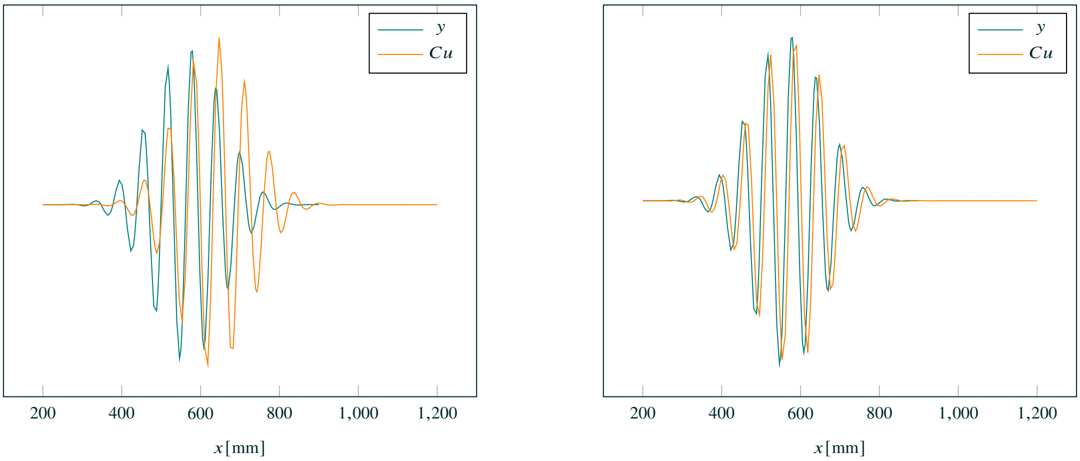


Figure 3.8 – Snapshots of the observed data and the observation operator applied to the model are shown for $\theta = 47°$ and $\theta = 243°$ at time 620 μs.

Algorithm 2: Implementation of the Gradient Descent method with fixed descent step

```
Input: The finite element spaces (model and observation), observed data $y$, total simulation time, the
       descent step $\alpha^k$, initial $\theta_0$, the residue tolerance $r$ and the maximum number of iterations $k_{max}$.
Output: The temperature components $\theta_k$.
assemble $\mathbb{M}$, $\mathbb{M}_\omega$, $\mathbb{K}_\omega$
$k \leftarrow 0$
while $\frac{|\theta_k - \theta_{k-1}|}{|\theta_k|} > r$ and $k < k_{max}$ do
    compute $\Delta t$, $N_t$
    $\nabla \mathfrak{J}_T \leftarrow 0$
    for $n \leftarrow 1$ to $N_t$ do
        $u^{n+1} \leftarrow \Delta t^2 \mathbb{M}^{-1}\left(f^n - \mathbb{K}_{c_k^2} u^n\right) + 2u^n - u^{n-1}$
    end
    for $i \leftarrow 1$ to $N_\theta$ do
        for $n \leftarrow 1$ to $N_t$ do
            $s_i^{n+1} \leftarrow -\Delta t^2 \mathbb{M}^{-1}\left(\mathbb{K}_{\beta_{i,k}} u^n + \mathbb{K}_{c_k^2} s_i^n\right) + 2s_i^n - s_i^{n-1}$
        end
    end
    for $i \leftarrow 1$ to $N_\theta$ do
        for $n \leftarrow 1$ to $N_t$ do
            $\frac{\partial \mathfrak{J}_T}{\partial \theta_i} \leftarrow \frac{\partial \mathfrak{J}_T}{\partial \theta_i} - \Delta t\left(\mathbb{C}_u s_i^n\right)^\intercal \left(\mathbb{M}_\omega + \mathbb{K}_\omega\right)\left(y^n - \mathbb{C}_u u^n\right)$
        end
    end
    $k \leftarrow k + 1$
    $\theta_k \leftarrow \theta_{k-1} - \alpha^k \nabla \mathfrak{J}_T$
end
```

at several different θ and Newton's method uses the tangent of the current θ_k to redirect the algorithm towards the zero. For instance, from $\theta_0 = 320$, instead of directing towards $\theta* = 250$, the algorithm would direct it towards the zero at $\theta = 350$. Whenever the algorithm falls near a local maximum, it diverges, as the tangent tends to be parallel to the horizontal axis, crossing the zero at distant θ. The pseudo-code of the implemented method is written in Algorithm 3.

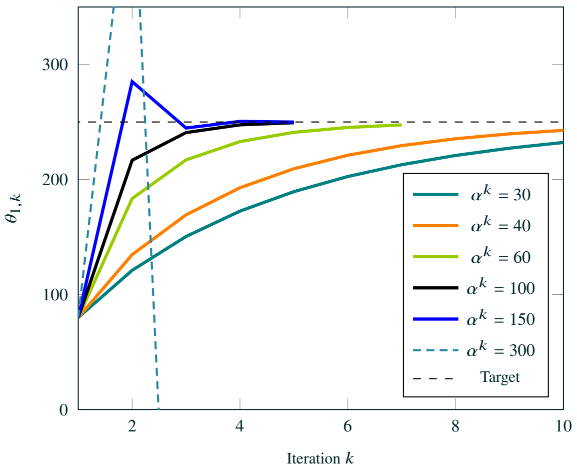


Figure 3.9 – Evolution of the descent method for local temperature variation using different step sizes.

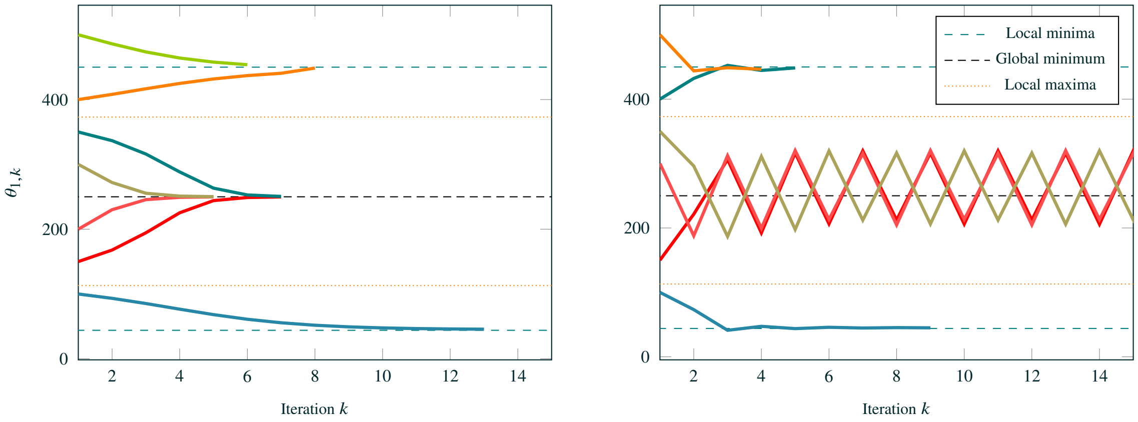


Figure 3.10 – Convergence of the descent method for different initial guesses. Results for global temperature variation with $\alpha^k = 10^{-3}$ (left) and $\alpha^k = 4 \times 10^{-3}$ (right).

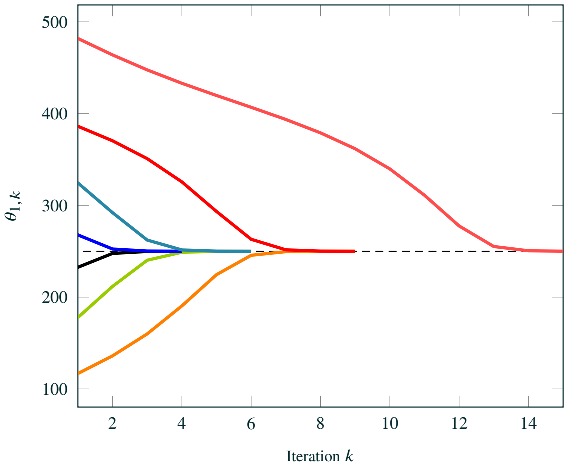


Figure 3.11 – Convergence of the descent method for different initial guesses using a Ricker wavelet as the source time signature. Results for global temperature variation with $\alpha^k = 4 \times 10^{-3}$.

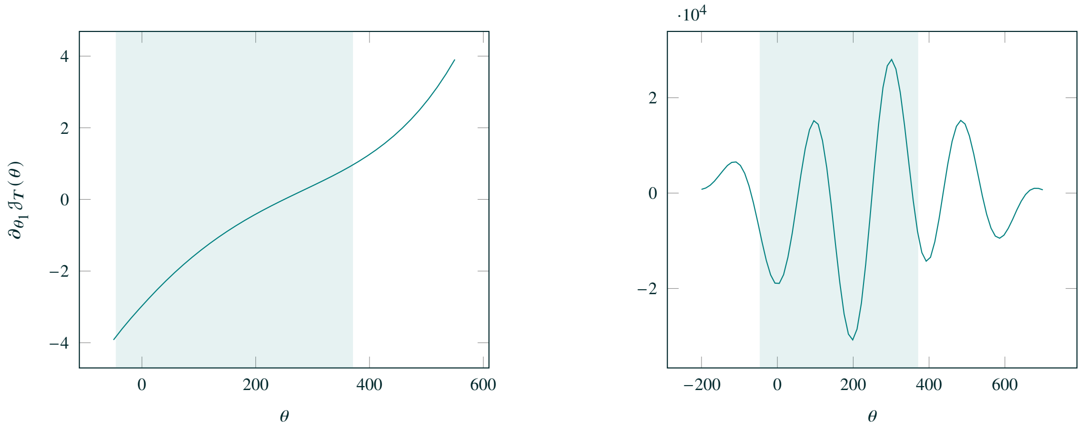


Figure 3.12 – Derivative of the cost function evaluated over a θ range for the local (left) and global (right) varying temperature cases.

Gauss-Newton method. In the implementation of the Gauss-Newton method, the *for* loop in line 17 can be removed and instead of using the Hessian matrix, we will use its approximation $\widetilde{H}$. Lines 33 and 39 change to

$$\widetilde{H}_{ij} \leftarrow \widetilde{H}_{ij} + \Delta t \left(\mathbb{C}_u s_i^n\right)^\mathsf{T} \mathbb{N}_\omega \left(\mathbb{C}_u s_j^n\right)$$

and

$$\theta_k \leftarrow \theta_{k-1} - \widetilde{H}_k^{-1} \nabla \mathcal{J}_T,$$

respectively. The iteration evolution for the Newton and Gauss-Newton methods is presented in Figure 3.13. For the local varying case (left) the gradient-descent method with $\alpha^k = 100$ is also plotted to illustrate the differences. The Newton and Gauss-Newton methods have better convergence, as expected, and respond very similarly. In the global varying case, we see clearer the limitations of the proposed Newton method as $\nabla \mathcal{J}_T$ crosses zeros for several θ. Although it converges faster, the region of convergence is narrower and the method is sensible to the initial guess.

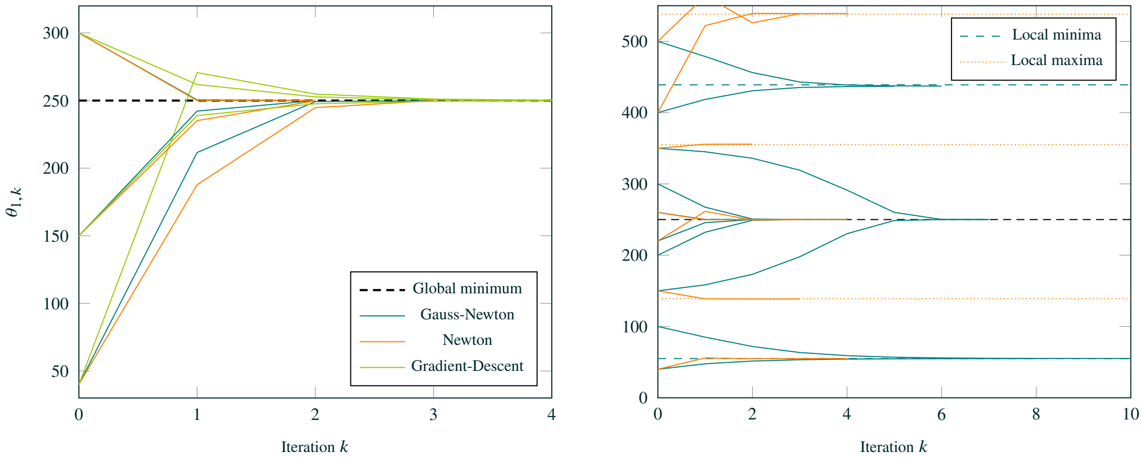


Figure 3.13 – Convergence of different methods starting from different initial guesses. Results for local (left) and global (right) varying temperature.

Discussion. We presented a nonlinear inverse problem to estimate the temperature field. First, we assessed the inverse problem when the perturbations to be reconstructed are localized in a region with dimensions close to the wavelength, leading to reflections of the wave. The functional was convex with

Algorithm 3: Implementation of the Newton-Raphson method with fixed descent step

```
Input: The finite element spaces (model and observation), observation data $y$, total simulation time, the
       descent step $\alpha^k$, initial $\theta_0$, the residue tolerance $r$ and the maximum number of iterations $k_{max}$.
Output: The temperature components $\theta_k$.
assemble $\mathbb{M}$, $\mathbb{M}_\omega$, $\mathbb{K}_\omega$
$\mathbb{N}_\omega \leftarrow \mathbb{M}_\omega + \mathbb{K}_\omega$
$k \leftarrow 0$
while $\frac{|\theta_k - \theta_{k-1}|}{|\theta_k|} > r$  and  $k < k_{max}$ do
    compute $\Delta t$ , $N_t$
    $\nabla\mathfrak{J}_T \leftarrow 0$
    for $n \leftarrow 1$ to $N_t$ do
        $u^{n+1} \leftarrow \Delta t^2 \mathbb{M}^{-1}\left(f^n - \mathbb{K}_{c_k^2} u^n\right) + 2u^n - u^{n-1}$;
    end
    for $i \leftarrow 1$ to $N_\theta$ do
        for $n \leftarrow 1$ to $N_t$ do
            $s_i^{n+1} \leftarrow -\Delta t^2 \mathbb{M}^{-1}\left(\mathbb{K}_{\beta_{i,k}} u^n + \mathbb{K}_{c_k^2} s_i^n\right) + 2s_i^n - s_i^{n-1}$
        end
    end
    for $i \leftarrow 1$ to $N_\theta$ do
        for $j \leftarrow 1$ to $i$ do
            for $n \leftarrow 1$ to $N_t$ do
                $h_{ij}^{n+1} \leftarrow -\Delta t^2 \mathbb{M}^{-1}\left(\mathbb{K}_{\beta_{i,k}} s_i^n + \mathbb{K}_{\beta_{i,k}} s_j^n + \mathbb{K}_{c_k^2} h_{ij}^n + \mathbb{K}_{\gamma_{ij,k}} u^n\right) + 2h_{ij}^n - h_{ij}^{n-1}$
            end
        end
    end
    for $n \leftarrow 1$ to $N_t$ do
        for $i \leftarrow 1$ to $N_\theta$ do
            $\frac{\partial\mathfrak{J}_T}{\partial\theta_i} \leftarrow \frac{\partial\mathfrak{J}_T}{\partial\theta_i} - \Delta t\left(\mathbb{C}_u s_i^n\right)^\intercal \mathbb{N}_\omega \left(y^n - \mathbb{C}_u u^n\right)$
        end
    end
    for $n \leftarrow 1$ to $N_t$ do
        for $i \leftarrow 1$ to $N_\theta$ do
            for $j \leftarrow 1$ to $i$ do
                $\frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j} \leftarrow \frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j} + \Delta t\left(\mathbb{C}_u s_i^n\right)^\intercal \mathbb{N}_\omega \left(\mathbb{C}_u s_j^n\right) - \Delta t\left(\mathbb{C}_u h_{ij}^n\right)^\intercal \mathbb{N}_\omega \left(y^n - \mathbb{C}_u u^n\right)$
            end
        end
    end
    complete symmetry of $\frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j}$
    $k \leftarrow k + 1$
    $\theta_k \leftarrow \theta_{k-1} - \left(\frac{\partial^2\mathfrak{J}_T}{\partial\theta_i\partial\theta_j}\right)^{-1} \nabla\mathfrak{J}_T$
end
```

the L^2 misfit, leading to good convergence for all presented methods. The choice of the descent step is crucial for good results. Secondly, we used the same methods to estimate changes in the temperature in the whole domain, leading to no reflections but changes in the wave speed. This causes the issue of cycle skipping, where the estimator falls in local minima. Changing the waveform helps but it may be an unavailable solution to this issue. Gradient descent methods were used with the gradient and Hessian was computed by sensitivity and cross-sensitivity, respectively. These computations require solving several wave problems per iteration, with a potentially high number of iterations depending on observability.

3.4.3 Nonlinear problem: Extended and Unscented Kalman Filters

In this section, we focus on the “Reduced-Order” version of the sequential methods to solve the inverse problem of reconstructing the temperature described in the previous section. Additionally, regularization will be added to the problem. For the sake of simplicity, we consider only local perturbations on the temperature with a single degree of freedom, $N_\theta = 1$, and

$$\psi_1^{\Theta}(x) = e^{-\left(\frac{x-x_0}{2\sigma}\right)^2}$$

with $x_0 = 200$ and $\sigma = 10$. The initial guess is set $\theta_0 = -30$ and the target $\theta^* = 250$. We implement the Reduced-Order Extended (ROEKF) and Unscented (ROUKF) Kalman Filters for the described inverse problem. The Extended version requires tangent dynamics while the Unscented version requires only the forward procedure. For both methods, we provide an initial covariance $\mathbb{R}_\Theta^{-1} = (80)^2$ with 80 being the estimated standard deviation slightly constraining the ROUKF. The simulation is run up to 600μs. Snapshots of the target system are shown in Figure 3.14.

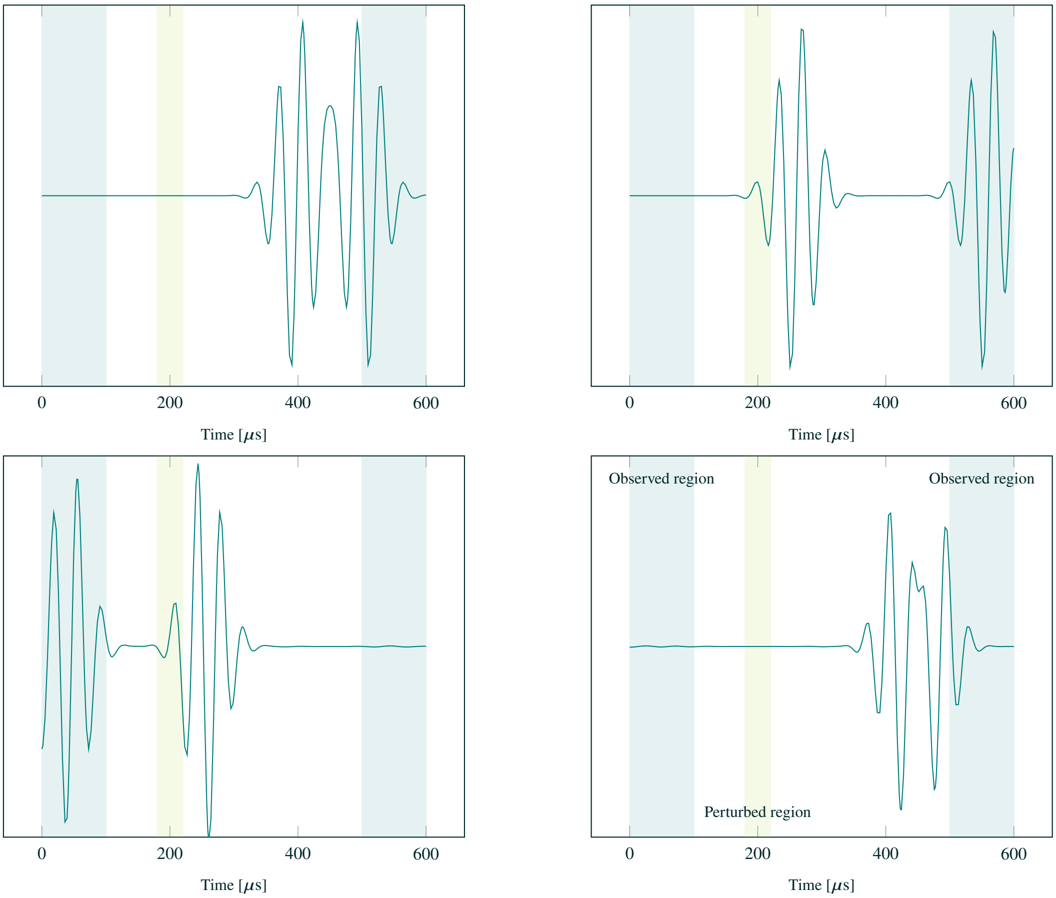


Figure 3.14 – Wavefield snapshots of the target system. Regions of observation and local temperature perturbations are highlighted. Snapshots taken at 60, 110, 220 and 440 μs.

Reduced-Order Extended Kalman Filter. To implement the ROEKF algorithm (3.73), we need the tangent dynamics represented by $\mathrm{D}\Phi_{n+1|n}(\widehat{\theta}^n_+)$. We recall the wave equation written in the first order,

$$\begin{cases} z^{n+1}_\theta = \Phi_1^{-1}\Phi_0(\theta) z^n_\theta + \Phi_1^{-1} b^n, & \forall n \in [\![0; N-1]\!], \\ z^0 = 0, \end{cases} \tag{3.93}$$

for

$$z^n_\theta = \begin{pmatrix} u^n_\theta \\ v^n_\theta \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} \mathbb{M} & -\Delta t \mathbb{M} \\ 0 & \mathbb{M} \end{pmatrix}, \quad \Phi_0(\theta) = \begin{pmatrix} \mathbb{M} & 0 \\ -\Delta t \mathbb{K}(\theta) & \mathbb{M} \end{pmatrix}, \quad b^n = \begin{pmatrix} 0 \\ \Delta t f^n \end{pmatrix}.$$

Then, for a given $\widetilde{\theta} \in \Theta$, the tangent operation reads

$$\left(\mathrm{D}\Phi_{n+1|n}(\theta)\widetilde{\theta}\right) z^n_\theta = \Phi_1^{-1}\left(\Phi_0(\theta) + \begin{pmatrix} 0 & 0 \\ -\Delta t \mathrm{D}\mathbb{K}(\theta)\widetilde{\theta} & 0 \end{pmatrix}\right) z^n_\theta.$$

Reduced-Order Unscented Kalman Filter. For the Unscented filter, no tangent is needed and the algorithm is implemented as in (3.76).

Discussion. The use of the reduced-order version is motivated by the fact that the uncertainty is not present in the state of the system but only in the parameter θ. Contrary to the variational methods, the parameter changes during the estimation from one time step to another, requiring an updated operator. This potentially means to compute the stiffness matrix at each time step, which adds excessive costs. Especially in our case, as presented in Chapter 1, we use a non-assembled application of the stiffness matrix, overcoming this issue as the change of the parameter does not change the forward procedure. The dynamic change of parameters can also affect numerical stability. As the time step Δt and mesh step h are computed during the initialization, it may not satisfy the stability conditions for the updated parameter during an estimation. To overcome this issue, we set a safety factor for Δt and h considering the range in which the parameter θ is expected to operate. Additionally to the safety factor, the covariance of the Unscented filter must be constrained so the parameter at each particle does not go beyond the range of operation, otherwise, the estimation can be unstable. Due to a tendency of monotone convergence for this illustration, this was not a particular issue. In more complex cases such as the ones presented in the next chapter, we ensure the stability of the estimation by constraining the initial covariance.

The evolution of the estimation for both methods is presented in Figure 3.15 with a region highlighted corresponding to a standard deviation for the estimated component, namely $\sqrt{(\mathbb{U}^{-1})_{11}}$. The ROUKF presented similar results to the ROEKF, while not requiring any tangent operator.

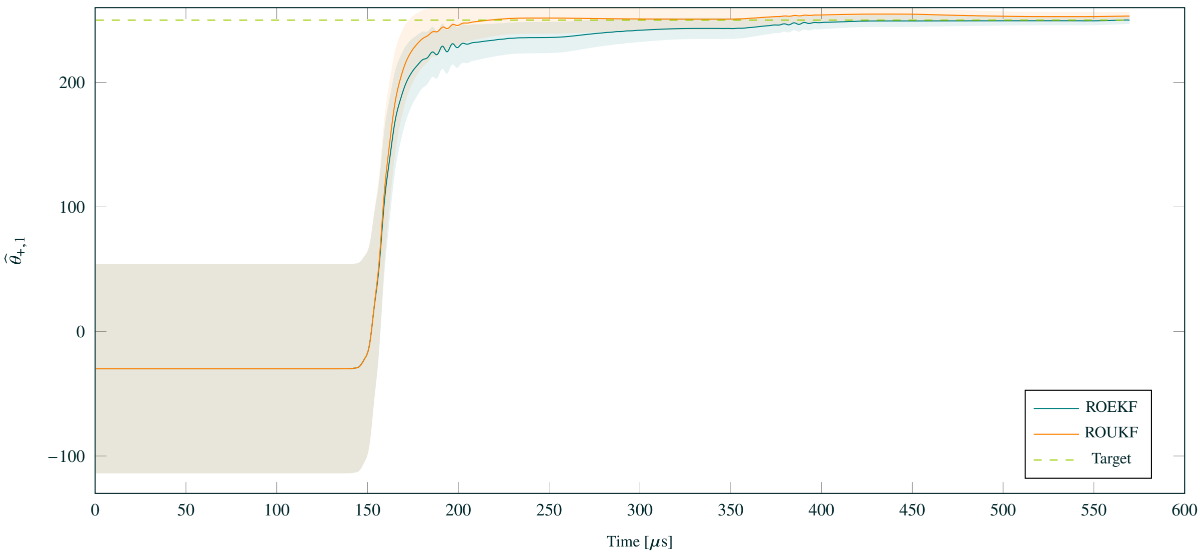


Figure 3.15 – Evolution of the ROEKF and ROUKF observers. Both present similar estimation performance, while the ROUKF does not require the tangent dynamics. The highlighted region represents the ~68% standard deviation.

3.5 Conclusion

In this chapter, we introduced a general framework for inverse problems associated with wave propagation and presented some traditional variational and sequential methods. The methods are presented in their time-continuous and time-discrete form. Finally, we implement those methods in linear and nonlinear one-dimensional wave propagation problems. The issue of cycle skipping is illustrated. The methods were implemented and compared in their original form, without specificities or acceleration. For the illustration of a linear-quadratic problem, the variational methods presented a higher cost to achieve a given convergence compared to the sequential methods. Sequential methods were robust and optimal however performing considerably less forward problems. For the nonlinear case, we first use the sensitivity approach for illustration purposes. Then, we apply the ROEKF and ROUKF, where they had similar performances while the ROUKF does not require a tangent model. As our objective is to use the direct model presented in Chapter 1, using tangent dynamics would require the unpractical third differentiation of the hyperelastic potential. Hence, tangent-free methods are preferred. Considering the presented results and being able to reduce the cost of changing the parameter at each time step by using a non-assembled stiffness matrix, the ROUKF seems appropriate for the task of estimating the structural deformation. One may note that a good convergence is not ensured for one run of the Kalman-Filter in the nonlinear case, which leads to the iterative strategy proposed in the next chapter.

Chapter 4

Kalman-based estimation of loading conditions from ultrasonic guided wave measurements

This chapter aims to present our data assimilation strategy to estimate structural deformation using guided ultrasonic measurements. The estimation problem is interpreted as a least squares minimization problem, which is then solved by an original combination of variational and sequential approaches. Our resulting strategy is an iterative version of the Unscented Kalman Filter, which has a link with the Levenberg-Marquardt algorithm. The direct problem presented previously is used as the forward model in the inversion loop. The effectiveness of our algorithm is shown by implementing it to model realistic industrial SHM configurations. The chapter takes the form of a pre-print article co-authored by André Dalmora, Alexandre Imperiale, Sébastien Imperiale and Philippe Moireau.

Contents

Kalman-based estimation of loading conditions from ultrasonic guided wave measurements

Andre Dalmora[1,2,3], *Alexandre Imperiale*[1], *Sébastien Imperiale*[2,3], *Philippe Moireau*[2,3]

1 - CEA Saclay, LIST - DIN, Université Paris Saclay, 91191 Gif-sur-Yvette, France
2 - Inria, Team-MΞDISIM, Inria-Saclay Ile de France 91120 Palaiseau, France
3 - LMS, Ecole Polytechnique, CNRS – Institut Polytechnique de Paris

Abstract. Ultrasonic guided wave-based Structural Health Monitoring (SHM) of structures can be perturbed by Environmental and Operations Conditions (EOCs) that alter wave propagation. In this work, we present an estimation procedure to reconstruct an EOC-free baseline of the structure suitable for SHM from the only available Ultrasonic guided wave measurements. Our approach is model-based, *i.e.* we use a precise modeling of the wave propagation altered by structure loading conditions. This model is coupled with the acquired data through a data assimilation procedure to estimate the deformation caused by the unknown loading conditions. From a methodological point of view, our approach is original since we have proposed an iterated Reduced-Order Unscented Kalman strategy, which we justify as an alternative to a Levenberg-Marquardt strategy for minimizing the non quadratic least-squares estimation criteria. Therefore, from a data assimilation perspective, we provide a quasi-sequential strategy that can valuably replace more classical variational approaches. Indeed, our resulting algorithm proves to be computationally very effective, allowing us to successfully apply our strategy to realistic 3D industrial SHM configurations.

4.1 Introduction

In various cutting-edge industrial fields, *e.g.* nuclear power generation, transportation, or aeronautics, the safe and reliable use of critical parts of structures is of paramount importance. To meet safety regulations in these areas, one must often be able to assess the integrity of the materials or equipment that make up these critical parts. To this end, numerous Non-Destructive Techniques (NDT) have been developed over the years. They are means of examining the material in question and obtaining quantitative information about its integrity without damaging it. Among them, Structural Health Monitoring (SHM) is an approach that – compared to other NDT techniques – generates a continuous stream of field data by incorporating actuators and sensors *in situ*. In other words, SHM systems monitor structures as it is used. One way to implement such systems is to rely on ultrasonic Guided Waves (GWs) [Mitra and Gopalakrishnan, 2016; Ricci et al., 2022] because of their attractive properties, such as propagation over long distances to study a large volume of material, or sensitivity to local thickness variations due to dispersion phenomena. Nevertheless, there are a number of challenges in the actual implementation of GW-based SHM systems in realistic structures. To name just two of them that drive the goals of our work, let us mention that the monitoring system should be minimal to avoid overloading the structure of interest, and the Environmental and Operations Conditions (EOCs) have a non-negligible impact on GW propagation [Gorgin, Luo, and Wu, 2020]. Typical EOCs that can affect wave propagation include temperature variations and the mechanical loading sustained by the structure during its use. We focus on the latter because it is more general in terms of mathematical formulation and can indeed lead to a change in wave velocities or even induce anisotropy. These effects are often referred to by the term “acoustoelastic” propagation [Shams, Destrade, and Ogden, 2011; Abiza, Destrade, and Ogden, 2012]. These acoustoelastic effects can alter the data registered by ultrasonic sensors, potentially affecting the precision of GW-based monitoring systems. In this context, the main objective of our work is the following: using the available data – the ultrasonic measurements – we aim to remove the bias caused by the mechanical loading conditions in order to reconstruct an EOC-free baseline. This objective solves the above two problems since the influence of mechanical loading is captured without additional sensors. This EOC-free baseline can then be used to find evidence of potential defects or damage within the structure in the ultrasonic signals – assuming here that the contributions of the defects and the EOCs are separable in the time or frequency domain.

In essence, our goal is to reconstruct a pre-deformation of the structure using only the GW measurements. From the wave propagation point of view, this is an inverse problem that we solve as a minimization problem of a fidelity-to-data functional under the condition that wave propagation dynamics is satisfied. An important feature arising from our EOC context is that the propagation model is in fact the one obtained by linearizing the nonlinear elastodynamics model around the sought pre-deformation, as presented in details in previous works [Dalmora et al., 2022] and recalled later in this article.

Solving such a non-linear optimization problem can be carried out through different approaches. A first one is to resort to gradient descent iterations or quasi-Newton processes [Luenberger and Ye, 2008; Courtier, Thépaut, and Hollingsworth, 1994]. At each iteration, the gradient of the cost function can be obtained by solving the so-called forward and backward adjoint problems. This approach is arguably the most common one to address this type of problem, it is referred to as the variational approach (4D-Var) [Dimet and Talagrand, 1986; Blum, Le Dimet, and Navon, 2009] in the data assimilation community with now extension to mechanical systems [Haik, Maday, and Chamoin, 2023], while in the geophysics community, it has been labeled as Full Waveform Inversion (FWI) [Virieux et al., 2017] with adaptation to tomography [Bernard et al., 2017]. A strong advantage of FWI is its robustness *w.r.t.* the size of the parametric space, *i.e.* the space in which lies the (discrete) solution of the minimization problem. In fact, numerous successful applications of this method have led to the reconstruction of wave velocity maps over large propagation domains. However, one significant difficulty of this approach is managing the adjoint dynamics, which contains the tangent of the propagation model around the forward trajectory. In the context of our work, this tangent model is intricate. In particular, it entails the third derivative of the hyperelastic potential ruling the constitutive behavior of the material. Also, storing the forward trajectory to evaluate this tangent is prohibitive, since the state space is very large in the context of high-frequency time-domain wave propagation. Note that however, in the case of inviscid wave propagation, one can save storage space by simply back-propagating the forward trajectory – a technique exploited in other wave propagation inverse problems [Ramdani, Tucsnak, and Weiss, 2010; De Buhan and Kray, 2013]. However, this is done at the cost of yet another call to the wave propagation solver.

A second approach is to resort to sequential methods and, in particular, to Kalman filtering approaches that can be developed in a stochastic or a deterministic context [Bensoussan, 1971; Simon, 2006]. The term sequential stems from the fact that the main building blocks of this approach are (exclusively) forward problems, where the dynamics are modified by the addition of a feedback loop. This feedback loop is proportional to the discrepancy between the synthetic data generated from the model's current trajectory and the actual measurements. The resulting modified dynamics is often referred to as a sequential estimator (in the stochastic context) or an observer (in the deterministic context) of the target trajectory. In Kalman-based filtering, the gain operator in the feedback loop is computed from a covariance operator satisfying a Riccati equation [Aussal and Moireau, 2022]. For Linear-Quadratic (LQ) problems, *i.e.* linear state dynamics with linear parameter-state coupling and a quadratic misfit functional, the observer at the final time corresponds exactly to the solution of the minimization problem [Moireau, 2022]. Also, in the special case where the uncertainty is limited exclusively to the parameter space, this method leads to a specific application of the Reduced-Order Kalman Filter [Moireau, Chapelle, and Le Tallec, 2008; Moireau, 2022]. This method can be generalized to cases with non-linear parameter-state coupling, leading to the so-called Reduced Order Extended Kalman Filter (EKF) or its gradient-free version, the Reduced-Order Unscented Kalman Filter (UKF) [Moireau, Chapelle, and Le Tallec, 2008; Moireau and Chapelle, 2011]. One of the main advantages of these sequential estimators lies in their ability to provide a solution to the minimization problem in one pass, consisting of embarrassingly parallel forward problems. Moreover, thanks to their tangent-free alternative, *e.g.* UKF, they are easy to interface with legacy code used in parallel as a black box propagating each UKF *sigma-point*, also referred to as *particle*. Nevertheless, the usual computational bottleneck of these sequential methods is to store and invert the (dense) covariance matrix whose dimension corresponds to the size of the parameter space. This limits the application of Kalman filters to relatively small parameter spaces compared to the typical configurations handled by variational or FWI methods. Moreover, although Kalman filters are exactly equivalent to the minimization problem LQ problems their extension to nonlinear minimization problems is either prohibitively expensive [Moireau, 2019; Moireau, 2022] or approximate when relying on Extended or Unscented Kalman filters [Moireau, Chapelle, and Le Tallec, 2008; Moireau and Chapelle, 2011; Moireau, 2022].

In our work, we develop a new estimation procedure that combines elements of the two approaches. Namely, we first apply a Levenberg-Marquardt (LM) algorithm [Hanke, 1997; Bal et al., 2013] to derive from the initial minimization problem a set of LQ subproblems satisfied by parameter increments. Each of these LQ subproblems is then solved using a Kalman filter approach, a sequential strategy already studied for wave or elasticity problems [Moireau, Chapelle, and Le Tallec, 2008; Marchand, Chamoin, and Rey, 2016; Furuya and Potthast, 2022]. We then revisit the UKF approach, to obtain a tangent-free algorithm with increasing convergence at each iteration of the outer loop of LM descent. Thus, we avoid the differentiation of the acoustoelastic wave propagation model and provide an estimation algorithm that can be easily interoperated with blackbox industrial codes. To be compatible with the dimensionality constraints of Kalman filtering, the parameter space is built from a modal decomposition of the pre-deformation, which in practice leads to the estimation of tens to hundreds of components on a modal basis. Moreover, the number of LM iterations to achieve convergence is rather small, leading overall to an almost sequential estimation approach. To illustrate the power of our approach, both in computation efficiency and parameter estimation, we provide extensive 3D results with synthetic noisy data in configurations associated with realistic SHM applications.

The structure of this article is as follows. In Section 4.2, we provide details on the direct problem, *i.e.* the time-domain acoustoelasticity propagation model. This leads us to the definition of the inverse problem we consider. In particular, we give a precise definition of the observation operator that generates the GW measurements and the modal basis that forms the parameter space. In Section 4.3, we develop our new method for identifying parameters using a combination of the LM algorithm and Kalman-based filtering. In a first step, we provide meaningful insights into this approach by considering continuous-time dynamics. In a second step, we give extensive details leading to the final fully discrete observer. Finally, in Section 4.4, we give relevant numerical illustrations of the estimation of the 3D pre-deformation with our approach and noisy synthetic data in realistic industrial configurations.

4.2 Problem setting

We here consider a deformable system classically modeled in continuum mechanics in a Lagrangian referential defined from an initial stress-free configuration. Defining by Ω the reference domain with Lipschitz boundaries and the material position $\mathbf{x}$ in the reference configuration, the dynamics principle reads

$$\begin{cases} \varrho_0 \partial^2_{tt} \boldsymbol{u}_{\text{tot}}(\mathbf{x},t) - \boldsymbol{\nabla} \cdot \boldsymbol{T}(\mathbf{x},t) = \varrho_0 \boldsymbol{f}_{\text{tot}}(\mathbf{x},t) & (\mathbf{x},t) \in \Omega \times (0,T), \\ \boldsymbol{u}_{\text{tot}}(\mathbf{x},t) = 0 & (\mathbf{x},t) \in \Gamma_D \times (0,T), \\ \boldsymbol{T} \cdot \boldsymbol{n}(\mathbf{x},t) = 0 & (\mathbf{x},t) \in \Gamma_N \times (0,T), \\ \boldsymbol{u}(\mathbf{x},0) = 0 & \mathbf{x} \in \Omega, \end{cases} \tag{4.1}$$

where, for the sake of simplicity, we consider homogeneous boundary conditions defined in the reference configuration. Moreover, we scale the volume loading with the volume mass for consistency. In (4.1) the first Piola-Kirchhoff stress tensor $\boldsymbol{T}$ is here defined from the derivative of an hyperelastic potential $\mathscr{W}$, with respect to the deformation gradient $\boldsymbol{F}(\mathbf{x},t) = \mathbf{Id} + \boldsymbol{\nabla} \boldsymbol{u}_{\text{tot}}(\mathbf{x},t)$, namely $\boldsymbol{T} = \mathrm{D}_{\boldsymbol{F}} \mathscr{W}(\mathbf{x},\boldsymbol{F})$.

As an alternative to the strong form, the system is defined using the weak form of the dynamics principle, also known as the principle of virtual work, by defining a space of admissible displacements, typically $\mathcal{K} \subset \mathcal{V} = \{\boldsymbol{w} \in \mathrm{H}^1(\Omega)^3 \mid \boldsymbol{w}|_{\Gamma_D} = 0\}$ such that

$$\forall \boldsymbol{w} \in \mathcal{V}, \quad \int_\Omega \varrho_0 \partial_{tt}^2 \boldsymbol{u}_{\text{tot}} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x} + \int_\Omega \boldsymbol{T} : \boldsymbol{\nabla} \boldsymbol{w} \, \mathrm{d}\mathbf{x} = \int_\Omega \varrho_0 \boldsymbol{f}_{\text{tot}} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x}. \tag{4.2}$$

In this general framework, we now consider that our system is inspected while being loaded by external unknown forces. These structural loading forces $\boldsymbol{f}_0$ are considered to be volume distributed, for the sake of simplicity, and quasi-static – namely not depending on time – albeit of possibly strong amplitude, as opposed to the ultrasonic excitation $\boldsymbol{f}$ which is of high-frequency and low amplitude. In fact, the quasi-static assumption is considered with respect to the time scale of ultrasonic excitation so the total loading decomposes into

$$\boldsymbol{f}_{\text{tot}}(\mathbf{x}) = \boldsymbol{f}_0(\mathbf{x}) + \delta \boldsymbol{f}(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \Omega \times (0, T), \tag{4.3}$$

where δ is a small parameter representing the fact that the amplitude of the ultrasonic excitation is small compared to the external loading. This allows us to separate scales and to consider that the resulting displacement is decomposed into

$$\boldsymbol{u}_{\text{tot}}(\mathbf{x}, t) = \boldsymbol{u}_0(\mathbf{x}) + \delta \boldsymbol{u}(\mathbf{x}, t) + O(\delta^2). \tag{4.4}$$

Injecting such anszatz into the principle of virtual work (4.2) and identifying the zero order terms *w.r.t.* δ, we formally showed in [Dalmora et al., 2022] that the displacement can be reconstructed from the solution $\boldsymbol{u}_0 \in \mathcal{K}$ of a large displacement static problem

$$\forall \boldsymbol{w} \in \mathcal{V}, \quad \int_\Omega \boldsymbol{T}(\mathbf{x}, \boldsymbol{F}_0) : \boldsymbol{\nabla} \boldsymbol{w} \, \mathrm{d}\mathbf{x} = \int_\Omega \varrho_0 \boldsymbol{f}_0 \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x}, \tag{4.5}$$

where $\boldsymbol{F}_0 = \mathbf{Id} + \boldsymbol{\nabla} \boldsymbol{u}_0$ is the deformation gradient associated with $\boldsymbol{u}_0$. Then, identifying the first order term in δ, we find that $\boldsymbol{u} \in \mathrm{L}^2((0,T); \mathcal{V})$ should be a solution of the wave-propagation problem given by

$$\forall \boldsymbol{w} \in \mathcal{V}, \quad \int_\Omega \varrho_0 \partial_{tt}^2 \boldsymbol{u} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x} + \int_\Omega \boldsymbol{\nabla} \boldsymbol{u} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \boldsymbol{F}_0) : \boldsymbol{\nabla} \boldsymbol{w} \, \mathrm{d}\mathbf{x} = \int_\Omega \varrho_0 \boldsymbol{f} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x}. \tag{4.6}$$

From a mathematical viewpoint, the existence of a solution for the general problem formulation (4.1) is still an open problem. However, under suitable conditions, we can consider that problem (4.5) and (4.6) admits one and only one solution. For (4.5), this is typically the case when defining a hyperelastic law using a polyconvex potential [Ciarlet, 1988b], giving the existence of a displacement $\boldsymbol{u}_0$ in the space of admissible displacements $\mathcal{K} \subset \mathcal{V}$.

Then, moving to (4.6), this formulation can be recast into a general second-order in time weak formulation. Denoting by $\mathcal{H}$ the space $\mathrm{L}^2(\Omega)^3$ equipped with the scalar product

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{H}^2, \quad (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} = \int_\Omega \varrho_0 \boldsymbol{u} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x},$$

Moreover, we equip $\mathcal{V}$ with the scalar product

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{V}^2, \quad (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{V}} = \int_\Omega \boldsymbol{\nabla} \boldsymbol{u} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \mathbf{Id}) : \boldsymbol{\nabla} \boldsymbol{w} \, \mathrm{d}\mathbf{x} + (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}}.$$

Note that $\mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \mathbf{Id})$ actually corresponds to the standard Hooke's law in linear elasticity. We can identify $\mathcal{H}$ with its dual so that we have the following Gelfand triple

$$\mathcal{V} \subset \mathcal{H} \equiv \mathcal{H}' \subset \mathcal{V}'.$$

Then, introducing the linear operator $A_0(\boldsymbol{u}_0) \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$ such that

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{V}^2, \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w}\rangle_{\mathcal{V}',\mathcal{V}} = \int_\Omega \boldsymbol{\nabla u} : \mathrm{D}^2_{\boldsymbol{F}}\mathscr{W}(\mathbf{x}, \boldsymbol{F}_0) : \boldsymbol{\nabla w}\, \mathrm{d}\mathbf{x},$$

the weak formulation (4.6) can be written in the following form

$$\forall \boldsymbol{w} \in \mathcal{V}, \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}(\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} + \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w}\rangle_{\mathcal{V}',\mathcal{V}} = (\boldsymbol{f}, \boldsymbol{w})_{\mathcal{H}}. \tag{4.7}$$

We now make the following assumption: there exists $\lambda \geq 0$ and $\alpha > 0$ such that

$$\forall \boldsymbol{w} \in \mathcal{V}, \langle A_0(\boldsymbol{u}_0)\boldsymbol{w}, \boldsymbol{w}\rangle_{\mathcal{V}',\mathcal{V}} + \lambda\|\boldsymbol{w}\|^2_{\mathcal{H}} \geq \alpha\|\boldsymbol{w}\|^2_{\mathcal{V}}. \tag{4.8}$$

Such assumption is satisfied for $\boldsymbol{u}_0$ small enough whereas for large displacement $\boldsymbol{u}_0$, it may be violated. This is however very dependent on the potential $\mathscr{W}$ that describes the elastic behavior of the medium. Some choices tend to make the assumption above more restrictive than others – see *e.g.* the examples in [Dalmora et al., 2022]. Thanks to this assumption – see for instance [Duvaut and Lions, 1976] – there exists one, and only one, *variational* solution of (4.7), namely $\boldsymbol{u} \in \mathcal{W}_T$ with

$$\mathcal{W}_T = \Big\{\boldsymbol{u} \in \mathrm{L}^2((0,T);\mathcal{V}), \partial_t\boldsymbol{u} \in \mathrm{L}^2((0,T);\mathcal{H}), \partial^2_{tt}\boldsymbol{u} \in \mathrm{L}^2((0,T);\mathcal{V}')\Big\}.$$

Moreover, we can extend the operator $A_0(\boldsymbol{u}_0) \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$ into an unbounded operator $(A_0(\boldsymbol{u}_0), \mathcal{D}(A_0))$ by defining

$$\mathcal{D}(A_0) = \Big\{\boldsymbol{u} \in \mathcal{V} \text{ such that } \exists \boldsymbol{r} \in \mathcal{H} : \forall \boldsymbol{w} \in \mathcal{V}, \quad \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w}\rangle_{\mathcal{V}',\mathcal{V}} = (\boldsymbol{r}, \boldsymbol{w})_{\mathcal{H}}\Big\}.$$

and

$$\forall \boldsymbol{u} \in \mathcal{D}(A_0), \quad \forall \boldsymbol{w} \in \mathcal{V}, \quad (A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} = \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w}\rangle_{\mathcal{V}',\mathcal{V}}.$$

Then, for a given $\boldsymbol{u}_0 \in \mathcal{K}$, by introducing the operator

$$A(\boldsymbol{u}_0) = \begin{pmatrix} 0 & \mathrm{Id} \\ -A_0(\boldsymbol{u}_0) & 0 \end{pmatrix} \text{ with } \mathcal{D}(A) = \mathcal{D}(A_0) \times \mathcal{V} \subset \mathcal{Z} := \mathcal{V} \times \mathcal{H},$$

we can rewrite the wave dynamics in the state-space form

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}z = A(\boldsymbol{u}_0)z + r, & \text{in } (0,T) \\ z(0) = 0, \end{cases} \quad \text{with } z = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}, \text{ and } r = \begin{pmatrix} 0 \\ \boldsymbol{f} \end{pmatrix}, \tag{4.9}$$

and $\boldsymbol{v}$ denotes the velocity unknown. In the sequel, we will denote $\dot{z} = \frac{\mathrm{d}}{\mathrm{d}t}z$. This state-space form allows to define solutions using semi-group theory [Bensoussan et al., 2007]. Namely for $r \in \mathrm{L}^2((0,T),\mathcal{Z})$ there exists one and only one mild solution of (4.9) in $\mathrm{C}^0((0,T),\mathcal{Z})$ which is also a variational solution in $\mathcal{W}_T$ of (4.7). We here underline that the variational formalism is necessary to further justify finite element discretization of (4.7), whereas the mild solution formulation in the sense of semi-group will simplify the presentation of the estimation problem.

In this modeling context, we consider that we have at our disposal some recorded measurements on a subregion ω_i – of the boundary $\partial\Omega$ – from d sensors of the wave propagation of a target wave solution $\check{\boldsymbol{u}}$. Typically we consider recordings $y = (y_i)_{1\leq i\leq d}$ given by

$$[0,T] \ni t \mapsto \check{y}_i(t) = \frac{1}{|\omega_i|}\int_{\omega_i} \check{\boldsymbol{u}}(t,\mathbf{x}) \cdot \boldsymbol{d}_i \,\mathrm{d}\mathbf{x} \in \mathbb{R}, \quad 1 \leq i \leq d,$$

up to certain measurement errors and where the fields $\boldsymbol{d}_i$ represent the sensitivity of the sensors to a displacement field. In fact, we do not have at our disposal $[0,T] \ni t \mapsto \check{y}(t) \in \mathbb{R}$ but rather the perturbed measurement $y_\gamma \in \mathrm{L}^2((0,T);\mathcal{Y})$ with $\mathcal{Y} = \mathbb{R}^d$ such that for a given noise level $\gamma > 0$,

$$\int_0^T \|\check{y}_i - y_{\gamma,i}\|^2_{\mathcal{Y}} \, \mathrm{d}t \lesssim \gamma^2 T.$$

From a state-space point of view, this allows to define an observation operator $C \in \mathcal{L}(\mathcal{Z},\mathcal{Y})$ by

$$C : z = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \mapsto \left(\frac{1}{|\omega_i|} \int_{\omega_i} \boldsymbol{u}(\mathbf{x}) \cdot \boldsymbol{d}_i \, \mathrm{d}\mathbf{x} \right)_{1 \leq i \leq d}. \tag{4.10}$$

Remark. *We would like to point out that in an alternative modeling approach, we could have considered that the measurements are recorded from the deformed configuration as the sensors are operating in this configuration. When rewriting the integrals with respect to the reference configuration, this should lead to the definition of an observation operator $C(\boldsymbol{u}_0)$ that depends on the deformation.*

We can now introduce the inverse problem that we want to solve. We assume that $\boldsymbol{u}_0$ is unknown and we want to reconstruct it from the available measurements. More precisely, we would like to specify the true $\check{\boldsymbol{u}}_0$ from an *a priori* displacement $\hat{\boldsymbol{u}}_0$ assuming that the true displacement is a regular perturbation with respect to the a priori, typically there exists a constant M such that

$$\|\check{\boldsymbol{u}}_0 - \hat{\boldsymbol{u}}_0\|^2_{\mathcal{V}} \leq M^2.$$

Moreover in practice, we propose to decompose any $\boldsymbol{u}_0 \in \mathcal{H}$ on the basis made of the eigenvectors $(\boldsymbol{\varphi}_j)_{j\geq 0}$ of the compact operator $\Lambda_0 = A_0(\mathbf{0})^{-1} \in \mathcal{L}(\mathcal{V}',\mathcal{V})$. We recall that there exists $(\mu_j)_{j\geq 0} \in (\mathbb{R}^*_+)^{\mathbb{N}}$ such that

$$A_0(\mathbf{0})\boldsymbol{\varphi}_j = \mu_j \boldsymbol{\varphi}_j \text{ with } \|\boldsymbol{\varphi}_j\|^2_{\mathcal{H}} = 1 \tag{4.11}$$

and $\lambda_j = \sqrt{\mu_j}^{-1}$ tends to 0 as j tends to infinity. Therefore, we can decompose

$$\boldsymbol{u}_0(\mathbf{x}) = \sum_{j\geq 0} \theta_j \boldsymbol{\varphi}_j(\mathbf{x}) \tag{4.12}$$

and we can enforce

$$\boldsymbol{u}_0 \in \mathcal{V} \Leftrightarrow \sum_{j\geq 0} \frac{\theta_j^2}{\lambda_j^2} < +\infty.$$

In practice, we may even want to consider that $\boldsymbol{u}_0$ belongs to a more regular space $\mathcal{V}^m$ leading to increased convergence rate of the sequence $(\theta_j)_{k\geq 0}$ to 0, typically

$$\boldsymbol{u}_0 \in \mathcal{V}^m \Rightarrow \sum_{j\geq 0} \frac{\theta_j^2}{\lambda_j^{2m}} < +\infty. \tag{4.13}$$

However, as a first step, we remain with our choice of estimating $\boldsymbol{u}_0 \in \mathcal{V}^m$ which is equivalent to estimating $\theta \in \mathcal{P} \subset \ell_2(\mathbb{R})$, equipped with a norm

$$\|\theta\|^2_{\mathcal{P}} = (\theta, \Lambda_0^{-m}\theta)_{\ell_2}, \text{ with } (\Lambda_0)_{ij} = \lambda_j^{2m}\delta_{ij} \text{ and } \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases} \tag{4.14}$$

Moreover, when we need to discretize $\boldsymbol{u}_0$, we will consider a finite-dimensional space corresponding to a finite number N_p of modes, hence $\theta \in \mathbb{R}^{\mathrm{N}_p}$ and Λ_0 becomes typically a diagonal matrix.

To conclude our problem setting, our objective is, therefore, to estimate $\check{\theta} \in \mathcal{P}$ from the available measurements, namely to invert the following operator

$$\Psi_T : \begin{cases} \mathcal{P} \to \mathrm{L}^2((0,T);\mathcal{Y}) \\ \theta \mapsto \left[t \mapsto \int_0^t C e^{A(\theta)(t-s)} r(s)\, \mathrm{d}s \right] \end{cases}$$

where we replace the $\boldsymbol{u}_0$ dependency by the θ dependency in the definition of the operator A_0. In the following, this inversion is based on a least-squares minimization using Levenberg-Marquardt strategy.

4.3 Identification method

4.3.1 From Levenberg-Marquardt minimization scheme to an iterated Extended Kalman strategy

Following [Bal et al., 2013], we approximate pseudo-inverse of Ψ_T with a Levenberg-Marquardt minimization scheme [Hanke, 1997] which consists of the following iterative procedure.

$$\begin{aligned} \theta^{k+1} &= \theta^k + \alpha^k \\ &= \theta^k + \left[\mathrm{D}\Psi_T(\theta^k)^* \mathrm{D}\Psi_T(\theta^k) + \varepsilon \frac{\gamma^2}{M^2} \Lambda_0^{-m} \right]^{-1} \mathrm{D}\Psi_T(\theta^k)^* (y_\gamma - \Psi_T(\theta^k)), \end{aligned} \tag{4.15}$$

where the number of iterations of this scheme is finite, thus acting as a regularization, and typically controlled by classical Morozov-like criteria, as advised in [Bal et al., 2013]. The parameter $\frac{\gamma^2}{M^2}$ is a scaling parameter taking into account the prior over noise ratio, while ε will give us an additional degree of freedom for weighting the regularization in the LM algorithm, see below and in particular Remark 4.3.1. Then, we can see the increment as

$$\alpha^k = \arg\min_{\alpha \in \mathcal{P}} \left\{ \frac{\varepsilon}{2M^2} \|\alpha\|_{\mathcal{P}}^2 + \frac{1}{2\gamma^2} \|y_\gamma - \Psi_T(\theta^k) - \mathrm{D}\Psi_T(\theta^k)\alpha\|_{\mathrm{L}^2((0,T);\mathcal{Y})}^2 \right\}, \tag{4.16}$$

Remark. *Note again that by penalizing* $\|\alpha\|_{\mathcal{P}}^2$ *we penalize an incremental displacement* $\|\tilde{\boldsymbol{u}}_0(\alpha)\|_{\mathcal{V}^m}^2$*. Therefore, we can force the first increments of the seeking displacement to belong to* $\mathcal{K}$ *since a sufficiently smooth displacement with a sufficiently small amplitude will necessarily be admissible. However, this may come at the price of too much regularization, which can only be compensated by a higher number of iterations of the LM algorithm. This is illustrated in the numerical section.*

Let us now specify the tangent operator

$$\forall \theta \in \mathcal{P}, \quad \mathrm{D}\Psi_T(\theta) : \left| \begin{aligned} \mathcal{H} &\to \mathrm{L}^2((0,T);\mathcal{Y}), \\ \alpha &\mapsto y = C\zeta_{|\theta,\alpha} \end{aligned} \right.$$

and $\zeta_{|\theta,\alpha}$ is a mild solution of

$$\begin{cases} \dot{\zeta}_{|\theta,\alpha}(t) = A(\theta)\zeta_{|\theta,\alpha}(t) + B(\theta, z_{|\theta}(t))\alpha, \quad t \in [0,T] \\ \zeta_{|\theta,\alpha}(0) = 0, \end{cases}$$

where we have introduced the linear operator representing the tangent of $A(\theta)$ *w.r.t.* the parameter θ

$$\forall \theta \in \mathcal{P}, \quad \forall z \in \mathcal{D}(A), \quad \mathcal{L}(\mathcal{P}, \mathcal{Z}) \ni B(\theta, z) : \mathcal{P} \ni \alpha \mapsto (\mathrm{D}_\theta A(\theta)\alpha) z \in \mathcal{Z}.$$

Therefore, we face a linear-quadratic optimal control problem, that can be minimized using the following Reduced-Order Kalman Filter (ROKF) sequential estimator reformulated from the initial design found in [Moireau, Chapelle, and Le Tallec, 2008] – see also the more recent review [Moireau, 2022]. To this end, let us first introduce the set of sensitivity operators $(L_{|\theta}(t))_{t\geq 0}$ in $\mathcal{L}(\mathcal{P},\mathcal{Z})$ defined for all time $t \geq 0$ by

$$L_{|\theta}(t) : \alpha \mapsto \zeta(t) \text{ the mild solution of}$$
$$\begin{cases} \dot{\zeta}(s) = A(\theta)\zeta(s) + B(\theta, z_{|\theta}(s))\alpha, & s \in [0,t] \\ \zeta(0) = 0. \end{cases} \tag{4.17}$$

We easily verify that $L_{|\theta} \in \mathrm{C}^0([0,T];\mathcal{L}(\mathcal{P},\mathcal{Z}))$ the space of continuous mapping from $[0,T]$ to $\mathcal{L}(\mathcal{P},\mathcal{Z})$ endowed with the uniform convergence topology [Bensoussan et al., 2007].

Then, we introduce the time-dependent Riccati operator $(\Lambda_{|\theta})_{t\geq 0}$ – which can be interpreted as a parameter-covariance operator [Bensoussan et al., 2007] – solution in $\mathrm{C}^0([0,T];\mathcal{S}_+^*(\mathcal{P}))$ – with $\mathcal{S}_+^*(\mathcal{P})$ the space of symmetric positive definite bounded operators – of

$$\begin{cases} \dot{\Lambda}_{|\theta}(t) = -\frac{1}{\gamma^2}\Lambda_{|\theta}(t)L_{|\theta}(t)^*C^*CL_{|\theta}(t)\Lambda_{|\theta}(t), & t \in [0,T] \\ \Lambda_{|\theta}(0) = \frac{M^2}{\varepsilon}\Lambda_0^m \end{cases} \tag{4.18}$$

We would like to underline the fact that here the adjoint operator $L_{|\theta}(t)^*$ is here defined with respect to the ℓ^2-norm as $\theta \in \mathcal{P} \subset \ell^2(\mathbb{R})$.

We then define the sequential estimator

$$\begin{cases} \dot{\hat{\zeta}}^k(t) = A(\theta^k)\hat{\zeta}^k(t) + B(\theta^k, z_{|\theta^k}(t))\hat{\alpha}^k(t) + L_{|\theta^k}(t)\dot{\hat{\alpha}}^k(t), & t \in [0,T] \\ \dot{\hat{\alpha}}^k(t) = \frac{1}{\gamma^2}\Lambda_{|\theta^k}(t)L_{|\theta^k}(t)^*C^*(y_\gamma(t) - Cz_{|\theta^k}(t) - C\hat{\zeta}^k(t)), & t \in [0,T] \\ \hat{\zeta}^k(0) = 0 \\ \hat{\alpha}^k(0) = 0 \end{cases} \tag{4.19}$$

which, ultimately, sequentially solve the minimization problem (4.16) as recall in the next theorem justified in [Moireau, 2022].

Theorem 4.3.1. *At every iteration k, the mild solution $\hat{\alpha}^k \in \mathrm{C}^0([0,T];\mathcal{P})$ and $\hat{\zeta}^k \in \mathrm{C}^0([0,T];\mathcal{Z})$ of* (4.19) *satisfies*

$$\hat{\alpha}^k(T) = \alpha^k \text{ and } \hat{\zeta}^k(T) = \zeta_{|\theta^k,\alpha^k}(T). \tag{4.20}$$

In the previous theorem, we understand that we can sequentially compute $t \mapsto \hat{\alpha}^k$ but also $t \mapsto \hat{\zeta}^k$ which is interpreted as a sequential estimator of the associated trajectory $t \mapsto \zeta_{|\theta^k,\alpha^k}$. Note that this estimator can be computed together with $\hat{\alpha}^k$ in only one coupled forward dynamics.

In fact, instead of a parameter increment, we are more interested in reconstructing the parameter itself, namely $t \mapsto \bar{\theta}^k(t) = \theta^k + \hat{\alpha}^k(t)$, and its associated trajectory $t \mapsto \bar{z}^k(t) = z_{|\bar{\theta}^k(t)}(t)$. As $t \mapsto \hat{\alpha}^k(t)$ is a time-dependent trajectory, we have

$$\dot{\bar{z}}^k = \frac{\mathrm{d}}{\mathrm{d}t}\left[z_{|\bar{\theta}^k(\cdot)}\right] = \dot{z}_{|\bar{\theta}^k(\cdot)} + \mathrm{D}_{\theta(\cdot)}z_{|\theta(\cdot)}\Big|_{\theta(\cdot)=\bar{\theta}^k(\cdot)}\dot{\hat{\alpha}}^k,$$

and the sensitivity operator $L_{|\bar{\theta}^k(\cdot)} = \mathrm{D}_{\theta(\cdot)}z_{|\theta(\cdot)}\Big|_{\theta(\cdot)=\bar{\theta}^k(\cdot)}$ is nothing else than, for all time $t \geq 0$,

$$L_{|\bar{\theta}^k(\cdot)}(t) : \alpha \mapsto \zeta(t) \text{ is a mild solution of}$$
$$\begin{cases} \dot{\zeta}(s) = A(\bar{\theta}^k(s))\zeta(s) + B(\bar{\theta}^k(s), z_{|\bar{\theta}^k(t)}(s))\alpha, & s \in [0,t] \\ \zeta(0) = 0 \end{cases} \tag{4.21}$$

Here note that in the previous dynamics, $s \mapsto \bar{\theta}^k(s)$ is a function of time, implying a time-dependent operator $s \mapsto A(\bar{\theta}^k(s))$. The existence of a solution for such dynamics is fundamentally based on the study of evolution equation operator in general [Pazy, 1983, Chapter 5].

As $\dot{\hat{\alpha}}^k = \dot{\bar{\theta}}^k$, and by approximating $z_{|\theta^k} + \hat{\zeta}^k \simeq z_{|\theta^k+\hat{\alpha}^k} = \bar{z}^k$ we could compute the sequential estimator

$$\begin{cases} \dot{\bar{z}}^k = A(\bar{\theta}^k)\bar{z}^k + r + L_{|\bar{\theta}^k}\dot{\bar{\theta}}^k, & t \in [0,T], \\ \dot{\bar{\theta}}^k = \frac{1}{\gamma^2}\Lambda_{|\theta^k}(t)L_{|\theta^k}(t)^* C^*(y_\gamma - C\bar{z}^k), & t \in [0,T], \\ \bar{z}^k(0) = 0 \\ \bar{\theta}^k(0) = \bar{\theta}^k, \end{cases} \tag{4.22}$$

such that at final time $\bar{\theta}^k(T) \simeq \bar{\theta}^{k+1}$.

One drawback in solving (4.22) is that it necessitates to propagate two sensitivity operators $L_{|\theta^k}$ and $L_{|\bar{\theta}^k}$. Moreover $L_{|\theta^k}$ necessitates to store the complete trajectory $z_{|\theta^k}$. To circumvent these drawbacks, one could imagine to instead solve

$$\begin{cases} \dot{\hat{z}}^k = A(\hat{\theta}^k)\hat{z}^k + r + L^k\dot{\hat{\theta}}^k, & t \in [0,T], \\ \dot{\hat{\theta}}^k = \frac{1}{\gamma^2}\Lambda^k(t)L^k(t)^* C^*(y_\gamma - C\hat{z}^k), & t \in [0,T], \\ \hat{z}^k(0) = 0 \\ \hat{\theta}^k(0) = \hat{\theta}^{k-1}(T) \end{cases} \tag{4.23}$$

where, from now on, we simply use the notation $L^k = L_{|\hat{\theta}^k}$ and $\Lambda^k = \Lambda_{|\hat{\theta}^k}$ and we see that the sensitivity and the covariance operator are updated through time. Indeed, Λ^k is an operator that should be seen as a *mild* solution of the dynamics

$$\begin{cases} \dot{\Lambda}^k(t) = -\frac{1}{\gamma^2}\Lambda^k(t)L^k(t)^* C^* C L_{|\theta}(t)\Lambda^k(t), & t \in [0,T] \\ \Lambda^k(0) = \frac{M^2}{\varepsilon}\Lambda_0^m, \end{cases} \tag{4.24}$$

while L^k can be seen from (4.21) as a mild solution in $C^0([0,T], \mathcal{L}(\mathcal{P}, \mathcal{Z}))$ of

$$\begin{cases} \dot{L}^k(t) = A(\hat{\theta}^k)L^k(t) + B(\hat{\theta}^k(t), \hat{z}^k(t)) & t \in [0,T] \\ L^k(0) = 0. \end{cases} \tag{4.25}$$

We here recognize in (4.23) – combined with (4.24)-(4.25) – an iterated version of the Reduced-Order-Extended-Kalman-Filter (ROEKF) estimator proposed in [Moireau, Chapelle, and Le Tallec, 2008] for joint state and parameter estimation for wave-like equations, here formally generalized to infinite dimensional systems. One very strong advantage of such an estimator is that it is fully sequential in the sense that it does not require storing any trajectory in the iteration procedure. Note that the existence of a solution of (4.23) is much more intricate and is based on justifying that L^k is well-defined even for time-dependent parameter $\hat{\theta}^k$. This question is typically covered by the definition of stable families of generators in evolution equation [Pazy, 1983], see also [Afshar and Germ, 2020] for similar questions when defining the Extended Kalman Filter for infinite dimensional systems.

In this work, we call iROEKF the proposed iterated ROEKF. However, we must warn the reader that in our case the iteration k is to be considered as an outer loop of the LM descent. This iROEKF is therefore different from the classical iterated Extended Kalman Filter presented in the literature [Särkkä, 2013].

4.3.2 Space-time-discretized version of the iterated reduced-order Extended Kalman Filter

Since we ultimately solve a discretized version of (4.23), we now present a discretization strategy based on a stable discretization of the Levenberg-Marquardt increment estimator (4.19). Our discretization strategy is based on the fundamental principle *discretize-then-optimize*, which means that we first discretize the direct problem and then reapply the equivalence of Levenberg-Marquardt and dynamic programming at the discretized level, as recommended in [Moireau, 2022]. For (4.7), we consider a finite-element discretization based on the Spectral Finite Element Method (SFEM) [Cohen, 2002] leading to the formulation

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \quad (\partial^2_{tt}\boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} + (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} = (\boldsymbol{f}, \boldsymbol{w}_h)_{\mathcal{V}_h}. \tag{4.26}$$

To construct the finite dimensional space $\mathcal{V}_h$ we assume given a partition $\mathcal{T}_h$ of quadrangles (in 2D) or hexahedra (in 3D) of the domain Ω, namely

$$\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K, \quad \forall\,(K, L) \in \mathcal{T}_h \times \mathcal{T}_h \quad \mathring{K} \cap \mathring{L} = \emptyset,$$

with maximum diameter given by h. Then $\mathcal{V}_h \subset \mathcal{V}$ is obtained using a $\mathbb{Q}_k$-Lagrangian basis on a set of nodes $\{\xi_i\}_{i=1}^{N_h}$,

$$\mathcal{V}_h = \left(\mathrm{span}\{\varphi_i\}_{i=1}^{N_h}\right)^3, \quad \forall 1 \le i, j \le N_h, \quad \varphi_i(\xi_j) = \delta_{ij}, \quad \forall\, K \in \mathcal{T}_h, \quad \varphi_{i|K} \circ F_K \in \mathbb{Q}_k.$$

where the mapping from the reference element to any element $K \in \mathcal{T}_h$ is denoted F_K – see [Cohen, 2002] for more details. The nodes $\{\xi_i\}$ are obtained using Gauss-Lobatto integration points on a reference square (in 2D) or cube (in 3D). One fundamental aspect of the efficiency of the SFEM is that $\mathcal{V}_h$ is equipped with a scalar product that – using a quadrature formula on the nodes $\{\xi_i\}$ – leads at the algebraic level to a diagonal mass matrix. This is called a mass-lumping strategy and preserves a near-optimal accuracy [Duruflé, Grob, and Joly, 2009]. The mentioned quadrature formula is also employed to compute the operator $A_h(\boldsymbol{u}_{h0})$, thus given by

$$\forall (\boldsymbol{u}_h, \boldsymbol{w}_h) \in \mathcal{V}_h^2, \quad (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} = \oint_\Omega \boldsymbol{\nabla}\boldsymbol{u}_h : \mathrm{D}^2_{\boldsymbol{F}}\mathscr{W}(\mathbf{x}, \boldsymbol{F}_{0h}) : \boldsymbol{\nabla}\boldsymbol{w}_h \,\mathrm{d}\mathbf{x},$$

where $\oint$ stands for an integral computed from the quadrature formula, and $\boldsymbol{F}_{0h} = \mathbf{Id} + \boldsymbol{\nabla}\boldsymbol{u}_{h0}$ with $\boldsymbol{u}_{h0}$ an interpolation in $\mathcal{V}_h$ of $\boldsymbol{u}_0$. Note that the operator inherits some properties of the operator $A(\boldsymbol{u}_{h0})$, in particular it is self-adjoint. However it is not clear that it satisfies a positivity property of the form (4.8) for at least two reasons, the interpolation of $\boldsymbol{u}_{0h}$ and the use of the quadrature formula. Therefore we are led to the following assumption that is sufficient to have a well-posed discrete problem,

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \quad (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{w}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} \ge 0.$$

The spatial discretization is then followed by an explicit time-discretization. Being given a time step Δt, the solution $\boldsymbol{u}_h(t)$ is approximated at time $t^n = n\Delta t$ by solving

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \left(\frac{\boldsymbol{u}_h^{n+1} - 2\boldsymbol{u}_h^n + \boldsymbol{u}_h^{n-1}}{\Delta t^2}, \boldsymbol{w}_h\right)_{\mathcal{V}_h} + (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h^n, \boldsymbol{w}_h)_{\mathcal{V}_h} = (\Pi_h \boldsymbol{f}(t^n), \boldsymbol{w}_h)_{\mathcal{V}_h}, \tag{4.27}$$

where $\Pi_h \boldsymbol{f}$ is a projection of $\boldsymbol{f}$ in $\mathcal{V}_h$. Thanks to the mass-lumping strategy the computation of $\boldsymbol{u}_h^{n+1}$ involves – at the algebraic level – only the inversion of a diagonal mass matrix. This discretization is stable for small enough time step Δt, namely the time step must satisfy the CFL condition

$$\Delta t \le 2 \left(\sup_{\boldsymbol{w}_h \in \mathcal{V}_h} \frac{(A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{w}_h, \boldsymbol{w}_h)_{\mathcal{V}_h}}{\|\boldsymbol{w}_h\|^2_{\mathcal{V}_h}} \right)^{-1/2}. \tag{4.28}$$

The CFL condition depends on the mesh size h (in particular Δt behaves like $O(h)$) and on the mesh quality but also on the gradient of the displacement field $\boldsymbol{u}_{h0}$. We can finally rewrite the time-discretization (4.27) as

$$\begin{cases} \dfrac{\boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n}{\Delta t} = \boldsymbol{v}_h^{n+1} \\ \dfrac{\boldsymbol{v}_h^{n+1} - \boldsymbol{v}_h^n}{\Delta t} + A_{0h}(\theta_h)\boldsymbol{u}_h^n = \Pi_h \boldsymbol{f}(t^n) \end{cases} \tag{4.29}$$

where we replace the $\boldsymbol{u}_{h0}$ dependency by the $\theta_h \in \mathcal{P}_h \simeq \mathbb{R}^{\mathrm{N}_p}$ dependency such that

$$\boldsymbol{u}_{h0} = \sum_{j=1}^{\mathrm{N}_p} \theta_{h,j} \boldsymbol{\varphi}_{h,j} \quad \text{with} \quad A_{0h}(0)\boldsymbol{\varphi}_{h,j} = \mu_{h,j}\boldsymbol{\varphi}_{h,j},$$

namely the $(\mu_{h,j}, \boldsymbol{\varphi}_{h,j})$ are the eigenelements of the discrete operators $A_{0h}(0)$ (ordered increasingly with respect to the eigenvalues). Then, by defining $z_h^n = (\boldsymbol{u}_h^n, \boldsymbol{v}_h^n)^\intercal$, we rewrite (4.29) in the following abstract state-space form

$$\begin{cases} z_h^{n+1} = \Phi_{h,\Delta t}(\theta_h) z_h^n + r_{h,\Delta t}^{n+1} \\ z_h^0 = 0 \end{cases} \tag{4.30}$$

where

$$\Phi_{h,\Delta t}(\theta_h) : z_h \mapsto \begin{pmatrix} \mathrm{Id}_{\mathcal{V}_h} - \Delta t^2 A_{0h}(\theta_h) & \Delta t \mathrm{Id}_{\mathcal{V}_h} \\ \Delta t A_{0h}(\theta_h) & \mathrm{Id}_{\mathcal{V}_h} \end{pmatrix} z_h \text{ and } r_{h,\Delta t}^{n+1} = \begin{pmatrix} \Delta t^2\, \Pi_h \boldsymbol{f}(t^n) \\ \Delta t\, \Pi_h \boldsymbol{f}(t^n) \end{pmatrix}.$$

Using the time-discretized system (4.30), we propose to solve the Levenberg-Marquardt procedure

$$\begin{aligned} \theta_h^{k+1} &= \theta_h^k + \alpha_h^k \\ &= \theta_h^k + \left[\mathrm{D}\Psi_{h,N}(\theta^k)^* \mathrm{D}\Psi_{h,N}(\theta^k) + \varepsilon \frac{\gamma^2}{M^2} \Lambda_{0h}^{-m} \right]^{-1} \mathrm{D}\Psi_{h,N}(\theta_h^k)^* (y_\gamma - \Psi_{h,N}(\theta_h^k)), \end{aligned}$$

where

$$\Psi_{h,N} : \begin{cases} \mathcal{P}_h \to \ell^2((1:N);\mathcal{Y}) \\ \theta_h \mapsto \left\{ \displaystyle\sum_{j=1}^{n} C_h \Phi_{h,\Delta t}^{n-j}(\theta_h) r_{h,\Delta t}^j \right\}_{n=1}^{N} \end{cases}$$

which can be equivalently rewritten in the following form

$$\alpha_h^k = \arg\min_{\alpha_h \in \mathcal{P}} \left\{ \frac{\varepsilon}{2M^2} \|\alpha_h\|_{\mathcal{P}}^2 + \frac{\Delta t}{2\gamma^2} \|y_\gamma - \Psi_{h,N}(\theta^k) - \mathrm{D}\Psi_{h,N}(\theta_h^k)\alpha_h\|_{\ell^2((1:N);\mathcal{Y})}^2 \right\}. \tag{4.31}$$

Then, we define the fully discrete sensitivity operator, for all $n \in [0:N]$, as

$$\begin{aligned} &L_{h|\theta_h}^{k,n} : \alpha_h \mapsto \zeta_h^n \text{ the solution of} \\ &\qquad \begin{cases} \zeta_h^{j+1} = \Phi_{h,\Delta t}(\theta_h) z_h^j + B_{h,\Delta t}(\theta_h, z_h^j)\alpha_h, & j \in [0:n-1] \\ \zeta_h^0 = 0 \end{cases} \end{aligned}$$

with

$$B_{h,\Delta t}(\theta_h, z_h)\alpha_h = \begin{pmatrix} \Delta t^2 \mathrm{D}_\theta A_{0h}(\theta_h)\alpha_h & 0 \\ -\Delta t \mathrm{D}_\theta A_{0h}(\theta_h)\alpha_h & 0 \end{pmatrix} z_h.$$

Moreover, we define the fully discrete covariance operator $\Lambda_{h|\theta_h}^{k,n}$ satisfying the dynamics

$$\begin{cases} (\Lambda_{h|\theta_h}^{k,n+1})^{-1} = (\Lambda_{h|\theta_h}^{k,n})^{-1} + \dfrac{\Delta t}{\gamma^2} L_{h|\theta_h}^{k,n+1} C_h^* C_h L_{h|\theta_h}^{k,n+1}, & n \in [0:N] \\ \Lambda_{h|\theta_h}^{k,0} = \dfrac{M^2}{\varepsilon} \Lambda_{h0}^{m}. \end{cases}$$

If we now choose to solve from $\hat{\zeta}_h^{k,0} = 0$, and $\hat{\alpha}_h^{k,0} = 0$, for all $k \leq 0$ and $n \in [0\!:\!N-1]$

$$\begin{cases} \hat{\zeta}_h^{k,n+1-} = \Phi_{h,\Delta t}(\theta_h)\hat{\zeta}_h^{k,n} + B_{h,\Delta t}(\theta_h^k, z^n_{h|\theta_h^k})\hat{\alpha}_h^{k,n} \\ \hat{\alpha}_h^{k,n+1} = \hat{\alpha}_h^{k,n} + \dfrac{\Delta t}{\gamma^2}\Lambda_{h|\theta_h}^{k,n+1} L_{h|\theta_h}^{k,n+1*} C_h^*(y_\gamma^{n+1} - C_h z^{n+1}_{h|\theta_h^k} - C_h\hat{\zeta}_h^{k,n+1-}) \\ \hat{\zeta}_h^{k,n+1} = \hat{\zeta}_h^{k,n+1-} + L_{h|\theta_h}^{k,n+1}(\hat{\alpha}_h^{k,n+1} - \hat{\alpha}_h^{k,n}) \end{cases} \tag{4.32}$$

where the "$n-$" and "n" exponents correspond to the two steps of a splitting time-scheme for $\hat{\zeta}_h^{k,n}$ – where the "$n-$" step is often referred as the prediction (or forecast) step and the "n" step is referred as the correction (or analysis) step – we have at the discrete level the following equivalence proved in [Moireau, 2022].

Theorem 4.3.2. *At every iteration k of the Levenberg-Marquardt minimization, we have*

$$\hat{\alpha}_h^{k,N} = \alpha_h^k \text{ and } \hat{\zeta}_h^{k,N} = \zeta^N_{h|\theta^k,\alpha^k} \tag{4.33}$$

The last theorem helps to understand that we keep at the discrete-time level the same equivalence between the least square minimization of a discretized criterion and the sequential approach. Therefore pursuing the same strategy than for the continuous-time level leads us to defining the Discrete-Time iterated Reduced Order Extended Kalman Filter (DT-iROEKF) as an approximation of the Levenberg-Marquardt minimization strategy for discrete-time dynamics. The resulting joint state-parameter sequential estimator based on this DT-iROEKF is initialized from $\hat{z}_h^k = 0, \hat{\theta}_h^{k,0} = \hat{\theta}_h^{k-1,N}$ if $k > 1$ and $\hat{\theta}_h^{1,0} = \hat{\theta}_{0h}$, while $L^{k,0} = 0, \Lambda^{k,0} = \varepsilon^{-1}M^2\Lambda_{h0}^m$. Then the recursive dynamics reads for all $k \geq 1$ and for all $n \in [0\!:\!N-1]$

— Prediction / Forecast:

$$\begin{cases} \hat{z}_h^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})\hat{z}_h^{k,n} + r_{h,\Delta t}^{n+1}, \\ L_h^{k,n+1} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})L_h^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}). \end{cases}$$

— Correction / Analysis:

$$\begin{cases} (\Lambda_h^{k,n+1})^{-1} = (\Lambda_h^{k,n})^{-1} + \dfrac{\Delta t}{\gamma^2} L_h^{k,n+1} C_h^* C_h L_h^{k,n+1}, \\ \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2}\Lambda_h^{k,n+1} L_h^{k,n+1*} C_h^*(y_\gamma^{n+1} - C_h\hat{z}_h^{k,n+1-}), \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + L_h^{k,n+1}(\hat{\theta}_h^{k,n+1} - \hat{\theta}_h^{k,n}). \end{cases}$$

Here again, the approximation relies on the fact that the sensitivity and the covariances are computed sequentially from the current estimation $\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}$.

4.3.3 From a square root to a tangent-free formulation

We are now ready to formulate a practical implementation of our DT-iROEKF. In this respect, we recast the DT-iROEKF formulation into a so-called square-root form as commonly done in data assimilation for the robust implementation of the Kalman approaches [Simon, 2006]. Then we show how we can replace, up to second-order terms, the tangent computations by finite difference scheme over wave solution, hence leading to an original tangent-free approach.

The SEEK algorithm for parameter identification. In this section, we start by formulating a square-root approach of the reduced-order Extended Kalman filter that we are going to apply to our state-parameter decomposition. This strategy is an adaption to parameter estimation of the square root algorithm initially presented in [Rozier et al., 2007]. Following their strategy, we choose to store not $L_h^{k,n}$ but $L_h^{k,n}D_h^{k,n}$ with $D_h^{k,n}$ a Cholesky factorization of $\Lambda_h^{k,n}$. More precisely, we define

$$D_h^{k,n} := \mathrm{chol}(\Lambda_h^{k,n}) \text{ namely } \Lambda_h^{k,n} = D_h^{k,n}D_h^{k,n*} \tag{4.34}$$

$$S_h^{k,n} := L_h^{k,n}D_h^{k,n} \tag{4.35}$$

$$S_h^{k,n+1-} := L_h^{k,n+1}D_h^{k,n} \tag{4.36}$$

Note that $D_h^{k,n}$ can be equivalently represented as a collection of N_p members

$$\{D_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_p} \in \mathcal{P}_h,$$

that can be seen as the "columns" of the matrix representations of the operator. Each member is comparable to a parameter increment around the observer trajectory $\hat{\theta}_h^{k,n}$, hence is associated with a reconstructed displacement

$$\boldsymbol{u}_{0h} = \sum_{1\leq i\leq \mathrm{N}_p} D_{h,ij}^{k,n}\boldsymbol{\varphi}_{h,i},$$

with $\boldsymbol{\varphi}_{h,i}$ the normalized eigenvectors of the operator A_{0h}. In the same manner $S_h^{k,n}$ and $S_h^{k,n+1-}$ can be equivalently represented as a collection of N_p members

$$\{S_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_p} \in \mathcal{Z}_h.$$

Each member is comparable to the variation of a wave solution for the specific parameter increment $D_{h,j}^{k,n}$. We are going to prove the following recursive dynamics, defining again a splitting time-scheme with one prediction step followed by a correction step.

Theorem 4.3.3. *For all $k \geq 1$ and for all $n \in [0:N-1]$, we define the Gramian operator $G^{k,n}$ by*

$$G^{k,n} = \mathrm{Id}_{\mathcal{Y}} + \frac{\Delta t}{\gamma^2}C_hS_h^{k,n-}S_h^{k,n-*}C_h^*,$$

and we denote by $Q^{k,n}$ the Cholesky decomposition of $(G^{k,n})^{-1} = Q^{k,n}Q^{k,n*}$. With these notations, we obtain the following dynamics:

— *Prediction / Forecast:*

$$S_{h,j}^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n})D_{h,j}^{k,n}, \quad \forall j \in [1:\mathrm{N}_p]. \tag{4.37}$$

— *Correction / Analysis:*

$$S_h^{k,n+1} = S_h^{k,n+1-}Q^{k,n+1}, \quad \textit{and} \quad D_h^{k,n+1} = D_h^{k,n}Q^{k,n+1}. \tag{4.38}$$

Proof. We proceed by induction. This is true at $n = 0$ since $L_h^{k,0} = 0$. Let us now assume that (4.37) and (4.38) are valid at iteration n. For the prediction step, we have for all $j \in [1:\mathrm{N}_p]$,

$$\begin{aligned}
S_{h,j}^{k,n+1-} &= L_{h,j}^{k,n+1}D_{h,j}^{k,n} \\
&= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})L_{h,j}^{k,n}D_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1-})D_{h,j}^{k,n} \\
&= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1-})D_{h,j}^{k,n}
\end{aligned}$$

so that (4.37) is satisfied at iteration n. From Woodbury inversion formula, we have that

$$\begin{aligned}
(G^{k,n+1})^{-1} &= \left(\mathrm{Id}_{\mathcal{Y}} + \frac{\Delta t}{\gamma^2}C_hS_h^{k,n+1-}S_h^{k,n+1-*}C_h^*\right)^{-1} \\
&= \mathrm{Id}_{\mathcal{P}_h} - \Delta t\, S_h^{k,n+1-*}C_h^*[\gamma^2\mathrm{Id}_{\mathcal{Y}} + \Delta t\, C_hS_h^{k,n+1-}S_h^{k,n+1-*}C_h^*]^{-1}C_hS_h^{k,n+1-} \\
&= Q^{k,n+1}Q^{k,n+1*}
\end{aligned}$$

Therefore if we now define $D_h^{k,n+1} = D_h^{k,n} Q^{k,n+1}$ then

$$\begin{aligned} D_h^{k,n+1} D_h^{k,n+1*} &= D_h^{k,n} Q^{k,n+1} Q^{k,n+1*} D_h^{k,n} \\ &= \Lambda_h^{k,n} \\ &\quad - \Delta t \Lambda_h^{k,n} L_h^{k,n+1*} C_h^* [\gamma^2 \mathrm{Id}_{\mathcal{Y}} + \Delta t C_h L_h^{k,n+1} \Lambda_h^{k,n} L_h^{k,n+1*} C_h^*]^{-1} C_h L_h^{k,n+1} \end{aligned}$$

which, with a second use of Woodbury inversion formula gives $D_h^{k,n+1} D_h^{k,n+1*} = \Lambda_h^{k,n+1}$. Finally, we have that

$$S_h^{k,n+1} = L_h^{k,n+1} D_h^{k,n+1} = L_h^{k,n+1} D_h^{k,n} Q^{k,n+1} = S_h^{k,n+1-} Q^{k,n+1},$$

which concludes the proof. □

Note that in practice, one can use the component-wise expression of the Gramian matrix, namely

$$G_{ij}^{k,n} = \delta_{ij} + \frac{\Delta t}{\gamma^2} (C_h S_{h,i}^{k,n-}, C_h S_{h,j}^{k,n-})_{\mathcal{Y}}, \quad \forall 1 \leq i, j \leq \mathrm{N}_p.$$

In the same fashion, we have

$$S_{h,j}^{k,n+1} = \sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} S_{h,i}^{k,n+1-}, \quad \text{and} \quad D_{h,i}^{k,n+1} = \sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} D_{h,i}^{k,n}.$$

From the previous theorem, we then easily verify that the DT-iROEKF estimator can be computed using only $S_h^{k,n-}$, $S_h^{k,n}$ and $D_h^{k,n}$ with

— Prediction / Forecast:

$$\hat{z}_h^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) \hat{z}_h^{k,n} + r_{h,\Delta t}^{n+1}, \tag{4.39}$$

— Correction / Analysis:

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} D_h^{k,n+1} S_h^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-}), & (4.40) \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} S_h^{k,n+1} S_h^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-}). & (4.41) \end{cases}$$

And again in practice, solving (4.40) and (4.41) is performed by developing

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1 \leq j \leq \mathrm{N}_p} (C_h S_{h,j}^{k,n+1}, y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-})_{\mathcal{Y}} D_{h,j}^{k,n+1}, \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1 \leq j \leq \mathrm{N}_p} (C_h S_{h,j}^{k,n+1}, y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-})_{\mathcal{Y}} S_{h,j}^{k,n+1}. \end{cases}$$

A tangent-free alternative. Ultimately, (4.37) can be seen as the tangent equation of

$$\begin{aligned} S_{h,j}^{k,n+1-} &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}) D_{h,j}^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}) D_{h,j}^{k,n} - \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) \hat{z}_h^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n} + D_{h,j}^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) - \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) \hat{z}_h^{k,n} + O(\|D_{h,j}^{k,n}\|^2) \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n} + D_{h,j}^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) + r_{h,\Delta t}^{n+1} - \hat{z}_h^{k,n+1-} + O(\|D_{h,j}^{k,n}\|^2). \end{aligned} \tag{4.42}$$

Therefore, by computing the forecast of each member $S_{h,j}^{k,n+1-}$ by a finite difference between two wave equations, we do not need to compute the operator $B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1})$. This means that the forecast step can be performed by simply forecasting $\hat{z}_h^{k,n}$ and each member $S_{h,j}^{k,n}$.

4.3.4 The Reduced-Order UKF alternative

In the DT-iROEKF approach, we have seen that we can propagate the estimator $\hat{\theta}_h^{k,n}$, $\hat{z}_h^{k,n}$ and the increment set $\{S_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_p}$ and $\{D_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_p}$. These increments give the sensitivity direction around the estimator $\hat{\theta}_h^{k,n}$, $\hat{z}_h^{k,n}$. Revisiting [Pham, Verron, and Gourdeau, 1998; Moireau and Chapelle, 2011], the Reduced-Order Unscented Kalman alternative (ROUKF) to ROEKF, we are going to replace the computation of the increment sets by sets of so-called *sigma-points* around $\hat{\theta}_h^{k,n}$ and $\hat{z}_h^{k,n}$ so that the estimator is computed by an averaging formula like in a finite difference stencil. In essence, while EKF propagates a point-estimator and N_p directions around this point-estimator, UKF computes a stencil of sigma-points such that the point-estimator is a resulting average of this point. In practice we will need at least $\mathrm{N}_s = \mathrm{N}_p + 1$ sigma-points when using a simplex stencil, but more complex stencils with $\mathrm{N}_s \geq \mathrm{N}_p + 1$ can also be proposed with additional trajectories to be computed. Note that our presentation clarifies the generalization of ROUKF to general complex stencil initially proposed in [Moireau and Chapelle, 2011; Moireau, Philippe and Chapelle, Dominique, 2011]. This may be of particular interest since increasing the number of sigma points results in a potentially more precise sampling and rendering of the sensitivity directions around the estimator.

Empirical mean and empirical covariance. Let us introduce a set of strictly positive weights $\beta = \{\beta_j\}_{j=1}^{\mathrm{N}_s} \in (\mathbb{R}_+^*)^{\mathrm{N}_s}$ such that

$$\sum_{1 \leq j \leq \mathrm{N}_s} \beta_j = 1.$$

From this set of weights, we introduce the following weight averaging operator defined, for all sample set $\{s_j\}_{j=1}^{\mathrm{N}_s} \in \mathcal{X}^{\mathrm{N}_s}$ of elements of a Hilbert space $\mathcal{X}$ (typically $\mathcal{Z}_h$, $\mathcal{P}_h$ or $\mathcal{Y}$), by

$$\mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) = \sum_{1 \leq j \leq \mathrm{N}_s} \beta_j s_j \in \mathcal{Z}.$$

This can be understood as an empirical mean of the $\{s_j\}_{j=1}^{\mathrm{N}_s}$. In the same fashion, we define the empirical covariance as the following linear operator

$$\mathbb{P}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) : \mathcal{X} \ni s \mapsto \sum_{1 \leq j \leq \mathrm{N}_s} \beta_j \left(s, s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s})\right)_{\mathcal{X}} \left(s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s})\right) \in \mathcal{X},$$

where $(\cdot,\cdot)_{\mathcal{X}}$ is the scalar product in $\mathcal{X}$. As soon as we are considering a finite-dimensional Euclidian space $\mathcal{X} = \mathbb{R}^{N_x}$, then

$$\mathbb{P}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) = \sum_{1 \leq j \leq \mathrm{N}_s} \beta_j (s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}))(s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}))^\intercal.$$

Sampling around the observer trajectory. Let us now introduce a scaling factor $\varrho \in \mathbb{R}_+^*$ and define a particular finite sequence of sigma-points $\{e_j\}_{j=1}^{\mathrm{N}_s}$ in the parameter space $\mathcal{P}_h$ such that

$$\mathbb{E}_\beta(\{e_j\}_{j=1}^{\mathrm{N}_s}) = 0 \text{ and } \mathbb{P}_\beta(\{e_j\}_{j=1}^{\mathrm{N}_s}) = \frac{1}{\varrho^2}\mathrm{Id}_{\mathcal{P}_h}.$$

Note that by choosing $\{e_j\}_{j=1}^{\mathrm{N}_s}$ such that $\|e_j\|_{\mathcal{P}}^2 = O(1)$ then $\varrho = O(\mathrm{N}_p^{-\frac{1}{2}})$. Again, defining $D_h^{k,n}$ such as $\Lambda_h^{k,n} = D_h^{k,n} D_h^{k,n*}$ – namely $D_h^{k,n}$ can be the square root of $\Lambda_h^{k,n}$ or a Cholesky decomposition – and using $L_h^{k,n}$ at iteration n, we define the *sampling* operation as constructing the following set of trajectories

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{z}_{h,j}^{k,n} = \hat{z}_h^{k,n} + \varrho L_h^{k,n} D_h^{k,n} e_j, \\ \hat{\theta}_{h,j}^{k,n} = \hat{\theta}_h^{k,n} + \varrho D_h^{k,n} e_j, \end{cases} \tag{4.43}$$

such that, by construction,

$$\mathbb{E}_\beta(\{\hat{z}^{k,n}_{h,j}\}_{j=1}^{\mathrm{N}_s}) = \hat{z}^{k,n}_h \quad \text{and} \quad \mathbb{E}_\beta(\{\hat{\theta}^{k,n}_{h,j}\}_{j=1}^{\mathrm{N}_s}) = \hat{\theta}^{k,n}_h,$$

while

$$\mathbb{P}_\beta(\{\hat{z}^{k,n}_{h,j}\}_{j=1}^{\mathrm{N}_s}) = L^{k,n}_h \Lambda^{k,n}_h L^{k,n*}_h \quad \text{and} \quad \mathbb{P}_\beta(\{\hat{\theta}^{k,n}_{h,j}\}_{j=1}^{\mathrm{N}_s}) = \Lambda^{k,n}_h.$$

Different kinds of sigma-points can be used to perform this sampling operation – see for instance [Moireau and Chapelle, 2011]. In Figure 4.1 we illustrate three different (albeit standard) sets of sigma-points, and we compare it with the sampling strategy derived from the tangent-free EKF approach presented in the previous section.

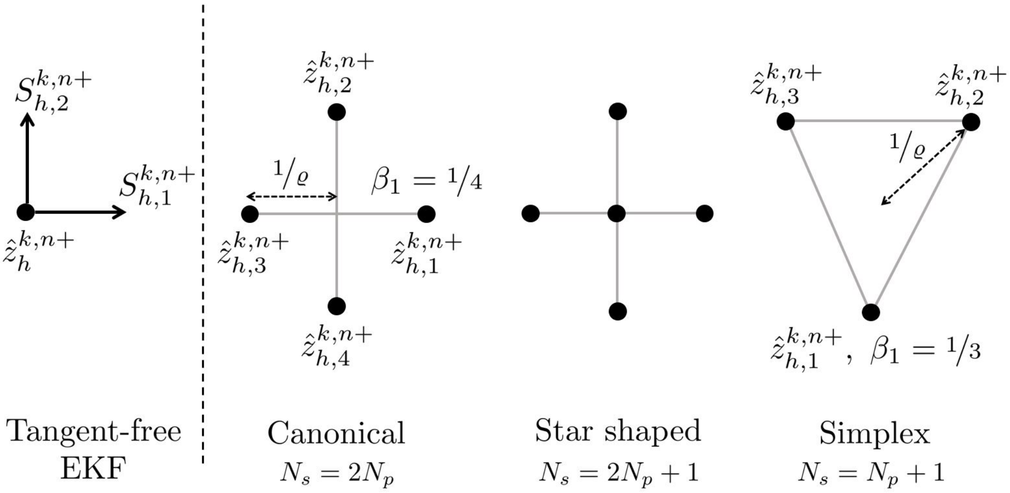


Figure 4.1 – Definition of standard sigma-points for the simple case of $\mathcal{P} = \mathbb{R}^2$ and comparison with the state and sensitivity trajectories associated with the tangent-free EKF approach.

Prediction step of the reduced-order UKF. The prediction step of the reduced-order UKF is simply obtained by propagating $\hat{z}^{k,n}_{h,j}$ using the discretized dynamics, namely

$$\hat{z}^{k,n+1-}_{h,j} = \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t}.$$

Using the empirical average operator, we can derive the state and parameter sensitivity around the estimator trajectory, namely we define the predictions

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \Sigma^{k,n+1-}_{h,j} & = \hat{z}^{k,n+1-}_{h,j} - \mathbb{E}_\beta(\{\hat{z}^{k,n+1-}_{h,j}\}_{j=1}^{\mathrm{N}_s}), \\ \Delta^{k,n}_{h,j} & = \hat{\theta}^{k,n}_{h,j} - \mathbb{E}_\beta(\{\hat{\theta}^{k,n}_{h,j})\}_{j=1}^{\mathrm{N}_s}. \end{cases} \tag{4.44}$$

From its definition, and using the definition of the sampling operation (4.43), one can easily verify that the parameter sensitivity satisfies

$$\Delta^{k,n}_{h,j} = \varrho D^{k,n}_h e_j, \quad \forall j \in [1\!:\!\mathrm{N}_s]. \tag{4.45}$$

The next lemma gives us a similar result for the state sensitivity.

Lemma 4.3.4. *The predicted state sensitivity defined in* (4.44) *satisfies*

$$\Sigma^{k,n+1-}_{h,j} = \varrho L^{k,n+1}_h D^{k,n}_h e_j + O(\varrho^2). \tag{4.46}$$

Proof. On the one hand, the state prediction of each sigma-point satisfies for any $j \in [1\!:\!\mathrm{N}_s]$,

$$\begin{aligned}\hat{z}^{k,n+1-}_{h,j} &= \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t} \\ &= \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})\hat{z}^{k,n}_{h} + r^{n+1}_{h,\Delta t} + \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})(\hat{z}^{k,n}_{h,j} - \hat{z}^{k,n}_{h}) \\ &\qquad\qquad + B_{h,\Delta t}(\hat{\theta}^{k,n}_{h}, \hat{z}^{k,n}_{h})(\hat{\theta}^{k,n}_{h,j} - \hat{\theta}^{k,n}_{h}) + O(\varrho^2).\end{aligned}$$

On the other hand, we can introduce

$$\hat{z}^{k,n+1-}_{h} = \mathbb{E}_\beta(\hat{z}^{k,n+1-}_{h,j}) = \mathbb{E}_\beta(\{\Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t}\}^{\mathrm{N}_s}_{j=1}),$$

and a Taylor expansion gives

$$\hat{z}^{k,n+1-}_{h} = \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})\hat{z}^{k,n}_{h} + r^{n+1}_{h,\Delta t} + O(\varrho^2).$$

Using the sampling operations defined in (4.43) we obtain

$$\begin{aligned}\Sigma^{k,n+1-}_{h,j} &= \hat{z}^{k,n+1-}_{h,j} - \hat{z}^{k,n+1-}_{h} \\ &= \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})(\hat{z}^{k,n}_{h,j} - \hat{z}^{k,n}_{h}) + B_{h,\Delta t}(\hat{\theta}^{k,n}_{h}, \hat{z}^{k,n}_{h})(\hat{\theta}^{k,n}_{h,j} - \hat{\theta}^{k,n}_{h}) + O(\varrho^2) \\ &= \varrho\big(\Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})L^{k,n}_{h} + B_{h,\Delta t}(\hat{\theta}^{k,n}_{h}, \hat{z}^{k,n}_{h})\big)D^{k,n}_{h} e_j + O(\varrho^2) \\ &= \varrho L^{k,n+1}_{h} D^{k,n}_{h} e_j + O(\varrho^2).\end{aligned}$$

□

Correction step of the reduced-order UKF. From the relations (4.45) and (4.46) we see that – up to second order terms – storing $\Sigma^{k,n+1-}_{h,j}$ and $\Delta^{k,n}_{h,j}$ is equivalent to the storing $S^{k,n+1-}_{h}$ and $D^{k,n}_{h}$ except that $\mathrm{N}_s > \mathrm{N}_p$. Then, following the definition of the correction (or analysis) step of the SEEK algorithm, namely (4.38), we propagate the state and parameter sensitivity through

$$\Sigma^{k,n+1}_{h,j} = \sum_{1\le i\le \mathrm{N}_s} Q^{k,n+1}_{ij}\Sigma^{k,n+1-}_{h,i} \text{ and } \Delta^{k,n+1}_{h,j} = \sum_{1\le i\le \mathrm{N}_s} Q^{k,n+1}_{ij}\Delta^{k,n}_{h,i},$$

where for all n, $Q^{k,n} \in \mathbb{M}_{\mathrm{N}_s}(\mathbb{R})$ is given by the Cholesky decomposition of $(G^{k,n})^{-1} = Q^{k,n}Q^{k,n*}$ a Gramian matrix defined by

$$G^{k,n}_{ij} = \delta_{ij} + \frac{\Delta t}{\gamma^2}(C_h\Sigma^{k,n-}_{h,i}, C_h\Sigma^{k,n-}_{h,j})_{\mathcal{Y}}, \quad \forall i,j \in [1\!:\!\mathrm{N}_s].$$

Using the same analogy, the corrected estimator can be rewritten in the form

$$\begin{cases}\hat{\theta}^{k,n+1}_{h} = \hat{\theta}^{k,n}_{h} + \dfrac{\Delta t}{\gamma^2}\displaystyle\sum_{1\le j\le \mathrm{N}_s}(C\Sigma^{k,n+1}_{h,j}, y^{n+1}_{\gamma} - C\hat{z}^{k,n+1-}_{h})_{\mathcal{Y}}\Delta^{k,n+1}_{h,j}, \\ \hat{z}^{k,n+1}_{h} = \hat{z}^{k,n+1-}_{h} + \dfrac{\Delta t}{\gamma^2}\displaystyle\sum_{1\le j\le \mathrm{N}_s}(C\Sigma^{k,n+1}_{h,j}, y^{n+1}_{\gamma} - C\hat{z}^{k,n+1-}_{h})_{\mathcal{Y}}\Sigma^{k,n+1}_{h,j}.\end{cases}$$

Furthermore, with a proof very similar to the proof of Theorem 4.3.3 (introducing in finite dimension the matrix $E = [e_1|\cdots|e_{\mathrm{N}_s}]$ and concatenating in columns the sigma-points $\{e_j\}^{\mathrm{N}_s}_{j=1}$), one can verify that for all $j \in [1\!:\!\mathrm{N}_s]$,

$$\Sigma^{k,n+1}_{h,j} = \varrho L^{k,n+1}_{h} D^{k,n+1}_{h} e_j + O(\varrho^2) \quad \text{and} \quad \Delta^{k,n+1}_{h,j} = \varrho D^{k,n+1}_{h} e_j + O(\varrho^2). \tag{4.47}$$

Hence, $\Sigma^{k,n+1}_{h,j}$ and $\Delta^{k,n+1}_{h,j}$ are equivalent to $S^{k,n+1}_{h}$ and $D^{k,n+1}_{h}$, up to second order terms, as for the prediction step. Once again, the only – but fundamental – difference compared to the EKF version is the fact that the Gramian is defined in $\mathbb{M}_{\mathrm{N}_s}(\mathbb{R})$ instead of $\mathbb{M}_{\mathrm{N}_p}(\mathbb{R})$. Note that, one important aspect of the relations (4.47) is that it enables us to rewrite the sampling operation (4.43) exclusively in terms of the computed state and parameter sensitivities. Namely, we recreate the sample at $n+1$ by discarding second-order terms in the following relations

$$\hat{z}_{h,j}^{k,n+1} = \hat{z}_h^{k,n+1} + \Sigma_{h,j}^{k,n+1} + O(\varrho^2), \quad \text{and} \quad \hat{\theta}_{h,j}^{k,n+1} = \hat{\theta}_h^{k,n} + \Delta_{h,j}^{k,n+1} + O(\varrho^2).$$

We provide in Figure 4.2 an illustration of its main steps, in order to fully apprehend the various unknowns appearing in the algorithm.

Final algorithm. We can now summarize the complete algorithm used in this work.

The Discrete-Time iterated Reduced-Order Unscented Kalman Filter (DT-iROUKF)

- Choice of the UKF sigma-points $\{e_j\}_{j=1}^{\mathrm{N}_s}$ and the associated weights $\{\beta_j\}_{j=1}^{\mathrm{N}_s}$.
- Initialization of the parameter prior from the modal decomposition

$$\hat{\boldsymbol{u}}_0(\mathbf{x}) = \sum_{1 \leq j \leq \mathrm{N}_p} \hat{\theta}_{0h,j} \boldsymbol{\varphi}_j(\mathbf{x}).$$

- For all $k \geq 1$ until convergence:
  - Initialization at $n = 0$:

  $\hookrightarrow$ Parameter initialization:

$$\hat{\theta}_h^{k,0} = \hat{\theta}_h^{k-1,N} \quad \text{if } k > 1, \quad \hat{\theta}_h^{1,0} = \hat{\theta}_{0h} \quad \text{otherwise.}$$

  $\hookrightarrow$ Parameter and state particles initialization:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{\theta}_{h,j}^{k,0} = \hat{\theta}_h^{k,0} + \varrho\, D^{k,0} e_j, \quad \text{with} \quad D^{k,0} = \mathrm{chol}(\Lambda_0^m), \\ \hat{z}_{h,j}^{0} = 0. \end{cases}$$

  - For $0 \leq n \leq N$:

  $\hookrightarrow$ Prediction step:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{z}_{h,j}^{k,n+1-} & = \Phi_{h,\Delta t}(\hat{\theta}_{h,j}^{k,n})\hat{z}_{h,j}^{k,n} + r_{h,\Delta t}^{n+1}, \\ \Sigma_{h,j}^{k,n+1-} & = \hat{z}_{h,j}^{k,n+1-} - \mathbb{E}_\beta(\{\hat{z}_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_s}), \\ \Delta_{h,j}^{k,n} & = \hat{\theta}_{h,j}^{k,n} - \mathbb{E}_\beta(\{\hat{\theta}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}). \end{cases}$$

  $\hookrightarrow$ Correction step:

    – Compute Gramian matrix and Cholesky decomposition of its inverse:

$$\begin{cases} G^{k,n+1} = (\delta_{ij} + \dfrac{\Delta t}{\gamma^2}(C_h \Sigma_{h,i}^{k,n+1-}, C_h \Sigma_{h,j}^{k,n+1-})_{\mathcal{Y}})_{1 \leq i,j \leq \mathrm{N}_s}, \\ Q^{k,n+1} = \mathrm{chol}((G^{k,n+1})^{-1}). \end{cases}$$

    – Compute corrected sensitivity matrices:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \Sigma_{h,j}^{k,n+1} = \displaystyle\sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} \Sigma_{h,i}^{k,n+1-}, \\ \Delta_{h,j}^{k,n+1} = \displaystyle\sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} \Delta_{h,i}^{k,n}. \end{cases}$$

    – Compute corrected state and parameter:

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1 \leq j \leq \mathrm{N}_s} (C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}} \Delta_{h,j}^{k,n+1} \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1 \leq j \leq \mathrm{N}_s} (C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}} \Sigma_{h,j}^{k,n+1} \end{cases}$$

  $\hookrightarrow$ Sampling step:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{\theta}_{h,j}^{k,n+1} = \hat{\theta}_h^{k,n+1} + \Delta_{h,j}^{k,n+1}, \\ \hat{z}_{h,j}^{k,n+1} = \hat{z}^{k,n+1} + \Sigma_{h,j}^{k,n+1}. \end{cases}$$

Additionally, for output purposes, we can re-define up to second-order terms the parameter covariance operator from the stored $\{\Delta_{h,j}^{k,n}\}_{j=1}^{N_s}$ and $\{\hat{z}_{h,j}^{n+1-}\}_{j=1}^{N_s}$:

$$\Lambda_h^{k,n} : \mathcal{P}_h \ni \theta \mapsto \sum_{1 \le j \le N_s} \beta_j (\Delta_{h,j}^{k,n}, \theta)_{\mathcal{P}} \Delta_{h,j}^{k,n} \in \mathcal{P}_h.$$

In the same fashion, the parameter sensitivity operator reads

$$L_h^{k,n+1} : \mathcal{P}_h \ni \theta \mapsto \sum_{1 \le j \le N_s} \beta_j ((\Lambda_h^{k,n})^{-1} \theta, \Delta_{h,j}^{k,n})_{\mathcal{P}} (\hat{z}_{h,j}^{k,n+1-} - \hat{z}_h^{k,n+1-}) \in \mathcal{Z}_h.$$

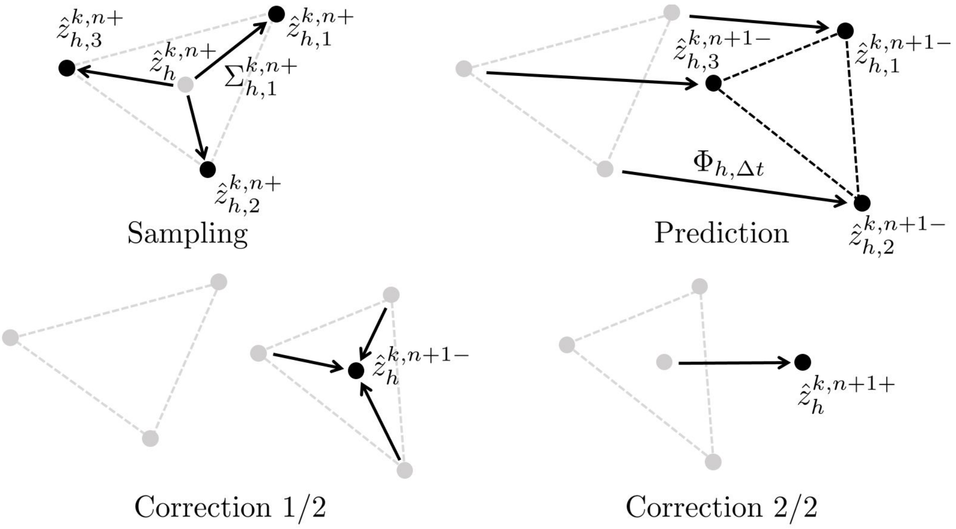


Figure 4.2 – Illustration of the main steps of the reduced-order UKF algorithm in the specific case of $\mathcal{P} = \mathbb{R}^2$ and focusing specifically on the state estimation part.

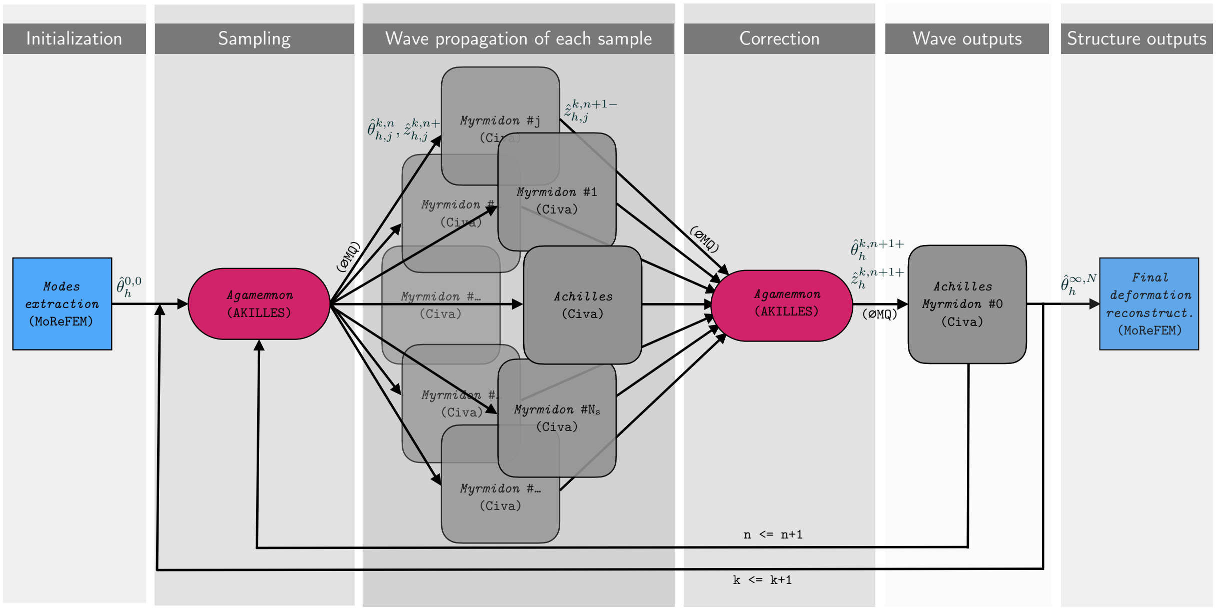


Figure 4.3 – Parallel implementation of the DT-iROUKF in AKILLES.

Finally let us give some implementation details, see also Figure 4.3. We implemented the DT-iROUKF algorithm in a specific Python library called AKILLES [1] based on the message passing library

---

[1] https://gitlab.inria.fr/AKILLES/AKILLES

`ØMQ`[2]. The objective of this library is to couple the estimator library with any model, written in any language that can receive and send vector quantities through `ØMQ`. In `AKILLES`, the Gramian, Cholesky decomposition, corrected states and parameters are computed by the main process of `AKILLES` called *Agamemnon* who communicates through `ØMQ` with *Achilles*' army of *Myrmidons* – here they represent wave solvers built upon the SFEM kernel of the commercial software `CIVA`[3]. The physical outputs are finally processed after the correction step by the process *Achilles*, a specific *Myrmidon* wave solver in `CIVA` in charge of extra output computations. Note that the non-linear mechanical deformation and the corresponding modal decomposition are initially computed using the Finite Element Library `MoReFEM`[4] for solving large deformation mechanical problems.

4.4 Numerical results

In this section, we illustrate the use of the presented algorithm for estimating the deformation caused by the mechanical loading in structures using GW measurements. Three different cases are presented in increasing order of complexity. The first case is a small-scale problem with the objective of testing the robustness and analyzing the estimation results of the proposed strategies. The second case is an aluminum plate under traction forces based on an experiment [Gandhi, Michaels, and Lee, 2012] done in the context of acoustoelasticity to retrieve the material third-order elastic constants. The third case is a pipe subjected to a 4-point bending test based on an experiment [Tschöke et al., 2017] related to the detection of welding defects under operational conditions employing ultrasound. All the illustrated cases are presented with the structure described below.

Noisy synthetic data generation. First, let us describe the numerical experimental setup for generating noisy synthetic data, and comprising mechanical loading, ultrasonic excitation and signal aquisition. A quasi-static structural problem is defined to compute the structural deformation caused by mechanical loading. For all the cases, the constitutive behavior of the material is modeled by the Compressible Neo Hookean (CNH) hyperelastic law. With λ and μ being the Lamé parameters, its hyperelastic potential reads

$$\mathscr{W}^{\text{CNH}} = \frac{\lambda}{2}(\sqrt{I_3} - 1)^2 + \frac{\mu}{2}(I_1 - 3 - \log(I_3)),$$

with $I_1 = \operatorname{tr}\boldsymbol{C}$ and $I_3 = \det\boldsymbol{C}$ being invariants of the right Cauchy-Green deformation tensor $\boldsymbol{C} = \boldsymbol{F}^\intercal\boldsymbol{F}$. The mesh, type of forces and boundary conditions configuring the quasi-static problem are given for every configurations considered in the following. After defining the configuration for the structural deformation problem, we define the configuration for the excitation, propagation and measurement of ultrasonic waves. A force acting upon the outer surface generates a wavefield radially with respect to the structure surface. It represents a force $\boldsymbol{f}$ defined in a sufficiently thin layer on the outer surface. Its ring geometry has an inner radius of 10mm and an outer radius of 20mm. The excitation signal is a 5-cycle cosine Hanning windowed at a specific frequency for each case. We consider zero initial conditions for the wavefield. The observation data are generated by using the operator in (4.10) with the domain ω_i defined upon the outer surface as a sufficiently thin layer in which the wavefield does no vary in the thickness direction. This amounts to modeling a point or surface probe with specific sensitivity. One may note that during the estimation procedure, as the estimated deformation changes while the wave propagation problem is running, the CFL condition (4.28) also changes. However, the time step is computed once at the initialization and this changes may cause numerical instability. To avoid such an issue, the time step is computed to satisfy the CFL condition for the expected deformation extrema.

To emulate signals obtained from an acquisition system in real scenarios, we generate synthetic observation data using the target deformation and the wave propagation solver. Gaussian noise is added

---

[2] `https://zeromq.org`

[3] `https://www.extende.com`

[4] `https://gitlab.inria.fr/MoReFEM`

to the simulated data to represent potential electronic and environmental noise. The Gaussian noise is added to the simulated measurements as

$$y^n_{\gamma,i} = \check{y}^n_i + \frac{\gamma}{\sqrt{\Delta t}} \|\check{y}_i\|_{\ell_2} \chi^n_i,$$

where χ^n_i is a realization of a random variable with distribution $\mathcal{N}(0,1)$. The noise level is set by adjusting γ.

Reducing the parametric space using mode decomposition. To represent the parametric space on a reduced basis, we use the eigenmodes of the quasi-static problem as defined in (4.11). First, we compute a finite set of eigenmodes $\mathcal{I}$ ordered in ascending eigenvalue order, *i.e.* from the lower to highest spatial frequency. A reduced parametric space for reconstructing the deformation is selected as a subset $\mathcal{I}^\star \subset \mathcal{I}$. This selection is made by having a guess $\boldsymbol{u}^\star_0$ of the deformation decomposed in the eigenbasis of $\mathcal{I}$ and selecting the smallest subset $\mathcal{I}^\star$ that satisfy a representation error criterion. More precisely, by decomposing $\boldsymbol{u}^\star_0$ in the eigenmodes of $\mathcal{I}$ we select a set of modes associated with the components $\theta^\star_i$ that have a minimum relative ℓ_2-error $\tau^\star$ defined as

$$\tau^\star = \frac{\|\boldsymbol{u}^\star_0 - \sum_{j \in \mathcal{I}^\star} \theta^\star_j \boldsymbol{\varphi}_j\|}{\|\boldsymbol{u}^\star_0\|}.$$

To better visualize the relevance of the selected modes we plot the representation error for the first m modes in $\mathcal{I}$, namely $1 - \tau_m$, where

$$\tau_m = \frac{\|\boldsymbol{u}^\star_0 - \sum_{j \in \mathcal{I}_m} \theta^\star_j \boldsymbol{\varphi}_j\|}{\|\boldsymbol{u}^\star_0\|}$$

and $\mathcal{I}_m$ being the set of the m lowest frequency modes in $\mathcal{I}$.

Initializing the estimator. To initialize the estimator, we must give an *a priori* parameter and its covariance. Even though one could use $\theta^\star$ as an *a priori* parameter value, we set $\hat{\theta}^{1,0}_h$ to zero in order to test the robustness of our estimator with respect to the initial guess. Regarding the initial covariance, we consider that $\max_{i \in \mathcal{I}^\star} \theta^\star_i$ is a relevant estimation of the standard deviation of the lowest frequencies, assuming that the lowest frequency is the most relevant for reconstruction. Then, the initial covariance Λ_0 is computed as, based on (4.14),

$$(\Lambda_0)_{ij} = \frac{\lambda_i^2}{\lambda_{\min}^2} (\max_{i \in \mathcal{I}^\star} \theta^\star_i)^2 \delta_{ij}, \quad \forall i, j \in \mathcal{I}^\star,$$

where the subscript min is associated with the lowest-frequency component in $\mathcal{I}^\star$. The constant M used to normalize the regularization term is computed as

$$M = \sum_{i \in \mathcal{I}^\star} \frac{\lambda^2_{\min}}{\lambda^2_i}.$$

Finally, the constant ε, which weights the regularization and misfit terms, is adjusted for each case. The parameter m appearing in the parametric space norm is set as $m = 1$.

Presentation of the estimation results. At each outer iteration k of our estimator, the last estimated parameter $\hat{\theta}^{k-1,N}_h$ is used as an initial parameter for the next one $\hat{\theta}^{k,0}_h$ while the initial covariance restarts at Λ_0. We plot for each case the evolution in $n \in [1 : N]$ of the estimated parameters $\hat{\theta}^{k,n}_h$ and a region representing an estimation of the associated standard deviation $\sqrt{(\Lambda^{k,n+1})_{ii}}$. Regarding the later, the initial covariance is propagated through the estimation process while retrieving information from the observations. It provides qualitative information about how the uncertainty evolves, although a quantitative interpretation of this uncertainty must be considered carefully.

Sensitivity analysis. Instead of using the subset $\mathcal{I}^\star$, the estimation could be done using a larger set of eigenmodes, $\mathcal{I}$ for instance, eliminating the need for $\boldsymbol{u}_0^\star$. However, the ill-posedness of the problem combined with the potential instability of the forward model, as the estimated pre-deformation varies in time, makes it unviable. To circumvent this difficulty, we assume to know a set $\mathcal{I}^\star$ of eigenmodes that best represents the target deformation. Nonetheless, one may like to gather $\mathcal{I}^\star$ directly from the measurements, instead of using $\boldsymbol{u}_0^\star$. As an attempt to do so, we perform a sensitivity analysis using one Levenberg-Marquardt iteration of the proposed estimator and the set $\mathcal{I}$. As we run the estimator in the whole set $\mathcal{I}$, we must avoid instability by constraining the trajectory $\hat{\theta}_h^{1,n}$. This can be done by setting the covariance as

$$(\Lambda_0)_{ii} = \gamma^2 = 10^{-8}$$

with a parameter $\varepsilon = 1$, initially. Although at the end of the iteration the estimated $\hat{\theta}_h^{1,N}$ have no quantitative significance, it still retains qualitative information relating the parametric space to the observations. To visualize and assess this relation, we plot the normalized estimated parameters at the end of the estimation,

$$\frac{\hat{\theta}_{h,i}^{1,N}}{\max_{i\in\mathcal{I}} \hat{\theta}_{h,i}^{1,N}}, \qquad \forall i \in \mathcal{I},$$

as well as the estimated parameters $\hat{\theta}_{h,i}^{1,n}$ over time. Additionally to the estimated values, the covariance matrix is also an output of the estimation process and can be used for this purpose. We plot its normalized inverse,

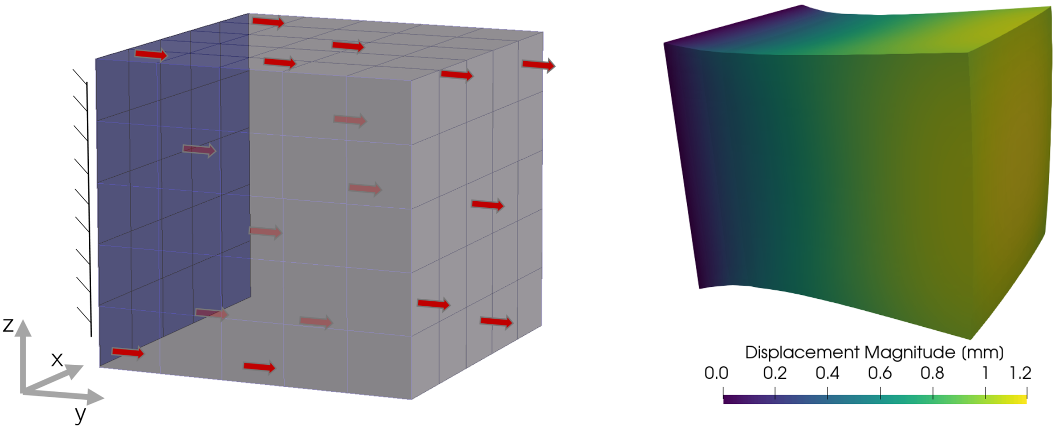
$$(\Lambda_0^{-\frac{1}{2}}) \left((\Lambda^{0,N})^{-1} - \Lambda_0^{-1} \right) (\Lambda_0^{-\frac{1}{2}}),$$

representing the normalized Gramian matrix, related to the system's observability [Moore, 1981].

4.4.1 Results and discussion

4.4.1.1 A test case on robustness to noise

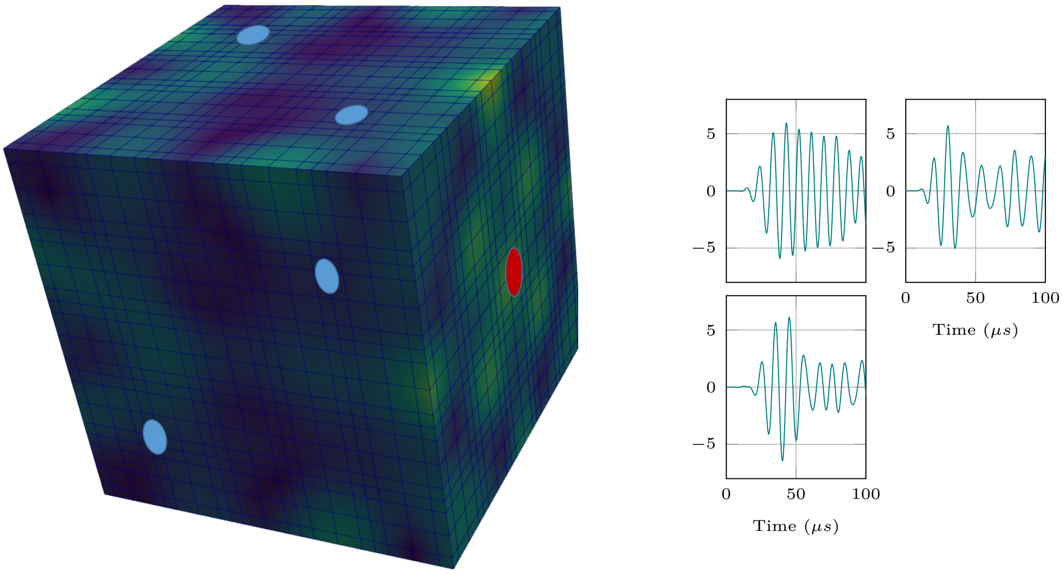
To illustrate and test the estimator robustness to noise, we model an aluminum cube of dimensions $60 \times 60 \times 60$ mm^3. The configuration for the quasi-static problem is shown in Figure 4.4a. The target deformation shown in Figure 4.4b is computed by solving the quasi-static problem for a body force of 50N/mm^3 in the Y direction.



(a) Configuration of forces the quasi-static problem. (b) Target deformation (visualization scaled 10x).

Figure 4.4 – Quasi-static problem configuration for the illustration on the cube.

From zero initial conditions, the wave field is excited at one of its faces at 100kHz in a ring region as described previously. The configuration for the wave propagation model is depicted in Figure 4.5a, where the characteristic Gauss-Lobatto points of the SFEM discretization are represented. Examples of extracted signals are plotted in Figure 4.5b. In this case, the observations are point measurements at ten positions, where the three components of the displacement field are obtained. Considering the definition of the observation operator (4.10), ω_i is a sufficiently small volume for which the displacement field is considered constant. The three components are taken as sensitivity $\boldsymbol{d}_i$, resulting in $d = 30$, the number of measured signals. Each face, except the excited one, has two measurement points on the surface at a distance of 15 mm from the edges. We run the wave propagation solver up to 100μs to generate the synthetic observation data. In Figure 4.5b we show the extracted field components for the point located at (15,15,0).



(a) Configuration of the wave propagation problem. An example of the wave field is illustrated.

(b) Example of measured signals (three components) at (15,15,0).

Figure 4.5 – Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

To test the effect of noise in the estimation, different sets of observed data are constructed with different levels of noise $\gamma = [10^{-4}, 0.01, 0.1, 0.2]$. Examples of the different levels of added noise are shown in Figure 4.6.

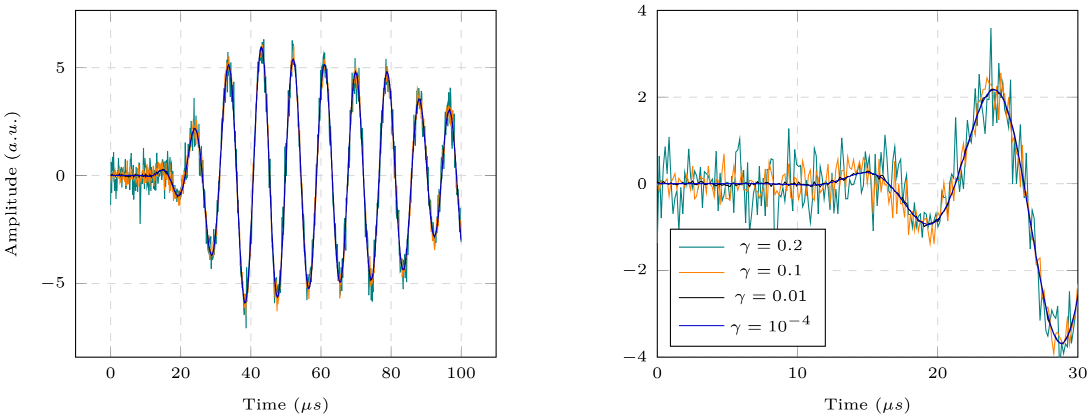
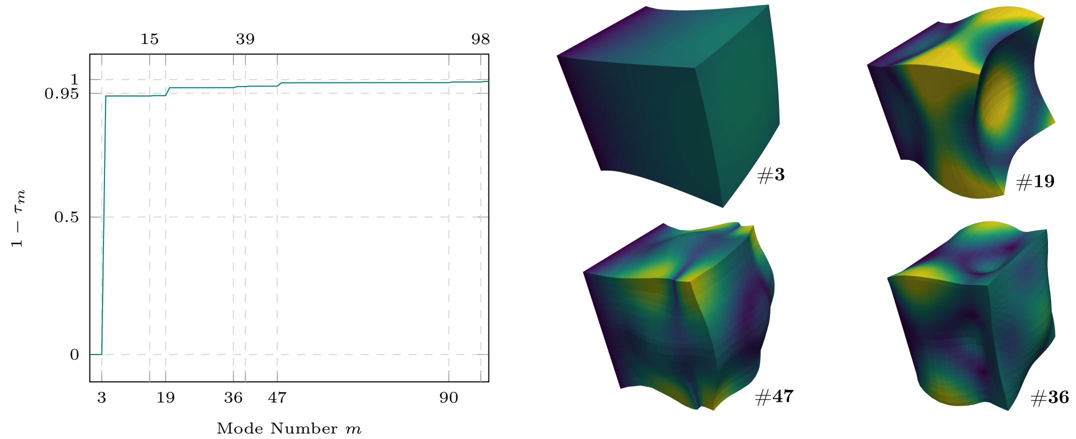


Figure 4.6 – Example of different noise levels added to the synthetic data for the cube case.

The eigenmodes associated with the hundred lowest eigenvalues are computed, embodying $\mathcal{I}$. We decompose a guess of the deformation generated by an axial body force with low amplitude, $\boldsymbol{u}_0^\star$, in the basis associated with $\mathcal{I}$. Using this decomposition we plot $1 - \tau_m$ and the associated relevant modes in Figure 4.7. Then, the most relevant modes can be selected and are therefore highlighted, making, for $\tau^\star = 0.01$, $\mathcal{I}^\star = \{3, 15, 19, 36, 39, 47, 90, 98\}$, meaning that with 8 selected modes we can reconstruct 99% of its deformation.



(a) Reconstruction error per number of modes.

(b) Deformation associated with the relevant eigenmodes.

Figure 4.7 – Reconstruction of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

We run the estimation algorithm for the different noise levels γ. We first use $\varepsilon = 1$ as the regularization and misfit terms in function (4.16) are normalized. The iterative process is shown in Figure 4.8 with the time evolution of the estimated parameters $\hat{\theta}_h^k$. To illustrate the evolution of the estimated deformation, we reconstruct it at given iterations and time and compare it with the target deformation in Figure 4.9, for the case with $\gamma = 0.2$.

Discussion. At relatively low noise levels ($\gamma \leq 0.01$), the estimate converges in one iteration, and at each subsequent iteration, the estimate deviates slightly from the converged value and converges again. At higher noise levels ($\gamma \geq 0.1$), a few iterations may be required to converge with significant errors *w.r.t.* the target, although it presents a good estimation of the reconstructed deformation, as shown in Figure 4.9. In addition, the estimator updates the estimated parameter at a lower rate as γ increases, as we can see in the case of $\gamma = 0.2$. This is due to the fact that γ weights the mismatch term, as can be seen in (4.16), and gives less credibility to the observed data when noise increases. In such a case, the parameter ε can be adjusted to increase the rate at which the estimator updates the parameter based on the observed data. For example, we set $\varepsilon = 0.025$ and show the corresponding estimate in Figure 4.10. We can see that when the update rate is increased, the noise causes larger instabilities in the estimated parameter during an iteration. The variable ε can be adjusted on a case-by-case basis. As discussed earlier, the tangent stiffness operator in (4.6) can be ill-posed, *i.e.* the coercivity assumption (4.8) for the forward problem is not guaranteed. Therefore, ε must be adjusted taking into account the direct-problem stability issues due to strong oscillations during the estimation.

As mentioned earlier, we can restrict the estimation to a reduced parametric space without resorting to an initial deformation prior, but by retrieving information from the observations. For this case, we evaluate the possibility of doing this through a sensitivity analysis using a low noise level, $\gamma = 10^{-4}$, and $\varepsilon = 1$. The normalized estimated parameters are shown in Figure 4.11a and the quantity $\mathcal{I}^\star$ is highlighted. Based on these results, the most important modes cannot be completely distinguished, but one could empirically set a threshold to select them. At the risk of overlooking some important modes, an iterative selection process can be proposed but it is not discussed here. The evolution of these parameters during estimation is shown in Figure 4.11b, with the $\mathcal{I}^\star$ set highlighted. The main modes do not appear to exhibit any particular dynamic behavior that could be used to distinguish them. Finally, Figure 4.11c plots the Gramian for the 50 first modes in $\mathcal{I}$. The Gramian quantifies how observable and distinguishable the different modes are from each other and information about the overall observability of the inverse problem can be retrieved. High observability can lead to overestimation, as in the case of the higher frequency modes, for example, mode 47. Inversely, low observability can lead to underestimation, as in the case of mode 3. The diagonal values quantifies the observability of the associated modes while off-diagonal terms quantifies how indistinguishable the associated modes are from each other.

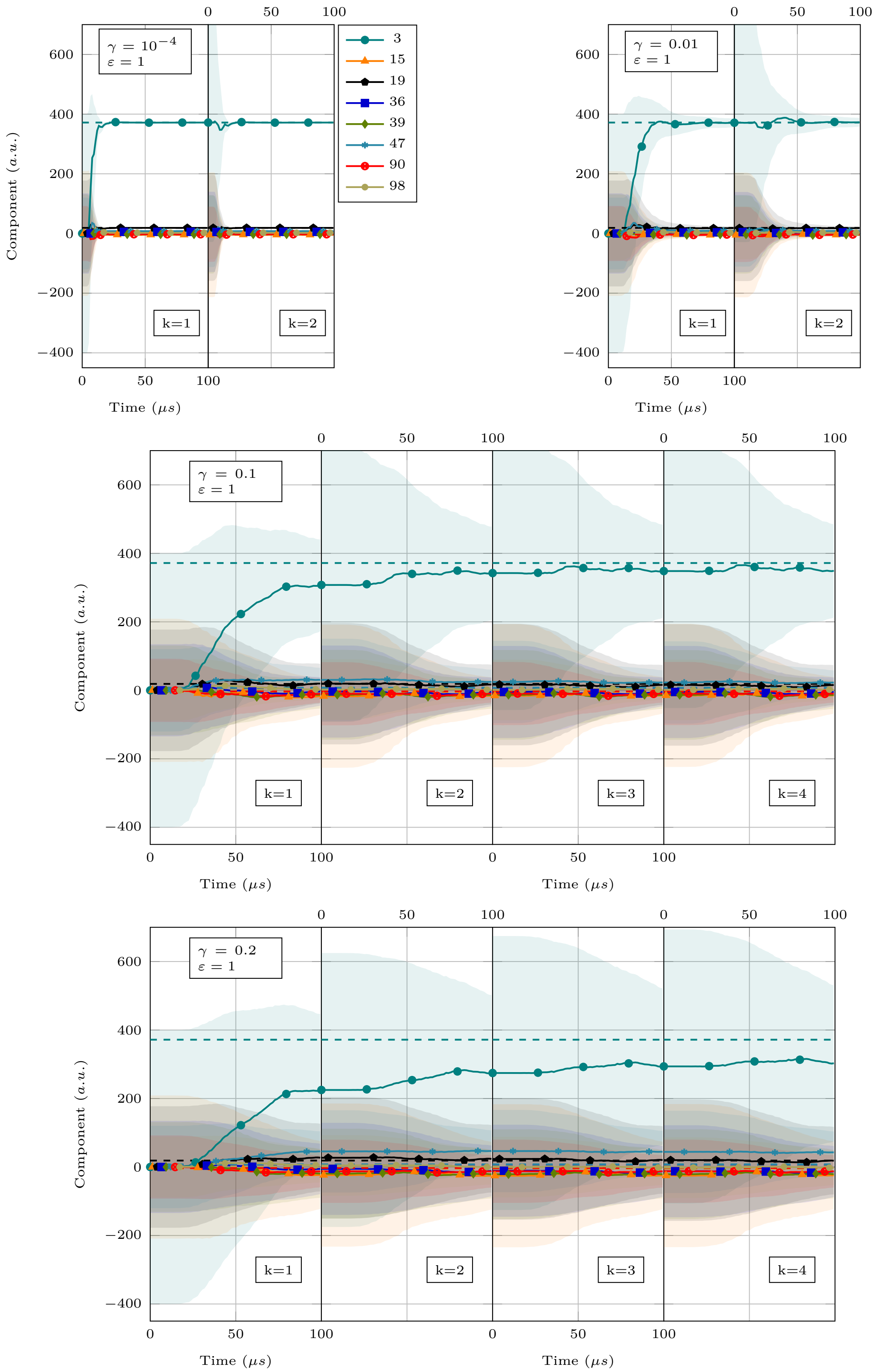


Figure 4.8 – Evolution in time for the estimation $\hat{\theta}_h^k$ for the cube case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a regularization $\varepsilon = 1$.

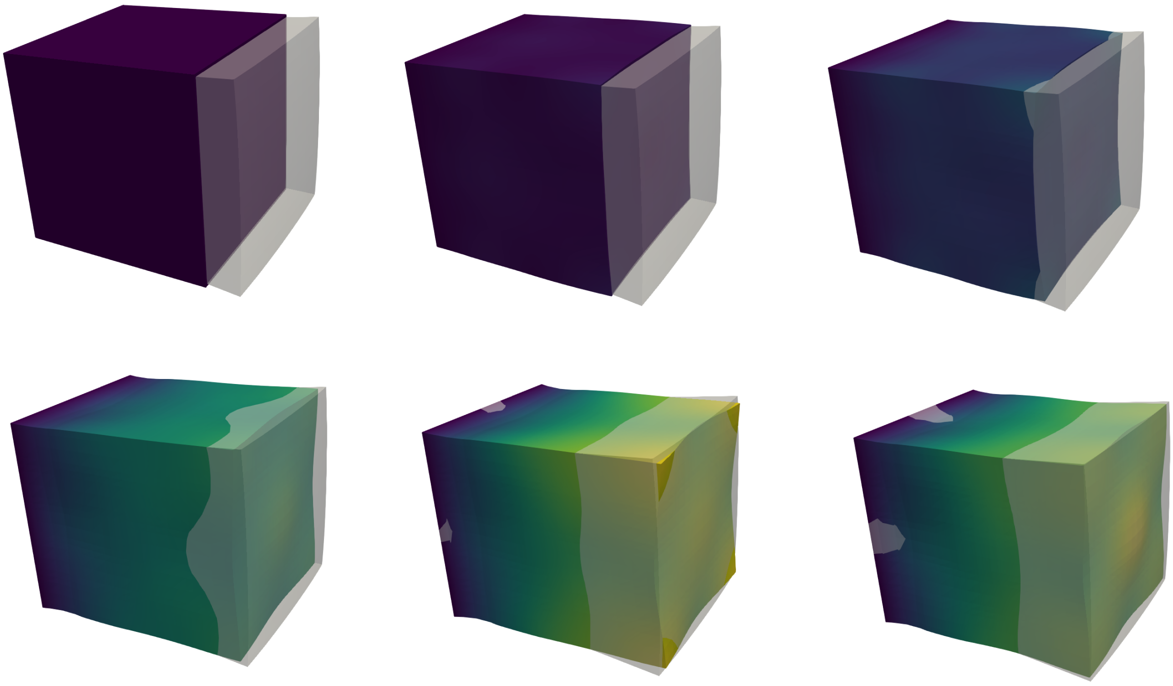


Figure 4.9 – Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent) for a high level of noise, $\gamma = 0.2$. From left to right, top to bottom we have (k=0, t=0μs) (k=0, t=25μs) (k=0, t=35μs) (k=1, t=25μs) (k=2, t=75μs) and (k=3, t=100μs).

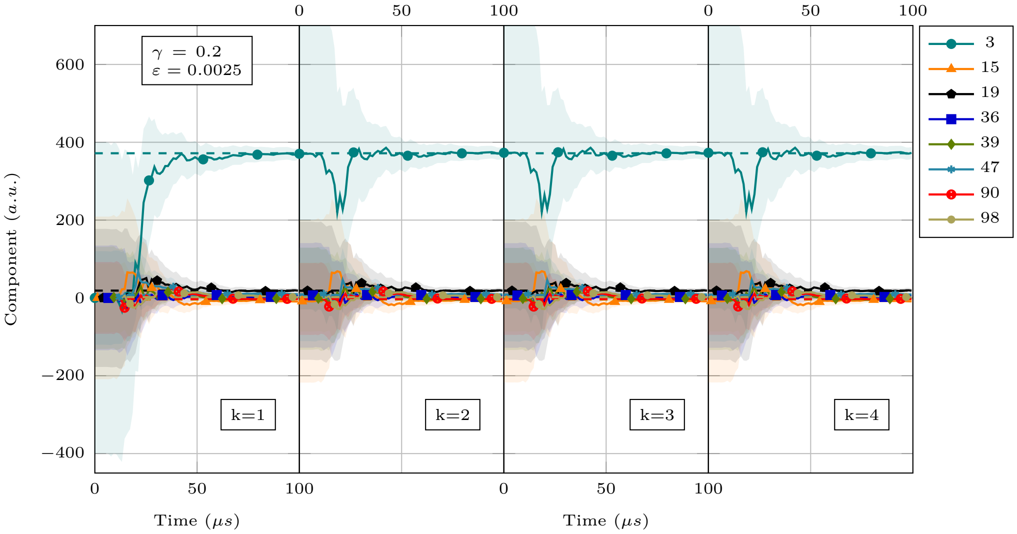
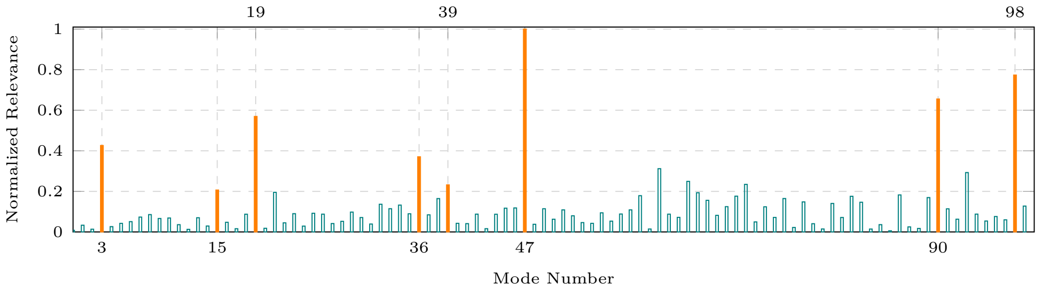
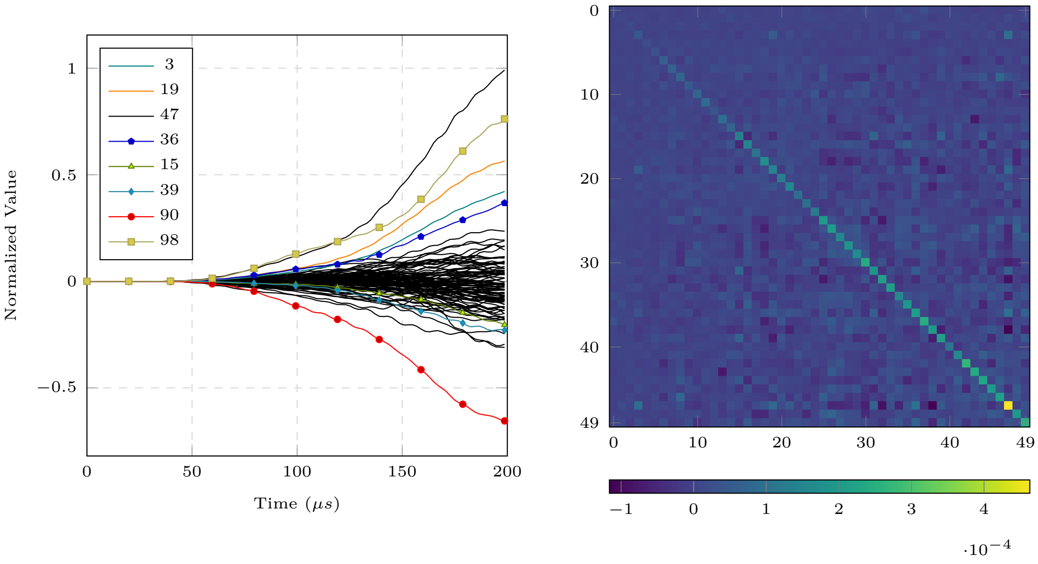


Figure 4.10 – Evolution in time for the estimation $\hat{\theta}_h^k$ for the cube case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. A weak regularization, $\varepsilon = 0.0025$, and noise $\gamma = 0.2$ is used.



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.



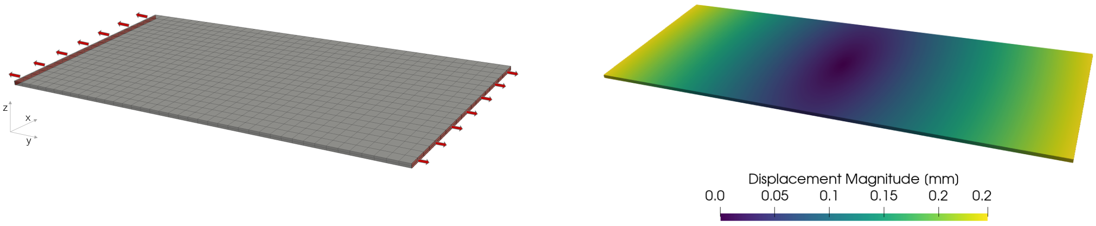
(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$.

(c) Normalized Gramian at the end of the analysis.

Figure 4.11 – The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the cube illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted.

4.4.1.2 Aluminum plate under traction

We present here a case based on the experiment done in [Gandhi, Michaels, and Lee, 2012] to model guided wave propagation in axially loaded structures. The specimen is an aluminum plate of dimensions $610 \times 305 \times 6.35$ mm^3 under traction forces. The configuration for the quasi-static problem is shown in Figure 4.12a. We compute the target deformation with an axial stress of 57.5 MPa, resulting in the deformation illustrated in Figure 4.12b.

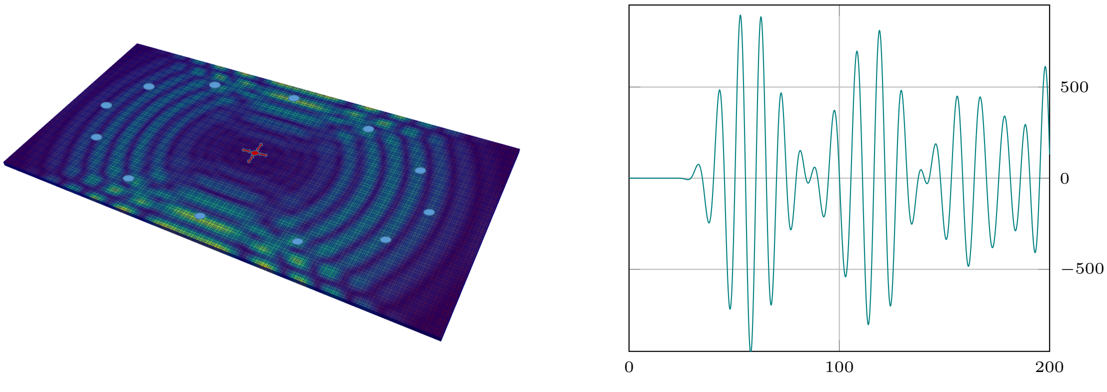


(a) Configuration.

(b) Deformation for an axial stress of 57.5MPa.

Figure 4.12 – Quasi static problem configuration and target deformation for the plate illustration.

The ultrasonic excitation is done as previously described at the center of the plate (upper surface) at 100kHz. The observation data are obtained by defining the observation operator in twelve surface regions ω_i distributed in an ellipse as depicted in Figure 4.13a. The mesh used in the wave propagation problem is also depicted. The synthetic observed data are generated by running the simulation up to $200\mu s$. An example of the obtained signal is shown in Figure 4.13b.

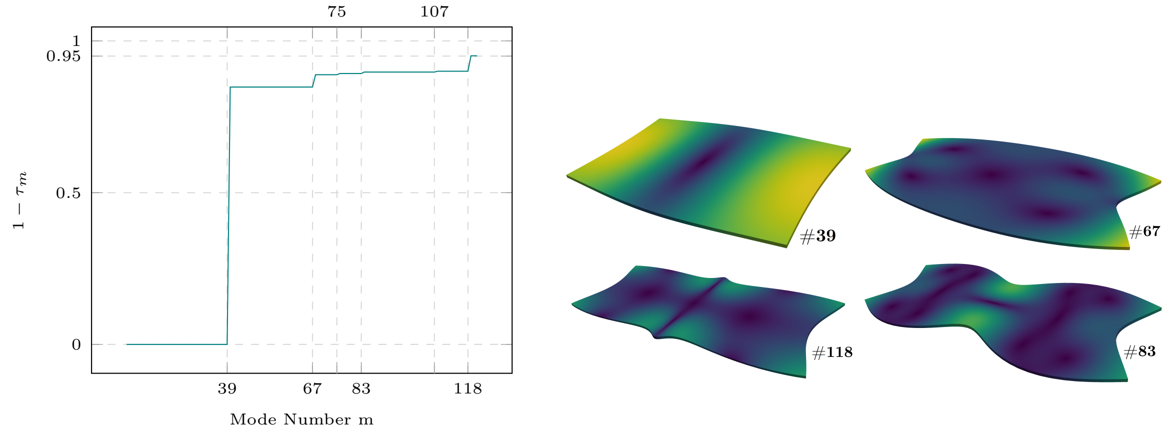


(a) Wave propagation problem configuration, mesh, excitation and measurement regions.

(b) Example of measure signals.

Figure 4.13 – Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

For the larger set of eigenmodes $\mathcal{I}$ we compute those associated with the 120 lowest eigenvalues. As the experiment has no essential boundary conditions, rigid body movements are penalized and the first 6 eigenmodes, associated to them, are removed. From a guess of the deformation $\boldsymbol{u}_0^\star$ generated by the traction surface forces with lower amplitude, we can select a smaller set of eigenmodes by analyzing its decomposition on a parametric basis. In Figure 4.14 we plot $1 - \tau_m$ from which the most relevant modes can be selected. For $\tau^\star = 0.05$, we have $\mathcal{I}^\star = \{39, 67, 75, 83, 107, 118\}$.



(a) Reconstruction error per number of modes. (b) Deformation associated with the relevant eigenmodes.

Figure 4.14 – Reconstruction of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

After acquiring the observed data from the simulated measurements on the target deformation, we add different levels of noise, $\gamma = [0.01, 0.1]$, as plotted in Figure 4.15.

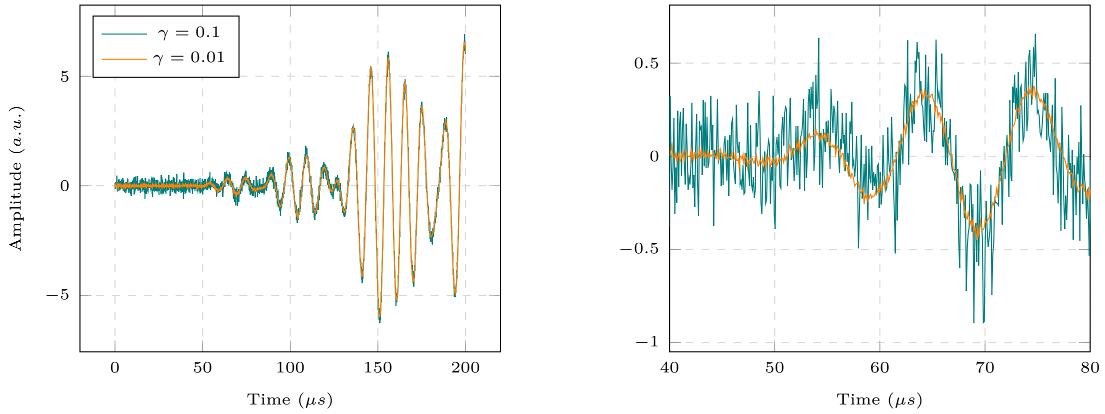


Figure 4.15 – Example of different noise levels for the aluminum plate illustration.

Discussion. As in the previous case of the cube, a few specific modes are important to the reconstruction of the deformation although the first relevant mode is of relatively high frequency. This must be taken into account when computing $\mathcal{I}$ as we risk not including the relevant modes. The estimation shown in Figure 4.16 required more iterations to achieve apparent convergence, even with the lowest noise. As noise increases, the differentiation between modes becomes more difficult although the most important mode, 39, remains pronounced. For visualization purposes, the estimated deformation is reconstructed at different times of the estimation and depicted in Figure 4.17. As previously done, by adjusting ε we can achieve a faster response of the estimate by reducing the number of iterations needed until apparent convergence, as shown in the estimation for this adjusted case, Figure 4.18.

Regarding performance, the model for the quasi-static problem has 7750 degrees of freedom (DoFs) and the computation of the first 120 eigenmodes takes 9 minutes. Finite elements of order one is used for the quasi-static problem as numerical locking is negligible in pure extensional deformation. The wave

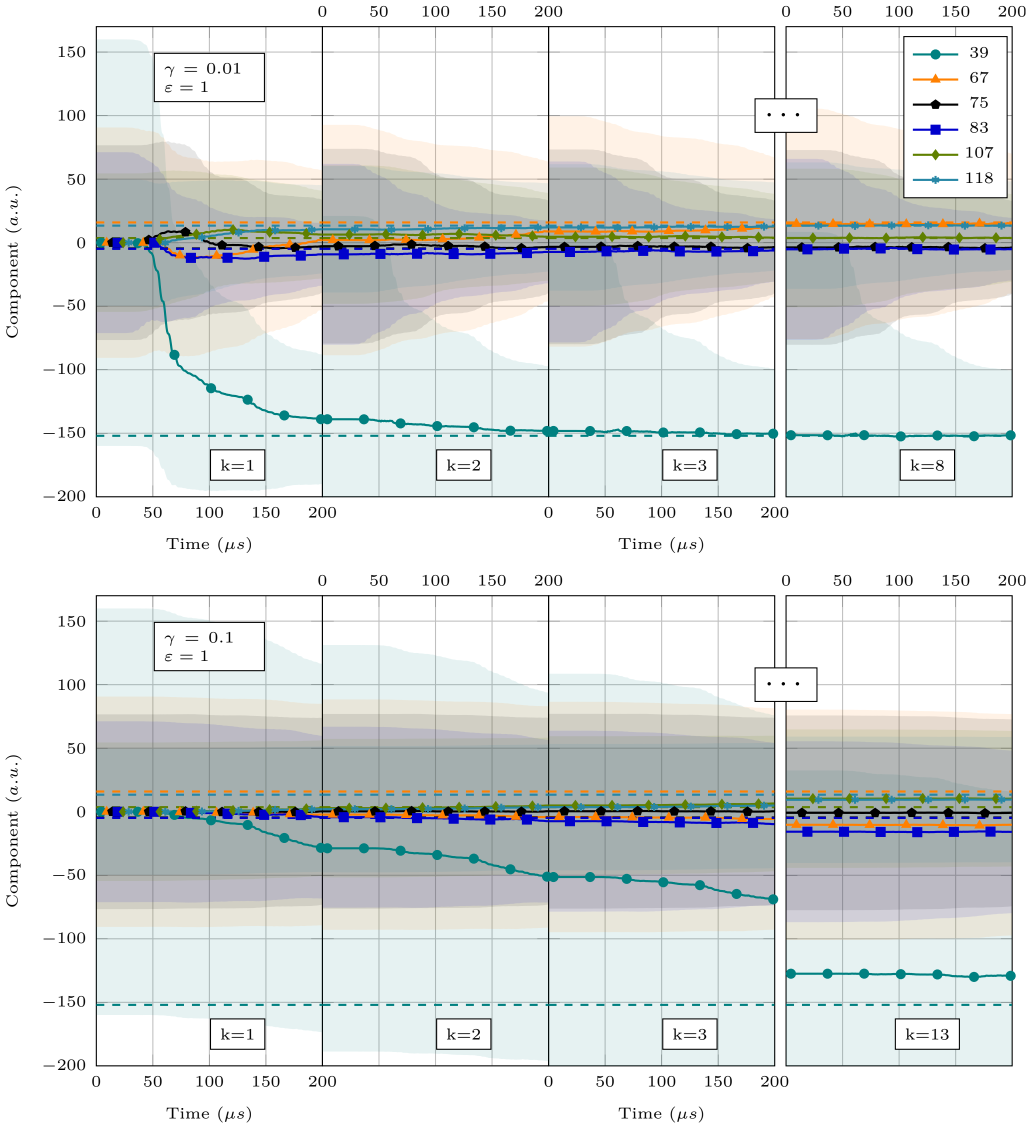


Figure 4.16 – Evolution in time for the estimation $\hat{\theta}_h^k$ for the plate case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a neutral regularization $\varepsilon = 1$.

propagation problem has 420k DoFs and a total of 2223 time steps. One run of the wave propagation problem to generate the observations takes less than a minute and requires 180MB of RAM. During an estimation of 6 mode components, 7 wave propagation problems are run in parallel and each iteration of the estimation procedure takes 12 minutes. The computational cost of solving a wave propagation problem is higher during estimation due to parameter change at each time step, requiring extra computations related to the update of the tangent stiffness operator. We use a laptop computer equipped with a *Intel i9-9880H* CPU and 32GB of RAM for computing the presented results.

Finally, the sensitivity analysis is done for a negligible noise $\gamma = 10^{-4}$ and $\varepsilon = 1$. The final estimated values for the sensitivity analysis are plotted in Figure 4.19a. In this case, the relevant modes are well distinguishable. The observability can be analyzed in the Gramian matrix, Figure 4.19c. Due to its sparsity, we conclude that there are only a few observable modes and the fact that the diagonal is not pronounced means that the modes are less indistinguishable from each other. This can be seen in the results with the reduced set of eigenmodes, where the estimation of less relevant modes are highly affected by noise. The high observability shown in Figure 4.19a is due to the fact that the deformation of the

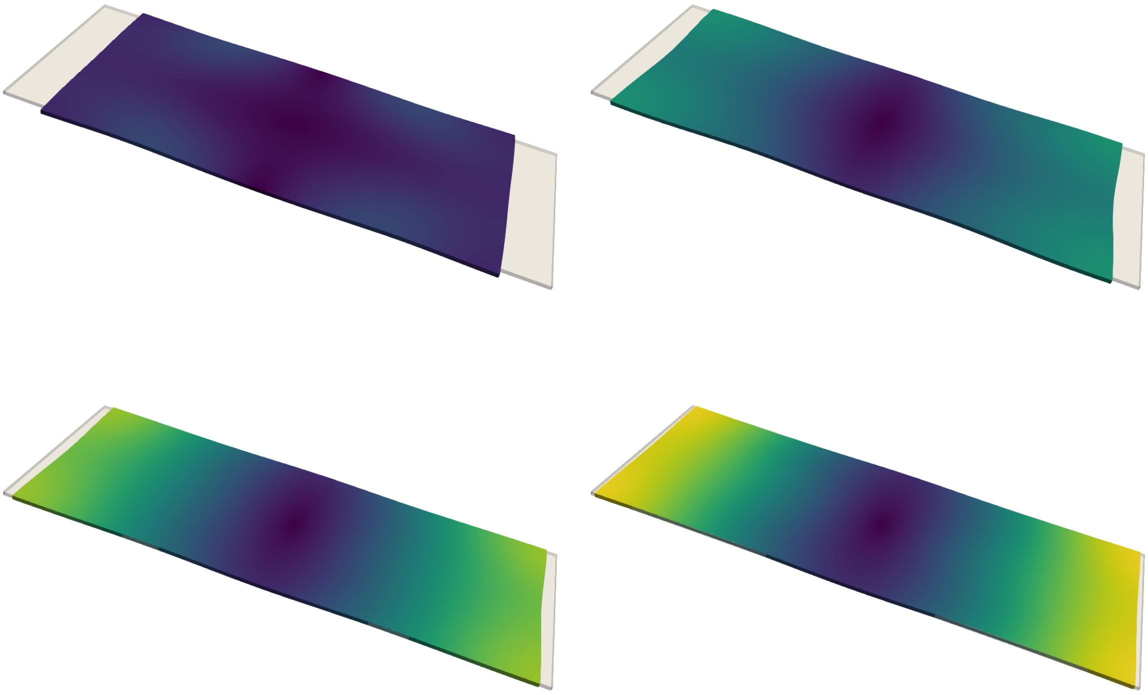


Figure 4.17 – Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent). Estimation for $\gamma = 0.1$. From left to right, top to bottom we have (k=0, t=60μs) (k=0, t=130μs) (k=0, t=180μs) and (k=1, t=200μs). Visualization is scaled 300x.

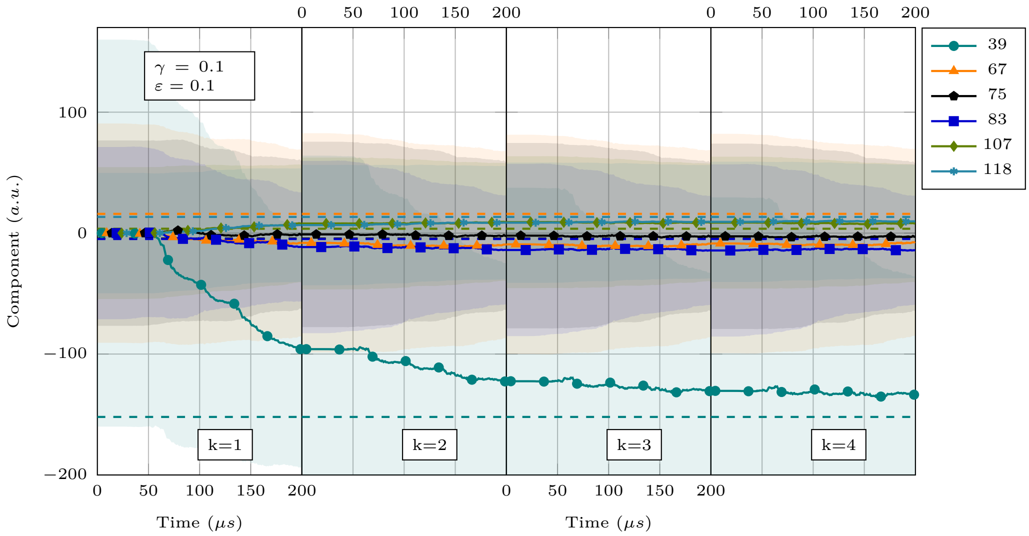
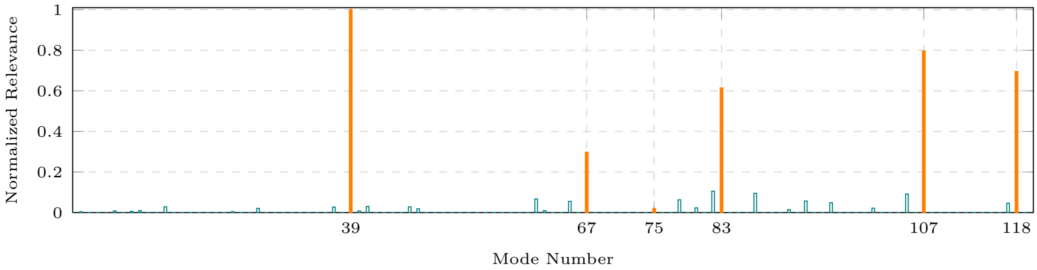


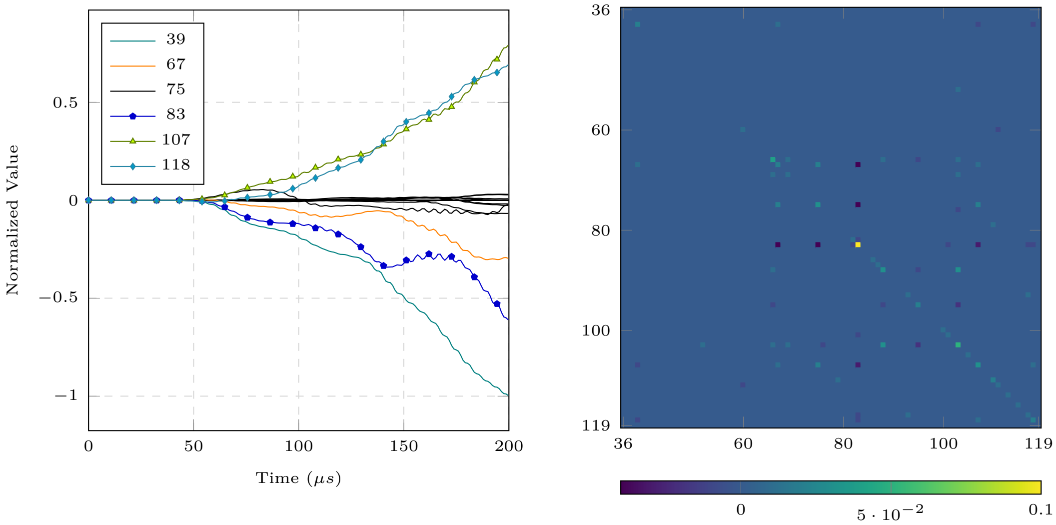
Figure 4.18 – Evolution in time for the estimation $\hat{\theta}_h^k$ for the plate case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. A weaker regularization, $\varepsilon = 0.1$, is used.

structure is of relatively high frequency , hence more observable according to the Gramian. Also, the relevant modes are not clustered around a certain frequency, which helps their spatial differentiation. The evolution of the parameters shown in Figure 4.19b can be analyzed and provide some insights into the process of estimation. Modes 67 and 87 present non-monotonic behavior since the deformation related to these modes is concentrated farther from the center where the waves were excited (see Figure 4.14), hence the wave takes longer to travel through these regions and its carried information to be considered.

Regarding performance of the sensitivity analysis, 115 wave propagation problems had to be run in parallel, taking 5 hours to complete and requiring 271GB of RAM, in total. The memory usage for each problem was significantly higher than when estimating only 6 modes as each problem had to store the deformation gradient of all 114 modes. Other strategies for accessing such gradients can be proposed, leaving room for improvement. A desktop workstation equipped with 2**Intel Xeon Platinum 8276* CPU and 512GB of RAM was used to run these problems. Note that, although we use only one workstation, our implementation with `ØMQ` allows parallelization through workstations and clusters.



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.



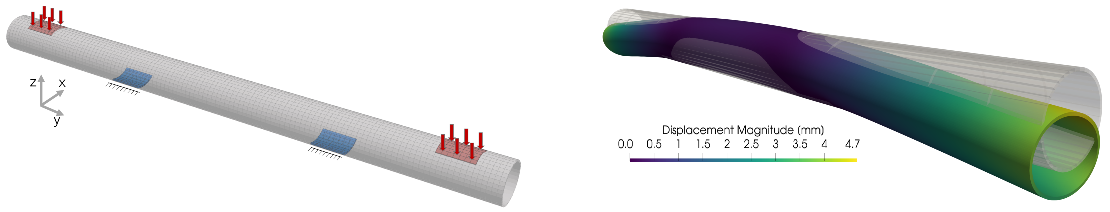
(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$.

(c) Normalized Gramian at the end of the analysis.

Figure 4.19 – The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the plate illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted..

4.4.1.3 Experiment-based 4-point bending on a steel pipe

To increase complexity, we consider the experiment done in [Tschöke et al., 2017], where the authors use ultrasonic guided waves to detect weld damage in a steel pipe while it was subjected to a 4-point bending test. They assess the performance of traditional imaging techniques to detect the damage while the specimen is subjected to mechanical load. To overcome the bias introduced by the mechanical load to the ultrasonic signal, they perform signal acquisitions in pristine conditions for different loads and use them as baselines to, through comparison, detect changes in the signal due to potential damage. The baseline should be acquired for every expected load condition, which requires experimental manipulation and reduces the robustness of a detection system. An alternative to acquiring several baselines is to estimate the deformation of the structure using the already available signals. For that purpose, we illustrate the use of our method in this more complex and application-related configuration. The modeled pipe has a length of 2.94 meters, a diameter of 0.1973m and is 8mm thick. The quasi-static problem is depicted in Figure 4.20a. For a total applied force of 220kN, the target deformation is depicted in Figure 4.20b.

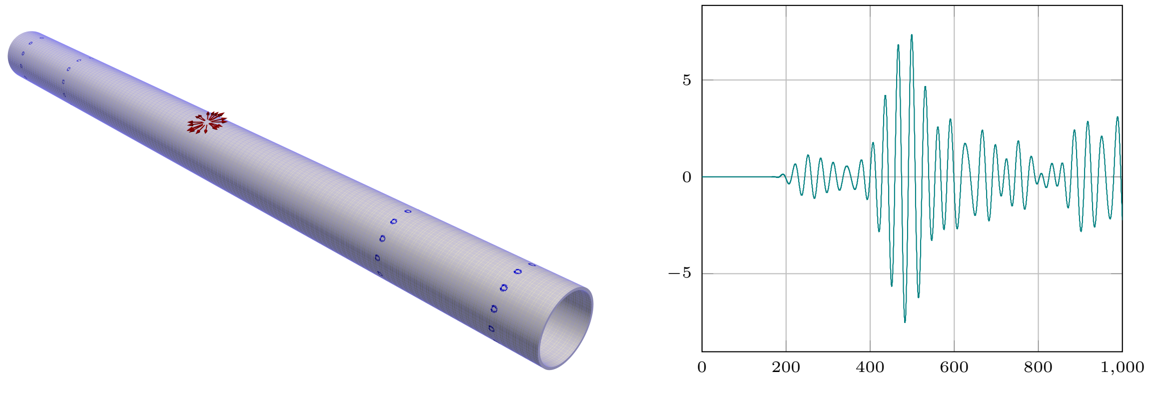


(a) Quasi static problem configuration.

(b) Target Deformation (visualization 20x).

Figure 4.20 – Quasi static problem configuration and target deformation for the pipe illustration.

The excitation is done at the top outer-upper surface of the steel pipe at 30kHz, the observations are obtained from applying the observation operator, as previously described, with radial sensitivity in 48 regions distributed in 4 circular evenly distributed sections as depicted in Figure 4.21a. Each section of transducers is rotated 15 degrees with respect to each other. The synthetic observed data are obtained from simulating the wave propagation problem up to $1000\mu s$. An example of the obtained signal is shown in Figure 4.21b.



(a) Wave propagation problem configuration, mesh, excitation and measurement regions.

(b) Example of measured signal.

Figure 4.21 – Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

The set $\mathcal{I}$ is embodied by the eigenmodes associated with the 60 lowest eigenvalues. From a guess of the deformation $\boldsymbol{u}_0^\star$ generated by the bending surface forces, we can analyze its decomposition in the eigenbasis. In Figure 4.22 we plot $1 - \tau_m$ from which the most relevant modes can be selected and are therefore highlighted, making, for $\tau^\star = 0.05$, $\mathcal{I}^\star = \{0, 1, 2, 4, 6, 7, 8, 9\}$. Differently from the previous cases, the relevant modes are concentrated in the lowest frequencies and are poorly distributed.

Due to the stability issues associated with the potential non-coercivity of the wave propagation problem, the estimation has to be adjusted with $\varepsilon = 10^4$, setting a relatively low update rate for the estimation, hence avoiding instabilities during estimation. The evolution of the parameter can be seen in Figure 4.23 for two different levels of noise, $\gamma = 10^{-4}$ and $\gamma = 0.1$. For both levels of noise, the estimation succeeds in differentiating the modes. The estimated deformation are reconstructed at different steps of

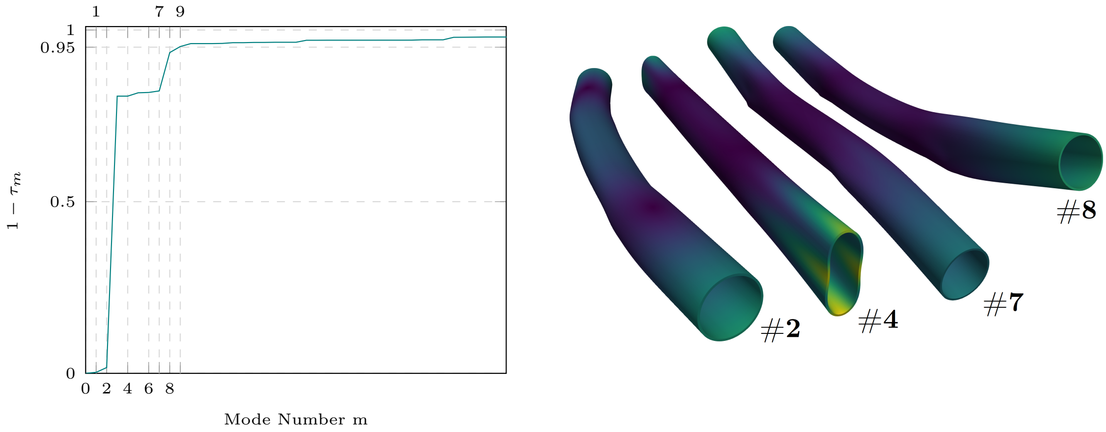


Figure 4.22 – Reconstruction error of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

the estimation procedure and is depicted in Figure 4.24.

Discussion. Using high regularization and a large amount of observed data, the estimate is stable and shows good convergence to the target in a few iterations, even when noise increases. By comparison with the plate case, we believe that this is because sensors are well positioned and the acquisition time is long enough so that this case benefits from rather strong observability, allowing it to well distinguish the deformation modes. This illustrates the importance of a robust and well-designed acquisition configuration for a correct estimation.

Regarding performance of the estimation, The model for the quasi-static problem had 148k DoFs and the computation of the first 60 eigenmodes takes 8 hours to complete using a *Intel(R) Xeon(R) W-3245* CPU. Finite elements of order two is used for the quasi-static problem. The wave propagation problem has $2M$ DoFs and a total of 4662 time steps. One run of the wave propagation problem to generate the observations takes 5 minutes and requires 1GB of RAM. During an estimation of 8 mode components, 9 wave propagation problems are computed in parallel and each iteration of the estimation procedure takes 2 hours. A workstation equipped with a *Intel i9-9880H* CPU and 32GB of RAM is used for the estimation.

Note, however, that while we obtain satisfactory results when the estimation is performed in a reduced set of modes, the sensitivity analysis performs poorly when it comes to selecting the relevant modes from a larger set, as can be seen in Figure 4.25a. The high-frequency modes are overestimated because they are more observable, as seen in the Gramian (Figure 4.25c). This overestimation of the less relevant modes due to observability is a cause of instability when trying to estimate the deformation using the larger set of modes. The evolution of the sensitivity analysis in Figure 4.25b shows that some irrelevant modes compensate for each other when trying to minimize the misfit because they are less distinguishable. Without the potential instability issue of the wave propagation problem, the Gramian diagonal indicates that the modes are readily observable. The Gramian diagonal can also be used to group the modes that conflict with each other in minimizing the mismatch, allowing selection of distinguishable modes and adaptive estimation. This is typical of ill-posed inverse problems, where estimation is efficient once we accept a regularization that here comes from a reconstruction in a parameter space with a small dimension.

Regarding performance of the sensitivity analysis, 61 wave propagation problems are computed in parallel, taking 5 hours to complete and using 400GB of RAM, in total. A desktop workstation equipped with 2×*Intel Xeon Platinum 8276* and 512GB of RAM was used. The same remarks done for the plate case regarding performance are valid here.

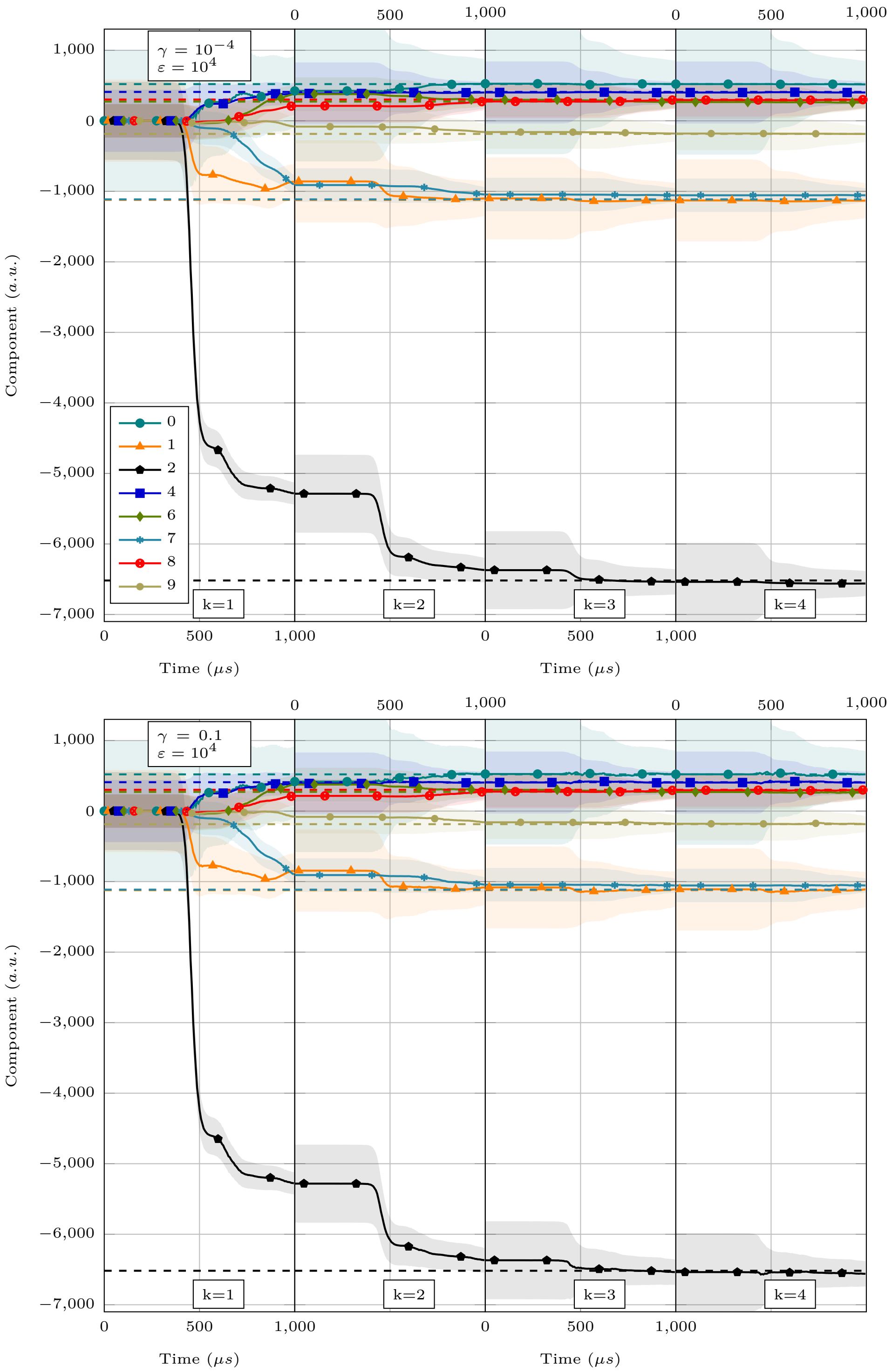


Figure 4.23 – Evolution in time for the estimation $\hat{\theta}_h^k$ for the pipe case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a strong regularization $\varepsilon = 10^4$.

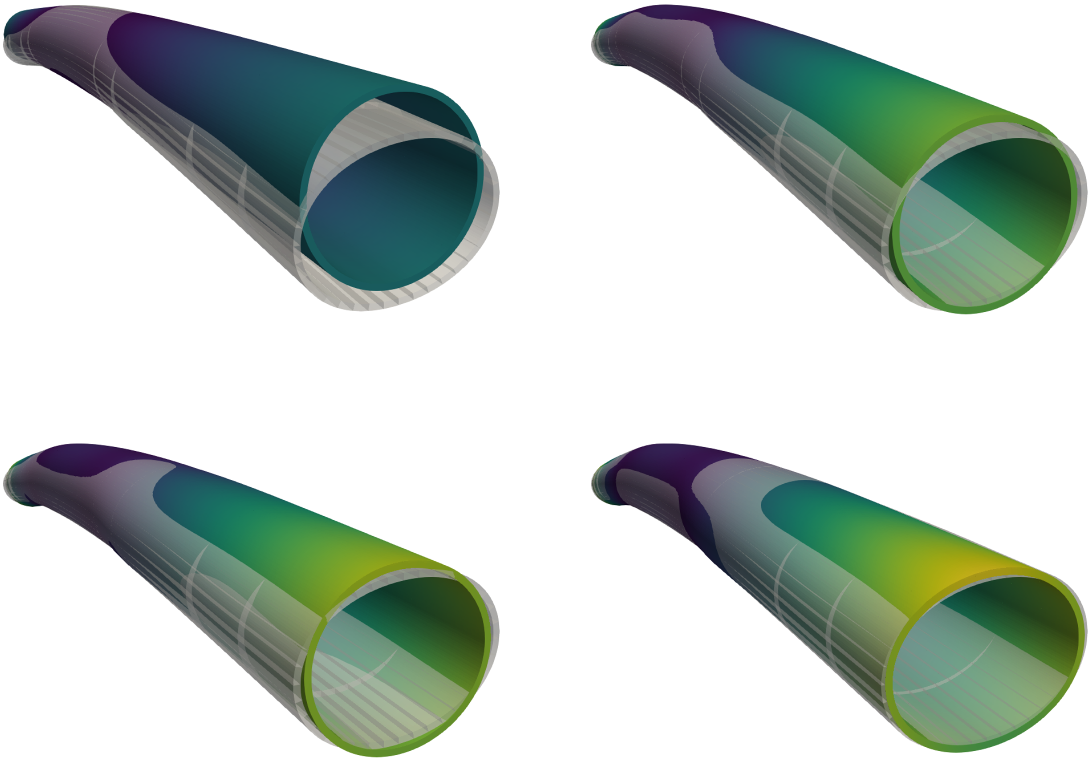
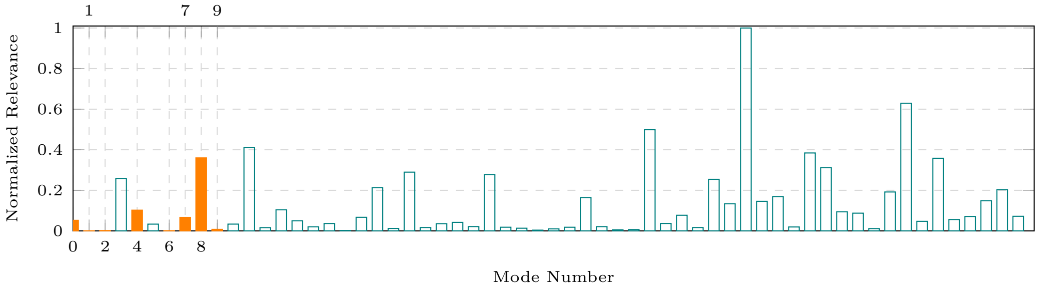
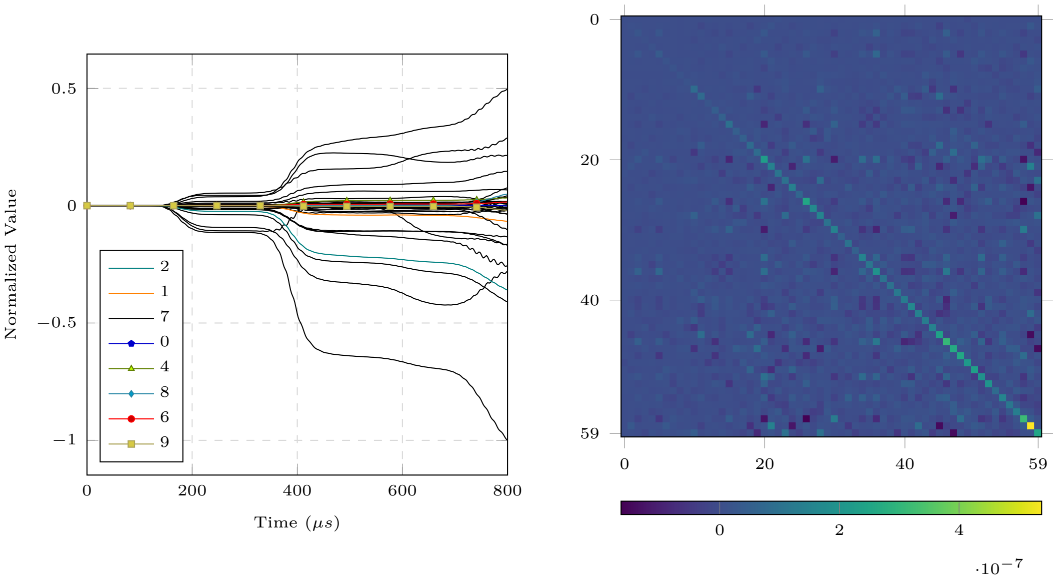


Figure 4.24 – Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent). Estimation for $\gamma = 0.01$. From left to right, top to bottom we have (k=0, t=60μs) (k=0, t=130μs) (k=0, t=180μs) and (k=1, t=200μs). Visualization is scaled 20x.



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.



(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$.

(c) Normalized Gramian at the end of the analysis.

Figure 4.25 – The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the pipe illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted.

4.5 Conclusion and perspectives

In this paper, we show how to reconstruct in nondestructive tests a baseline free of environmental loading conditions using available SHM measurements. Our model-based approach is inspired by Full-Wave Inversion (FWI) strategies, but here we take advantage of dynamic programming principles to avoid multiple iterations of adjoint minimization. In fact, we still rely on an iterative strategy, the Levenberg-Marquardt algorithm, to transform the initial nonlinear inverse problem into successive linear-quadratic estimation problems. Each of these linear-quadratic estimation problems is then solved using a Kalman-based approach. By relying on the Unscented-Kalman filter, we also avoid the computation of the tangent operator required by the Levenberg-Marquardt algorithm. The final algorithm converges in a limited number of iterations compared to standard FWI and is therefore mostly sequential and inherently parallel with respect to the sigma-points covering the reconstruction space. As a result, our algorithm efficiency remains comparable to solving the direct guided wave propagation problem. The main limitation of our approach is that, due to followed dynamic programming point of view, we are limited to a low-dimensional representation of the estimated deformation, i.e., only about a hundred modes can be estimated. Moreover, and this will be the goal of future developments, our approach first considers the reconstruction of the displacement field, while from the identification perspective, it might be better suited to reconstruct a strain tensor, for instance, the Cauchy-Green strain or its invariants. In particular, we believe that this research direction will help us to better control some deformation constraints related to a potentially ill-posed wave propagation problem that limits the trust region neighborhood of our current method. Moreover, our method now needs to be extended to more general observational operators, some of which may depend on the estimated deformation, making the presented approach more complex. Moreover, our approach suffers from the same limitation as the classical FWI method in terms of the choice of the discrepancy measure between the model and the data. Recently, [Engquist, Froese, and Yang, 2016; Engquist, Ren, and Yang, 2020] has been shown that the use of Wasserstein-based distances can overcome cycle-skipping effects in wave field inversions.

Aknowledgment and fundings

The authors would like to deeply thank Laurent Steff for his implementation of the `AKILLES` Library and Jérôme Diaz for his help with the `MoReFEM` Library. The authors would also like to thank the Isaac Newton Institute for Mathematical Sciences for support (EPSRC grant number EP/R014604/1) and hospitality during the programme “The mathematical and statistical foundation of future data-driven engineering” when part of this work on this paper was undertaken. This research was funded by the following project: “GW4SHM”(gw4shm.eu) project from the European Union’s Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie, grant number 860104.

Conclusion and perspectives

Overview

In the first part of this thesis, we treat numerical methods for acoustoelasticity. With a *small-on-large* deformation assumption, the nonlinear elastodynamics formulation was linearized around a deformed state of the structure originating from mechanical loading. This resulted in two decomposed problems: the quasi-static problem, related to the structural deformation, and the linearized wave propagation problem. Both problems are solved in the Total Lagrangian Formulation, facilitating the definition of boundary conditions such as transducer positioning. Using Murnaghan's hyperelastic law, we link the presented formulation with a commonly used formulation in the context of acoustoelasticity. The quasi-static problem is solved using 3D Shell Finite Elements, allowing any hyperelastic law to be used without additional assumptions while mitigating numerical locking in thin structures. To solve the linearized wave propagation problem we use the high-order spectral finite elements method, which showed through numerical illustrations to be an efficient choice for guided wave propagation due to negligible numerical dispersion and anisotropy. Such a combination of numerical methods allows it to model arbitrary geometries, loading configurations, and constitutive behavior based on hyperelastic laws. We assess the potential instability of the model due to coercivity assumptions related to the hyperelastic potential by presenting and analyzing unstable cases. Finally, we model realistic cases for applications in SHM with complex loading configurations and validate our numerical methods, from the structural deformation computation to the linearized solver, using experimental data.

In the second part of the thesis, we aim to apply the proposed direct model in the context of inversion. First, we present a framework and notations for inverse problems associated with wave propagation. In this framework, the inverse problem is interpreted as a least squares minimization problem. For the sake of illustration, we introduce generic problems of interest related to estimating initial conditions and parameters. Some traditional variational methods and sequential methods are detailed and discussed. To assess the different presented methodologies, we apply these methods to one-dimensional wave propagation problems and compare their different advantages and performances. For exemplification, a linear and a nonlinear problem were implemented and discussed. In the last chapter, we present our original data assimilation strategy to estimate the deformation of the structure under loading conditions by means of limited ultrasonic data. The structural deformation is represented in a reduced parametric space based on the eigen decomposition associated with the operator of the quasi-static problem, effectively reducing the amount of information needed to represent the deformation to be estimated. Combining variational and sequential approaches, our strategy results in an iterated Reduced Order Unscented Kalman Filter algorithm and a direct link can be made with the Levenberg-Marquardt method. Moreover, from the sequential method, we obtain a time-varying estimation of the parameters and their respective uncertainty. We apply our approach to realistic cases with different configurations for transducers, including a 4-point bending test steel pipe, while showing robustness to different levels of noise.

Concluding discussion

To our knowledge, no solver currently available in the literature is able to, at hand, compute guided wave propagation in structures under arbitrary loading conditions with corresponding genericity. Although semi-analytical implementations are fast and can be used with arbitrary geometry and complex stress profiles, they are limited to their underlying analytical assumptions. Such methodologies are mainly able to compute dispersion curves, *i.e.* frequency-dependent velocity and phase of guided waves. When using them for inverse problems, they require post-processing methods to properly isolate guided modes in the signal. This post-processing may result in loss of information and add uncertainties, issues that become more relevant with higher frequencies when the amount of propagating modes increases significantly. Moreover, high-frequency modes are particularly of interest in our context due to their sensitivity to stresses [Pei and Bond, 2016]. Our direct model, by solving the elastodynamic problem, computes a wavefield including complex wave interactions such as diffraction and reflection by complicated obstacles and boundaries. Additionally, in terms of acoustoelastic effects, mode conversion due to stress gradients is also represented. Moreover, as any hyperelastic law can be easily implemented, our solver is capable of modeling wave propagation in a wide variety of materials in loading conditions, such as the illustrated Carbon Fiber-Reinforced Polymer (CFRP) stratified plate.

The direct numerical solver is placed in the context of commercial software for NDT and SHM-related simulations, `CIVA`, which allows this work to be integrated into a rich environment for numerical tools in such applications. It includes the efficient implementation of the spectral finite elements method, modules for meshing, boundary conditions definitions, and output for visualization in dedicated software, among others. Additionally, with the presented solver, several useful outputs can be made for the analysis of wave propagation in specific loading configurations. For instance, the Kelvin-Christoffel tensor can be computed at each degree of freedom, giving an insight into bulk wave polarization and velocities for any propagation direction in the whole structure. By evaluating the tensor for different directions of propagation, one can plot "local" slowness curves for bulk waves. Moreover, as we use the linearized model for an arbitrary hyperelastic law, we have the 9×9 elasticity tensor "C_{ijkl}" at each degree of freedom. The elasticity tensor can be used as input to semi-analytical models to compute "local" dispersion curves in complex stress profiles. Where, by "local" we mean localized in the structure, as their values may vary spatially.

Due to the modeling capabilities of our transient finite elements approach, raw ultrasonic signal can extracted from the wavefield by modeling the characteristics of a transducer, for instance, its sensitivity to displacements on the surface it is bonded. In particular, the output signal can be compared directly with measurements, which allowed us to use a performant and robust data assimilation strategy analogous to the *Full-Waveform Inversion*. Moreover, by using a wave model-based assimilation able to model complex wave interactions, more of the measured data is expected to be useful as these interactions tend to be increasingly present with longer acquisitions. This is possible assuming that an accurate description of the geometry and boundaries is given.

Our data assimilation strategy method was used for estimating the structural deformation using limited ultrasonic waves. This can be particularly useful in the context of SHM as it can be done with the already-available acquired data and, from such estimation procedure, one can remove the bias in this data caused by loading conditions. Furthermore, by solving the transient elastodynamic problem, the multi-modal and dispersive aspects of guided waves do not limit the acquisition setup and do not require specific post-processing of the signal for mode distinction. For instance, in a case where the acquisition sensor is close to obstacles, the acquired signal would contain complicated interference from reflected and/or diffracted waves, difficulting in the identification of the first arrival, hence, the measure of wave speed and phase. While a semi-analytical strategy would require extra effort, it comes with no additional cost when using wave-based modeling.

An important aspect of our proposed data assimilation strategy is that one can retrieve insights on the observability of the system as the Gramian and its evolution through the estimation process is given at no extra cost. As we show, one can assess which estimated parameters are more observable and if their estimation interfere with each other. Also, the sensitivity of a given parameter to the signal obtained by different sensors can be extracted. These analyses are useful for enhancing and understanding the experimental design. Using a strategy based on filtering, the estimation was demonstrated to be robust to different levels of noise with the possibility of adjustment of the regularization. One can increase the regularization with the trade-off of decreasing convergence rate, while decreasing the regularization results in faster convergence but may cause instability.

Perspectives

The two main objectives described in the introduction of this work were tackled and satisfactory results in realistic scenarios for SHM applications were obtained. Nevertheless, the development of such a project also gives us insights into the directions of newer studies and developments that could be done.

Among those that can be envisaged by the reader, we find it important to pursue the following problems:

- The stability of the model cannot be ensured for any hyperelastic law and any deformation due to the potential failure to satisfy the coercivity properties of the tangent operator, and due to buckling issues. This issue is known in mechanical modeling [Ogden, 1984; Ndanou, Favrie, and Gavrilyuk, 2014]. By proposing new hyperelastic potentials, one can mitigate such instability problems. In particular, some works dealing with hyperelastic laws in the framework of a polyconvexity analysis could be used in that matter [Schröder and Neff, 2003; Itskov and Aksel, 2004].
- The data assimilation strategy could be modified to other types of estimation. For instance, the calibration of hyperelastic laws can be done using ultrasonic raw data with few adjustments in our model. For this purpose, our method would significantly reduce uncertainties when compared to traditional pitch and catch approaches that use wave velocity to calibrate constitutive models [Muir, Michaels, and Michaels, 2009]. Additionally, the same procedures could be adjusted to optimize sensor placement, for better observing the pertinent unknown.
- Our data assimilation strategy is limited to reduced parametric spaces. In this specific implementation we reduce them by selecting the lowest frequency modes of the structural dynamics, *i.e.* around 100 modes. Other approaches can be proposed for this selection of modes as well as an adaptative selection of those between estimation iterations, as in [De Buhan and Kray, 2013]. Moreover, other representations of the deformation can be used instead of the displacement field, for instance, the strain tensor or its invariants. A different representation could help to better penalize unlikely deformations during the estimation process and therefore help to keep it stable. Furthermore, other constructions of the parametric space could be proposed to estimate, for instance, localized residual stresses or concentration of stresses in certain regions by using level-set methods [Aghasi, Kilmer, and Miller, 2011], among others.
- In our implementation of the sequential approach, model errors were not considered. These errors may be present in addition to noise, hence determining the relevancy of such errors can be fundamental for a reliable implementation in pair with SHM systems. Existing traditional techniques are used for this purpose.
- A mathematical analysis concerning observability/identifiability of the proposed data assimilation strategy can be done, for instance, using Carleman methods as in [Baudouin et al., 2021] and in references within.
- When working with experimental data of axially loaded CFPR plates from Airbus, isolating a wavepacket was not straightforward due to the experimental setup that included reflections. Retrieving guided mode velocity with the accuracy required to capture acoustoelastic effects was not possible with the available information. Using a suitable transversely isotropic constitutive law and our inverse framework, we could first propose to calibrate the constitutive law parameters by using the raw signals at zero loading, and then being able to refine the third-order parameters using the signals in the loaded configuration.
- Here, we use a transversely isotropic law that does not include the third-order elastic constants. First, a corresponding constitutive law can be proposed, including these constants. This allows a useful modeling of composite materials, where experimental data can be used to calibrate the model. Then, as aforementioned, our algorithm can be used for such calibration with slight changes.

From a SHM application point of view, although our inverse strategy is not suitable for real-time application due to computational time, our validated and generic method can be used as a reference to reduced models or fast data-driven models in complex scenarios. Also, performance assessment and analysis of SHM systems in loading conditions can be done.

Appendices

Appendix A

Properties and Definitions

Definitions and properties used throughout this work are presented here. A brief description of the concepts of functional analysis used is also given.

A.1 Tensors

In this section we define real-valued tensors and associated useful properties. Here, we use tensors to represent multilinear operations between elements of finite-dimensional vector spaces. Let $\boldsymbol{A}$ be a third-order tensor in $\mathbb{R}^2$, it is defined as

$$\boldsymbol{A} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} A_{ijk}\, \widehat{e}_i \otimes \widehat{e}_j \otimes \widehat{e}_k,$$

where $\{\widehat{e}_i\}_{i=1}^{2}$ is a basis of $\mathbb{R}^2$. For a given basis, the tensor $\boldsymbol{A}$ can be represented by its scalar components $\{A_{ijk}\}_{i,j,k=1}^{2}$. This definition can be extended to a n-th-order tensor in $\mathbb{R}^d$. The operation $\otimes$ is the tensor product. The tensor product generates, from two n-th-order tensor a $(n \times n)$-th-order tensor. The standard summation of repeated indexes may be used, such that

$$\boldsymbol{A} = A_{ijk}\, \widehat{e}_i \otimes \widehat{e}_j \otimes \widehat{e}_k.$$

A.1.1 Definitions

Definition A.1.1. ***Tensor products.*** *Considering* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *arbitrary tensors in* $\mathbb{R}^2$ *as*

$$\boldsymbol{A} = A_{ij}\widehat{e}_i \otimes \widehat{e}_j \quad \text{and} \quad \boldsymbol{B} = B_{ij}\widehat{e}_i \otimes \widehat{e}_j,$$

we have the following tensor products:

$$\begin{aligned}(\boldsymbol{A} \otimes \boldsymbol{B})_{ijkl} &= A_{ij}B_{kl},\\ (\boldsymbol{A} \overline{\otimes} \boldsymbol{B})_{ijkl} &= A_{ik}B_{jl},\\ (\boldsymbol{A} \underline{\otimes} \boldsymbol{B})_{ijkl} &= A_{il}B_{jk}.\end{aligned}$$

Definition A.1.2. ***Tensor symmetries.*** *Considering a fourth-order tensor* $\boldsymbol{C}$*, it is said to have* ***major*** *symmetry if*

$$C_{ijkl} = C_{klij},$$

and ***minor*** *symmetries if*

$$C_{ijkl} = C_{jikl} = C_{ijlk}.$$

Definition A.1.3. *Let* $\boldsymbol{a}$ *and* $\widetilde{\boldsymbol{a}}$ *be vector fields and* $\boldsymbol{A}$, $\boldsymbol{B}$ *be tensors, the following chain-rule for tensor derivative can be written:*

$$\mathrm{D}_{\boldsymbol{a}}\, \boldsymbol{A}[\boldsymbol{B}(\boldsymbol{a})](\widetilde{\boldsymbol{a}}) = \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{B}} : \mathrm{D}_{\boldsymbol{a}}\, \boldsymbol{B}[\boldsymbol{a}](\widetilde{\boldsymbol{a}}),$$

where $\dfrac{\partial \boldsymbol{A}}{\partial \boldsymbol{B}}$ *is a unique fourth-order tensor for this operation.*

A.1.2 Properties

Property A.1.1. ***Tensor contraction (symmetric).*** *Considering a second-order tensor* $\boldsymbol{A}$ *and a* *fourth-order tensor* $\boldsymbol{C}$ *with minor symmetries, we have*

$$\boldsymbol{C} : \boldsymbol{A} = \boldsymbol{C} : \frac{1}{2}(\boldsymbol{A}^\intercal + \boldsymbol{A})$$

Proof. The contraction in indicial notation is

$$\frac{1}{2}(C_{ijkl}A_{lk} + C_{ijkl}A_{kl}),$$

By using the minor symmetry, $C_{ijkl} = C_{ijlk}$, we have

$$\frac{1}{2}(C_{ijlk}A_{lk} + C_{ijkl}A_{kl}),$$

As the sum is done in repeated indices,

$$\sum_{l,k} C_{ijlk}A_{lk} = \sum_{k,l} C_{ijkl}A_{kl},$$

yielding

$$\frac{1}{2}(C_{ijkl}A_{lk} + C_{ijkl}A_{kl}) = C_{ijkl}A_{kl}.$$

□

Property A.1.2. *Given* $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}$ *and* $\boldsymbol{E}$ *as second-order tensors,* $\boldsymbol{C}$ *as a fourth-order tensor with the symmetries*

$$C_{ijkl} = C_{jikl}, \quad C_{ijkl} = C_{ijlk} \qquad \textit{and} \qquad C_{ijkl} = C_{klij}$$

the term

$$\boldsymbol{A}^\intercal \boldsymbol{B} : \boldsymbol{C} : \boldsymbol{A}^\intercal \boldsymbol{D} + \boldsymbol{E} : \boldsymbol{D}^\intercal \boldsymbol{B}$$

can be written as

$$\boldsymbol{B} : \widetilde{\boldsymbol{C}} : \boldsymbol{D} \qquad \textit{with} \qquad \widetilde{C}_{ijkl} = A_{im}C_{mjnl}A_{kn} + E_{lj}\delta_{ik}.$$

Proof. We start by writing the terms using Einstein notation

$$A_{mi}B_{mj}C_{ijkl}A_{nk}D_{nl} + E_{ij}D_{mi}B_{mj}.$$

Rearranging the components and permuting index $m \leftrightarrow i$, $n \leftrightarrow k$ we have

$$B_{ij}\Big(A_{im}A_{kn}C_{mjnl}\Big)D_{kl} + B_{ij}E_{mj}D_{im}.$$

As $\boldsymbol{C}$ has minor symmetries, the first term can be written as

$$\boldsymbol{B} : \boldsymbol{A}\,\boldsymbol{C}\,\boldsymbol{A}^\intercal : \boldsymbol{D}.$$

We aim now to write the term $B_{ij}E_{mj}D_{mi}$ as two double contractions analogue to the first term, so

$$B_{ij}E_{mj}D_{im} = B_{ij}E_{kj}D_{ik} = B_{ij}(E_{lj}\delta_{ik})_{ijkl}D_{kl},$$

obtaining then the intended relation. □

A.2 Functional Analysis

In this appendix, we deal with **real-valued finite-dimensional Hilbert spaces** and we present here the used notations, definitions and properties related to them. Particularly we will consider function spaces. A Hilbert space is defined as a complete metric vector space in which distance between elements are defined by its inner product.

A.2.1 Definitions

Definition A.2.1. $\mathrm{H}^k(\Omega)$ ***spaces.*** *A functional space denoted* $\mathrm{H}^k(\Omega)$ *for a given domain* Ω *is a Hilbert space of functions defined as*

$$\mathrm{H}^k(\Omega) = \left\{ f \in \mathrm{L}^2(\Omega) : \forall |\alpha| \leq k, \partial^\alpha f \in \mathrm{L}^2(\Omega) \right\},$$

where $\mathrm{L}^2(\Omega)$ *is the space of functions where,* $\forall f \in \mathrm{L}^2(\Omega)$

$$\left(\int_\Omega |f|^2 \mathrm{d}\Omega \right)^{\frac{1}{2}} < \infty.$$

Definition A.2.2. ***Operator.*** *An operator is defined by its domain and image, i.e. the space of its inputs and the space of its outputs, respectively. Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be Hilbert spaces, we define a function* f *with domain* $\mathcal{X}$ *and image* $\mathcal{Y}$ *as*

$$\begin{aligned} f \ : \ \mathcal{X} &\longmapsto \mathcal{Y} \\ x &\longrightarrow f(x) = y, \end{aligned}$$

for $x \in \mathcal{X}$ *and* $y \in \mathcal{Y}$.

Definition A.2.3. ***Inner product and norm.*** *Let* $\mathcal{X}$ *be a Hilbert space, we denote its inner product and norm as*

$$(x, \widetilde{x})_\mathcal{X} \in \mathbb{R} \quad \text{and} \quad \|x\|_\mathcal{X}^2 = (x, x)_\mathcal{X} \quad \forall x, \widetilde{x} \in \mathcal{X},$$

respectively.

Definition A.2.4. ***Linear operators.*** *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be Hilbert spaces, the space* $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ *is defined as the space of linear operators of* $\mathcal{Y}$ *with values in* $\mathcal{X}$*, such that,* $\forall A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$

$$A(x + \widetilde{x}) = Ax + A\widetilde{x}, \quad \forall x, \widetilde{x} \in \mathcal{X}.$$

When the image and domain are the same, the notation is shortened as $\mathcal{L}(\mathcal{X}) \coloneqq \mathcal{L}(\mathcal{X}, \mathcal{X})$.

Definition A.2.5. ***Adjoint operators.*** *In the space of linear operators, we can define the adjoint of a linear operator* $A \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ *as* $A^* \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *such that*

$$(x, Ay)_\mathcal{X} = (A^* x, y)_\mathcal{Y}, \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}.$$

Definition A.2.6. ***Dual space.*** *Let $\mathcal{X}$ be a Hilbert space, we define its dual space as $\mathcal{X}' = \mathcal{L}(\mathcal{X}, \mathbb{R})$, the space of linear functionals. We also introduce the notation for its corresponding duality product*

$$\langle A, x\rangle_{\mathcal{X}',\mathcal{X}} = A(x) \in \mathbb{R}, \quad \forall A \in \mathcal{X}', x \in \mathcal{X}.$$

Definition A.2.7. ***Fréchet differential.*** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hilbert spaces, an operator $A : \mathcal{X} \mapsto \mathcal{Y}$ is Fréchet-differentiable at $x \in \mathcal{X}$ if there exists with*

$$\mathrm{D}_x A(x) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}),$$

such that $\forall (x, \widetilde{x}) \in \mathcal{X} \times \mathcal{X}$

$$A(x + \widetilde{x}) = A(x) + \mathrm{D}_x A(x)\widetilde{x} + o(\|\widetilde{x}\|_{\mathcal{X}}).$$

Remark. *In Definition A.2.7, if A has only one argument, we omit x in the differentiation symbol such that $\mathrm{D}A(x) = \mathrm{D}_x A(x)$.*

Remark. *In Definition A.2.7, if A is twice-differentiable, we have*

$$A(x + \widetilde{x}) = A(x) + \mathrm{D}_x A(x)\widetilde{x} + O(\|\widetilde{x}\|^2_{\mathcal{X}}).$$

Definition A.2.8. ***Second Fréchet differential.*** *For x, $\widetilde{x}$, $\widetilde{x}' \in \mathcal{X}$ and a operator $A : \mathcal{X} \mapsto \mathcal{Y}$ twice-differentiable, with $\mathcal{X}$ and $\mathcal{Y}$ being Hilbert spaces, we define the second-order (Fréchet) differentiation of A at x as*

$$\mathrm{D}^2_x A(x) \in \mathcal{L}\big(\mathcal{X}, \mathcal{L}(\mathcal{X}, \mathcal{Y})\big),$$

if it exists, such that

$$\mathrm{D}_x A(x + \widetilde{x}')\widetilde{x} = \mathrm{D}_x A(x)\widetilde{x} + \left[\mathrm{D}^2_x A(x)\widetilde{x}\right] \widetilde{x}' + o(\|\widetilde{x}'\|_{\mathcal{X}}).$$

The space $\mathcal{L}\big(\mathcal{X}, \mathcal{L}(\mathcal{X}, \mathcal{Y})\big)$ is the space of bilinear mappings. If $\mathcal{Y} = \mathbb{R}$, leading to $\mathrm{D}^2_x A(x) \in \mathcal{L}(\mathcal{X}, \mathcal{X}')$, the operator $\mathrm{D}^2_x A(x)$ becomes a bilinear form and its application can be written as

$$\mathrm{D}^2_x A(x)(\widetilde{x}, \widetilde{x}') = \langle \mathrm{D}^2_x A(x)\widetilde{x}, \widetilde{x}'\rangle_{\mathcal{X}',\mathcal{X}} = \left[\mathrm{D}^2_x A(x)\widetilde{x}\right] \widetilde{x}'.$$

Remark. *In Definition A.2.8, if A has only one argument, we omit x in the differentiation symbol such that $\mathrm{D}^2 A(x) = \mathrm{D}^2_x A(x)$.*

Remark. *The Definition A.2.7 and Definition A.2.8 can be generalized for higher-order derivatives, namely D^3_x, D^4_x, etc.*

A.2.2 Properties

Property A.2.1. ***Riesz representation theorem.*** *Let $\mathcal{X}$ be a Hilbert space and $\mathcal{X}'$ its dual. For a given linear map $A \in \mathcal{X}'$, there exists an unique $x' \in \mathcal{X}$ such that*

$$\langle A, \widetilde{x}\rangle_{\mathcal{X}',\mathcal{X}} = (x', \widetilde{x})_{\mathcal{X}}, \quad \forall \widetilde{x} \in \mathcal{X}.$$

Property A.2.2. ***Vectorial representation of elements.*** *Let $\mathcal{X}$ be a Hilbert space of dimension N, then there exists $\{e_i\}_{i=1}^N \in \mathcal{X}^N$ such that*

$$\mathcal{X} = \textbf{span}\{e_i\}_{i=1}^N,$$

therefore every element $x \in \mathcal{X}$ can be associated with an element $\overrightarrow{x} \in \mathbb{R}^N$, such that

$$x = \sum_{i=1}^N x_i e_i,$$

where x_i is the i-th component of $\overrightarrow{x}$.

Property A.2.3. ***Matricial representation of norms.*** *Let $\mathcal{X}$ be a Hilbert space of dimension N with*

$$\mathcal{X} = \textbf{span}\{e_i\}_{i=1}^N,$$

there exists a unique symmetric matrix $\mathbb{X} \in \mathcal{M}(\mathbb{R})_{N\times N}$ with components

$$(\mathbb{X})_{ij} = (e_i, e_j)_{\mathcal{X}}$$

such that

$$(x, \widetilde{x})_{\mathcal{X}} = \overrightarrow{x}^{\mathsf{T}} \mathbb{X} \overrightarrow{\widetilde{x}}, \quad \forall x, \widetilde{x} \in \mathcal{X},$$

where $\overrightarrow{x}$ and $\overrightarrow{\widetilde{x}}$ are the $\mathbb{R}^N$ representations of x and $\widetilde{x}$, respectively.

Property A.2.4. ***Matricial representation of linear operators.*** *Let $\mathcal{X}$, $\mathcal{Y}$ be Hilbert spaces of dimensions $N_{\mathcal{X}}$ and $N_{\mathcal{Y}}$, respectively, with*

$$\mathcal{X} = \textbf{span}\{e_i\}_{i=1}^{N_{\mathcal{X}}} \quad \textit{and} \quad \mathcal{Y} = \textbf{span}\{n_j\}_{j=1}^{N_{\mathcal{Y}}}.$$

Given an operator $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ there exists a unique matrix $\mathbb{A} \in \mathcal{M}(\mathbb{R})_{N_{\mathcal{Y}}\times N_{\mathcal{X}}}$ with components

$$(\mathbb{A})_{ij} = (Ae_i, n_j)_{\mathcal{Y}}$$

such that, $\forall x \in \mathcal{X}$ and $\forall y = Ax \in \mathcal{Y}$,

$$\mathbb{Y}\overrightarrow{y} = \mathbb{A}\overrightarrow{x},$$

with $\mathbb{Y}$ begin the norm of $\mathcal{Y}$ in its matricial form and where $\overrightarrow{x}$ and $\overrightarrow{y}$ are the $\mathbb{R}^{N_{\mathcal{X}}}$ and $\mathbb{R}^{N_{\mathcal{Y}}}$ representations of x and y, respectively.

Property A.2.5. ***Inversion Lemma.*** *Inversion Lemma. Let A, B, C and D be four invertible operators, we have that*

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}.$$

Bibliography

Abderahmane, A., A. Lhémery, and L. Daniel (2021). “Effects of multiaxial pre-stress on Lamb and shear horizontal guided waves”. In: *The Journal of the Acoustical Society of America* 149.3, pp. 1724–1736 (cit. on pp. 19, 53).

— (2022). “Stress imaging by guided wave tomography based on analytical acoustoelastic model”. In: *The Journal of the Acoustical Society of America* 151.5, pp. 2863–2876 (cit. on p. 20).

Abiza, Z., M. Destrade, and R. W. Ogden (2012). “Large acoustoelastic effect”. In: *Wave Motion* 49.2, pp. 364–374 (cit. on pp. 8, 19, 52, 146).

Adams, R. A. and J. J. Fournier (2003). *Sobolev spaces*. Elsevier (cit. on p. 57).

Afshar, S and K. Germ F.and Morris (2020). “Well-posedness of Extended Kalman Filter equations for semilinear infinite-dimensional systems”. In: *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 1210–1215 (cit. on p. 154).

Aghamiry, H., A. Gholami, K. Aghazade, M. M. Sonbolestan, and S. Operto (2022). *Large-scale highly-accurate extended full waveform inversion using convergent Born series* (cit. on p. 96).

Aghasi, A., M. Kilmer, and E. L. Miller (2011). “Parametric Level Set Methods for Inverse Problems”. In: *SIAM Journal on Imaging Sciences* 4.2, pp. 618–650 (cit. on p. 190).

Agudo, O. C., L. Guasch, P. Huthwaite, and M. Warner (2018). “3D imaging of the breast using full-waveform inversion”. In: *Proc. Int. Workshop Med. Ultrasound Tomogr*, pp. 99–110 (cit. on p. 96).

Alexandre Imperiale and E. Demaldent (2019). “A macro-element strategy based upon spectral finite elements and mortar elements for transient wave propagation modeling. Application to ultrasonic testing of laminate composite materials”. In: *International Journal for Numerical Methods in Engineering* 119.10, pp. 964–990 (cit. on pp. 20, 44, 69).

Arun Sundaram, B., K. Kesavan, and S. Parivallal (2018). “Recent Advances in Health Monitoring and Assessment of In-service Oil and Gas Buried Pipelines”. In: *Journal of The Institution of Engineers (India): Series A* 99.4, pp. 729–740 (cit. on pp. 7, 17).

Asay, J. R. and A. H. Guenther (1967). “Ultrasonic Studies of 1060 and 6061-T6 Aluminum”. In: *Journal of Applied Physics* 38.10, pp. 4086–4088 (cit. on pp. 19, 64).

Aussal, M. and P. Moireau (2022). “Kernel representation of Kalman observer and associated H-matrix based discretization”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 28, p. 78 (cit. on p. 147).

Bal, G., W. Naetar, S. O., and J. Schotland (2013). “The Levenberg–Marquardt iteration for numerical inversion of the power density operator”. In: *Journal of Inverse and Ill-Posed Problems* 21.2, pp. 265–280 (cit. on pp. 148, 152).

Basabe, J. D. (2007). “Grid dispersion and stability criteria of some common finite-difference and finite-element methods for acoustic and elastic wave propagation”. In: *Seg Technical Program Expanded Abstracts* (cit. on pp. 9, 20, 39, 68).

Bathe, K.-J. and E. N. Dvorkin (1986). “A formulation of general shell elements—the use of mixed interpolation of tensorial components”. In: *International journal for numerical methods in engineering* 22.3, pp. 697–722 (cit. on pp. 53, 65).

Bathe, K.-J., A. Iosilevich, and D. Chapelle (2000). “An evaluation of the MITC shell elements”. In: *Computers & Structures* 75.1, pp. 1–30 (cit. on pp. 53, 65).

Bathe, K.-J. (2006). *Finite Element Procedures*. Klaus-Jurgen Bathe (cit. on pp. 9, 19, 20, 53, 65).

Baudouin, L., M. De Buhan, S. Ervedoza, and A. Osses (2021). “Carleman-based reconstruction algorithm for waves”. In: *SIAM Journal on Numerical Analysis* 59.2, pp. 998–1039 (cit. on p. 190).

Bensoussan, A. (1971). *Filtrage optimal des systèmes linéaires*. Vol. 3. Dunod (cit. on p. 147).

Bensoussan, A., G. Da Prato, M. C. Delfour, and S. K. Mitter (2007). *Representation and Control of Infinite Dimensional Systems*. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston (cit. on pp. 21, 102, 105, 150, 153).

Bergman, R. H. and R. A. Shahbender (1958). “Effect of Statically Applied Stresses on the Velocity of Propagation of Ultrasonic Waves”. In: *Journal of Applied Physics* 29.12, pp. 1736–1738 (cit. on pp. 20, 64).

Bernard, S., V. Monteiller, D. Komatitsch, and P. Lasaygues (2017). “Ultrasonic computed tomography based on full-waveform inversion for bone quantitative imaging”. In: *Physics in Medicine & Biology* 62.17, p. 7011 (cit. on p. 147).

Bertino, L., G. Evensen, and H. Wackernagel (2003). “Sequential Data Assimilation Techniques in Oceanography”. In: *International Statistical Review* 71.2, pp. 223–241 (cit. on pp. 21, 105).

Bhatia, R. (2015). *Positive definite matrices*. Princeton series in applied mathematics. Princeton Oxford: Princeton University Press (cit. on p. 90).

Biot, M. A. (1940). “The Influence of Initial Stress on Elastic Waves”. In: *Journal of Applied Physics* 11.8, pp. 522–530 (cit. on pp. 18, 62).

Blum, J., F.-X. Le Dimet, and I. M. Navon (2009). “Data assimilation for geophysical fluids”. In: *Handbook of numerical analysis*. Vol. 14. Elsevier, pp. 385–441 (cit. on p. 147).

Boller, C (2001). “Ways and options for aircraft structural health management”. In: *Smart Materials and Structures* 10.3, pp. 432–440 (cit. on pp. 7, 17).

Bonet, J and A. Burton (1998). “A simple orthotropic, transversely isotropic hyperelastic constitutive equation for large strain computations”. In: *Computer methods in applied mechanics and engineering* 162.1-4, pp. 151–164 (cit. on pp. 36, 62).

Bonet, J. and R. D. Wood (1997). *Nonlinear continuum mechanics for finite element analysis*. Cambridge ; New York, NY, USA: Cambridge University Press (cit. on pp. 19, 33).

Bonnet, M. and W. Aquino (2015). “Three-dimensional transient elastodynamic inversion using an error in constitutive relation functional”. In: *Inverse Problems* 31.3, p. 035010 (cit. on pp. 21, 96).

Brown, R. J. (1989). “Relationships between the velocities and the elastic constants of an anisotropic solid possessing orthorhombic symmetry”. In: *CREWES Research Report* (cit. on p. 36).

Brugger, K. (1965). “Pure Modes for Elastic Waves in Crystals”. In: *Journal of Applied Physics* 36.3, pp. 759–768 (cit. on p. 18).

Bucalem, M. L. and K.-J. Bathe (1993). “Higher-order MITC general shell elements”. In: *International Journal for Numerical Methods in Engineering* 36.21, pp. 3729–3754 (cit. on pp. 53, 65).

Bulletti, A. and L. Capineri (2015). “Interdigital Piezopolymer Transducers for Time of Flight Measurements with Ultrasonic Lamb Waves on Carbon-Epoxy Composites under Pure Bending Stress”. In: *Journal of Sensors* 2015, pp. 1–11 (cit. on p. 64).

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995). “Multiscale seismic waveform inversion”. In: *GEOPHYSICS* 60.5, pp. 1457–1473 (cit. on p. 21).

Caiazzo, A., F. Caforio, G. Montecinos, L. O. Muller, P. J. Blanco, and E. F. Toro (2017). “Assessment of reduced-order unscented Kalman filter for parameter identification in 1-dimensional blood flow models using experimental data”. In: *International Journal for Numerical Methods in Biomedical Engineering* 33.8, e2843 (cit. on pp. 22, 105).

Carcione, J. M. (2001). *Wave Fields in Real Media Wave propagation in anisotropic, anelastic and porous media*. Kidlington: Pergamon (cit. on p. 36).

Castanie, B., C. Bouvet, and M. Ginot (2020). “Review of composite sandwich structure in aeronautic applications”. In: *Composites Part C: Open Access* 1, p. 100004 (cit. on p. 135).

Cea, J. (1986). “Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 20.3, pp. 371–402 (cit. on pp. 21, 96).

Chabassier, J. and S. Imperiale (2013). “Stability and dispersion analysis of improved time discretization for simply supported prestressed Timoshenko systems. Application to the stiff piano string”. In: *Wave Motion* 50.3, pp. 456–480 (cit. on p. 68).

— (2021). “Construction and convergence analysis of conservative second order local time discretisation for linear wave equations”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 55.4, pp. 1507–1543 (cit. on p. 68).

Chabiniok, R., P. Moireau, P.-F. Lesault, A. Rahmouni, J.-F. Deux, and D. Chapelle (2012). “Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model”. In: *Biomechanics and Modeling in Mechanobiology* 11.5, pp. 609–630 (cit. on p. 105).

Chakrapani, S. K. (2017). “Nonlinear laminated plate theory for determination of third order elastic constants and acoustic nonlinearity parameter of fiber reinforced composites”. In: *Composite Structures* 180, pp. 276–285 (cit. on p. 19).

Chapelle, D. and K.-J. Bathe (2011). *The Finite Element Analysis of Shells - Fundamentals*. 2nd ed. Computational Fluid and Solid Mechanics. Berlin Heidelberg: Springer-Verlag (cit. on pp. 8, 9, 18, 20, 53, 65).

Chapuis, B. (2023). *Guided waves based SHM system for rail monitoring and its environmental impact* (cit. on p. 17).

Chaves, E. W. V. (2013). *Notes on Continuum Mechanics*. Ed. by E. Oñate. Lecture Notes on Numerical Methods in Engineering and Sciences. Dordrecht: Springer Netherlands (cit. on pp. 8, 18, 30, 31, 33–35).

Chen, F. and P. D. Wilcox (2007). “The effect of load on guided wave propagation”. In: *Ultrasonics* 47.1, pp. 111–122 (cit. on pp. 19, 53).

Cheng, C., H. Mei, V. Giurgiutiu, S. Yuan, F. Fang, and R. James (2020). “Simulation of guided wave under varying temperature and load conditions”. In: *Health Monitoring of Structural and Biological Systems IX*. Online Only, United States: SPIE, p. 19 (cit. on pp. 19, 53).

Ciarlet, P. G. (1988a). *Introduction à l’analyse numérique matricielle et à l’optimisation* (cit. on pp. 20, 125).

— (1988b). *Mathematical elasticity. Volume I, Three-dimensional elasticity*. Amsterdam, New York: North-Holland ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co. (cit. on pp. 8, 18, 33, 54, 56, 57, 60, 61, 69, 149).

Clayton, J. D. and K. M. Bliss (2014). “Analysis of intrinsic stability criteria for isotropic third-order Green elastic and compressible neo-Hookean solids”. In: *Mechanics of Materials* 68, pp. 104–119 (cit. on p. 19).

Cohen, G. C. (2002). *Higher-order numerical methods for transient wave equations*. 1st edition. Scientific computation. Berlin Heidelberg: Springer-Verlag (cit. on pp. 9, 20, 38, 53, 66–68, 155).

Courtier, P., J.-N. Thépaut, and A. Hollingsworth (1994). “A strategy for operational implementation of 4D-Var, using an incremental approach”. In: *Quarterly Journal of the Royal Meteorological Society* 120.519, pp. 1367–1387 (cit. on pp. 21, 96, 147).

Courtier, P., J. Derber, R. Errico, J.-F. Louis, and T. Vukićević (1993). “Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology”. In: *Tellus A: Dynamic Meteorology and Oceanography* 45.5, p. 342 (cit. on pp. 21, 105).

Crecraft, D. I. (1967). “The measurement of applied and residual stresses in metals using ultrasonic waves”. In: *Journal of Sound and Vibration* 5.1, pp. 173–192 (cit. on pp. 20, 64).

Croxford, A., P. Wilcox, B. Drinkwater, and G Konstantinidis (2007). “Strategies for guided-wave structural health monitoring”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 463 (cit. on pp. 17, 84, 135).

Cîndea, N. and A. Münch (2015). “A mixed formulation for the direct approximation of the control of minimal L2-norm for linear type wave equations”. In: *Calcolo* 52.3, pp. 245–288 (cit. on p. 102).

Dalmora, A., A. Imperiale, S. Imperiale, and P. Moireau (2022). “A Generic Numerical Solver for Modeling the Influence of Stress Conditions on Guided Wave Propagation for SHM Applications”. In: *QNDE 2022 - 49th Annual Review of Progress in Quantitative Nondestructive Evaluation*. San Diego, CA, United States: American Society of Mechanical Engineers Digital Collection (cit. on pp. 147, 149, 150).

Dautray, R. and J.-L. Lions (1992). *Evolution problems I, volume 5 of Mathematical analysis and numerical methods for science and technology* (cit. on p. 60).

De Buhan, M. and M. Kray (2013). “A new approach to solve the inverse scattering problem for waves: combining the TRAC and the adaptive inversion methods”. In: *Inverse Problems* 29.8, p. 085009 (cit. on pp. 147, 190).

Degtyar, A. D. and S. I. Rokhlin (1995). “Absolute stress determination in orthotropic materials from angular dependences of ultrasonic velocities”. In: *Journal of Applied Physics* 78.3, pp. 1547–1556 (cit. on p. 20).

Dehghan-Niri, E., A. Farhidzadeh, and S. Salamone (2013). “Nonlinear Kalman Filtering for acoustic emission source localization in anisotropic panels”. In: *Ultrasonics* 54 (cit. on pp. 22, 105).

Demailly, J.-P. (2006). *Analyse numérique et équations différentielles*. Nouvelle éd. Collection Grenoble sciences. Les Ulis: EDP sciences (cit. on p. 107).

Dimet, F.-X. L. and O. Talagrand (1986). “Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects”. In: *Tellus A* 38A.2, pp. 97–110 (cit. on pp. 21, 96, 147).

Dodson, J. C. and D. J. Inman (2014). “Investigating the thermally induced acoustoelastic effect in isotropic media with Lamb waves”. In: *The Journal of the Acoustical Society of America* 136.5, pp. 2532–2543 (cit. on p. 17).

Dubuc, B., A. Ebrahimkhanlou, and S. Salamone (2017a). “Effect of pressurization on helical guided wave energy velocity in fluid-filled pipes”. In: *Ultrasonics* 75, pp. 145–154 (cit. on pp. 17, 19, 53).

— (2017b). “The effect of applied stress on the phase and group velocity of guided waves in anisotropic plates”. In: *The Journal of the Acoustical Society of America* 142.6, pp. 3553–3563 (cit. on pp. 19, 53, 64).

— (2018). “Computation of propagating and non-propagating guided modes in nonuniformly stressed plates using spectral methods”. In: *The Journal of the Acoustical Society of America* 143.6, pp. 3220–3230 (cit. on pp. 19, 53).

Dubuget, M., R. El Guerjouma, S. Dubois, J. C. Baboux, and A. Vincent (1996). “Characterization of the Non-Linear Elastic Properties of Aluminium Alloys Using Ultrasonic Evaluation under Load”. In: *Materials Science Forum* 217-222, pp. 951–956 (cit. on p. 19).

Duruflé, M., P. Grob, and P. Joly (2009). “Influence of Gauss and Gauss-Lobatto quadrature rules on the accuracy of a quadrilateral finite element method in the time domain.” In: *Numerical Methods for Partial Differential Equations* 25.3, pp. 526–551 (cit. on pp. 40, 67, 155).

Duvaut, G. and J. L. Lions (1976). *Inequalities in mechanics and physics*. Springer-Verlag (cit. on pp. 60, 150).

Eikrem, K. S., G. Nævdal, and M. Jakobsen (2019). “Iterated extended Kalman filter method for time-lapse seismic full-waveform inversion”. In: *Geophysical Prospecting* 67.2, pp. 379–394 (cit. on pp. 22, 105).

Engquist, B., B. D. Froese, and Y. Yang (2016). “Optimal transport for seismic full waveform inversion”. In: *Communications in Mathematical Sciences* 14.8, pp. 2309–2330 (cit. on pp. 21, 96, 185).

Engquist, B., K. Ren, and Y. Yang (2020). “The quadratic Wasserstein metric for inverse data matching”. In: *Inverse Problems* 36.5, p. 055001 (cit. on pp. 137, 185).

Evensen, G. (2003). “The Ensemble Kalman Filter: theoretical formulation and practical implementation”. In: *Ocean Dynamics* 53.4, pp. 343–367 (cit. on pp. 21, 105).

Farrar, C. R. and K. Worden (2010). “An Introduction to Structural Health Monitoring”. In: *New Trends in Vibration Based Structural Health Monitoring*. CISM Courses and Lectures. Vienna: Springer, pp. 1–17 (cit. on pp. 7, 17).

Feissel, P. and O. Allix (2007). “Modified constitutive relation error identification strategy for transient dynamics with corrupted data: The elastic case”. In: *Computer Methods in Applied Mechanics and Engineering* 196.13, pp. 1968–1983 (cit. on pp. 21, 96).

Furuya, T. and R. Potthast (2022). “Inverse medium scattering problems with Kalman filter techniques”. In: *Inverse Problems* 38.9, p. 095003 (cit. on pp. 22, 148).

Gandhi, N., J. E. Michaels, and S. J. Lee (2012). “Acoustoelastic Lamb wave propagation in biaxially stressed plates”. In: *The Journal of the Acoustical Society of America* 132.3, pp. 1284–1293 (cit. on pp. 19, 53, 64, 76, 77, 166, 175).

Gandhi, M. (2010). *Determination of dispersion curves for acoustoelastic Lamb Wave propagation* (cit. on p. 76).

Gorgin, R., Y. Luo, and Z. Wu (2020). “Environmental and operational conditions effects on Lamb wave based structural health monitoring systems: A review”. In: *Ultrasonics*, p. 106114 (cit. on pp. 8, 17, 52, 146).

Guasch, L., O. Calderón Agudo, M.-X. Tang, P. Nachev, and M. Warner (2020). “Full-waveform inversion imaging of the human brain”. In: *npj Digital Medicine* 3.1, pp. 1–12 (cit. on p. 96).

Haik, W., Y Maday, and L. Chamoin (2023). “A real-time variational data assimilation method with data-driven model enrichment for time-dependent problems”. In: *Computer Methods in Applied Mechanics and Engineering* 405, p. 115868 (cit. on p. 147).

Hanke, M. (1997). “A regularizing Levenberg - Marquardt scheme, with applications to inverse groundwater filtration problems”. In: *Inverse Problems* 13.1, p. 79 (cit. on pp. 148, 152).

Hayes, M. and A. E. Green (1963). “Wave propagation and uniqueness in prestressed elastic solids”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 274.1359, pp. 500–506 (cit. on p. 19).

Helnwein, P. (2001). “Some remarks on the compressed matrix representation of symmetric second-order and fourth-order tensors”. In: *Computer Methods in Applied Mechanics and Engineering* 190.22, pp. 2753–2770 (cit. on p. 35).

Hikata, A., R. Truell, A. Granato, B. Chick, and K. Lücke (1956). “Sensitivity of Ultrasonic Attenuation and Velocity Changes to Plastic Deformation and Recovery in Aluminum”. In: *Journal of Applied Physics* 27.4, pp. 396–404 (cit. on pp. 18, 64).

Hiki, Y. (1981). “Higher Order Elastic Constants of Solids”. In: *Annual Review of Materials Science* 11.1, pp. 51–73 (cit. on p. 18).

Holzapfel, G. A. (2000). *Nonlinear solid mechanics: a continuum approach for engineering*. Chichester ; New York: Wiley (cit. on pp. 8, 18).

Houtekamer, P. L. and F. Zhang (2016). “Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation”. In: *Monthly Weather Review* 144.12, pp. 4489–4532 (cit. on pp. 21, 105).

Hsu, N. N. (1974). “Acoustical birefringence and the use of ultrasonic waves for experimental stress analysis”. In: *Experimental Mechanics* 14.5, pp. 169–176 (cit. on pp. 20, 64).

Huang, C. and T. Zhu (2019). “Time-lapse full waveform inversion plus extended Kalman filter for high-resolution seismic models and uncertainty estimation”. In: *SEG Technical Program Expanded Abstracts 2019*. SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 5239–5244 (cit. on p. 105).

Hughes, D. S. and J. L. Kelly (1953). “Second-Order Elastic Deformation of Solids”. In: *Physical Review* 92.5, pp. 1145–1149 (cit. on pp. 8, 18, 19, 33, 62–64).

Huthwaite, P. (2014). “Accelerated finite element elastodynamic simulations using the GPU”. In: *Journal of Computational Physics* 257, pp. 687–707 (cit. on p. 45).

Itskov, M. and N. Aksel (2004). “A class of orthotropic and transversely isotropic hyperelastic constitutive models based on a polyconvex strain energy function”. In: *International Journal of Solids and Structures* 41.14, pp. 3833–3848 (cit. on pp. 19, 190).

Joly, P. (2003). “Variational Methods for Time-Dependent Wave Propagation Problems”. In: *Topics in Computational Wave Propagation: Direct and Inverse Problems*. Lecture Notes in Computational Science and Engineering. Berlin, Heidelberg: Springer, pp. 201–264 (cit. on pp. 38, 67, 68).

Joly, P. (2007). “Numerical Methods for Elastic Wave Propagation”. In: *Waves in Nonlinear Pre-Stressed Materials*. CISM Courses and Lectures 495. Vienna: Springer, pp. 181–281 (cit. on pp. 9, 20, 38, 53, 66, 67).

Jones, R. H. (1965). “Optimal Estimation of Initial Conditions for Numerical Prediction”. In: *Journal of the Atmospheric Sciences* 22.6, pp. 658–663 (cit. on pp. 21, 105).

Julier, S., J. Uhlmann, and H. Durrant-Whyte (2000). “A new method for the nonlinear transformation of means and covariances in filters and estimators”. In: *IEEE Transactions on Automatic Control* 45.3, pp. 477–482 (cit. on pp. 21, 114).

Kijanka, P., P. Packo, X. Zhu, W. J. Staszewski, and F. L. d. Scalea (2015). “Three-dimensional temperature effect modelling of piezoceramic transducers used for Lamb wave based damage detection”. In: *Smart Materials and Structures* 24.6, p. 065005 (cit. on p. 17).

Komatitsch, D., J.-P. Vilotte, R. Vai, J. M. Castillo-Covarrubias, and F. J. Sanchez-Sesma (1999). “The spectral element method for elastic wave equations-application to 2-D and 3-D seismic problems”. In: *International Journal for numerical methods in engineering* 45.9, pp. 1139–1164 (cit. on pp. 9, 20, 53, 66, 67).

Komatitsch, D. and J.-P. Vilotte (1998). “The spectral element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures”. In: *Bulletin of the Seismological Society of America* 88.2, pp. 368–392 (cit. on p. 38).

Kubrusly, A. C., N. Pérez, T. F. de Oliveira, J. C. Adamowski, A. M. B. Braga, and J. P. von der Weid (2016). “Mechanical Strain Sensing by Broadband Time Reversal in Plates”. In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 63.5, pp. 746–756 (cit. on p. 64).

Lamb, H. (1917). “On waves in an elastic plate”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 93.648, pp. 114–128 (cit. on pp. 7, 16, 48).

Landau, L. D., E. M. Lifšic, and L. D. Landau (1986). *Theory of elasticity*. 2. ed. Course of theoretical physics Vol. 7. Oxford: Pergamon Press (cit. on pp. 8, 18, 61).

Le Tallec, P. (1994). “Numerical methods for nonlinear three-dimensional elasticity”. In: *Handbook of Numerical Analysis*. Vol. 3. Elsevier, pp. 465–622 (cit. on pp. 19, 54, 56).

— (2009). *Modélisation et calcul des milieux continus*. Palaiseau: les Éd. de l’ École polytechnique (cit. on pp. 8, 18, 34).

Léger, S, J Deteix, and A Fortin (2015). “A Moore–Penrose continuation method based on a Schur complement approach for nonlinear finite element bifurcation problems”. In: *Computers & Structures* 152, pp. 173–184 (cit. on p. 64).

Li, H. and J. Ou (2016). “The state of the art in structural health monitoring of cable-stayed bridges”. In: *Journal of Civil Structural Health Monitoring* 6.1, pp. 43–67 (cit. on pp. 7, 17).

Lions, J. L. (1971). *Optimal control of systems governed by partial differential equations*. Vol. 170. Springer (cit. on pp. 21, 60, 96, 102).

Liu, C., F. Gao, X. Feng, Y. Liu, and Q. Ren (2015). “Memoryless quasi-Newton (MLQN) method for 2D acoustic full waveform inversion”. In: *Exploration Geophysics* 46.2, pp. 168–177 (cit. on p. 100).

Lonné, S. (2003). *Modélisation de la propagation ultrasonore dans les matériaux composites obtenus par le procédé de fabrication RTM (Resin Transfer Molding)* (cit. on p. 74).

Loveday, P., C. Long, and P. Wilcox (2012). “Semi-analytical finite element analysis of the influence of axial loads on elastic waveguides”. In: (cit. on pp. 19, 53).

Lucka, F., M. Pérez-Liva, B. E. Treeby, and B. T. Cox (2021). “High resolution 3D ultrasonic breast imaging by time-domain full waveform inversion”. In: *Inverse Problems* 38.2, p. 025008 (cit. on p. 96).

Luenberger, D. (1971). “An introduction to observers”. In: *IEEE Transactions on Automatic Control* 16.6, pp. 596–602 (cit. on p. 106).

Luenberger, D. G. and Y. Ye (2008). *Linear and nonlinear programming*. 3rd ed. Vol. 2. International series in operations research and management science. New York, NY: Springer (cit. on pp. 20, 85, 96–99, 147).

Ma, X., G. Liang, S. Xu, Z. Li, and H. Feng (2020). “Exact line search method for using the L1-norm misfit function in full waveform inversion”. In: *Studia Geophysica et Geodaetica* 64.4, pp. 483–503 (cit. on p. 96).

Ma, Y., Z. Yang, J. Zhang, K. Liu, Z. Wu, and S. Ma (2019). “Axial stress monitoring strategy in arbitrary cross-section based on acoustoelastic guided waves using PZT sensors”. In: *AIP Advances* 9.12, p. 125304 (cit. on p. 20).

Maday, Y. and A. T. Patera (1989). *Spectral element methods for the incompressible Navier-Stokes equations* (cit. on pp. 9, 20, 38).

Mainini, E. (2012). “A description of transport cost for signed measures”. In: *Journal of Mathematical Sciences* 181.6, pp. 837–855 (cit. on p. 137).

Malekloo, A., E. Ozer, M. AlHamaydeh, and M. Girolami (2022). “Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights”. In: *Structural Health Monitoring* 21.4, pp. 1906–1955 (cit. on p. 17).

Man, C.-S. and W. Y. Lu (1987). “Towards an acoustoelastic theory for measurement of residual stress”. In: *Journal of Elasticity* 17.2, pp. 159–182 (cit. on p. 18).

Marchand, B., L. Chamoin, and C. Rey (2016). “Real-time updating of structural mechanics models using Kalman filtering, modified constitutive relation error, and proper generalized decomposition”. In: *International Journal for Numerical Methods in Engineering* 107.9, pp. 786–810 (cit. on pp. 22, 105, 148).

Mariani, S. and A. Ghisi (2007). “Unscented Kalman filtering for nonlinear structural dynamics”. In: *Nonlinear Dynamics* 49.1, pp. 131–150 (cit. on pp. 22, 105).

Marsden, J. E. and T. J. R. Hughes (1978). “Topics in the mathematical foundations of elasticity”. In: (cit. on pp. 54, 60).

Marzani, A., N. Testoni, L. De Marchi, M. Messina, E. Monaco, and A. Apicella (2020). “An open database for benchmarking guided waves structural health monitoring algorithms on a composite full-scale outer wing demonstrator”. In: *Structural Health Monitoring* 19.5, pp. 1524–1541 (cit. on pp. 8, 17).

Mesnil, O., A. Recoquillay, T. Druet, V. Serey, H. T. Hoang, A. Imperiale, and E. Demaldent (2021). “Experimental validation of transient spectral finite element simulation tools dedicated to guided wave based structural health monitoring”. In: *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, pp. 1–14 (cit. on p. 20).

Methenni, H. (2021). “Modélisation mathématique et méthode numérique pour la simulation du contrôle santé intégré par ultrasons de plaques composites stratifiées”. PhD thesis (cit. on p. 75).

Michaels, J., S.-J. Lee, and T. Michaels (2011). “Impact of Applied Loads on Guided Wave Structural Health Monitoring”. In: *AIP Conference Proceedings* 1335 (cit. on p. 17).

Michaels, J. E., T. E. Michaels, R. S. Martin, D. O. Thompson, and D. E. Chimenti (2009). “Analysis Of Global Ultrasonic Sensor Data From A Full Scale Wing Panel Test”. In: *AIP Conference Proceedings*. Chicago (Illinois): AIP, pp. 950–957 (cit. on p. 17).

Mitra, M. and S Gopalakrishnan (2016). “Guided wave based structural health monitoring: A review”. In: *Smart Materials and Structures* 25.5, p. 053001 (cit. on pp. 7, 17, 52, 84, 146).

Mohabuth, M., A. Kotousov, C.-T. Ng, and L. R. F. Rose (2018). “Implication of changing loading conditions on structural health monitoring utilising guided waves”. In: *Smart Materials and Structures* 27.2, p. 025003 (cit. on pp. 19, 53).

Moireau, P. (2019). “A Discrete-time Optimal Filtering Approach for Non-linear Systems as a Stable Discretization of the Mortensen Observer”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 24.4, pp. 1815 –1847 (cit. on p. 148).

Moireau, P. (2022). “Discrete-time formulations as time discretization strategies in data assimilation”. In: *Handbook of Numerical Analysis*. Handbook of Numerical Analysis. Elsevier (cit. on pp. 22, 84, 93, 103, 105, 110, 147, 148, 153, 155, 157).

Moireau, P. and D. Chapelle (2011). “Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 17.2, pp. 380–405 (cit. on pp. 114, 120, 147, 148, 160, 161).

Moireau, P., D. Chapelle, and P. Le Tallec (2008). “Joint state and parameter estimation for distributed mechanical systems”. In: *Computer Methods in Applied Mechanics and Engineering* 197.6-8, pp. 659–677 (cit. on pp. 22, 105, 147, 148, 153, 154).

Moireau, Philippe and Chapelle, Dominique (2011). “Erratum of article “Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems””. In: *ESAIM: COCV* 17.2, pp. 406–409 (cit. on p. 160).

Molent, L. and B. Aktepe (2000). “Review of fatigue monitoring of agile military aircraft”. In: *Fatigue and Fracture of Engineering Materials and Structures* 23.9, pp. 767–785 (cit. on pp. 7, 17).

Moore, B. (1981). “Principal component analysis in linear systems: Controllability, observability, and model reduction”. In: *IEEE Transactions on Automatic Control* 26.1, pp. 17–32 (cit. on p. 168).

Mora, P. (1987). “Nonlinear two-dimensional elastic inversion of multioffset seismic data”. In: *GEOPHYSICS* 52.9, pp. 1211–1228 (cit. on pp. 21, 96).

Muir, D. D., T. E. Michaels, and J. E. Michaels (2009). “One-sided liquid coupled ultrasonic method for recovery of third order elastic constants”. In: *AIP Conference Proceedings* 1096.1, pp. 1275–1282 (cit. on pp. 19, 64, 190).

Murnaghan, F. D. (1937). “Finite Deformations of an Elastic Solid”. In: *American Journal of Mathematics* 59.2, pp. 235–260 (cit. on pp. 8, 18).

Murnaghan, F. D. (1951). *Finite Deformation of an Elastic Solid*. John Wiley & Sons (cit. on pp. 8, 18, 33, 62).

Musgrave, M. J. P. (1970). *Crystal Acoustics: Introduction to the Study of Elastic Waves and Vibrations in Crystals*. Acoustical Society of America (cit. on p. 36).

Métivier, L., R. Brossier, F. Kpadonou, J. Messud, and A. Pladys (2022). “A review of the use of optimal transport distances for high resolution seismic imaging based on the full waveform”. In: *arXiv:2204.08514 [physics]* (cit. on pp. 21, 96).

Nakayama, S. (2019). “Memoryless quasi-Newton methods for large-scale unconstrained optimization”. In: (cit. on p. 98).

Ndanou, S., N. Favrie, and S. Gavrilyuk (2014). “Criterion of Hyperbolicity in Hyperelasticity in the Case of the Stored Energy in Separable Form”. In: *Journal of Elasticity* 115.1, pp. 1–25 (cit. on pp. 19, 190).

Nguyen, H. N., L. Chamoin, and C. Ha Minh (2022). “mCRE-based parameter identification from full-field measurements: Consistent framework, integrated version, and extension to nonlinear material behaviors”. In: *Computer Methods in Applied Mechanics and Engineering* 400, p. 115461 (cit. on pp. 21, 96).

Nocedal, J. and S. J. Wright (2006). *Numerical optimization*. 2nd ed. Springer series in operations research. New York: Springer (cit. on pp. 20, 85, 96–98, 128).

Ogden, R. W. (1984). “Non-linear elastic deformations”. In: *Engineering Analysis* 1.2, p. 119 (cit. on pp. 8, 18, 30, 32, 62, 190).

Ogden, R. W. (2007). “Incremental Statics and Dynamics of Pre-Stressed Elastic Materials”. In: *Waves in Nonlinear Pre-Stressed Materials*. CISM Courses and Lectures. Vienna: Springer, pp. 1–26 (cit. on pp. 8, 19).

Pao, Y. (1984). “Acoustoelasticity and Ultrasonic Measurement of Residual Stress”. In: *undefined* (cit. on pp. 18, 20).

Pao, Y. and U. Gamer (1985). “Acoustoelastic waves in orthotropic media”. In: *The Journal of the Acoustical Society of America* 77.3, pp. 806–812 (cit. on p. 18).

Pau, A. and F. Lanza di Scalea (2015). “Nonlinear guided wave propagation in prestressed plates”. In: *The Journal of the Acoustical Society of America* 137.3, pp. 1529–1540 (cit. on pp. 19, 53).

Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Vol. 44. Applied Mathematical Sciences. New York: Springer-Verlag, pp. viii+279 (cit. on pp. 153, 154).

Peddeti, K. and S. Santhanam (2018). “Dispersion curves for Lamb wave propagation in prestressed plates using a semi-analytical finite element analysis”. In: *The Journal of the Acoustical Society of America* 143.2, pp. 829–840 (cit. on pp. 19, 53).

Pei, N. and L. J. Bond (2016). “Higher order acoustoelastic Lamb wave propagation in stressed plates”. In: *The Journal of the Acoustical Society of America* 140.5, pp. 3834–3843 (cit. on pp. 17, 19, 53, 188).

Pham, D. T., J. Verron, and L. Gourdeau (1998). “Filtres de Kaiman singuliers évolutifs pour l’assimilation de données en océanographie”. In: *Comptes Rendus de l’Académie des Sciences - Series IIA - Earth and Planetary Science* 326.4, pp. 255–260 (cit. on pp. 21, 105, 160).

Pica, A., J. P. Diet, and A. Tarantola (1990). “Nonlinear inversion of seismic reflection data in a laterally invariant medium”. In: *GEOPHYSICS* 55.3, pp. 284–292 (cit. on pp. 21, 96).

Plessix, R.-E. (2006). “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications”. In: *Geophysical Journal International* 167.2, pp. 495–503 (cit. on pp. 21, 96).

Ponschab, M., D. A. Kiefer, and S. J. Rupitsch (2019). “Towards an Inverse Characterization of Third Order Elastic Constants Using Guided Waves”. In: *2019 IEEE International Ultrasonics Symposium (IUS)*, pp. 1264–1268 (cit. on pp. 19, 64).

Prosser, W. H. and R. E. Green (1990). “Characterization of the Nonlinear Elastic Properties of Graphite/Epoxy Composites Using Ultrasound”. In: *Journal of Reinforced Plastics and Composites* 9.2, pp. 162–173 (cit. on pp. 19, 64).

Rabier, F., J.-N. Thépaut, and P. Courtier (1998). “Extended assimilation and forecast experiments with a four-dimensional variational assimilation system”. In: *Quarterly Journal of the Royal Meteorological Society* 124.550, pp. 1861–1887 (cit. on pp. 21, 96).

Ramdani, K., M. Tucsnak, and G. Weiss (2010). “Recovering the initial state of an infinite-dimensional system using observers”. In: *Automatica* 46.10, pp. 1616–1625 (cit. on p. 147).

Rao, J., M. Ratassepp, and Z. Fan (2016). “Guided wave tomography based on full waveform inversion”. In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 63.5, pp. 737–745 (cit. on p. 96).

Ratassepp, M., J. Rao, X. Yu, and Z. Fan (2021). “Modeling the Effect of Anisotropy in Ultrasonic-Guided Wave Tomography”. In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 69.1, pp. 330–339 (cit. on p. 96).

Rauter, N. and R. Lammering (2018). “A constitutive model for the analysis of second harmonic Lamb waves in unidirectional composites”. In: *International Journal of Solids and Structures* 135, pp. 184–196 (cit. on p. 19).

Rayleigh, L. (1885). “On Waves Propagated along the Plane Surface of an Elastic Solid”. In: *Proceedings of the London Mathematical Society* s1-17.1, pp. 4–11 (cit. on p. 16).

Ricci, F., E. Monaco, N. Boffa, L. Maio, and V. Memmolo (2022). “Guided waves for structural health monitoring in composites: A review and implementation strategies”. In: *Progress in Aerospace Sciences* 129, p. 100790 (cit. on pp. 8, 17, 52, 146).

Rozier, D., F. Birol, E. Cosme, P. Brasseur, J. M. Brankart, and J. Verron (2007). “A Reduced-Order Kalman Filter for Data Assimilation in Physical Oceanography”. In: *SIAM Review* 49.3, pp. 449–465 (cit. on pp. 21, 22, 105, 158).

Santos Jr, A. and D. Bray (2002). “Comparison of Acoustoelastic Methods to Evaluate Stresses in Steel Plates and Bars”. In: *Journal of Pressure Vessel Technology-transactions of The Asme - J PRESSURE VESSEL TECHNOL* 124 (cit. on pp. 19, 20, 64).

Schröder, J. and P. Neff (2003). “Invariant formulation of hyperelastic transverse isotropy based on polyconvex free energy functions”. In: *International Journal of Solids and Structures* 40.2, pp. 401–445 (cit. on pp. 19, 190).

Semenov, A. A. and Y. M. Beltukov (2020). “Nonlinear elastic moduli of composite materials with nonlinear spherical inclusions dispersed in a nonlinear matrix”. In: *International Journal of Solids and Structures* 191-192, pp. 333–340 (cit. on p. 19).

Seriani, G. and S. P. Oliveira (2008). “Dft modal analysis of spectral element methods for acoustic wave propagation”. In: *Journal of Computational Acoustics* 16.04, pp. 531–561 (cit. on pp. 9, 20, 39, 68).

Seriani, G. and E. Priolo (1994). “Spectral element method for acoustic wave simulation in heterogeneous media”. In: *Finite Elements in Analysis and Design* 16.3-4, p. 337 (cit. on pp. 9, 20).

Shams, M., M. Destrade, and R. Ogden (2011). “Initial stresses in elastic solids: Constitutive laws and acoustoelasticity”. In: *Wave Motion* 48.7, pp. 552–567 (cit. on pp. 8, 18, 19, 52, 62, 146).

Shi, F., J. E. Michaels, and S. J. Lee (2013). “In situ estimation of applied biaxial loads with Lamb waves”. In: *The Journal of the Acoustical Society of America* 133.2, pp. 677–687 (cit. on pp. 20, 64).

Simon, D. (2006). *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J: Wiley-Interscience (cit. on pp. 21, 105, 106, 108, 114, 147, 157).

Smith, R. T., R. Stern, and R. W. B. Stephens (1966). “Third-Order Elastic Moduli of Polycrystalline Metals from Ultrasonic Velocity Measurements”. In: *The Journal of the Acoustical Society of America* 40.5, pp. 1002–1008 (cit. on pp. 19, 64).

Sohn, H. (2007). “Effects of environmental and operational variability on structural health monitoring”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1851, pp. 539–560 (cit. on p. 17).

Song, F., G. L. Huang, and G. K. Hu (2013). “Coupled piezo-elastodynamic modeling of guided wave excitation and propagation in plates with applied prestresses”. In: *Journal of Intelligent Material Systems and Structures* 24.5, pp. 598–611 (cit. on pp. 17, 64).

Spencer, A. J. M. (1982). “The Formulation of Constitutive Equation for Anisotropic Solids”. In: *Mechanical Behavior of Anisotropic Solids / Comportment Méchanique des Solides Anisotropes*. Dordrecht: Springer Netherlands, pp. 3–26 (cit. on pp. 19, 32).

Särkkä, S. (2013). *Bayesian Filtering and Smoothing*. 1st ed. Cambridge University Press (cit. on pp. 105, 106, 154).

Takahashi, S. (2018). “Measurement of third-order elastic constants and stress dependent coefficients for steels”. In: *Mechanics of Advanced Materials and Modern Processes* 4.1, p. 2 (cit. on p. 19).

Tarantola, A. (1984). “Inversion of seismic reflection data in the acoustic approximation”. In: *GEOPHYSICS* 49.8, pp. 1259–1266 (cit. on pp. 9, 21, 96).

Temam, Roger and Miranville, Alain (2005). *Mathematical Modeling in Continuum Mechanics*. 2nd ed. Cambridge: Cambridge University Press (cit. on p. 30).

Thompson, R. B., S. S. Lee, and J. F. Smith (1986). “Angular dependence of ultrasonic wave propagation in a stressed, orthorhombic continuum: Theory and application to the measurement of stress and texture”. In: *The Journal of the Acoustical Society of America* 80.3, pp. 921–931 (cit. on pp. 19, 64).

Thurston, R. N. and K. Brugger (1964). “Third-Order Elastic Constants and the Velocity of Small Amplitude Elastic Waves in Homogeneously Stressed Media”. In: *Physical Review* 133.6A, A1604–A1610 (cit. on p. 18).

Toupin, R. A. and B. Bernstein (1961). “Sound Waves in Deformed Perfectly Elastic Materials. Acoustoelastic Effect”. In: *The Journal of the Acoustical Society of America* 33.2, pp. 216–225 (cit. on p. 18).

Tschöke, K., B. Weihnacht, E. Schulze, T. Gaul, L. Schubert, and R. Neubeck (2017). “Determination of Defect Sizes with the help of Structural-Health-Monitoring Methods based on Guided Waves”. In: *2017-12* (cit. on pp. 17, 71, 72, 166, 179).

Villares Holguin, C. D., H. V. Hultmann Ayala, and A. C. Kubrusly (2021). “Improved Stress Estimation with Machine Learning and Ultrasonic Guided Waves”. In: *Experimental Mechanics* (cit. on p. 17).

Virieux, J. and S. Operto (2009). “An overview of full-waveform inversion in exploration geophysics”. In: *GEOPHYSICS* 74.6, WCC1–WCC26 (cit. on pp. 21, 96).

Virieux, J., A. Asnaashari, R. Brossier, L. Métivier, A. Ribodetti, and W. Zhou (2017). “An introduction to full waveform inversion”. In: *Encyclopedia of exploration geophysics*. Society of Exploration Geophysicists, R1–1 (cit. on pp. 21, 135, 147).

Wilcox, P. D., A. J. Croxford, G. Konstantinidis, and B. W. Drinkwater (2007). “Sensitivity limitations for guided wave structural health monitoring”. In: 6532, 65320A (cit. on p. 17).

Willberg, C., S. Duczek, J. M. Vivar-Perez, and Z. A. B. Ahmad (2015). “Simulation Methods for Guided Wave-Based Structural Health Monitoring: A Review”. In: *Applied Mechanics Reviews* 67.1, p. 010803 (cit. on p. 19).

Worden, K., C. R. Farrar, G. Manson, and G. Park (2007). “The fundamental axioms of structural health monitoring”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 463.2082, pp. 1639–1664 (cit. on pp. 7, 17).

Workman, G. L., D. Kishoni, and P. O. Moore, eds. (2007). *Ultrasonic testing*. 3rd ed. Nondestructive testing handbook v. 7. Columbus, OH: American Society for Nondestructive Testing (cit. on pp. 7, 16).

Wu, J.-D., C.-W. Huang, and R. Huang (2004). “An application of a recursive Kalman filtering algorithm in rotating machinery fault diagnosis”. In: *NDT & E International* 37.5, pp. 411–419 (cit. on pp. 22, 105).

Xu, W., M. Yuan, W. Xuan, X. Ji, and Y. Chen (2021). “Quantitative Inspection of Complex-Shaped Parts Based on Ice-Coupled Ultrasonic Full Waveform Inversion Technology”. In: *Applied Sciences* 11.10, p. 4433 (cit. on p. 96).

Yue, N. and M. Aliabadi (2020). “A scalable data-driven approach to temperature baseline reconstruction for guided wave structural health monitoring of anisotropic carbon-fibre-reinforced polymer structures”. In: *Structural Health Monitoring* 19.5, pp. 1487–1506 (cit. on p. 17).

Zhao, C., X. Chen, J. Li, and Y. Liu (2022). “Stress inversion in waveguides with arbitrary cross sections with acoustoelastic guided waves”. In: *Journal of Applied Physics* 131.24, p. 245102 (cit. on p. 20).

Zhu, Y., J. Tan, and C. Song (2020). “Finite Element Analysis and Ultrasonic Evaluation of Stresses in Q345 Plates”. In: *IOP Conference Series: Materials Science and Engineering* 740.1, p. 012005 (cit. on p. 20).

Zou, Y., D. Du, B. Chang, L. Ji, and J. Pan (2015). “Automatic weld defect detection method based on Kalman filtering for real-time radiographic inspection of spiral pipe”. In: *NDT & E International* 72, pp. 1–9 (cit. on pp. 22, 105).

Zuazua, E. (2005). “Propagation, Observation, and Control of Waves Approximated by Finite Difference Methods”. In: *SIAM Review* 47.2, pp. 197–243 (cit. on p. 93).

All those daring birds that soar far and ever farther into space, will somewhere or other be certain to find themselves unable to continue their flight, and they will perch on a mast or some narrow ledge—and will be grateful even for this miserable accommodation! But who could conclude from this that there was not an endless free space stretching far in front of them, and that they had flown as far as they possibly could? In the end, however, all our great teachers and predecessors have come to a standstill, and it is by no means in the noblest or most graceful attitude that their weariness has brought them to a pause: the same thing will happen to you and me! but what does this matter to either of us? Other birds will fly farther! Our minds and hopes vie with them far out and on high; they rise far above our heads and our failures, and from this height they look far into the distant horizon and see hundreds of birds much more powerful than we are, striving whither we ourselves have also striven, and where all is sea, sea, and nothing but sea! And where, then, are we aiming at? Do we wish to cross the sea? whither does this over-powering passion urge us, this passion which we value more highly than any other delight? Why do we fly precisely in this direction, where all the suns of humanity have hitherto set? Is it possible that people may one day say of us that we also steered westward, hoping to reach India—but that it was our fate to be wrecked on the infinite? Or, my brethren? or—?
- Friedrich Nietzsche, 1881

I Just Wanted to Make You Something Beautiful
Industries of the Blind

ECOLE DOCTORALE

Titre : Modélisation et assimilation de données pour les ondes guidées ultrasonores pour le *contrôle de santé intégré* dans des conditions de charges opérationnelles

Mots clés : contrôle santé intégré, conditions opérationnelles et environnementales, assimilation de données, ondes guidées ultrasonores, éléments finis spectraux transitoires, filtres de Kalman

Résumé : Le contexte applicatif de ce travail est l'évaluation de l'état des structures des avions, des ponts, des pipelines et bien d'autres. Plus précisément, nous nous concentrons ici sur les méthodes fondées sur la propagation d'ondes ultrasonores guidées dans le contexte du *contrôle de santé intégré* (SHM), où les parties critiques de la structure sont monitorées par un système distribué de capteurs afin de détecter les anomalies. Dans ce contexte, les *Conditions Environnementales et Opérationnelles* (EOCs) peuvent conduire à une mauvaise interprétation des données mesurées et doivent donc être intégrées à l'analyse des données acquises. En particulier, la propagation des ultrasons est affectée par les éventuels chargements mécaniques affectant la structure. D'abord, ce travail propose de modéliser numériquement la propagation d'ondes guidées dans des structures préchargées. Les effets du chargement sur la propagation des ondes sont modélisés dans le cadre de l'acoustoélasticité. Les ondes ultrasonores se superposent de manière incrémentale à la déformation structurelle causée par le chargement. Cette hypothèse nous permet de définir une formulation élastodynamique linéarisée autour de la (éventuellement) grande déformation préinduite quasi statiquement par le chargement. Nous proposons des méthodes d'éléments finis (EF) adaptées pour résoudre les deux problèmes résultants. La grande déformation statique de la structure est calculée à l'aide des EF de coque 3D, qui devient ensuite une entrée du problème élastodynamique linéarisée résolue par EF d'ordre élevé en transitoire. Cela permet pour les systèmes SHM de modéliser génériquement tout type de géométries, de chargement et de lois de comportement, ici supposées hyperélastiques. Une deuxième contribution est le développement d'une stratégie originale d'assimilation de données pour identifier la déformation à partir des mesures des capteurs. Le problème d'estimation est interprété comme un problème de minimisation aux moindres carrés, qui est ensuite résolu par une combinaison astucieuse d'approches variationnelles et séquentielles, conduisant in fine à une version itérative du filtre de Kalman Unscented. Pour les deux contributions, des illustrations fondées sur de scénarios réalistes démontrent l'efficacité de la stratégie proposée.

Title : Modeling and data assimilation for ultrasonic guided waves in *Structural Health Monitoring* under operational loading conditions

Keywords : structural health monitoring, operational loading conditions, data assimilation, ultrasonic guided waves, transient spectral finite elements, Kalman filters

Abstract : The application context of this work is the assessment of the structural health of aircraft, bridges, pipelines, and many others. More specifically, we focus here on methods based on guided ultrasonic wave propagation in the context of *Structural Health Monitoring* (SHM), where critical parts of the structure are monitored during operation by a distributed system of sensors to detect anomalies. In this context, *Environmental and Operational Conditions* can lead to misinterpretation of the measured data and must therefore be considered when analyzing the acquired data. In particular, it is known that the mechanical loading of the structure affects ultrasound propagation. As a first contribution, this work proposes a numerical method for modeling efficiently the propagation of guided waves in loaded structures. The effects of mechanical loading on wave propagation are modeled using concepts of acoustoelasticity. The ultrasound waves are assumed to be incremental to the structural deformation caused by the loading. This assumption allows us to define a linearized elastodynamics formulation around the (potentially) large deformation caused by the loading. We then propose adapted finite element methods to solve both formulations. First, the structural deformation is computed using 3D Shell FE, which is then an input of the linearized elastodynamics solver based on the transient High-Order Spectral FE. Such a combination allows a suitable application for SHM systems as it can model arbitrary geometries, loading configurations, and constitutive behavior based on hyperelastic laws. Using this direct model, a second contribution of this dissertation is the development of an original data assimilation strategy to recover the deformation from sensor measurements. The estimation problem is interpreted as a least squares minimization problem, which is then solved by a clever combination of variational and sequential approaches, resulting in an iterative version of the Unscented Kalman Filter. Both the presentation of the direct model and its inversion are complemented by illustrations based on large realistic scenarios to demonstrate the efficiency of the proposed approaches.

Institut Polytechnique de Paris
91120 Palaiseau, France